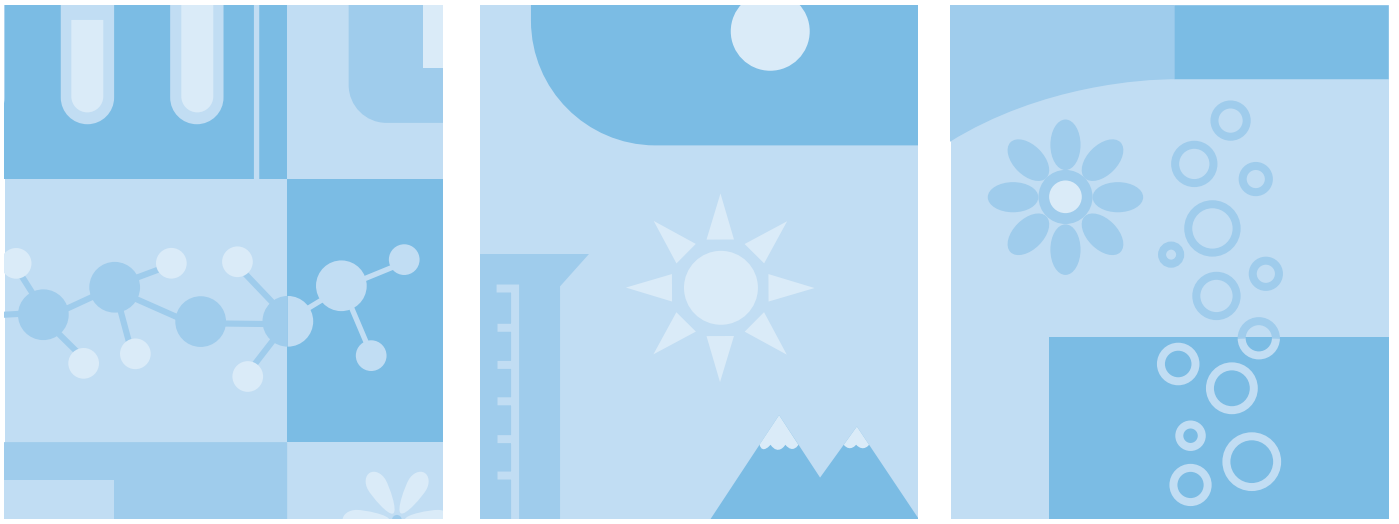
PM
2/13

Is it possible to determine thresholds for the effects of endocrine disruptors?

– A summary of scientific argumentation from 15 relevant publications on endocrine disruption



Anna Beronius och Annika Hanberg IMM, Karolinska Institutet

Swedish Chemicals Agency – www.kemi.se

KEMI
Swedish Chemicals Agency

Is it possible to determine thresholds for the effects of endocrine disruptors?

– A summary of scientific argumentation from 15 relevant publications on endocrine disruption

Order No. 511 086
Sundbyberg, February 2013
Publisher: Swedish Chemicals Agency©
Order address: CM Gruppen, P.O. Box 11063, SE-161 11 Bromma, Sweden
Phone: +46 8 5059 33 35, Fax +46 8-5059 33 99, E-mail: kemi@cm.se
The report is available as a downloadable pdf on www.kemi.se

Preface

The Swedish Chemicals Agency (KemI) has been assigned by the Swedish Government to produce a national action plan for a toxic-free everyday environment: Action plan for a toxic-free everyday environment 2011 – 2014 – protect the children better.

Efforts are now going on in several areas, both in Sweden, within the EU and internationally and often in cooperation with other authorities. Reducing chemical risks in the everyday environment is one step towards attaining the Swedish Parliament´s environment quality objective A Non-Toxic Environment, which is the objective that KemI is responsible for.

Within the framework of the action plan, KemI compiles knowledge in KemI´s report and PM series elaborated by experienced colleagues, researchers or consultants. In this way, KemI presents new and essential knowledge in publications which can be downloaded from the website www.kemikalieinspektionen.se

Endocrine disruptors are one of the areas prioritised in the action plan. Sweden has ambitions to actively contribute to the processes going on in the EU and internationally in this area. The EU is making active efforts to develop criteria for endocrine disruptors and to agree on a new EU strategy on endocrine disrupters. In addition, the Commission under Article 138 (7) of REACH is by 1 June 2013, to make a review of the scientific situation to determine if it believes that it is possible to establish thresholds (PNEC) for endocrine disruptors.

In order to take a position on the question of the existence of a threshold or not for endocrine disruptors from a regulatory perspective and to provide input to and guide the Commission on their decision in relation to Article 138, paragraph 7 of REACH, the Swedish Chemicals Agency commissioned the Institute of Environmental Medicine (IMM) to produce a compilation of available scientific arguments about whether it is possible to establish a threshold or not for endocrine disruptors.

The report contains a summary of the arguments for or against a threshold for safe levels of endocrine disruptors to be determined and a summary (Appendix 1) of scientific publications with arguments that support or reject the hypothesis that it is not possible to determine a safe level of endocrine disruptors.

The report was written by Anna Beronius and Annika Hanberg IMM, Karolinska Institutet.

Förord

Kemikalieinspektionen (KemI) har på uppdrag av regeringen tagit fram en handlingsplan ”Handlingsplan för en giftfri vardag 2011–2014 – Skydda barnen bättre”. Insatser sker nu på flera områden både nationellt, inom EU och internationellt och ofta i samarbete med andra myndigheter.

Att minska kemiska risker i vardagen är ett steg på vägen att nå riksdagens miljökvalitetsmål Giftfri miljö – det mål KemI ansvarar för.

Inom ramen för handlingsplanen tar KemI fram kunskapssammanställningar, som publiceras i KemI:s rapport- respektive PM-serie. Bakom publikationerna står egna medarbetare, forskare eller konsulter. KemI vill på detta sätt dela med sig av ny och angelägen kunskap. Publikationerna, som är kostnadsfria, finns på webbplatsen www.kemikalieinspektionen.se

Hormonstörande ämnen är ett av de områden som prioriterats i handlingsplanen. Sverige har

höga ambitioner att aktivt bidra till de processer som pågår både inom EU och internationellt på området. Inom EU pågår ett aktivt arbete med att ta fram kriterier för hormonstörande ämnen samt att enas om en ny EU-strategi för hormonstörande ämnen. Dessutom ska Kommissionen enligt artikel 138(7) i Reach, senast den 1 juni 2013, göra en översyn av det vetenskapliga läget för att avgöra om man anser att det går att fastställa gränsvärden (DNEL eller PNEC) för hormonstörande ämnen.

För att besluta om hur KemI ska ställa sig till frågan om förekomst av tröskel eller ej ur ett regulatoriskt perspektiv och för att bidra med underlag till och påverka Kommissionen inför deras beslut i relation till artikel 138 punkt 7 i Reach har KemI uppdragit åt Institutet för Miljömedicin (IMM) att ta fram en sammanställning av tillgänglig vetenskaplig argumentationen kring frågan om det går att fastställa en tröskel eller ej för hormonstörande ämnen.

Rapporten innehåller en sammanfattning av argument som talar för eller emot att en tröskel för säkra nivåer av hormonstörande ämnen går att fastställa samt en sammanställning (bilaga 1) av vetenskapliga publikationer med argument som stödjer eller förkastar hypotesen att det inte går att bestämma en säker nivå för hormonstörande ämnen.

Rapporten har skrivits av Anna Beronius och Annika Hanberg IMM, Karolinska Institutet.

Content

Scope and purpose

In December 2012 the Swedish Chemicals Agency (KemI) asked the Institute of Environmental Medicine (IMM) at Karolinska Institutet to review selected scientific literature and summarize arguments for and against the existence of a threshold for effect, i.e. below which no adverse effects are observed, for endocrine disrupting chemicals (EDCs). Reports and scientific articles to be included in the review were identified and ordered in terms of priority by KemI.

What is a threshold?

The threshold for effect may be defined in different ways, which may be relevant to the arguments made for or against a threshold. Slob (1999) provided three different definitions for "threshold":

1. *Biological definition*: The dose below which the organism does not suffer from any (adverse) effects from the compound considered.
2. *Experimental definition*: The dose below which no effects are observed.
3. *Mathematical definition*: The dose below which the response is zero, and above which it is nonzero.

Review of selected literature

Fifteen references were reviewed (Table 1). These ranged from toxicological and ecotoxicological research studies to reviews and state-of-the art reports. The studies by Sheenan et al. (1999) and Calabrese (2008) were not on the initial list provided by KemI, these studies were added to the review by IMM since they were referred to in many of the other articles and reports.

Arguments for or against assuming a threshold stated in the reviewed articles and reports for EDCs were summarized in a table in Microsoft Excel. Arguments were given as direct citations whenever possible (indicated by citation marks). However, in some cases paraphrasing was necessary.

Table 1. References included in the review and ordered in terms of priority.

Reference	**Title**
Kortenkamp et al. 2012	State of the Art Assessment of Endocrine Disrupters (SoA)
Blair et al. 2001	Threshold analysis of selected dose-response data for endocrine active chemicals.
Sheehan 2006	No-threshold dose–response curves for nongenotoxic chemicals: Findings and applications for risk assessment.
Conolly and Lutz 2004	Nonmonotonic Dose-Response Relationships: Mechanistic Basis, Kinetic Modeling, and Implications for Risk Assessment.
Fawcett et al. 1996	Is there a no-effect dose for corticosteroid-induced cleft palate? The contribution of endogenous corticosterone to the incidence of cleft palate in mice.
Rhomberg and Goodman 2012	Commentary: Low-dose effects and nonmonotonic dose–responses of endocrine disruptingchemicals: Has the case been made?
Bärlocher et al. 2011	Effects of 4-n-nonylphenol on aquatic hyphomycetes.
Vandenberg et al. 2012	Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses.
Scholze and Kortenkamp 2007	Statistical power considerations show the endocrine disrupter low dose issue in a new light.
Gross et al. 2010	Thresholds of toxicological concern for endocrine active substances in the aquatic environment.
Calabrese 2008	Hormesis and medicine.
Slob 1999	Thresholds in toxicology and risk assessment
White et al. 2009	State-of-the-science workshop report: issues and approaches in low-dose-response extrapolation for environmental health risk assessment.
Welshons et al. 2003	Large effects from small exposures. I. Mechanisms for endocrine-disrupting chemicals with estrogenic activity.
Sheenan et al. 1999	No threshold dose for estradiol-induced sex reversal of turtle embryos: how little is too much?

Conclusions

The presence or absence of a threshold can never be experimentally proven (e.g. Kortenkamp et al. 2012; Sheehan et al. 1999; Slob 1999). All experiments have a limit of detection below which effects cannot be observed, i.e. no conclusion regarding the shape of the dose-response curve can be made below this detection limit. Also, to generate an exact dose-response curve would require an infinite number of doses and infinitely precise measures (Slob 1999).

The literature reviewed here provides arguments both for and against assuming a threshold for EDCs. The general argument for assuming no threshold for EDCs was that compounds that act by the same mechanism as endogenous factors, e.g. hormones, just add to the actions of these factors and increase the response of already ongoing biological processes. This "additivity-to-background" argument has also been made to defend a no-threshold-approach for genotoxic carcinogens (Slob 1999).

On an animal experiment or population level thresholds may be "masked" by individual variation. In other words, even if a threshold *does* exist for a certain endpoint the threshold-dose may vary significantly between individuals and the threshold may therefore not be observable. This has been discusses for e.g. epidemiological studies (White et al. 2009).

The arguments made in support of a threshold for EDCs were mainly that, as stated for example by Blair et al. (2001): "…a threshold could be expected if there is no endogenous hormone, if the endogenous hormone induces no adverse effect, or if there is effective homeostatic control." Also, Conolly and Lutz (2004) state that the first interaction of a toxic agent with its primary biological target molecule is likely to have no threshold but imply that the complexity of a biological system makes non-threshold dose-response curves unlikely for many "higher" endpoints, such as behavior, reproduction, organ weights and growth.

The scientific support for assuming a threshold may thus depend on the endpoint under investigation and what is known about its mechanism of action. It follows that the discussions regarding a threshold in risk assessment of EDCs is tightly connected with discussions concerning what types of effects should be considered "adverse".

Based on this literature review it seems that the decision on whether or not to accept a non-threshold model for EDCs has to be based on considerations of mechanism of action. Thus, the assumption of no threshold may be as valid, or questionable, for EDCs as for genotoxic carcinogens.

References

Barlocher, F., Guenzel, K., Sridhar, K.R., Duffy, S.J., 2011. Effects of 4-n-nonylphenol on aquatic hyphomycetes. Sci Total Environ 409, 1651-1657.

Blair, R.M., Fang, H., Gaylor, D., Sheehan, D.M., 2001. Threshold analysis of selected dose-response data for endocrine active chemicals. APMIS 109, 198-208.

Calabrese, E.J., 2008. Hormesis and medicine. Br J Clin Pharmacol 66, 594-617.

Conolly, R.B., Lutz, W.K., 2004. Nonmonotonic dose-response relationships: mechanistic basis, kinetic modeling, and implications for risk assessment. Toxicol Sci 77, 151-157.

Fawcett, L.B., Buck, S.J., Beckman, D.A., Brent, R.L., 1996. Is there a no-effect dose for corticosteroid-induced cleft palate? The contribution of endogenous corticosterone to the incidence of cleft palate in mice. Pediatr Res 39, 856-861.

Gross, M., Daginnus, K., Deviller, G., de Wolf, W., Dungey, S., Galli, C., Gourmelon, A., Jacobs, M., Matthiessen, P., Micheletti, C., Nestmann, E., Pavan, M., Paya-Perez, A., Ratte, H.T., Safford, B., Sokull-Kluttgen, B., Stock, F., Stolzenberg, H.C., Wheeler, J., Willuhn, M., Worth, A., Comenges, J.M., Crane, M., 2010. Thresholds of toxicological concern for endocrine active substances in the aquatic environment. Integr Environ Assess Manag 6, 2-11.

Kortenkamp, A., Martin, O., Faust, M., Evans, R., McKinlay, R., Orton, F., Rosivatz, E., 2012. State of the Art Assessment of Endocrine Disrupters - Final Report. EU Commission DG Environment.

Rhomberg, L.R., Goodman, J.E., 2012. Low-dose effects and nonmonotonic dose-responses of endocrine disrupting chemicals: has the case been made? Regul Toxicol Pharmacol 64, 130-133.

Scholze, M., Kortenkamp, A., 2007. Statistical power considerations show the endocrine disruptor low-dose issue in a new light. Environ Health Perspect 115 Suppl 1, 84-90.

Sheehan, D.M., 2006. No-threshold dose-response curves for nongenotoxic chemicals: findings and applications for risk assessment. Environ Res 100, 93-99.

Sheehan, D.M., Willingham, E., Gaylor, D., Bergeron, J.M., Crews, D., 1999. No threshold dose for estradiol-induced sex reversal of turtle embryos: how little is too much? Environ Health Perspect 107, 155-159.

Slob, W., 1999. Thresholds in toxicology and risk assessment. International Journal of Toxicology 18, 259-268.

Vandenberg, L.N., Colborn, T., Hayes, T.B., Heindel, J.J., Jacobs, D.R., Jr., Lee, D.H., Shioda, T., Soto, A.M., vom Saal, F.S., Welshons, W.V., Zoeller, R.T., Myers, J.P., 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. Endocr Rev 33, 378-455.

Welshons, W.V., Thayer, K.A., Judy, B.M., Taylor, J.A., Curran, E.M., vom Saal, F.S., 2003. Large effects from small exposures. I. Mechanisms for endocrine-disrupting chemicals with estrogenic activity. Environ Health Perspect 111, 994-1006.

White, R.H., Cote, I., Zeise, L., Fox, M., Dominici, F., Burke, T.A., White, P.D., Hattis, D.B., Samet, J.M., 2009. State-of-the-science workshop report: issues and approaches in low-dose-response extrapolation for environmental health risk assessment. Environ Health Perspect 117, 283-287.

	A	B	C	D	E	F
1	**Reference**	**Abstract**	**Arguments against threshold**	**Arguments for threshold**	**Additional comments by the author(s)**	**IMM comment**
2	Kortenkamp et al. 2012		"Because of pre-existing internal exposures to steroidal estrogens, it can be inferred that any quantum of externally added estrogenic agent adds to the internal load, thereby exhibiting activity in a threshold-independent fashion." The threshold concept may be questioned for non-carcinogenic chemicals in general. Several epidemiological studies have reported that human health risks increased linearly with dose in the low dose range for ozone, tobacco smoke, nitric oxide and sulphur dioxide, particulate matter and lead.	In epidemiological studies a threshold may theoretically exist on the individual level but be obscured at the population level because of inter-individual variations in sensitivity or by background (exogenous or endogenous) exposures.	"The existence of dose thresholds cannot be proven or ruled out by experimental approaches, because all methods for measuring effects have their limits of detection which will obscure thresholds, if they exist." Only assumptions regarding the absence of a threshold can be made, and this could e.g. be based on a compound's mode of action, such as for genotoxic carcinogens.	
3	Blair et al. 2001	Using a biologically relevant mathematical model, the Michaelis-Menten equation, we examined published data from endocrine active chemicals for evidence of no-threshold dose-response curves. Data were fit to a modified Michaelis-Menten equation which accounted for total background response. Subsequently, the data sets were analyzed using non-linear regression in order to estimate the four parameters of interest (non-hormone controlled background (Bnh), maximum response (Rmax), endogenous hormone level (D0), and the dose at which a half-maximal response was observed (ED50)) and to determine the fit to the fully modified Michaelis-Menten equation. Subsequently, response data were adjusted to account for Bnh and then normalized to Rmax, while dose data were adjusted to account for D0 and then normalized to the ED50. This data set was combined into a single, composite data set and fit to the fully modified Michaelis-Menten equation. We examined 31 data sets (24 endpoints) from studies on 9 different chemical/hormone treatments. Twenty-six of the data sets fit the modified Michaelis-Menten equation with high multiple correlation coefficients (r>0.90). The normalized data demonstrated a good fit to the modified Michaelis-Menten equation. These results indicate that a variety of biological responses fit the modified Michaelis-Menten equation, which does not have a threshold dose term.	Twenty-six out of 31 data sets reviewed in this study could be fitted to the adjusted Michaelis-Menten equation with high correlation coefficients (>0.90), suggesting that they did not show a threshold. The data sets were collected from 12 different studies investigating a wide variety of "physiological" and "adverse" responses to estradiol, TCDD, nafoxidine, PCBs, oxazepam, phenylhydroxylamine, DES, endosulfan and "conjugated estrogens". The authors state that "...because endocrine disruptors can alter normal function by binding to hormone receptors, both theory and data suggest that a threshold will not exist."	"...a threshold could be expected if there is no endogenous hormone, if the endogenous hormone induces no adverse effect, or if there is effective homeostatic control."	While there is no threshold term in the Michaelis-Menten equation the authors argue that a threshold would still be recognized using this method since in that case the curve at zero response would strike the X-axis at some positive value.	
4	Sheehan 2006	We tested the hypothesis that no threshold exists when estradiol acts through the same mechanism as an active endogenous estrogen. A Michaelis-Menten (MM) equation accounting for response saturation, background effects, and endogenous estrogen level fit a turtle sex-reversal data set with no threshold and estimated the endogenous dose. Additionally, 31 diverse literature dose-response data sets were analyzed by adding a term for nonhormonal background; good fits were obtained but endogenous dose estimations were not significant due to low resolving power. No thresholds were observed. Data sets were plotted using a normalized MM equation; all 178 data points were accommodated on a single graph. Response rates from approximately 1% to >95% were well fit. The findings contradict the threshold assumption and low-dose safety. Calculating risk and assuming additivity of effects from multiple chemicals acting through the same mechanism rather than assuming a safe dose for nonthresholded curves is appropriate.	"Because many endocrine disruptors act against a background of endogenous hormones and bind to receptors... data suggest that a threshold will not exist ... Thus, not only for endocrine disruptors but also for all other toxicants that act against a background of endogenous chemicals (or one or more exogenous chemicals acting via the same mechanism), the threshold hypothesis and its derived NOAEL/RfD appear inappropriate."	"...a threshold could occur if a particular hormone is absent at some life stage, or if the hormone is innocuous, or if homeostasis, if it exists, is essentially perfect."	"Given that there is evidence that supports both threshold and nonthreshold models for risk assessments, an appropriate approach is to regress the curve to the x-axis using the best-fit relevant equation after accounting for additive and nonadditive sources of background. If the lower 95% confidence interval strikes the x-axis at a positive value then a threshold model is appropriate. On the other hand, if the upper confidence interval regresses to the origin or the negative x-axis then, a nonthreshold model should be used. If the confidence intervals bracket the origin, then other information, such as the mechanism of action, should be used to determine the appropriate model. Rather than an ideological one-size-fits all policy, the approach suggested here is pragmatic and data driven... Genetically heterogenous populations, such as humans and wildlife, are expected to have a broader distribution of thresholds than most laboratory animal strains... Thus, there will be a population distribution of sensitivity to mimics of endogenous chemicals, such as hormones, and no population threshold is possible"	Sheenan challenges the statement that a threshold can never be proven in an animal study. But the argument given is that this would not apply if an endogenous chemical acts through a mechanism that leads to an adverse effect because then the threshold is already exceeded. This seems to not address the issue of whether a threshold can be shown but is rather a statement saying that it would be irrelevant to try to show a threshold.
5	Conolly and Lutz 2004	Dose-response curves for the first interaction of a chemical with a biochemical target molecule are usually monotonic; i.e., they increase or decrease over the entire dose range. However, for reactions of a complex biological system to a toxicant, nonmonotonic (biphasic) dose-effect relationships can be observed, showing a decrease at low dose followed by an increase at high dose, or vice versa. We present four examples to demonstrate that nonmonotonic dose-response relationships can result from superimposition of monotonic dose responses of component biological reactions. Examples include (i) a membrane-receptor model with receptor subtypes of different ligand affinity and opposing downstream effects (adenosine receptors A1 vs. A2), (ii) androgen receptor-mediated gene expression driven by homodimers, but not mixed-ligand dimers, (iii) repair of background DNA damage by enzymatic activity induced by adducts formed by a xenobiotic, (iv) rate of mutation as a consequence of DNA damage times rate of cell division, the latter being modulated by cell-cycle delay at low-level DNA damage, and cell-cycle acceleration due to regenerative hyperplasia at cytotoxic dose levels. Quantitative analyses based on biological models are shown, and factors that affect the degree of nonmonotonicity are identified. It is noted that threshold-type dose-response curves could in fact be nonmonotonic. Our analysis should promote a scientific discussion of biphasic dose responses and the concept termed "hormesis," and of default procedures for low-dose extrapolation in toxicological risk assessment.	For molecular events, e.g. the first interaction of a toxic agent with its primary biological target molecule, "a linear default extrapolation to low dose appears to be appropriate".	No arguments made clearly as this is not the purpose of the paper but implied that even though response may be linear at a molecular level the complexity of a biological system makes linear (i.e. no threshold) dose-response curves unlikely.	Non-monotonicity may not always be evident and the dose-response curve may appear to have a threshold: "For a low-control activity and a small decrease at low dose, the nonmonotonic shape of the dose response might not be evident. The curve can appear as threshold. Numerous toxicological data sets show this type of dose response, which is also seen in some of the curves shown in our figures. Changing critical model parameters can lead to a group of curves that often include a threshold-like example."	This paper mainly explains different mechanisms that could explain non-monotonicity on a molecular level, i.e. cAMP formation, gene transcription, repair of DNA-damage and mutation rate. The authors also challenge the default assumption that genotoxic carcinogens have a linear dose-response curve because although linearity is plausible on the "first interaction" molecular level "the linear default assumption cannot reflect the complexity of responses of a biological system". There is no specific discussion concerning whether or not exposure to substances such as EDCs could result in dose-response without threshold.

	A	B	C	D	E	F
6	Fawcett et al. 1996	Teratology and genetic counselors are frequently asked whether very low exposures of drugs and chemicals can cause a child's congenital malformations. One critical factor on which the counseling is based is the dose. Because teratogenic effects follow a toxicologic dose-response curve with a no-effect dose, frequently counselors can refute a causal relationship because the dose was far below the no-observable-effect dose. Recently, some investigators have suggested that some teratogens which are present in physiologic levels such as cortisone, glucose, insulin, or sex steroids may contribute to the background incidence of congenital malformations and, therefore, there is no safe dose. Using corticosteroid-induced cleft palate in mice as the model, we conducted experiments to test this hypothesis. Adrenalectomy of A/J or CD-1 dams resulted in a reduction of endogenous corticosterone, but did not reduce the spontaneous incidence of cleft palate in the offspring. In A/J mice, the incidence of isolated cleft palate increased with adrenalectomy indicating that the spontaneous incidence of this defect is not due to endogenous corticosterone. Adrenalectomy did not affect the susceptibility of CD-1 mice to cortisone induced cleft palate demonstrating that endogenous corticosterone did not contribute significantly to the incidence of cleft palate induced by the exogenous corticosteroid. Finally, results in CD-1 mice clearly indicate that cortisone, like other teratogens, has a no-effect level for teratogenesis. These studies support the concept of a threshold in the dose-response relationship for corticosteroid-induced cleft palate in mice.	*none made*	In this study sham adrenalectomized and adrenalectomized pregnant mice were injected with 12.5, 25, 50 and 100 mg cortisone/kg bw. Litters from both sham adrenalectomized and adrenalectomized dams injected with 25 mg of cortisone/kg or greater had an increased incidence of CP when compared with controls. The authors conclude that there is a threshold for corticosteroid-induced cleft palate in mice and that the background exposure of endogenous corticosterone does not contribute to the effect of the exogenous exposure of cortisone.		In this study, even though no teratogenic effects were observed, the lowest dose of cortisone (12.5 mg/kg bw) significantly suppressed maternal corticosterone secretion in sham adrenalectomized dams, presumably due to negative feedback of the pituitary-adrenal axis. It can be argued that teratogenicity may be a rather insensitive endpoint and also that a linear dose-response curve is unlikely for such an effect. Thus this study may not be very useful in discussing a threshold in the dose-response for EDCs.
7	Rhomberg and Goodman 2012	Vandenberg et al. (2012) claim that "most if not all [endocrine-disrupting chemicals (EDCs)] are likely to have low-dose effects" and "nonmonotonicity is a common occurrence after exposures to hormones and EDCs in cell culture and animals and across human populations." They present examples as anecdotes without attempting to review all available pertinent data, selectively citing studies without evaluating most of them or examining whether their putative examples are consistent and coherent with other relevant information. They assume that any statistically significant association indicates causation of an adverse effect, and their limited evaluation of specific studies is not done uniformly (i.e., studies with positive results are evaluated differently than those with null results). They also do not evaluate whether exposures in studies are truly "low-dose" and relevant to humans. They propose a number of different nonmonotonic dose-response curves, but do not consider reasons for why they should be expected to apply generally across species. Many of their examples would be - and indeed have been - questioned by many scientists. Overall, Vandenberg et al. put forth many asserted illustrations of their two conclusions without providing sufficient evidence to make the case for either and while overlooking evidence that suggests the contrary.	*none made*	*none made*	"In our view, the case for widespread nonmonotonicity leading to undetected toxicity at low doses has not been made..."	This commentary is a response to the review by Vandenberg et al (2012) and mainly discusses the issues of the existence of low-dose effects and non-monotonicity. While it brings up important criticisms of that review it does not contribute any arguments for or against a threshold for EDCs.
8	Bärlocher et al. 2011	We measured the removal of 4-n-nonylphenol (between 50 and 500 µg L(-1)) from an aqueous solution with or without linden and oak leaf disks. More 4-n-NP was removed when the leaves were first exposed for 3 weeks in a stream, which allowed colonization by aquatic hyphomycetes. The response of fungal sporulation rates from beech, linden, maple and oak leaves to increasing levels of 4-n-NP was complex. Linear regressions were non-significant, arguing against a no-threshold model. The response at the lowest concentration (50 µg L(-1)) was between 7% (beech) and 67% (maple) higher than in the absence of 4-n-NP, however, the difference was not significant. The number of sporulating species of aquatic hyphomycetes was significantly higher at the lowest concentration than in the control treatment without 4-n-NP. The composition of the fungal community was affected by leaf species but not by 4-n-NP concentration. The results suggest the presence of a weak hormeotic effect. The known ability of aquatic hyphomycetes and other fungi to degrade nonylphenols and related substances, combined with fungal resilience in their presence, makes decaying leaves potential candidates for bioremediation.	*none made*	In this study the effects of 4 different concentrations (50, 100, 250 and 500 µg/L) of 4-n-NP on fungal sporulation from linden, maple, oak and beech leaves, respectively, was investigated. No significant effects compared to controls were observed at the lowest dose for any leaf type, indicating a threshold for effect.	For all leaf types a slight induction (not statistically significant) of sporulation was observed at the lowest or second lowest 4-n-NP dose. The authors point out that these findings are consistent with a hormetic effect as it has been described by Calabrese (2008).	The stated objective of this study was to "investigate the impact of aquatic hyphomycetes on the fate of nonylphenol". However, the effects of 4-n-NP on fungal sporulation were also investigated and discussed and is the part of the study that has relevance for this literature review. Significance of effects was not very clearly reported in this study, presumably because of the complexity of the dose-response curves observed. Significant results are only described in text. Interpretation of results by the reader would be facilitated by some indication of significant results in the figures. It should be noted that no significant effect seemed to have been observed in any of the leaf types for any dose compared to controls. Significant differences were only described between the highest and lowest (maple, oak and beech leaves) or second to lowest (linden leaves) dose levels.

	A	B	C	D	E	F
9	Vandenberg et al. 2012	For decades, studies of endocrine-disrupting chemicals (EDCs) have challenged traditional concepts in toxicology, in particular the dogma of "the dose makes the poison," because EDCs can have effects at low doses that are not predicted by effects at higher doses. Here, we review two major concepts in EDC studies: low dose and nonmonotonicity. Low-dose effects were defined by the National Toxicology Program as those that occur in the range of human exposures or effects observed at doses below those used for traditional toxicological studies. We review the mechanistic data for low-dose effects and use a weight-of-evidence approach to analyze five examples from the EDC literature. Additionally, we explore nonmonotonic dose-response curves, defined as a nonlinear relationship between dose and effect where the slope of the curve changes sign somewhere within the range of doses examined. We provide a detailed discussion of the mechanisms responsible for generating these phenomena, plus hundreds of examples from the cell culture, animal, and epidemiology literature. We illustrate that nonmonotonic responses and low-dose effects are remarkably common in studies of natural hormones and EDCs. Whether low doses of EDCs influence certain human disorders is no longer conjecture, because epidemiological studies show that environmental exposures to EDCs are associated with human diseases and disabilities. We conclude that when nonmonotonic dose-response curves occur, the effects of low doses cannot be predicted by the effects observed at high doses. Thus, fundamental changes in chemical testing and safety determination are needed to protect human health.	"...the assumption of a threshold does not take into account situations where an endogenous hormone is already above the dose that causes detectable effects and that an exogenous chemical (whether an agonist or antagonist) will modulate the effect of the endogenous hormone at any dose above zero (Fig. 4B). There can thus be no threshold or safe dose for an exogenous chemical in this situation. Forced identification of NOAEL or threshold doses based on the assumption that dose-response curves are always monotonic without considering the background activity of endogenous hormones and the limitations of analytical techniques supports the misconception that hormonally active agents do not have any significant biological effects at low doses... Assumptions used in chemical risk assessments to estimate a threshold dose below which daily exposure to a chemical is estimated to be safe are false for EDCs... experimental data provide evidence for the lack of a threshold for EDCs [Sheehan et al. 2006]."	*none made*		The purpose of this review was to summarize evidence for the presence of low-dose effects and non-monotonic dose-response relationships for EDCs. Specific arguments against a threshold are not made very consistently (see "Arguments against threshold"). For example, in one place in the review it is stated that it is very plausible that dose-response curves for EDCs may lack a threshold under certain circumstances depending on the mechanism of action for these compounds (similar argument as made by Blair et al., 2001). However, later in the article a very strong statement is made that the assumption of a threshold is completely false for EDCs. It also seems like the authors erroneously equated monotonicity with the presence of a threshold: "...identification of NOAEL or threshold doses based on the assumption that dose-response curves are always monotonic ...". Clearly, monotonicity does not require that there is a threshold, just as non-monotonicity does not, in theory, preclude a threshold.
10	Scholze and Kortenkamp 2007	BACKGROUND: The endocrine disruptor field has been vexed by difficulties in reproducing various claims of effects at unusually low doses. In previous analyses, variations in control responses from experiment to experiment and problems with observing effects in positive controls have been identified as possible explanations of the resulting impasse. OBJECTIVE: In this article, we argue that both of these viewpoints fail to take sufficient account of the problems that exist in estimating low effects and low-effect doses. We have carried out post hoc power analyses on selected published data to illustrate that claims of low-dose effects (or their absence) are often compromised by insufficient statistical power of the chosen experimental design. CONCLUSIONS: We demonstrate that low-dose estimates such as the no observed adverse effect levels derived from statistical hypothesis-testing procedures are dependent on the specific experimental conditions used for testing. Thus, below the statistical detection limit of the experiment, the presence of effects can neither be proven nor ruled out. Common practice is to attempt to establish "doses without effect." However, low-dose estimations in the endocrine-disruptor field could be improved if decisions regarding the toxicologic effect size of relevance formed the starting point of testing procedures. Statistical power considerations could then reveal the resources necessary to demonstrate effect magnitudes of concern.	*none made*	*none made*	Toxicity testing, and establishing thresholds for effect, of EDCs is often hampered by the fact that a relevant or critical effect level is not clearly defined. Establishing a relevant effect level is a prerequisite for designing a study with the appropriate statistical power to detect that effect. This is especially important since low dose effects of EDCs can be expected to be very subtle. "Taking this analysis to its logical conclusion, hypothesis testing leads to an irresolvable dilemma. Below the detection limit of a specific experimental system, the presence of effects can neither be proven nor ruled out. We suggest that this, rather than bias due to sources of research funding (vom Saal and Welshons 2006), is at the root of the ED "low-dose" impasse." "Because the mathematical features of most regression models mean that zero effects are approached asymptotically without the regression line ever crossing the dose axis, the estimation of any possible effect, even down to infinitesimally small values, is feasible in principle. However, because of the lack of power, the models themselves cannot give any indications as to when estimates become unreliable. Statisticians have attempted to overcome this problem by including an additional model parameter that allows the estimation of a mathematical dose threshold associated with a zero response (Cox 1987; Hunt and Bowman 2006; Schwartz et al. 1995). However, considerable confusion has arisen because these modeling outcomes are often interpreted as toxicologic thresholds, but not as what they really are, that is, descriptive model parameters with little predictive power, strongly dependent on the selected model and estimation method. Even with the same set of data, widely differing threshold estimates can be obtained (Slob 1999)."	This paper discusses the difficulty of investigating effects of low doses of EDCs and establishing a threshold for those effects. The issues raised are mainly study design and that a relevant or critical effect level is often not clearly defined. The importance of making appropriate statistical considerations when designing experiments and interpreting results from investigations of effects of EDCs at low doses is highlighted. The issues raised are e.g. the probability of making type I and II errors and statistical power, as well as the choice of statistical method. It becomes clear that the expertise of statisticians is crucial in the investigation of subtle effects at low doses, and potential thresholds, of EDCs. An integrated approach using multiple comparisons and regression modeling to estimate dose-response relationships and establishing thresholds for relevant/critical effects is suggested.

	A	B	C	D	E	F
11	Gross et al. 2010	The threshold of toxicological concern (TTC) concept proposes that an exposure threshold value can be derived for chemicals, below which no significant risk to human health or the environment is expected. This concept goes further than setting acceptable exposure levels for individual chemicals, because it attempts to set a de minimis value for chemicals, including those of unknown toxicity, by taking the chemical's structure or mode of action (MOA) into consideration. This study examines the use of the TTC concern concept for endocrine active substances (EAS) with an estrogenic MOA. A case study formed the basis for a workshop of regulatory, industry and academic scientists held to discuss the use of the TTC in aquatic environmental risk assessment. The feasibility and acceptability, general advantages and disadvantages, and the specific issues that need to be considered when applying the TTC concept for EAS in risk assessment were addressed. Issues surrounding the statistical approaches used to derive TTCs were also discussed. This study presents discussion points and consensus findings of the workshop.	*none made*	*none made*		The purpose of this study was to investigate if the TTC concept can be used for the assessment of EDCs in aquatic ecosystems. The presence of a threshold for EDCs was a basic assumption in this paper. NOEC data were collected for 69 estrogenic compounds from different authority data bases and risk assessments and were used in modeling: "Cumulative functions of NOEC data for the most sensitive endpoints for each substance were analyzed ... by fitting a lognormal distribution." No arguments why a threshold is feasible for these substances were given. However, the authors state that "There also needs to be further consideration of potential low dose effects and uncertainties connected with this type of effect". The endpoints used were "apical (whole organism), higher tier test endpoints, particularly growth, development and reproduction as these are demographically relevant and suitable for use in risk assessment." It can be argued that a threshold would be likely for these "apical" endpoints given the complexity of biological responses to EDC exposure in an intact organism (Connolly and Lutz, 2004).
12	Calabrese 2008	Evidence is presented which supports the conclusion that the hormetic dose–response model is the most common and fundamental in the biological and biomedical sciences, being highly generalizable across biological model, endpoint measured and chemical class and physical agent. The paper provides a broad spectrum of applications of the hormesis concept for clinical medicine including anxiety, seizure, memory, stroke, cancer chemotherapy, dermatological processes such as hair growth, osteoporosis, ocular diseases, including retinal detachment, statin effects on cardiovascular function and tumour development, benign prostate enlargement, male sexual behaviours/dysfunctions, and prion diseases.	*none made*	*none made*		This paper mainly discusses the implications of a hormetic dose-response for medical treatment and drug development. Thus, Calabrese discusses the size of therapeutic windows and beneficial effects below a threshold of toxicity for different pharmaceuticals but does not provide any arguments for or against a threshold for toxicity. Also, hormonally active substances are not discussed. However, it seems to be a basic assumption throughout the paper that thresholds can exist for hormetic (or non-monotonic) dose-response curves and it can also be deduced from the many dose-response curves presented. This paper may be useful for discussing the shape of hormetic dose-response curves, e.g. the average maximum response and width of the "hormetic zone" (the area of the curve showing a stimulatory response at low doses). It should be noted that "hormesis" traditionally implies that the response observed at low doses is a beneficial effect compared to the response observed at higher doses, e.g. an initial decrease in anxiety at low doses but increased anxiety at high doses. This is different from what is discussed for EDCs where the argument is that the effect observed at low doses, although opposite from effects observed at higher doses, is potentially adverse

	A	B	C	D	E	F
13	Slob 1999	It is argued that dose-thresholds in a strict quantitative sense cannot exist in dose-response relationships that are studied by toxicologists. The biological, more qualitative, notion of thresholds can only be translated into quantitative terms by defining thresholds in the effect-size of continuous end-points: What size of an effect can the organismcope with and when is the effect adverse? These are key questions that toxicologists should address, for the various (continuous) end-points that may be observed in toxicological studies. Avoiding this question by resorting to a "formal" statistical test is deceptive and makes the outcome of a risk assessment fortuitous. There is no defendable rationale for different procedures of data analysis based upon assumed underlying mechanisms. The approach of data analysis should be determined by the nature and quality of the data. If the data allow for dose-response modelling, this should be done whatever the underlying biological mechanism. Quantal dose-response relationships (relating to noncancer endpoints) are problematic, because their slopes are largely determined by the experimental noise. Therefore, the concept of "extra risk" is not well defined in this case. In the case of tumor incidence data, on the other hand, the definition and estimation of an extra risk level are defendable, if it can be assumed that the stochastic component of the carcinogenicity process itself is large compared to the interindividual variation and experimental noise.	<u>These arguments are based on the mathematical definition of threshold.</u> *On the receptor level* : "Under normal circumstances receptors are to a certain degree occupied by "natural" factors, and therefore the toxic compound increases the percentage occupancy to a certain degree… in a strict sense, a dose-threshold cannot exist at the receptor level." *On the organ level* : "even though a threshold mechanism is obvious in general biological terms, the associated dose-response relationship cannot have a breaking point between zero and nonzero response in a living organism continuously interacting with its environment." *On the animal experiment level* : assuming that there is a threshold on the individual organism level "it is unlikely that such a threshold is exactly the same for individual animals… In other words, on the level of an animal experiment, thresholds in the dose-response relationship of a continuous end-point disappear unless all individuals have exactly the same threshold."	<u>Based on the biological and experimental definitions of threshold.</u> "From a biological point of view, the effect is nil, even though it is in fact nonzero. In other words, the meaning of "no response" in a biological sense is not the same as zero response in a strict quantitative sense. Apparently, thresholds do exist, but rather in the minds of toxicologists: it is tacitly assumed that very small responses (of any continuous end-point) can be considered as "no-effect," and that only responses of some magnitude do in fact represent an effect in a biological sense."	"Thresholds (in the strict sense [i.e. according to the mathematical definition]) cannot be measured: it would require infinitely precise measurements, both of effects and of applied doses. Therefore, the existence of thresholds cannot be proven, nor disproven, on the basis of data. All we can do is contemplate on what is most likely, using basic reasoning and general knowledge." "the existence of dose-thresholds in a strict quantitative sense, and the associated approach of analyzing dose-response data, is hard to defend." "dose-response data should not be characterized by dose-threshold models, even when it can be assumed that a biological threshold mechanism is involved."	One of the author's main conclusions seems to be that the choice of an effect-size that can be considered adverse and the corresponding dose causing that change in response should be the starting point of deducing "safe" levels for humans. I.e. a threshold below which there is NO effect (strictly speaking) should not be assumed.
14	White et al. 2009	Low-dose extrapolation model selection for evaluating the health effects of environmental pollutants is a key component of the risk assessment process. At a workshop held in Baltimore, Maryland, on 23-24 April 2007, sponsored by U.S. Environmental Protection Agency and Johns Hopkins Risk Sciences and Public Policy Institute, a multidisciplinary group of experts reviewed the state of the science regarding low-dose extrapolation modeling and its application in environmental health risk assessments. Participants identified discussion topics based on a literature review, which included examples for which human responses to ambient exposures have been extensively characterized for cancer and/or noncancer outcomes. Topics included the need for formalized approaches and criteria to assess the evidence for mode of action (MOA), the use of human versus animal data, the use of MOA information in biologically based models, and the implications of interindividual variability, background disease processes, and background exposures in threshold versus nonthreshold model choice. Participants recommended approaches that differ from current practice for extrapolating high-dose animal data to low-dose human exposures, including categorical approaches for integrating information on MOA, statistical approaches such as model averaging, and inference-based models that explicitly consider uncertainty and interindividual variability.	"In studied populations, thresholds have not generally been observed for cancer or, more notably, noncancer outcomes." "The complex molecular and cellular events that underlie the actions of agents that lead to cancer and noncancer outcomes are likely to be both linear and nonlinear. At the human population level, however, biological and statistical attributes tend to smooth and linearize the dose–response relationship, obscuring thresholds that might exist for individuals. Most notable of these attributes are population variability, additivity to preexisting disease or disease processes, and background exposure–induced disease processes."	*none made*	"Participants generally concurred that modeling approaches using a linear, no-threshold assumption improved consideration of the population-level factors (noted above) for both cancer and noncancer end points." "Almost all workshop participants preferred a linear, no-threshold approach to low-dose extrapolation, combined with modeled estimates of the low range of the observed data (e.g., benchmark dose modeling), for both cancer and noncancer outcomes. We discuss this in more detail below. A small minority of participants expressed some reservation regarding selection of a linear nonthreshold dose–response function as the default model assumption for cancer and noncancer outcomes given information on human biologic processes such as reversibility and repair."	This workshop report mainly discusses the presence of thresholds at the population level. The workshop participants concluded that assuming a threshold for effects of both carcinogens and non-carcinogens does not seem prudent.
15	Welshons et al. 2003	Information concerning the fundamental mechanisms of action of both natural and environmental hormones, combined with information concerning endogenous hormone concentrations, reveals how endocrine-disrupting chemicals with estrogenic activity (EEDCs) can be active at concentrations far below those currently being tested in toxicological studies. Using only very high doses in toxicological studies of EEDCs thus can dramatically underestimate bioactivity. Specifically: a) The hormonal action mechanisms and the physiology of delivery of EEDCs predict with accuracy the low-dose ranges of biological activity, which have been missed by traditional toxicological testing. b) Toxicology assumes that it is valid to extrapolate linearly from high doses over a very wide dose range to predict responses at doses within the physiological range of receptor occupancy for an EEDC; however, because receptor-mediated responses saturate, this assumption is invalid. c) Furthermore, receptor-mediated responses can first increase and then decrease as dose increases, contradicting the assumption that dose-response relationships are monotonic. d) Exogenous estrogens modulate a system that is physiologically active and thus is already above threshold, contradicting the traditional toxicological assumption of thresholds for endocrine responses to EEDCs. These four fundamental issues are problematic for risk assessment methods used by regulatory agencies, because they challenge the traditional use of extrapolation from high-dose testing to predict responses at the much lower environmentally relevant doses. These doses are within the range of current exposures to numerous chemicals in wildlife and humans. These problems are exacerbated by the fact that the type of positive and negative controls appropriate to the study of endocrine responses are not part of traditional toxicological testing and are frequently omitted, or when present, have been misinterpreted.	"when xenoestrogen activity is added to a natural system that is already responding to endogenous estrogen such as estradiol, any threshold in estrogenic response must already be exceeded by the endogenous hormone."	*none made*	The authors also argue that epigenetic mechanisms during fetal development, which allow for significant changes in phenotype in response to very small fluctuations in endogenous hormone signals, is one explanation for effects at very low doses of EDCs. In other words, exposure to very low doses of EDCs during critical periods of fetal development may lead to permanent changes in gene activity and organ function. However, they do not clearly argue that this would be an explanation for the absence of thresholds.	This paper specifically argues the presence of low-dose effects and non-monotonic dose-response relationships, as well as a lack of threshold, for EDCs with estrogenic activity.

	A	B	C	D	E	F
16	Sheehan et al. 1999	Risk assessments for nongenotoxic chemicals assume a threshold below which no adverse outcomes are seen. However, when an endogenous chemical, such as 17ss-estradiol (E2), occurs at a concentration sufficient to cause an effect, the threshold is already exceeded. Under these circumstances, exogenous estradiol is not expected to provide a threshold dose. This principle is demonstrated for E2 in the red-eared slider, a turtle with temperature-dependent sex determination. In this species, gonadal sex is determined by egg incubation temperature; female development requires endogenous estrogen produced by elevated temperature. While normal production of females by endogenous estrogens is not an adverse effect, exogenous estrogens can sex reverse presumptive males, which can be an adverse effect. A large dose-response study was conducted using seven doses and a vehicle control (starting n = 300/group); a single E2 dose was applied to the eggshell of recently laid eggs. Animals were sexed after hatching. The incubation temperature chosen, 28.6 degrees C, generates a minority of females. Thus, the criteria for testing the threshold hypothesis were met, i.e., there is evidence that there is endogenous estrogen and that it generates an irreversible response. The lowest E2 dose tested, 400 pg/egg (40 ng/kg), sex reversed 14.4% of the animals, demonstrating very low dose sensitivity. The data were fit with a modified Michaelis-Menten equation, which provided an estimate of 1.7 ng/egg for endogenous estradiol. The median effective dose (ED50) was 5.0 +/- 2.0 ng/egg (95% confidence limits), of which 1.7 ng/egg was endogenous estradiol and 3.3 ng/egg came from the applied estradiol. There was no apparent threshold dose for E2. A smaller replication confirmed these results. These results provide a simple biologically based dose-response model and suggest that chemicals which act mechanistically like E2 may also show no threshold dose. If so, even low environmental concentrations of such chemicals may carry risk for sex reversal.	"The BBDR [biologically based dose-response] model demonstrates that no exogenous E2 is without risk (as opposed to the threshold assumption)... Just as a threshold is not predicted to lie between two exogenous doses on the ascending portion of the dose-response curve, it is illogical to assume that a threshold exists between the control endogenous dose and the lowest exogenous dose in our model. The fit of the Michaelis-Menten model was slightly better with a Hill coefficient of 1 compared to 2, but because most of the curvature in the latter case is in the low-dose region of the dose-response curve (i.e. in the endogenous dose region), good resolution was not feasible. However, in the replication, the control value was lower and the curve showed no low-dose curvature. These considerations, taken with the fit of the large data set to the Michaelis-Menten equation, the three analyses in Figure 1, and the replication study, argue for our conclusions that no threshold dose exists."	*none made*		This paper is often referred to when arguing that there can be no threshold for (estrogenic) EDCs. The authors mean to challenge the assumption of threshold in risk assessment and state that this assumption cannot be made in general for EDCs

www.kemikalieinspektionen.se

Swedish Chemicals Agency, P.O. Box 2, SE-172 13 Sundbyberg
Phone: +46 8 519 41 100, Fax: +46 8 735 76 98, E-mail: kemi@kemi.se