AOP: 25 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

	Defining Question	High (Strong)	Moderate	Low (Weak)
1. Support for Biological Plausibility of KERS				
	a) Is there a mechanistic relationship between KE_{up} and KE_{down} consistent with established biological knowledge?	Extensive understanding of the KER based on extensive previous documentation and broad acceptance.	KER is plausible based on analogy to accepted biological relationships, but scientific understanding is incomplete	Empirical support for association between KEs , but the structural or functional relationship between them is not understood.
KE1 (MIE) => KE2: Aromatase inhibition directly leads to 17β-estradiol synthesis by ovarian granulosa cells, reduction	**STRONG.** It is well established that aromatase is rate limiting for 17β-estradiol synthesis and that the granulosa cells of the ovary are the primary site of expression and production.			
KE2 => KE3: 17β-estradiol synthesis by ovarian granulosa cells, reduction directly leads to plasma 17β-estradiol concentrations, reduction	**STRONG.** The biochemistry of steroidogenesis and the predominant role of the gonad in synthesis of the sex steroids is well established			
KE3 => KE4: plasma 17β-estradiol concentrations, reduction directly leads to vitellogenin production in liver (transcription, translation), reduced	**STRONG.** The role of E2 as the major regulator of hepatic vitellogenin production is widely documented in the literature			
KE4 => KE5: vitellogenin production in liver (transcription, translation), reduced directly leads to plasma vitellogenin concentrations, reduced	**STRONG.** It is well established that hepatic synthesis is the major source of plasma vitellogenin in oviparous vertebrates. The central dogma of molecular biology dictates that transcription and translation are needed for protein production.			
KE5 => KE6: plasma vitellogenin concentrations, reduced directly leads to vitellogenin uptake into oocytes and oocyte growth/development, reduction.	**STRONG.** It is well established that the circulation is the primary source of egg yolk proteins in fish.			
KE6 => KE7 (AO): vitellogenin uptake into oocytes and oocyte growth/development, reduction directly leads to cumulative fecundity and spawning, reduction	**MODERATE.** The direct connection between reduced VTG accumulation and impaired spawning/reduced cumulative fecundity is somewhat tentative. It is not clear, for instance whether impaired VTG accumulation limits oocyte growth and failure to reach a critical size in turn impairs physical or inter-cellular signaling processes that promote release of the oocyte from the surrounding follicles. In at least one experiment, oocytes with similar size to vitellogenic oocytes, but lacking histological staining characteristic of vitellogenic oocytes was observed (R. Johnson, personal communication). At present, the link between reductions in circulating VTG concentrations and reduced cumulative fecundity are best supported by the correlation between those endpoints across multiple experiments, including those that impact VTG via other molecular initiating events (Miller et al. 2007). Reference: Miller DH, Jensen KM, Villeneuve DL, Kahl MD, Makynen EA, Durhan EJ, Ankley GT. Linkage of biochemical responses to population-level effects: a case study with vitellogenin in the fathead minnow (Pimephales promelas). Environ Toxicol Chem. 2007 Mar;26(3):521-7.			

AOP: 25 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

<table>
<tr><td>KE7 (AO) => KE8 (AO): cumulative fecundity and spawning, reduction directly leads to population trajectory, decrease</td><td colspan="4">MODERATE.
Using a relatively simple density-dependent population model and assuming constant young of year survival with no immigration/emigration, reductions in cumulative fecundity have been predicted to yield declines in population size over time (Miller and Ankley 2004). Under real-world environmental conditions, outcomes may vary depending on how well conditions conform with model assumptions. Nonetheless, cumulative fecundity can be considered one vital rate that contributes to overall population trajectories.

Reference: Miller DH, Ankley GT. Modeling impacts on populations: fathead minnow (Pimephales promelas) exposure to the endocrine disruptor 17beta-trenbolone as a case study. Ecotoxicol Environ Saf. 2004 Sep;59(1):1-9.</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td rowspan="2">2. Support for Essentiality of KEs</td><td>Defining Question</td><td>High (Strong)</td><td>Moderate</td><td>Low (Weak)</td></tr>
<tr><td>Are downstream KEs and/or the AO prevented if an upstream KE is blocked?</td><td>Direct evidence from specifically designed experimental studies illustrating essentiality for at least one of the important KEs</td><td>Indirect evidence that sufficient modification of an expected modulating factor attenuates or augments a KE</td><td>No or contradictory experimental evidence of the essentiality of any of the KEs.</td></tr>
<tr><td>Essentiality of the KEs was assessed for the AOP as a whole – rationale for the individual KE calls is provided.</td><td colspan="4">Support for the essentiality of a number of key events in the AOP was provided by several time-course, stop-reversibility, experiments with fathead minnows exposed to aromatase inhibitors.

1. Villeneuve et al. 2009 and 2013 examined a time-course of key event responses to fadrozole as well as the time-course of recovery following cessation of fadrozole delivery. Once fadrozole was removed from the system, ex vivo E2 production increased, followed by increases in plasma E2 concentrations, and then increases in plasma vitellogenin concentrations. Additionally, while exposure to the chemical was on-going, compensatory up-regulation of CYP19a1a gene expression resulted in increases in ex vivo E2 production, followed by increased plasma E2 and plasma VTG. The essentiality of aromatase inhibition relative to impaired E2 production was further supported by the observation of an "overshoot" in E2 production, relative to controls, shortly after cessation of fadrozole delivery.
2. Similar support was provided in a study by Ankley et al. (2009a). Cessation of prochloraz delivery resulted in rapid recovery of ex vivo E2 production and plasma E2 concentrations, with recovery of vitellogenin concentrations lagging slightly behind. Increased expression of cyp19a1a mRNA during the exposure period aligned with increased ex vivo E2 production, and increased plasma E2, compared to the first day of exposure.

Rationale for essentiality calls:

• Aromatase, inhibition: [Strong] There is good evidence from stop/reversibility studies that ceasing delivery of the aromatase inhibitor leads to recovery of the subsequent key events.
• 17beta-estradiol synthesis by ovarian granulosa cells, reduction: [Strong] In both exposure studies and stop/reversibility studies, when ex vivo E2 production (as measure of this KE) recovers either through compensation or due to removal of the stressor, subsequent KEs have been shown to recover after a lag period.
• plasma 17beta-estradiol concentrations, reduction: [Strong] In both exposure studies and stop/reversibility studies, when plasma E2 concentrations recover either through compensation or due to removal of the stressor, subsequent KEs have been shown to recover after a lag period.
• vitellogenin production in liver (transcription, translation), reduction: [Moderate] This endpoint was not specifically examined in stop/reversibility studies with aromatase inhibitors, but biological plausibility provides strong support for the essentiality of this event.
• plasma vitellogenin concentrations, reduction: [Strong] Shown to recover in a predictable fashion consistent with the order of events in the AOP in stop/recovery studies.
• vitellogenin accumulation into oocytes and oocyte growth/development, reduction: [Weak] Some contradictory evidence regarding the essentiality of this event. No stop/reversibility studies have explicitly considered this key event.
• cumulative fecundity and spawning, reductions: [Moderate] By definition, some degree of spawning is required to maintain population.

REFERENCES
Villeneuve DL, Breen M, Bencic DC, Cavallin JE, Jensen KM, Makynen EA, Thomas LM, Wehmas LC, Conolly RB, Ankley GT. Developing predictive approaches to characterize adaptive responses of the reproductive endocrine axis to aromatase inhibition: I. Data generation in a small fish model. Toxicol Sci. 2013 Jun;133(2):225-33. doi: 10.1093/toxsci/kft068.

Villeneuve DL, Mueller ND, Martinović D, Makynen EA, Kahl MD, Jensen KM, Durhan EJ, Cavallin JE, Bencic D, Ankley GT. Direct effects, compensation, and recovery in female fathead minnows exposed to a model aromatase inhibitor. Environ Health Perspect. 2009 Apr;117(4):624-31. doi: 10.1289/ ehp.11891.</td></tr>
</table>

AOP: 25 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

	Ankley GT, Bencic DC, Cavallin JE, Jensen KM, Kahl MD, Makynen EA, Martinovic D, Mueller ND, Wehmas LC, Villeneuve DL. Dynamic nature of alterations in the endocrine system of fathead minnows exposed to the fungicide prochloraz. Toxicol Sci. 2009 Dec;112(2):344-53. doi: 10.1093/toxsci/kfp227.			
3. Empirical Support for KERs	Defining Questions	High (Strong)	Moderate	Low (Weak)
	Does empirical evidence support that a change in KE_{up} leads to an appropriate change in KE_{down}? Does KE_{up} occur at lower doses and earlier time points than KE_{down} and is the incidence of KE_{up} > than that for KE_{down}? Inconsistencies?	Multiple studies showing dependent change in both events following exposure to a wide range of specific stressors. No or few critical data gaps or conflicting data	Demonstrated dependent change in both events following exposure to a small number of stressors. Some inconsistencies with expected pattern that can be explained by various factors.	Limited or no studies reporting dependent change in both events following exposure to a specific stressor; and/or significant inconsistencies in empirical support across taxa and species that don't align with hypothesized AOP
KE1 (MIE) => KE2: Aromatase inhibition directly leads to 17β-estradiol synthesis by ovarian granulosa cells, reduction	**MODERATE** Direct measurement of aromatase inhibition following in vivo exposures are difficult to achieve, therefore identification of aromatase inhibition as a relevant MIE is most often based on in vitro experiments. Reductions in the rate of E2 production by ovary tissue or steroid producing cells following exposure to chemicals identified in vitro as aromatase inhibitors provides support. **Dose-response:** There is little direct support for dose-response concordance of these key events in vivo. However, using in vitro systems concentrations that reduce aromatase activity tend to elicit reductions in E2 production. **Temporality:** E2 production by ovary explants obtained from fish exposed to known aromatase inhibitors declines rapidly, following exposure, and has also been shown to recover rapidly upon cessation of the delivery of known aromatase inhibitors. **Uncertainties**: Because E2 synthesis is at the fairly terminal end of the steroid biosynthesis pathway, impacts of chemicals on other enzymes in the steroid biosynthesis pathway can lead to reduced E2 synthesis. There is also compelling evidence for fairly rapid in vivo compensation for aromatase inhibition via up-regulated transcription of aromatase mRNA expression. Consequently, complementary data from multiple types of in vitro assays are likely superior to in vivo evidence for establishing this KER.			
KE2 => KE3: 17β-estradiol synthesis by ovarian granulosa cells, reduction directly leads to plasma 17β-estradiol concentrations, reduction	**STRONG** The rate of E2 production by ovarian explants and circulating concentrations of estradiol can generally both be measured for individual animals exposed in an experiment. Therefore, there is a fair amount of concurrent data for these endpoints. **Dose Response:** Effects on KE2 are generally observed at or near the same concentrations that impact KE3. There are exceptions, but these are typically explained by the higher variability (and thus lower statistical power) associated with the ex vivo steroid production assays often used to measure KE2. **Temporality:** Data from several time course studies, with at least two different aromatase inhibitors, support the idea that impacts on KE2 are detected (statistically) at earlier time-points than impacts on KE3. Data from these studies also show that KE2 recovers before KE3 both as the result of compensatory responses during an exposure period and following cessation of delivery of an aromatase inhibitor. **Incidence:** Particularly for experiments of longer duration (> 4 d), there are cases where impacts on KE3 are detected without concurrent effects on KE2. These are plausibly explained by the fact that compensatory responses in vivo lead to more rapid "recovery" of KE2 than KE3. It also reflects the fact that measures of KE2 represent a rate of steroid production per unit mass of tissue, while KE3 reflects total output of the whole organ into circulation. Small reductions in the rate of production per unit mass of tissue, which are not statistically detectable, can still lead to statistically detectable reductions in circulating concentrations.			
KE3 => KE4: plasma 17β-estradiol concentrations, reduction directly leads to vitellogenin production in liver (transcription, translation), reduced	**WEAK** Circulating E2 concentrations and the relative abundance of hepatic vitellogenin transcripts can generally be concurrently measured for individual animals from the same experiment. Although methodologically more challenging, hepatic vitellogenin protein abundance can also be measured from the same fish. However, based on the empirical evidence currently assembled, relatively few studies have included a measurement of either VTG mRNA abundance or VTG protein abundance as an endpoint (see Tables 1 and 2). **Dose Response:** In one study that examined both KE3 and KE4, impacts on KE4 were observed at much lower concentrations. However, the measurement technology (mass spectroscopy-based proteomics) employed for measuring KE4 may be significantly more quantitative and precise than that employed for measuring KE3. **Temporality:** There are currently no time-course studies in which KE3 and KE4 were both measured. **Incidence:** In the only study that examined both KE3 and KE4, effects on both KEs were observed.			
KE4 => KE5: vitellogenin production in liver (transcription, translation),	**WEAK** Few studies with aromatase inhibitors have reported impacts on hepatic vitellogenin transcription or translation, thus empirical data for evaluating this KER are limited.			

reduced directly leads to plasma vitellogenin concentrations, reduced	**Dose Response:** In the one study that examined both KE4 and KE5, the impact on the upstream event (KE4) occurred at a lower concentration than that at which the effect on the downstream KE (KE5) was observed. However, it should be noted that the measurement methods may not be comparable in terms of precision; mass-spectroscopy-based proteomics for KE4 versus an ELISA for KE5. **Temporality:** There are not sufficient empirical data to evaluate the temporal concordance of these key events. **Incidence:** In the only study that examined both KE3 and KE4, effects on both KEs were observed.
KE3 => KE5: plasma 17β-estradiol concentrations, reduction indirectly leads to plasma vitellogenin concentrations, reduced	**STRONG** Circulating E2 (KE3) and VTG concentrations (KE5) are readily measured in plasma samples collected from the same individual animals exposed in an experiment. Measurements of both KEs have frequently been made. **Dose Response:** Generally speaking effects on the downstream KE (KE5) were observed at concentrations equal to or greater than those at which effects on the upstream event (KE3) were reported. There were several exceptions. However exceptions are plausibly explained by a number of factors. First, vitellogenin concentrations in plasma have a much greater dynamic range (i.e., often change by orders of magnitude) than circulating steroid concentrations (changes are typically within 1-2 orders of magnitude). Second, compensatory responses elicited in response to aromatase inhibition have been shown to impact KE3 more rapidly than KE5, which can lead to a disconnect in the apparent dose needed to elicit a response at a given time-point. **Temporality:** In several independent time-course studies with multiple aromatase inhibitors, impacts on KE3 reliably precede those on KE5. Likewise, "recovery" of KE3 as a result of compensatory responses during exposure or cessation of chemical delivery consistently precede that of KE5. **Incidence:** Taking the temporal relationship between the two KEs into account, there is strong concordance in the incidence of KE3 and KE5 across several studies.
KE5 => KE6: plasma vitellogenin concentrations, reduced directly leads to vitellogenin accumulation into oocytes and oocyte growth/development, reduction.	**WEAK** Conceptually, both plasma vitellogenin concentrations and ovarian histology measurements can be made in the same individuals exposed in a given experiment. However, among the studies available to date, examination of both endpoints has generally been limited to the longer duration studies. Given that ovulation and spawning are the major routes through which oocytes containing vitellogenin are lost from the ovary, one or more spawning events may need to occur in order for existing vitellogenic oocytes to be "cleared" from the ovary or to undergo atresia, before the impacts on KE6 can be detected. **Dose Response:** For the one study in which both plasma vitellogenin and ovarian histology were examined, effects on uptake of VTG into oocytes were detected at concentrations greater than those that impacted plasma steroid concentrations. **Temporality:** Impacts on circulating vitellogenin have been observed at time points earlier than those at which significant histological evidence of reduced VTG uptake into oocytes has been detected. **Incidence:** Given the limited data set, incidence concordance cannot be thoroughly evaluated.
KE6 => KE7 (AO): vitellogenin accumulation into oocytes and oocyte growth/development, reduction directly leads to cumulative fecundity and spawning, reduction	**WEAK** There are only a few studies in which KE6 and KE7 were examined concurrently. **Dose Response**: In the one study in which concurrent measures for KE6 and KE7 were reported, effects were detected at the same concentration. **Temporality**: At present, there are no time-course data that directly address the temporal concordance between KE6 and KE7. **Incidence**: Given the limited data set, incidence concordance cannot be robustly evaluated.
KE7 (AO) => KE8 (AO): cumulative fecundity and spawning, reduction directly leads to population trajectory, decrease	**WEAK** There is limited direct evidence in the literature that population size will decrease if fecundity is decreased. There are no empirical data suitable for evaluating the dose-response, temporal, or incidence concordance between KE7 and KE8.

AOP: 25 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

KER	Integrative Assessment leading to the final weight of evidence call for each KER
KE1 (MIE) => KE2: Aromatase inhibition directly leads to 17β-estradiol synthesis by ovarian granulosa cells, reduction	STRONG Strong biological plausibility supported by moderate empirical support and well established essentiality for both KEs.
KE2 => KE3: 17β-estradiol synthesis by ovarian granulosa cells, reduction directly leads to plasma 17β-estradiol concentrations, reduction	STRONG Strong biological plausibility supported by strong empirical support.
KE3 => KE4: plasma 17β-estradiol concentrations, reduction directly leads to vitellogenin production in liver (transcription, translation), reduced	STRONG Even though the empirical support available is quite limited, plausibility provides a very strong basis upon which to build confidence in this key event relationship. Estrogen-dependent regulation of vitellogenin production is very well established. Additionally, there is strong support for the indirect relationship linking KE3 and KE5, which together with plausibility lends strong support for this KER.
KE4 => KE5: vitellogenin production in liver (transcription, translation), reduced directly leads to plasma vitellogenin concentrations, reduced	STRONG Even though the empirical support available is quite limited, plausibility provides a very strong basis upon which to build confidence in this key event relationship. Estrogen-dependent regulation of vitellogenin production is very well established. Additionally, there is strong support for the indirect relationship linking KE3 and KE5, which together with plausibility lends strong support for this KER.
KE5 => KE6: plasma vitellogenin concentrations, reduced directly leads to vitellogenin accumulation into oocytes and oocyte growth/development, reduction.	MODERATE While plausibility is fairly strong, the empirical support for the relationship is relatively weak. There are few studies in which both plasma VTG and ovarian histology have been examined. Because VTG is the only major source of VTG to the developing oocytes the connection is highly plausible. However, it remains unclear how much decreases in plasma VTG impacts accumulation if the decrease happens after oocytes have already reached vitellogenic stage. Presumably, the more rapid the oocyte turn over, the tighter the linkage, but uncertainties remain.
KE6 => KE7 (AO): vitellogenin accumulation into oocytes and oocyte growth/development, reduction directly leads to cumulative fecundity and spawning, reduction	MODERATE The plausibility is only moderate and only a few studies have examined KE6 and KE7 concurrently in the same experiment.
KE7 (AO) => KE8 (AO): cumulative fecundity and spawning, reduction directly leads to population trajectory, decrease	MODERATE The relationship is plausible, but not necessarily generalizable to real-world situation or a diversity of life histories and reproductive strategies. Direct evidence is quite limited.