# OPINION

## Likelihood Ratios in Assessing the Safety of New Medicines

**Robert A. Coleman**

*Until we know the true predictive value of animal-based methods for predicting clinical safety issues, it is impossible to assess the advantage or otherwise of non-animal based approaches*

The use of animals in the discovery and development of new medicines has generated debate for decades. For much of this time, contrasting views have been primarily polarised on the basis of practicality *versus* ethics, with proponents arguing that for the development of new medicines to treat human disease, the end justifies the means, while the opponents' chief objection has been the associated animal suffering. Commercial and public health pressures over the years have ensured that those in the 'practicality camp' have held sway.

More recently, however, the issue has become increasingly complex, with growing concerns that, irrespective of ethical considerations, data generated in animal (i.e. non-human) models are not necessarily or sufficiently relevant to human patients.[1-3] There is now general consensus that inter-species variability is a real issue, and that animal models are far from perfect for the purpose of ensuring either the efficacy or the safety of potential new medicines intended for human subjects. Supporters of the continued use of animals argue that, while they do not provide an absolute indication of either efficacy or safety, in the absence of any other approach, one that is somewhat unreliable is better than none at all.

Such an argument has some merit, if indeed it is valid. However, in this field, all may not be entirely as it seems. Firstly, human-based *in vitro* and *in silico* alternatives are becoming ever more sophisticated,[4] thus overcoming many of the criticisms originally directed toward them. For example, it has long been held that it would be impossible to model the complexity of the intact patient through a study of isolated cells and tissues, and while this problem may never be wholly overcome, the gap gets ever smaller. Secondly, it is important to understand that we really don't know how good existing animal-based methods are. In the field of efficacy, there is a wealth of evidence that results obtained by using experimental animals can be hugely misleading.[5-11] Here, we have the advantage that drugs that promise efficacy in patients on the basis of animal data can advance into clinical testing, and their utility can be directly assessed. For the majority of these drugs, the clinical outcome has been disappointing. With safety, the issue is different, as drugs with identified safety issues in animals will seldom, if ever, advance to clinical testing, thus the relevance of the animal data to safety in humans may never be determined. However, what we do know is that many drugs identified as safe in pre-clinical profiling eventually prove to cause serious and use-limiting side effects in human subjects.[12] The key question is, "Could such failures have been avoided, had we relied on human-based test methods?" Until we know how frequently non-animal methods could have identified safety issues that were missed by animal tests, it is impossible to assess the advantage or otherwise of those methods. It is a fact that, despite the continued use of animals as human surrogates in pharmaceutical research, there has never been a solid, published, peer-reviewed study demonstrating fitness for purpose, whereas reviews identifying the shortcomings are abundant.

## Assessing the value of animal studies

It is for this reason that any information that sheds light on the actual value of animal-based testing for its intended purpose is of inestimable worth. Until recently, much evidence, while valuable, has been indirect. For example, a recent telling study demonstrated that pre-clinical fast-tracking (i.e. abbreviated safety testing) of potential new medicines resulted in no increase in the proportion of candidates that subsequently proved toxic in human subjects.[13] In another study, the ability of animal studies to detect serious post-marketing adverse events was demonstrably poor.[14] While such reports add to the

volume of data providing witness to the shortcomings of animal-based approaches to ensuring clinical safety, they do not provide a robustly measurable metric of predictive efficiency. In view of the colossal amount of data generated over the years in pre-clinical safety studies on thousands of new potential medicines, many of which have progressed to clinical testing and even to market, it is amazing that, until recently, no comprehensive analysis of such data has been applied in order to explore the value of the current approach to safety testing.

## Likelihood ratios

In the light of such a background, it is of considerable significance that serious attempts are now being made to extract intelligence from the wealth of information available in publicly accessible sources, in order to shed more light on the actual predictive power of animal-based safety testing. A particular example is the utilisation of the Safety Intelligence Programme (SIP),[15] which overcomes semantic issues to extract valid information from all available data sources. SIP has been used, for example, to explore the predictive power of animal models for the detection of liver toxicity associated with a wide range of human medicines,[16] highlighting the highly variable efficiency of different models in combination with different drug toxicities. More recently, SIP has been used to particular effect in two studies that have explored directly the value of dogs, mice, rats and rabbits in predicting safety issues in human subjects.[17,18] While most previous studies have relied on determining 'concordance' between animal and human data, that tells only a part of the story, and is too simplistic a measure to be of much real value. Its problem is that it only deals with positive correlation, i.e. the frequency that toxicity in experimental animals and in human subjects coincide, ignoring the issue of true prediction. What is needed is a determination of likelihood ratios (LRs),[19] both positive (PLRs) and negative (NLRs), to gain a more complete picture. What emerged when LRs were determined was that, although there was indeed some measure of concordance between positive toxicity data between animals and humans, in terms of LRs, none of the species proved to offer any useful level of real predictive power. Although the studies and their conclusions did not escape criticism from some quarters,[20] the suggested limitations, real or perceived, are arguably irrelevant to its overall validity.[21]

What did emerge from the application of this approach were absolute values for both PLRs and NLRs for a wide range of specific drugs. The importance of this is that, for the first time, such measures can provide a robust yardstick against which to evaluate the relative merits of alternative approaches to toxicity testing.

## Conclusions

To summarise, the use of more-rigorous approaches to the evaluation of animal models as predictors of the likelihood that any chemical will be similarly toxic *or non-toxic* in human subjects provides not only a realistic measure of their actual fitness for purpose, but crucially, also a basis by which the efficiency of other, ideally human-based, approaches can be evaluated through their exposure to the same range of drugs. The use of the same drugs ensures that any criticisms related to potential bias, or other potentially confounding factors, are negated. Such a prospective study would be of inestimable value. Dare we hope that the government and pharmaceutical companies will take up the challenge and fund such a study?

*Dr Robert A. Coleman*
*Independent Consultant*
*UK*
*E-mail: robt.coleman@btinternet.com*

## References

1 Wall, R.J. & Shani, M. (2008). Are animal models as good as we think? *Theriogenology* **69**, 2–9.
2 Hartung, T. (2013). Food for Thought… Look back in anger — What clinical studies tell us about preclinical work. *ALTEX* **30**, 275–291.
3 Pippin, J.J. & Sullivan, K. (undated). *Dangerous medicine: Examples of animal-based "safety" tests gone wrong*. Washington, DC, USA: The Physicians Committee for Responsible Medicine. Available at: http://www.pcrm.org/research/animaltestalt/animaltesting/dangerous-medicine-examples-of-animal-based-tests (Accessed 03.11.14).
4 Ashton, R., De Wever, B., Fuchs, H.W., Gaça, M., Hill, E., Krul, C., Poth, A. & Roggen, E.L. (2014). State of the art on alternative methods to animal testing from an industrial point of view: Ready for regulation? *ALTEX* **31**, 357–363.
5 Pound, P., Ebrahim, S., Sandercock, P., Bracken, M.B. & Roberts, I. (2004). Where is the evidence that animal research benefits humans? *British Medical Journal* **328**, 514-517.
6 Kaste, M. (2005). Use of animal models has not contributed to development of acute stroke therapies: Pro. *Stroke* **36**, 2323–2324.
7 Pippin, J.J. (2005). *The Need for Revision of Pre-Market Testing: The Failure of Animal Tests of COX-2 Inhibitors*, 23pp. Washington, DC, USA: The Physicians Committee for Responsible Medicine. Available at: pcrm.org/pdfs/research/testing/exp/COX2Report.pdf (Accessed 04.11.14).
8 Hackam, D.G. & Redelmeier, D.A. (2006). Translation of research evidence from animals to humans. *Journal of the American Medical Association* **296**, 1731–1732.
9 Knight, A. (2008). Systematic reviews of animal experiments demonstrate poor contributions towards human healthcare. *Reviews on Recent Clinical Trials* **3**, 89–96.

[10] Matthews, R.A.J. (2008). Medical progress depends on animal models — doesn't it? *Journal of the Royal Society of Medicine* **101**, 95–98.

[11] Seok, J., Warren, H.S., Cuenca, A.G., Mindrinos, M.N., Baker, H.V., Xu, W., Richards, D.R., McDonald-Smith, G.P., Gao, H., Hennessy, L., Finnerty, C.C., López, C.M., Honari, S., Moore, E.E., Minei, J.P., Cuschieri, J., Bankey, P.E., Johnson, J.L., Sperry, J., Nathens, A.B., Billiar, T.R., West, M.A., Jeschke, M.G., Klein, M.B., Gamelli, R.L., Gibran, N.S., Brownstein, B.H., Miller-Graziano, C., Calvano, S.E., Mason, P.H., Cobb, J.P., Rahme, L.G., Lowry, S.F., Maier, R.V., Moldawer, L.L., Herndon, D.N., Davis, R.W., Xiao, W., Tompkins, R.G. & Inflammation and Host Response to Injury, Large Scale Collaborative Research Program (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences of the USA* **110**, 3507–3512.

[12] Li, A.P. (2004). Accurate prediction of human drug toxicity: A major challenge in drug development. *Chemico-biological Interactions* **150**, 3–7.

[13] Arnardottir, A.H., Haaijer-Ruskamp, F.M., Straus, S.M., Eichler, H.G., de Graeff, P.A. & Mol, P.G. (2011). Additional safety risk to exceptionally approved drugs in Europe? *British Journal of Clinical Pharmacology* **72**, 490–499.

[14] van Meer, P.J., Kooijman, M., Gispen-de Wied, C.C., Moors, E.H. & Schellekens, H. (2012). The ability of animal studies to detect serious post marketing adverse events is limited. *Regulatory Toxicology & Pharmacology* **64**, 345–349.

[15] Berry, S. (2012). *Safety Intelligence Program Provides Insight into Drug-Induced Cardiac Effects*. Stone, UK: Instem plc. Available at: http://www.biowisdom.com/content/safety-intelligence-program (Accessed 04.11.14).

[16] Fourches, D., Barnes, J.C., Day, N.C., Bradley, P., Reed, J.Z. & Tropsha, A. (2010). Cheminformatics analysis of assertions mined from literature that describe drug-induced liver injury in different species. *Chemical Research in Toxicology* **23**, 171–183.

[17] Bailey, J., Thew, M. & Balls, M. (2013). An analysis of the use of dogs in predicting human toxicology and drug safety. *ATLA* **41**, 335–350.

[18] Bailey, J., Thew, M. & Balls, M. (2014). An analysis of the use of animal models in predicting human toxicology and drug safety. *ATLA* **42**, 181–199.

[19] Altman, D.G. & Bland, J.M. (1994). Diagnostic tests 2: Predictive values. *British Medical Journal* **309**, 102.

[20] Brooker, P. (2014). The use of second species in toxicology testing. *ATLA* **42**, 147–149.

[21] Bailey, J. (2014). A response to the ABPI's *Letter to the Editor* on the use of dogs in predicting drug toxicity in humans. *ATLA* **42**, 149–153.