An Analysis of the Use of Dogs in Predicting Human Toxicology and Drug Safety

Jarrod Bailey,[1] Michelle Thew[1] and Michael Balls[2]

[1]British Union for the Abolition of Vivisection (BUAV), London, UK; [2]c/o Fund for the Replacement of Animals in Medical Experiments (FRAME), Nottingham, UK

Summary — Dogs remain the main non-rodent species in preclinical drug development. Despite the current dearth of new drug approvals and meagre pipelines, this continues, with little supportive evidence of its value or necessity. To estimate the evidential weight provided by canine data to the probability that a new drug may be toxic to humans, we have calculated Likelihood Ratios (LRs) for an extensive dataset of 2,366 drugs with both animal and human data, including tissue-level effects and Medical Dictionary for Regulatory Activities (MedDRA) Level 1–4 biomedical observations. The resulting LRs show that the absence of toxicity in dogs provides virtually no evidence that adverse drug reactions (ADRs) will also be absent in humans. While the LRs suggest that the presence of toxic effects in dogs can provide considerable evidential weight for a risk of potential ADRs in humans, this is highly inconsistent, varying by over two orders of magnitude for different classes of compounds and their effects. Our results therefore have important implications for the value of the dog in predicting human toxicity, and suggest that alternative methods are urgently required.

Key words: *canine, dog, drug development, preclinical testing, toxicology.*

Address for correspondence: *Jarrod Bailey, British Union for the Abolition of Vivisection (BUAV), 16a Crane Grove, London N7 8NN, UK.*
E-mail: *jarrod.bailey@mac.com*

Introduction

It is generally assumed that testing new pharmaceuticals on animals helps to ensure human safety and efficacy. Regulatory agencies worldwide require preclinical trials (e.g. 1, 2), which involve at least two species — typically one rodent and one non-rodent species — to determine toxicity and pharmacokinetics. The expectation is that additional data from the non-rodent will detect adverse effects not detected by rodent tests. Despite the current dearth of new drug approvals and meagre pipelines (e.g. 3, 4), this practice continues, with little supportive evidence of its value or necessity (5).

Dogs are used in significant numbers in science — approximately 90,000 are used per annum across the EU and the USA, according to the latest available figures (6–8). About 80% of this use is as the non-rodent species in the evaluation of pharmaceutical safety and efficacy (6). However, only limited evaluations of the reliability of the canine model for this purpose have been conducted, chiefly due to the difficulty of accessing relevant data, most of which are unpublished and proprietary to pharmaceutical companies. Those evaluations that have been conducted have usually employed 'concordance' metrics (e.g. 9), which various authors have interpreted as the true positive rate ('sensitivity') or the Positive Predictive Value (PPV). While these metrics are appropriate for assessing the reliability of a diagnostic test for a specific disorder (e.g. HIV infection), the insights they provide depend critically on the question being asked of the diagnostic test. However, they are not appropriate for assessing the salient question at issue with animal models, which is *whether or not they contribute significant weight to the evidence for or against the toxicity of a given compound in humans*. Overcoming this key problem — almost entirely overlooked by previous authors — requires a precise specification of the various terms used (see *Methods*). Briefly, the appropriate metrics are Likelihood Ratios (LRs; 10): the Positive Likelihood Ratio (PLR) and the inverse Negative Likelihood Ratio (iNLR). Therefore, there is clearly a need for the kind of statistically-appropriate critical analysis that we provide here. The dataset we have used is unique, in that it is large and allows the conditional probabilities required for the LRs (PLR/iNLR) to be calculated.

Methods

Animal models are widely used to assess the risk that a given compound will prove toxic in humans. As with any diagnostic test, their reliability can only be assessed by performing tests in which the same compound is given to both animals and humans, and the presence or absence of toxicity recorded. This leads to a 2 × 2 matrix of results, as shown in Figure 1 (11).