AOP DEVELOPERS' HANDBOOK: SUPPLEMENT TO THE GUIDANCE DOCUMENT FOR DEVELOPING AND ASSESSING AOPs

FOREWORD

This document is the AOP Developers' Handbook supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways (AOPs) [ENV/JM/MONO(2013)6, Second Edition]. The Guidance Document provides a historical background for the AOP development programme, and outlines the elements required to construct an AOP as well as the principles of the AOP framework.

The AOP Developers' Handbook (previously "Users' Handbook") supplement was prepared initially in June 2014 by a subgroup of the Extended Advisory Group on Molecular Screening and Toxicogenomics (EAGMST). At that time it was acknowledged that the Handbook should be revised as expert groups and member countries acquire experience in developing, assessing, and applying AOPs. The present version of the AOP Developers' Handbook reflects the most recent principles, practices, and recommendations pertaining to AOP development as implemented and supported via Release 2.7 of the adverse outcome pathway Wiki (AOP-Wiki; aopwiki.org) and overseen by the Emerging Science for Chemical Assessment (ESCA) advisory group.

Contributing Authors:
Daniel Villeneuve, Bette Meek, Barbara Viviani, Tanja Burgdorf, Carlie LaLone, Jason O'Brien, Dries Knapen, Michelle Angrish, Rex FitzGerald, Shihori Tanabe, Travis Karschnik

TABLE OF CONTENTS

AOP DEVELOPERS' HANDBOOK: SUPPLEMENT TO THE GUIDANCE DOCUMENT FOR DEVELOPING AND ASSESSING ADVERSE OUTCOME PATHWAYS (AOPs)

ABOUT THIS DOCUMENT

This document, the OECD AOP Developers' Handbook, is a supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways (AOPs) [ENV/JM/MONO(2013)6, Second Edition] (AOP guidance hereafter).

The AOP Guidance, originally published in 2013 and revised in 2017, provides an introduction to the terminology and concepts of AOP development, including the identification and use of relevant scientific data and resulting knowledge. The Guidance also briefly outlines some potential applications of AOPs.

While the AOP Guidance document provides a set of definitions and the conceptual background behind AOP development, this AOP Developers' Handbook is designed to provide focused, in-depth, and practical instructions concerning development and review of AOP descriptions in the **AOP knowledgebase (AOP-KB)**, generally accessed via the **AOP-Wiki (aopwiki.org)**. The AOP Developers' Handbook can be thought of as being analogous to the "instructions for authors" used in preparing a journal article. However, rather than describing the preparation of a technical manuscript, this Handbook (organized into sections) details how to develop, structure, and document an AOP description in the AOP-Wiki. Each section corresponds to "pages" in the AOP-Wiki which are presented as standardized template forms to be filled in during developer AOP construction within the AOP-Wiki environment. The guidance provided in each section of this Handbook include descriptions of documentation strategies for AOP development i.e. AOP component descriptions, and organisation of that information into each section of the template Wiki AOP pages. This Handbook also provides more explicit guidance on documentation of the information and the factors considered during collection of the evidence relevant to the AOP and evaluating overall weight of evidence (WoE) considerations that inform both the potential fit-for-purpose applications of the AOP and its relevance to different life stages, sex, taxa, susceptible populations, etc.

AOP Knowledgebase (AOP-KB) refers to the accumulated machine-readable text and data organized and stored in a MySQL database in accordance with the current AOP Data Model and compiled in the AOP XML.

AOP-Wiki (aopwiki.org) is a web-based interface that provides read/write access to the AOP-KB and serves as the official and primary tool for entering new AOP information in accordance with OECD guidance.

A variety of other tools have read access to the AOP-KB via the XML downloads and can make use of the information contained therein for a variety of purposes. At present, the AOP-Wiki is the only portal for entry of new information into the AOP-KB.

Although there is no one-size-fits-all approach to AOP development, the sections of the handbook are organized according to a generalized workflow that applies to many AOP development projects (Figure 1). As with the AOP Guidance itself, this Handbook is not intended to provide a review or summary of the literature informing the AOP concept. It focuses on practical aspects of AOP development and assessment and is intended to promote consistency and ensure all AOP developers and contributors understand the approach for AOP development and contribution within the AOP-Wiki. The template and practices outlined herein, to the extent feasible, are intended to support efficient assembly of information pertinent to an AOP and its components (the focus of Handbook Sections 1-3), as well as transparent documentation of

information considered during evaluation of evidence confidence and the overall assessment, including WoE, of the AOP (the focus of Section 4) along with critical gaps and uncertainties that are relevant to decisions regarding appropriate regulatory applications.

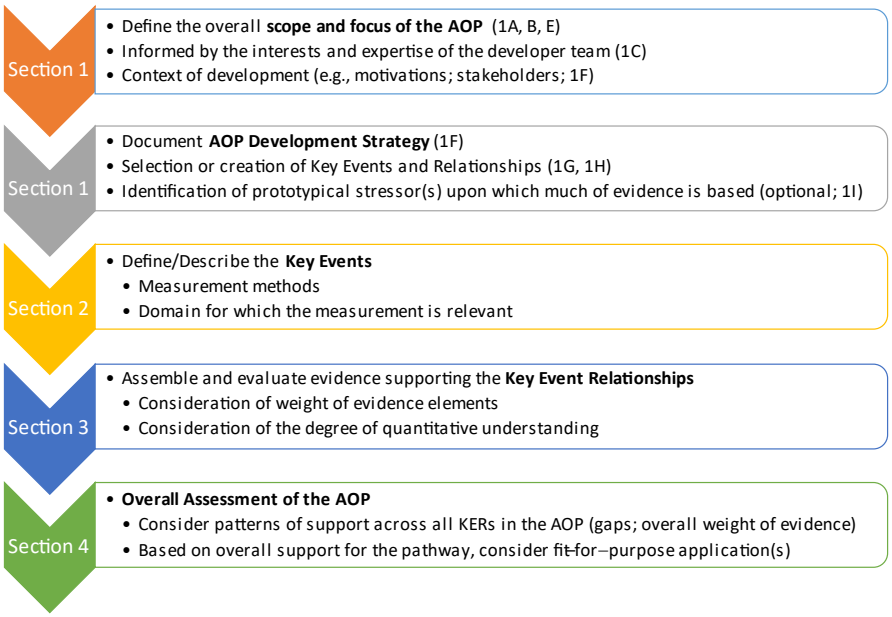

Figure 1. A generalized workflow for AOP development that has informed the organization of the Developer's Handbook.

Developers are encouraged to consult **Annex 1** which outlines a set of guiding questions for evaluating the evidence considered in the overall support for an AOP. Familiarity with these questions before starting an AOP development project can guide the initial scoping including expert solicitation and review of existing literature and/or the design of novel studies toward the data that best inform and support AOPs. Review of the guiding questions and weight of evidence considerations are intended to cue developers on the types of studies that are most influential in providing support for regulatory applications. AOPs are generally best supported by studies that consider multiple key events where comparisons of the concentration, time, or incidence of biological effect in the sample population is not confounded by variations in experimental design. Essentiality of any given key event along the pathway is best evaluated by examining the effects of its prevention or modulation on all downstream events. Searching for or designing studies that best address the guiding questions in **Annex 1** can be expected to lead to both efficient, and high quality AOP development.

AOP descriptions developed as part of the OECD AOP Development Programme are peer-reviewed according to procedures outlined by the OECD [Guidance Document for the Scientific Review of AOPs; ENV/CBC/MONO(2021)22]. Because AOP descriptions within the AOP-Wiki are viewed as living documents, they are expected to continue to evolve over time, as new evidence may increase or decrease the overall confidence and certainty in an AOP or its component(s). Consequently, AOPs that are reviewed and endorsed by the OECD will have multiple versions, namely, a static pdf version created at the time of the review or endorsement (termed a "snapshot"), and the current version in the AOP-Wiki, which can continue to change over time. Reviews are performed on these static versions which are permanently stored in the AOP-KB. In this way, users can distinguish content that has been peer-reviewed and endorsed from that which may have been added or modified afterward. The time-stamped, static versions corresponding to the endorsed version of the AOP are also published in the OECD series on Adverse Outcome Pathways.

INTRODUCTION TO ADVERSE OUTCOME PATHWAYS (AOPs)

An AOP describes a sequence of events commencing with initial interaction(s) of a stressor with a biomolecule within an organism that causes a perturbation in its biology (i.e., molecular initiating event, MIE), which can progress through a dependent series of intermediate key events (KEs) and culminate in an adverse outcome (AO) considered relevant to risk assessment or regulatory decision-making (Table 1). AOPs are composed of a causal sequence of upstream to downstream KEs, representing a cascading series of measurable biological changes that can be expected to occur if the perturbation is sufficiently severe (i.e., in terms of potency, duration, frequency) to drive the pathway all the way to the AO. **Importantly, AOPs do not describe every detail of the biology but instead focus on describing critical steps or check-points along the path to adversity, which are both measurable and have potential predictive value for regulatory application.** While the focus of AOP development is to capture and organise what is known, the process of AOP development may also identify current knowledge gaps which, if filled, could further improve predictive utility.

Table 1: Definitions of key terms and abbreviations used in this Handbook (see AOP guidance for additional terminology relevant to the AOP framework and its application).

Molecular initiating event	MIE	A specialised type of key event that represents the initial point of chemical/stressor interaction at the molecular level within the organism that results in a perturbation that starts the AOP.
Key event	KE	A change in biological or physiological state that is both measurable and essential to the progression of a defined biological perturbation leading to a specific adverse outcome.
Key event relationship	KER	A scientifically-based relationship that connects one key event to another, defines a causal and predictive relationship between the upstream and downstream event, and thereby facilitates inference or extrapolation of the state of the downstream key event from the known, measured, or predicted state of the upstream key event.
Adverse Outcome	AO	A specialised type of key event that is generally accepted as being of regulatory significance on the basis of correspondence to an established protection goal or equivalence to an apical endpoint in an accepted regulatory guideline toxicity test.

KEs are measurable biological changes that are essential to the progression along an AOP. Essentiality indicates that the KEs play a causal role in the pathway, such that if a given KE is prevented or fails to occur, progression to subsequent KEs in the pathway will not occur. While KEs are essential to progression along the AOP, they are not necessarily sufficient. The extent of triggering of the pathway (influenced by intensity and duration of exposure to a stressor) determines whether it will progress all the way to the AO The conditions under which progression can be expected are described as quantitatively as possible, in the KERs that link an upstream to a downstream KE.

The suitability of a given AOP for application in different regulatory contexts is influenced by (1) the confidence and precision with which the KEs can be measured, (2) the level of confidence in the relationships between the KEs linked in an AOP (KERs) based on biological plausibility and empirical support for the KERs; and (3) WoE for the overall hypothesised pathway, taking into account additional considerations including any uncertainties and inconsistencies. Therefore, overall assessment of AOPs is best supported by providing thorough descriptions of the KEs [Section 2], relationships between those KEs [i.e., KERs, Section 3], and by final consideration of the overall patterns of support including plausibility and other direct and indirect empirical evidence of causal relationships across the key events defined for the pathway that increase or decrease overall confidence in the AOP [Section 4].

The overall patterns of support, ultimately inform the suitability (i.e., fit-for-purpose) for various types of applications. Consequently, both the Handbook and AOP-Wiki are structured in a manner that include structured pages and prompts for AOP developers to provide relevant types of supporting documentation.

Principles of AOP Development and their Implications for AOP Description

As a pragmatic convention, AOPs are conceptualised as a single sequence of events proceeding from the MIE to the AO via a series of intermediate KEs (Villeneuve et al. 2014a). That is, they describe how one particular molecular perturbation may cause one AO, not every possible AO that perturbation may cause, nor every perturbation leading to a particular AO. MIEs, KEs, and AOs may be shared by more than one AOP to form an AOP network. Consequently, KEs should be constructed as discrete (modular) units without reference to a specific MIE, AO, or other KEs. Likewise, it is important that KERs describing relationships between discrete pairs of KEs are independent of other elements of the AOP. This facilitates generation of self-contained KE and KER descriptions that can be linked to multiple other AOPs. Such an approach both fosters consistency and increases efficiencies in the AOP development process, by eliminating the need for AOP developers to completely re-describe biological measurements (KEs) or evidence supporting the relationship between two KEs (KERs) that another developer may have already detailed. Maintaining KE and KER descriptions as discrete units that avoid reference to other elements of the AOP also facilitates the updating of KE and KER descriptions as new methods for measuring KEs or new evidence supporting KERs are developed. Finally, it facilitates the construction and conceptualisation of AOP networks.

An AOP network is defined as an assembly of two or more AOPs that share one or more KEs (Knapen et al. 2018). Because the components of an AOP (KEs and KERs) are described in the AOP-Wiki, in a modular fashion, AOP networks emerge from the description of individual AOPs that share KEs. AOP networks capture broader knowledge concerning the range of possible AOs which a perturbation may cause, or the variety of upstream KEs which can lead to a given AO. AOP networks are also suited to address exposures to multiple stressors that lead to the same AO or individual stressors that activate multiple MIEs (Knapen et al., 2015; Villeneuve et al., 2014a, b; Knapen et al. 2018).

In describing the KEs and KERs of an AOP, the content of each information field of the KE or KER description should be completed where possible and supported by citation of primary literature and other relevant sources. Nevertheless, AOP descriptions reflect current knowledge and will evolve as additional information becomes available, so AOP descriptions should be regarded as "living documents" that reflect the state of knowledge at the time they were last updated. It is expected that, as "living documents", AOPs may have gaps that may be addressed over time as the science progresses or as other researchers contribute. This also encourages collaboration and contributions between experts in various areas of research and the regulatory risk assessment community.

AOPs thus provide a relevant construct to promote collaboration and better coordinate and tailor research to practical application, such as the development of KE-based testing strategies. The AOP-Wiki facilitates this by providing a tool to organise and share the relevant data and information. Consequently, it is recommended that descriptions are structured using presentation of bullets or tables and organised into topical subsections rather than as extensive narrative text.

In this Handbook, particular emphasis is placed on sections related to the description of the MIE, KEs and AO in an AOP (Section 2), the assembly of available scientific evidence supporting the KERs (Section 3) and the overall support for the AOP as a whole (Section 4) and may additionally consider its potential application (Figure 1).

AOP descriptions should be supported with well documented and transparent citation of the appropriate peer-reviewed literature and/or other relevant sources. Authors are encouraged to provide references formatted according to the OECD Style Guide (https://www.oecd.org/about/publishing/OECD-Style-Guide-Third-Edition.pdf).

REVIEW AND ENDORSEMENT of AOP-WIKI CONTENT

AOPs developed and evaluated according to the guidance in the Handbook may be submitted for technical review via the OECD AOP Development Programme, submitted for potential publication in a partner journal[1], or have a review managed by an approved third party organization, provided the reviews are managed as described in the Guidance Document for the scientific review of Adverse Outcome Pathways. AOPs that are accepted after review and revision according to the guidance are then eligible to be added to the OECD AOP Development Workplan and considered for endorsement by the OECD Working Party on Hazard Assessment (WPHA) and/or Working Group of the National Coordinators for the Test Guidelines Program (WNT).

[1] Link to current listing of partner journals that have signed a memorandum of understanding (MOU) to review AOPs in the AOP-Wiki as per the guidance document.

OBTAINING AUTHOR ACCESS TO THE AOP-Wiki

Read-access to all contents of the AOP-KB is publicly available via the AOP-Wiki (aopwiki.org) without need to create a user profile, login ID, or password.

Commentor access: A self-created user account, with a verified email address, grants the user the ability to comment on all pages in the AOP-Wiki including AOPs, KEs, and KERs. Users can create an account on the AOP-Wiki by clicking the "Register" button on the AOP-Wiki home page.

Author Access: In order to create or edit AOPs, KEs, or KERs, the user must request author access to the AOP-Wiki by following the instructions here.

A NOTE ON AOP DESCRIPTIONS IN THE AOP-Wiki

AOP descriptions in the AOP-Wiki consist of both structured information and free text.

Structured information fields in the AOP-Wiki employ standardised ontologies or controlled vocabularies available through look-up tables or by making selections from a drop-down list. Structured information fields within the AOP-Wiki populate a back-end database and can be exported in a machine-readable format (i.e., XML) that can be used in a variety of computational analyses, and more complex querying, and searching of the AOP-KB. For example, construction of AOP networks from the modular units of individual AOP descriptions relies on these structured annotation fields.

Free text sections in the AOP-Wiki provide AOP developers with much greater descriptive flexibility than structured information fields. While free text is searchable, it is not standardised and machine-readable and is not part of the XML download, thus limiting its use from a computational standpoint.

CONTENT LICENSING

By default, all content in the AOP-Wiki is licensed under a Creative Commons, Attribution, Share Alike (CC BY-SA) license. This license stipulates the following:

- Users must not **restrict access** to the work using technical measures, or otherwise attempt to impose limitations on the freedoms to use, study, apply, redistribute, or distribute derivative works.
- Users must **give proper attribution to the author and retain the license notice.**
- Users must **release derivative works under identical license terms**.

Any reuse of AOP-Wiki content or derivative of AOP-Wiki content requires appropriate attribution including a link to the license and indication of any changes made. AOPs are, however, represented by pages within the AOP-Wiki that have page-specific accessibility properties. AOP page licensing options (Table 2) are described below.

Key **Event** and Key Event **Relationship** pages in the AOP-Wiki are shared pages that any author can edit. Consequently, at present, only a BY-SA license can be applied. Authors wishing to protect unpublished content on an Event or Relationship page, are encouraged to develop their content on an external pre-print server, and then cite the appropriate DOI on the relevant Event or Relationship pages in the AOP-Wiki. To facilitate attribution, authors may also want to "tag" content they have added to these shared pages with their name or initials.

AOP Pages have restricted author access in the AOP-Wiki. They can only be edited by authors listed as contributors. Consequently, there is an option to directly protect content of an AOP page, if desired. At the time an AOP page is first created in the AOP-Wiki (**and only at that time**), authors have the option to override the default CC BY-SA license and instead select a "©; Copyright, All Rights Reserved" license. A © license indicates that the author retains all rights provided by copyright law, and prohibits others from reproducing, distributing, and/or adapting any part of the work without the copyright holder's permission. Conceptually, this allows AOP-pages on the AOP-Wiki to function as a pre-print server. While the content under development is visible to other authors and potential users, the content is restricted and protected by law. This option is provided to encourage transparent AOP development on the AOP-Wiki, while protecting the intellectual property of the authors and the effort they have put into developing the AOP.

To ensure the ultimate accessibility and usability of information in the AOP-Wiki, All Rights Reserved licenses in the AOP-Wiki automatically revert to CC BY-SA after 12 months from the AOP page creation date, unless the authors take action to extend the All Rights Reserved license, prior to its expiration. The All Rights Reserved License can be extended at any time, prior to its expiration by clicking the "Edit" button on the AOP page and then clicking the "Extend current All Rights Reserved License" button from the Editing page. An active All Rights Reserved License can be extended multiple times. However, it is the authors responsibility to monitor the All Rights Reserved expiration date and take action to extend the term before the expiration date. The current expiration date for an All Rights Reserved License can be found on the Editing page, in blue highlighted text positioned directly above the "Extend current All Rights Reserved License" button.

Once the All Rights Reserved license expires, the AOP page defaults automatically to a CC BY SA license. The authors can also switch to a CC BY-SA license at any time by clicking the "Edit" button on the AOP page, then making a new license selection. Note, **any switch to a CC BY-SA license is irreversible**. Once an AOP page defaults or is switched by the authors to a CC BY-SA license, it cannot be changed back to an All Rights Reserved license.

In addition to the default CC BY-SA license, authors also have the option to select a CC BY-SA License with an "Open for Adoption" tag. This option applies the same license terms as the CC BY-SA license, however, it is used to signal that the original authors are no longer actively

developing the AOP and invite new authors to take over development. New authors wishing to take over development of the AOP can do so by contacting the AOP-Wiki gardening team at aopwiki@googlegroups.com. Note, an All Rights Reserved License cannot be applied to an AOP page that has been opened for adoption.

Table 2: AOP page License Options Overview[1]

License Option	Terms	Implementation Notes
All Rights Reserved	Re-use of the content of the AOP page, in any form, requires advanced, written permission from the authors.	Must be selected at the time the AOP page is first created. Expires after 12 months unless extended by the authors. Once an All Rights Reserved license expires or a different license type is selected, it is **not** possible to revert back to an All Rights Reserved license.
BY-SA	This license allows users to distribute, remix, adapt and build upon the material in any medium or format so long as attribution is given to the creator(s). The license allows for commercial use. However, if you remix, adapt, or build upon the material all derivative works must be licensed under identical terms.	This is the default license applied at the time of AOP page creation, unless an All Rights Reserved license was selected at that time. Authors can switch from All Rights Reserved (if applicable) to BY-SA at any time. However, it is not possible to revert back to All Rights reserved once a BY-SA selection has been made.
BY-SA Open for Adoption	This license allows users to distribute, remix, adapt and build upon the material in any medium or format so long as attribution is given to the creator(s). The license allows for commercial use. However, if you remix, adapt, or build upon the material all derivative works must be licensed under identical terms.	This option is available on the Editing Page, accessed by clicking the "Edit" button on the AOP page. This selection is used to signal that the original authors are no longer developing the page and invite other developers to take over. An All Rights Reserved license cannot be applied to an AOP page that was opened for adoption.

1 License options described apply only to AOP pages in the AOP-Wiki. Key event and key event relationship pages are BY-SA only.

SECTION 1 – AOP DESCRIPTION

This section is for information on the AOP to be. entered on the upper portion of an AOP page within the AOP-Wiki. Here the overall structure of the AOP is introduced, the motivation and strategy for its development described and the component KEs and KERs are listed.

1A. AOP Identifier and Title

This subsection provides guidance for naming the AOP.

i. AOP Identifier

Each AOP is automatically given a numerical AOP identifier by the AOP-Wiki when it is created (e.g., AOP: ###).

ii. (AOP) Title

Each AOP should be given a descriptive title that takes the form "MIE leading to AO via distinctive KE". For example, "Aromatase inhibition [MIE] leading to reproductive dysfunction

[AO] via reduced vitellogenin production" or "Thyroperoxidase inhibition [MIE] leading to decreased cognitive function [AO] via decreased circulating thyroid hormone concentrations". While each AOP is distinguished in the AOP-KB and AOP-Wiki by their AOP page ID numbers and unique URL, in a growing number of cases where AOPs linking the same MIE to the same AO are being entered into the AOP-Wiki, the "via distinctive KE" descriptor makes it easier to distinguish different AOPs within a network of closely releated AOPs.

In cases where the MIE is unknown or undefined, the earliest known KE in the sequence (i.e., furthest upstream) should be used in lieu of the MIE and it should be made clear that the stated event is a KE and not the MIE.

iii. Short Name

A short name should also be provided that succinctly summarises the information from the title. This name should not exceed 90 characters.

1B. Graphical Representation of the AOP

A graphical summary of the AOP listing all the KEs in sequence, including the MIE (if known) and AO, and the pair-wise relationships (links or KERs) between those KEs should be provided. This is easily achieved using the standard box and arrow AOP diagram (Figure 2).

Figure 2. Generic AOP diagram, where boxes represent KEs and arrows represent KERs.

Development tip 1 – Graphical Representation: The graphical representation (AOP diagram) serves as a useful road-map to guide AOP development in the AOP-Wiki. For this reason, it is recommended that an AOP diagram be developed prior to creating an AOP description in the AOP-Wiki. Starting with the graphical summary provides a useful overview of the KE and KER pages that will need to be included. Ideally, development of a graphical overview of the AOP should be followed by a search of existing content to determine whether analogous AOPs and/or KEs or KERs already exist in the knowledgebase. This prevents duplicated effort and help to ensure that KEs and KERs are shared among AOPs, allowing for de facto creation of AOP networks. Once existing KE and KER pages relevant to the AOP have been identified, the developer then knows which pages in the AOP-KB will need to be edited or created de novo.

The graphical summary is prepared and uploaded by the user is often included as part of the proposal when AOP development projects are submitted to the OECD AOP development workplan.

The graphical representation, or AOP diagram, provides a useful and concise overview of the KEs that are included in the AOP, and the sequence in which they are linked together. This can aid both the process of development, as well as review and use of the AOP.

***Development tip 2 – Number of KEs to include*:** Determining the number of KEs to include in an AOP and the specificity with which they are defined is one of the more challenging aspects of AOP development. In describing KEs within an AOP, it is important to recognise their distinction from "mechanism of action". AOPs provide a description of a limited number of essential, measurable events (check-points or nodes of convergence of mechanistic pathways most relevant to informing application) leading to induction of the relevant toxicity endpoint. They do not necessarily provide a comprehensive molecular description of every aspect of the biology involved. With that in mind, the following "rules of thumb" can help guide the process of KE definition (Villeneuve et al. 2014a, b):

- Where possible and appropriate for application, try to include at least one KE at each major level of biological organisation (molecular, cellular, tissue, organ, individual).
- Where feasible/appropriate, focus on KEs that can be measured in a relatively routine manner over those that require highly specialised expertise, equipment, or supplies to measure. These will tend to be the KEs for which empirical evidence to support KERs is more likely to be available to support the WoE evaluation.
- Select a limited number of KEs that are measurable and for which evidence supports plausibility and potential predictive utility. Where relevant, more detailed description of the underlying biology involved can be incorporated into the descriptions of the biological plausibility linking two KEs (see section 3 – KER descriptions).

Development tip 3 – Branching of AOPs captured on a single AOP page

In principle, an individual AOP is defined as a single, non-branching sequence of KEs, linked by KERs that connect a single MIE to an AO (Villeneuve et al. 2014a). Consequently, most AOPs pages should define a single, non-branching, sequence of KEs linked by KERs. However, it is recognized that in some cases there may be exceptions for which representation of a simple AOP network on an AOP page is a more pragmatic unit of development and evaluation (see Leist et al. 2017 for examples and further explanation). In such cases, representation of a branched structure on an AOP page is acceptable, so long as the principles of modularity of the KEs and KERs and overall coherence to the framework is maintained.

For example, representation of branching on an AOP is acceptable when there are multiple KEs, causally linked to the MIE and AO that are occurring concurrently and acting in concert to drive the downstream effects. In such cases, the various KEs cannot be placed neatly into a single temporal sequence because they are effectively occurring simultaneously. Likewise it cannot be determined which of the concurrent KEs is most essential or critical, because there are multiple KEs contributing jointly such that it cannot be effectively determined whether one could cause the pathway to progress without the other. This is contrasted with cases where KEs act independently such that one event or the other, alone, would allow progression toward the outcome.

In cases where an additive ("and") relationship must be assumed, representation of a simple AOP network on a single AOP page within the AOP-KB may be more practical from both a development and use stand-point than breaking those multiple highly related pathways into separate AOP descriptions. As long as KEs and associated KERs are each represented as separate modular pages in the AOP-KB (as described below), capturing such networks on single AOP pages does not create problems for modular AOP network building. Indeed, it can strengthen the overall AOP by capturing the evidence for pleiotropic effects of the same MIE that ultimately contribute to the same outcome.

Note, such branched AOP structures should only be included on a single AOP page when all the branches diverge from a common MIE (or MIEs in the case that two or more MIEs MUST occur to drive the pathway) and converge to a common AO (Figure 3A) and two or more of the KEs contributing causally to the AO occur concurrently such that it is experimentally intractable to isolate and identify which is playing the dominant causal role and all KEs have predictive value.

Branched structures should not be included on a single AOP page when they diverge to independent adverse outcomes (e.g., Figure 3B) and/or are operating largely independent of one another and can be experimentally resolved from one another in space or time. Following this logic, two or more MIEs may occur on an AOP page, when two or more MIEs MUST occur simultaneously in order for the pathway to be triggered (Figure 3C).

448

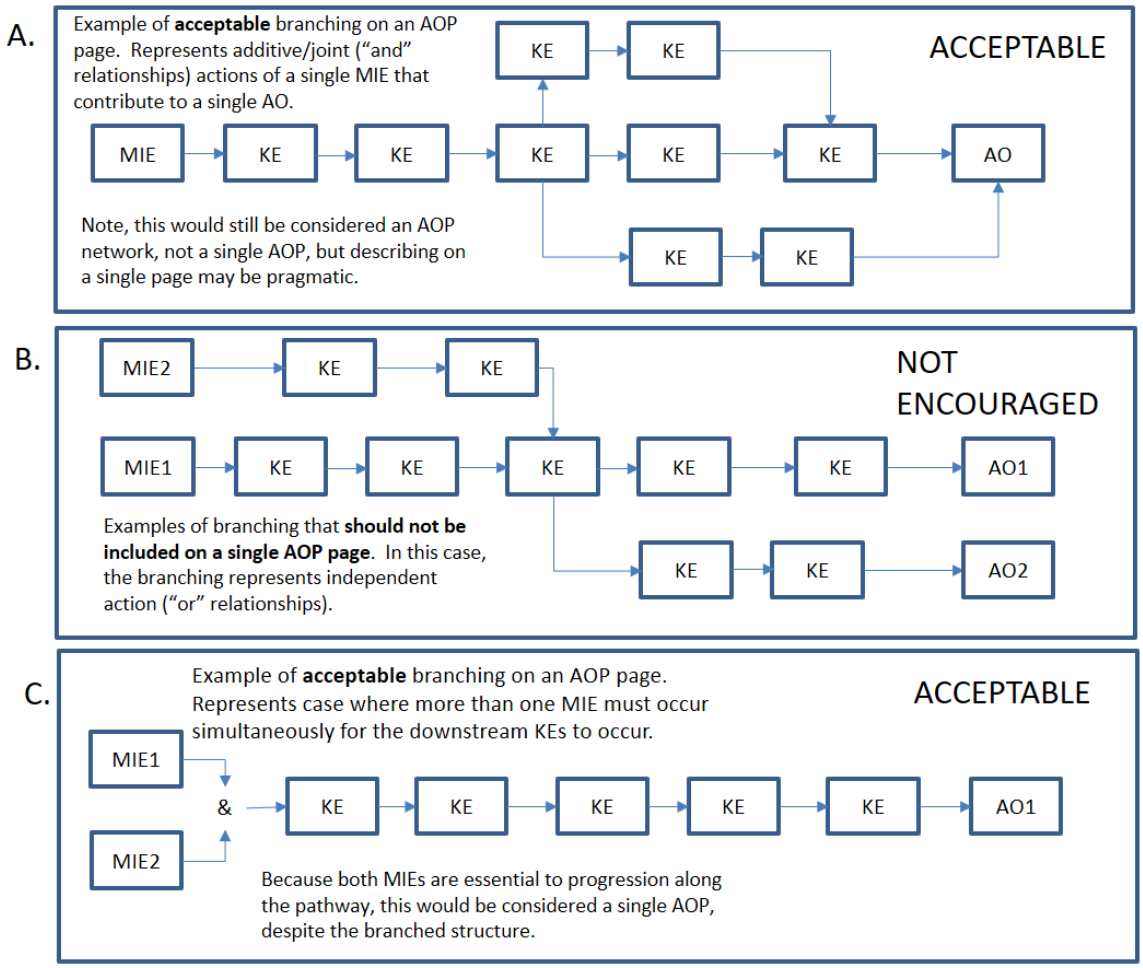

Figure 3. Illustration of general guidance regarding inclusion of simple AOP networks or branched AOP structures (A) on a single AOP page. Branching representing independent actions leading to more than AO should not be included in an AOP description (B). Branching indicating multiple KEs (including MIEs) that MUST occur for the pathway to progress downstream should be included in an AOP description. In case multiple MIEs are essential, branching of MIEs are acceptable (C).

1C. Authors of the AOP

This section provides guidance on author identification.

i. Authors and Affiliations

List the name and affiliation information of the individual(s)/organisation(s) that created/developed the AOP. In the context of the OECD AOP Development Workplan, this would typically be the individuals and organisation that submitted an AOP development proposal to ESCA and further considered under an OECD working party (e.g., WPHA, WNT). Significant contributors to the AOP should also be listed. A corresponding author with contact information may be provided here. This author does not need an account on the AOP-Wiki and can be distinct from the point of contact below. The list of authors will be included in any snapshot made from an AOP.

ii. Point of Contact

Indicate the point of contact for the AOP-Wiki entry itself. This person is responsible for managing the AOP entry in the AOP-Wiki and controls write access to the page by defining the contributors as described below. Clicking on the name will allow any wiki user to correspond with the point of contact via the email address associated with their user profile in the AOP-

Wiki. This person can be the same or vary from the corresponding author listed in the authors section. In cases where the individuals are different, the corresponding author would be the appropriate person to contact for scientific issues whereas the point of contact would be the appropriate person to contact about technical issues with the AOP-Wiki entry itself.

Corresponding authors and the point of contact are encouraged to monitor comments on their AOPs and develop or coordinate responses as appropriate. Selecting the "Watch" (

)option on the AOP page will allow an e-mail alert to be sent whenever changes to the AOP page or linked KE or KER pages are made.

iii. AOP-Wiki Contributors

List user names of all authors contributing to or revising pages in the AOP-Wiki that are linked to the AOP description. Identification of contributors in this section controls write access to the AOP page. Only contributors listed here, with author rights in the AOP-Wiki, can edit the AOP page.

iv. Coach(es)

This field is used to identify coaches who supported the development of the AOP.

Coaches are experienced AOP developers that are familiar with the guidance document, AOP development principles, and navigation within the AOP-Wiki. They assist AOP developers by answering questions about the framework, the organization of information in the AOP-Wiki and facilitate compliance with the guidance document and best practices. Upon acceptance of the AOP development project under the OECD workplan, a coach will be assigned. AOP developers without an OECD workplan – related project can request a coach from the SAAOP (Society for the Advancement of AOPs) via aopwiki@googlegroups.com.

Identification of coaches in this section provides acknowledgement of the volunteer contributions made by the coach(es) and professional recognition.

1D. Handbook Versioning and OECD Status

i. Handbook Version

- As the AOP framework evolves and information fields, features, or functions are added or modified in the AOP-Wiki, the AOP Developers' Handbook (this document) is updated to reflect the current state of the AOP-Wiki. In many cases, the AOP-Wiki and Handbook may undergo several updates over the duration of an AOP development project. Newly added AOPs are required to comply with the version of the Handbook that was current on the date the AOP was created, or newer. Where feasible, authors are encouraged to update their AOPs for consistency with the current Handbook version. However, this in not always possible or practical. Consequently, the "Handbook Version" column of the "Status" table is used to indicate the version of the Handbook that the authors used to guide their development.
- When a developer creates an AOP, the current version of the Handbook, on the date of creation, will be automatically populated into the "Handbook Version" column of the "Status" table, along with a link to that version of the Handbook. This information will also display in the "Title" section of the AOP page, right under the "Short name". As newer versions are released, the authors have the option to switch to a newer Handbook version by selecting from a drop down menu on the Edit page. However, they cannot select versions that pre-date the creation date of their AOP. Both archived handbook versions and release notes summarizing the major changes can be found on the Developers' Handbooks archive page (https://aopwiki.org/handbooks).

ii. OECD Status

For AOPs that are included in a project that has been accepted into the OECD AOP Development Workplan (see http://www.oecd.org/chemicalsafety/testing/projects-adverse-outcome-pathways.htm), the status with regard to progress through OECD review and

endorsement processes is indicated . 'OECD status' tracks the level of review/endorsement of the AOP . This designation is managed and updated by the OECD. It cannot be changed by the AOP author(s). AOPs in the AOP-Wiki can be filtered by their OECD status either using the table heading filters on the AOP listing page, or by clicking the "With OECD status" button on the listing page, which toggles to the OECD view and contains only those AOP development projects that are part of the OECD workplan. The OECD status designations for filtering purposes are the following:

- WPHA/WNT Endorsed
- ESCA Approved
- Under Review
- Under Development

iii. OECD Project Number

The OECD project number is assigned upon acceptance into the OECD AOP development workplan and indicated along with the current OECD status of the AOP This designation is managed and updated by the OECD. It cannot be changed by the AOP author(s). OECD project numbers are listed in the all AOPs listing table (blank for AOP development projects not on the OECD workplan), and are displayed on the "OECD View" page, which is accessed by clicking the "With OECD status" button on the AOP listing page.

iv. Date Modified

The date the AOP was last modified is automatically tracked by the AOP-Wiki. The date modified field can be used to evaluate how actively the page is under development and how recently the version within the AOP-Wiki has been updated compared to any snapshots that were generated.

1E. ABSTRACT

In the abstract section, authors should provide a concise and informative summation of the AOP under development. Abstracts should typically be 200-400 words in length (similar to an abstract for a journal article). Suggested content for the abstract includes the following: (1) the background/purpose for initiation of the AOP's development (if there was a specific intent); (2) a brief description of the MIE, AO, and/or major KEs that define the pathway; (3) a short summation of the overall WoE supporting the AOP and identification of major knowledge gaps (if any); (4) a brief statement about how the AOP may be applied (optional). The aim is an "executive summary" to capture the highlights of the AOP and its potential scientific and regulatory relevance.

1F. AOP Development Strategy

This subsection describes key elements of "Why" (Context) and "How" (Strategy) the AOP was developed. The content informs other developers, reviewers and users about the strategy and focus for identification and assimilation of the relevant evidence base for KEs and KERs in the AOP.

Context:

This subsection describes key elements of *why* the AOP was developed and for whom (e.g., funding sources; stakeholders; etc.).

Below are examples of the *types* of information to include:

- Key research question(s) or regulatory needs being addressed
- Scope and basis for the evidence gathering/literature search scope
    - e.g., focused on a specific taxonomic group?
    - adding new branches to an existing AOP?
    - development of an additional KE/KER?

- Acknowledgement of the source of funding (if applicable)
- The overall objective/envisaged use of the AOP that informed its development, e.g., to
    - document biology based on specialized expertise,
    - establish the relevance and utility of an assay,
    - develop an organizing construct in stressor specific (quantitative) hazard characterization,
    - contribute to development of an integrated approach to testing and assessment, etc.
    - indication of interesting biology encompassed by the AOP that is not necessarily evident from the KE and KER descriptions;
    - as part of a network-guided approach to AOP development, noting other AOP(s) developed as part of the effort
- Other information that may be useful to the AOP developer and/or user that facilitates understanding of motivation/objective/scope for AOP development.

Strategy

This subsection describes *how* the AOP was developed. Specifically, what was the strategy, focus and workflow for identification and assembly of relevant evidence to meet the objective/envisaged application? This information is critical to facilitate the reuse of components and expansion of AOPs. Transparency of the rationale for identification and selection of supporting data also contributes to confidence for regulatory application of AOPs and/or their components.

Developers should tailor the contents of this section to their particular AOP context and approach, depending e.g., on the scope, nature of prior documentation of the pathway, the starting point for development (e.g., the MIE or AO), complexity, and/or envisaged application(s). For example, it may build on previously well-documented and accepted pathways, with focus on particular aspects of uncertainty or particular components of the pathway.

Content may include:

- **Overall data search and identification strategy/ies,** including general strategies (i.e., workflow) for information search, retrieval, and screening (and possibly assessment). Example content includes:
    - reliance on prior knowledge and/or documentation of the pathway, e.g.,
        - expert knowledge
        - previously conducted stressor specific (systematic) reviews documenting KEs and KERs
        - previous AOP descriptions
    - overview of data identification and search strategies, including initial and refined approaches, e.g.,
        - search terms, search strings, etc. and databases searched, the time period of searching, and returned results,
    - novel data – describe the type(s) of experiments that were conducted, specialized software and tools used for assimilation, screening and assessment of information for relevance to the AOP.

- **Level of resolution / detail in terms of the KEs and KERs represented** in the pathway. A goal in AOP development is to identify notable milestones or checkpoints in the progression of an adverse biological response that are both measurable and have predictive utility relevant to regulatory application. It is not, necessarily, to describe every detail of the biology. Consequently, there is often a considerable degree of fit-for-purpose judgement that goes into determination of how many KEs to include, which to include, and the degree of resolution with which they are defined (e.g., lumping versus splitting). As different developers and users of the content may have different perspectives on what level of detail/resolution to employ, a description of how the authors arrived at their

decisions regarding the level of resolution to include when defining the KEs can also be helpful.

- Example: KE #___ was included to align with a specific high throughput assay that has been developed.
- Example: KE #___ was already included in several AOPs and re-use of this KE in the present AOP helps to link this AOP to a broader network of related AOPs.
- Example: Although a multistep signal transduction pathway plays a role in mediating the relationship between X & Y, the steps involved are not easily and readily measured – consequently they are just described in the KER description(s) rather than represented as separate KEs.

The description in this section provides an *overview* of the search strategy relevant to inclusion of the KEs and KERs in the AOP. Considerations for documentation of more detailed information on search and assimilation strategies for individual KERs is presented in Section 3.

1G. KE and KER Tables

Tables listing each KE and KER are automatically created in the AOP-KB as KE pages to link to the AOP are selected or created and as KERs are defined.

- **KE Table**: This table summarises all of the KEs of the AOP, including the MIE and AO. This table is populated in the AOP-Wiki as KEs are added to the AOP. Each table entry acts as a link to the individual KE description page. For guidance on completing the KE descriptions see Section 2.
- **Relationship Table**: This table summarises all of the KERs of the AOP and is populated in the AOP-Wiki as KERs are added to the AOP. Each table entry acts as a link to the individual KER description page. For guidance on completing the KER descriptions see Section 3.

1H. Network View

The AOP-Wiki automatically generates a network view of the AOP (Figure 4). This network graphic is based on the information provided in the MIE, KEs, AO, KERs, and WoE summary tables. The width of the arrows (representing the KERs) is determined by its WoE confidence level, with thicker lines representing higher degrees of confidence. This network view also shows which KEs are shared with other AOPs. Visibility of non-adjacent relationships and/or other AOPs that share KEs with the AOP in question can be toggled on and off, as can the names of KEs. Users can customize the layout of network representation of the viewer. If logged in, that customized view is retained when returning to the AOP-Wiki.

With AOP-Wiki release 2.6 there is also an option to display the AOP in third party tools that allow for alternative visualization of the AOP in an AOP network context. These third party options can be accessed via the “Explore in a Third Party Tool” button.

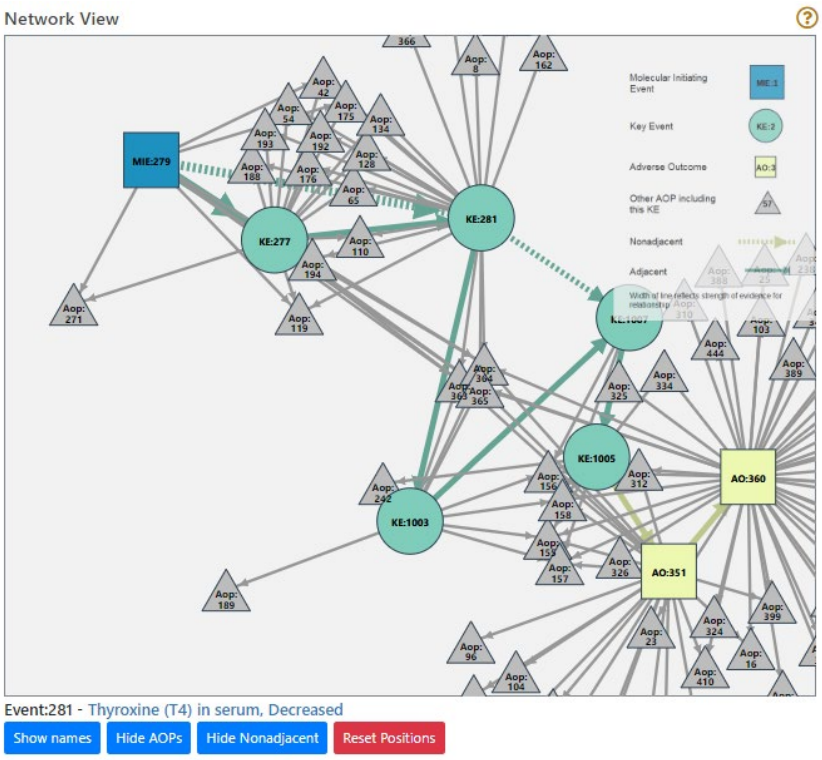

Figure 4. Example of the default network view in the AOP-Wiki. Note the option to hide or show AOPs that share one of more or the same KEs, non-adjacent relationships, and event names.

1I. Prototypical Stressor(s)

The Prototypical Stressor field is a structured data field that can be used to identify one or more "prototypical" stressors that act through this AOP. However, please recall that an AOP should not be stressor-specific. Prototypical stressors are stressors for which responses at multiple KEs in addition to the MIE have been well documented. Experiments with the prototypical stressor(s) may have provided much of the empirical support for the AOP and/or quantitative understanding of the KERs. Thus, prototypical stressors identified may serve as useful "positive controls" for evaluating responses of other stressors that may act on this pathway and/or provide insights into the types of structures or properties that may be relevant to the stressor domain that is relevant to this AOP. The relative potency of various other stressors, compared to the prototypical stressor(s) may also be informative relative to quantitative understanding of the KERs and associated applications of the AOP.

Please note:

- This field is NOT intended to provide a comprehensive listing of all stressors known to act through this AOP.
- It is NOT intended that AOPs will be searchable by prototypical stressor(s)
- Identification of a prototypical stressor does NOT indicate the AOP is stressor specific.

In the case of prototypical stressors that are chemicals, chemical names can be selected from established chemical ontologies. However, non-chemical stressors such as radiation, genetic or environmental factors, disease vectors or viruses, etc. may also be identified. Authors are encouraged to utilize appropriate ontologies wherever possible.

1J. Life Stage/Taxonomic/and Sex Applicability

See Section 4 on Overall Assessment of the AOP

716
717 **1K. Overall Assessment of the AOP**
718 See Section 4
719
720

SECTION 2 – KE DESCRIPTIONS

Development tip 4 – Sharing of KEs:

- **Use existing KEs when possible** - when adding KEs to an AOP it is strongly recommended to use KEs that already exist in the AOP-Wiki as much as possible. When adding a new KE in the AOP-Wiki, the system will identify events using related terms to aid in reviewing whether suitable KEs already exist.
- **Existing KE requires modification** - If an existing KE requires modification to make it suitable, changes to the content on that page should be coordinated with the point(s) of contact for other AOPs sharing the KE to ensure that the original meaning is not altered.
- **AOP-KB Etiquette** – When using an existing KE, it is the responsibility of the person making changes to ensure that KEs used in multiple AOPs are not altered in such a way as to diminish the applicability of that KE for the existing AOPs. Please be courteous to your fellow AOP developers.
- **Creating new KEs** - If no suitable KEs are available in the AOP-Wiki, or if the revisions needed to make an existing KE description suitable for the AOP under-development would make it unsuitable for use in AOPs it is already linked to, then a new KE should be created.

2A. Event ID

When a KE is created, an ID number is automatically assigned to it (Event: ###). This number is used for tracking the KE in the AOP-KB and corresponds with a unique URL of the form https://aopwiki.org/events/###.

2B. KE Title

The KE title should describe a discrete biological change that can be measured. It should generally define the biological object or process being measured and whether it is increased, decreased, or otherwise definably altered relative to a control state. For example “enzyme activity, decreased”, “hormone concentration, increased”, or “growth rate, decreased”, where the specific enzyme or hormone being measured is defined.

2C. Short Name

The KE short name should be a reasonable abbreviation of the KE title and is used in labelling this object throughout the AOP-Wiki. The short name should be less than 80 characters in length.

2D. Level of Biological Organisation

Structured terms, selected from a drop-down menu, are used to identify the level of biological organisation for each KE (e.g. molecular, cellular, organ). Note that KEs should be defined within a particular level of biological organisation. Only KERs should be used to transition from one level of organisation to another. Selection of the level of biological organisation defines which structured terms will be available to select when defining the Event Components (below).

2E. KE Components and Biological Context

Because one of the aims of the AOP-Wiki is to facilitate generation of AOP networks through the use of shared KE and KER elements, authors are strongly encouraged to define their KEs using a set of structured ontology terms (Event Components); in the absence of structured terms, the same KE could have a variety of titles. In order to make synonymous KEs more machine-readable, they should be defined by one or more “event components” consisting of a **biological process**, **object**, and **action** with each term originating from one of 22 biological

ontologies (Ives, et al., 2017). **Biological process** describes dynamics of the underlying biological system (e.g., receptor signalling). The biological **object** is the subject of the perturbation (e.g., a specific biological receptor that is activated or inhibited). **Action** represents the direction of perturbation of this system (generally increased or decreased; e.g., 'decreased' in the case of a receptor that is inhibited to indicate a decrease in the signalling by that receptor).

> ***Development tip 5– How specifically should my KE be defined*:** The following are some general recommendations and "rules of thumb" concerning how specifically to define a KE (see also Villeneuve et al. 2014a, b):
>
> - Define the KE with enough specificity that it is clear what to measure to determine the state of the KE. For example "histological changes" is too broad; "oocyte atresia" or "hyperplasia" would be better.
> - KEs should refer to/focus on a single measurable event within a specific biological level of organisation, rather than compounding events together. For example, it would be better to define a KE as "enzyme activity, increased" (if that can be measured), rather than "transcription and translation leading to enzyme activity, increased".
>
> The biological context of the KE (e.g., the tissue type/taxa/life stage/sex etc.) should only be restricted (e.g., "enzyme activity in liver, decreased" or "hormone concentration in females, increased") to the extent that function changes with context. If the function is equivalent in both sexes, do not restrict the context by sex. If the function is equivalent in all cell types, do not restrict to a specific cell type.

2F. Other AOPs that use this KE

All of the AOPs that are linked to this KE will automatically be listed in this subsection. This table can be particularly useful for identifying AOP networks which include the KE.

2G. KE Description

A description of the biological state being observed or measured, the biological compartment in which it is measured, and its general role in the biology should be provided. For example, the biological state being measured could be the activity of an enzyme, the expression of a gene or abundance of an mRNA transcript, the concentration of a hormone or protein, neuronal activity, heart rate, etc. The biological compartment may be a particular cell type, tissue, organ, fluid (e.g., plasma, cerebrospinal fluid), etc. The "role in the biology" could describe the reaction that an enzyme catalyses and the role of that reaction within a given metabolic pathway; the protein that a gene or mRNA transcript codes for and the function of that protein; the function of a hormone in a given target tissue, physiological function of an organ, etc. Care should be taken to avoid reference to other KEs, KERs or AOPs. Only describe this KE as a single isolated measurable event/state. This will ensure that the KE is modular and can be used in other AOPs, thereby facilitating construction of AOP networks. Additionally, avoid the use of semi-quantitative terms that suggest an undefined threshold (e.g., insufficient, inadequate, sustained). Quantitative understanding of the magnitude or duration of change in the KE required to impact a downstream event should be defined in the KER (see Section 3G), not in the KE description or title.

2H. How it is Measured or Detected

One of the primary considerations in evaluating AOPs is reliability and relevance of the methods used to measure the KEs. The aim of this section of the KE description is not to provide detailed protocols, but rather to capture, in a sentence or two, per method, the type(s) of measurements that can be employed to evaluate the KE and the relative level of scientific confidence in those measurements. These can range from citation of specific validated test guidelines, to citation of specific methods published in the peer reviewed literature, to outlines of a general protocol or approach (e.g., a protein may be measured by ELISA).

Key considerations regarding scientific confidence in the measurement approach include

whether the assay is fit for purpose, whether it provides a direct or indirect measure of the biological state in question, evidence that it is reproducible, and the extent to which it is accepted in the scientific and/or regulatory community. Information can be obtained from the OECD Test Guidelines website and the EURL ECVAM Database Service on Alternative Methods to Animal Experimentation (DB-ALM).

2I. Biological Domain of Applicability

The biological domain(s) of applicability of the KE in terms of sex, life-stage, taxa, and other aspects of biological context are defined in this section. In essence, the taxa/life-stage/sex applicability is defined based on the species or groups of organisms for which the measurements represented by the KEs can be made based on direct evidence from the literature (i.e., empirical domain of applicability) or based on one or more lines of scientific reasoning (i.e., biologically plausible domain of applicability) [see Development tip 6]. Defining the taxonomic, life stage and sex relevance of each KE helps to bound the domain of applicability of the AOP as a whole and provides an understanding of how broadly data represented by a KE measurement may be applied.

> ***Development tip 6 – Domain of applicability:*** When defining domain of applicability, it is useful to think about it in two ways
>
> **Empirical domain of applicability**: Species, sexes, life stages, for which there is already demonstrable evidence that the measurement can be made (KEs), the relationship applies (KERs) or the AOP in its entirety is relevant (AOPs).
>
> **Biologically plausible domain of applicability**: The broad range of species, sexes, life stages for which the measurement (KE), relationship (KER), or AOP is likely to apply based on scientific reasoning (i.e., molecular conservation of targets/pathways; phylogenetic releatedness; similarity in life history; analogy).
>
> Authors are encouraged to present both, and to clearly distinguish between the two based on the "evidence calls" made in the structured table and/or the explanatory text provided in the free text field.

As a general guide, whether defining the domain of applicability empirically or based on biological plausibility, there are two primary considerations for a KE:

1. Structure: Is there evidence that the biological object being measured/observed is present/conserved in the taxa/sex/life-stage of interest? Here biological object may refer to a protein, a cell type, an organ, etc.
2. Function: Is there evidence that the function of that biological object and the process being measured via the KE are conserved and relevant in the taxa/sex/life-stage of interest. Does it play the same role?

For example, if the KE involves binding to the estrogen receptor, but invertebrates lack a functional homolog of the estrogen receptor, one could reasonably conclude that the AOP is not relevant to invertebrates on the basis of a lack of conserved structure. Evidence supporting this biologically plausible taxonomic domain of applicability could be collected from bioinformatics approaches and existing toxicity data across species to support this broad extrapolation to all invertebrates. Depending on the evidence supporting the taxonomic domain of applicability, the specific (common or Latin) species name or taxonomic group (e.g., class, order, family) may be reported with the appropriate NCBI taxonomy ID in the "Taxonomic Applicability" table of the AOP-Wiki. Likewise, if the KE involves a measurement in ovarian tissue, its applicability domain in terms of sex would be restricted to females. Such information would be captured in the "Sex Applicability" table of the AOP-Wiki using predefined terms like: male, female, mixed, asexual, hermaphrodite, or unspecific. If a KE involved altered organogenesis (e.g., heart formation), the KE would only be relevant to the life-stage during which the heart is actually formed, not adult life stages in which organ development has already

completed. Life-stage can be described in the "Life Stage" table of the AOP-Wiki by selecting from structured ontology terms. If an applicable life-stage term cannot be found, new terms may be added on request by the AOP-Wiki administrators.

Biological domain of applicability is defined in the AOP-KB using a combination of structured fields and free text. Selection of structured terms to describe the applicability domain can aid AOP network construction as well as facilitating other types of computational processing and searching of information captured in the AOP-KB.

When the developer selects structured ontology terms to help define the domain of applicability of the KE, there is also an option to make evidence calls related to applicability of the specific KE for that category term. These calls should be based on expert knowledge of the biology and the extent of supporting evidence. Recommendations for these calls are:

- Low: With the understanding that by definition a KE must be measurable in the species/taxonomic group/lifestage/sex defined, no such measurements have been reported or shown experimentally *in vitro* or *in vivo* to date; however, there are one or more scientifically-based lines of evidence suggesting that measurement could plausibly be made (e.g., in silico or bioinformatic evidence of protein or pathway conservation).
- Moderate: The measurement associated with the KE can plausibly be made for the species/taxonomic group/lifestage/sex, and there is at least some supporting *in vitro* or *in vivo* experimental evidence, although though it may not involve direct measurement of the KE.
- High: The measurement associated with the KE has been made repeatedly *in vitro* or *in vivo* and/or with multiple orthogonal methods for the species/taxonomic group/lifestage/sex.

i. Taxonomic Applicability

Latin or common names of a species or broader taxonomic grouping (e.g., class, order, family) can be selected from an ontology. In many cases, individual species identified in these structured fields will be those for which the evidence used in constructing the AOP was strongest in relation to this KE.

ii. Life Stage Applicability

The structured ontology terms for life-stage are more comprehensive than those for taxa, but may still require further description/development and explanation in the free text section.

iii. Sex Applicability

The authors must select from one of the following: Male, female, mixed, asexual, third gender, hermaphrodite, or unspecific.

iv. Evidence for Biological Domain of Applicability

This free text section should be used to elaborate on the scientific basis for the indicated domains of applicability and the WoE calls (if provided). While structured terms may be selected to define the taxonomic, life stage and sex applicability (see structured applicability terms, above) of the KE, the structured terms may not adequately reflect or capture the overall biological applicability domain (particularly with regard to taxa). Likewise, the structured terms do not provide an explanation or rationale for the selection. The free-text section on evidence for taxonomic, life stage, and sex applicability can be used to elaborate on why the specific structured terms were selected, and provide supporting evidence, references and background information. This information should also indicate the type of data used as evidence (e.g., in silico, in vitro, in vivo).

2J. AO-Specific Content

An AO is a specialised KE that represents an adverse outcome of regulatory significance, ("apical endpoint"). For KEs that are designated as an AO, one additional field of information (regulatory significance of the AO) should be completed, to the extent feasible. If the KE is being described is not an AO, simply indicate "not an AO" in this section.

Regulatory Significance of the AO

A key criterion for defining an AO is its relevance for regulatory decision-making (i.e., it corresponds to an accepted protection goal or common apical endpoint in an established regulatory guideline study). For example, in humans this may constitute increased risk of disease-related pathology in a particular organ or organ system in an individual or in either the entire or a specified subset of the population. In wildlife, this will most often be an outcome of demographic significance, e.g., population sustainability. In addition to describing the biological state associated with the AO, how it can be measured, and its taxonomic, life stage, and sex applicability, it is useful to describe regulatory examples using this AO.

2K. References

List of the literature that was cited for this KE description. References should either be numbered [#], and cited by number, or cited in (Author, Year) style at locations on the Event page corresponding to the statement(s) they support. Ideally, the list of references, should conform, with the OECD Style Guide (https://www.oecd.org/about/publishing/OECD-Style-Guide-Third-Edition.pdf) (OECD, 2015).

SECTION 3 – KER DESCRIPTIONS

The utility of AOPs for regulatory application is defined, to a large extent, by the confidence and precision with which they facilitate extrapolation of data measured at low levels of biological organisation to predicted outcomes at higher levels of organisation and the extent to which they can link biological effect measurements to their specific causes. Within the AOP framework, the predictive relationships that facilitate extrapolation are represented by the KERs. Consequently, the overall WoE for an AOP is a reflection in part, of the level of confidence in the underlying series of KERs it encompasses. Evidence related to determination of confidence in the supporting data for the KER as part of the AOP is included here. The confidence in the overall AOP pathway is considered in Section 4, taking into account the KER specific evidence and patterns of support across all levels of biological organization in the AOP.

Describing the KERs in an AOP involves assembling and organising the types of information and evidence that defines the scientific basis for inferring the probable change in, or state of, a downstream KE from the known or measured state of an upstream KE. Before describing a KER, developers should carefully consider the following:

KERs are always described in the form of a directed relationship (one-way arrow) linking an upstream "causing" event to a downstream "responding" event. The pair of KEs linked via a KER may either be adjacent to one another in the sequence of KEs that define a given AOP, or non-adjacent (Figure 5). Regardless of adjacency, one event is always positioned upstream of the other. By convention (and for clarity), KERs linking adjacent KEs in an AOP are represented using solid arrows, while KERs that link KEs that are not adjacent to one another in sequence

are linked via dashed arrows (e.g., Figure 5). This is a graphical convention only which has no bearing on the type of content to include in the KER description.

A KER description must be created for each adjacent upstream-downstream pair of KEs in the pathway. Graphically speaking, there should always be at least one solid arrow path connecting each KE in the pathway into a sequence. There should be no KEs that are unconnected or are only connected via a non-adjacent path (represented as a dashed arrow) only.

Inclusion and description of non-adjacent KERs within an AOP can be particularly useful for assembling evidence supporting the AOP and in the consideration of the overall support across the entire AOP (section 4). For example, some KE measurements may be fairly difficult to make, such that they are rarely made in routine studies. While there may be sufficient data or plausibility to establish an intermediate KE as part of the AOP, much of the available WoE may ignore or "leap over" that particular KE. Including KER descriptions for non-adjacent KE pairs allows the WoE for these relationships to be readily described and linked to other AOPs without compromising the principle of modularity with regard to the KER descriptions. With this in mind, the upstream-downstream pair of KEs linked via a KER may be adjacent in one AOP and non-adjacent in another (Figure 6).

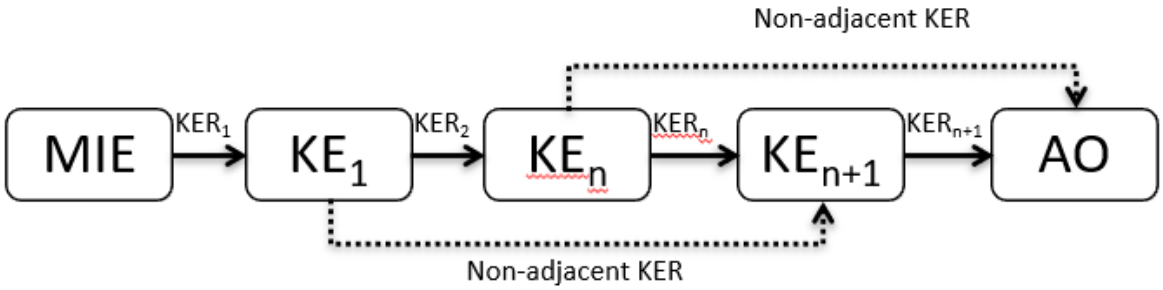

Figure 5. Generic AOP diagram illustrating the graphical convention for depicting KERs linking adjacent (solid arrow) versus non-adjacent (dashed arrow) upstream-downstream KE pairs within an AOP. Regardless of adjacency, each KER represents a predictive relationship between a pair of KEs and can be supported by WoE.

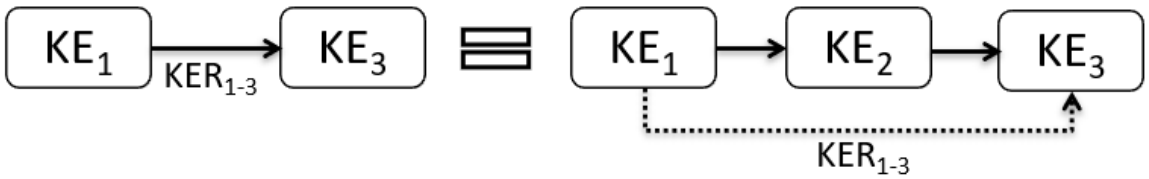

Figure 6. Graphical depiction of the modular functionality of KERs connecting KE1 to KE3. The content of KER1-3 is identical despite the fact that the KE1 and KE3 are adjacent in one AOP and non-adjacent in the other.

Overall, the subsections of the KER descriptions are intended to aid the user in collecting relevant information that will support evaluation of the level of confidence in each KER, which in turn contributes to the assessment of the WoE of the AOP overall (section 4).

3A. Relationship ID

When a KER is created, an ID number is automatically assigned to it (Relationship: ###). This number is used for tracking the KER in the AOP-KB and corresponds with a unique URL of the form https://aopwiki.org/relationships/###.

3B. KER Title

All KER titles take the form "upstream KE leads to downstream KE". KER titles are generated automatically by selecting an upstream KE and downstream KE to link in the AOP-Wiki (Figure 7).

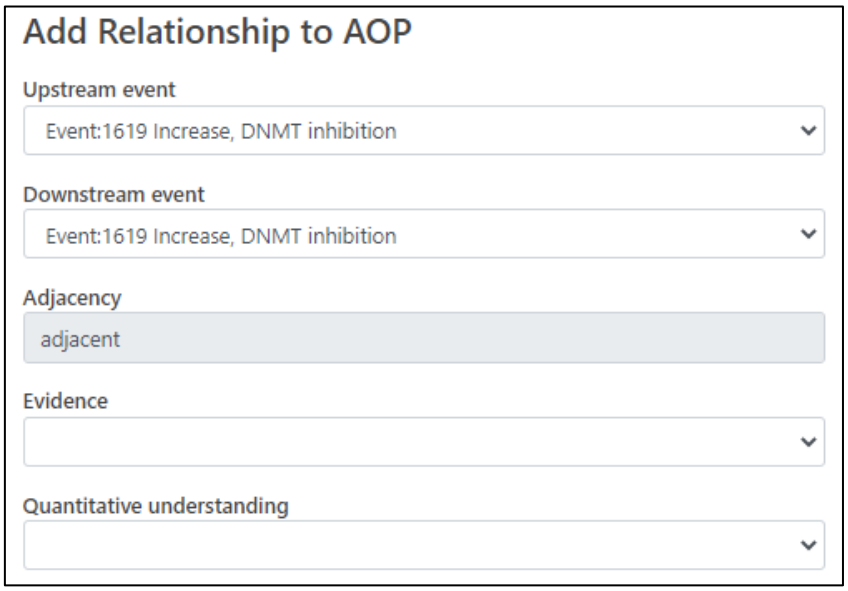

Figure 7. Add Relationship dialog from AOP-Wiki. Note, user will select KEs from a drop-down menu of options, therefore the KER title is created automatically. This also means that the KEs must be created before a KER can be defined.

3C. AOPs Referencing Relationship

All of the AOPs that are linked to this KER will automatically be listed in this subsection.

3D. Biological Domain of Applicability

Developers have the option to select one or more structured terms that help to define the biological applicability domain of the KER. As a rule, the biological domain of applicability of a KER can never be broader than the more restrictive of the two KEs it links together. For example, if the upstream KE is relevant to all vertebrates but the downstream KE is relevant only to sexually mature, egg-laying female vertebrates, the KER would be relevant to sexually mature egg-laying female vertebrates. This concept applies whether considering the empirical domain of applicability, or the biologically plausible domain of applicability and once again authors should clearly indicate both.

Thus, the biological applicability domains of the two KEs being linked is a strong determinant of the biological domain of applicability of a KER (Figure 8).

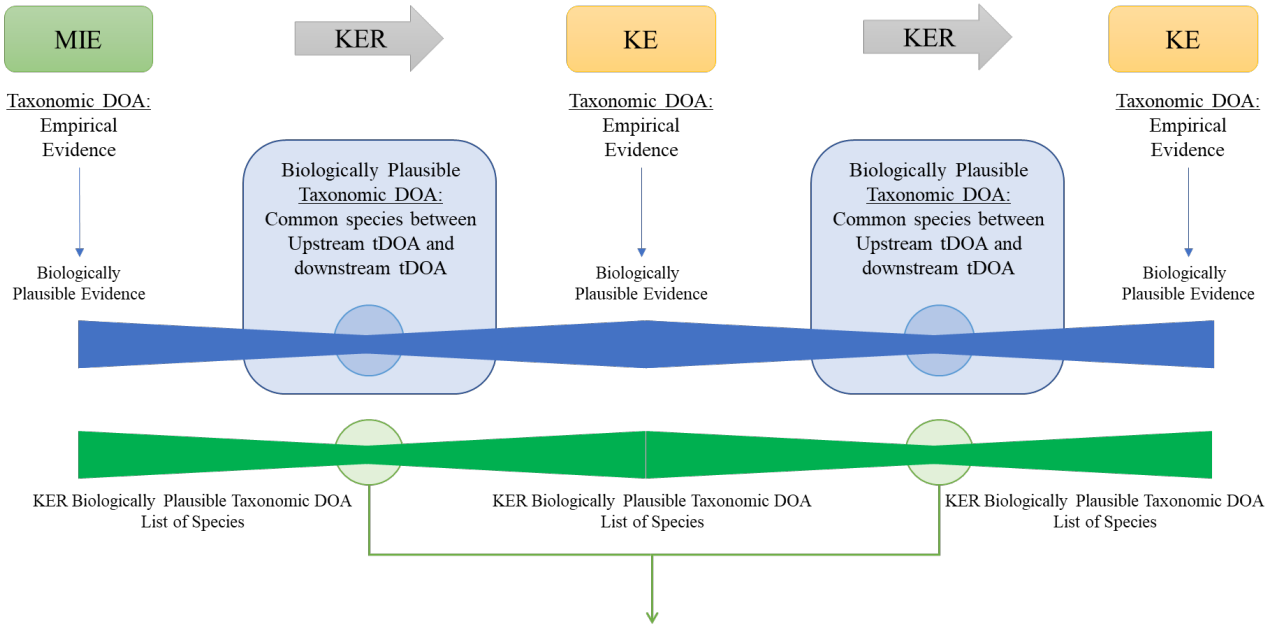

Figure 8. Example for determining the taxonomic domain of applicability (tDOA) considering both the empirical evidence and biologically plausible evidence and combining upstream KE and downstream KE tDOA to determine KER tDOA. Further, considering the KER tDOAs across the AOP the most restrictive tDOA across all KERs defines the tDOA for the AOP. The blue horizontal line considers each KE to define the biologically plausible tDOA of the KER, whereas the green horizontal line considers each KER to define the biologically plausible tDOA for the entire AOP. Figure modified from Jensen et al. 2022.

However, in some cases, the biological applicability domain of the KER may be even more restrictive. This is because in addition to structural and functional conservation, the KER also considers the conservation of a biological relationship between two KEs. The three considerations that generally guide definition of the biological domain of applicability are thus:

1. Structure: Is there evidence that the biological object(s) being measured/observed in the context of the two KEs being linked present/conserved in the taxa/sex/life-stage of interest?

2. Function: Is there evidence that the functions of those biological objects and the processes being measured in the two KEs are conserved and relevant in the taxa/sex/life-stage of interest? Does the object/process play the same role in both KEs?

3. Regulation: Is there evidence that the regulation of the KEdownstream by KEupstream is conserved and relevant in the taxa/sex/life-stage of interest?

Selection of structured terms to describe the biological domain of applicability can aid AOP network construction as well as facilitating other types of computational processing and searching of information captured in the AOP-Wiki.

Upon selection of structured biological applicability domain terms, developers have the option to classify the extent of the supporting evidence for the terms they have selected:

- **Low** the relationship is biologically plausible, but has not been shown experimentally *in vitro* or *in vivo* in this species/taxonomic group/lifestage/sex; evidence may be computationally derived by models or other available tools for evaluating structural and functional conservation (e.g., in silico or bioinformatic evidence of protein or pathway conservation).
- **Moderate** the relationship is biologically plausible, and there is some limited supporting *in vitro* and/or *in vivo* experimental evidence in the species/taxonomic group/lifestage/sex of interest; computationally derived data to support the biologcially plausible domain of applicability could be included as evidence toward structural conservation and used for extrapolation.
- **High** the relationship is biologically plausible, and there is considerable supporting evidence in the species/taxonomic group/lifestage/sex, including evidence of temporal, dose-response, and/or incidence concordance between the two KEs for the group in question.

i. Taxonomic Applicability

Authors can indicate the relevant taxa for this KER in this subsection. The process is similar to that described for KEs (Section 2).

ii. Life Stage Applicability

Authors can indicate the relevant life stage for this KER in this subsection. The process is similar to that described for KEs (Section 2).

iii. Sex Applicability

Authors can indicate the relevant sex for this KER in this subsection. The process is similar to that described for KEs (Section 2).

iv. Evidence Supporting the Biological Domain of Applicability

As for the KEs, there is also a free-text section of the KER description that the developer can use to explain his/her rationale for the structured terms selected with regard to taxonomic, life stage, or sex applicability, or provide a more exact description of the applicability domain than may be feasible using standardised terms. Developers are also encouraged to distinguish the empirical domain of applicability from the more expansive biologically plausible domain of applicability (see *Development tip 5*). Here developers can indicate what type(s) of evidence were used to support the domain of applicability (e.g., *in silico*, *in vitro*, *in vivo*) and cite the methods if relevant.

3E. KER Description

Provide a brief, descriptive summation of the KER. While the title itself is fairly descriptive, this section can provide details that are not inherent in the description of the KEs themselves (see Section 2, recommendations regarding number of KEs to include). For example, if the upstream KE was antagonism of a specific receptor, the description could stipulate that "persistent antagonism of the receptor for a period of days" will trigger the downstream KE. Shorter term antagonism of the same receptor (i.e., same upstream KE) may trigger a different downstream KE, and thus would be described in a different KER. This description section can be viewed as providing the increased specificity in the nature of upstream perturbation (KEupstream) that leads to a particular downstream perturbation (KEdownstream), while allowing the KE descriptions to remain generalised so they can be linked to different AOPs. The description is also intended to provide a concise overview for readers who may want a brief summation, without needing to read through the detailed support for the relationship (covered below). Care should be taken to avoid reference to other KEs that are not part of this KER, other KERs or other AOPs. This will ensure that the KER is modular and can be used by other AOPs.

3F. Evidence Collection Strategy

Include a description of the approach for identification and assembly of the evidence base for the KER. For the literature searches and surveys, include, for example:

i. Sources and dates of information consulted including expert knowledge, databases searched and associated search terms/strings,
ii. Study screening criteria and methodology (e.g., inclusion/exclusion criteria, specialized software tools, number of reviewers); any constraints on the search.
iii. Study quality assessment considerations including links to existing resources (e.g., existing tools applied)
iii. Data extraction strategy, specialized software tools and/or data management strategy, and
iv. Links to any repositories/databases of relevant references

Tabular summaries and links to relevant supporting documentation are encouraged, wherever possible.

Alternatives to literature search-based approaches include, but are not limited to, novel experimentation, application of biologically-based models, identification of sources of canonical knowledge, etc.

3G. Evidence Supporting this KER

Assembly and description of the scientific evidence supporting KERs in an AOP is an important step in the AOP development process that sets the stage for overall assessment of the AOP relevant to regulatory application (Section 4). To do this, biological plausibility, empirical support, and the current quantitative understanding of the KER are evaluated with regard to the predictive relationships/associations between defined pairs of KEs as a basis for considering WoE (Section 4). In addition**,** uncertainties and inconsistencies are considered.

i. Biological Plausibility

Define, in free text, the biological rationale for a connection between KEupstream and KEdownstream. What are the structural or functional relationships between the Kes (see Annex 1)? For example, there is a functional relationship between an enzyme's activity and the product of a reaction it catalyses.

Contextual citation of supporting references should be included. However, it is recognised that there may be cases where the biological relationship between two KEs is very well established, to the extent that it is widely accepted and consistently supported by so much literature that it is unnecessary and impractical to cite the relevant primary literature (i.e.,canonical knowledge). Citation of review articles or other secondary sources, like text books, may be reasonable in such cases. The primary intent is to provide scientifically credible support for the structural and/or functional relationship between the pair of KEs if one is known.

In general, the structural and/or functional relationship supporting biological plausibility is based on understanding of "normal" biological function, rather than response to a specific stressor. The description of biological plausibility can also incorporate additional mechanistic detail that helps inform the relationship between KEs, but is not practical/pragmatic to represent as separate KEs due to the difficulty or relative infrequency with which it is likely to be measured. For example, in the case of G protein coupled receptor activation (KEupstream) leading to increased activity of a specific enzyme (KEdownstream), there may be numerous mechanistic steps between these KEs (e.g., alterations in signal transduction pathways, transcriptional regulation, post-translational modifications, etc.). These underlying details, if known, can be captured in the description of biological plausibility (if desired) rather than represented as independent KEs. The KER descriptions are the appropriate place for"embedding" this type of biological detail without compromising the reusability of the KE

descriptions within the AOP-Wiki. However, it should be kept in mind that added detail should only be included to the extent that it enhances the predictive utility of the AOP for regulatory application. Detail may be particularly useful in considering the differences across taxonomic groups or species that may dictate the broad utility of the AOP (i.e., the taxonomic domain of applicability). In part, the AOP is intended to filter through much of the mechanistic detail to focus on what important causal events for the adverse outcome have predictive value for regulatory application. Thus, efforts should be made to keep the descriptions focused and concise.

ii. Empirical Evidence

In this section authors are encouraged to cite specific evidence relevant to assessment of changes in the upstream KE (KEupstream) leading to, or being associated with, a predictable subsequent change in the downstream KE (KEdownstream).

In particular, it is useful to cite direct evidence showing that stressors that perturb KEupstream also perturb KEdownstream. Because this section of the KER description cites evidence from specific studies, it is also helpful to provide as much detail as possible about the toxicological and biological context in which the measurements were made. While the KER itself is not intended to be stressor-specific, this information addresses whether supporting data on quantitative patterns of relationships between key events is consistent with what's expected, if the KER is operative. Expected patterns are that the upstream KE is impacted at doses/concentrations of the stressor that are equal to or lower than those that impact the downstream KE (dose concordance; Figure 9), that at any given dose of stressor, the upstream is impacted earlier in the time-course of exposure than the downstream event (temporal concordance; Figure 9), and likewise for any given dose and duration of exposure to the stressor, the upstream event is observed in an equal to or greater proportion of the sample population than the downstream event (incidence concordance; Figure 9). Deviations from these expected patterns may be due to factors such as experimental design, the relative sensitivity of methods for measuring KEs, and other factors; experimental details that could influence apparent concordance or lack thereof, should therefore be considered when assembling and presenting evidence.

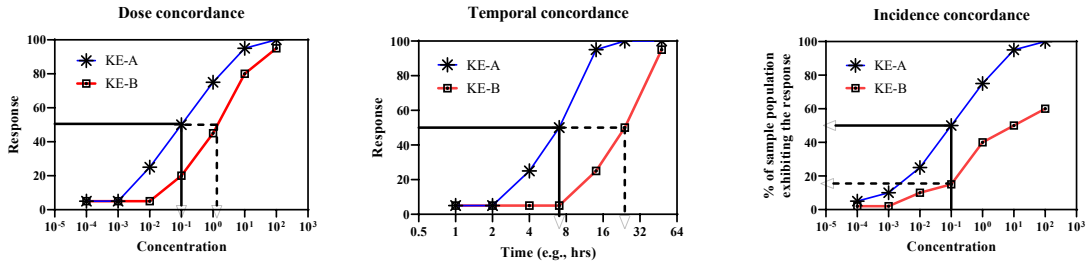

Figure 9. Examples of dose concordance, temporal concordance, and incidence concordance. Note that dose concordance and temporal concordance are comparing the relative dose or time at which a defined level of response is observed for KE_A compared to KE_B. Incidence concordance compared the fraction of the population impacted at the same dose and time point for KE_A versus KE_B.

The consideration of empirical support in the form of bulleted lists or tables that include a short description of the nature of the observed empirical support along with the corresponding reference(s) is preferred as a basis to consider whether available data consistently supports expected patterns. An example is provided below (Table 3). However, authors are free to modify the format to best suit their approach to support the consideration of weight of evidence for the pathway. To the extent possible, entries in the table should be based on benchmark

doses (BMDs; for additional background see US EPA, 2012, EFSA 2022) to facilitate comparative assessment of effect, thus normalizing for groupsizes and dose spacing.

Table 3. Example of an empirical evidence table assembled for a KER[1].

Species, life-stage, sex tested	Stressor(s)	Upstream Effect (Y/N)	Downstream Effect (Y/N)	Effect on Upstream Event (descriptive)	Effect on Downstream Event (descriptive)	Citation
Adult, female, rainbow trout	Gemfibrozil	Y	Y	Benchmark dose (BMD) 15 µg/L	BMD 45 µg/L	Smith et al. 1978
Adult, F, Sprague Dawley rat	Low fat diet	Y	N	Significant decrease at 100 mg/kg/day, after 3 days	No effect at concentrations up to 2 g/kg/d, fed up to 10 days	Zonk 2018
Juvenile, M, mouse	Clofibric acid	N	Y	BMD 45 mg/kg/d, measured 5 d post-injection	BMD 5 mg/kg/d, measured 5 d post-injection	Doe et al. 2012
Larval zebrafish	UV radiation @ UV index = 90	Y	Y	Significant decrease in 80% of sampled population after 48 h	Significant increase in 22% of sampled population after 96 h	Lee et al. 1994

[1] *Entries in this table are for illustrative purposes only. They do not refer to results from real studies. Any resemblance to existing scientific results or authors is coincidental.*

a. Dose Concordance

In the case of dose-response concordance, the aim is not to consider dose-dependence of a single KE in the pair, but rather to assess the extent of the evidence that KE upstream is generally impacted at doses (or stressor severities) equal to or less than those at which KE downstream is impacted (data row 2 of Table 3 shows an example of dose concordance; row 3 does not follow the expected pattern for dose concordance).

b. Temporal Concordance

In the case of temporal concordance, it is desirable to assemble evidence relevant to assessing whether effects on KE upstream are observed earlier in a time-course than effects on the downstream KE (data row 3 of Table 3 shows an example of temporal concordance, as well as dose concordance).

c. Incidence Concordance

In the case of incidence concordance, evidence should be assembled that addresses whether, at an equivalent dose or stressor severity, KEupstream occurs more frequently than Kedownstream (data row 4 of Table 3 shows an example of incidence concordance, as well as temporal concordance).

d. Other Evidence (optional)

Although evidence that demonstrates dose, temporal or incidence concordance is preferred, other evidence that empirically supports the relations that a sufficient change in KEupstream will lead to a change in KEdownstream, but do not fall into the above three categories, can

be cited in this subsection.

iii. Uncertainties and Inconsistencies

In addition to outlining the evidence supporting a particular linkage, it is also important to identify inconsistencies or uncertainties in the relationship. This could include, for example, empirical evidence showing changes in KEupstream that did not elicit alterations in KEdownstream. It could also include descriptions of gaps in biological understanding that lend to uncertainties in understanding of the exact nature of the structural or functional relationship between the two KEs. Additionally, while there are expected patterns of concordance that support a causal linkage between the KEs in the pair, it is also helpful to identify experimental details that may explain apparent deviations from the expected patterns of concordance. An example of this would be a case where methods for measuring the upstream KE are relatively insensitive compared to those for measuring the downstream KE, leading to the appearance of dose-response or incidence discordance that is simply an artefact of the measurement techniques employed. In this regard, when assembling information from multiple disparate studies, it is important to capture variables that directly influence how well concordance can be assessed (i.e., information regarding the doses tested in various experiments and the time-points at which various KE measurements were made). Identification of uncertainties and inconsistencies contributes to evaluation of the overall WoE supporting the AOPs that contain a given KER (see Section 4), and to the identification of research gaps that warrant investigation.

Given that AOPs are intended to support regulatory applications, AOP developers should focus on those inconsistencies or gaps that would have a direct bearing or impact on the confidence in the KER and its use as part of an AOP for inference or extrapolation in a regulatory setting. Uncertainties that would have little impact on regulatory application do not need to be described. In general, this section details evidence that may raise questions regarding the overall validity and predictive utility of the KER (including consideration of both biological plausibility and empirical support). It also contributes, along with other elements, to the overall evaluation of the WoE for the KER (see, Section 4).

3H. Known Modulating Factors

This section presents information regarding modulating factors/variables known to alter quantitative aspects of the response-response function that describes the relationship between the two KEs (for example, an iodine deficient diet causes a significant increase in the sensitivity of the downstream event to changes in the upstream event [alters the slope of the relationship]; a particular genotype doubles the sensitivity of KEdownstream to changes in KEupstream). Information on these known modulating factors should be listed in this subsection, along with relevant information regarding the manner in which the modulating factor alters the relationship (if known). Note: this section should focus on those modulating factors for which solid evidence supported by relevant data and literature are available. It should NOT list all possible/plausible modulating factors. In this regard, it is useful to bear in mind that many risk assessments conducted through conventional apical guideline testing-based approaches generally consider few if any modulating factors.

It is recommended that information regarding known modulating factors be captured in a tabular format (Table 4), providing the following information about each:

- What it is – the modulating factor for which there is solid evidence that it influences this KER.
- Details of the modulating factor – specify which features (classes or subsets?) of this modulating factor are relevant for this KER.
- Describe the known effect(s) of the modulating factor on the KER.
    - **i.** E.g., increases magnitude of effect on downstream KE by two-fold
    - **ii.** E.g., reduces the probability of effect on the downstream event by 40%
    - **iii.** E.g., delays onset of the downstream event by 12-18 h

iv. E.g., increases sensitivity to the upstream event by a factor of four

- Reference(s) – provide one or more references that provide supporting scientific evidence that establishes the effect of the modulating factor on the KER.

Table 4. Recommended tabular format for capturing information regarding known modulating factors[1].

Modulating Factors	MF details	Effects on the KER	References
Age	>55 years old (human)	Sensitivity of downstream event to change in upstream event increased by factor of 4	Smith et al. 1978
Genotype	BRCA1 truncation mutation in nucleotides 2401-4109)	Probability of downstream event increased by 40%	Zonk 2018
Diet	Iodine deficient	Delays onset of downstream effect by 5-10 d	Doe et al. 2012
Disease state	Type 2 diabetes	Increases risk of downstream event by 10 fold	Lee et al. 1994
Previous exposure	Within 3 years of Covid 19 infection	Magnitude of effect on downstream event increased 2-fold Delay	Walla Walla and Grant, 2022

[1] *Entries in this table are for illustrative purposes only. They do not refer to results from real studies. Any resemblance to existing scientific results or authors is coincidental.*

3I. Quantitative Understanding

The quantitative understanding section of the KER description is intended to capture information that helps to define how much change in the upstream KE, and/or for how long, is needed to elicit a detectable and defined change in the downstream KE. While empirical support (see previous section F Evidence Supporting this KER) addresses whether data on the relationship between the two KEs are consistent with the patterns that are expected if the upstream event is causing the downstream event, the quantitative understanding section helps to define the precision with which the state of the downstream KE can be predicted from knowledge of the state of the upstream KE. The higher the confidence in empirical support for a KER, the greater the likelihood that the response response relationship can be quantified. These quantitative relationships may be defined in terms of correlations, response-response relationships, dose-dependent transitions or points of departure (i.e., a threshold of change in KEupstream needed to elicit a change in KEdownstream), etc. They may take the form of simple mathematical equations or sophisticated biologically-based computational models that consider other modulating factors such as compensatory responses, or interactions with other biological or environmental variables. Regardless of form, the idea is to briefly describe what is known regarding the quantitative relationship between the KEs and cite appropriate literature that defines those relationships and/or provides support for them.

Data that confer quantitative understanding of a KER are not necessarily independent of those addressing other weight of evidence considerations. Rather, the quantitative understanding section collects additional detail about the nature of the quantitative relationship generally from the same studies used to establish empirical support. These further details are intended to support quantitative prediction of the probability or magnitude of change in KEdownstream based on a known state of KEupstream. For transparency, the toxicological and biological context in which the quantitative relationships were defined should be indicated within the description. The ultimate goal is to identify quantitative relationships that generalise across the entire applicability domain of the two KEs being linked via the KER.

Based on recommendations from workshops held in September 2015 (Wittwehr et al. 2016) and April 2017 (LaLone et al. 2017), description of the quantitative understanding of the KER has been organised into subsections in order to more consistently capture information useful for both quantitative AOP and AOP network applications. As with other areas of the AOP descriptions, authors are encouraged to complete the subsections to the extent feasible, but it is recognized that supporting information may not be adequate to address all.

i. Response-response relationship

This subsection should be used to define sources of data that define the response-response relationships between the KEs. A response-response relationship is a mathematical function that describes the magnitude, probability, or severity of change in the downstream KE (B) as a function of the measured (or predicted) state of the upstream KE (A). Information regarding the general form of the relationship (e.g., linear, exponential, sigmoidal, threshold, etc.) should be captured if possible. If there are specific mathematical functions or computational models relevant to the KER in question that have been defined, those should also be cited and/or described where possible, along with information concerning the approximate range of certainty with which the state of the KEdownstream can be predicted based on the measured state of the KEupstream (i.e., can it be predicted within a factor of two, or within three orders of magnitude?). For example, a regression equation may reasonably describe the response-response relationship between the two KERs, but that relationship may have only been validated/tested in a single species under steady state exposure conditions. It is important to note such uncertainties.

ii. Time-scale

This sub-section should be used to provide information regarding the approximate time-scale of the changes in KEdownstream relative to changes in KEupstream (i.e., do effects on KEdownstream lag those on KEupstream by seconds, minutes, hours, or days?). This can be useful information both in terms of modelling the KER, as well as for analysing the critical or dominant paths through an AOP network (e.g., identification of an AO that could kill an organism in a matter of hours will generally be of higher priority than other potential AOs that take weeks or months to develop). Identification of time-scale can also aid the assessment of temporal concordance. For example, for a KER that operates on a time-scale of days, measurement of both KEs after just hours of exposure in a short-term experiment could lead to incorrect conclusions regarding dose-response or temporal concordance if the time-scale of the upstream to downstream transition was not considered.

iii. Known Feedback loops influencing this KER

KERs are depicted in a manner that suggests that the upstream event is independent of the downstream event. However, in biological systems, feedback relationships are common. This subsection should define whether there are known positive or negative feedback loops involved and what is understood about their time-course and homeostatic limits. In some cases where feedback processes are measurable and causally linked to the outcome, they may be represented as KEs (see development tip 5). However, in most cases these features are expected to predominantly influence the shape of the response-response and time-

course, behaviours between selected KEs (i.e., the KER). For example, if a feedback loop acts as an auto-regulatory loop designed to maintain a homeostatic range of concentrations between some upper and lower limit, the feedback loop will directly shape the response-response relationship between the KEs. It is recommended that an annotation indicating a positive or negative feedback loop (Figure 10) in a KER be added to the graphical representation, and that details be provided in this subsection of the KER description.

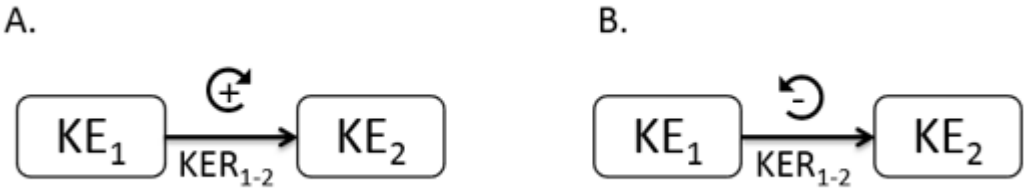

Figure 10. Recommended graphical annotation to indicate that a known (A) positive feedback (i.e., feedforward) or (B) negative feedback loop is involved in the transition from one KE to the next in the AOP. Note, this is an optional annotation. See Development tip 7 for more information on describing positive and negative feedback processes using the AOP framework.

Development tip 7 – Capturing information on positive or negative feedback loops.
Ways to capture/represent known positive or negative feedback loops have emerged as a frequently asked question in relation to use of the AOP framework. Thus, a few general guidelines are provided here.

- In cases where feedback loops play a direct causal role in the progression of a biological perturbation leading to an AO, they can be included as KEs as long as they are measurable. For example, for an AOP in which a negative feedback process results in decreased hormone signalling that leads to the AO, a measurable event indicative of or involved in the activation of the negative feedback could be included as a KE.
- In cases where a feedback loop may act as a key compensatory or adaptive mechanism that dictates how severely the KEupstream needs to be impacted in order to effect the KEdownstream, but does not play a direct causal role in the AOP (other than defining the relevant point of departure), the feedback should not be included as a separate KE. Rather it should be detailed as part of the quantitative understanding section of the KER description. In the user supplied graphical representation, a forward or backward looping symbol could be added above the arrow linking the two KEs to indicate that a known positive or negative feedback loop is involved in the transition (Figure 10B).
- In cases where two measurable KEs in an AOP are part of a positive feedback loop, it can be challenging to define which should be upstream and which downstream, as they are amplifying or altering one another in a cycle. A two headed arrow is undesirable as it can incorrectly suggest that the AOP is reversible. However, in practice an AOP with a positive feedback loop could be accurately represented as two different AOPs in the AOP-Wiki, in which the KEs involved in the positive feedback are presented in either order. This effectively creates a bi-directional arrow when the AOP network is assembled. Rather than creating two nearly identical AOP pages with the KE order reversed for each, the current recommendation is to select either order for the KEs and connect them with a unidirectional arrow, but add a forward looping symbol above the arrow in the user-supplied graphical representation to indicate that a known feedforward loop is involved (Figure 10A).

iv. Classification of quantitative understanding

To aid in overall assessment of the AOP and whether it is fit-for-purpose for various applications, developers are also asked to classify the extent of quantitative understanding of the KER as low, moderate, or high, taking into account the extent of data and resulting confidence in empirical

support, but also the extent to which quantitative impact of relevant modulating factors is understood. General guidance for classification of the level of quantitative understanding of a KER as low, moderate, or high (Annex 2) is based on several key considerations:

- The accuracy and precision with which a change in KEdownstream can be predicted based on KEupstream.
- The precision with which uncertainty in the prediction of KEdownstream can be quantified.
- The extent to which known modulating factors or feedback mechanisms are accounted for.
- The extent to which the relationships described can be reliably generalised across the biological applicability domain of the KER.

3J. References

List of the literature that was cited for this KER description using the appropriate format. Ideally, the list of references, should conform, with the OECD Style Guide (https://www.oecd.org/about/publishing/OECD-Style-Guide-Third-Edition.pdf) (OECD, 2015).

SECTION 4 – OVERALL ASSESSMENT OF THE AOP

This section addresses the relevant biological domain of applicability of the AOP as a whole (i.e., in terms of taxa, sex, life stage, etc.) and WoE for the overall AOP. Both are critical for determining the AOP's fit-for-purpose for various applications. This overall assessment is captured on the lower portion of the AOP pages within the AOP-Wiki. **The goal of the overall assessment is not to reproduce or reiterate all the content assembled as part of sections 1-3, but rather to provide a high level synthesis and overview of the relative confidence in the AOP and any significant gaps or weaknesses .** While description and evaluation of modular components facilitate development through sharing, regulatory applications, such as integrated approaches to testing and assessment and stressor specific mode of action evaluation, require integrated, pathway-level, analyses. Assimilation and assessment of the extent to which experimental data support expected patterns across all the KERs for the AOP informs relative confidence relevant to consideration of its suitability for specific regulatory applications. For example, the confidence required for prioritizing testing is normally less than that for screening assessment or full assessment to inform risk management.

Determination of confidence in the overall AOP is based on the biological plausibility, empirical support, and extent of quantitative understanding for the KERs (Section 3) and the evidence supporting essentiality of the KEs.
Assessment of the AOP is organised into a number of steps. Guiding questions that inform evaluation at each step are included in Annex 1. The questions are designed to facilitate assignment of categories of high, moderate, or low confidence for each consideration. While it is not necessary to repeat lengthy text that appears elsewhere in the AOP description (or related KE and KER descriptions), a brief explanation or rationale for the selection of high, moderate, or low confidence should be made, based on the guiding questions detailed below.

4A. Define the Biological Domain of Applicability of the AOP

The relevant biological domain(s) of applicability in terms of sex, life-stage, taxa, and other aspects of biological context are defined in this section. Biological domain of applicability is informed by the "Description" and "Biological Domain of Applicability" sections of each KE and KER description (see sections 2G and 3E for details). In essence the taxa/life-stage/sex applicability is defined based on the groups of organisms for which the measurements represented by the KEs are relevant and the structural, functional, and regulatory relationships represented by the KERs are operative.

The relevant biological domain of applicability, including the biologically plausible domain of applicability of the AOP as a whole will nearly always be defined based on the most narrowly restricted of its KEs and KERs. For example, if most of the KEs apply to either sex, but one is relevant to females only, the biological domain of applicability of the AOP as a whole would be limited to females. While much of the detail defining the domain of applicability may be found in the individual KE and KER descriptions, the rationale for defining the relevant biological domain of applicability of the overall AOP should be briefly summarised on the AOP page.

4B. Assess the Essentiality of All KEs

An important aspect of assessing an AOP is evaluating the essentiality of its KEs. This normally entails assessment of the impact of manipulating a given KE (e.g., experimentally blocking or exacerbating the event) on the downstream sequence of KEs defined for the AOP. Consequently, evidence supporting essentiality is collated on the AOP page, rather than on the independent KE pages that are as stand-alone modular units that do not reference other KEs in the sequence. That said, such evidence can also be captured through the description of adjacent and non-adjacent KERs.

The nature of experimental evidence that is relevant to assessing essentiality relates to the impact on downstream KEs and the AO if upstream KEs are prevented or modified. This includes:

- Direct evidence: directly measured experimental support that blocking or preventing a KE prevents or impacts downstream KEs in the pathway in the expected fashion. Depending on the nature of the KE, could also be evidence that overexpression of the object of the KE prevents or impacts the downstream KEs in a manner consistent with its causal, and essential, role in the pathway.
- Indirect evidence: evidence that modulation or attenuation in the magnitude of impact on a specific KE (increased effect or decreased effect) is associated with corresponding changes (increases or decreases) in the magnitude or frequency of one or more downstream KEs.

When evaluating the overall support for essentiality of the KEs, authors may want to summarize their evaluation of relative levels of support in a tabular format (e.g., Table 5). The objective is to summarise briefly investigations in which the essentiality of KEs has been experimentally explored either directly or indirectly. In some cases, the impact of blocking or modifying an early KE on all downstream KEs in the pathway has been determined; in other cases, the impact only on a single adjacent or non-adjacent downstream KE has been measured.

When assembling support for essentiality of the KEs, it is not necessary to repeat lengthy text on the design or results of relevant investigations that may appear in other parts of the AOP description (e.g., as biological plausibility or empirical support for a KER). Rather, the entries should briefly address the extent of the supporting and contradictory data through a short description of the nature of the direct or indirect evidence addressing essentiality, along with relevant references. The objective is to provide an overview of the extent and nature of supporting and inconsistent data on essentiality of the KEs in a format that will facilitate a "call" on the overall degree of support for essentiality across the AOP. Some examples of brief narratives addressing support for essentiality are included here. The specific nature of these narratives necessarily vary, depending on the nature of key events in the AOP. See https://aopwiki.org/info_pages/2/info_linked_pages/6 for additional examples:

For direct evidence:

- Knock-out of KE1 or early KEs leads to blockage of all downstream KEs
- Overexpression or underexpression of KE1 leads to effect on all downstream KEs
- One or more downstream KEs is blocked or reversed by inhibiting (or allowing recovery of) upstream KEs
- Overexpression or underexpression in repair enzyme for early KEs leads to decreased or increased incidence of downstream KEs
- Antagonism or agonism of upstream KE leads to expected pattern of effects on downstream KEs

For indirect evidence:

- Impact of a known modulating factor for early KEs leads to expected pattern of effects on later KEs

Table 5: Example of a Table Format for summarizing the relative evidence supporting the

Essentiality of KEs in the AOP.

Event	Direct Evidence	Indirect Evidence	No experimental evidence	Contradictory experimental evidence
MIE	****	**		
KE1	*	****		
KE2			****	
KE3…… KEn	**			*

Uncertainties or Inconsistencies:
In addition to outlining the evidence supporting essentiality, it is also important to identify inconsistencies or uncertainties. This could include, for example, evidence in specific studies that did not support that blockage or attenuation of an early KE impacted later KEs in the AOP. Discordance with the results of other studies should be considered based on evaluation of the adequacy of study design, taking into account, for example, the sensitivity of the detection of impact. It could also include, for example, gaps in knowledge concerning the essentiality of the MIE or particular KEs where there are data on essentiality only for one or a few. To the extent possible, inconsistencies and uncertainties should focus on data gaps important for potential envisaged regulatory applications as a basis for indicating priorities for further research.

Based on the assembled evidence on essentiality for the KEs, confidence in the supporting data on essentiality is considered for the entire AOP, including KERs and KEs. This is commonly based on the extent of direct and/or indirect evidence for one, several or all of the KEs.

Confidence in the supporting data for essentiality of KEs within the AOP is considered:

- **High** if there is direct evidence from specifically designed experimental studies illustrating prevention or corresponding impact on downstream KEs and/or the AO if upstream KEs are blocked or modified [e.g., via stop exposure/reversibility studies, antagonism, knock out models, etc.];
- **Moderate** if there is indirect evidence that modification of one or more upstream KEs is associated with a corresponding (increase or decrease) in the magnitude or frequency of downstream KEs [e.g., augmentation of proliferative response (KEupstream) leading to increase in tumour formation (KEdownstream or AO)];
- **Low** if there is no or contradictory experimental evidence that blocking or modulating/attenuating any of the KEs influences the KEs downstream or AO (Annex 1).

4C. Evidence Assessment.
The biological plausibility, empirical support, and quantitative understanding from each KER in an AOP are assessed together:

i. Review the Biological Plausibility of Each KER
Biological plausibility of each of the KERs in the AOP is the most influential consideration in assessing WoE or degree of confidence in an overall hypothesised AOP for potential regulatory application (Meek et al., 2014; 2014a). The defining question for biological plausibility (Annex 1) is: Is there a mechanistic (i.e., structural or functional) relationship between KEupstream and KEdownstream consistent with established biological knowledge? Confidence in the WoE for the biological plausibility of the KERs would be considered:

- **High** if it is well understood based on extensive previous documentation and has an established mechanistic basis and broad acceptance (canonical knowledge; e.g., increased follicle stimulating hormone signalling leading to increased estrogen

synthesis, increased incidence of alkylated DNA leading to increased incidence of mutations)

- **Moderate** if the KER is plausible based on analogy to accepted biological relationships but scientific understanding is not completely established
- **Low** if there is empirical support for a statistical association between KEs but structural or functional relationship between them is not understood.

ii. Review the Empirical Support for Each KER

Empirical support entails consideration of experimental data in terms of the associations between KEs – namely dose-response concordance and temporal relationships between and across multiple KEs. It is examined most often in studies of dose-response/incidence and temporal relationships for stressors that impact the pathway at multiple levels of biological organization These patterns are most evident when considered across all KERs of the AOP with experimental protocols optimally designed to address incidence and severity of key events in the AOP at multiple or all levels of biological organization. While less influential than biological plausibility and essentiality (Meek et al., 2014; 2014a), empirical support contributes to the assessment of confidence in in an AOP for regulatory application.

It is important to recognise that empirical support relates to the "concordance" of dose response, temporal and incidence relationships for KERs; the defining question is not whether or not there is a dose response relationship for a specific KE but rather, whether there is expected concordance with the dose-response relationships for KERs – i.e., between KEs (Figure 9).

The defining questions for empirical support (Annex 1) are: Does KEupstream occur at lower doses and earlier time points than KEdownstream; is the incidence or frequency of KEupstream greater than that for KEdownstream for the same dose of tested stressor? Inconsistencies in empirical support across taxa, species and stressors that don't align with the expected pattern for the hypothesised AOP as described in Section 3 should be identified and their basis considered.

Empirical support for each of the KERs would be considered:

- **High** if there is dependent change in both events following exposure to a wide range of specific stressors (extensive evidence for temporal, dose-response and incidence concordance) and no or few data gaps or conflicting data'
- **Moderate** if there is demonstrated dependent change in both events following exposure to a small number of specific stressors and some evidence inconsistent with the expected pattern that can be explained by factors such as experimental design, technical considerations, differences among laboratories, etc.;
- **Low** if there are limited or no studies reporting dependent change in both events following exposure to a specific stressor (i.e., endpoints never measured in the same study or not at all), and/or lacking evidence of temporal or dose-response concordance, or identification of significant inconsistencies in empirical support across taxa and species that don't align with the expected pattern for the hypothesised AOP.

Although developers should evaluate the support for each KER, most critically for the Overall Assessment of the AOP is to consider the overall level of support across all of the KERs. It may not be uncommon that the degree of supporting evidence for some KERs in the pathway are quite limited. However, when there is strong plausibility for the pathway as a whole, and there are well supported non-adjacent relationships that bridge across some of the weaker intermediate KERs, the support for the pathway as a whole may still be quite strong. While evidence assembly may be done in a highly modular fashion, the Overall Assessment of the AOP should once again step back and evaluate the evidence supporting the pathway as a whole. It is that more integrated and wholistic view that really informs application.

Tables summarising the relevant experimental data for tested stressors across all the KEs may be helpful in considering the extent of empirical support and to the extent possible should be based on benchmark doses. For example, points of departure (e.g., benchmark doses, LOAELs, EC_x, etc.) for specified similar increases in each of the KEs are entered in the cells of the table. If the hypothesised linkages in the AOP are supported by empirical data, there is a pattern of increasing points of departure from the top lefthand corner to the bottom right hand corner for each of the tested stressors. Presentation in this manner readily identifies any exceptions to the expected patterns that are considered as inconsistencies and diminish from the overall weight of empirical support (see Table 6).

Table 6. Generic example of a concordance table for evaluating overall empirical support for an AOP.

Point of departure[a] (mg/kg/d)	KE 1	KE 2	KE 3	KE 5	KE 6	KE 7
0.01	----	----	----	----	----	----
0.05	+++	++	---	++	----	----
0.1		+	+++	+++	----	----
0.5					++	----
1.0					+	++++

a. Point of departure at which a specified level of change in the KE relative to controls was inferred, based on the empirical results. (Note, depending on the type of point of departure employed, characteristics of the underlying study designs such as dose selection, dose-spacing, statistics employed, etc. must be taken into account).

4D. Known Modulating Factors

The evidence supporting the influence of various modulating factors is assembled within the individual KERs. As part of the Overall Assessment of the AOP, authors should list the known modulating factors that have been identified, briefly note their expected influence on the outcome, and list the specific KER(s) involved. This can be captured in a simple table (e.g., Table 7). Additional details or notes can be supplied as free text below the table.

Table 7. Example of suggested tabular format for identifying critical information concerning known modulating factors that may be expected to influence the AOP.

Modulating Factor	Influence on Outcome	KER(s) Involved

4E. Review the Quantitative Understanding of the KERs

The extent of quantitative understanding of the KERs in an AOP is critical with regard to potential regulatory application. For some applications (e.g., dose- response analysis in an in-depth risk assessment), quantitative characterization of downstream KERs may be essential, while for others quantitative understanding of upstream KERs may be most important (e.g., QSAR modelling for category formation for testing). Because evidence that contributes to quantitative understanding of the KER is generally not mutually exclusive with the empirical support for the KER (i.e., expected patterns of quantitative relationships), evidence that contributes to quantitative understanding will generally be considered to some extent as part of the evaluation of the WoE supporting the KER (see Section 3.E. and Annex 1, footnote b). However, specific attention is also given to how precisely and accurately one can potentially predict an impact on KEdownstream based on some measurement of KEupstream. This is captured in the form of quantitative understanding calls for each KER, i.e. as low, moderate, or high (Annex 2). As noted in section 3, general guidance for characterising the level of quantitative understanding of a KER is based on several key considerations:

- The extent to which a change in KEdownstream can be precisely predicted based on KEupstream.
- The precision with which uncertainty in the prediction of KEdownstream can be quantified.
- The extent to which known modulating factors or feedback mechanisms are accounted for.
- The extent to which the relationships described can be reliably generalized across the applicability domain of the KER.

As with the other parts of the overall assessment of the AOP, it is not necessary to repeat all the details provided in the KER descriptions. The overall evaluation of the quantitative understanding should briefly explain the rationale for the assigned level of quantitative understanding of each KER. It should then consider the overall pattern of quantitative understanding across all KERs to indicate how precisely outcomes along the entire pathway may be predicted for a given exposure scenario. If certain parts of the pathway can be predicted with quantitative precision, while others cannot, the potential implications for application may be discussed.

4F. Considerations for Potential Applications of the AOP (optional)

The Overall Assessment of the AOP is intended to help inform decisions about an AOP's fit-for-purpose for different types of applications. Consequently, at their discretion, following their assessment of the AOP, the developers may want to discuss the type(s) of application(s) they feel the AOP would be suited for, based on their evaluation. This may include, for example, possible utility for test guideline development or refinement, development of integrated testing and assessment approaches, development of (Q)SARs / or chemical profilers to facilitate the grouping of chemicals for subsequent read-across, screening-level hazard assessments or even risk assessment. This section can consider whether the AOP assembled can support the intended application that was outlined previously in the "AOP Development Strategy" section. It may also be that new potential applications or limitations which become apparent when developing the AOP and assessing the evidence could also be noted in this section.

It is further recognized, that developers may not be aware of all the potential applications for any given AOP. Consequently, users of the AOP-Wiki are encouraged to leave comments on the discussion pages, or via the AOP Forum if they identify suitable applications for a given AOP. Listing these applications can aid others in using the AOP.

4G. References

References cited elsewhere on the AOP page should be listed here. This is not a compilation of all references cited on the linked KE and KER pages. Ideally, the list of references, should conform with the OECD Style Guide (https://www.oecd.org/about/publishing/OECD-Style-Guide-Third-Edition.pdf) (OECD, 2015).

REFERENCES

Becker RA, Ankley GT, Edwards SW, Kennedy SW, Linkov I, Meek B, Sachana M, Segner H, Van Der Burg B, Villeneuve DL, Watanabe H, Barton-Maclaren TS. (2015). Increasing scientific confidence in adverse outcome pathways: application of tailored Bradford-Hill considerations for evaluating weight of evidence. Regul Toxicol Pharmacol 72: 514-537.

Collier ZA, Gust KA, Gonzalez-Morales B, Gong P, Wilbanks MS, Linkov I, Perkins EJ. (2016). A weight of evidence assessment approach for adverse outcome pathways. Regul Toxicol Pharmacol 75: 46-57.

European Food Safety Authority. (2022). Guidance on the use of the benchmark dose approach in risk assessment. https://doi.org/10.2903/j.efsa.2022.7584.

Jensen, MA, Blatz, D J, LaLone, CA. (2022) Defining the Biologically Plausible Taxonomic Domain of Applicability of an Adverse Outcome Pathway: A Case Study Linking Nicotinic Acetylcholine Receptor Activation to Colony Death. Environmental Toxicology and Chemistry. 42:71-87

Knapen, D., Vergauwen, L., Villeneuve, D.L. and Ankley GT. (2015) The potential of AOP networks for reproductive and developmental toxicity assay development. Reprod Toxicol. 56: 52-55.

Knapen D, Angrish MM, Fortin MC, Katsiadaki I, Leonard M, Margiotta-Casaluci L, Munn S, O'Brien JM, Pollesch N, Smith LC, Zhang X, Villeneuve DL. Adverse outcome pathway networks I: Development and applications. Environ Toxicol Chem. 2018 Jun;37(6):1723-1733. doi: 10.1002/etc.4125. Epub 2018 May 7. PMID: 29488651; PMCID: PMC6004608.

Krewski D, Acosta D Jr., Andersen M, Anderson H, Bailar J.C. 3rd, Boekelheide K, Brent R, Charnley G, Cheung VG, Green S Jr, Kelsey KT, Kerkvliet NI, Li AA, McCray L, Meyer O, Patterson RD, Pennie W, Scala RA, Solomon GM, Stephens M, Yager J, Zeise L. (2010). Toxicity testing in the 21st century: a vision and strategy. J Toxicol Environ Health B Crit Rev. 13: 51-138.

LaLone CA, Ankley GT, Belanger SE, Embry MR, Hodges G, Knapen D, Munn S, Perkins EJ, Rudd MA, Villeneuve DL, Whelan M, Willett C, Zhang X, Hecker M. (2017.) Advancing the adverse outcome pathway framework – an international horizon scanning approach. Environ Toxicol Chem. 36: 1411-1421.

Leist M, Ghallab A, Graepel R, Marchan R, Hassan R, Bennekou SH, Limonciel A, Vinken M, Schildknecht S, Waldmann T, Danen E, van Ravenzwaay B, Kamp H, Gardner I, Godoy P, Bois FY, Braeuning A, Reif R, Oesch F, Drasdo D, Höhme S, Schwarz M, Hartung T, Braunbeck T, Beltman J, Vrieling H, Sanz F, Forsby A, Gadaleta D, Fisher C, Kelm J, Fluri D, Ecker G, Zdrazil B, Terron A, Jennings P, van der Burg B, Dooley S, Meijer AH, Willighagen E, Martens M, Evelo C, Mombelli E, Taboureau O, Mantovani A, Hardy B, Koch B, Escher S, van Thriel C, Cadenas C, Kroese D, van de Water B, Hengstler JG. (2017) Adverse outcome pathways: opportunities, limitations, and open questions. Regulat. Toxicol. DOI: 10.1007/s00204-017-2045-3

Meek ME, Klaunig JE. (2010). Proposed mode of action of benzene induced leukemia: interpreting available data and identifying critical data gaps for risk assessment. Chem. Biol. Interact. 184: 279-285.

Meek ME, Boobis AR, Cote I, Dellarco V, Fotakis G, Munn S, Seed J, Vickers C. (2014a). New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. J Appl Toxicol. 34: 1-18.

Meek ME, Palermo CM, Bachman AN, North, CM, Lewis RJ. (2014b). Mode of Action Human Relevance (MOA/HR) Framework – Evolution of the Bradford Hill Considerations and Comparative Analysis of Weight of Evidence. J Appl Toxicol. 34: 595-606.

OECD (2015), OECD Style Guide third edition, OECD Publishing, Paris. https://www.oecd.org/about/publishing/OECD-Style-Guide-Third-Edition.pdf

United States Environmental Protection Agency. (2012). Benchmark dose technical guidance. EPA/100/R-12/001. https://www.epa.gov/sites/default/files/2015-01/documents/benchmark_dose_guidance.pdf

Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L, Whelan M. (2014a) Adverse outcome pathway (AOP) development I: strategies and principles. Toxicol Sci. 142: 312-320.

Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L, Whelan M. (2014b) Adverse outcome pathway development II: best practices. Toxicol Sci. 142: 321-330.

Wittwehr C, Aladjov H, Ankley G, Byrne HJ, de Knecht J, Heinzle E, Klambauer G, Landesmann B, Luijten M, MacKay C, Maxwell G, Meek ME, Paini A, Perkins E, Sobanski T, Villeneuve D, Waters KM, Whelan M. (2017) How Adverse Outcome Pathways can aid the development and use of computational prediction models for regulatory toxicology. Toxicol Sci. 155: 326-336.

ANNEX 1: Guidance for Assessing Relative Level of Confidence in the Overall AOP

Examples of complete tables for selected AOPs are available:

AOP	Assessment Summary File
https://aopwiki.org/aops/15	https://aopwiki.org/system/dragonfly/production/2017/05/19/7s1ibrunwt_RevisedAssessmentSummaryAop_15.pdf
https://aopwiki.org/aops/23	https://aopwiki.org/system/dragonfly/production/2017/03/20/3usvv7naq8_Annex1_for_AOP_23_AR_reproductive_dys_2017_03_20.pdf
https://aopwiki.org/aops/38	https://aopwiki.org/aops/38#evidence
https://aopwiki.org/aops/42	https://aopwiki.org/system/dragonfly/production/2017/03/24/6u60jhkjp8_TPO_AOP_Summary_Tables.pdf

1. Support for Biological Plausibility of KERs[1]	Defining Question	High [2,3]	Moderate	Low
	Is there a mechanistic (i.e., structural or functional) relationship between KEup and KEdown consistent with established biological knowledge?	Extensive understanding based on extensive previous documentation and broad acceptance -Established mechanistic basis	The KER is plausible based on analogy to accepted biological relationships but scientific understanding is not completely established.	There is empirical support for a statistical association between KEs (See 3.), but the structural or functional relationship between them is not understood.
[4]MIE => KE1: (copy and paste the KER description into this cell)	Biological Plausibility of the MIE => KE1 is xxx. Rationale:			
KE1 => KE2: (copy and paste the KER description into this cell)	Biological Plausibility of KE1 => KE2 is xxx Rationale:			
KE2 => KE3 (copy and paste the KER description into this cell)	Biological Plausibility of KE2 => KE3 is xxx. Rationale:			

[1]Rank ordered Bradford Hill considerations adapted from Meek et al. (2014b)

[2]The guidance for "high", "moderate" and "low" draws on limited current experience. Additional delineation of the nature of relevant evidence in these broadly defined categories requires more experience with larger numbers of documented AOPs.

[3]"Direct evidence" implies specifically designed experiments to consider the relevant element. "Indirect evidence" may overlap with other elements.

[4]To the extent possible, each of the relevant Bradford Hill considerations is addressed for each of the KERs (biological plausibility and empirical support) and KEs (essentiality) and separate rationales provided.

2. Support for Essentiality of KEs[5]	Defining Question	High	Moderate	Low
	What is the impact on downstream KEs and/or the AO if an upstream KE is modified or prevented?	Direct evidence from specifically designed experimental studies illustrating prevention or impact on downstream KEs and/or the AO if upstream KEs are blocked or modified	Indirect evidence that modification of one or more upstream KEs is associated with a corresponding (increase or decrease) in the magnitude or frequency of downstream KEs	No or contradictory experimental evidence of the essentiality of any of the KEs.
AOP	Rationale for Essentiality of KEs in the AOP is xxx:			

[5]While the extent of the supporting data on the essentiality of each of the KEs is addressed separately (Table 3), delineation of the degree of confidence is based on consideration of evidence for all of the KEs within the AOP and therefore, only one rationale is required. This call is normally based on the extent of the available evidence for a range of KEs in the AOP.

3. Empirical Support for KERs	Defining Questions	High	Moderate	Low
	Does KEup occur at lower doses and earlier time points than KE down and at the same dose of stressor, is the incidence of KEup > than that for KEdown?[6,7]. Are there inconsistencies in empirical support across taxa, species and stressors that don't align with expected pattern for hypothesised AOP?	Multiple studies showing dependent change in both events following exposure to a wide range of specific stressors. (Extensive evidence for temporal, dose-response and incidence concordance) and no or few critical data gaps or conflicting data	Demonstrated dependent change in both events following exposure to a small number of specific stressors and some evidence inconsistent with expected pattern that can be explained by factors such as experimental design, technical considerations, differences among laboratories, etc.	Limited or no studies reporting dependent change in both events following exposure to a specific stressor (i.e., endpoints never measured in the same study or not at all); and/or significant inconsistencies in empirical support across taxa and species that don't align with expected pattern for hypothesised AOP
MIE => KE1: (copy and paste the KER description into this cell)b	Empirical Support of the MIE => KE1 is xxx. Rationale:			
KE1 => KE2: (copy and paste the KER description into this cell)	Empirical Support of the KE1 => KE2 is xxx. Rationale:			
KE2 => KE3 (copy and paste the KER description into this cell)	Empirical Support of the KE2 => KE3 is xxx. Rationale:			
b In many cases, evidence that contributes to quantitative understanding (Section 4 of a KER description) will also provide empirical support for the relationship. Consequently, relevant information from the "Quantitative Understanding" section of the KER description should be considered as part of the overall weight of evidence				

[6]This is normally considered on the basis of tabular presentation of available data on temporal and dose-response aspects, in a template that documents the extent of support. See, for example, Table 4.

[7]Note that this relates to concordance of dose response, temporal and incidence relationships for KERs rather than the KEs; the defining question is not whether or not there is a dose response relationship for the KE but whether there is concordance with that for earlier and later KEs. This is normally demonstrated in studies with different types of stressors.

ANNEX 2: General guidance for characterizing the level of quantitative understanding of a KER as low, moderate, or high.

Extent of Quantitative Understanding	Characteristics
High	Change in KEdownstream can be precisely predicted based on a relevant measure of KEupstream. Uncertainty in the quantitative prediction can be precisely estimated from the variability in the relevant measure of KEupstream. Known modulating factors and feedback/feedforward mechanisms are accounted for in the quantitative description. There is evidence that the quantitative relationship between the KEs generalizes across the relevant applicability domain of the KER.
Moderate	Change in KEdownstream can be precisely predicted based on a relevant measure of KEupstream. Uncertainty in the quantitative prediction is influenced by factors other than the variability in the relevant measure of KEupstream. Quantitative description does not account for all known modulating factors and/or known feedback/feedforward mechanisms. The quantitative relationship has only been demonstrated for a subset of the overall applicability domain of the KER (e.g., based on a single species).
Low	Only a qualitative or semi-quantitative prediction of the change in KEdownstream can be determined from a measure of KEupstream. Known modulating factors and/or known feedback/feedforward mechanisms are not accounted for. The quantitative relationship has only been demonstrated for a narrow subset of the overall applicability domain of the KER (e.g., based on a single species).