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).

Figure 1: A 2 × 2 matrix of results

	Compound toxic in humans	Compound not toxic in humans
Compound toxic in animal model	a: true positives (TPs)	b: false positives (FPs)
Compound not toxic in animal model	c: false negatives (FNs)	d: true negatives (TNs)

The basis of this matrix is that the human data are correct, and the dog data are true/false, if they do/do not match them. The various cells in this matrix allow a variety of diagnostic metrics to be deduced, of which the most familiar and widely used are the true positive rate for the test (or 'sensitivity' = a/[a + c]), and the true negative rate (or 'specificity' = d/[d + b]). In previous research into the reliability of animal models as predictors of toxicity in humans, some authors (e.g. 9) have focused on the sensitivity, expressed as the 'true positive concordance rate', or the so-called Positive Predictive Value (PPV), given by a/(a+b), which reflects the probability that human toxicity was correctly identified by the animal model, given that toxicity was observed in the animal model (e.g. 12). However, neither of these metrics is suitable for the role of assessing the evidential weight provided by any toxicity test. In the case of animal models, the sensitivity addresses only the ability of such models to detect toxicity that will subsequently manifest itself in humans. This is a necessary, but not sufficient, measure of evidential weight. Suppose, for example, that the animal model always indicates toxicity found in humans; it would then have a sensitivity of 100%. However, if, in addition, the model always indicates toxicity, even in humans, its evidential value is no better than simply dismissing *every* compound as toxic from the outset. Thus, a useful toxicity test must also be able to give insight into when toxicity seen in the animal model is not observed in humans, which requires knowledge of the *specificity* of the test.

There is, of course, an obvious reason for the focus on sensitivity in animal model evaluation: if a compound is found to be positive in an animal model, it is unlikely to go into human evaluation. Nevertheless, the fact remains that sensitivity alone cannot be an adequate guide to the value of animal models.

The case of the PPV is more subtle. This metric is a measure of the probability that human toxicity will be correctly identified, given that the animal model detected toxicity. As such, PPVs are conditional probabilities, the condition being the pre-existence of a positive animal test result. This makes PPVs dependent on the prevalence of toxicity in compounds, and thus an inappropriate measure of the reliability of the test with any specific compound (e.g. 10, 13).

Thus, any appropriate metric of the evidential value of animal models requires knowledge of *both* the sensitivity *and* the specificity of the model. This, in turn, implies that the appropriate metrics for the evidential weight provided by an animal model are LRs (e.g. 13). In general, these are ratios of functions of the sensitivity and specificity, which can be extracted from the 2 × 2 matrix given above. In the specific case of animal models in general, two LRs are relevant. The first is the so-called PLR, which is given by:

$$\text{PLR} = \text{sensitivity}/(1 - \text{specificity}) = (a / a + c)/(b/ b + d)$$

This LR captures the ability of an animal model to add evidential weight to the belief that a specific compound is toxic. Any animal model that gives a PLR that is statistically significantly higher than 1.0, can be regarded as contributing evidential weight to the probability that the compound under test will be toxic in humans.

The other relevant LR is the so-called iNLR, given by:

$$\text{iNLR} = \text{specificity}/(1 - \text{sensitivity}) = (d / b + d)/(c / a + c)$$

This LR captures the ability of an animal model to add evidential weight to the belief that a specific compound is not toxic: any animal model that gives an iNLR that is statistically significantly higher than 1.0, can be regarded as contributing evidential weight to the probability that the compound under test will not be toxic in humans.

It is worth noting at this point that the above definitions imply that a good animal model for detecting human toxicity is not necessarily also good for detecting an absence of toxicity. That is, a high PLR does not guarantee a high iNLR; this will emerge as a key issue in this study.

The above definitions also underscore the need for data on the human toxicity of compounds that

fail initial animal tests. Again, a key feature of the current study is that this issue has been overcome via data mining methods. Data were obtained from a leading pharmaceutical safety consultancy, Instem Scientific Limited (Harston, Cambridge, UK; http://www.instem-lss.com 'Safety Intelligence Programme'), with funding provided by FRAME. All the information stemmed from publicly accessible sources, including: PubMed (http://www.ncbi.nlm.nih.gov/pubmed), the FDA Adverse Event Reporting System (FAERS), DrugBank (http://www.drugbank.ca), and the National Toxicology Program (http://ntp.niehs.nih.gov). Data were available for more than 2,300 drug compounds in humans and preclinical species.

Inference of the good quality of the data used in this evaluation is outlined in the *Discussion*. Compounds were selected that feature in the FAERS, FDA New Drug Applications (FDA NDAs) and DrugBank. Thus, the drugs selected for this analysis are in clinical use, and have undergone preclinical testing: human and animal data are therefore available for them. A non-redundant list of parent moieties was created, for example, by normalising therapeutic products to their generic names (e.g. Lipitor to Atorvastatin). This yielded 2,366 compounds.

A signature of the effects of each compound was created, focusing on tissue-level effects (e.g. bradycardia and arrhythmic disorder would both be considered to be effects on heart tissues), as well as the individual observations, which were mapped to their MedDRA (Medical Dictionary for Regulatory Activities; http://www.meddramsso.com) counterparts. MedDRA observations are classified into four levels, Level 1 being the most specific and Level 4 providing a more generic 'System Organ Class'. These classifications help to eliminate false positives that may arise from species-specific observations, and help the identification of concordant observations that might otherwise have been missed, by their 'rolling up' into more-generic terms.

LRs were derived for broad tissue-level effects (n = 52), and more-specific biomedical observations (BMOs; n = 384), mapped to MedDRA classifications (Levels 1 [most specific] to Level 4 [more generic 'organ class']). Fourteen BMO classifications not involving dogs were eliminated from the study. A total of 3,275 comparisons were made between the human and the dog, for 2,366 compounds, involving 436 (52 + 384) classifications of effects. The Instem Scientific data on which our analysis was based are shown in the Appendix, and the full set of data, including 95% Confidence Intervals, are available on the FRAME website (www.frame.org.uk).

With regard to potential bias: FNs are more common than FPs, since there is a bias resulting from a 'precautionary principle' not to progress positives to human administration. This has been mitigated by limiting the dataset to compounds reported in the FAERS database. Therefore, all the compounds are certain to have proceeded to market, and animal preclinical data are available for these compounds. Specific details of how the FPs that were identified arose were not sought, because they were not pertinent to this analysis, and this was not feasible, given the nature of the dataset. It must be assumed that the dog data were correlated with the human data retrospectively, and/or the human data arose from post-marketing studies, and/or clinical trials were applied for and approved, since the adverse effect(s) in dogs were minor and/or mitigated by other data.

Results

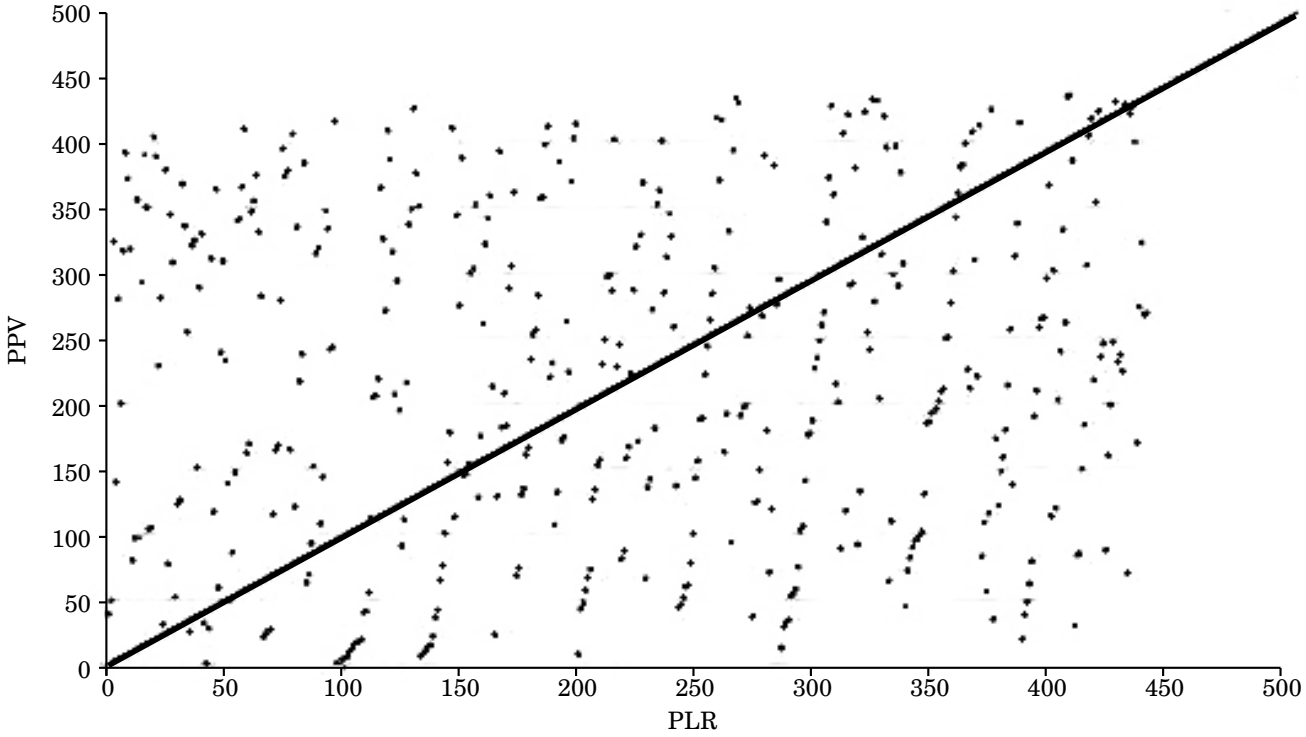
The inappropriate nature of PPVs is demonstrated in Figure 2, which shows a scatter plot of 'ranked' PPVs against equivalent ranked PLRs. Each PPV and PLR was ranked according to its value for each of the 436 classifications of effects, and these ranks were plotted against each other. The disparity is evidenced by the scatter of points, few of which lie close to the $y = x$ line that shows an ideal correlation. The misclassifications and misplaced assumptions of the accuracy of canine data for the prediction of human adverse drug reactions (ADRs) are clear. For example, MedDRA 'Level 4, Vascular Disorder' was ranked 20/436 with regard to the most favourable classifications for human predictivity based on PPV, but its cognate PLR ranked 404/436 — one of the least predictive. Conversely, MedDRA classification 'Level 2, Ventricular Conduction' ranked 30/436 by PLR, but 406/436 by PPV.

Dog PLRs were generally high (median ~28), implying that compounds that are toxic in dogs are likely also to be toxic in humans. However, because the PLRs vary considerably (range 4.7–548.7), with no obvious pattern regarding the form of toxicity, the reliability of this aspect of canine models cannot be generalised or regarded with confidence.

In contrast, the calculated inverse negative LRs (iNLRs) are substantially more consistent, but their median value of 1.11 (range 1.01–1.92) supports the view that dogs provide essentially no evidential weight to this aspect of toxicity testing. Specifically, the fact that a compound shows no toxic effects in dogs provides essentially no insight into whether the compound will also show no toxic effects in humans.

This lack of evidential weight has important implications for the role of dogs in toxicity testing, especially for the pharmaceutical industry. The critical observation for deciding whether a candidate drug can proceed to testing in humans is the absence of toxicity in tests on animals. However,

Figure 2: Scatter plot illustrating the lack of correlation of PPVs and PLRs of biomedical observations (BMOs) and tissue effects in humans and dogs



PPVs and PLRs for all 436 results were ordered according to their value, with the highest ranking first and the lowest last. For each BMO and tissue effect, the corresponding PPV and PLR rank were plotted against each other. If a perfect correlation exists, all points should lie on the line, where, for example, the 10th, 50th, and 100th highest PPV value would also be the 10th, 50th, and 100th highest PLR values. However, the significant scatter of the data points demonstrates that little correlation exists between PPV and PLR. For example: the 20th highest PPV ranks only 404/436 for PLR, whereas the 30th highest PLR ranks only 406/436 for PPV.

our findings show that the predictive value of the animal test in this regard is barely greater than that that would be obtained by chance (see below).

Discussion

The analysis presented here is urgently required, to support informed debate about the worth of animal models in preclinical testing. It is acknowledged among some stakeholders (if not universally among all stakeholders) that assessment of the scientific value of animal data in drug development is necessary, has been scarce, and has been thwarted for decades by the unavailability of relevant data for analysis (e.g. 14). Nevertheless, primarily due to concerns over privacy and commercial interests, data sharing and making data available continue to be resisted, in spite of assurances to the contrary from industry (14).

Those few analyses that have been done, tend to reflect unfavourably on animal models, including the dog. In 2012, a study that expressly set out to minimise bias, showed that 63% of serious ADRs had no counterparts in animals, and less than 20% of serious ADRs had a true positive corollary in animal studies (15). Other similar examples exist for testing generally (e.g. 16–18) and more-specifically, for example, in teratology (e.g. 19, 20) and drug-induced liver injury (e.g. 5, 21). One notable study claimed a good concordance for dog and human toxicology (10), though neither the predictive nature of the animal data for humans, nor the evidential weight provided by those data, were addressed (22).

We have, for the first time, addressed the salient question of *contribution of evidential weight for or against the toxicity of a given compound in humans* by data from dog tests, by using the appropriate metrics of LRs. Furthermore, we have applied the apposite LRs to a dataset of unprecedented scale, to critically question the value of the use of the dog as a preclinical species in the testing of new pharmaceuticals.

Substantiation of data quality is evidenced by: the methods used to source the data and the assured quality of the databases supplying them (listed above); the ways in which the data had been used recently as a basis for scientific publications and presentations (e.g. 23–26); and the international corporate and academic clients that have used the consultancy and its data (e.g. AstraZeneca; see 23–26). In addition, the impact of 'missing data' (i.e. unpublished data held by pharmaceutical companies) was mitigated by strictly limiting the dataset to drugs "with the greatest chance of having been evaluated in all the species included in the study" (here, dogs and humans). In other words, "…lack of evidence for an association between a compound and a specific BMO demonstrates a real absence of effect, and is not due to missing data" (Instem Scientific Ltd. Analysis Report, unpublished).

Naturally, there must be caveats. Our analysis was limited to data that are published and publicly available. It is widely acknowledged that many animal experimental results/preclinical data remain unpublished and/or proprietary, for a variety of reasons (e.g. 15, 27–30). Such publication bias is a major problem (e.g. 31–34), and, compounded by other factors such as size and quality of the animal studies, variability in the requirements for reporting animal studies, 'optimism bias', and lack of randomisation and blinding (28, 35), it means that gauging the true contribution of animal data to human toxicology is impossible — at least for third parties without access to pharmaceutical company files. All datasets are imperfect to varying degrees. However, it is only possible to use data which are available, and to ensure that, as far as feasible, those data are of good quality and as free from biases as possible, and that their analysis and derived conclusions are as objective as possible.

It must be made abundantly clear that we, the authors of this report, did not make decisions regarding the toxicity/non-toxicity of drugs, or decide upon or apply any criteria to such decisions. The mining of the data, and the decisions on toxicity of the drugs, were independent of the authors of this paper, and were made by one or both of the authors of the drug/toxicity papers and/or database submissions used, and the data-mining consultancy/curators of the Safety Intelligence Programme, Instem Scientific Limited. Therefore, if any pharmaceutical industry stakeholders have issues or concerns with our conclusions, we would encourage them to conduct further analyses by using their own proprietary data, and/or to facilitate such investigations by making available anonymised data, in accordance with the promotion of transparency encouraged by EU *Directive 2010/63/EU* (36), as well as to engage fully in constructive discussion and debate with us and our colleagues in animal protection organisations.

Our findings have practical implications for the use of animal models for toxicity testing, especially in the pharmaceutical industry. Reliance on flawed models of toxicity testing leads to two types of failure. If the models have poor PLRs, then there is a risk that many potentially useful compounds will be wrongly discarded, because of 'false positives' produced by the toxicity model. On the other hand, if the models have poor iNLRs, then many toxic compounds will wrongly find their way into human tests, and will fail in clinical trials. The relatively high PLRs found in this study show that animal models may not be leading to the loss of many potentially valuable candidate drugs through false positives. However, our results do imply that many toxic drugs are not being detected by animal models, leading to the risk of unnecessary harm to humans.

In this regard, our findings are entirely consistent with the acknowledged failure of animal models in general to provide guidance on likely toxicity ahead of the entry of compounds into human trials. Drug attrition has increased significantly over the past two decades (e.g. 3, 4, 37–42): 92–94% of all drugs that pass preclinical tests fail in clinical trials, mostly due to unforeseen toxicities (43–45), and half of those that succeed may be subsequently withdrawn or re-labelled due to ADRs not detected in animal tests (46). ADRs are a major cause of premature death in developed countries (47). A major contributing factor is the inadequacy of preclinical animal tests: one recent study showed that 63% of ADRs had no counterpart in animals, and less than 20% had a positive corollary in animal studies (15).

With specific regard to the dog, the most extensive study prior to the report we present here, concluded that 92% of dog toxicity studies did not provide relevant information in addition to that provided by the rat, and that the other 8% did not result in the immediate withdrawal of drugs from development, indicating that dog studies are not required for the prediction of safe doses for humans (17). There is a scientific basis for this: among several notable species differences which confound the extrapolation of data from dogs to humans, significant differences between humans and dogs in their cytochrome P450 enzymes (CYPs) — the major enzymes involved in drug metabolism — have been acknowledged for some time, compelling the conclusion that, "…it is readily seen that the dog is frequently not a good metabolic model for man and is poorly comparable to the rat and mouse" (for references, see 46). The lack of knowledge of canine CYPs has been highlighted, which is surprising, considering the extent of the use of dogs in preclinical testing. This problem is likely to be amplified by intra-species differences, as well as by inter-species differences (49). It may therefore be argued that, if many differ-

ences exist between different breeds or strains of the same species, then extrapolating pharmacokinetic data from that highly variable species to humans must not only be difficult, but must also be patently unreliable.

Conclusions

This analysis of the most comprehensive quantitative database of publicly-available animal toxicity studies yet compiled, suggests that dogs are highly inconsistent predictors of toxic responses in humans, and that the predictions they can provide are little better than those that could be obtained by chance — or tossing a coin — when considering whether or not a compound should proceed to testing in humans. In other words: "…for any putative source of evidential weight to be deemed useful, its specificity and sensitivity must be such that LR+ [PLR] >1. Tossing a coin contributes no evidential weight to a given hypothesis, as the sensitivity and specificity are the same — 50% — and thus the LR+ [PLR] is equal to 1" (22).

Dog PLRs were generally high, showing that a drug which is toxic in the dog is likely to be toxic in humans. However, they were extremely variable and with no obvious pattern, suggesting this aspect of dog tests cannot be considered particularly reliable or helpful. Further, though not within the scope of this analysis, it is of great interest whether the dog revealed any significant toxicities, that were also present in humans, that other species such as the rat did not. In other words, did the dog 'catch' any true human toxicities not caught by the rat? It has been previously argued that such toxicities are relatively low in number (e.g. the development of just 11% of new compounds was terminated due to effects uniquely seen in dogs, though the human significance of these could not be determined), which would further diminish any value the canine model may have in this respect (50).

More importantly, while iNLRs were much more consistent, they revealed that dogs provide essentially no evidential weight to this aspect of toxicity testing. Specifically, if a compound shows no toxic effects in dogs, this provides essentially no insight into whether the compound will also show no toxic effects in humans. This is crucial: the critical observation for deciding whether a candidate drug can proceed to testing in humans is the absence of toxicity in tests on animals, and our findings show that the predictive value of the dog test in this regard is barely greater than by chance.

A quantitative example illustrates this. Suppose researchers wish to investigate a candidate compound belonging to a family which prior experience indicates has a 70% probability of freedom from ADRs in humans. Before conducting tests in humans, the drug is tested in dogs. By using the median iNLR figure found by our study, if the compound shows no sign of toxicity in the dog, the probability that the compound will also show no toxic effects in humans will have been increased by the animal testing from 70% to 72%. The testing thus contributes essentially no additional confidence in the outcome, but at considerable extra cost, both in monetary terms and in terms of animal welfare. This also has obvious practical relevance to the issue of high attrition rates in clinical trials on new drug candidates.

It is argued that a comprehensive suite of more reliable alternative methods is now available (14, 51, 52). Combined with considerable public concern over the use of dogs in science (53), the high ethical costs of doing so, given the sensitive nature of dogs (e.g. 15, 54), and the expressed desire for the use of dogs as a second species in drug testing to have a scientific, rather than a habitual, basis (14), we conclude that the preclinical testing of pharmaceuticals in dogs cannot currently be justified on scientific or ethical grounds.

Acknowledgements

The authors are grateful to the British Union for the Abolition of Vivisection (BUAV), the Fund for the Replacement of Animals in Medical Experiments (FRAME), and The Kennel Club (via FRAME), for funding. They thank Robert Matthews for advice on inferential issues, Bob Coleman for his help and encouragement during the inception of this undertaking, and Instem Scientific (previously BioWisdom; Harston, Cambridge, UK) for scientific consultancy and for data analysis on integrated data relating to adverse events in model animal species. The research described in this article is based on the analysis and conclusions of the authors: it has not been subjected to each agency's peer review and policy review; therefore, it does not necessarily reflect the views of the organisations, and no official endorsement should be inferred.

Received 23.05.13; received in final form 11.09.13; accepted for publication 19.09.13.

References

1. Anon. (2004). *Directive 2004/27/EEC* of the European Parliament and the Council of 31 March 2004, amending *Directive 2001/83/EC* on the Community code relating to medicinal products for human use. *Official Journal of the European Union* **L136**, 30.04.2004, 34–57.
2. Anon. (2010). *Federal Food, Drug and Cosmetics Act*. Silver Spring, MD, USA: US Food and Drug Administration. Available at: http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/default.htm (Acc-

essed 11.10.13).
3. Duyk, G. (2003). Attrition and translation. *Science, New York* **302**, 603–605.
4. Kola, I. & Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nature Reviews Drug Discovery* **3**, 711–715.
5. Aithal, G.P. (2010). Mind the gap. *ATLA* **38**, Suppl. 1, 1–4.
6. UK Home Office (2013). *Statistics of Scientific Procedures on Living Animals — Great Britain 2012.* HC 549, 60pp. London, UK: The Stationery Office.
7. USDA (2011). *Annual Report. Animal Usage by Fiscal Year — Fiscal Year 2010,* 2pp. Riverdale, MD, USA: United States Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS). Available at: http://www. aphis. usda.gov/animal_welfare/efoia/downloads/2010_ Animals_Used_In_Research.pdf (Accessed 05.09. 13).
8. Anon. (2010). *Sixth Report on the Statistics on the Number of Animals Used for Experimental and Other Scientific Purposes in the Member States of the European Union. SEC(2010) 1107*, 14pp. Brussels, Belgium: European Commission. Available at: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0511:REV1:EN:PDF (Accessed 10. 10.13).
9. Olson, H., Betton, G., Robinson, D., Thomas, K., Monro, A., Kolaja, G., Lilly, P., Sanders, J., Sipes, G., Bracken, W., Dorato, M., Van Deun, K., Smith, P., Berger, B. & Heller, A. (2000). Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regulatory Toxicology & Pharmacology* **32**, 56–67.
10. Altman, D.G. & Bland, J.M. (1994). Diagnostic tests 2: Predictive values. *British Medical Journal* **309**, 102.
11. Anon. (2012). *Likelihood Ratios*. Oxford, UK: Centre for Evidence Based Medicine (CEBM). Available at: http://www.cebm.net/index.aspx?o=1043 (Accessed 10.10.13).
12. Greek, R. & Menache, A. (2013). Systematic reviews of animal models: Methodology *versus* epistemology. *International Journal of Medical Sciences* **10**, 206–221.
13. Grimes, D.A. & Schulz, K.F. (2005). Refining clinical diagnosis with likelihood ratios. *Lancet* **365**, 1500–1505.
14. Hasiwa, N., Bailey, J., Clausing, P., Daneshian, M., Eileraas, M., Farkas, S., Gyertyan, I., Hubrecht, R., Kobel, W., Krummenacher, G., Leist, M., Lohi, H., Miklosi, A., Ohl, F., Olejniczak, K., Schmitt, G., Sinnett-Smith, P., Smith, D., Wagner, K., Yager, J.D., Zurlo, J. & Hartung, T. (2011). Critical evaluation of the use of dogs in biomedical research and testing in Europe. *ALTEX* **28**, 326–340.
15. 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.
16. Igarashi, T., Nakane, S. & Kitagawa, T. (1995). Predictability of clinical adverse reactions of drugs by general pharmacology studies. *Journal of Toxicological Sciences* **20**, 77–92.
17. Broadhead, C.L., Jennings, M. & Combes, R. (1999). *A Critical Evaluation of the Use of Dogs in the Regulatory Toxicity Testing of Pharmaceuticals,* 106pp. Nottingham, UK: Fund for the Replacement of Animals in Medical Experiments (FRAME).
18. Litchfield, J.T.J. (1962). Symposium on clinical drug evaluation and human pharmacology. XVI. Evaluation of the safety of new drugs by means of tests in animals. *Clinical Pharmacology & Therapeutics* **3**, 665–672.
19. Bailey, J. (2008). Developmental toxicity testing: Protecting future generations? *ATLA* **36**, 718–721.
20. Schardein, J. (2000). *Chemically Induced Birth Defects*, 3rd edn, 1019pp. Boca Raton, FL, USA: CRC Press.
21. Spanhaak, S., Cook, D., Barnes, J. & Reynolds, J. (2008). *Species Concordance for Liver Injury*, 6pp. Cambridge, UK: Biowisdom Ltd. Available at: http://www.biowisdom.com/files/SIP_Board_Species_Concordance.pdf (Accessed 10.10.13).
22. Matthews, R.A. (2008). Medical progress depends on animal models — doesn't it? *Journal of the Royal Society of Medicine* **101**, 95–98.
23. Barnes, J.C., Matis, S., Kenna, G., Swinton, J., Bradley, P.M., Day, N.C., Reed, J.Z., Reynolds, J. & Cook, D. (2008). *The Safety Intelligence Program: An Intelligence Network for Drug-induced Liver Injury*, 2pp. Cambridge, UK: Biowisdom Ltd. Available at: http://bioblog.instem.com/downloads/Carboxylic_acids_A4.pdf (Accessed 10.10.13).
24. Sidaway, J.R.M., Roberts, S., Huby, R., Nicholson, A., Pemberton, J., South, M., Noeske, T., Engkvist, O., Bradley, P. & Reed, J. (2012). *Drug Toxicities Associated With Pharmacological Activity: Using Harmonised Data to Make the 'Known' Visible*, 2pp. Macclesfield, UK: Safety Assessment, AstraZeneca. Available at: http://bioblog.instem.com/wp-content/uploads/downloads/2012/03/SOT2012_make-the-known-visible_poster.pdf (Accessed 10.10.13).
25. 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.
26. Greco, I., Day, N., Riddoch-Contreras, J., Reed, J., Soininen, H., Kloszewska, I., Tsolaki, M., Vellas, B., Spenger, C., Mecocci, P., Wahlund, L.O., Simmons, A., Barnes, J. & Lovestone, S. (2012). Alzheimer's disease biomarker discovery using *in silico* literature mining and clinical validation. *Journal of Translational Medicine* **10**, 217.
27. Wandall, B., Hansson, S.O. & Ruden, C. (2007). Bias in toxicology. *Archives of Toxicology* **81**, 605–617.
28. Hackam, D.G. (2007). Translating animal research into clinical benefit. *British Medical Journal* **334**, 163–164.
29. ter Riet, G., Korevaar, D.A., Leenaars, M., Sterk, P.J., Van Noorden, C.J., Bouter, L.M., Lutter, R., Elferink, R.P. & Hooft, L. (2012). Publication bias in laboratory animal research: A survey on magnitude, drivers, consequences and potential solutions. *PLoS One* **7**, e43404.
30. Briel, M., Muller, K.F., Meerpohl, J.J., von Elm, E., Lang, B., Motschall, E., Gloy, V., Lamontagne, F., Schwarzer, G. & Bassler, D. (2013). Publication bias in animal research: A systematic review protocol. *Systematic Reviews* **2**, 23.
31. van der Worp, H.B., Howells, D.W., Sena, E.S., Porritt, M.J., Rewell, S., O'Collins, V. & Macleod, M.R. (2010). Can animal models of disease reliably inform human studies? *PLoS Medicine* **7**, e1000245.

32. Sena, E.S., van der Worp, H.B., Bath, P.M., Howells, D.W. & Macleod, M.R. (2010). Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS Biology* **8**, e1000344.
33. Perel, P., Roberts, I., Sena, E., Wheble, P., Briscoe, C., Sandercock, P., Macleod, M., Mignini, L.E., Jayaram, P. & Khan, K.S. (2007). Comparison of treatment effects between animal experiments and clinical trials: Systematic review. *British Medical Journal* **334**, 197.
34. Schott, G., Pachl, H., Limbach, U., Gundert-Remy, U., Ludwig, W.D. & Lieb, K. (2010). The financing of drug trials by pharmaceutical companies and its consequences. Part 1: A qualitative, systematic review of the literature on possible influences on the findings, protocols, and quality of drug trials. *Deutsches Arzteblatt International* **107**, 279–285.
35. Kilkenny, C., Parsons, N., Kadyszewski, E., Festing, M.F., Cuthill, I.C., Fry, D., Hutton, J. & Altman, D.G. (2009). Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLoS One* **4**, e7824.
36. Anon. (2010). *Directive 2010/63/EU* of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union* **L276**, 20.10.2010, 33–79.
37. US FDA (2004). *Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products*, 12pp. Silver Spring, MD, USA: US Department of Health and Human Services, Food and Drug Administration. Available at: http://www.fda.gov/ScienceResearch/Special Topics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm077262.htm (Accessed 10.10.13).
38. Issa, A.M., Phillips, K.A., Van Bebber, S., Nidamarthy, H.G., Lasser, K.E., Haas, J.S., Alldredge, B.K., Wachter, R.M. & Bates, D.W. (2007). Drug withdrawals in the United States: A systematic review of the evidence and analysis of trends. *Current Drug Safety* **2**, 177–185.
39. Bennani, Y.L. (2011). Drug discovery in the next decade: Innovation needed ASAP. *Drug Discovery Today* **16**, 779–792.
40. Eichler, H.G., Aronsson, B., Abadie, E. & Salmonson, T. (2010). New drug approval success rate in Europe in 2009. *Nature Reviews Drug Discovery* **9**, 355–356.
41. Hughes, B. (2008). 2007 FDA drug approvals: A year of flux. *Nature Reviews Drug Discovery* **7**, 107–109.
42. Hartung, T. (2009). Toxicology for the twenty-first century. *Nature, London* **460**, 208–212.
43. Harding, A. (2004). More compounds failing phase I. FDA chief warns that high drug attrition rate is pushing up the cost of drug development. *The Scientist*, 6 August 2004. Available at: http://www.the-scientist.com/?articles.view/articleNo/23003/title/More-compounds-failing-Phase-I/ (Accessed 10.10.13).
44. Okie, S. (2006). Access before approval — a right to take experimental drugs? *New England Journal of Medicine* **355**, 437–440.
45. Aurup, P. (2012). *Er Danmark et Attraktivt Land for Klinisk Forskning?* (*Is Denmark an Attractive Country for Clinical research?*), 23pp. Ballerup, Denmark: MSD Laboratories. Available at: http://di.dk/SiteCollectionDocuments/Opinion/Sundhed/Høring/Præsentation%20-%20Peter%20Aurup,%20Merck.pdf (Accessed 10.10.13).
46. Anon. (1990). *FDA Drug Review: Post Approval Risks 1976–1985*. GAO/PEMD-90-15, 132pp. Washington, DC, USA: US General Accounting Office. Available at: http://161.203.16.4/d24t8/141456.pdf (Accessed 10.10.13).
47. Lazarou, J., Pomeranz, B.H. & Corey, P.N. (1998). Incidence of adverse drug reactions in hospitalized patients: A meta-analysis of prospective studies. *Journal of the American Medical Association* **279**, 1200–1205.
48. Gad, S.C. (2006). *Animal Models in Toxicology,* 952pp. Boca Raton, FL, USA: CRC Press.
49. Martinez, M.N., Antonovic, L., Court, M., Dacasto, M., Fink-Gremmels, J., Kukanich, B., Locuson, C., Mealey, K., Myers, M.J. & Trepanier, L. (2013). Challenges in exploring the cytochrome P450 system as a source of variation in canine drug pharmacokinetics. *Drug Metabolism Reviews* **45**, 218–230.
50. Broadhead, C.L., Betton, G., Combes, R., Damment, S., Everett, D., Garner, C., Godsafe, Z., Healing, G., Heywood, R., Jennings, M., Lumley, C., Oliver, G., Smith, D., Straughan, D., Topham, J., Wallis, R., Wilson, S. & Buckley, P. (2000). Prospects for reducing and refining the use of dogs in the regulatory toxicity testing of pharmaceuticals. *Human & Experimental Toxicology* **19**, 440–447.
51. Spielmann, H., Kral, V., Schäfer-Korting, M., Seidle, T., McIvor, E., Rowan, A. & Schoeters, G. (2011). *The AXLR8 Consortium. Alternative Testing Strategies, Progress Report 2011*, 364pp. Berlin, Germany: Institute of Pharmacy, Free University of Berlin. Available at: http://scrtox.eu/~scrtox/images/stories/AXLR8-2011.pdf (Accessed 10.10.13).
52. Anon. (2007). *Toxicity Testing for the 21st Century: A Vision and a Strategy*, 216pp. Washington, DC, USA: National Academies Press.
53. Anon. (2009). *Public Opinion*. London, UK: European Coalition to End Animal Experiments. Available at: http://www.eceae.org/en/what-we-do/campaigns/12-million-reasons/public-opinion (Accessed 10.10.13).
54. Hare, B., Brown, M., Williamson, C. & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science, New York* **298**, 1634–1636.

Appendix

Table A1: Raw data from Instem Scientific's 'Safety Intelligence Programme', showing the number of drugs associated with ADRs in humans and dogs

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	**Human/Dog a (TP)**	**Dog b (FP)**	**Human c (FN)**	**Neither d (TN)**	**Dog total**	**Human total**
1	Level 1 — qrs prolongation	10	0	22	2334	10	32
2	Level 2 — glomerulonephritis and nephrotic syndrome	13	0	104	2249	13	117
3	Level 2 — kidney neoplastic disorder	8	0	57	2301	8	65
4	Level 1 — ocular hypertension	6	0	17	2343	6	23
5	Level 3 — glaucoma and ocular hypertension	6	0	19	2341	6	25
6	Level 2 — glaucomas (excluding congenital)	6	0	19	2341	6	25
7	Level 1 — nephrotic syndrome	8	0	68	2290	8	76
8	Level 2 — renal neoplasms (malignant)	7	0	49	2310	7	56
9	Level 3 — renal and urinary tract neoplasms (malignant and unspecified)	7	0	50	2309	7	57
10	Level 2 — urinary tract neoplasms (unspecified malignancy not elsewhere classified; nec)	6	0	27	2333	6	33
11	Level 1 — kidney neoplastic disorder	6	0	27	2333	6	33
12	Level 1 — orthostatic hypotension	10	0	170	2186	10	180
13	Level 2 — hepatic peroxisome proliferation	5	0	16	2345	5	21
14	Level 1 — cholestatic jaundice	7	0	131	2228	7	138
15	Brain	78	1	854	1433	79	932
16	Level 2 — renal failure and impairment	62	2	510	1792	64	572
17	Bodily fluid	474	18	980	894	492	1454
18	Skin	154	7	1066	1139	161	1220
19	Large intestine	44	2	810	1510	46	854
20	Kidney	214	11	696	1445	225	910
21	Heart	509	27	727	1103	536	1236
22	Liver	515	31	853	967	546	1368
23	Vasculature	465	28	842	1031	493	1307
24	Nervous tissue	165	10	1001	1190	175	1166
25	Level 2 — cholestasis and jaundice	47	3	418	1898	50	465
26	Nerve	43	3	459	1861	46	502
27	Level 4 — cardiac disorder	480	34	698	1154	514	1178
28	Muscle	168	12	829	1357	180	997
29	Level 4 — blood and lymphatic system disorders	28	2	465	1871	30	493
30	Level 4 — hepatobiliary disorder	472	34	787	1073	506	1259
31	Level 3 — hepatic and hepatobiliary disorders	338	26	764	1238	364	1102
32	Level 3 — renal disorder	75	6	394	1891	81	469
33	Ear	12	1	444	1909	13	456
34	Level 4 — vascular disorder	415	35	835	1081	450	1250
35	Bodily fluid liver	304	26	848	1188	330	1152
36	Level 3 — renal disorders (excluding nephropathies)	113	10	591	1652	123	704
37	Level 4 — skin and subcutaneous tissue disease	195	18	619	1534	213	814
38	Level 2 — hepatic lipid peroxidation	32	3	129	2202	35	161
39	Level 4 — investigation	339	32	723	1272	371	1062
40	Level 1 — kidney failure	21	2	292	2051	23	313

*All entries are numbered for identification only (column 1). The second column (**parameters**) indicates the specific biomedical observation (BMO) in question (e.g. 'bradycardia' or 'arrhythmic disorder'), or tissue-level effects (e.g. 'heart', which would encompass these two BMOs). The BMOs were mapped to their MedDRA (Medical Dictionary for Regulatory Activities) counterpart, which are classified into four levels, level 1 being the most specific and level 4 providing a more generic 'System Organ Class'. The number of drugs for which ADRs were observed in each species is shown in columns 3–8. **Human/Dog** represents drugs for which an ADR was reported in both humans and dogs: these are True Positives (TPs), and correspond to cell 'a' in the 2 × 2 matrix (see* Methods, Figure 1*). **Dog** represents drugs for which an ADR was reported in dogs, but not in humans: these are False Positives (FPs), and correspond to cell 'b' in the 2 × 2 matrix. **Human** represents drugs for which an ADR was reported in humans, but not in dogs: these are False Negatives (FNs), and correspond to cell 'c' in the 2 × 2 matrix. **Neither** represents drugs for which an absence of ADRs was evident in both humans and dogs: these are True Negatives (TNs), and correspond to cell 'd' in the 2 × 2 matrix. Notably, lack of an association between a compound and a specific BMO was assumed (by the data provider) to demonstrate a real absence of effect, and not be due to missing data. To minimise the impact of missing data, the group of compounds in the dataset were chosen with the greatest chance of having been evaluated in all the species included in the study (see* Methods*). The total number of drugs exhibiting ADRs in each species, regardless of the presence or absence of ADRs in the other species, is given in the final two columns: **Dog** = a + b (TP + FP); **Human** = a + c (TP + FN).*

nec = not elsewhere classified.

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	**Human/Dog a (TP)**	**Dog b (FP)**	**Human c (FN)**	**Neither d (TN)**	**Dog total**	**Human total**
41	Level 4 — injury, poisoning and procedural complications	276	27	670	1393	303	946
42	Level 4 — renal and urinary disorders	222	22	630	1492	244	852
43	Breast	20	2	434	1910	22	454
44	Adrenal gland	10	1	224	2131	11	234
45	Pancreas	19	2	421	1924	21	440
46	Level 3 — cardiac and vascular investigations (excluding enzyme tests)	320	34	647	1365	354	967
47	Level 3 — arrhythmic disorder	294	32	577	1463	326	871
48	Level 3 — skin vascular abnormalities	182	20	482	1682	202	664
49	Level 2 — trophic disorders	9	1	110	2246	10	119
50	Level 4 — musculoskeletal disorder	105	12	706	1543	117	811
51	Neuromuscular tissue	104	12	712	1538	116	816
52	Level 4 — neurological disorder	85	10	627	1644	95	712
53	Haemolymphoid tissue	68	8	748	1542	76	816
54	Joint	17	2	420	1927	19	437
55	Level 1 — cardiac hypertrophy	25	3	106	2232	28	131
56	Level 2 — hepatic microsomal lipid peroxidation	8	1	26	2331	9	34
57	Level 1 — interstitial nephritis	8	1	108	2249	9	116
58	Respiratory tissue	70	9	653	1634	79	723
59	Level 2 — haemolysis	15	2	330	2019	17	345
60	Level 4 — respiratory, thoracic and mediastinal disorders	112	15	717	1522	127	829
61	Level 1 — muscle twitching	29	4	162	2171	33	191
62	Level 2 — non-site specific vascular disorders nec	137	19	382	1828	156	519
63	Level 3 — vascular disorder	150	21	525	1670	171	675
64	Level 2 — skin vasomotor conditions	171	24	331	1840	195	502
65	Level 1 — vasodilation	113	16	246	1991	129	359
66	Level 2 — nephropathies and tubular disorders nec	56	8	313	1989	64	369
67	Level 1 — congestive heart failure	7	1	158	2200	8	165
68	Prostate gland	7	1	234	2124	8	241
69	Level 3 — chemical injury and poisoning	168	25	544	1629	193	712
70	Level 2 — poisoning and toxicity	168	25	544	1629	193	712
71	Level 2 — vascular test	193	29	580	1564	222	773
72	Stomach	86	13	863	1404	99	949
73	Level 2 — electrocardiogram observation	59	9	306	1992	68	365
74	Level 2 — rate and rhythm disorders nec	208	32	530	1596	240	738
75	Level 3 — muscular disorder	78	12	632	1644	90	710
76	Level 3 — epidermal and dermal conditions	13	2	354	1997	15	367
77	Urogenital tissue	70	11	695	1590	81	765
78	Level 3 — decreased and non-specific blood pressure disorders and shock	153	25	498	1690	178	651
79	Bodily fluid kidney	61	10	479	1816	71	540
80	Level 3 — general system disorders nec	30	5	502	1829	35	532
81	Level 3 — central nervous system vascular disorders	18	3	266	2079	21	284
82	Level 1 — jaundice	12	2	327	2025	14	339
83	Level 2 — dermal and epidermal conditions nec	12	2	329	2023	14	341
84	Level 1 — hypertrophy	6	1	68	2291	7	74
85	Level 3 — respiratory and mediastinal neoplasms (malignant and unspecified)	6	1	72	2287	7	78
86	Level 2 — lower respiratory tract neoplasms	6	1	79	2280	7	85
87	Level 3 — respiratory tract neoplastic disorder	6	1	79	2280	7	85
88	Level 2 — hepatocellular damage and hepatitis nec	215	36	582	1533	251	797
89	Level 2 — heart rate and pulse investigations	203	34	529	1600	237	732
90	Level 3 — arteriosclerosis, stenosis, vascular insufficiency and necrosis	171	29	564	1602	200	735
91	Level 1 — acute kidney failure	34	6	289	2037	40	323
92	Lung	96	17	733	1520	113	829
93	Level 2 — myocardial disorder	73	13	198	2082	86	271
94	Level 2 — vascular hypotensive disorders	145	26	470	1725	171	615
95	Level 3 — myocardial disorder	100	18	321	1927	118	421
96	Level 2 — muscle observation	50	9	336	1971	59	386
97	Bodily fluid cardiovascular	105	19	608	1634	124	713
98	Level 4 — immunological disorder	44	8	527	1787	52	571
99	Level 2 — cerebrovascular and spinal necrosis and vascular insufficiency	11	2	138	2215	13	149
100	Level 2 — inflammatory kidney disease	11	2	155	2198	13	166

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	**Human/Dog a (TP)**	**Dog b (FP)**	**Human c (FN)**	**Neither d (TN)**	**Dog total**	**Human total**
101	Uterus	11	2	201	2152	13	212
102	Bladder	11	2	293	2060	13	304
103	Level 3 — cardiac disorder signs and symptoms	65	12	509	1780	77	574
104	Level 3 — hypertension	96	18	520	1732	114	616
105	Pituitary gland	16	3	221	2126	19	237
106	Level 2 — hepatic failure and associated disorders	16	3	274	2073	19	290
107	Level 1 — hypotension	140	27	442	1757	167	582
108	Level 2 — bronchospasm and obstruction	31	6	376	1953	37	407
109	Level 3 — bronchial disorders (excluding neoplasms)	31	6	376	1953	37	407
110	Eye	77	15	762	1512	92	839
111	Level 3 — injury	120	24	511	1711	144	631
112	Level 1 — hepatic apoptosis	15	3	111	2237	18	126
113	Level 2 — central nervous system vascular disorders	10	2	90	2264	12	100
114	Level 3 — haemolyses and related conditions	15	3	339	2009	18	354
115	Level 2 — hepatic collagen synthesis	5	1	23	2337	6	28
116	Level 1 — atrial flutter	5	1	26	2334	6	31
117	Level 1 — glutathione depletion	5	1	29	2331	6	34
118	Level 2 — haematological analyses nec	5	1	29	2331	6	34
119	Level 1 — liver failure	10	2	221	2133	12	231
120	Level 1 — cardiomyocyte apoptosis	5	1	32	2328	6	37
121	Level 2 — cardioprotection	5	1	35	2325	6	40
122	Level 3 — haematology investigations (including blood groups)	5	1	40	2320	6	45
123	Level 1 — lung neoplastic disorder	5	1	49	2311	6	54
124	Level 2 — encephalopathies (toxic and metabolic)	5	1	49	2311	6	54
125	Level 2 — respiratory tract and pleural neoplasms (malignancy unspecified nec)	5	1	50	2310	6	55
126	Level 2 — diabetic complications (renal)	5	1	52	2308	6	57
127	Level 1 — abnormal liver function	10	2	323	2031	12	333
128	Level 2 — hepatic enzymes and function abnormalities	10	2	323	2031	12	333
129	Level 1 — toxic hepatitis	5	1	94	2266	6	99
130	Level 3 — cardiac physiological observation	229	46	322	1769	275	551
131	Level 2 — cardiac disorder	64	13	378	1911	77	442
132	Digestive tissue	49	10	731	1576	59	780
133	Level 4 — general disorders and administration site conditions	58	12	599	1697	70	657
134	Level 2 — ventricular arrhythmias and cardiac arrest	118	25	390	1833	143	508
135	Level 2 — peripheral vasoconstriction, necrosis and vascular insufficiency	101	22	209	2034	123	310
136	Salivary gland	9	2	440	1915	11	449
137	Level 1 — vasoconstriction	100	23	185	2058	123	285
138	Level 1 — cardiotoxicity	26	6	142	2192	32	168
139	Level 1 — thrombosis	13	3	186	2164	16	199
140	Mouth	13	3	297	2053	16	310
141	Level 2 — ischaemic coronary artery disease	89	21	413	1843	110	502
142	Level 1 — tachycardia	112	27	402	1825	139	514
143	Level 3 — hepatic physiological phenomenon	267	65	519	1515	332	786
144	Level 2 — hypertension	73	18	480	1795	91	553
145	Level 3 — coronary arterial disease	93	23	413	1837	116	506
146	Level 1 — nephrotoxicity	36	9	212	2109	45	248
147	Level 1 — injury	20	5	150	2191	25	170
148	Level 1 — left ventricular hypertrophy	8	2	62	2294	10	70
149	Level 1 — asthma	12	3	218	2133	15	230
150	Level 1 — kidney papillary necrosis	4	1	27	2334	5	31
151	Level 2 — atrial natriuretic factor secretion	4	1	28	2333	5	32
152	Level 2 — renal disorders (congenital)	4	1	32	2329	5	36
153	Level 3 — renal and urinary tract disorders (congenital)	4	1	37	2324	5	41
154	Level 1 — acute lung injury	4	1	39	2322	5	43
155	Level 1 — pulmonary toxicity	4	1	46	2315	5	50
156	Level 2 — herg current	4	1	77	2284	5	81
157	Level 2 — inflammation	4	1	109	2252	5	113
158	Level 2 — non-site specific gastrointestinal haemorrhages	4	1	125	2236	5	129
159	Level 2 — hepatic and hepatobiliary disorders	138	36	524	1668	174	662
160	Level 1 — torsade de pointes	19	5	163	2179	24	182

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	**Human/Dog a (TP)**	**Dog b (FP)**	**Human c (FN)**	**Neither d (TN)**	**Dog total**	**Human total**
161	Level 1 — arrhythmic disorder	79	21	383	1883	100	462
162	Level 2 — heart failure	30	8	326	2002	38	356
163	Level 1 — hypertension	71	19	477	1799	90	548
164	Level 2 — renal and urinary tract injuries nec	26	7	185	2148	33	211
165	Level 1 — renal injury	26	7	185	2148	33	211
166	Level 2 — renal structural abnormalities and trauma	26	7	189	2144	33	215
167	Level 2 — allergic conditions	26	7	374	1959	33	400
168	Level 3 — allergic conditions	26	7	377	1956	33	403
169	Level 2 — coronary necrosis and vascular insufficiency	62	17	409	1878	79	471
170	Level 2 — left ventricular failure	21	6	165	2174	27	186
171	Level 4 — congenital, familial and genetic disorders	21	6	294	2045	27	315
172	Level 2 — central nervous system haemorrhages and cerebrovascular accidents	7	2	209	2148	9	216
173	Level 3 — lower respiratory tract disorders (excluding obstruction and infection)	38	11	368	1949	49	406
174	Bone	31	9	398	1928	40	429
175	Level 2 — signs and symptoms	24	7	439	1896	31	463
176	Level 2 — non-site specific embolism and thrombosis	17	5	230	2114	22	247
177	Level 2 — non-site specific injuries nec	20	6	158	2182	26	178
178	Level 1 — ventricular tachyarrhythmia	10	3	32	2321	13	42
179	Cardiovascular tissue	73	22	445	1826	95	518
180	Level 1 — long qt interval	23	7	186	2150	30	209
181	Level 1 — cardiac arrest	23	7	211	2125	30	234
182	Level 2 — renal disorder	19	6	185	2156	25	204
183	Level 2 — cardiac conduction abnormality	44	14	278	2030	58	322
184	Level 1 — bradycardia	78	25	271	1992	103	349
185	Level 4 — metabolism and nutrition disorders	62	20	368	1916	82	430
186	Level 3 — tissue disease	24	8	218	2116	32	242
187	Level 1 — pulmonary oedema	18	6	145	2197	24	163
188	Level 4 — eye disorder	9	3	88	2266	12	97
189	Level 2 — muscular disorder	15	5	286	2060	20	301
190	Level 3 — immunological disorder	12	4	258	2092	16	270
191	Level 3 — cardiac and vascular disorders congenital	9	3	173	2181	12	182
192	Level 2 — renal observation	6	2	85	2273	8	91
193	Oesophagus	9	3	258	2096	12	267
194	Level 3 — gastrointestinal haemorrhage	6	2	159	2199	8	165
195	Level 2 — gastrointestinal haemorrhage	6	2	177	2181	8	183
196	Level 1 — hepatic injury	61	21	341	1943	82	402
197	Level 2 — abdominal injury	61	21	343	1941	82	404
198	Level 1 — hepatotoxicity	92	32	386	1856	124	478
199	Musculoskeletal tissue	23	8	525	1810	31	548
200	Level 2 — pulmonary oedema	20	7	197	2142	27	217
201	Level 1 — liver disorder	17	6	211	2132	23	228
202	Level 1 — muscular disorder	8	3	113	2242	11	121
203	Level 1 — liver cirrhosis	8	3	139	2216	11	147
204	Nose	8	3	401	1954	11	409
205	Level 1 — hypoxia	13	5	150	2198	18	163
206	Level 2 — conditions associated with abnormal gas exchange	13	5	151	2197	18	164
207	Level 2 — muscle tone abnormalities	13	5	232	2116	18	245
208	Level 1 — ventricular arrhythmia	41	16	159	2150	57	200
209	Level 3 — neuromuscular disorder	23	9	351	1983	32	374
210	Level 3 — heart failure	53	21	388	1904	74	441
211	Bodily fluid cardiac	53	21	443	1849	74	496
212	Level 4 — endocrine disorder	10	4	125	2227	14	135
213	Level 2 — hepatic monooxygenase system	5	2	15	2344	7	20
214	Level 1 — sudden cardiac death	5	2	48	2311	7	53
215	Level 3 — encephalopathy	5	2	52	2307	7	57
216	Thyroid gland	10	4	266	2086	14	276
217	Level 3 — diabetes-related disorder	5	2	63	2296	7	68
218	Level 2 — vascular smooth muscle cell proliferation	5	2	69	2290	7	74
219	Level 1 — acute liver failure	5	2	76	2283	7	81
220	Level 1 — muscle rigidity	5	2	134	2225	7	139

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	Human/Dog a (TP)	Dog b (FP)	Human c (FN)	Neither d (TN)	Dog total	Human total
221	Level 1 — vascular disorder	5	2	144	2215	7	149
222	Level 2 — vascular anomalies congenital nec	5	2	147	2212	7	152
223	Level 2 — purpuras (excluding thrombocytopenic)	5	2	198	2161	7	203
224	Level 3 — coagulopathies and bleeding diatheses (excluding thrombocytopenic)	5	2	220	2139	7	225
225	Level 3 — infections (pathogen unspecified)	5	2	270	2089	7	275
226	Level 2 — cardiomyopathy	32	13	227	2094	45	259
227	Level 4 — neoplasms benign, malignant and unspecified (including cysts and polyps)	66	27	359	1914	93	425
228	Level 2 — parenchymal lung disease	12	5	182	2167	17	194
229	Level 2 — myocardial contraction	81	34	130	2121	115	211
230	Level 2 — pulmonary vascular resistance	28	12	61	2265	40	89
231	Level 2 — musculoskeletal and connective tissue signs and symptoms nec	28	12	115	2211	40	143
232	Level 3 — musculoskeletal disorder	28	12	117	2209	40	145
233	Level 1 — endothelial dysfunction	7	3	77	2279	10	84
234	Level 2 — vascular malformations and acquired anomalies	7	3	175	2181	10	182
235	Level 1 — cardiac disorder	7	3	240	2116	10	247
236	Level 4 — gastrointestinal disorder	7	3	280	2076	10	287
237	Level 3 — heart valve disorder	16	7	128	2215	23	144
238	Level 3 — vascular physiological observation	91	40	201	2034	131	292
239	Level 2 — non-site specific necrosis and vascular insufficiency nec	52	23	285	2006	75	337
240	Level 1 — cardiac lesion	9	4	56	2297	13	65
241	Level 2 — cell metabolism disorders nec	9	4	117	2236	13	126
242	Level 1 — pulmonary fibrosis	9	4	121	2232	13	130
243	Level 2 — neuromuscular disorder	9	4	216	2137	13	225
244	Bodily fluid vascular	22	10	301	2033	32	323
245	Level 1 — hepatitis	22	10	367	1967	32	389
246	Level 3 — metabolic disorder	11	5	133	2217	16	144
247	Level 3 — vasculitis	11	5	298	2052	16	309
248	Level 3 — pulmonary vascular disorder	24	11	206	2125	35	230
249	Level 1 — bronchial spasm	13	6	246	2101	19	259
250	Level 3 — respiration disorder	15	7	277	2067	22	292
251	Level 2 — supraventricular arrhythmia	34	16	238	2078	50	272
252	Level 1 — cholestasis	19	9	181	2157	28	200
253	Level 1 — contracture	27	13	91	2235	40	118
254	Level 1 — aortic valve insufficiency	8	4	60	2294	12	68
255	Level 2 — diseases of aortic valve	8	4	79	2275	12	87
256	Level 3 — fatal outcomes	6	3	50	2307	9	56
257	Level 2 — death and sudden death	6	3	50	2307	9	56
258	Gallbladder	10	5	201	2150	15	211
259	Level 3 — miscellaneous and site unspecified neoplasms (malignant and unspecified)	8	4	171	2183	12	179
260	Cartilage	4	2	39	2321	6	43
261	Level 2 — biliary excretion	4	2	41	2319	6	45
262	Level 1 — arterial thrombosis	4	2	46	2314	6	50
263	Level 1 — fibrosis	6	3	132	2225	9	138
264	Level 2 — fibrosis	6	3	132	2225	9	138
265	Level 2 — vasculitides	8	4	228	2126	12	236
266	Level 1 — acute hepatitis	6	3	151	2206	9	157
267	Level 2 — muscle weakness	6	3	188	2169	9	194
268	Level 1 — myoclonus	4	2	113	2247	6	117
269	Level 2 — neurological signs and symptoms nec	4	2	119	2241	6	123
270	Integumentary tissue	6	3	428	1929	9	434
271	Level 3 — renal physiological observation	72	37	225	2032	109	297
272	Level 3 — embolism and thrombosis	23	12	392	1939	35	415
273	Level 1 — cardiomyopathy	9	5	104	2248	14	113
274	Level 1 — hepatocyte damage	9	5	117	2235	14	126
275	Level 3 — biliary tract neoplasm	48	27	188	2103	75	236
276	Level 3 — haemorrhage	28	16	502	1820	44	530
277	Level 1 — vasospasm	7	4	65	2290	11	72
278	Endocrine tissue	7	4	299	2056	11	306
279	Bodily fluid muscle	40	23	405	1898	63	445
280	Level 1 — liver neoplastic disorder	24	14	96	2232	38	120

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	**Human/Dog a (TP)**	**Dog b (FP)**	**Human c (FN)**	**Neither d (TN)**	**Dog total**	**Human total**
281	Level 2 — biliary tract neoplasm	24	14	99	2229	38	123
282	Level 2 — hepatobiliary neoplasms (malignancy unspecified)	24	14	99	2229	38	123
283	Level 1 — renal disorder	12	7	148	2199	19	160
284	Level 1 — ventricular tachycardia	27	16	147	2176	43	174
285	Level 2 — l-type calcium current	10	6	90	2260	16	100
286	Level 1 — haemorrhage	15	9	269	2073	24	284
287	Level 1 — cerebral vasospasm	5	3	27	2331	8	32
288	Level 3 — ancillary infectious topics	5	3	86	2272	8	91
289	Level 2 — inflammatory disorders following infection	5	3	86	2272	8	91
290	Level 2 — lower respiratory tract inflammatory and immunologic conditions	5	3	163	2195	8	168
291	Level 1 — muscle spasm	8	5	98	2255	13	106
292	Level 1 — myocardial disorder	22	14	60	2270	36	82
293	Level 2 — cardiac afterload	11	7	40	2308	18	51
294	Level 2 — liver cirrhosis	14	9	239	2104	23	253
295	Level 1 — ischaemic disease	20	13	141	2192	33	161
296	Level 1 — ischaemic cardiomyopathy	20	13	156	2177	33	176
297	Level 1 — ventricular fibrillation	27	18	127	2194	45	154
298	Level 3 — hepatobiliary neoplasms (malignant and unspecified)	24	16	116	2210	40	140
299	Level 1 — st-segment elevation	9	6	30	2321	15	39
300	Level 1 — necrosis	9	6	59	2292	15	68
301	Level 2 — necrosis	9	6	60	2291	15	69
302	Level 2 — hemorrhage	18	12	372	1964	30	390
303	Level 1 — apoptosis	6	4	92	2264	10	98
304	Level 2 — arterial and aortic injuries	3	2	12	2349	5	15
305	Level 1 — neoplastic disorder	6	4	118	2238	10	124
306	Level 2 — neoplasms unspecified (malignancy and site unspecified nec)	6	4	122	2234	10	128
307	Level 2 — hepatic lipid level	3	2	20	2341	5	23
308	Level 1 — renal fibrosis	3	2	22	2339	5	25
309	Level 1 — right ventricular hypertrophy	3	2	22	2339	5	25
310	Level 1 — hepatic mitochondrial swelling	3	2	36	2325	5	39
311	Level 1 — glomerular sclerosis	3	2	39	2322	5	42
312	Level 2 — pericardium disorder	3	2	47	2314	5	50
313	Small intestine	6	4	233	2123	10	239
314	Level 1 — hypersensitive vasculitis	3	2	85	2276	5	88
315	Level 1 — inflammation	3	2	96	2265	5	99
316	Level 2 — cutaneous vasculitis	3	2	102	2259	5	105
317	Level 2 — peripheral vascular disorder	3	2	143	2218	5	146
318	Connective tissue	3	2	204	2157	5	207
319	Level 2 — muscle tone abnormal	10	7	208	2141	17	218
320	Level 1 — pulmonary hypertension	17	12	94	2243	29	111
321	Level 2 — systemic vascular resistance	24	17	79	2246	41	103
322	Level 2 — hepatic cytochrome p450 level	7	5	50	2304	12	57
323	Level 1 — renal impairment	7	5	276	2078	12	283
324	Level 2 — hepatic protein biosynthesis	11	8	61	2286	19	72
325	Bone-marrow	15	11	302	2038	26	317
326	Level 3 — lipid metabolism disorder	19	14	137	2196	33	156
327	Level 2 — lipid metabolism and deposit disorders nec	19	14	137	2196	33	156
328	Level 1 — hepatic necrosis	46	34	177	2109	80	223
329	Level 1 — premature cardiac complex	4	3	51	2308	7	55
330	Level 3 — pericardium disorder	4	3	72	2287	7	76
331	Level 1 — muscle weakness	4	3	174	2185	7	178
332	Level 1 — myocardial infarction	21	16	296	2033	37	317
333	Level 2 — pulmonary hypertension	17	13	108	2228	30	125
334	Level 3 — electrolyte and fluid balance conditions	9	7	107	2243	16	116
335	Level 4 — infections and infestations	9	7	381	1969	16	390
336	Level 1 — hepatomegaly	16	13	116	2221	29	132
337	Ovary	11	9	327	2019	20	338
338	Level 1 — hepatic steatosis	18	15	135	2198	33	153
339	Level 1 — cardiogenic shock	6	5	83	2272	11	89
340	Testis	19	16	260	2071	35	279

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	Human/Dog a (TP)	Dog b (FP)	Human c (FN)	Neither d (TN)	Dog total	Human total
341	Level 1 — heart failure	13	11	208	2134	24	221
342	Level 1 — centrilobular hepatic necrosis	7	6	27	2326	13	34
343	Level 3 — acid-base disorders	7	6	50	2303	13	57
344	Level 2 — action potential duration	38	35	91	2202	73	129
345	Level 3 — procedural related injuries and complications nec	13	12	193	2148	25	206
346	Level 2 — cardiovascular injuries	15	14	107	2230	29	122
347	Level 1 — atrial fibrillation	16	15	162	2173	31	178
348	Level 2 — action potential	28	27	101	2210	55	129
349	Level 1 — hyperaemia	7	7	34	2318	14	41
350	Level 2 — hepatic glutathione level	7	7	37	2315	14	44
351	Level 2 — cardiac function diagnostic procedures	13	13	180	2160	26	193
352	Level 2 — liver respiration	5	5	32	2324	10	37
353	Level 1 — coronary artery vasospasm	5	5	38	2318	10	43
354	Level 2 — hepatic RNA synthesis	4	4	18	2340	8	22
355	Level 2 — renal vascular and ischaemic conditions	6	6	65	2289	12	71
356	Level 1 — ECG abnormality	9	9	161	2187	18	170
357	Level 2 — circulatory collapse and shock	8	8	135	2215	16	143
358	Level 1 — complete atrioventricular block	5	5	52	2304	10	57
359	Level 2 — effective renal plasma flow	4	4	29	2329	8	33
360	Level 1 — hepatic fibrosis	7	7	116	2236	14	123
361	Level 1 — oedema	5	5	64	2292	10	69
362	Level 2 — total fluid volume increased	5	5	72	2284	10	77
363	Level 2 — metabolic acidoses (excluding diabetic acidoses)	4	4	41	2317	8	45
364	Level 2 — oedema	5	5	73	2283	10	78
365	Level 2 — bile acid biosynthesis	3	3	40	2320	6	43
366	Level 3 — increased intracranial pressure and hydrocephalus	4	4	92	2266	8	96
367	Level 2 — increased intracranial pressure disorders	4	4	92	2266	8	96
368	Level 3 — neurological disorder	5	5	153	2203	10	158
369	Thymus gland	5	5	164	2192	10	169
370	Level 2 — cerebrospinal fluid tests (excluding microbiology)	3	3	69	2291	6	72
371	Level 1 — intracranial hypertension	3	3	69	2291	6	72
372	Level 3 — neurological, special senses and psychiatric investigations	3	3	70	2290	6	73
373	Level 2 — vascular tissue neoplasm	3	3	71	2289	6	74
374	Level 2 — nervous system haemorrhagic disorders	3	3	105	2255	6	108
375	Level 2 — coronary blood flow	40	41	41	2244	81	81
376	Level 2 — bile flow	14	15	59	2278	29	73
377	Level 2 — cardiac and vascular procedural complications	11	12	183	2160	23	194
378	Level 1 — infarct size	20	22	95	2229	42	115
379	Level 2 — coronary arterial disease	7	8	64	2287	15	71
380	Level 3 — bile duct disorder	7	8	192	2159	15	199
381	Level 2 — hepatic portal blood flow	6	7	26	2327	13	32
382	Level 2 — heart weight	6	7	36	2317	13	42
383	Level 2 — hepatic glycogen level	5	6	18	2337	11	23
384	Level 2 — hepatocyte proliferation	5	6	39	2316	11	44
385	Level 2 — liver weight	53	66	156	2091	119	209
386	Level 2 — hepatic glucose release	4	5	10	2347	9	14
387	Level 1 — left ventricular dysfunction	4	5	49	2308	9	53
388	Level 2 — atrioventricular conduction	9	12	32	2313	21	41
389	Level 1 — myocardial infarct size	9	12	42	2303	21	51
390	Level 2 — hepatic drug metabolism	6	8	54	2298	14	60
391	Level 1 — muscle hypotonia	3	4	43	2316	7	46
392	Level 2 — left ventricular mass	3	4	45	2314	7	48
393	Level 2 — hepatic blood flow	8	11	64	2283	19	72
394	Level 1 — myocardial fibrosis	5	7	37	2317	12	42
395	Level 2 — renal blood flow	28	40	63	2235	68	91
396	Level 1 — atrioventricular block	7	10	94	2255	17	101
397	Level 2 — hepatobiliary signs and symptoms	17	25	142	2182	42	159
398	Level 1 — coronary artery thrombosis	4	6	9	2347	10	13
399	Level 2 — site specific embolism and thrombosis nec	4	6	12	2344	10	16
400	Level 2 — renal function	6	9	104	2247	15	110
401	Level 2 — ventricular refractory period	2	3	6	2355	5	8
402	Level 2 — hepatic cytochrome p450 function	4	6	76	2280	10	80
403	Level 1 — high cardiac output	2	3	17	2344	5	19
404	Level 1 — chronic hepatitis	2	3	22	2339	5	24
405	Level 1 — myocarditis	2	3	46	2315	5	48

	Parameters	Number of drugs					
	Adverse effect: tissue-level or BMO (MedDRA Level 1–4)	**Human/Dog a (TP)**	**Dog b (FP)**	**Human c (FN)**	**Neither d (TN)**	**Dog total**	**Human total**
406	Level 2 — non-infectious myocarditis	2	3	50	2311	5	52
407	Level 4 — pregnancy, puerperium and perinatal conditions	2	3	65	2296	5	67
408	Level 1 — st-segment depression	2	3	67	2294	5	69
409	Level 1 — inflammatory kidney disease	2	3	68	2293	5	70
410	Level 2 — cardiac conduction	5	8	28	2325	13	33
411	Level 1 — infarction	5	8	74	2279	13	79
412	Level 1 — hepatocyte hypertrophy	15	25	40	2286	40	55
413	Level 2 — site-specific vascular disorders nec	3	5	29	2329	8	32
414	Level 2 — bile duct infections and inflammations	3	5	42	2316	8	45
415	Level 2 — liver vascular disorder	4	7	78	2277	11	82
416	Level 2 — renal vascular resistance	9	16	48	2293	25	57
417	Level 3 — hepatobiliary investigations	6	11	241	2108	17	247
418	Level 2 — hepatic function test	6	11	241	2108	17	247
419	Level 2 — cardiac function	8	15	88	2255	23	96
420	Level 1 — centrilobular hypertrophy	9	18	22	2317	27	31
421	Level 1 — ventricular premature complex	8	16	117	2225	24	125
422	Level 3 — vascular injury	5	10	59	2292	15	64
423	Level 2 — ventricular conduction	2	4	6	2354	6	8
424	Level 2 — mixed acid-base disorders	2	4	17	2343	6	19
425	Level 1 — acidosis	2	4	17	2343	6	19
426	Level 1 — tachyarrhythmia	2	4	33	2327	6	35
427	Level 1 — hepatic enzyme level	3	6	104	2253	9	107
428	Level 2 — coronary vascular resistance	13	28	25	2300	41	38
429	Level 3 — hepatic and biliary neoplasms (benign)	15	33	106	2212	48	121
430	Level 1 — ischaemia-reperfusion injury	5	11	67	2283	16	72
431	Level 2 — biliary secretion	4	9	36	2317	13	40
432	Level 2 — left ventricular pressure	7	17	33	2309	24	40
433	Level 1 — altered hepatic foci	2	5	14	2345	7	16
434	Level 1 — cholangitis	2	5	24	2335	7	26
435	Level 2 — hepatic enzyme function	2	5	28	2331	7	30
436	Level 1 — arteriosclerosis	2	5	37	2322	7	39
437	Level 2 — hepatobiliary neoplasms benign	13	33	105	2215	46	118
438	Level 2 — renal plasma flow	3	8	30	2325	11	33
439	Level 2 — left ventricular function	3	8	31	2324	11	34
440	Level 2 — myocardial blood flow	7	19	24	2316	26	31
441	Level 1 — hepatocellular adenoma	12	33	97	2224	45	109
442	Level 2 — hepatic vacuolation	2	6	9	2349	8	11
443	Level 2 — cerebral blood flow	5	16	79	2266	21	84
444	Level 3 — neurological physiological observation	5	16	83	2262	21	88
445	Level 2 — peripheral vascular resistance	7	24	53	2282	31	60
446	Level 2 — ventricular repolarisation	2	7	17	2340	9	19
447	Level 2 — left ventricular systolic blood pressure	5	19	24	2318	24	29
448	Level 2 — hepatocyte vacuolation	5	26	22	2313	31	27
449	Level 1 — hepatocyte degeneration	2	12	16	2336	14	18
450	Level 2 — left ventricular contraction	2	12	16	2336	14	18