

Are Animal Tests Inherently Valid?

Michael Balls

FRAME, Russell & Burch House, 96–98 North Sherwood Street, Nottingham NG1 4EE, UK
E-mail: michael.balls@frame.org.uk

Summary — The proposition that animal tests are inherently valid, merely because they are animal tests, is discussed and is rejected. It is concluded that there is no justifiable reason for subjecting new or substantially modified animal test procedures or testing strategies to a validation process that is any less stringent than that applied to non-animal tests and testing strategies.

Key words: *animal test, regulatory acceptance, replacement alternative test, validation.*

Introduction

This question arises from a discussion session during the OECD *Conference on Validation and Regulatory Acceptance of New and Updated Methods in Hazard Assessment*, which took place in Stockholm on 6–8 March 2002, when a participant from North America said that animal tests did not need to be subjected to the same rigorous validation process as non-animal tests, since, being animals tests, they are inherently valid. It would be true to say that this proposition raised the eyebrows of most, if not all, of the other participants, and there is no suggestion that such a position is held by the OECD. Nevertheless, that even one person believes it, justifies its discussion at this Congress. Unfortunately, however, nobody could be found to defend it in what was supposed to be a point-counterpoint debate.

Before the question can be examined, some other questions need to be answered, relating to the use of terminology in this context.

What is an animal?

An animal is a “living vertebrate” used in the laboratory as a model for humans or other animals, and, in particular, a rat, mouse, guinea-pig, rabbit, dog or non-human primate.

What is a test?

A test is composed of a “test system”, with a protocol defining an endpoint, an endpoint assay, an exposure regimen, and a procedure for calculating and expressing a result, together with a “prediction model” for converting the result into a prediction relevant to an *in vivo* toxicological hazard.

What does “valid” mean?

Valid means scientifically established as “relevant and reliable for a particular purpose”.

What does “inherent” mean?

Inherent is used to convey such meanings as “intrinsic, innate, inbuilt, essential, natural,

real, genuine” and “existing as an inseparable part of”.

What is validation?

Validation is a process for independently evaluating the validity of a test, which follows “test development” and precedes consideration for “regulatory acceptance” and application.

How is validation carried out?

“Prevalidation” is widely accepted as a means of optimising a test protocol and maximising its interlaboratory transferability, before a formal validation study is undertaken. A “formal validation study” is characterised by management, test item selection, coding and distribution, and data collection and analysis, which are independent of the laboratories conducting the test. Where such a study is not possible (e.g. because of a lack of test items backed by sufficient knowledge of sufficiently high quality), or is not considered necessary (e.g. because of long experience in the use of a test), “a weight-of-evidence” approach may be appropriate.

Why is validation necessary?

Primarily because of the importance of the decisions made when using hazard predictions in risk assessment, and also for scientific, commercial, economic and administrative reasons.

The Validation Process and Animal Tests

There could be a number of reasons for the general assumption that animal tests could be considered relevant and reliable for predicting hazard to humans. For example:

- They are very similar to humans.
- They possess all the functioning, integrated and interacting body systems.

- They can provide unexpected information concerning unknown hazards.
- Toxicologists understand their limitations.
- Regulators understand their use in risk assessment.
- Their formal validation has not been considered necessary in the past.
- There is long experience with their use in practice.

However, these reasons are not a sufficient basis for assuming inherent validity. The truth is that, while animal tests have been, and, for the foreseeable future, will continue to be, useful in certain circumstances, they suffer from a number of major and insuperable disadvantages, and, in the medium-term to long-term, they will inevitably be replaced by more-modern, more scientifically sound, independently validated, non-animal tests and testing strategies.

The biggest problem with animal tests is that of species differences of a general nature and involving specific responses at the molecular, cell and tissue levels. This was brilliantly summed up by Russell & Burch in *The Principles of Humane Experimental Technique* (1), in their discussion on models and what they called the “high-fidelity fallacy”. Recognising that models, by definition, cannot be precisely the same as what is being modelled, they concluded there are two main factors governing the ways in which a model differs from the original: “fidelity” means “overall similarity”, and “discrimination” means the extent to which a model reproduces “one particular property” of the original.

The ideal model provides high fidelity and high discrimination, but a low fidelity/high discrimination model is more useful in answering specific questions than is a high fidelity/low discrimination model.

A bacterial mutagenicity test is an example of a low fidelity/high discrimination model, since the mechanistic basis of damage to DNA is considered to be sufficiently similar in bacterial cells and human cells for it to be useful. However, in the case of peroxisome proliferation, the rat is an example of a high fidelity/low discrimination model, since, despite many general similarities between rats and humans, the kind of peroxisome formation that occurs in some rat cells does not occur in human cells.

The “high fidelity fallacy” refers to the dangerous assumption that, because laboratory mammals are, in general, similar to ourselves, the data tests on them will, therefore, be relevant and reliable for use in human risk assessment.

The Answer to the Question

The real question before us is: “Are animal tests inherently valid, merely because they are animal tests?”. The logical and scientific answer is clear — it is: *No, they certainly are not!*

Test Development and Validation

One of the lessons learned from validation studies conducted on *in vitro* tests at the beginning of the 1990s was that tests must be properly developed before they can be considered suitable for inclusion in a formal, interlaboratory validation study. As a result, the European Centre for the Validation of Alternative Methods (ECVAM) established a set of criteria for test development and asks proponents of new test methods to provide evidence that the following information can be provided (2):

- A definition of the scientific purpose of the method, and of its proposed practical application.
- A description of the basis of the method.
- The case for its relevance.
- An explanation of the need for the method in relation to existing *in vivo* methods and other non-animal methods.
- An optimised protocol, including: any necessary standard operation procedures; a specification of endpoints and endpoint measurements; the method for deriving and expressing results; the interpretation of the results in terms of one or more *in vivo* pharmacotoxicological endpoints, by means of a prediction model; and the use of adequate controls.
- A statement about the limitations of the test.
- Evidence of intralaboratory reproducibility, and if available, interlaboratory transferability.

Similarly, if a formal validation study is conducted, an independent evaluation of the outcome is conducted before a scientifically validated method is put forward for consideration for regulatory acceptance for a particular application. Again, ECVAM has established a set of questions that need to be considered when such an evaluation is conducted (2):

- Clarity of defined goals.
- Quality of overall design.
- Independence of management.

- Independence of selection, coding and distribution of test materials.
- Independence of data collection and analysis.
- Number and properties of test materials.
- Quality and interpretation of results.
- Performance of the method(s) in relation to the predetermined goals of the study.
- Compliance with the principles of Good Laboratory Practice (GLP) and Good Cell Culture Practice (GCCP).
- Reporting of outcome in the peer-review literature.
- Availability of raw data.

These two sets of criteria are now being applied to establish the relevance and reliability of non-animal test methods for particular purposes, to the satisfaction of the regulatory authorities.

This leads to the next question: “If new animal tests cannot be considered to be inherently valid, should they be subjected to validation and evaluation, according to similar criteria?”

The answer must be an emphatic *yes*, since the purpose of the validation process is to independently establish the relevance and reliability of a method for its particular purpose, bearing in mind the importance of the decisions that will be based on the information it provides.

This inevitably leads to another question: “How can *anybody* advocate the acceptance and application of test methods in regulatory toxicology that have not been independently shown to be scientifically valid?”

Are Dual Standards being Applied to the Acceptance of Animal Tests and Non-animal Tests?

There is a wealth of anecdotal evidence that this is true, and it is no secret that regulators “feel more comfortable” with animal test data than with data from non-animal tests (3).

There is also a widespread feeling that the US Environmental Protection Agency considers that the application of the formal validation process, as accepted internationally as a result of discussions between ECVAM, the (US) Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the OECD, would delay the introduction of its multi-test strategy for testing for “endocrine disruptors”.

Also, the National Coordinators for the OECD Test Guidelines Programme have recently

accepted draft guidelines for an *in vivo* skin absorption test (TG 427) and for *in vitro* methods for skin absorption (TG 428) — but have the same evaluation criteria been applied to these two draft guidelines?

A more concrete case is the OECD validation study on the uterotrophic assay for “endocrine disruption” (4), which, however well intentioned, does not meet the ECVAM/ICCVAM/OECD (Solna 1; 5) criteria for test validation and acceptance, for the following reasons:

- The relevance of the method was not convincingly established before the study began.
- No test protocols were agreed or optimised in a prevalidation phase, before the study began.
- No prediction model was provided before the study began.
- No performance criteria were agreed before the study began.
- An insufficient range of test items was used, especially with respect to negative controls.

Therefore, this has to be seen as a test development exercise, which was unnecessarily expensive in terms of human and economic resources, as well as animal lives.

Conclusions

There is no justifiable reason for subjecting new or substantially modified animal tests and testing strategies to a validation process and to the application of test development, validation and acceptance criteria and standards that are any less stringent than those applied to new or substantially modified non-animal tests and testing strategies.

No test should be accepted for regulatory use and application until it has been independently established as relevant and reliable for its intended purpose.

Note Added in Proof

This important topic is further considered in: Combes, R.D. & Balls, M. (2003). How much flexibility is possible when validating new *in vivo* and *in vitro* toxicity test methods? *ATLA* **31**, 225–232.

References

1. Russell, W.M.S. & Burch, R.L. (1959). *The Principles of Humane Experimental Technique*, 238pp. London, UK: Methuen.

2. Worth, A. & Balls, M. (2002). The principles of validation and the ECVAM validation process. *ATLA* **30**, Suppl. 2, 15–21.
3. O'Connor, A.M. (1997). Barriers to regulatory acceptance. In *Animal Alternatives, Welfare and Ethics* (ed. L.F.M. van Zutphen & M. Balls), pp. 1173–1176. Amsterdam, The Netherlands: Elsevier Science B.V.
4. Combes, R.D. (2000). Endocrine disruptors: a critical review of *in vitro* and *in vivo* testing strategies for assessing their toxic hazard to humans. *ATLA* **28**, 81–118.
5. OECD (1996). *Final Report of the OECD Workshop on Harmonisation of Validation and Acceptance Criteria for Alternative Toxicological Test Methods*, 60pp. Paris, France: OECD.