Environment International xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Environment International

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

Guidance on assessing the methodological and reporting quality of toxicologically relevant studies: A scoping review

Gbeminiyi O. Samuel [a], Sebastian Hoffmann [b], Robert A. Wright [c], Manoj Mathew Lalu [d], Grace Patlewicz [e,1], Richard A. Becker [f], George L. DeGeorge [g], Dean Fergusson [d], Thomas Hartung [a], R. Jeffrey Lewis [h], Martin L. Stephens [a,*]

[a] Johns Hopkins Center for Alternatives to Animal Testing, 615 N. Wolfe St., Baltimore, MD 21205, USA
[b] seh consulting + services, Stembergring 15, 33106 Paderborn, Germany
[c] William H. Welch Medical Library, Johns Hopkins University, 2024 E. Monument St., Suite 1-200, Baltimore, MD 21287, USA
[d] The Ottawa Hospital, The Ottawa Hospital Research Institute, Ottawa, Ontario K1Y 4E9, Canada
[e] DuPont Haskell Global Centers, 1090 Elkton Rd., Newark, DE 19711, USA
[f] Science and Research Division, American Chemistry Council, 700 2nd St., NE, Washington, DC 20002, USA
[g] MB Research Labs, 1765 Wentz Rd., Spinnerstown, PA 18968, USA
[h] ExxonMobil Biomedical Sciences, Inc., 1545 U.S. Highway 22 East, Room LA 350, Annandale, NJ 08801, USA

ARTICLE INFO

Article history:
Received 6 August 2015
Received in revised form 9 March 2016
Accepted 9 March 2016
Available online xxxx

Keywords:
Methodological quality
Reporting quality
Risk of bias
Scoping review
Toxicity studies

ABSTRACT

Assessments of methodological and reporting quality are critical to adequately judging the credibility of a study's conclusions and to gauging its potential reproducibility. To aid those seeking to assess the methodological or reporting quality of studies relevant to toxicology, we conducted a scoping review of the available guidance with respect to four types of studies: *in vivo* and *in vitro*, (quantitative) structure-activity relationships ([Q]SARs), physico-chemical, and human observational studies. Our aims were to identify the available guidance in this diverse literature, briefly summarize each document, and distill the common elements of these documents for each study type. In general, we found considerable guidance for *in vivo* and human studies, but only one paper addressed *in vitro* studies exclusively. The guidance for (Q)SAR studies and physico-chemical studies was scant but authoritative. There was substantial overlap across guidance documents in the proposed criteria for both methodological and reporting quality. Some guidance documents address toxicology research directly, whereas others address preclinical research generally or clinical research and therefore may not be fully applicable to the toxicology context without some translation. Another challenge is the degree to which assessments of methodological quality in toxicology should focus on risk of bias – as in clinical medicine and healthcare – or be broadened to include other quality measures, such as confirming the identity of test substances prior to exposure. Our review is intended primarily for those in toxicology and risk assessment seeking an entry point into the extensive and diverse literature on methodological and reporting quality applicable to their work.

1. Introduction

Research in toxicology, as in other fields, should be well-designed, rigorously conducted, and appropriately analyzed. These are key components of methodological quality. In clinical medicine, assessments of methodological or study quality typically focus on "risk of bias," *i.e.*, the degree to which the design, conduct, and analysis of a study could potentially compromise confidence in its results by introducing systematic error in the magnitude or direction of the results (Higgins and Green, 2008). Risks of bias include, for example, failure to randomize study subjects to treatment groups or failure to "blind" outcome assessors to the treatment groups being assessed. Beyond risk of bias, methodological quality can also include other considerations. Within toxicology, these include adherence to standardized test guidelines and Good Laboratory Practices. Methodological quality is sometimes referred to as "reliability" in toxicology (Klimisch et al., 1997).

Striving for high standards of methodological quality should be coupled with similar rigor for reporting the results of research in the literature. Research should be reported accurately, thoroughly, and

* Corresponding author.
E-mail addresses: gsamuel4@jhu.edu (G.O. Samuel), Sebastian.hoffmann@seh-cs.com (S. Hoffmann), rwrigh32@jhmi.edu (R.A. Wright), manojlalu@gmail.com (M.M. Lalu), grace.y.tier@usa.dupont.com, Tier.Grace@epa.gov (G. Patlewicz), rick_becker@americanchemistry.com (R.A. Becker), degeorge@mbresearch.com (G.L. DeGeorge), dafergusson@ohri.ca (D. Fergusson), thartun1@jhu.edu (T. Hartung), r.jeffrey.lewis@exxonmobil.com (R.J. Lewis), msteph14@jhu.edu (M.L. Stephens).
[1] Present address: US EPA, National Center for Computational Toxicology (NCCT), 109 T. W. Alexander Dr., Research Triangle Park, NC 27711, USA.

http://dx.doi.org/10.1016/j.envint.2016.03.010

transparently. Reporting quality (sometimes referred to as "completeness of reporting") (Moher, 2015) is distinct from methodological quality but the two concepts overlap in a number of ways. Thorough reporting helps in the assessment of the methodological quality of a study. For instance, including only statistically significant results in a research paper is an example of both poor reporting and a risk of bias ("selective outcome reporting") (Guyatt et al., 2011a). Consequently, an appraisal of both methodological and reporting quality is essential to ensure that accurate information is derived from published research.

A considerable body of literature has addressed methodological and reporting quality, providing guidance not only on retrospectively assessing the quality of published studies but also on prospectively designing, conducting, analyzing, and reporting new studies. In this paper, we summarize both types of guidance. We address several types of studies of direct relevance to assessing the hazards and risks of environmental chemicals, namely, *in vivo* and *in vitro* (or mechanistic) studies, *in silico* studies (represented here by studies of (quantitative) structure–activity relationships ((Q)SARs), studies of physico-chemical properties, and observational human studies. *In vivo* studies examine effects on living animals, whereas *in vitro* studies examine effects on biomolecules, cells or tissues, from animals or humans. (Q)SARs are approaches that relate the properties of a chemical encoded in its molecular structure to a physical property or to a biological effect, *e.g.*, toxicity. Studies of physico-chemical properties investigate, for example, a chemical's octanol–water partition coefficient, providing information that can guide subsequent toxicity testing. Human observational studies may explore the relation between human exposure to an environmental agent and a health effect. Such studies include various types (*e.g.*, case–control, cohort, and cross-sectional).

For each study type, our aims were (1) to identify and summarize the available guidance on prospectively ensuring or retrospectively assessing methodological and reporting quality, and (2) to distill the common elements from this guidance. We adopted a scoping review approach. A scoping review "is a form of knowledge synthesis that addresses an exploratory research question aimed at mapping key concepts, types of evidence, and gaps in research related to a defined area or field by systematically searching, selecting, and synthesizing existing knowledge" (Colquhoun et al., 2014). Frameworks for the conduct of scoping reviews are emerging, and reporting guidelines are still in preparation (Colquhoun et al., 2014). Broadly speaking, scoping reviews identify the research topic; identify and select relevant studies; chart the data; collate, summarize, and report the results; and consult with relevant stakeholders (Arksey and O'Malley, 2005).

The literature on methodological and reporting quality has a rich history in clinical medicine and healthcare, thanks in part to an emphasis on evidence-based medicine. Our review emphasizes the relevance of this literature to toxicology and its diverse study types. It is intended primarily as an entry point into this literature for those in toxicology and risk assessment who wish to assess the methodological and reporting quality of research. Such assessments are usually retrospective (*e.g.*, evaluating published studies) but can also be prospective (*e.g.*, evaluating grant proposals). Apart from the assessment context, toxicologists have an obvious interest in ensuring the methodological and reporting quality of their own planned research.

Although toxicologists have grappled with issues of methodological and reporting quality over the years, some of the relevant terminology that has emerged primarily from other fields may be unfamiliar to toxicologists. Consequently, we provide a glossary of key terms in Table 1.

2. Methods

To retrieve published guidance on assessing or ensuring the quality of various types of studies relevant to toxicology, literature searches were devised and conducted with the aid of an information specialist (Appendix). Search strategies used a combination of controlled vocabulary and keywords adapted to each database searched. They were designed to achieve a balance of precision and recall in the results. There was no restriction on publication dates. Experts in toxicity research were consulted to identify any additional guidance.

Table 1
Glossary of key terms.

Allocation concealment: A process that it used to prevent selection bias. The person allocating subjects to experimental arms is unaware of which arm the subjects are being allocated until the moment of assignment. This prevents researchers from (unconsciously or otherwise) influencing the allocation of subjects (National Research Council, 2014; http://www.consort-statement.org/resources/glossary).
Attrition bias: Systematic differences in excluding study units between groups
Bias: Systematic deviation of the estimated intervention/exposure effect away from the "truth." This can be caused by inadequacies in the design, conduct, or analysis of an experiment, and produce deviations in either direction (*i.e.* under or over-estimate) (http://www.consort-statement.org/resources/glossary; handbook.cochrane.org/chapter_8/8_2_2_risk_of_bias_and_quality.htm).
Blinding (*or masking*): A set of procedures that keeps the participants and personnel involved in a study unaware of which intervention/exposure was received; this reduces the risk of performance bias. Similarly, outcome assessment can be blinded, so that personnel who assess outcome measures are unaware of the treatment allocation; this reduces the risk of detection bias (National Research Council, 2014).
Confounding bias: Systematic differences in factors potentially influencing the results between groups.
Detection bias: Systematic differences in the outcome assessment between groups
External validity: The extent to which a study provides a correct basis to generalize to other circumstances (Henderson et al., 2013).
Good Laboratory Practices (*GLPs*): A framework for study design, conduct, and oversight that reduces the risk of bias that can be associated with the adequacy of temperature, humidity, and other environmental conditions; experimental equipment and facilities; animal care; health status of animals; animal identification; separation from other test systems; and presence of contaminants in feed, soil, water, or bedding (National Research Council, 2014).
Internal validity: The extent to which the design and conduct of study minimizes bias and systematic error (Grimes and Schulz, 2002; http://www.consort-statement.org/resources/glossary).
Methodological quality: The extent to which the design and conduct of a study is likely to have prevented systematic errors (bias) (Olivo et al., 2008) and, as a result, identified "the truth" in its results and inferences. This term is quite similar to risk of bias.
Performance bias: Systematic differences introduced during the study.
Randomization: Randomly allocating an intervention under study across the comparison groups to ensure that group assignment cannot be predicted (National Research Council, 2014).
Reporting bias: Systematic omission of results in the study documentation/publication.
Reporting quality: Providing a complete and transparent description of the design, conduct, and analysis of a study (Moher et al., 1995). Also known as "reporting completeness."
Risk of bias: The risk of a systematic error or deviation from the truth in results or inferences. This term is interchangeable with internal validity (handbook.cochrane.org/chapter_8/8_2_2_risk_of_bias_and_quality.htm)
Scoping review: A form of knowledge synthesis that incorporates a range of study designs to comprehensively summarize and synthesize evidence with the aim of informing practice, programs, and policy and providing the direction for future research priorities (Colquhoun et al., 2014).
Selection bias: Systematic differences in the comparison groups.
Selective outcome reporting: The reporting of only selected results, not all results.

Please cite this article as: Samuel, G.O., et al., Guidance on assessing the methodological and reporting quality of toxicologically relevant studies: A scoping review, Environ Int (2016), http://dx.doi.org/10.1016/j.envint.2016.03.010