Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Pharmacology

Bridging the gap between academic research and regulatory health risk assessment of Endocrine Disrupting Chemicals

Anna Beronius[1], Annika Hanberg[2], Johanna Zilliacus[2] and Christina Rudén[1]

Regulatory risk assessment is traditionally based primarily on toxicity studies conducted according to standardized and internationally validated test guidelines. However, health risk assessment of endocrine disrupting chemicals (EDCs) is argued to rely on the efficient integration of findings from academic research. The aim of this review was to provide an overview of current developments to facilitate the use of academic research in regulatory risk assessment of chemicals and how certain aspects of study design and reporting are particularly important for the risk assessment process. By bridging the gap between academic research and regulatory health risk assessment of EDCs, scientific uncertainty in risk assessment conclusions can be reduced, allowing for better targeted policy decisions for chemical risk reduction.

Addresses
[1] Department of Applied Environmental Science, Stockholm University, 106 91 Stockholm, Sweden
[2] Institute of Environmental Medicine, Karolinska Institutet, PO Box 210, 171 77 Stockholm, Sweden

Corresponding author: Beronius, Anna (anna.beronius@itm.su.se)

Current Opinion in Pharmacology 2014, **19**:99–104

This review comes from a themed issue on **Endocrine and metabolic diseases**

Edited by **Scott M Belcher**

http://dx.doi.org/10.1016/j.coph.2014.08.005

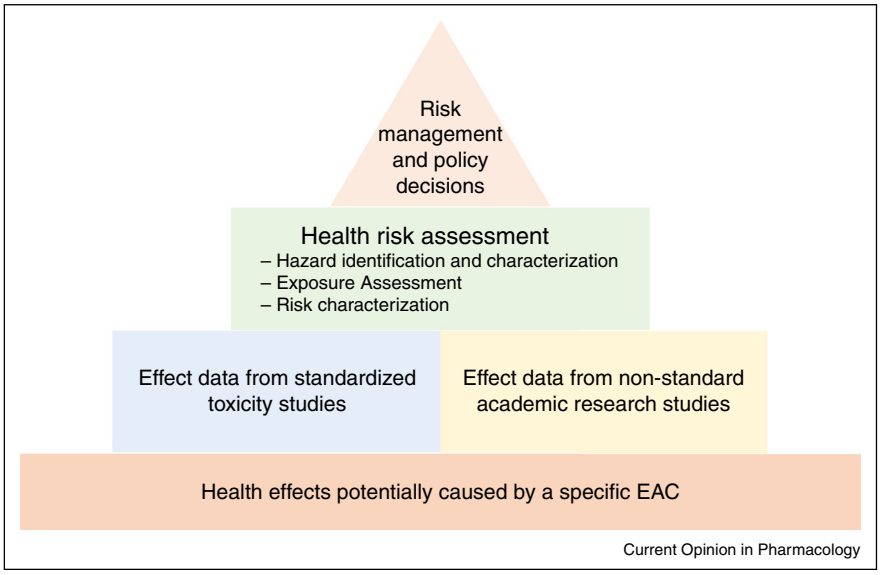
Introduction

Regulatory risk assessment of chemicals is conducted for the purpose of protecting human health and the environment against the negative effects of hazardous chemicals. It is carried out as basis for decisions on approving, restricting or phasing out the use of chemicals (Figure 1). As such, it is important that the scientific data on which the risk assessment, and consequently regulatory decision making, is to be based are reliable and relevant for this purpose. The evaluation of toxicity data is thus an integral and critical part of the regulatory process.

Guidance documents for risk assessment of chemicals issued by different authorities and organizations generally require or recommend that *all* relevant toxicity data should be considered in the risk assessment process [1–5]. However, toxicity studies conducted in accordance with standardized and internationally validated test guidelines, such as the Organisation for Economic Co-operation and Development (OECD) test guidelines, and Good Laboratory Practices (GLP) have often been preferred in regulatory health risk assessment. Test guidelines provide standardized requirements as well as recommendations for the design, performance and, to some extent, the reporting of toxicity studies and have been implemented to ensure the reliability of results. The major advantages of using standardized test methods are that the results are directly comparable across substances and that the data they generate will be accepted across jurisdictions. The major disadvantage is that standard methods do not always represent the most relevant testing approach and cannot cover all relevant adverse endpoints given the substance under investigation. Available standardized methods are, for example, criticized for being inadequate when it comes to identifying and evaluating endocrine disrupting chemicals (EDCs), *e.g.* for not including the most sensitive endpoints or covering sensitive windows of exposure [6•,7,8,9•,10,11]. The aim of GLP is to ensure the quality of the laboratory practices by specifying standard operational laboratory procedures and extensive requirements for data reporting. Notably, neither standardized test guidelines nor GLP will automatically ensure the relevance of a study for the health risk assessment purpose in question.

The importance of reducing the potential risks to human health, as well as to the environment, posed by EDCs have been highlighted in several recent reports and is high on the political agenda [8,12,13]. A number of standardized *in vitro* and *in vivo* test guidelines have been enhanced or developed by the OECD [14] as well as by the US Environmental Protection Agency (EPA) [15] for the detection and characterization of some types of EDCs, primarily substances interacting with estrogen, androgen, and thyroid hormone signaling pathways. But for many other endocrine pathways standardized methods are still lacking and a common and comprehensive strategy for regulatory identification and risk assessment of EDCs remains to be implemented.

Figure 1



Health risk assessment is carried out as a basis for policy decisions to protect human health against the harmful effects of chemicals. Non-standard research studies of EDCs often use novel methods and investigate sensitive endpoints, which are currently not covered in standardized test protocols. Thus, toxicological information from non-standard studies complement information from standard studies to generate a more complete picture of the effects and risks of EDCs and enable better targeted risk management and policy decisions.

Over the last decades EDCs has been an area where research has advanced rapidly. Hence, academic research studies investigating the toxicity of EDCs often include novel methods and endpoints that may complement the information provided by standard studies. Although often not conducted according to any internationally standardized test guidelines, such studies have in many cases been discussed as being more sensitive and relevant than standard studies for identifying and evaluating endocrine-related effects of chemicals [6•,7,9•,11,16]. Also, as implied above, it is acknowledged that standard studies in general cannot cover all relevant adverse endpoints, especially not for EDCs. Thus, it could be argued that the health risk assessment of EDCs relies on the efficient integration of findings from academic research studies [17] (Figure 1). However, academic research is seldom conducted for the specific purpose of chemical risk assessment and, in a regulatory setting, such studies are often criticized for having methodological limitations, such as only investigating a single dose, failing to control for litter effects and using inappropriate statistical methods, and/or suffering from insufficient reporting, which negatively impact their reliability and hamper their usability for this purpose [18,19,20,21•,22,23].

Although the highest quality should always be the ambition in scientific research, regulatory risk assessment may set specific requirements on study design and reporting. This review aims to provide an overview of current developments to facilitate the use of academic research in regulatory risk assessment of chemicals and how certain aspects of study design and reporting are particularly important for the risk assessment process. The intention is not to promote a general template for how to design and conduct research but merely to increase awareness of how research studies can contribute valuable information to regulatory decision-making. While *in vivo* studies are currently required in order to establish actual endocrine disruption, *i.e.* adverse endocrine-related effects occurring in an intact organism, *in vitro* studies contribute important information about the mechanisms for endocrine activity, as well as for the investigation of adverse outcome pathways (AOP) and mode of action-based assessments. This review therefore aims to include aspects relevant for both *in vivo* and *in vitro* studies.

Evaluation of toxicity data for regulatory risk assessment and decision making

There is in general a need to improve the scientific basis for regulatory decision making and risk assessment [2,5,23,24,25•,26,27], *i.e.* to improve methods for integrating available evidence to support conclusions about health risks from chemicals. The evaluation of the ''quality'', ''validity'' or ''adequacy'' of toxicity studies is a central step of this integration process. In the field of clinical research methodologies for study evaluation and synthesizing science into evidence-based decisions have been discussed and developed for decades. However, in the area of environmental health these issues are still emerging. It has been noted in several reports that study

Table 1

Recently developed methods for scientific evidence evaluation and integration for decision-making.

Method	Stated purpose	Source/affiliation
OHAT Approach [30]	Collecting and synthesizing scientific evidence to answer environmental health questions.	National Institute of Environmental Health Sciences (NIEHS) Office of Health Assessment and Translation (OHAT), US
Navigation Guide [31]	Evaluation of the quality of evidence on hazardous environmental exposures to support evidence-based decision making by clinicians and patients, as well as professional societies, health care organizations and government agencies.	University of Southern California San Francisco, US
SCENIHR approach [29]	Transparent and consistent health risk assessment within SCENIHR.	European Commission Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR)
EFSA systematic review [2]	Identifying, selecting and critically appraising research relevant for risk assessment within the food and feed safety areas.	European Food Safety Authority (EFSA)
Science in Risk Assessment and Policy (SciRAP) tool for evaluation of individual studies [32]	Assessment of the reliability and relevance of toxicity studies used in risk assessment of chemicals.	Stockholm University, Sweden Karolinska Institutet, Sweden
Qualichem *in vivo* [33]	Assessment of the quality of *in vivo* studies used in risk assessment of chemicals.	Centre National de la Recherche Scientifique, France Utrecht University, Netherlands
SYRCLE's risk of bias tool for animal studies [34]	Critical appraisal of evidence from animal studies and translation of animal research into clinical practice.	Radboud University, Netherlands University of Amsterdam, Netherlands
ToxRTool [35]	Reliability evaluations of *in vitro* and *in vivo* toxicological data.	Forschungs- und Beratungsinstitut Gefahrstoffe GmbH (FoBiG), Germany German Cancer Research Center ECVAM, E.C. Joint Research Centre, Italy Johns Hopkins University, US TÜV Rheinland BioTech GmbH, Germany

evaluation procedures applied by different authorities for the purpose of health risk assessment or decision making are seldom very transparent or systematic [3,23,28,29].

Table 1 summarizes a number of methods for evaluating *in vitro* and/or *in vivo* toxicity studies, as well as for data integration for decision-making that have been developed in the past few years. These different methods have been developed to make data evaluation, synthesis and integration more systematic and transparent. Thus, they do not in most cases (excepting SciRAP, see Box 1) specifically identify and present principles for good study design and reporting aimed at researchers. However, some guidance for researchers may be derived from these methods.

Box 1 Published guidelines for reporting in scientific articles.

- The Nature Journal reporting checklist for life sciences articles (including human, animal and cell studies) [21•]. Available online at http://www.nature.com/authors/policies/checklist.pdf
- The SciRAP checklist for reporting *in vivo* studies [32]. Available on-line at www.scirap.org
- The Animal Research: Reporting of *In Vivo* Experiments (ARRIVE) guidelines [39]. Available online at http://www.nc3rs.org.uk/page.asp?id=1357
- Landis et al., 2012 [42].
- The US National Research Council Institute for Laboratory Animal Research Guidance for the description of animal research in scientific publications [44].

Relevance, reliability and risk of bias

Different terms may be used to define the adequacy of a toxicity study for regulatory risk assessment and decision-making, and these vary between jurisdictions. In terms of describing the extent to which toxicity studies are adequate for use in risk assessment within European chemicals regulations the terms ''relevance'' and ''reliability'' are often used [36].

Relevance is defined as ''covering the extent to which data and tests are appropriate for a particular hazard identification or risk characterization'' [36]. I.e. does the study investigate effects and include exposure scenarios that are relevant for evaluating the risks to human health? The relevance of a study is thus context-dependent and may change with time as the scientific knowledge evolves, *e.g.* regarding relevant exposure sources and routes. The aspect of relevance also refers to the relevance of the methods used for investigating the selected endpoints. As noted above, relevance of a study cannot be ensured by adhering to GLP or standardized test guidelines.

Reliability, is commonly defined as ''the inherent quality of a test report or publication relating to preferably standardized methodology and the way the experimental procedure and results are described to give evidence of the clarity and plausibility of the findings. Reliability of data is closely linked to the reliability of the test method used to generate the data'' [36]. Importantly, non-standardized research studies can be just as reliable as studies performed under strict implementation of GLP and/or standardized test guidelines. A non-standard study should fulfill general quality criteria for scientific investigations, *e.g.* the control of relevant variables, comparison to appropriate control groups, and proper reporting of the results, etc.

Some of the methods presented above focus on the evaluation of ''risk of bias'' as part of the evaluation of a study's adequacy for risk assessment. In statistical terms bias is similar to ''systematic error''. A systematic error will remain the same regardless of the number of repetitions of the experiment or sample size. This is in contrast to random errors that are reduced with repeated measurements. Risk of bias thus relates to the study's internal validity, reflecting characteristics in the study design that might introduce a systematic error and affect the magnitude and/or direction of study results [3,23,30,31]. Some sources of bias are, for example, the lack of randomization, blinding, specification of inclusion and exclusion criteria or statistical power. For a more in depth discussion on these factors the reader is referred to the recent review of the US EPA Integrated Risk Information System (IRIS) conducted by the US National Research Council [23].

Critical study design issues

Adequate planning and good study design is essential to the reliability and relevance of any research study and principles of sound study design have been proposed and published previously [[*e.g.* 17,20]]. However, regulatory risk assessment of chemicals is based on specific principles, such as characterizing a dose–response relationship and identifying a no observed adverse effect level (NOAEL) or lower confidence limit of a Benchmark Dose (BMDL), which requires that toxicity studies include certain design aspects or information to be useful for this purpose. Further, in the testing of EDCs the specific characteristics of such compounds need also be accounted for in the study design [[*e.g.* 6,8,9,11,37]]. These include, for example 1) that effects may occur at very low doses, 2) the potential occurrence of non-monotonic dose–response relationships, and 3) that effects of EDCs vary depending on life stage. It is not within the scope of this review to provide an exhaustive list of factors important to the design and execution of scientific studies or to describe an ''optimal'' design of an *in vivo* or *in vitro* study to investigate EDCs, since this will vary from case to case. For more detail the reader is referred to the methods and reports reviewed here. However, some important design factors to consider include:

- Including the appropriate sample size and basing the number of animals or replicates per dose group on statistical power calculations.
- Inclusion of enough dose groups to demonstrate any dose-related effect. Traditionally, at least three doses and a negative control group have been required. However, in order to evaluate low dose effects and non-monotonic dose–response relationships more dose groups may be necessary. Thus, it should be noted that the choice and spacing of doses may be especially important when investigating EDCs.
- Using a sensitive and relevant test model, *i.e.* the relevant animal species and strain, tissue or cells, appropriate for the endpoints investigated.
- Inclusion of endpoints that are relevant and sensitive in terms of endocrine activity as well as relevant to human health.
- Applying an appropriate exposure regimen *in vivo*, *i.e.* covering the most sensitive and relevant exposure window and using the relevant exposure route.
- In the case of *in vivo* studies, carefully considering the appropriate housing conditions that could influence study outcomes such as circulating hormone levels or behavioral endpoints, *e.g.* number of animals per cage, temperature, light–dark cycle and choosing the appropriate cage, water bottle, bedding and enrichment materials.
- Reducing additional confounders, *e.g.* analyzing feed and water for contaminants that could impact study objectives, such as phytoestrogens, pesticide residues, persistent organic pollutants, heavy metals and mycotoxins.
- Randomization of the assignment of animals to different treatment groups.
- Blinding of investigators as to the allocation of animals to different treatment groups. However, it is acknowledged that blinding may not always be feasible or desired in all research studies, depending on the type of study and study conditions [38].

Importantly, it should be clearly reported that the above factors have been considered in the planning and execution of the study.

It is worth noting that several of the design factors listed here are not yet covered in standardized test guidelines, adding further to the importance of integrating well conducted and reported non-standard studies in the risk assessment process.

Sufficient reporting for regulatory risk assessment

From a regulatory risk assessment perspective clear and detailed reporting of the research aim, design, performance and results of a study is critical. If these aspects are not sufficiently reported it may not be possible to evaluate, and thereby ensure, sufficient reliability and

relevance of the study for health risk assessment. Weaknesses in reporting of research and its impact on study reproducibility and reliability, and especially consequences for clinical research and policy-making, have been extensively discussed in the literature for decades [39–42] and have been recently highlighted by prominent scientific journals, such as Nature [21•] and Environmental Health Perspectives [43].

The information reported in research studies may reflect different standards related to the intended use of the data and what the researcher considers important in relation to the hypothesis tested. While adequately reported for its research purpose the study may still be insufficiently reported in the eyes of a risk assessor. Unawareness regarding what information is required to meet the demands that regulatory agencies put on data intended for risk assessment may be one reason for such discrepancies. In addition, space provided for individual articles in scientific journals is often limited, forcing investigators to reduce the amount of information reported.

Risk assessors may sometimes contact individual study authors to ask for clarifications in cases where a study seems specifically critical to the risk assessment. However, if research studies would comply with the reporting requirements of regulatory risk assessment to a larger extent, it would increase the number of studies useful for health risk assessment and decision making as well as facilitate study evaluation. To this end, different guidelines and checklists for reporting toxicity studies, such as the ARRIVE guidelines [39] have been recently proposed [21•,32,39,42,44]. Some of these recent proposals have been listed in Box 1. These include many of the same items to be reported but vary somewhat in detail and format and the reader is referred to the original publications or reports for further detail.

Conclusions

The health risk assessment of EDCs relies on the efficient integration of academic research to fill information gaps. However, the use of non-standard academic research studies in regulatory risk assessment has often been hampered because of limitations in study design or reporting.

In this review we have provided an overview of study design and reporting aspects that are critical to the evaluation of data for regulatory purposes. These are based on recently developed systems to facilitate transparent and systematic evidence integration in regulatory decision making and on proposed reporting guidelines available from different sources. These systems can help bridging the gap between scientists and regulators and improve the scientific basis for regulatory decision making. In our view, the issues covered here should also be considered in the wider context as important for reproducibility and reliability of research studies in general and the integrity of life sciences as a whole.

Implementing systems along these lines has the potential to facilitate the use of academic research in regulatory health risk assessment of EDCs to reduce scientific uncertainty in risk assessment conclusions, and in extension contribute to better targeted policy decisions for chemical risk reduction (Figure 1).

Conflict of interest statement

The authors have no competing interests to declare.

Acknowledgements

This work was funded by faculty grants from Stockholm University and by grants from the Swedish Research Council Formas (grant no: 2007-2056).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. European Chemicals Agency (ECHA): *Evaluation under REACH, Progress Report 2013*. 2012:. Available online at: echa.europa.eu.

2. European Food Safety Authority (EFSA): **Guidance of EFSA Application of systematic review methodology to food and feed safety assessments to support decision making**. *EFSA J* 2010, **8**:90.

3. US National Research Council of the National: *Academies review of EPA's integrated risk information system (IRIS) process*. Washington D.C.: The National Academies Press; 2014, .

4. Organisation for Economic Co-operation Development (OECD): *Manual for the assessment of chemicals. Chapter 3: Data evaluation*. 2005:. Available online at www.oecd.org/.

5. United Stated Environmental Protection Agency: *Framework for Human Health Risk Assessment to Inform Decision Making*. 2014:. Available online at: www.epa.gov/.

6. • Kortenkamp A, Martin O, Faust M, Evans R, McKinlay R, Orton F, Rosivatz E: *State of the Art Assessment of Endocrine Disrupters - Final Report*. . Available online at: http://ec.europa.eu/environment/chemicals/endocrine/pdf/sota_edc_final_report.pdf 2012.
This report summarizes advances in the state of the science for EDCs since 2002 and highlights issues critical for identification and evaluation of these compounds

7. Myers JP, vom Saal FS, Akingbemi BT, Arizono K, Belcher S, Colborn T, Chahoud I, Crain DA, Farabollini F, Guillette LJ Jr *et al.*: **Why public health agencies cannot depend on good laboratory practices as a criterion for selecting data: the case of bisphenol A**. *Environ Health Perspect* 2009, **117**:309-315.

8. United Nations Environment Programme the World Health Organization: *State of the science of endocrine disrupting chemicals - 2013*. 2012:. Available online at: http://www.who.int/ceh/publications/endocrine/en/.

9. • Zoeller RT, Brown TR, Doan LL, Gore AC, Skakkebaek NE, Soto AM, Woodruff TJ, Vom Saal FS: **Endocrine-disrupting chemicals and public health protection: a statement of principles from The Endocrine Society**. *Endocrinology* 2012, **153**:4097-4110.
A review of the principles of endocrinology that need to be considered in the health risk assessment of EDCs.

10. European Food Safety Authority: *Scientific opinion on the hazard assessment of endocrine disruptors: Scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these*

substances on human health and the environment 2013.. 2013:. Available online at: http://www.efsa.europa.eu/en/efsajournal/doc/pdf/3132.

11. Schug TT, Abagyan R, Blumberg B, Collins TJ, Crews D, DeFleur PL, Dickerson SM, Edwards TM, Gore AC, Guillette LJ *et al.*: **Designing endocrine disruption out of the next generation of chemicals**. *Green Chem* 2013, **15**:181-198.

12. European Environment Agency: *The impacts of endocrine disrupters on wildlife, people and their environments. The Weybridge +15 (1996-2011) report*. 2012:. Available online at: http://www.eea.europa.eu/publications/the-impacts-of-endocrine-disrupters.

13. Parliament: EU: *MEPs call for clampdown on chemicals affecting hormones Press release 14 March*. 2013:. Available on-line at www.europarl.europa.eu/.

14. Organisation for Economic Co-operation and Development (OECD). **Detailed Review Paper on the State of the Science on novel in vitro and in vivo screening and testing methods and endpoints for evaluating endocrine disruptors. Series on Testing & Assessment No. 178**. 2012:. Available online at: http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmenttestingforendocrinedisrupters.htm.

15. Juberg DR, Borghoff SJ, Becker RA, Casey W, Hartung T, Holsapple MP, Marty MS, Mihaich EM, Van Der Kraak G, Wade MG *et al.*: **t4 workshop report–lessons learned, challenges, and opportunities: the U.S. Endocrine Disruptor Screening Program**. *ALTEX* 2014, **31**:63-78.

16. Beronius A, Johansson N, Ruden C, Hanberg A: **The influence of study design and sex-differences on results from developmental neurotoxicity studies of bisphenol A, implications for toxicity testing**. *Toxicology* 2013, **311**:13-26.

17. Schug TT, Heindel JJ, Camacho L, Delclos KB, Howard P, Johnson AF, Aungst J, Keefe D, Newbold R, Walker NJ *et al.*: **A new approach to synergize academic and guideline-compliant research: the CLARITY-BPA research program**. *Reprod Toxicol* 2013, **40**:35-40.

18. Alcock RE, Macgillivray BH, Busby JS: **Understanding the mismatch between the demands of risk assessment and practice of scientists—the case of Deca-BDE**. *Environ Int* 2011, **37**:216-225.

19. Hengstler JG, Foth H, Gebel T, Kramer PJ, Lilienblum W, Schweinfurth H, Volkel W, Wollin KM, Gundert-Remy U: **Critical evaluation of key evidence on the human health hazards of exposure to bisphenol A**. *Crit Rev Toxicol* 2011, **41**:263-291.

20. Harris CA, Scott AP, Johnson AC, Panter GH, Sheahan D, Roberts M, Sumpter JP: **Principles of sound ecotoxicology**. *Environ Sci Technol* 2014, **48**:3100-3111.

21. • Nature Editorial: **Announcement: reducing our irreproducibility**. *Nature* 2013, **496**:398.
Emphasizes and motivates the importance of sufficient reporting of scientific articles and provides clear and concise guidance for authors.

22. Batke M, Aldenberg T, Escher S, Mangelsdorf I: **Relevance of non-guideline studies for risk assessment: the coverage model based on most frequent targets in repeated dose toxicity studies**. *Toxicol Lett* 2013, **218**:293-298.

23. National Research Council: *Review of EPA's Integrated Risk Information System (IRIS) Process*. Washington DC: The National Academies Press; 2014, .

24. Bishop PL, Willett CE: **The use and acceptance of Other Scientifically Relevant Information (OSRI) in the U.S Environmental Protection Agency (EPA) Endocrine Disruptor Screening Program**. *Birth Defects Res B Dev Reprod Toxicol* 2014, **101**:3-22.

25. • Silbergeld E, Scherer RW: **Evidence-based toxicology: strait is the gate, but the road is worth taking**. *ALTEX* 2013, **30**:67-73.
Describes the importance of improving the use of academic research for regulatory decision-making and the challenges this process faces.

26. Birnbaum LS, Thayer KA, Bucher JR, Wolfe MS: **Implementing systematic review at the National Toxicology Program: status and next steps**. *Environ Health Perspect* 2013, **121**:A108-A109.

27. McCarty LS, Borgert CJ, Mihaich EM: **Information quality in regulatory decision making: peer review versus good laboratory practice**. *Environ Health Perspect* 2012, **120**:927-934.

28. Beronius A, Rudén C, Håkansson H, Hanberg A: **Risk to all or none?. A comparative analysis of controversies in the health risk assessment of Bisphenol A**. *Reprod Toxicol* 2010, **29**:132-146.

29. Scientific Committee on Emerging: *Newly Identified Health Risks (SCENIHR): Memorandum on the use of the scientific literature for human health risk assessment purposes - weighing of evidence and expression of uncertainty*. 2012:. Available online at: http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_s_001.pdf.

30. Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA: **Systematic review and evidence integration for literature-based environmental health science assessment**. *Environ Health Perspect* 2014. Epub ahead of print.

31. Woodruff TJ, Sutton P: **An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences**. *Health Aff (Millwood)* 2011, **30**:931-937.

32. Beronius A, Molander L, Ruden C, Hanberg A: **Facilitating the use of non-standard in vivo studies in health risk assessment of chemicals: a proposal to improve evaluation criteria and reporting**. *J Appl Toxicol* 2014, **34**:607-617.

33. Maxim L, van der Sluijs JP: **Qualichem in vivo: a tool for assessing the quality of in vivo studies and its application for bisphenol a**. *PLoS One* 2014, **9**:e88773.

34. Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW: **SYRCLE's risk of bias tool for animal studies**. *BMC Med Res Methodol* 2014, **14**:43.

35. Schneider K, Schwarz M, Burkholder I, Kopp-Schneider A, Edler L, Kinsner-Ovaskainen A, Hartung T, Hoffmann S: **ToxRTool, a new tool to assess the reliability of toxicological data**. *Toxicol Lett* 2009, **189**:138-144.

36. European Chemicals Agency (ECHA): *Guidance on information requirements and chemical safety assessment. Chapter R.4: Evaluation of available information*. 2011:. Available on-line at: echa.europa.eu.

37. Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR Jr, Lee DH, Shioda T, Soto AM, Vom Saal FS, Welshons WV *et al.*: **Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses**. *Endocr Rev* 2012, **33**:378-455.

38. Morris R: **Reporting standards: rigid guidelines may restrict research**. *Nature* 2012, **491**:192.

39. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG: **Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research**. *PLoS Biol* 2010, **8**:e1000412.

40. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG: **Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network**. *BMC Med* 2010, **8**:24.

41. Ågerstrand M, Edvardsson L, Rudén C: **Bad reporting or Bad Science? Systematic Data Evaluation as a Means to Improve the Use of Peer-Reviewed Studies in Risk Assessments of Chemicals**. *Hum Ecol Risk Assessment: Int J* 2014, **20**:1427-1445.

42. Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, Crystal RG, Darnell RB, Ferrante RJ, Fillit H *et al.*: **A call for transparent reporting to optimize the predictive value of preclinical research**. *Nature* 2012, **490**:187-191.

43. Tilson HA, Schroeder JC: **Reporting of results from animal studies**. *Environ Health Perspect* 2013, **121**:A320-A321.

44. US National Research Council of the National Academies: *Guidance for the Description of Animal Research in Scientific Publications*. Washington, DC: National Research Council (US) Institute for Laboratory Animal Research; 2011, .