Environment International xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Review article

A critical review of frameworks used for evaluating reliability and relevance of (eco)toxicity data: Perspectives for an integrated eco-human decision-making framework

N. Roth [a,*], P. Ciffroy [b]

[a] Swiss Centre for Applied Human Toxicology (SCAHT) Directorate, Regulatory Toxicology Unit, Missionsstrasse 64, 4055 Basel, Switzerland
[b] Electricité de France (EDF) R&D, National Hydraulic and Environment Laboratory, 6 quai Watier, 78400 Chatou, France

ARTICLE INFO

Article history:
Received 3 May 2016
Received in revised form 16 July 2016
Accepted 20 July 2016
Available online xxxx

Keywords:
Reliability
Relevance
Data evaluation
Quality assessment
Integrated hazard assessment
Integrated risk assessment

ABSTRACT

Considerable efforts have been invested so far to evaluate and rank the quality and relevance of (eco)toxicity data for their use in regulatory risk assessment to assess chemical hazards. Many frameworks have been developed to improve robustness and transparency in the evaluation of reliability and relevance of individual tests, but these frameworks typically focus on either environmental risk assessment (ERA) or human health risk assessment (HHRA), and there is little cross talk between them. There is a need to develop a common approach that would support a more consistent, transparent and robust evaluation and weighting of the evidence across ERA and HHRA. This paper explores the applicability of existing Data Quality Assessment (DQA) frameworks for integrating environmental toxicity hazard data into human health assessments and vice versa. We performed a comparative analysis of the strengths and weaknesses of eleven frameworks for evaluating reliability and/or relevance of toxicity and ecotoxicity hazard data. We found that a frequent shortcoming is the lack of a clear separation between reliability and relevance criteria. A further gaps and needs analysis revealed that none of the reviewed frameworks satisfy the needs of a common eco-human DQA system. Based on our analysis, some key characteristics, perspectives and recommendations are identified and discussed for building a common DQA system as part of a future integrated eco-human decision-making framework. This work lays the basis for developing a common DQA system to support the further development and promotion of Integrated Risk Assessment.

Contents

* Corresponding author.
E-mail address: nicolas.roth@scaht.org (N. Roth).

http://dx.doi.org/10.1016/j.envint.2016.07.011
0160-4120/

1. Introduction

When deciding whether a chemical is safe for the scenarios of its use, human health and environmental risk assessors have to rank and weight data according to (often implicit) quality criteria such as reliability and relevance, before integrating all available data and finally providing results which are understandable for risk managers and decision-makers. The quality of any risk assessment (RA) is driven by the quality of the input data. Laboratory in vitro and in vivo (eco)toxicity[1] data generated in studies conforming to internationally accepted guidelines and standards (e.g. OECD Test Guidelines, regional or national standards such as CEN or DIN) are often seen as 'gold standards' for hazard assessment because they are viewed as being of high quality, but they may not necessarily be relevant for specific RA scenarios. Concerns over the quality of data derived from in vitro or in vivo studies often prevent their use if they have been generated by non-standard approaches, even if they can be a relevant source of information for the context under investigation.

Information derived from both 'gold standard' and non-standard (eco)toxicity studies is heterogeneous, depending on e.g. the target organism (e.g. human, rat, fish, algae), the endpoint (e.g. mortality, reproductive toxicity), the study design (e.g. acute or chronic exposure, life stage) or the methodology used (e.g. 'gold standard' test vs non-standard test); data can be quantitative, semi-quantitative or qualitative; and quality may vary markedly among non-standardized methods. This heterogeneity contributes, among other factors (e.g. lack of study relevance or concordance between studies, inference gaps), to uncertainty in the RA process, which can be reduced if data quality is evaluated by scientifically sound, consistent and transparent processes. Therefore, the development of clear guidelines and unambiguous criteria for the evaluation of data quality is necessary for both standard and non-standard test methods.

Many frameworks have already been developed to assess quality (incl. reliability and relevance) of individual test data (as well as of non-testing data), sometimes as part of broader guidance documents describing the overall Weight-of-Evidence (WoE) analysis process, by regulatory bodies (e.g. EFSA, 2009; OECD, 2005; USEPA, 2003, 2011), scientific organizations (e.g. EURL-ECVAM/Schneider et al., 2009; SCENIHR, 2012) or industry consortia (e.g. ECETOC, 2009). However, these frameworks have usually been developed for either Environmental Risk Assessment (ERA) or Human Health Risk Assessment (HHRA), and there is little cross talk between them. This is because, for historical and practical reasons, ERA and HHRA have generally developed independently using different terminologies and largely separate data, models and assumptions (Bridges, 2003; Suter et al., 2005). This situation prevents the mutual use and sharing of information and best practices across ERA and HHRA, which may reduce the efficiency and confidence in the way human health and environmental risks are evaluated and managed (Suter, 1997; Vermeire et al., 2007; Wilks et al., 2015).

The concept of integrated risk assessment (IRA) has been proposed as a potential solution because it brings together independent sources of (eco)toxicity data to enable a more harmonized, comprehensive, informative and efficient risk analysis process (WHO/IPCS, 2001; Wilks et al., 2015). Two key components for successful IRA implementation that yet remained to be developed are integrated assessment frameworks for 'Data Quality Assessment' (DQA) and for WoE analysis. For effective integration, one has to consider appropriate methods for managing heterogeneity in data quality and for assessing the confidence level (or uncertainty level) of individual data as well as of their combined content (Péry et al., 2013). A common and consistent approach to data quality evaluation, in particular reliability, and to relevance, that has equal application to both ERA and HHRA, and that facilitates the comparison and interpretation of the evidence across the two scientific disciplines, could play a pivotal role in building more robust and relevant WoE approaches to support the development of an integrated eco-human decision-making framework for IRA. These needs have been addressed by the EU FP7 HEROIC[2] Coordination action, which aimed to consolidate knowledge gained from previous integration-related initiatives at EU level and promote the further development and implementation of IRA (Péry et al., 2013; Wilks et al., 2015).

2. Aim and method

This paper is the first of two papers that explores the applicability of existing DQA frameworks and WoE frameworks and methods for integrating environmental toxicity data into human health assessments and vice versa, to support the development of an integrated eco-human decision-making framework for IRA. The present work reviews eleven frameworks built for evaluating reliability and relevance of (eco)toxicity data in environmental and human health hazard assessments. Our objective was to understand to what extent these frameworks could be adapted for developing a common data quality evaluation system that would equally apply to both human and environmental targets, as an input for further WoE approaches. Therefore, the primary focus of this paper is the DQA component of the overall decision-making process, not the WoE itself, which will be addressed in a second related paper (manuscript in preparation).

Selection criteria for including a framework in the review were: (i) explicit focus on reliability and/or relevance of (eco)toxicity studies; (ii) quality evaluation scheme based on a method or a formalized procedure. Consideration was given to include in our review not only DQA

[1] In this paper, (eco)toxicity refers to both human related toxicity and ecotoxicity data and studies.

[2] Health and Environmental Risks: Organisation, Integration and Cross-fertilisation of Scientific Knowledge, www.heroic-fp7.eu.

Please cite this article as: Roth, N., Ciffroy, P., A critical review of frameworks used for evaluating reliability and relevance of (eco)toxicity data: Perspectives for an integrate..., Environ Int (2016), http://dx.doi.org/10.1016/j.envint.2016.07.011