Comment

Validating Animal Models for Preclinical Research: A Scientific and Ethical Discussion

Orsolya E. Varga,[1,2] Axel K. Hansen,[3] Peter Sandøe[2] and I. Anna S. Olsson[1,2]

[1]Laboratory Animal Science, Institute for Molecular and Cell Biology (IBMC), University of Porto, Porto, Portugal; [2]Danish Centre for Bioethics and Risk Assessment, Faculty of Life Sciences, University of Copenhagen, Copenhagen, Denmark; [3]Department of Veterinary Disease Biology, Faculty of Life Sciences, University of Copenhagen, Copenhagen, Denmark

Summary — The use of animals to model humans in biomedical research relies on the notion that basic processes are sufficiently similar across species to allow extrapolation. Animal model validity is discussed in terms of the similarity between the model and the human condition it is intended to model, but no formal validation of models is applied. There is a stark contrast here with the use of non-animal alternatives in toxicology and safety studies, for which an extensive validation is required. We discuss both the potential and the limitations of validating preclinical animal models for proof-of-concept studies, by using an approach similar to that applied to alternative non-animal methods in toxicology and safety testing. A major challenge in devising a validation system for animal models is the lack of a clear gold standard with which to compare results. While a complete adoption of the validation approach for alternative methods is probably inappropriate for research animal models, key features, such as making data available for external validation and defining a strategy to run experiments in a way that permits meaningful retrospective analysis, remain highly relevant.

Key words: *animal models, ethics, predictive validity, validation.*

Address for correspondence: *Orsolya E. Varga, Laboratory Animal Science, Institute for Molecular and Cell Biology (IBMC), University of Porto, Rua do Campo Alegre 823, 4150-180 Porto, Portugal.*
E-mail: *ovarga@ibmc.up.pt*

Introduction

The use of animals to model humans in biomedical research relies on the notion that basic processes are sufficiently similar across species to allow extrapolation. We discuss both the potential and the limitations of validating preclinical animal models for proof-of-concept studies, by using an approach similar to that applied to alternative non-animal methods in toxicology and safety testing.

While studies which use animal models are an important part of biomedical research, the translation of results into treatments for human beings is far from straightforward (1). Both economic and ethical issues come into play, when a potential therapy fails first-in-human or later trials (2). Better (use of) animal models is one way of reducing the high attrition rate (3).

Animal model validity is discussed in terms of the similarity between the model and the human condition it is intended to model, but no formal validation of models is applied. There is a stark contrast here to the use of non-animal alternatives in toxicology and safety studies, for which an extensive validation is required.

Animal Models and Validity

Roughly speaking, the present approach to model development is based on similarities in the symptoms and/or aetiology of a disease in humans and animals. An animal model is considered to be valid, if it "resembles the human condition in aetiology, pathophysiology, symptomatology and response to therapeutic interventions" (4). Usually, this general validity is broken down into three aspects (5): predictive validity (performance in the test predicts performance in the modelled condition); face validity (phenomenological analogy with the modelled condition); and construct validity (the model has a sound theoretical rationale).

Over the last few years, several initiatives have been launched to encourage the use of more-accurate animal models in both industrial and aca-

demic research. European and US authorities have published guidelines which identify the key characteristics of an approved animal model (6, 7) and list criteria which, if met, demonstrate a model's suitability (cross-species comparison, taking into account the target, its structural homology, distribution, signal transduction pathways and the nature of pharmacological effects). These guidelines are intended for use by those seeking approval or a licence for drugs or biological products. Several voluntary initiatives from researchers and industry point in the same direction, including the STRAIT initiative for more-sophisticated, consensus-based validity criteria governing preclinical animal studies of stroke (8) and the ongoing MATRICS, TURNS and CNTRICS programmes to improve research into therapies for schizophrenia (9). Essentially, these initiatives promote a more-sophisticated way of delivering construct validity and face validity. However, when the results of an animal study are intended to be translated into human treatments (preclinical research), the ultimate proof of a model's value is its predictive validity.

While face and construct validity are primarily theoretical considerations, predictive validity involves the calculation of a number of statistical parameters in a validation process. In a simple case, predictive validity can be calculated in terms of reliability and relevance. Reliability is assessed by calculating inter-laboratory reproducibility and intra-laboratory repeatability. Relevance shows whether a model is meaningful and useful for a particular purpose, and the extent to which the model accurately measures or predicts the biological effect of interest (sensitivity and specificity; 10).

The Process of Validation — The Alternative Methods Approach

The predictive validity of an animal model can be tested by systematic examination of the data from animal model studies, and by comparing these data with reference data obtained in humans. One way of doing this would be to follow the validation process for alternative methods. The process described here is used by the European Centre for the Validation of Alternative Methods (ECVAM; 11); a similar system has been adopted by the Organisation for Economic Co-operation and Development (OECD) and North American organisations, which have harmonised their validation processes (12).

This process has five basic steps (10, 13). The first is test development. The fifth is formal regulatory acceptance. Actual validation, in the sense of generating, analysing and assessing data, takes place in steps two, three and four, outlined below:

— Step 2, Pre-validation: An inter-laboratory pre-validation study is conducted to optimise the protocol and assess its performance over three phases: phase I, where the protocol is refined in a single laboratory; phase II, where the transferability of the method to a second laboratory is assessed; and phase III, where the relevance and reliability of the test are assessed under blind conditions in two or more laboratories.

— Step 3, Validation: The formal validation study can be thought of as an extended version of the phase III stage of pre-validation, in which an inter-laboratory blind trial (involving at least three laboratories) is conducted to assess whether tests can be shown to be relevant and reliable for one or more specific purposes. This inter-laboratory trial is followed by data analysis and an evaluation of the outcome of the study in comparison with predefined performance criteria.

— Step 4, Independent Assessment: Validation study results are published in peer-review journals and considered by independent assessment panels working under the auspices of appropriate national or international organisations. The panel review of the data and peer review recommendations are published.

The validation process, from test development to regulatory acceptance, need not be unidirectional; retrospective data analysis is also common. This helps to reduce both economic and ethical costs — the repetition of animal research or human clinical trials is obviously wasteful, when the necessary data are already available. On the other hand, retrospective data is often less reliable, and its interpretation can be challenging (11, 14). Therefore, the prospective approach is usually preferred.

Could the Alternative Methods Approach be Used to Validate Animal Models?

Validation has two principal aspects: a) how well a test method compares with itself when repeated under identical conditions, as well as under different conditions (e.g. with different test substances and in different laboratories); and b) how well a test method compares with a reference method. These two aspects present somewhat different challenges in terms of the data required, but there is no theoretical obstacle to their application to animal models in biomedical research.

How well an animal model compares with itself under different conditions can be evaluated by using animal data alone. The evaluation requires data to be available from the use of the same

model, ideally from several replicates performed both under identical conditions (to estimate repeatability [15]) and under controlled conditions, where one factor is varied while others are kept constant (to estimate reproducibility [15]).

Evaluating how well an animal model compares with reference data is more challenging. This is a practical challenge, because it requires data from humans and is thus only possible when a substance, or other type of therapy, has advanced through preclinical stages to human trials. An even more fundamental challenge is presented by the difference between the repetitive nature of testing and the innovative nature of research. When non-animal alternatives in safety testing are validated, there is a clear gold standard in the form of the animal test to be replaced (although it should be remembered that this gold standard is only a proxy measure of the real parameter of interest — the human reaction to a substance — against which it has, in fact, never been validated). In proof-of-concept studies in research, there is no gold standard. Depending on the intended target of drug action, different types of research approach require different models, and a model with proven predictive validity for a particular compound may not, in fact, be sensitive to the effects of a different type of compound that acts on different targets (9). Efforts to validate against a standard in the form of a proven successful treatment may give rise to a system that will only detect 'me-too' treatments, i.e. those based on the same principle of action (16), and hence may unduly restrict necessary innovation. This does not mean that the analysis of the correspondence of results of animal and human experiments is impossible or of no value. Indeed, it is precisely this type of retrospective analysis, which, in recent studies, has helped to identify inconsistencies in animal and human studies (e.g. in dosage, administration method, parameters, and method of assessing effect) that are likely, at least in part, to underlie the poor translation of results.

The validation of animal models potentially carries monetary as well as ethical costs. Validation is time consuming (2–6 years for the alternative methods) and costly, and financial returns may be more difficult to secure, since intellectual property rights over animal models are more restricted than they are for alternative methods. Ethical concerns may also arise over the use of animals for the sole purpose of validation. However, validation that is based on the re-analysis of existing data may partly overcome these concerns, and if validation results in more-effective research, both animal numbers and costs may be offset by savings in subsequent research. Thus, we argue, there is reason to consider partial adoption of the validation procedure.

Conclusions and Suggestions

Over the last few years, a number of recommendations and guidelines have been published, to encourage the more-accurate use of animal models (6, 7). Against that background, what benefits would accompany the application of the alternative methods approach to the validation of animal models? We identify two key gains: retrospection and publication.

Guidelines for better animal experiments take a primarily prospective view, but, if lessons are to be learned from previous mistakes, retrospective analysis and the re-assessment of data are vital. A recurring obstacle here is the difficulty of accessing an unbiased and complete dataset. Data from many experiments simply do not enter the public domain, either because the results are negative and therefore difficult to publish (publication bias), or because they are compiled in pharmaceutical companies and only, if at all, presented to authorities for drug approval.

The type of prospective validation favoured for non-animal alternatives is ethically problematic, when living subjects — animals or humans — are involved. Therefore, the challenge is to produce a system in which data are made available for external validation, and to define strategies for running experiments which will permit more-meaningful retrospective analyses. Within the validation system for alternatives, there is unique experience in dealing with this in a systematic way. Making these analyses available in a peer-review journal — the fourth step in the alternative methods validation procedure — is also crucial, if knowledge is to be disseminated to the wider scientific community.

Successfully learning from experience also means being able to accept new data that challenge old paradigms. Old models and methods must be abandoned, or suitably revised, if the systematic analysis of replicability, repeatability and correspondence with reference data, indicate that their performance is not up to standard.

The validation of animal models and tests is a shared responsibility, in which academic research, the pharmaceutical industry, regulatory authorities and ethics committees/IACUCs all have a part to play. That this issue is taken seriously and validation is integrated into the research process are both scientific and ethical imperatives. As scientists, we need to reassure those who have concerns about animal use in research that we are using animals in the best possible way to make progress in the treatment of human diseases. Validation can underwrite that reassurance. Specifically, where an animal model results in the progress of a drug from preclinical to first-in-human trials, when in fact better preclinical trials could have prevented that progression, animals are used

needlessly, economic resources are wasted, and human volunteers are exposed to risks — all to no avail. Conversely, the abandonment of a drug development programme, where the drug would have proven effective in clinical trials, is not only a waste of resources, but also a loss for patients.

Acknowledgement

The research for this article was funded by The Danish Council for Strategic Research through its Food and Health Programme, NUTRIOMICS — functional foods for cloned, lean/obese pigs project.

References

1. Anders, H.J. & Vielhauer, V. (2007). Identifying and validating novel targets with *in vivo* disease models: Guidelines for study design. *Drug Discovery Today* **12**, 446–451.
2. US FDA (2004). *Innovation or Stagnation? Challenge and Opportunity on the Critical Path to New Medical Products*, 38pp. Silver Spring, MD, USA: US Department of Health and Human Services, Food and Drug Administration. Available at: http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm113411.pdf (Accessed 12.04.10).
3. Bhogal, N. & Balls, M. (2008). Translation of new technologies: from basic research to drug discovery and development. *Current Drug Discovery Technologies* **5**, 250–262.
4. Van Dam, D. & De Deyn, P.P. (2006). Drug discovery in dementia: The role of rodent models. *Nature Reviews Drug Discovery* **5**, 956–970.
5. van der Staay, F.J. (2006). Animal models of behavioral dysfunctions: Basic concepts and classifications, and an evaluation strategy. *Brain Research Reviews* **52**, 131–159.
6. US FDA (2009). *Guidance for Industry, Animal Models — Essential Elements to Address Efficacy Under the Animal Rule*, 19pp. Silver Spring, MD, USA: US Department of Health and Human Services, Food and Drug Administration. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM078923.pdf (Accessed 12.04.10).
7. Committee for Medicinal Products for Human Use (CHMP) (2007). *Guideline on Strategies to Identify and Mitigate Risks for First-in-human Clinical Trials with Investigational Medicinal Products* EMEA/CHMP/SWP/28367/07, 12pp. London, UK: European Medicines Agency. Available at: http://www.ema.europa.eu/pdfs/human/swp/2836707enfin.pdf (Accessed 12.04.10).
8. Regenberg, A., Mathews, D.J., Blass, D.M., Bok, H., Coyle, J.T., Duggan, P., Faden, R., Finkel, J., Gearhart, J.D., Hillis, A., Hoke, A., Johnson, R., Johnston, M., Kahn, J., Kerr, D., King, P., Kurtzberg, J., Liao, S.M., McDonald, J.W., McKhann, G., Nelson, K.B., Rao, M., Siegel, A.W., Smith, K., Solter, D., Song, H., Sugarman, J., Vescovi, A., Young, W., Greely, H.T. & Traystman, R.J. (2009). The role of animal models in evaluating reasonable safety and efficacy for human trials of cell-based interventions for neurologic conditions. *Journal of Cerebral Blood Flow & Metabolism* **29**, 1–9.
9. Markou, A., Chiamulera, C., Geyer, M.A., Tricklebank, M. & Steckler, T. (2009). Removing obstacles in neuroscience drug discovery: The future path for animal models. *Neuropsychopharmacology* **34**, 74–89.
10. Worth, A.P. & Balls, M. (2002). The principles of validation and the ECVAM validation process. *ATLA* **30**, Suppl. 2, 15–21.
11. Hoffmann, S., Edler, L., Gardner, I., Gribaldo, L., Hartung, T., Klein, C., Liebsch, M., Sauerland, S., Schechtman, L., Stammati, A. & Nikolaidis, E. (2008). Points of reference in the validation process: The report and recommendations of ECVAM Workshop 66. *ATLA* **36**, 343–352.
12. Balls, M., Amcoff, P., Bremer, S., Casati, S., Coecke, S., Clothier, R., Combes, R., Corvi, R., Curren, R., Eskes, C., Fentem, J., Gribaldo, L., Halder, M., Hartung, T., Hoffmann, S., Schechtman, L., Scott, L., Spielmann, H., Stokes, W., Tice, R., Wagner, D. & Zuang, V. (2006). The principles of weight of evidence validation of test methods and testing strategies. The report and recommendations of ECVAM workshop 58. *ATLA* **34**, 603–620.
13. Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell, D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H. & Zucco, F. (1995). Practical aspects of the validation of toxicity test procedures: The report and recommendations of ECVAM Workshop 5. *ATLA* **23**, 129–147.
14. Hartung, T., Bremer, S., Casati, S., Coecke, S., Corvi, R., Fortaner, S., Gribaldo, L., Halder, M., Hoffmann, S., Roi, A.J., Prieto, P., Sabbioni, E., Scott, L., Worth, A. & Zuang, V. (2004). A modular approach to the ECVAM principles on test validity. *ATLA* **32**, 467–472.
15. Barnhart, H.X., Haber, M.J. & Lin, L.I. (2007). An overview on assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics* **17**, 529–569.
16. Miczek, K.A. (2008). Challenges for translational psychopharmacology research: The need for conceptual principles. In *Animal and Translational Models for CNS Drug Discovery: Reward Deficit Disorders* (ed. F. Borsini & R. McArthur), pp. 359–375. Amsterdam, The Netherlands: Academic Press.