AOP: 23 – 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.
Relationship: 31 Agonism, Androgen receptor (Event 25) leads to gonadotropins circulating concentration (Event 129)	**WEAK** Negative feedback processes have been characterized for a number of vertebrates, but the extent to which those are directly mediated by agonism of the AR versus more indirectly is unknown. Furthermore, there are known differences among species an life stages in terms of the operability of the feedback mechanism.			
Relationship: 143 Reduction, gonadotropins, circulating concentrations (Event 129) leads to reduction, testosterone synthesis by ovarian theca cells (Event 274)	**STRONG.** The role of gonadotropins in stimulating ovarian steroid production is well established and is widely accepted in endocrinology. There is overwhelming evidence for a structural/functional relationship in vertebrates.			
Relationship: 302 Reduction, testosterone synthesis by ovarian theca cells (Event 274) leads to reduction, 17b-estradiol synthesis by ovarian granulosa cells (Event 3)	**STRONG.** It has been indisputably established the testosterone is a precursor for 17b-estradiol synthesis. The localization of T synthesis to the theca cells, and 17b-estradiol synthesis predominantly to the granulosa, in ovary has been well established for over 40 years.			
Relationship: 5 Reduction, 17b-estradiol synthesis by ovarian granulosa cells (Event 3) leads to reduction, plasma 17b-estradiol concentrations (Event 219)	**STRONG.** The biochemistry of steroidogenesis and the predominant role of the gonad in synthesis of the sex steroids is well established			
Relationship 252: Reduction, plasma 17b-estradiol concentrations (Event 219) leads to reduction, vitellogenin synthesis in liver (Event 285)	**STRONG** The role of E2 as the major regulator of hepatic vitellogenin production is widely documented in the literature			
Relationship 315: Reduction, vitellogenin synthesis in liver (Event 285) leads to Reduction, plasma vitellogenin concentrations (Event 221)	**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.			
Relationship: 255 Reduction, plasma vitellogenin concentrations (Event 221) leads to Reduction, vitellogenin accumulation into oocytes and oocyte	**STRONG.** It is well established that the circulation is the primary source of egg yolk proteins in fish.			

growth/development (Event 309)	
Relationship: 337 Reduction, vitellogenin accumulation in oocytes and oocyte growth/development (Event 309) leading to Reduction, cumulative fecundity and spawning (Event 78)	**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.
Relationship: 94 Reduction, cumulative fecundity and spawning (event 78) leads to decrease, population trajectory (event 360)	**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.
Relationship: 32 Agonism, androgen receptor (Event 25) leads to Reduction, testosterone synthesis by ovarian theca cells (Event 274)	**WEAK** While there is no direct structural or functional relationship between these two KEs, it is understood how these KEs may plausibly be linked together (see Relationships 31, 143). While the connection is plausible, there remain significant uncertainties about exactly how AR agonism may elicit a negative feedback response and whether such mechanisms operate across species.
Relationship: 1384 Agonism, androgen receptor (Event 25) leads to Reduction, 17beta-estradiol synthesis by ovarian granulosa cells (Event 3)	**WEAK** While there is no direct structural or functional relationship between these two KEs, it is understood how these KEs may plausibly be linked together (see Relationships 31, 143, 302). While the connection is plausible, there remain significant uncertainties about exactly how AR agonism may elicit a negative feedback response and whether such mechanisms operate across species. As a result, the specific mechanistic linkages between AR agonism, and reduced E2 production remains uncertain.
Relationship: 1385 Agonism, androgen receptor (Event 25) leads to Reduction, vitellogenin synthesis in liver (Event 285)	**WEAK** While there is no direct structural or functional relationship between these two KEs, it is understood how these KEs may plausibly be linked together (see Relationships 31, 143, 302, 5, 252). While the connection is plausible, there remain significant uncertainties about exactly how AR agonism may elicit a negative feedback response and whether such mechanisms operate across species. As a result, the specific mechanistic linkages between AR agonism, and reduced E2 production remains uncertain.
Relationship: 1386 Reduction, plasma 17b-estradiol concentrations (Event 219) leads to Reduction, plasma vitellogenin concentrations (Event 221)	**STRONG** While there are intermediate biological steps between these two KEs, the mechanisms of that intermediary biology is very well understood and convincingly established, meeting the criteria of extensive documentation and broad acceptance.

2. Support for Essentiality of KEs	Defining Question	High (Strong)	Moderate	Low (Weak)
	Are downstream KEs and/or the AO prevented if an upstream KE is blocked?	Direct evidence from specifically designed experimental studies illustrating essentiality for at least one of the important KEs	Indirect evidence that sufficient modification of an expected modulating factor attenuates or augments a KE	No or contradictory experimental evidence of the essentiality of any of the KEs.

AOP: 23 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

Essentiality of the KEs was assessed for the AOP as a whole – rationale for the individual KE calls is provided.	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. **STRONG** • Ekman et al. 2011 provide evidence that in fathead minnow, cessation of trenbolone exposure resulted in recovery of T production, E2 production (both of which were recovering during the exposure due to compensation), plasma E2 and VTG concentrations, with T and E2 production recovering earlier in the time-course, followed by plasma E2, followed by plasma VTG. This provides some support for the essentiality of these key events. • Essentiality of the proposed negative feedback key event is supported by experimental work that evaluated the ability of AR agonists to reduce T or E2 production in vitro. In vitro exposure of fathead minnow ovary tissue to 17β-trenbolone or spironolactone does not cause reductions in T or E2 synthesis at concentrations comparable to those that produce significant responses in vivo (i.e., at non-cytotoxic concentrations; D.L. Villeneuve, unpublished data; C.A. LaLone unpublished data), nor are there any known reports of 17β-trenbolone directly inhibiting steroid biosynthesis. When tested in an in vitro steroidogenesis assay using H295R adrenal carcinoma cells, trenbolone caused a concentration-dependent increase in estradiol production, as opposed to any reductions in steroid hormone concentrations, an effect that was concurrent with increased transcription of CYP19 (aromatase) in the cell line (Gracia et al. 2007). • 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. • 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. • plasma vitellogenin concentrations, have been shown to recover in a predictable fashion consistent with the order of events in the AOP in stop/recovery studies. • With regard to vitellogenin accumulation into oocytes and oocyte growth/development, there are some contradictory evidence regarding the essentiality of this event. No stop/reversibility studies have explicitly considered this key event. • By definition, some degree of spawning is required to maintain population.

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
Relationship: 31 Agonism, Androgen receptor (Event 25) leads to gonadotropins circulating concentration (Event 129)	**WEAK** There are quite a few basic endocrine studies that establish that gonadectomy and AR antagonists can cause increased gonadotropin concentrations. Relatively few that directly demonstrate the opposite (i.e., exposure to androgen agonists decreasing circulating gonadotropin concentrations). **Dose-response:** Data compiled thus far are insufficient to evaluate dose-response concordance for this KER. **Temporality:** Data compiled thus far are insufficient to evaluate temporal concordance for this KER. **Uncertainties**: Significant uncertainties remain regarding how applicable this proposed KER is across species employing different reproductive strategies.			

AOP: 23 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

Relationship: 143 Reduction, gonadotropins, circulating concentrations (Event 129) leads to reduction, testosterone synthesis by ovarian theca cells (Event 274)	**STRONG** Empirical support for this KER is so strong it was deemed impractical to list all the empirical support. This relationship is considered dogma in the field of vertebrate endocrinology.
Relationship: 302 Reduction, testosterone synthesis by ovarian theca cells (Event 274) leads to reduction, 17b-estradiol synthesis by ovarian granulosa cells (Event 3)	**MODERATE (based on what has been assembled in the AOP description)** Given the strength of biological plausibility, only a few examples of empirical support for the linkage were cited. They support both an association between the two KEs and temporal concordance of the KEs.
Relationship: 5 Reduction, 17b-estradiol synthesis by ovarian granulosa cells (Event 3) leads to reduction, plasma 17b-estradiol concentrations (Event 219)	**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 Event 3 are generally observed at or near the same concentrations that impact Event 219. 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 Event 3. **Temporality:** Data from several time course studies, with at least two different aromatase inhibitors, support the idea that impacts on Event 3 are detected (statistically) at earlier time-points than impacts on Event 219. Data from these studies also show that Event 3 recovers before Event 219 both as the result of compensatory responses during an exposure period and following cessation of delivery of an aromatase inhibitor. In a time-course study with a strong androgen, a significant effect on Event 219 was observed before a significant effect on Event 3, however, throughout the time course it appears compensatory responses could more effectively offset the effect on production per unit mass of ovary tissue, than that on plasma E2 concentrations. **Incidence:** Particularly for experiments of longer duration (> 4 d), there are cases where impacts on Event 219 are detected without concurrent effects on Event 3. These are plausibly explained by the fact that compensatory responses in vivo lead to more rapid "recovery" of Event 3 than Event 219. It also reflects the fact that measures of Event 3 represent a rate of steroid production per unit mass of tissue, while Event 219 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.
Relationship 252: Reduction, plasma 17b-estradiol concentrations (Event 219) leads to reduction, vitellogenin synthesis in liver (Event 285)	**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. **Dose Response:** In one study that examined both Event 219 and Event 285, impacts on Event 219 were observed at much lower concentrations. However, the measurement technology (mass spectroscopy-based proteomics) employed for measuring Event 285 may be significantly more quantitative and precise than that employed for measuring Event 219. **Temporality:** There are currently no time-course studies included in the evidence assembly in which Event 219 and Event 285 were both measured. **Incidence:** In the only study that examined both Event 219 and Event 285, effects on both KEs were observed.
Relationship 315: Reduction, vitellogenin synthesis in liver (Event 285) leads to Reduction, plasma vitellogenin concentrations (Event 221)	**WEAK** Relatively few studies with stressors have measured both impacts on hepatic vitellogenin transcription or translation or protein concentrations in liver as well as plasma. , thus empirical data for evaluating this KER are limited. **Dose Response:** There are not sufficient empirical data assembled to evaluated the dose-response concordance of these key events. **Temporality:** There are not sufficient empirical data to evaluate the temporal concordance of these key events. **Incidence:** In the only study cited that examined both Event 285 and Event 221, effects on both KEs were observed.
Relationship: 255	**WEAK**

AOP: 23 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

Reduction, plasma vitellogenin concentrations (Event 221) leads to Reduction, vitellogenin accumulation into oocytes and oocyte growth/development (Event 309)	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 Event 309 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.
Relationship: 337 Reduction, vitellogenin accumulation in oocytes and oocyte growth/development (Event 309) leading to Reduction, cumulative fecundity and spawning (Event 78)	**MODERATE** There are only a few studies in which Events 309 and 78 were examined concurrently. **Dose Response**: There are a couple studies in which effects on Event 78 were reported at concentrations lower than those of Event 309. However, given the difference in how the endpoints are measured (histology for Event 309; Egg counts for Event 78), equal sensitivity in the measurement methods cannot really be assumed. **Temporality**: At present, there are no time-course data that directly address the temporal concordance between Events 309 and 78. **Incidence**: There are a number of studies in which events 309 and 78 co-occur and no known cases where Event 309 was observed, but Event 78 was not.
Relationship: 94 Reduction, cumulative fecundity and spawning (event 78) leads to decrease, population trajectory (event 360)	**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 Events 78 and 360.
Relationship: 32 Agonism, androgen receptor (Event 25) leads to Reduction, testosterone synthesis by ovarian theca cells (Event 274)	**MODERATE** There are a number of studies with different AR agonists and different species that show a dependent association between Event 25 and Event 274
Relationship: 1384 Agonism, androgen receptor (Event 25) leads to Reduction, 17beta-estradiol synthesis by ovarian granulosa cells (Event 3)	**MODERATE** There are a number of studies with different AR agonists and different species that show a dependent association between Event 25 and Event 3. However, the data are not adequate to extensively evaluate dose-response or temporal concordance.
Relationship: 1385 Agonism, androgen receptor (Event 25) leads to Reduction, vitellogenin synthesis in liver (Event 285)	**STRONG** In over 16 independent experiments in fish, encompassing four species, and four different AR agonists, reductions in VTG have been observed, often along with evidence of masculinization of secondary sex characteristics in the female fish. Given that AR agonism cannot be measured in vivo, other than as tissue remodeling associated with altered secondary sex characteristics it is experimentally challenging to test temporal concordance, but dose dependence (a reasonable surrogate for dose-concordance when one KE is the MIE) of the VTG response is evident.
Relationship: 1386 Reduction, plasma 17b-estradiol concentrations (Event 219) leads to Reduction, plasma vitellogenin concentrations (Event 221)	**STRONG** There is an extensive amount of evidence across multiple species, and with multiple stressors that supports the relationship. In general, the data are concordant with the expected patterns, and in cases where that is not the case, there are reasonable technical/experimental reasons for the apparent lack of concordance.

AOP: 23 – 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
Relationship: 31 Agonism, Androgen receptor (Event 25) leads to gonadotropins circulating concentration (Event 129)	**WEAK** – While there is a foundation of plausibility, there are many uncertainties regarding the mechanistic linkage between these events and very little supporting evidence has been assembled, to date.
Relationship: 143 Reduction, gonadotropins, circulating concentrations (Event 129) leads to reduction, testosterone synthesis by ovarian theca cells (Event 274)	**STRONG** – There is extensive support for this relationship. So much, that assembly of empirical evidence is intractable.
Relationship: 302 Reduction, testosterone synthesis by ovarian theca cells (Event 274) leads to reduction, 17b-estradiol synthesis by ovarian granulosa cells (Event 3)	**STRONG** – This relationship is well documented and broadly accepted. Empirical evidence for specific stressor studies adds further weight.
Relationship: 5 Reduction, 17b-estradiol synthesis by ovarian granulosa cells (Event 3) leads to reduction, plasma 17b-estradiol concentrations (Event 219)	**STRONG** – Strong biological plausibility augmented by strong empirical support.
Relationship 252: Reduction, plasma 17b-estradiol concentrations (Event 219) leads to reduction, vitellogenin synthesis in liver (Event 285)	**STRONG** – This relationship is well document and broadly accepted. The established plausibility outweighs the meager assembly of empirical evidence.
Relationship 315: Reduction, vitellogenin synthesis in liver (Event 285) leads to Reduction, plasma vitellogenin concentrations (Event 221)	**STRONG** – Hepatic synthesis is the unquestioned primary source of circulating vitellogenin, therefore plausibility offsets the weak empirical support assembled.
Relationship: 255 Reduction, plasma vitellogenin concentrations (Event 221) leads to Reduction, vitellogenin accumulation into oocytes and oocyte growth/development (Event 309)	**MODERATE** Fairly well established plausibility but with some gaps in understanding, along with weak empirical support.
Relationship: 337 Reduction, vitellogenin accumulation in oocytes and oocyte growth/development	**MODERATE** Moderate plausibility and empirical support.

AOP: 23 – Annex 1, assessment of the relative level of confidence in the overall AOP based on rank ordered weight of evidence elements.

(Event 309) leading to Reduction, cumulative fecundity and spawning (Event 78)	
Relationship: 94 Reduction, cumulative fecundity and spawning (event 78) leads to decrease, population trajectory (event 360)	**MODERATE**: Strong plausibility, for the importance of cumulative fecundity and spawning as a vital rate influencing population. However, given the influence of other vital rates like survival to reproductive age, and different reproductive strategies, interplay relative to competition for resources and habitat etc. in the ambient environment, the overall plausibility is only moderate. Further, it is impractical to extensively test the relationship empirically.
Relationship: 32 Agonism, androgen receptor (Event 25) leads to Reduction, testosterone synthesis by ovarian theca cells (Event 274)	**MODERATE**: Some uncertainty regarding the exact mechanistic linkage involved, but the accumulation of correlative/associative evidence strengthens the overall AOP.
Relationship: 1384 Agonism, androgen receptor (Event 25) leads to Reduction, 17beta-estradiol synthesis by ovarian granulosa cells (Event 3)	**MODERATE**: Some uncertainty regarding the exact mechanistic linkage involved, but the accumulation of correlative/associative evidence strengthens the overall AOP.
Relationship: 1385 Agonism, androgen receptor (Event 25) leads to Reduction, vitellogenin synthesis in liver (Event 285)	**STRONG**: Some uncertainty regarding the exact mechanistic linkage involved, but the accumulation of correlative/associative evidence is quite compelling for this KE, as the effect has been observed for multiple species and multiple stressors and repeated many times.
Relationship: 1386 Reduction, plasma 17b-estradiol concentrations (Event 219) leads to Reduction, plasma vitellogenin concentrations (Event 221)	**STRONG**: Very well supported in terms of both plausibility and empirical support.