Review Hypothesis

Journal of Applied Toxicology

Received: 20 March 2014, Revised: 11 June 2014, Accepted: 11 June 2014 Published online in Wiley Online Library: 25 July 2014

(wileyonlinelibrary.com) DOI 10.1002/jat.3045

A proposal to improve clarity and communication in the EU Classification process for chemicals for carcinogenicity and reproductive and developmental toxicity

J. E. Doe*

ABSTRACT: There is an issue in the EU classification of substances for carcinogenicity and for reproductive or developmental toxicity which has brought difficulties to those involved in the process. The issue lies in the inability of the classification system to distinguish between carcinogens and reproductive toxicants with different levels of concern. This has its origins in the early years of toxicology when it was thought that a relatively small number of chemicals would be either carcinogens or reproductive toxicants, but this has turned out not to be the case. This can cause problems in communicating to the users of chemicals, including the public, the nature of the hazard presented by chemicals. Processes have been developed within the classification system for setting specific concentration limits which assess the degree of hazard for carcinogens and reproductive toxicants as high, medium or low. However these categories are not otherwise used in classification. It is proposed that their wider use would bring the advantages of transparency, clarity of communication, certainty of the process and would allow chemicals with a high degree of hazard to be identified and managed in an appropriate way. Copyright © 2014. The Authors. Journal of Applied Toxicology Published by John Wiley & Sons Ltd.

Keywords: carcinogenicity; reproductive toxicity; classification; degree of hazard; hazard characterization

What is the Issue?

The issue in the EU classification of substances for carcinogenicity and for reproductive toxicity is best summed up by the title of a paper by the inventor of the Ames test 'Chemical Carcinogenesis: Too many rodent carcinogens' (Ames and Gold, 1990). The system is seen to be too restrictive by many; it is seen to be too lenient by others; and it confuses the public. The origins of this issue are complicated and go back more than 40 years. The purpose of this article is to explore the issue and to suggest a way forward.

Background to Classification

Classification, labelling and packaging (CLP) in the EU was originally developed as a way of providing information for the packaging, labelling and sale of chemicals to occur (ECHA, 2012a). Before its introduction there was no agreed way of describing the potential hazardous properties of chemicals. Each company was at liberty to devise its own way of assessing and describing its products, making it difficult for purchasers to decide how to handle them. During the 1970s, individual countries started to develop classification schemes to harmonize activities within their own boundaries, but this did not address the problems of cross border trade. In the 1980s the EU developed a harmonized system across Europe, the 6th Amendment to Dangerous Substances Directive (EEC, 1979), which created one scheme for the whole European Community. In the 2000s attempts have been made to create one globally accepted scheme with the so-called Global Harmonization System (GHS, 2007) under the aegis of the United Nations. In turn, the GHS has been adopted by the EU into the new CLP Regulation introduced in 2008 (EC, 2008).

The range of hazardous properties that classification has embraced has expanded since its inception. Originally it focused on physicochemical hazards such as volatility, flammability and explosivity. The concept was then extended to the harm that chemicals could pose by their toxicity to humans or in the environment. This started with classification based on the results of acute toxicity tests for lethality and local toxicity tests for corrosivity, irritancy and sensitization. These tests have numerical outputs such as LD50, or scores from a rabbit skin or eye irritancy test, which made it possible to set criteria for classification which could be assessed objectively. There continues to be debate over whether the criteria are set in the correct place, but the classification can be determined from the data without relying on the judgement of the assessor.

Bringing Carcinogenicity, Mutagenicity and Reproductive Toxicity into Classification

During the 1980s and 1990s both the science of toxicology and the ambition of Classification grew. It was recognized that

*Correspondence to: J. E. Doe, Parker Doe Partnership LLP, Box 139, Frodsham, Cheshire, WA6 1AZ, UK.
E-mail: john.doe@parkerdoe.com

Parker Doe Partnership LLP, Box 139, Frodsham, Cheshire, WA6 1AZ, United Kingdom

Journal of Applied Toxicology

chemicals could cause more than acute and local toxicity and concerns extended to carcinogenicity, mutagenicity and reproductive toxicology (CMR). These properties were seen to pose a particular threat to human health because of their irreversible nature and because the result of exposure was not seen until sometime after the event. These are two major factors in increasing the perception of risk (Slovic, 1987) so it is perhaps not surprising that a view was taken to treat these properties differently from other adverse effects. The original hypothesis was that there were relatively small numbers of chemicals which were mutagenic, carcinogenic or toxic to reproduction and/or development and that the assays would identify them. Once identified, their use should be restricted to protect human health. These concerns were exacerbated once it was realized that mutagenicity was one of the mechanisms of action of the most potent carcinogens which react directly with DNA, and there was theoretically no threshold for these effects although this has been questioned in practice (Purchase and Auton, 1995).

Extending Classification into Risk Management

Full risk assessment requires a full range of toxicology studies and a full understanding of the range of uses of a chemical (Leeuwen and Vermieire, 2007). Most industrial chemicals do not have a full range of toxicology studies and they are used in many different ways depending on their chemistry. However, if a chemical was classified as having one of the CMR effects it seemed sensible to restrict their use and classification started to be used as the basis for risk management. At first this was seen in the way companies selected chemicals to use in their products and then it started to find its way into regulatory schemes such as that for Crop Protection Chemicals where there was a presumption against any Category 2 (now 1B) CMR being listed on annex 1 as available for use in EU after risk assessment (EEC, 1991).

CMR Classification Issues

However, issues emerged early in the Classification of CMRs which are still being addressed. Unlike acute and local toxicity, the data on which the judgements for CMR have to be made are not as clear cut. In the beginning of society's assessment of carcinogenicity, epidemiological studies were the main source of information. Bradford Hill (1965) developed a set of 'viewpoints' or 'features to be considered' to help to distinguish between association and causality in epidemiology studies and these were used to assess whether the results indicated that a chemical should be considered to be a carcinogen in humans. These considerations were usually performed by a panel of experts considering all the evidence and seeking consensus. In the original Classification schemes a special category was reserved for 'known human carcinogens' i.e. those chemicals which had gone through this rigorous process and were confirmed as carcinogens in humans, and this is still the case with the 2008 CLP Regulation (EC, 2008) where Category 1A is for 'known human carcinogens'.

The focus of these activities was to identify carcinogens, for it was thought that chemicals were either carcinogens or non-carcinogens. Once the carcinogens had been identified they could be removed and health would be safeguarded. The list of chemicals which had been identified as 'known human carcinogens' by the mid-1970s became a tool for toxicologists for the development of the understanding of the mechanism of carcinogenicity and for the development of assays which could predict carcinogenicity (Monro, 1993). The chemicals on the list were the subject of 2-year bioassays and they formed the positive controls for the development of the so-called short-term carcinogenicity assays. Most of the 'known human carcinogens' caused an increase in tumours in experimental animals in the 2-year bio-assays. In addition to the 'known human carcinogens', a number of other chemicals also caused an increase in tumours in rodents. Although the detection rate for carcinogenicity for the short-term tests was high for the 'known human carcinogens', their ability to detect the 'animal carcinogens' was not as good. As time went by it was discovered that about 50% of all the chemicals put through the 2-year bioassay were considered to have caused an increase in one or more tumours in experimental animals (Ames and Gold, 1990). It soon became obvious that the so-called short-term tests for carcinogenicity were actually tests for genotoxicity and their ability to predict genotoxic carcinogens is reasonable, provided that care is taken in interpreting the results of *in vitro* and *in vivo* assays.

Extrapolating the Results of Animal Studies to Humans

Interest in non-genotoxic carcinogens has centred on their relevance to humans. Questions were raised over the probability that 50% of all the chemicals in use were causing cancer in humans (Ames & Gold, 1990; Monro, 1993). The relevance of this type of animal carcinogen has been the subject of debate for many years. Some modes of action are considered to be non-relevant to humans, such as the induction of male rat kidney tumours by α-2-μ-globulin induction shown to be male rat specific and not relevant to humans (Swenberg, 1993). Others such as the induction of mammary tumours by chemicals which cause hyperprolactinaemia have been shown to have concordance in humans and are considered to be relevant (Harvey *et al*., 2006). However, the human relevance has not been established for most chemicals which cause tumours in 2-year bioassays. In the absence of evidence to the contrary, the findings are considered to be relevant to humans but they are generally recognized to be less of a hazard than genotoxic carcinogens because the animal tumours appear only after prolonged changes in the relevant organ or system (Monro, 1993). Unlike genotoxicity where there is a direct interaction with DNA and in theory one molecule of chemical can react with one molecule of DNA to cause a change which leads to cancer, the non-genotoxic carcinogens behave in a similar way to other forms of toxicity and a minimum dose or threshold is needed to cause the toxicity which has to be prolonged for carcinogenicity to be the end result. Protecting against the initial toxicity prevents the resulting carcinogenicity.

Degrees of Hazard

A convention grew up in the Classification of chemicals for carcinogenicity whereby chemicals with a genotoxic mode of action which caused tumours in animals were classified as presumed human carcinogens (category 1B). Chemicals which caused tumours in animals by a non-genotoxic mode of action were classified as suspect human carcinogens (category 2), or

not classified if either the evidence in laboratory animals was inconclusive or the mode of action was not relevant to humans. This convention resulted in the most potent carcinogens, that is those chemicals which could cause cancer at low doses or after relatively short exposure, being highlighted and subjected to stringent risk management, including withdrawal. The other chemicals which could only cause cancer after prolonged dosing at high levels were subjected to scrutiny, but could be used with care in the same way as other chemicals causing adverse effects after repeat dosing.

A similar situation existed with reproductive and developmental toxicity. Limited numbers of chemicals were determined as a result of epidemiological studies to cause adverse effects on fertility in humans or to cause abnormalities in human babies. These were classified as known human reproductive toxicants (category 1A) and they were subjected to stringent risk management. As a result of the findings of the epidemiology studies, animal studies were developed to identify chemicals which could have an adverse effect on reproduction and development. These studies are used to categorize chemicals for which there is no epidemiological evidence or before they are introduced, and they have become part of the mandatory studies for pharmaceuticals and plant protection chemicals. Experience over the past 30 years has shown that there are a limited number of chemicals which can cause effects on fertility or on development at low doses in the standard assays without other adverse effects. These were classified as presumed human reproductive toxicants (category 1B) and subjected to stringent risk management.

As with carcinogenicity, many more chemicals have been found to cause effects on fertility or development in laboratory animals at high doses. These have proved difficult to classify, but by convention they have been categorized as suspected human reproductive toxicants (category 2) or not classified and managed with care, as with chemicals causing adverse effects on other organs or systems.

Harmonization

The introduction of the GHS classification scheme was an opportunity to incorporate the conventions into the classification system and recognize the difference between chemicals based on their ability to cause adverse effects and the dose and duration of exposure required to cause the effects. However, the GHS continued the tradition of putting emphasis on the assumption that there are carcinogens or non-carcinogens, and reproductive toxicants and non-reproductive toxicants and guidance is given on how to distinguish between the two categories.

The revised CLP regulations pose the danger of an erosion of the conventions which were applied to make the distinctions within the categories of carcinogens and reproductive toxicants. Chemicals which previously would have been categorized as suspected human carcinogens or reproductive toxicants (category 2) could now be categorized as presumed (category 1B).

The Resulting Issue

This process leads to a lack of distinction between carcinogens and reproductive toxicants of high potency which require low exposure levels to cause their adverse effects and those of lower potency which require prolonged high exposure to cause cancer or reproductive toxicity. Chemicals which may have up to 5-7 orders of magnitude of difference in their potency (Gold *et al*., 1989; Muller *et al*., 2012) receive the same classification and warning statements. This can cause problems in communicating to the users of chemicals, including the public, the nature of the hazard presented by chemicals.

The original lists of known human carcinogens and reprotoxicants were identified because they are capable of causing the effects in humans at the doses to which people were exposed. In other words they were active at low enough doses for the effects to be seen. This concept of degree of hazard is recognized in the EU guidelines for setting specific concentration limits for carcinogens: 'Classified human carcinogens will generally be of high or medium potency in order to be recognized as such' (EC, 1999). Toxicology is based on the principle that all things are toxic; it is the dose which distinguishes what we consider to be toxic chemicals from those we consider to be non-toxic chemicals.

Although there are large numbers of chemicals capable of inducing tumours or reproductive or developmental toxicity in laboratory animals, the assumption that they will all produce such changes in a proportion of the population at much lower doses is not correct. This is the assumption implicit in classifying all of these chemicals in same way. The exception might be where the mode of action predicts that this is the case, such as in the case of genotoxicity (Purchase and Auton, 1995). For non-genotoxic modes of action the chemical has to be present at a high enough dose to induce an adverse effect which then leads to the carcinogenicity or reproductive or developmental toxicity. Many human studies, both experimental and epidemiological especially in the pharmaceutical area have reinforced the concept of the threshold for toxic effects in humans (Monro, 1993). The animal studies can be taken as an indication of where this threshold lies, using the conservative assumption that the human threshold lies below the animal threshold.

This concept has been incorporated into the classification of other toxic effects. Classification for acute toxicity has always used an estimate of potency to assign a substance to a category. The end point, death, is fixed and the dose required to cause death is determined and then the substance is ascribed to one of four categories on the basis of its acute lethal potency.

The classification system also incorporates potency in the way it deals with other types of toxicity resulting from repeat exposure, the so-called Specific Target Organ Toxicity (ECHA, 2012b). The system recognizes that many chemicals are capable of the hazard of causing damage or adverse effects to specific organs or systems. The distinction between the categories in Specific Target Organ Toxicity is based on the dose level used in the animal studies in which the adverse effects were seen, with the Category 1 being reserved for the chemicals which cause the adverse effects at low doses. The distinguishing dose levels are varied to account for the duration of dosing with lower levels being a cause for concern with longer dosing periods and higher levels with shorter dosing periods.

It is therefore very difficult to see why the same logic should not be applied to carcinogenicity and reproductive and developmental toxicity. The logic says that for a chemical to be a hazard to human health it must have two properties: it must have the capability to cause an adverse effect and it must have the potency to cause the effect at doses to which humans are likely to be exposed.

A Proposed Way Forward

These considerations lead to the conclusion that the terms 'carcinogen' and 'reproductive toxicant' on their own are no

Journal of Applied Toxicology

longer fit for the purpose. They must be qualified. The EU guidance document (ECHA, 2012b) on classification contains the following statement: 'Hazard classification is a process involving identification of the physical, health and environmental hazards of a substance or a mixture, followed by comparison of those hazards (including *degree of hazard*) with defined criteria in order to arrive at a *classification* of the substance or mixture'. It is the degree of hazard which is missing from the term 'carcinogen' or 'reproductive toxiciant'. The process which has been developed for setting specific concentration limits provides a possible answer.

Substances which cause carcinogenicity or reproductive or developmental toxicity in animal studies are subjected to an assessment to produce a value by which they can be compared. In the case of carcinogens (EC, 1999), the dose required to induce an increase in tumours of 25%, the T25 is calculated. For reproductive toxicants (ECHA, 2012c), the dose required to cause an increase of 10% in the relevant effect is calculated, the ED10. The substances are then placed into categories. Those with low T25s or ED_{10}s are considered to have high potency, those with high T25s or ED_{10}s are considered to have low potency, and those in between are considered to have medium potency.

The EU guidance for setting limits for carcinogenicity (EC, 1999) uses the terms:

- Carcinogens with high potency
- Carcinogens with medium potency
- Carcinogens with low potency.

Similarly, these terms could be used:

- Reproductive Toxicant with high potency
- Reproductive Toxicant with medium potency
- Reproductive Toxicant with low potency

Bringing this terminology into wider use is the way forward. This would communicate not only the nature of the hazard but also the degree of hazard. There are well-established and well-thought through methods for assigning substances to these potency groups contained within the regulatory guidelines (EC, 1999; ECHA, 2012c). The methodology described in the guidelines includes a review stage which considers other factors beyond the calculation of the T25 or ED_{10} which include mode of action, kinetics, shape of the dose response curve and the severity of the response. It is especially important to consider these factors when the T25 or ED_{10} is near to the boundary of the potency group. This avoids the boundary becoming a bright line. Adopting this approach more widely in classification and communication would have the following advantages:

- The approach would identify the most hazardous chemicals, but it would also be able to distinguish them from others with lower degrees of hazard.
- The approach would be based on peer reviewed and accepted science which recognizes the increasing knowledge gained since the classification scheme began.
- The approach would also be more objective and transparent in the way it is operated. The process for assessing the data would be clear and be more likely to give rise to the same answer no matter who is doing the assessment. This is important as within REACH the onus is on chemical producers to assess their own chemicals and they should be able to come to the same conclusions as ECHA. This helps consistency, removes uncertainty for business, reduces disputes, allows decisions to be explained and communicated and is much less resource intensive for all.
- The approach would allow appropriate downstream risk management measures to be put in place by being able to distinguish between the degrees of hazard presented by chemicals.

In conclusion, there is an issue in the classification of substances for carcinogenicity and for reproductive or developmental toxicity which has brought difficulties to all those involved in the process. The issue lies in the inability of the classification system to distinguish between carcinogens and reproductive toxicants with different levels of concern. This has its origins in the early years of toxicology when it was thought that a relatively small number of chemicals would be either carcinogens or reproductive toxicants. This can cause problems in communicating to the users of chemicals, including the public, the nature of the hazard presented by chemicals. Processes have been developed within the classification system for setting specific concentration limits which assess the degree of hazard for carcinogens and reproductive toxicants as high, medium or low. However these categories are not otherwise used in classification. Using them would bring the advantages of transparency, clarity of communication, certainty of process and would allow chemicals with a high degree of hazard to be identified and managed in an appropriate way.

Conflict of Interest

Dr John Doe was supported financially by Syngenta Ltd in undertaking this work. Syngenta Ltd had no role in the writing of the report and in the decision to submit the article for publication. The author has sole responsibility for the writing and contents of the paper.

References

Ames BN, Gold LW. 1990. Chemical carcinogenesis: Too many rodent carcinogens? *Proc. Natl. Acad. Sci. U. S. A.* **87**: 7272–7276.

EC. 1999. Guidelines for Setting Specific Concentration Limits for Carcinogens in Annex 1 of Directive 67/548/EEC inclusion of Potency Considerations. Commission Working Group on the Classification and Labelling of Dangerous Substances. European Commission: Belgium.

EC. 2008. Regulation (EC) No 1272/2008 on the Classification, Labelling and Packaging of Substances and Mixtures (CLP). European Commission: Belgium.

ECHA. 2012a. Guidance on the Application of the CLP Criteria. Guidance to Regulation (EC) No 1272/2008 on Classification, Labelling and Packaging (CLP) of Substances and Mixtures. Version 3.0 November 2012. European Chemicals Agency: Finland.

ECHA. 2012b. Guidance on the Application of the CLP Criteria. Guidance to Regulation (EC) No 1272/2008 on Classification, Labelling and Packaging (CLP) of Substances and Mixtures. Version 3.0 November 2012 Section 3.9. Specific Target Organ Toxicity – Repeated Exposure. European Chemicals Agency: Finland.

ECHA. 2012c. Section 3.7.2.5 Setting of Specific Concentration Limits in Guidance on the Application of the CLP Criteria. Guidance to Regulation (EC) No 1272/2008 on Classification, Labelling and Packaging (CLP) of Substances and Mixtures. Version 3.0 November 2012. European Chemicals Agency: Finland.

EEC. 1979. Council Directive 18 September 1979 Amending for the Sixth time Directive 67/548/EEC on the Approximation of Laws, Regulations and Administrative Provisions Relating to the Classification, Packaging and Labelling of Dangerous Substances (79/831/EEC). European Commission: Belgium.

EEC. 1991. Council Directive of 15 July 1991 concerning the Placing of Plant Protection Products on the Market (91/414/EEC). European Commission: Belgium.

GHS. 2007. Globally Harmonised System of Classification and Labelling of Chemicals (GHS) Second Revised Version, United Nations. ST/SG/AC.10/30/Rev.2

Gold L, Slone T, Bernstein L. 1989. Summary of carcinogenic potency and positivity for 492 rodent carcinogens in the carcinogenic potency database. *Environ. Health Persp.* **79**: 259–272.

Harvey PW, Everett DJ, Springall CJ. 2006. Hyperprolactinaemia as an adverse effect in regulatory and clinical toxicology: role in breast and prostate cancer. *Hum. Exp. Toxicol.* **25**: 395–404

Hill AB. 1965. The Environment and Disease: Association or Causation. *Proc. Royal Soc. Med.* **58**: 295–300.

Leeuwen CJ van, Vermieire TG. 2007. *Risk Assessment of Chemicals – An Introduction* (2nd edn). Springer: The Netherlands.

Monro A. 1993. How useful are chronic (life-span) toxicology studies in rodents in identifying pharmaceuticals that pose a carcinogenic risk to humans? *Adverse Drug React. Toxicol. Rev.* **12**: 5–34.

Muller A, Blaude M-N, Ihlemann C, Bjorge C, Ohlsson A, Gebel T. 2012. A regulatory approach to assess the potency of substances toxic to the reproduction. *Reg. Toxicol. Pharmacol.* **63**: 97–105.

Purchase IHF, Auton TR. 1995. Thresholds in chemical carcinogenesis. *Reg. Toxicol. Pharmacol.* **22**: 199–205.

Slovic P. 1987. Perception of risk. *Science* **236**: 280–255.

Swenberg JA. 1993. Alpha 2u-globulin nephropathy: review of the cellular and molecular mechanisms involved and their implications for human risk assessment. *Environ. Health Perspec.* **101**: 39–44.