Regulatory Toxicology and Pharmacology 67 (2013) 157–169

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

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

The OSIRIS Weight of Evidence approach: ITS for the endpoints repeated-dose toxicity (RepDose ITS)

Inga Tluczkiewicz [a,f,*], Monika Batke [a], Dinant Kroese [b], Harrie Buist [b], Tom Aldenberg [c], Eduard Pauné [d], Helvi Grimm [a], Ralph Kühne [e], Gerrit Schüürmann [e,f], Inge Mangelsdorf [a], Sylvia E. Escher [a]

[a] Fraunhofer Institute for Toxicology and Experimental Medicine, Nikolai-Fuchs Straße 1, 30625 Hannover, Germany
[b] TNO Quality of Life, Utrechtseweg 48, 3704 HE Zeist, The Netherlands
[c] RIVM, Antonie van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands
[d] Simpple, Cr. Joan Maragall 1, 43003 Tarragona, Spain
[e] UFZ Department of Ecological Chemistry, Helmholtz Centre for Environmental Research, Permoserstraße 15, 04318 Leipzig, Germany
[f] Institute for Organic Chemistry, Technical University Bergakademie Freiberg, Leipziger Str. 29, 09596 Freiberg, Germany

ARTICLE INFO

Article history:
Available online 21 February 2013

Keywords:
Integrated Testing Strategy
ITS
Repeated-dose
Risk Assessment
TTC
Non-testing method
OSIRIS
Weighting
Alternative
3R principle

ABSTRACT

In the FP6 European project OSIRIS, Integrated Testing Strategies (ITSs) for relevant toxicological endpoints were developed to avoid new animal testing and thus to reduce time and costs. The present paper describes the development of an ITS for repeated-dose toxicity called RepDose ITS which evaluates the conditions under which *in vivo* non-guideline studies are reliable. In a tiered approach three aspects of these "non-guideline" studies are assessed: the documentation of the study (reliability), the quality of the study design (adequacy) and the scope of examination (validity).

The reliability is addressed by the method "Knock-out criteria", which consists of four essential criteria for repeated-dose toxicity studies. A second tool, termed QUANTOS (Quality Assessment of Non-guideline Toxicity Studies), evaluates and weights the adequacy of the study by using intra-criterion and inter-criteria weighting. Finally, the Coverage approach calculates a probability that the detected Lowest-Observed-Effect-Level (LOEL) is similar to the LOEL of a guideline study dependent on the examined targets and organs of the non-guideline study. If the validity and adequacy of the non-guideline study are insufficient for risk assessment, the ITS proposes to apply category approach or the Threshold of Toxicological Concern (TTC) concept, and only as a last resort new animal-testing.

1. Introduction

For human risk assessment usually *in vivo* tests in rodents are needed to conclude on the toxicological potency of substances. Currently under REACH, the toxicological properties of thousands of chemicals have to be assessed. This also implies that numerous *in vivo* tests have to be performed to protect human health. Besides jeopardizing animal welfare, *in vivo* tests are time and cost intensive and laboratory testing facilities are limited.

In the European project OSIRIS (Optimized Strategies for Risk Assessment of Industrial Chemicals through Integration of Non-Test and Test Information) Integrated Testing Strategies (ITSs) were developed to give guidance on how to assess the quality of older non-guideline studies and how to integrate alternative methods and thereby to avoid animal tests. Four toxicologically relevant endpoints for humans were considered: repeated dose toxicity, mutagenicity and carcinogenicity, and skin sensitization (see Vermeire et al., 2013).

An ITS aims to gather all available information on the test substance first, including data from *in vivo* tests, alternative tests (*in vitro*, chemoassays e.g. Böhme et al., 2010 and Thaens et al., 2012, genomic tests) and non-testing information, such as structural alerts (e.g., Blaschke et al., 2010, 2012; Schramm et al., 2011), mathematical methods, QSAR models (e.g., Wondrousch et al., 2010; Mulliner et al., 2011), computational chemistry (Ji and Schüürmann 2012a, 2012b) or data on structurally similar substances for read across (e.g. Schüürmann et al., 2011). For a general description of the ITS approaches for human toxicity within OSIRIS (2007), refer to Vermeire et al. (2013) in this issue. The validity, relevance, and reliability of the alternative tests and non-testing information differ. A main challenge of the ITS is therefore the weighting of different types of information to conclude whether the compiled data are sufficient for risk assessment or whether a data gap exists. If a data gap is identified an ITS first prioritizes non-testing and alternative methods with regard to the 3R principle and animal welfare (Russel and Burch, 1959). New *in vivo* testing is foreseen if none of the alternative methods can be applied.

* Corresponding author Address: Fraunhofer Institute for Toxicology and Experimental Medicine, Nikolai-Fuchs Straße 1, 30625 Hannover, Germany. Tel.: +4905115350-347; fax: +49 5115350 335.
E-mail address: inga.tluczkiewicz@item.fraunhofer.de (I. Tluczkiewicz).

0273-2300/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.yrtph.2013.02.004

The present publication describes the development of the RepDose ITS, an Integrated Testing Strategy for repeated-dose toxicity (RDT) studies in rodents. RDT is assessed in study types such as subacute, subchronic and chronic and is therefore not a categorical endpoint such as mutagenicity or sensitization but a collection of continuous endpoints (e.g. changes in body weight, mortality increased). The overall result of a RDT study, a No Observed Effect Level (NOEL) or Lowest Observed Effect Level (LOEL), can be any value between the lowest dose and the highest dose tested in the respective study.

Repeated-dose toxicity *in vivo* can be caused by unspecific as well as numerous specific mechanisms of action, which are often not known. Because of this complex situation *in vivo*, non-testing methods such as QSAR models or alternative tests such as *in vitro* methods are currently not available to completely replace *in vivo* RDT studies. Many "old" non-guideline studies are however available, which provide to some extent relevant information on the toxicity of the target substance. The RepDose ITS therefore focuses on two major building blocks: the assessment of the validity of "old" non-guideline studies and the application of two alternative methods: the Threshold of Toxicological Concern (TTC) and read-across.

This article describes the development of the ITS concept and its building blocks and its application to subacute, subchronic and chronic studies is demonstrated. The application of the RepDose ITS concept to cancer studies is documented in more detail in Buist et al. (2013).

2. Overview of the RepDose ITS

Based on the general ITS scheme described by Vermeire et al. (2013), the ITS concept was adapted to the endpoint repeated-dose toxicity. The outline of the RepDose ITS is depicted in Fig. 1. In the following the main steps are described based on the building blocks of the ITS.

2.1. Step 1: Gather all relevant in vivo data

First the user has to gather already available information for the endpoint repeated-dose toxicity. Ideally these include all kinds of public or in-house repeated-dose toxicity studies with study durations from 10 days to lifetime. For the endpoint carcinogenicity in lifetime studies the ITS on carcinogenicity from Buist et al. (2013) can be used.

Actually, the RepDose ITS is applicable to oral studies, covering administration by feed, drinking water or gavage. Single parts of the ITS are also applicable to the evaluation of inhalation studies. The quality assessment tool QUANTOS is applicable to both routes of exposure – oral and inhalation as the tool evaluates general study parameters which are valid for both routes. Up to now the coverage approach is, however, only applicable to oral toxicity studies. Main target organs depend on the route of application and it has been shown that targets of the respiratory tract do frequently trigger the LOEC in inhalation studies (Escher et al. 2010).

2.2. Step 2: assess data quality of the gathered data. If non guideline studies are available examine the validity and adequacy of the non-guideline studies

In step 2 the quality of all gathered studies is assessed according to Klimisch et al. 1997 (see Vermeire et al., 2013). It can be assumed that studies conducted according to a guideline (Klimisch code 1) are sufficient to be used for risk assessment. Therefore further assessment of the validity and adequacy is not needed. For Klimisch code 4 studies the quality is not assignable. The RepDose ITS therefore focusses on the validity and adequacy of Klimisch code 2/3 studies and studies for which the user did not assign a Klimisch code. In this context the ITS evaluates three different aspects:

- Documentation of the study (reliability).
- Quality of the study design (adequacy).
- Scope of examination (validity).

The minimal requirements on the documentation of the study are given as a set of four knock-out criteria (Section 3.1). If any of these four criteria is not documented, the study is judged to be not sufficient for risk assessment and classification and labeling.

If the knock-out criteria are not applicable, subsequently the study design will influence the reliability of the study results. The QUANTOS (Quality Assessment of Non-guideline Toxicity

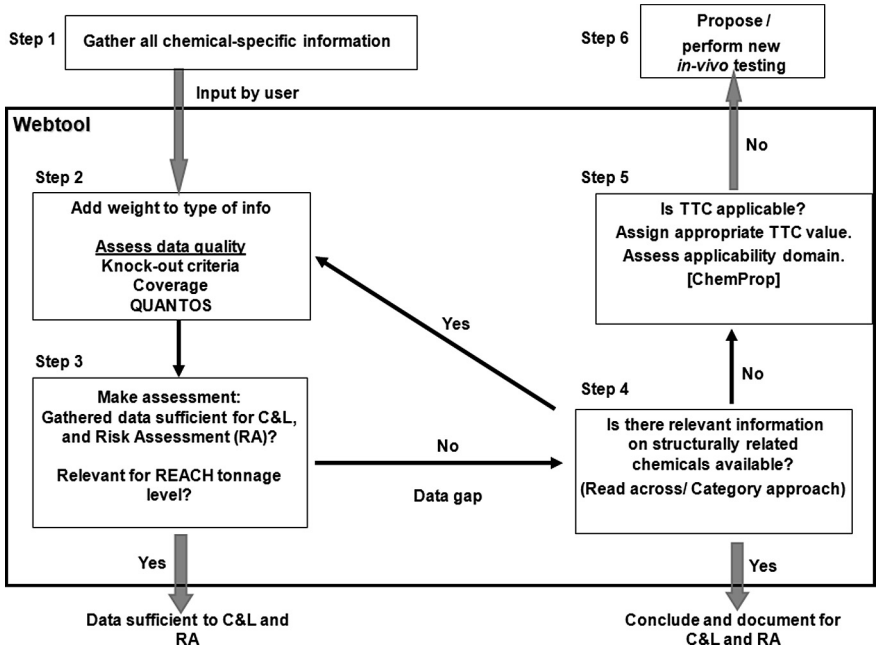

Fig. 1. RepDose ITS outline.