Available online at www.sciencedirect.com
ScienceDirect
Regulatory Toxicology and Pharmacology 50 (2008) 412–419
Regulatory Toxicology and Pharmacology
www.elsevier.com/locate/yrtph

Towards more efficient testing strategies—Analyzing the efficiency of toxicity data requirements in relation to the criteria for classification and labelling

Anna Nordberg *, Christina Rudén, Sven Ove Hansson

Department of Philosophy and the History of Technology, Royal Institute of Technology, Teknikringen 78B, SE-100 44 Stockholm, Sweden

Received 19 July 2007
Available online 1 February 2008

Abstract

This contribution is based on the assumption that the aim of toxicity testing as required by chemicals legislation is to identify as many chemicals of concern to human health and the environment as possible, given a limited amount of resources allocated to testing. Based on this assumption we propose a method for the optimization of test systems for industrial chemicals, based on the calculation of *efficiency ratios* for tests and test systems. The efficiency ratio of a toxicity test depends on the monetary cost of performing the test and the probability that the test will identify a chemical of concern, as estimated by the rules for classification and warning labelling. Efficiency ratios are estimated based on the results of basic standardized toxicity testing for acute toxicity, subacute toxicity, irritation and sensitization of 1409 industrial chemicals notified in the European Union between 1994 and 2004. This careful evaluation of the regulatory consequences of testing indicated that many of these substances are classified based on short-term testing of acute toxicity, irritation and sensitization and that most of the substances classified due to a 28-day study were also classified based on short-term testing. These results indicate that, within the classification and labelling system, it is currently more efficient to perform short-term testing of a larger number of substances rather than to perform subacute toxicity studies on substances already tested for acute toxicity.

Keywords: Testing strategies; Classification; Efficiency ratio; Acute toxicity; Subacute toxicity

1. Introduction

Toxicological and ecotoxicological testing as a part of regulatory risk assessment of chemicals aims at providing sufficiently accurate predictions of adverse effects on human health and the environment to guide risk management measures. Furthermore, decisions about testing regimes and regulatory test requirements need to take the costs of performing these tests into account.[1] Resource restrictions make it impossible, at least in the short run, to perform extensive testing programs for all industrial chemicals that are subject to regulation, although this would be needed to base risk management on the best possible scientific evidence. Testing has therefore to be prioritized and the choice of tests, and their combination into test systems, should also be judged according to how they support risk management. One promising approach to the prioritization of testing is therefore to focus on their expected regulatory impact. This means that a test that has low probability of making a difference for a risk management decision should be given lower priority.

For an example of this, consider a substance that has not yet been tested for either acute or subacute toxicity.[2]

* Corresponding author. Fax: +46 8 790 95 17.
E-mail address: anna.nordberg@infra.kth.se (A. Nordberg).

[1] In this paper, we only consider the monetary cost of performing the tests. In a subsequent paper, we will include also the cost of animal welfare for tests that use animal experiments.

[2] A well-informed risk management decision should of course be based on both these types of tests (and several others), but since resources are limited many existing industrial chemicals lack (publicly available) data from both these types of tests.

0273-2300/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.yrtph.2008.01.012

How important is it to perform both acute and subacute toxicity tests? Would it be sufficient to perform just the acute toxicity testing (and perhaps use the resources saved in this way to test some other substances that we could not otherwise afford testing)? The answer to this question will depend on the correlations between the results obtained from the acute and the subacute toxicity test. Suppose on the one hand that there is a good correlation between subacute toxicity and acute toxicity. An acute health effect may manifest at higher exposures compared to the exposures causing an adverse subacute effect (after repeated doses, usually tested in a 28-day study). Reducing the exposures below the levels causing acute effects will therefore not always protect against the long-term effects. Still, a reduction in exposures will contribute to reducing long-term risks. Then, for many substances risk management actions that protect us against acute effects will also contribute to the protection against the unknown subacute effects. But suppose instead that there is no correlation between these different types of hazards. Then measures against acute effects will not protect us against long-term effects. The need for subacute toxicity testing will then be larger. In both cases, risk management decisions based on short-term testing will provide direct protection against acute effects. In the former, but not the latter case, it may also provide us with indirect protection against subacute effects. Clearly, the priority that we should assign to the testing of substances for subacute toxicity that have already been tested for acute toxicity will depend on the size of this indirect protective effect.

There is a reasonably large body of research exploring the correlation between outcomes from tests of different toxicological endpoints. In one of the several studies that have found correlations between subacute and chronic toxicity data, Kramer et al. (1996) found the No Observed Adverse Effect Level for subacute studies ($NOAEL_{subacute}$) to be a predictor for the corresponding chronic value ($NOAEL_{chronic}$).

However, as mentioned above, knowledge about correlations between different types of toxicity needs to be supplemented with information about the relationship between the outcomes of these tests and risk management decisions. One recent study by Hoffmann et al. (2005) investigated the impact of the *in vivo* skin irritation test on the classification and labelling of substances according to the European regulations and to the GHS (Globally Harmonised System). They found the frequency of skin irritants to be 10% among a large set of industrial chemicals tested for this endpoint. Their study also indicates that the European system is biased towards overclassification of this effect. These authors propose that the prevalence of the effect tested for should be taken into consideration when designing testing strategies.

In this article, we will propose a method for the optimization of test systems for basic toxicity testing of industrial chemicals with the aim of increasing the cost efficiency in regulatory testing to enable testing of a large number of previously insufficiently tested chemicals. The purpose and methods of the study are stated in Section 2. This is followed by the results of the empirical study in Section 3. Finally, in Section 4, our conclusions from the study are presented.

2. Purpose and methods

We have explored the testing procedures and the decisions on toxicity classifications and warning labelling taken for industrial chemicals introduced on the European market between 1994 and 2004. In particular we have utilized the risk phrases for labelling assigned to substances classified according to data from studies of acute and subacute toxicity, irritation and sensitisation in accordance with the classification and labelling directive. Based on information on the prevalence of different risk phrases and the relative cost of the testing required to arrive at that particular risk labelling, an *efficiency ratio* was developed for the purpose of this article. The efficiency ratio is used to evaluate the usefulness of these different standardized tests from a risk management perspective. The overall purpose of this exercise is to improve the efficiency of regulatory testing of industrial chemicals.

This investigation is based on information obtained from the European classification and labelling system, and the former legislation for notification of new industrial chemicals (Commission Directive 92/32/EC, Commission Directive 2001/59/EC). The legislation on new chemicals has as of June 1, 2007 been replaced by the REACH system, but this does not affect this investigation as we have only studied substances regulated under the old legislation. The legislation for new industrial chemicals specified the testing that should have been performed before a new chemical was introduced on the European market. Data according to fixed test batteries were required, and the test requirements depended on the volume to be marketed. These testing requirements were implemented quite strictly, meaning that there was little room for manufacturers and importers to avoid performing any of the tests (information from the Swedish Chemicals Agency).

In Section 2.1, the legislation for new substances is outlined. In Section 2.2, the database we have used is described and in Section 2.3 we explain how we evaluate the efficiency of test requirements in relations to the classification and labelling criteria.

2.1. The classification and labelling system

Commission Directive 2001/59/EC regulates the criteria for classification and warning labelling of chemical substances and preparations within the European Union. This system classifies substances and preparations according to their chemical and toxicological properties into the following classes: Explosive, Extremely flammable, Highly flammable, Flammable, Oxidising, Very toxic, Toxic, Corrosive, Harmful, Irritant, Sensitisation, Carcinogenic, Mutagenic, Toxic to reproduction and Dangerous for the environment. The classification and labelling of a substance is constituted of danger symbols, safety phrases and risk phrases (R-phrases). For the purpose of this study we have only studied the use of tests for acute toxicity, subacute toxicity, irritation (skin and eyes) and sensitisation and the corresponding classifications as Very toxic, Toxic, Harmful, Irritant and Sensitisation (see Table 1).

The classification and labelling legislation does not make any testing mandatory. Instead this legislation defines criteria for the interpretation of available data, used to categorize chemicals (and mixtures) according to their toxicity and other properties. For each of the classification categories, there is a set of criteria for classification, sufficient to place the substance in a particular category. Hence, a substance can be classified as "toxic" either due to the outcome of a test of its acute toxicity or due to the outcome of a test of its toxicity after repeated doses (Commission Directive 2001/59/EC).