Environment International 92-93 (2016) 574–577

Contents lists available at ScienceDirect

Environment International

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

Systematic review in chemical risk assessment — A chemical industry perspective

Camilla K. Pease [a], Robinan P. Gentry [b]

[a] Ramboll Environ, 1 Broad Gate, The Headrow, Leeds, LS1 8EQ, UK
[b] Ramboll Environ, 1900 N. 18th Street, Suite 804, Monroe, LA, USA

article info

Article history:
Received 5 August 2015
Received in revised form 27 November 2015
Accepted 9 December 2015
Available online 22 December 2015

Keywords:
Systematic review
Chemical risk assessment

abstract

This commentary provides a perspective from the chemicals industry on the potential usefulness of systematic review approaches in chemical risk assessment.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Traditionally, the chemical industry, as well as regulatory agencies, has not used systematic review in chemical risk assessment. While the current approaches, using narrative methods, dose–response analysis, and risk evaluation frameworks, appear to be thorough and robust, hazard and risk assessments come under detailed scrutiny in the regulatory setting from decisions that seem to be subjective. Further difficulties can arise in communicating the risk assessment and resulting risk management decisions to the general public and politicians in a consistent and credible manner, when perhaps not all the data and choices have been documented as transparently as they could be. It is in the interest of the chemicals industry to assure people are safe and do not come to harm through exposure to hazardous chemicals, either in the environment or in consumer products. Here, we ask "can the process of a structured systematic review lead to increased transparency for clearer communications as to why certain decisions around chemical safety are reached?"

The process of systematic review was first documented in clinical research, under the US Cochrane Collaboration (Higgins and Green, 2011), to objectively and transparently determine whether a particular medical intervention is effective in treating a medical condition. This is a situation where human individuals and populations may react differently to a medical intervention and human evidence, from often well controlled studies with defined exposures, can give conflicting and unexpected results. Certain subjectivities and biases can creep into an evaluation that is not transparent. Interpreting the conflicts in the data can be open to misinterpretations and errors, and therefore the systematic review process arose with the aim to make transparent each step and choice in interpreting the data in the body of evidence, when drawing a conclusion on a particular clinical research question e.g. has the treatment been effective? Systematic review is a protocol-driven approach where each line of evidence is rigorously documented, the quality of data is evaluated against a pre-determined set of criteria, and if possible, meta-analysis and statistical evaluation of datasets is incorporated, with the goal of answering a specific pre-defined research question.

Recently, in the US, systematic review methods have been explored in the context of whether they are useful in the context of answering research questions on the potential for adverse health effects or characterizing hazard that may or may not be caused by environmental chemical exposures. This is usually in the context of whether anthropogenic (man-made industrially-derived) chemicals present in the workplace, environment or in consumer products can adversely affect the health of those exposed. If so, then policies and risk management strategies need to be implemented by regulators to either control, restrict or remove exposure to the offending chemical. The consequences of making erroneous decisions can lead to, most worryingly, an adverse human health effect (i.e. if a chemical hazard or risk is not appropriately identified and characterised); the consequences of making an excessively precautionary decision for a chemical (arguably for the benefit of public health protection) is on industrial innovation and economic growth. However, there is also a potentially indirect consequence that if a precautionary approach is taken on a substance that leads to a ban on the use of that substance, an evaluation also needs to be made on the chemical that is set to replace its function in society — the alternatives may lead to a greater health risk. It may be the case that precautionary restrictions and removal of chemicals that are used today and benefit society, are in reality used safely, but the evidence in the decision-making context has been biased in favour of taking a precautionary approach, for what could be a range of reasons that are not wholly based on

Corresponding author.
E-mail address: cpease@ramboll.com (C.K. Pease).

http://dx.doi.org/10.1016/j.envint.2015.12.007
0160-4120/© 2015 Elsevier Ltd. All rights reserved.

scientific evidence but on certain pre-conceived perceptions of risk. Systematic review could in principle make the decision-making more transparent in terms of the evidence base and the choices made, but there is a concern in industry that this will take significantly more resource than a narrative approach, and some might say, is prohibitively resource intensive. From our experience, this may not necessarily be the case if systematic review is used wisely and is focused on the 'high value' questions that are critical to the decision or where there is a lack of consensus.

Woodruff & Sutton (2014) from the University of California published the concept of the 'Navigation Guide' described as 'a systematic and rigorous approach to research synthesis that has been developed to reduce bias and maximize transparency in the evaluation of environmental health information'. The Office of Health Assessment and Translation (OHAT) at the US National Toxicology Program (Rooney et al. 2014) have also developed an approach to systematically review data and evidence for chemical hazard identification, which is an integral part of chemical risk assessment as outlined below. However, it is important to consider that the hazard identification step of risk assessment is typically a qualitative assessment, while quantitative steps are necessary to complete the risk assessment process. In this article, we will compare and contrast these two approaches for hazard identification, share experiences of the systematic review process as implemented in an industrial context, and consider whether systematic review processes have the potential to improve our ability to assure the protection of human health from chemicals that are present in the workplace, environment and consumer products. To discuss these two systematic review strategies in this context, it is important firstly to understand the chemical risk assessment process and how it is distinctly different from risk management, health regulation and decision making.

2. Chemical risk assessment & risk management

Chemical risk assessment is a scientific evaluation process that leads to an outcome, ideally quantitative on a risk scale, such that the risk can be communicated to policy- and safety decision-makers, in order that an appropriate risk management strategy can be devised and implemented. Risk is a function of the chemical (toxicological) hazard potency (that can be identified and characterised in toxicology studies) and the degree of chemical exposure (that is often modelled from a given exposure scenario). Toxicology studies are models of specific effects on the mammalian body; these can be in vivo or in vitro toxicology models, with guideline protocols specified by the Organisation of Economic Co-operation and Development (OECD) that are harmonized at an international level. If designed properly, a toxicology study can provide high confidence that there is no toxicological hazard or conversely identify a toxicological hazard, and, with careful dose selection, demonstrate a dose-effect relationship for a hazard that can be characterized quantitatively. From an industry perspective, as well as guideline toxicology studies that meet specified and internationally accepted regulatory standards, there may be peer reviewed literature from non-standard or academic research experiments that contribute valuable evidence and understanding, for example, to identifying a toxicological mode or mechanism of action. These data may be critical in understanding the relevance of animal studies to human health. In addition, human epidemiology or population-based studies may have been performed that can provide evidence or lack thereof of any adverse health effects arising from chemical exposure. The aim of the toxicity assessment step, which follows the hazard identification step, is to estimate the potential for hazard or risk and if possible define a point of departure (POD). This could either be a dose where no effects are likely to occur (i.e. a no observed (adverse) effect level (NO(A)EL)) or where there is good dose response information, a dose that has been shown to cause a specified level of effect (e.g. the benchmark dose at the 95th lowest percentile that shows a 10% increased incidence of effect ($BMDL_{10}$)). This toxicological benchmark can then be compared with real-life exposure to assess the level of risk that exposure could present to human health.

The third step of the risk assessment process is the exposure assessment or understanding the potential for exposure in the human population of concern. Workers, general public, and consumers are all exposed to chemicals differently, in terms of the sources, pathways and intake routes (oral, dermal, inhalation) of exposure and to different chemical concentration levels. There are a wide variety of exposure models that are used to estimate chemical exposure to the human body for example, CONSEXPO and CRÈME for consumer products, EFSA PRIMo and AOEL models for pesticides exposure, ChemSteer for environmental releases, E-FAST for consumer and general population exposure, to name but a few. The tools of physiologically-based pharmacokinetic (PBPK) modelling can also be used to estimate internal dose metrics (Campbell et al. 2012). Biomonitoring in humans can also be used as a measurement of exposure (Gentry et al. 2011). In short, there are many ways exposure levels can be modelled and estimated, that carry different levels of uncertainty associated with parameter input data and overall data integration into an exposure estimate.

The final step, the risk characterization, is then conducted by comparing the toxicological benchmark (toxicity assessment) and the exposure estimate (exposure assessment), to determine the difference between anticipated human exposure and a dose that is known not to cause adverse effects (NOAEL), or a 'Margin of Exposure' (MoE) away from a specified level of effect (e.g. a $BMDL_{10}$). Cancer risk estimates can also be estimated from human population-based studies. The data and the evidence can be used to map that scale of risk, but the scientist/risk assessor should not decide on what level of risk is acceptable for society. This is for policy-makers or designated safety-decision makers to take the output of the risk characterization and make those decisions in a transparent way, ideally with stakeholders, and with appropriate risk management strategies in place.

In any given situation, the level of risk that is deemed acceptable to society needs to be ascertained and agreed upon, ideally by (or on behalf of) those most affected by the chemical exposure. However, this should be accompanied by a clear explanation of the science so that stakeholders can understand levels of exposure that truly pose an increased risk of health effects. This is where clear risk communication and an understanding of risk perception come into play at the decision-making end; some people may say the risk should be zero (for a chemical with an identified hazard, this means exposure must be zero — and therefore this societal/political position can lead to a substance ban); some might say the risk can be 'negligible' or 'minimal' and then policy makers have to decide what these terms mean on the scale of risk e.g. is 1 in 1000,000 a negligible cancer risk; some might say the risk-benefit trade-off is significant enough that low risk can be tolerable/acceptable (e.g. 1 in 100,000 cancer risk estimate) as the benefits outweigh the risks. But already it can be seen that these decisions then move outside of the purely scientific and evidence-based assessment into the realms of political and societal considerations outside of the scope of the risk assessment itself. It is often in these areas that the chemical safety assessment becomes controversial, with the discussions to reach agreement on what constitutes 'acceptable risk' requiring greatest resource through stakeholder involvement.

3. The 'Navigation Guide' and the 'OHAT' approach — transparent hazard identification

The University of California 'Navigation Guide' and the OHAT approach are somewhat similar but there are some key process differences, as outlined in Table 1. Both are systematic review approaches that have been used to make hazard identification more transparent and objective. Neither approach goes as far as to consider the quantitative aspects of risk assessment (toxicity and exposure assessments) or uncertainty evaluations necessary in the risk characterization step of performing a chemical risk assessment. It is reported in Woodruff &

Table 1
A comparison of the Navigation Guide and OHAT approaches to systematic review.

Navigation Guide (Steps 1–3 are performed separately for human and non-human systems of evidence)	OHAT approach
1. Specify the study question — frame a specific question relevant to decision makers about whether human exposure to a chemical or class of chemicals or other environmental exposure is a health risk.	1. Problem formulation and protocol development — the objectives of the evaluation must be clearly stated, including the key questions to be addressed. The evaluation is structured to answer these key questions that guide the systematic-review process for the literature search, study selection, data extraction and synthesis. The questions define the populations, exposures, comparators, outcomes, timings, and settings of interest (PECOTS) eligibility criteria for the evaluation.
2. Select the evidence — conduct and document a systematic search for published and unpublished evidence.	2. Search for and select studies for inclusion — primary literature is searched. The search covers multiple databases (including, but not limited to, PubMed, TOXNET, Scopus, and Embase) with sufficient details of the search strategy documented in the protocol such that it could be reproduced. All references identified in the search are screened for relevance to the key question(s) of the evaluation based on the PECOTS eligibility criteria established when formulating the problem in step 1.
3. Rate the quality and strength of the evidence: rate the quality of individual studies and the quality of the overall body of evidence based on pre-specified and transparent criteria. The end result of Steps 1–3, is one of five possible statements about the overall strength of the evidence: • Known to be toxic • Probably toxic • Possibly toxic • Not classifiable • Probably not toxic	3. Extract data from selected studies — relevant data are copied or extracted from the publication. Human, animal and in vitro studies are evaluated separately. For each study, one member of the evaluation team performs the data extraction, and quality assurance procedures are undertaken as specified in the protocol (e.g., review and confirmation by another team member).
4. Grade the strength of the recommendations — involves integrating the strength of the evidence on toxicity (from Step 3) with information about exposure, the availability of less toxic alternatives, and patient values and preferences.	4. Assess the quality or risk of bias of individual studies — a) reporting quality (how well or completely a study was reported); b) internal validity or risk of bias (how credible the findings are based on the design and apparent conduct of a study); and c) external validity or directness and applicability (how well a study addresses the topic under review) 5. Rate the confidence in the body of evidence — methods based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group guidelines (GRADE 2014) 6. Translate the confidence ratings into levels of evidence — three descriptors used in step 6 ("high level of evidence," "moderate level of evidence," and "low level of evidence") directly translate from the confidence-in-the evidence ratings that exposure to the substance is associated with a heath effect, and a fourth designation ("evidence of no health effect") indicates confidence that the substance is not associated with a health effect 7. Integrate the information from different evidence streams (human, animal, and "other relevant data" including mechanistic or in vitro studies) to develop hazard identification conclusions: The five hazard identification conclusion categories are: • Known to be a hazard to humans • Presumed to be a hazard to humans • Suspected to be a hazard to humans • Not classifiable as a hazard to humans • Not identified to be a hazard to humans.

Sutton (2014) that the 'Navigation Guide' has only been applied to a case study (answering the question "does developmental exposure to perfluorooctanoic acid (PFOA) affect fetal growth?) in Steps 1–3. Interestingly, it is stated that resource was not available to complete Step 4 of the process in a case study setting, which is not surprising as this is arguably the most controversial step of the decision making process and where transparency is needed equally, if not more than for hazard identification.

In conducting systematic reviews in a truly transparent and structured manner, steps 4 (or 4 through 7 for OHAT) will require the development of a consistent scoring or rating approach that incorporates methods of reviewing the evidence that are agreed upon by the scientific and regulatory community. From an industry perspective, this will move us towards the elimination of the subjective nature of these steps that has been the source of controversy in hazard identification. The lack of agreed scientific methods for rating different data types and the use of mainly scientific judgment has resulted in criticisms of selected hazard identification steps as biased or focusing only on selected data which may not represent the body of evidence. The incorporation of "scoring" of all of the available data as part of the systematic review, to incorporate an evaluation of the quality of each line of evidence and 'risk of bias' in particular studies, will not only improve the transparency in integration of the evidence, but will allow for increased ease in characterizing the "weight" of the evidence. There are many scoring approaches in the literature, but a combination of these are likely to be necessary to improve hazard identification, as well as aid in the selection of the most appropriate, high quality studies for the quantitative steps of risk assessment (toxicity and exposure assessment).

4. Systematic review of a heterogeneous evidence-base including novel toxicology data types

With the advent of novel approaches, such as application of the mode of action framework, in toxicology now being structured around the concept of the 'adverse outcome pathway' (Villeneuve et al., 2014a, 2014b), it will become necessary in the near future to not only consider the classical observational toxicology and human epidemiology studies, but also more novel heterogeneous sources of data around mode/mechanisms of action from in vitro assays, high throughput approaches, computational predictive toxicology tools, categorization and read-across and systems biology approaches. Any systematic review framework derived for supporting chemical risk assessment must be future proofed to incorporate the new types of toxicology data that are emerging. As of now, the application of systematic review is more focused on traditional

toxicology assays that have been used for both hazard identification and toxicity assessment. However, it is foreseeable that the same generic principles listed in Table 1, can apply to the interpretation of new types of toxicology data as well as the traditional OECD guideline study types, but it may require specific quality and scoring criteria to be developed for these newer methods, particularly when there are known limitations, and applicability domains, as to how predictive single assay methods are at predicting human adverse health effects, or whether such assays only reveal information about potentiality and mechanism that needs to be considered in a holistic mode of action/adverse outcome pathway. This requires some careful detailed consideration going forward of how new relatively immature assays should be incorporated into a systematic review and considered in an overall weight-of-evidence assessment. Given the different technical specialisms that are required to interpret in vitro assay data, computational toxicology, systems biology, standard in vivo assays, epidemiology etc. it is likely that systematic reviews for substances where there are large bodies of heterogeneous evidence will require a multidisciplinary team effort, and to perform such assessments will have resource implications within industry for the future.

5. The path ahead for systematic review in supporting chemical safety decisions

In answer to the question posed in the introduction "can the process of a structured systematic review lead to increased transparency for clearer communications as to why certain decisions around chemical safety are reached?" We believe it can, if pragmatism is employed throughout the process. Exploring approaches to make better evidence-based and scientifically-informed decisions can only seek to improve the outcomes for society as a whole. The goal remains to maintain human health protection at the same time as enabling technological advances in the chemical industry that can bring real and tangible benefits to society, and striking that balance. Risk is only ever zero if either exposure is zero or the substance presents no toxicological hazard. We should continually consider to understand chemical risk in the context of the risks we encounter in daily activities and the fact that they will never be zero as well. In addition, there is always uncertainty around proving a negative, as well as with understanding a positive, and so in practice risk is not only almost never considered to be zero but should never be considered to be a single value. Systematic review can help structure the data and assess quality and confidence of the lines of evidence in a hazard evaluation into a clear and transparent record of the evidence. This moves us away from characterizing the data associated with potential hazard as either "positive" or "negative", but rather characterizing the available toxicity data as providing a certain "weight" of evidence for or lack of evidence for an adverse health outcome. It is feasible that a systemic approach to exposure modelling in any risk assessment can also be performed in a similar way to the structure provided by the systematic review of a hazard evaluation. However this still leaves the main issue, in that when bringing the toxicology and exposure information together, society (the public and/or their elected representatives or civil service) has to make choices about the level of calculated risk it is willing to accept in any given situation (Rodricks, 2007) and often policy makers find this difficult to address when not informed effectively by all stakeholders around a given decision. Clarity and transparency of evidence-based decision-making is crucial for decisions to be made that fit all of society in a considered and balanced way. Systematic review needs to go as far as making the whole risk-based safety decision-making process transparent. However, the resource implications of this should not be underestimated. The tools should be developed to ensure ease and speed of use, otherwise there is a risk that the time taken to perform a full and complete systematic review to come to decisions of increased confidence could hinder the speed to market of a new innovation and that will have consequences to industry, as well as the general public. Ideally, systematic review would be a tool that focused on critical and controversial questions in chemical risk assessment that are key to the safety decision or policy direction that needs to be set, and that these focused questions are identified as early as possible. There is scope for systematic review to increase the confidence in the hazard identification step of chemical risk assessment, however to be really effective in helping to increase confidence and transparency around the final safety decision, the transparency process needs to go further to cover exposure, the risk characterisation, the uncertainty evaluation and the value judgement on why a particular level of risk has been deemed acceptable according to the value judgements of the ultimate safety decision-maker. It is perfectly acceptable for different societies in different parts of the world to come to different conclusions using the same body of evidence, if their value judgments about the ultimate risk-benefit balance are different. However, at least the scientific data-driven and evidence-based components of chemical risk assessments should be consistent globally and can be documented as transparently and objectively as possible, incorporating a pragmatic process of systematic review to answer pre-defined questions using pre-defined criteria.

References

Campbell Jr., J.L., Clewell, R.A., Gentry, P.R., Andersen, M.E., Clewell III, H.J., 2012. Physiologically based pharmacokinetic/toxicokinetic modeling. Methods Mol. Biol. 929, 439–499.

Gentry, P.R., Clewell III, H.J., Clewell, R., Campbell, J., Van Landingham, C., Shipp, A.M., 2011. Challenges in the application of quantitative approaches in risk assessment: a case study with di-(2-ethylhexyl)phthalate Crit. Rev. Toxicol. 41 (Suppl. 2), 1–72.

Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M., 2014a. Adverse outcome pathway (AOP) development I: strategies and principles. Toxicol. Sci. 142 (2), 312–320.

Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M., 2014b. Adverse outcome pathway development II: best practices. Toxicol. Sci. 142 (2), 321–330.

Higgins, J.P.T., Green, S., 2011. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. Cochrane Collab. 2011 (Available from www.cochrane-handbook.org).

Rodricks, J.V. 2007. Calculated risks: the toxicity and human health risks of chemicals in our environment. Second edition, with Major Revisions and Additions. Cambridge University Press, New York, New York.

Rooney, A.A., Boyles, D.L., Wolfe, M.S., Bucher, J.R., Thayer, K.A., 2014. Systematic review and evidence integration for literature-based environmental health science assessments. Environ. Health Perspect. 122 (7), 711–718.

Woodruff, T.J., Sutton, P., 2014. The navigation guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. Environ. Health Perspect. 122, 1007–1014.