Importance of Systematic Reviews and Meta-analyses of Animal Studies: Challenges for Animal-to-Human Translation

Zahra Bahadoran,[1] Parvin Mirmiran,[2] Khosrow Kashfi,[3] and Asghar Ghasemi[4,*]

Results of animal experiments are used for understanding the pathophysiology of diseases, assessing safety and efficacy of newly developed drugs, and monitoring environmental health hazards among others. Systematic reviews and meta-analyses of animal data are important tools to condense animal evidence and translate the data into practical clinical applications. Such studies are conducted to explore heterogeneity, to generate new hypotheses about pathophysiology and treatment, to design new clinical trial modalities, and to test the efficacy and the safety of the various interventions. Here, we provide an overview regarding the importance of systematic reviews and meta-analyses of animal data and discuss common challenges and their potential solutions. Current evidence highlights various problems and challenges that surround these issues, including lack of generalizability of data obtained from animal models, failure in translating data obtained from animals to humans, poor experimental design and the reporting of the animal studies, heterogeneity of the data collected, and methodologic weaknesses of animal systematic reviews and meta-analyses. Systematic reviews and meta-analyses of animal studies can catalyze translational processes more effectively if they focus on a well-defined hypothesis along with addressing clear inclusion and exclusion criteria, publication bias, heterogeneity of the data, and a coherent and well-balanced assessment of studies' quality.

Animal experiments are scientific studies of animals for gaining new biologic knowledge or solving specific medical or biologic problems.[25] The use of animal models in biomedical research has a long history;[26,60] the first animal experiments are attributed to the Greek physician–scientists Aristotle (384 to 322 BC) and Erasistratus (304 to 258 BC).[32] Animal experiments have been involved in our understanding of human biology and have paved the way to improving human health.[3,8] Various medical achievements including the development of some vaccines (smallpox, anthrax, typhoid, cholera, and plague), treatment of ancient diseases (beriberi and rickets), development of blood transfusion, and discovery of insulin highlight the importance of animal studies.[3] About 90% of Nobel prizes awarded for physiology or medicine since 1900 have been dependent on research using animals.[72] Currently, NIH annually spends 12.0 to 14.5 billion dollars on animal experiments, and approximately 47% of grants awarded have animal-research–based components; this number has been fairly stable over the last decade.[9]

However, experiments on animals have been a frequent subject of debate.[59] Some believe that animal studies cannot appropriately and accurately predict human outcomes and are, therefore, unsuitable for assessing interventions or the toxicity of a substance in human.[39,49] In the evidence-based medicine hierarchy, animal research is considered to provide the lowest level of evidence for both prognostic and therapeutic studies; however, they can provide biologic plausibility.[14]

To provide more valid conclusions regarding the potential clinical usefulness of animal studies, systematic reviews and meta-analyses of existing animal experiments have been used increasingly during the past 3 decades and offer an important step forward.[74] However the questions of how and when systematic reviews and meta-analyses of animal experiments would be useful tools to address challenges of translational process have remained unclear. Our aim here is to provide an overview on the importance of systematic reviews and meta-analyses of animal experiments, to discuss the most relevant common challenges, and to offer some potential solutions.

Pros and Cons of Animal Experiments

An estimate of worldwide laboratory animal use in 2005 reports that 115.3 million research animals are used annually.[84] Despite policies toward minimizing animal use for research, overall numbers of animal use are increasing according to data from the United Kingdom[40] and United States.[27] Areas of animal experiments are basic and applied research, testing of drugs and other products, and educational purposes.[1,51] However, just how much animal experiments benefit human health has been a controversial issue for a long time. In 1976, the relative contribution of basic science research compared with clinical research in terms of important medical advances was reported to be 62%.[16] In 1978,[82] the *British Medical Journal* published a commentary implying that contribution to medical practice obtained from animal studies had not been as much as reported earlier[16] and, in 2004, the journal highlighted this challenge by publishing another commentary, entitled "Where is the evidence that animal research benefits humans?"[75]

Benefits from animal experiments can include making possible genetic manipulations,[29] studying effects of newly developed drugs,[2] and understanding disease pathophysiology.[33] Animal studies provide a framework that makes environmental and genetic manipulations a reality. This approach is rarely feasible in humans.[29] Preclinical evaluations are helpful for toxicity screening and establishing safety profiles of newly developed compounds and treatments[29] and can prevent testing of new treatment modalities in humans when the preliminary assessment in animals is not satisfactory. Several examples in the literature have shown that the adverse side effects and toxicity determined in animal studies have translated quite well into humans. The concordance rate between animal and human toxicity data is reported to be as high as 71%, and the sensitivity for predicting carcinogenic observations in animals that have applied to humans has been reported as 84%.[53] Moreover, depending on the species, types of drugs, and target organs, 37% to more than 70% of adverse drug reactions in humans were predicted from animal studies.[53] Despite concerns about reliability of animal toxicity testing and its failure regarding translation to humans,[6] this methodology has been the stalwart basis of ensuring safety of in-human clinical testing and use.[53,88] Moreover, regardless of the consensus regarding the urgent demand for human-focused alternative methods of predicting human toxicity[5,6] and the development of several in vitro substitutes (i.e., physicochemical methods, tissue culture, microbiologic systems, stem cells, DNA chips, micro fluidics, computer analysis models, epidemiologic surveys, and plant-tissue–based materials),[4] the lack of an appropriate alternative to study the whole-body pharmacokinetics and pharmacodynamics of drugs and their subsequent metabolites[4] implies the critical role of animals.

Animal experiments also provide a unique opportunity for evaluating the mechanistic actions of drugs, discovering new drug targets and biomarkers, and assessing pharmacodynamic and pharmacokinetic parameters of drugs.[2] Furthermore, animal studies provide unique insights into human biology, pathophysiology and the underlying mechanisms of diseases; new hypotheses are generated during animal experiments that can be used for conducting preventive and/or therapeutic clinical trials.[29,33,85]

Conversely, animal experiments have been criticized because therapeutic efficacy in animals often does not translate into the clinical domain.[30] Evidence shows that few results obtained from animal research is reproduced later in human trials; one study[30] reported that among highly-cited papers published in prestigious journals, only 37% (95% CI, 26% to 48%) were replicated in human randomized trials, 18% were contradicted by randomized trials, and 45% remained untested. Of the 700 effective treatments in animal models of acute ischemic stroke, only aspirin and very early intravenous thrombolysis or fibrinolytic therapy with alteplase (recombinant tissue plasminogen activator) have proved to be effective in humans.[69,76,87] The average rate of successful translation from various animal models of cancer to clinical cancer trials has also been reported to be less than 8%.[58] Focusing on 6 systematic reviews of animal studies that were published up to 2004 and evaluating the evidence for leading to clinical trials, one group concluded that "The contribution of animal studies to clinical medicine requires urgent formal evaluation."[75]

However, despite some limitations, well-designed and -conducted animal research likely will long remain an essential step for testing hypotheses at the whole-organism level and for validating human data.[8] Moreover, although no animal model is a complete replica for a process within a human body, intact animals do provide a useful model of the complex interaction of the physiologic processes.[4] The pros and cons of animal experiments from translational research point of view are summarized in Figure 1.

Systematic Review and Meta-analysis: Clinical Versus Preclinical Point of View

Definition and a brief history. Systematic review provides a transparent mean for gathering, synthesizing, and appraising the findings of studies on a particular topic or question by using scientific methodology.[45,83] Systematic reviews can help researchers and clinicians keep up with literature by summarizing a large body of evidence and helping to explain differences among studies on the same question.[17] Meta-analysis is the statistical method used to combine results from relevant studies, and the resultant larger sample size provides greater reliability of the estimates of a treatment effect.[34] The critical appraisal and synthesis of research findings by using a systematic method was developed in 1970s, and the term 'meta-analysis' was coined in 1976.[70] Now, meta-analysis is highly valued in evidence-based medicine and is considered the top of evidence hierarchy.[14]

Compared with clinical research, systematic reviews of animal studies have not been common practice in the past.[67] The scientific rationale for systematic reviews of animal studies was initially outlined in a commentary published in *Nature*,[78] where the authors called for systematic reviews to examine the human clinical utility of animal experiments and emphasized the comparison of such results with those of the corresponding clinical trials. Soon thereafter, various authors[37,73-75,77,79] further emphasized that systematic reviews and meta-analyses of experimental investigations can clarify whether and how translation from animal to clinical research could progress and can provide a unique opportunity to review the appropriateness of the animal models used.

Applications and usefulness. Meta-analyses of clinical trials are generally conducted to estimate overall effect sizes (Figure 2) of specific interventions and have a critical value for decision-making in clinical practice.[31] In contrast, systematic reviews and meta-analyses of preclinical studies are more exploratory and their findings are generally used to generate new hypotheses and also to help researchers to design human clinical trials.[36] The main applications of systematic reviews and meta-analyses of animal studies may include: 1) addressing gap of knowledge and generating new hypotheses; 2) providing new insights for clinical studies that help the translational process; and 3) addressing weaknesses and improving the quality of future animal studies (i.e., study design, reporting and ethical issues). Systematic reviews and meta-analyses of animal studies determine the extent of current knowledge in the field, provide new insights regarding unanswered questions, and minimize unnecessary duplication of animal studies.[36] Compared with clinical trials, animal studies are more heterogenous, and exploring possible sources of this heterogeneity using meta-analysis can provide seeds for thought[36] and therefore prompt new experiments.[18]

Meta-analysis combines findings from several experiments and increases the power of the analysis; consequently, knowledge about efficacy and side effects of a treatment or intervention may be more robust and thus provide new insights for clinical research.[36] Because animal studies are subjected to possible sources of bias, including selective analysis and outcome-reporting biases or excess significance bias,[11,74] using systematic review and meta-analysis combine the results

Pros	Cons
Have led to several medical advancements for humans	Therapeutic efficacy in animals often does not translate into clinical practice
Have led to new hypotheses	Are expensive
Provide an opportunity to investigate: • The toxicity and safety of new compounds and treatments • The mechanistic actions of drugs, discovering new drug targets and biomarkers • A complete life cycle • Environmental and genetic manipulations, which are not feasible in human • Human biology, pathophysiology, and the underlying mechanisms of diseases • Novel targets for clinical treatments	Differences in response to newly developed drugs leading to adverse human effects undetected in animal studies (e.g., TGN1412 [also known as CD28-Super-MAB]; the arthritis drug Vioxx; some NSAID, including Oraflex, Zomepirac, Bromfenac, and phenylbutazone)

Figure 1. The pros and cons of animal experiments from the point of view of translational research[76]

Step	Explanation
Framing the question	• Formulating a focused, unresolved and explicit research question consisting of Participants, Interventions, Comparisons, Outcomes, and Study design (PICOS)
Defining eligibility criteria	• Specifying study characteristics (such as PICOS, length of follow-up) • Specifying report characteristics (such as years of publication, language, publication status)
Transforming the research question into a search strategy	• Transforming research question into search strategy using key words or terms and Boolean operators including 'or,' 'and,' or 'not'
Searching databases and collecting references	• Identifying appropriate databases and sources of studies (general biomedical and topic-specific databases) and checking other sources, including reference lists • Saving references (titles or abstracts) in a reference manager software
Screening and selecting studies	• Unifying citations of all databases into one file in reference software and remove citations that appear more than once • Initial screening of titles and abstracts of the references and identifying papers based on potential relevance • Retrieving full texts of the potentially relevant reports • Assessing full-text reports for compliance of studies with eligibility criteria • Making final decisions on study inclusion and proceed to data collection
Extracting data	• Extracting information including details of methods, participants, setting, context, interventions, outcomes, results, publications, and investigators
Assessing the quality of the included studies	• Assessing quality of study report using published scales (e.g., Jadad scale, Newcastle-Otawa scale, and the STROBE statement for human studies)
Conducting meta-analysis on extracted data	• Determining the likely nature of outcome data (dichotomous, continuous, ordinal, counts and rates, time-to-event data) • Calculating an effect size for each comparison • Weighting the effects sizes • Calculating a summary of effect size • Calculating the heterogeneity • Deciding whether random-effects meta-analyses, fixed-effect meta-analyses or both methods will be used

Figure 2. Eight key steps to conducting a systematic review and meta-analysis[18,34,54]

of studies to provide a more reliable conclusion.[36] Systematic reviews and meta-analyses of animal research improve the precision of estimated effect sizes used in calculating the power of proposed human trials and reduce the risk of false-negative results.[75] Such analysis also provides a good estimation of the power of the reported effect size and reproducibility of the results.[15]

Systematic reviews and meta-analyses of animal studies contribute to improving the methodologic quality of animal experiments, evidence-based selection of animal models, evidence-based translation of animal data to the clinic, and implementation of the 3Rs (Replacement, Reduction, and Refinement) to conduct more ethical animal experiments.[18] Such studies can provide insights into limitations of animal models and help to improve the relevance of animal models to clinical trial design;[73] they also help researchers to identify and remedy deficiencies in conducting and reporting studies.[73]

Key steps to conducting a systematic review or meta-analysis. Despite differences in the purpose, design, and conduction of systematic reviews and meta-analyses of preclinical and clinical studies,[89] 8 key steps provided here[34] (Figure 3) are common to both; interested readers can consult other papers[30,46,77] for details

Effect size	A standardized scale free estimate of the relation between an exposure and an outcome. This term is applied to any measurement of the difference in the outcome between study groups
Heterogeneity	Any kind of variability among studies in a systematic review may be termed heterogeneity: • Clinical heterogeneity: variability in the participants, interventions and outcomes studied • Methodologic heterogeneity: variability in the study design and risk of bias • Statistical heterogeneity: variability in the intervention effects being evaluated in the different studies
Fixed-effects meta-analysis	Fixed-effects meta-analysis assumes there is a common treatment effect across all study settings Any differences between observed effect sizes are attributed to sampling error The summary treatment effect in a fixed-effects model is a weighted average of study-specific effect sizes
Random effects meta-analysis	Under the random-effects assumption, treatment effects are supposed to vary from study to study The differences in observed effect sizes are attributed not only to random error but also to variation in true treatment effects The summary treatment effect in a random-effects model is an estimate of the average of the distribution of treatment effects across various study settings
Publication bias	The tendency on the parts of investigators, reviewers, and editors to submit or accept manuscripts for publication based on the direction or strength of the study findings
Outcome reporting bias	A specific form of bias arising from the selective reporting of the set of study outcomes, defined as the selection for publication of a subset of the original recorded outcome variables on the basis of the results
Excess significant bias	Selective reporting biases favoring statistically significant or positive results Selective analysis and outcome reporting biases appear when there are many analyses that can be performed, but only the analysis with the best results is presented resulting in misleading findings

Figure 3. Common terms used in systematic review and meta-analysis[20,21,34,41,42,68,85]

regarding systematic review and meta-analysis of experimental studies.[36,54,89]

The first step is framing a clear, complete, and unresolved question; defining the question will guide the rest of the process[67] and, along with the rationale for conducting a systematic review, constitutes one of the most pivotal parts of the study.[64] The problem addressed should be specified in the form of an unambiguous and structured question (refer to PICOS structure, i.e., Participants, Interventions, Comparisons, Outcomes, and Study design).[46,56] A specific question for animal research generally includes 4 components: 1) intervention or exposure; 2) disease or health problem of interest; 3) animal species; and 4) outcomes.[54]

The second step is defining and prespecifying eligibility criteria for including and excluding studies, an step that distinguished a systematic from a narrative review.[56] Eligibility criteria are combinations of study characteristics (according to aspects of the study question, PICOS) and report characteristics (such as years of publication, language, and publication status).[18,34]

The third step is transforming the research question into a search strategy, which is obtained by splitting the research question into search components (PICOS), identifying relevant search terms for each component, and combining search terms by using Boolean operators.[34,54] The search strategy for systematic review needs to be comprehensive and include all relevant databases.[64] To be as comprehensive as possible, the search strategy must include synonyms, related terms, and variant spellings; using truncation and wildcard symbols (searching technique used in databases in which characters within or end of a word be substituted by a symbol such as * or ?) can also facilitate capturing keyword variations.[34] The search strategy needs to be written in such a way that it can be repeated by others.[64]

The fourth step for conducting a systematic review and meta-analysis is searching databases and collecting references; to prevent missing relevant studies, searching at least 2 databases (e.g., PubMed, Embase) equipped with thesaurus terms is recommended.[18,54] Some critical tips regarding designing and performing a comprehensive search strategy to identify potentially relevant animal studies on a specific research topic have been discussed previously.[54] Beyond the general medical databases, other literature-searching databases are available that contain specific information on animal science.[55,81] For more information on alternatives thesaurus terminology for animal science refer to https://pubs.nal.usda.gov/animal-use-alternatives-thesaurus-terminology-alphabetical-listing.

The fifth step is appraising the studies retrieved during database searching and making decisions regarding which studies to include in the review; a typical process[34] for selecting studies is summarized in Figure 3.

The sixth step to conduct a systematic review and meta-analysis is data extraction from the obtained references; the information includes details of methods (study design, study duration, allocation, blinding), participants (total numbers, setting, diagnostic criteria, age, sex), outcomes (definition, diagnostic method, name of scale, definition of threshold, timing, unit of measurement), interventions (doses, route of delivery), results (numbers that will be required to perform meta-analysis), publications, and investigators.[34] For preclinical studies, 2 types of information including the predefined study design characteristics and outcome data (including the outcome measure used, the number of animals, the aggregate value of effect such as mean, median, and, where applicable, a measure of group variance) must be extracted.[89]

The seventh step is formal assessment of the quality of the included studies; quality may include 3 components including internal validity (risk of bias), external validity (applicability/ variability), and reporting quality.[91] The GRADE (Grading of Recommendations, Assessment, Development and Evaluations) system for rating the overall quality of the papers included in a systematic review is recommended.[28] Several quality-assessment tools for preclinical and clinical studies have been developed;[34] these tools include scales that are scoring systems for various components of quality or checklists in which specific questions are asked.[34] To assess quality of preclinical studies some tools like SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE) and Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies (CAMARADES) facilitate and improve the critical appraisal of evidence from animal studies.[38,91]

The eighth and final step is conducting meta-analysis on extracted data, which can be dichotomous or continuous; the most common approaches for dichotomous data are pooled estimations of the risk ratio or odds ratio, whereas the common approach for continuous data is standardized mean difference (SMD) estimation.[31] The results of a meta-analysis are displayed in a forest plot, which allows readers to visualize and interpret findings.[36] Heterogeneity (variation across studies) analysis, sensitivity analysis, and evaluation of publication bias are further analyses.[31,34] Definition of common terms in systematic review and meta-analysis are presented in Figure 2.

Common Limitations and Pitfalls of Systematic Reviews and Meta-analyses of Animal Studies for Translational Purposes

Despite the critical role of systematic reviews and meta-analyses of animal experiments that link basic sciences to clinical practice, conducting such studies is not as straightforward as those in human clinical trials.[71] Overview of preclinical animal literature demonstrates that the concordance between treatment effect in animal experiments and humans clinical trials is limited[43,86] and that the translational success rate is unpredictable and ranges from 0% to 100%.[53] Potential explanations for the failure of animal models to capture treatment effects in humans can be summarized as: 1) intrinsic limitations of animal models; 2) poor design and reporting of animal studies; and 3) factors limiting the applicability of systematic review and meta-analysis of animal studies trying to obtain pooled treatment effect (benefit or harm).

Intrinsic limitations of animal models. The diversity of animal species studied, differences of experimental designs, and variations in animal characteristics are important limitations that may lead to translational failure;[36,75] in fact, human physiology and the pathophysiology of diseases are not adequately captured by animal models, a problem that makes animal studies mostly impractical for translating to similar results in human clinical studies.[58] The most important question prior to conducting a systematic review and meta-analysis, therefore, is whether the treatment effects or outcome of interest measured in animal models can be generalized to humans. Lack of external validity (generalizability) of results obtained from animal models that do not sufficiently reflect a disease in humans may be considered as one of the most important problems that makes animal data difficult to translate.[86] Pathophysiologies of diseases are not entirely simulated by animal models, a challenge that cannot easily be overcome, unless more suitable new models can be formulated.[43,74] In some cases, animals are incapable of providing a sufficiently good model for evaluating the safety (e.g., evaluating sensitivity to potential renal toxicants) or efficacy of drugs (e.g., for Alzheimer's disease) to yield relevant information that can be translated to human health benefit.[65]

Taken together, conducting a meta-analysis of animal studies to estimate efficacy and translate it to humans requires adequate consideration of relevance and limitations of the animal models used.

Poor design and reporting of animal studies. Shortcomings of the experimental design, conduct, and analysis and the reporting of the results[12,86] are other causes of poor concordance between preclinical and clinical outcomes; this issue can be resolved fairly easily.[43] Animal literature and the estimated effect sizes derived from the animal experiments are highly susceptible to the quality of experiments;[29,43] lack of randomization, allocation concealment, blinding of researchers regarding the treatment, and blind assessments of outcomes may lead to overestimation of treatment effects.[35,50] Overoptimistic conclusions regarding efficacy that is derived from methodologically flawed animal studies may cause failure when translating the experiments to clinical settings.[86]

In a systematic review of highly cited (median, 889; range, 639 to 2233) animal studies published between 1980 to 2000 in 7 leading scientific journals (*Science*, *Nature*, *Cell*, *Nature Medicine*, *Nature Genetics*, *Nature Immunology*, and *Nature Biotechnology*), only 49% were rated as being of good methodologic quality.[30] Dose–response gradients, clinically relevant outcomes, and long-term end points were determined in most studies, whereas few studies included random allocation of animals, adjustment for multiple-hypothesis testing, or blinded assessment of the outcomes.[30] Having a dose–response design is one of the most important criteria that can increase the probability of translation of study findings into human practice (odds ratio, 3.3; 95% CI = 1.1 to 10.1).[30]

Low sample size of animal experiments is one of the most challenging issues that may lead to a decreased power of estimated effect size in a meta-analysis and can severely disrupt the process of research progression from experimental into clinical settings.[75] Low sample size has been estimated to contribute to the low statistical power (median 21%) of observed effect sizes in several specific fields of animal studies, such as neuroscience;[15] such a low power of estimated effect size, in the absence of other biases, provides unreliable findings due to low probability of a true effect, low positive predictive value (PPV) of an observed effect, and an exaggerated estimate of the magnitude of a true effect.[15] In some cases, adequate power of an estimated effect size is only achieved once at least 100 experiments are included in a meta-analysis;[90] to address this issue, various available methods for determining the sample size of animal experiments can be used.[23]

The overestimated treatment effect and lack of reproducibility of animal experiments also contribute to inadequate reporting of experimental methods and materials and study results.[52] Current animal literature suffers severely from a lack of rigorous reporting of experimental details, including a lack of a statement of the hypothesis, animal characteristics, animal numbers, experimental design, and statistical analyses.[48] Given that the age of animals can affect disease phenotype and response to treatment, poor reporting of age can also be considered as an important confounding factor leading to translational failure.[24] Published papers that lacked information on randomization, concealment of group allocation, or blinded outcomes assessment resulted in overestimated treatment effect compared with papers that reported these points of information.[57]

Factors limiting the applicability of systematic reviews and meta-analyses of animal studies. *High heterogeneity of data.* Synthesis of data from multiple animal studies is subjected to bias due to high heterogeneity of the data. Sources of heterogeneity of data obtained from animal studies can be placed into 2 categories: 1) variations in the animals used for the experiments (species, age, and sex) and 2) variations in study design and methods.

The use of different species and different models for induction of diseases in animals, as well as variations in timing, doses of drugs, and route of administration, result in several challenges to conducting meta-analysis; these diversities lead to unreliable estimation of efficacy and toxicity in animals and uncertain extrapolation of the models to the condition in humans.[29,75] For example, cross-species differences have been reported to result in a 54% to 84% variation in extrapolated predictions of survival

in humans[22] and substantial (>90%) heterogeneity values.[63] Another source of bias and heterogeneity is sex bias, that is, favoring the use of male over female animals or failing to report the sex of the animals studied;[11,92] sex-associated physiologic differences confound experimental results due to differences in disease susceptibility, prognosis, and response to drugs.[10] A survey of published papers in the field of neuroscience noted that 47% of reports failed to report animal sex in 2010 and 40% demonstrated male bias in 2014.[92] Translating male-biased animal experiments to humans may result in adverse consequences for the health of women.[11,61] High heterogeneity can also be related to differences in animal age.[36]

Using different methods to select animals for study, including randomization methods, and the group chosen for comparison (i.e., no treatment, placebo, vehicle), can also contribute to increased heterogeneity of the studies and are a major source of bias in meta-analysis.[75]

Publication bias. The validity of systematic reviews and meta-analyses of animal studies is highly affected by publication bias.[13] Unpublished studies, studies published in the gray literature, and selective dissemination of positive or significant results are major threats to validity when performing systematic reviews and meta-analyses.[66] Publication bias in the systematic reviews of animal experiments may confound treatment efficacies and estimated pooled effects sizes;[86] for example, more than 30% of the efficacy reported in some stroke studies has been attributed to publication bias.[80]

Methodologic weaknesses in systematic reviews and meta-analyses. Another factor contributing to the failure of translating animal data to humans are methodologic weaknesses in systematic reviews and meta-analyses of animal studies.[62] An increasing trend of massive production of unnecessary, misleading, and often conflicting systematic reviews and meta-analyses[44] raises concerns about obtaining pooled effect size from animal experiments by meta-analysis and extrapolating to humans. One group[62] conducting a survey on the methodologic quality of systematic reviews of animal studies criticized this approach due to poor methodologic features, including nonexplicit testable hypothesis (30%); lack of assessment of publication bias (17%), study quality (50%), and heterogeneity (33%); and missing meta-analysis for quantitative data synthesis (40%). Other colleagues, evaluating 266 published systematic reviews between 2009 to 2013, reported that the majority of these reviews did not assess methodologic quality (71%), heterogeneity (81%), or publication bias (87%).[66]

Potential misunderstanding in meta-analysis. Although systematic reviews and meta-analyses of animal studies can be useful tools for translating animal research and data into human practice, some potential misunderstandings around these practices[7,19] need careful consideration. These misunderstandings include: 1) meta-analyses are generally better than narrative reviews; 2) meta-analyses are more objective procedures; 3) a meta-analysis provides the highest level of evidence; 4) a risk-of-bias analysis resolves the bias; 5) random-effects models solve heterogeneity; 6) assuming homogeneity between studies when the statistical test fails to show heterogeneity; 7) funnel-plot asymmetry proves publication bias; and 8) meta-analyses speak for themselves. Meta-analyses also need interpretation that considers validity, heterogeneity, and clinical relevance.[19]

Potential steps toward solving these problems. Better reporting of animal experiments, providing more transparency regarding animal studies (e.g., registration of studies and data availability), and conducting high-quality systematic reviews and meta-analyses of animal data are some suggestions for addressing current challenges regarding translational process of animal experiments and their clinical usefulness. Indeed, development of unified reporting requirements and a system for registering animal experiments has been initiated by some groups.[29,73] Practical guidelines have been developed to improve the methodologic reporting of animal studies; these guidelines emphasize reporting of the sample-size estimation, whether and how animals were randomized, whether investigators were blinded to the treatment protocols, and data handling.[52]

One well-known guideline, ARRIVE (Animals in Research: Reporting In Vivo Experiments), has been developed by using the CONSORT Statement as its foundation;[47] this guide (accessible at: http://www.equator-network.org) provides a 20-item checklist describing the minimal required information that should be reported for animal studies: sample size, specific characteristics of animals (species, strain, sex, developmental stage, weight, genetic background, and modification status), details of housing, the experimental procedures (drug formulation and dose, site and route of administration, anesthesia and analgesia used, surgical procedure, and method of euthanasia), statistical tests, and analytical methods (details of methods used to reduce bias such as randomization, blinding and matching, if applicable).[47]

CAMARADES provides practical tools, including quality assessment tools and reporting guidelines to reduce bias and improve the quality of methods and reporting in animal research (http://www.dcn.ed.ac.uk/camarades/default.htm). Other research groups including SYRCLE (www.umcn.nl/Research/Departments/cdl/SYRCLE) are also involved in promoting standards for systematic reviews of preclinical studies. In addition, the use of evidence-based medicine standards, defined for human clinical trials, has been suggested as a helpful guide in making the results of animal experiments more robust and broadly applicable.[50] Authors can consult a recent paper regarding guidelines related to animal research reports and the Harmonized Animal Research Reporting Principle (HARRP).[71]

To make animal research valid and improve its usefulness in human practice, careful documentation and inclusion of all collected data (both published and unpublished), may be as supplementary data, is essential.[43] Optimal documentation that enables access to raw data and analytical codes (computational models, bioinformatics algorithms, and statistical methods) increases transparency and allows researchers to integrate and analyze multiple studies on the same topic and enhance reliability of animal research.[43]

Conclusion

Systematic reviews and meta-analyses of animal studies are considered robust and informative tools for translating basic sciences into clinical practice; however, they have been criticized for some limitations. In addition to poor methodologic quality and publication bias that contribute to the poor estimated effect sizes obtained from meta-analysis of animal data, the failure of animal models to adequately represent human disease and poor reporting of animal experiments are other causes of translation failure. Methodologic weaknesses of systematic reviews and meta-analyses conducted to summarize data from original animal studies are another cause for the failure in the translational process.

Researchers, therefore, should be cautious when conducting systematic reviews and meta-analyses and extrapolating the findings of animal research to the human clinical practice. The methodology of systematic review and meta-analysis needs to

be designed carefully and must follow available guidelines. In addition to focusing on the quality of animal experiments and their meta-analysis, the quality of clinical studies must be considered, because they also may contribute to the current translational failure rate. Finally, animal experiments are not alternatives to human trials but are complementary. Although, animal experiments are done to benefit humans, the results obtained from these studies cannot directly be translated to that end.

References

1. **Akhtar A**. 2015. The flaws and human harms of animal experimentation. Cambridge quarterly of healthcare ethics: CQ: The international journal of healthcare ethics committees **24:**407–419. Cambridge University Press. https://doi.org/10.1017/S0963180115000079

2. **Amore BM, Gibbs JP, Emery MG.** 2010. Application of in vivo animal models to characterize the pharmacokinetic and pharmacodynamic properties of drug candidates in discovery settings. Comb Chem High Throughput Screen **13:**207–218. https://doi.org/10.2174/138620710790596808.

3. **Anonymous.** 2011. The contribution of animal science to the medical revolution. Heart Views **12:**43–44.

4. **Arora T, Mehta AK, Joshi V, Mehta KD, Rathor N, Mediratta PK, Sharma KK.** 2011. Substitute of Animals in drug research: an approach towards fulfillment of 4R's. Indian J Pharm Sci **73:**1–6. https://doi.org/10.4103/0250-474X.89750.

5. **Bailey J, Thew M, Balls M.** 2014. An analysis of the use of animal models in predicting human toxicology and drug safety. Altern Lab Anim **42:**181–199. https://doi.org/10.1177/026119291404200306.

6. **Bailey J, Thew M, Balls M.** 2015. Predicting human drug toxicity and safety via animal tests: can any one species predict drug toxicity in any other, and do monkeys help? Altern Lab Anim **43:**393–403. https://doi.org/10.1177/026119291504300607.

7. **Bangert-Drowns RL.** 1995. Misunderstanding meta-analysis. Eval Health Prof **18:**304–314. https://doi.org/10.1177/016327879501800305.

8. **Barré-Sinoussi F, Montagutelli X.** 2015. Animal models are essential to biological research: issues and perspectives. Future Sci OA **1:**1–3. https://doi.org/10.4155/fso.15.63.

9. **Bastasch M.** [Internet]. 2020. [Cited 25 January 2020]. Available at: https://www.globalanimal.org/2013/10/07/guilty-government-practices/111386/#.

10. **Becker JB, Arnold AP, Berkley KJ, Blaustein JD, Eckel LA, Hampson E, Herman JP, Marts S, Sadee W, Steiner M, Taylor J, Young E.** 2005. Strategies and methods for research on sex differences in brain and behavior. Endocrinology **146:**1650–1673. https://doi.org/10.1210/en.2004-1142.

11. **Beery AK, Zucker I.** 2011. Sex bias in neuroscience and biomedical research. Neurosci Biobehav Rev **35:**565–572. https://doi.org/10.1016/j.neubiorev.2010.07.002.

12. **Bracken MB.** 2009. Why animal studies are often poor predictors of human reactions to exposure. J R Soc Med **102:**120–122. https://doi.org/10.1258/jrsm.2008.08k033.

13. **Briel M, Muller KF, Meerpohl JJ, 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. Syst Rev **2:**1–5. https://doi.org/10.1186/2046-4053-2-23.

14. **Burns PB, Rohrich RJ, Chung KC.** 2011. The levels of evidence and their role in evidence-based medicine. Plast Reconstr Surg **128:**305–310. https://doi.org/10.1097/PRS.0b013e318219c171.

15. **Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR.** 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci **14:**365–376. https://doi.org/10.1038/nrn3475. Erratum: Nat Rev Neurosci 2013. **14:**451.

16. **Comroe JH Jr, Dripps RD.** 1976. Scientific basis for the support of biomedical science. Science **192:**105–111. https://doi.org/10.1126/science.769161.

17. **Cook DJ, Mulrow CD, Haynes RB.** 1997. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med **126:**376–380. https://doi.org/10.7326/0003-4819-126-5-199703010-00006.

18. **de Vries RB, Wever KE, Avey MT, Stephens ML, Sena ES, Leenaars M.** 2014. The usefulness of systematic reviews of animal experiments for the design of preclinical and clinical studies. ILAR J **55:**427–437. https://doi.org/10.1093/ilar/ilu043.

19. **Dekkers OM.** 2018. Meta-analysis: Key features, potentials and misunderstandings. Res Pract Thromb Haemost **2:**658–663. https://doi.org/10.1002/rth2.12153.

20. **Delgado-Rodríguez M.** 2001. Glossary on meta-analysis. J Epidemiol Community Health **55:**534–536. https://doi.org/10.1136/jech.55.8.534.

21. **Dickersin K.** 1990. The existence of publication bias and risk factors for its occurrence. JAMA **263:**1385–1389. https://doi.org/10.1001/jama.1990.03440100097014.

22. **Fay MP, Follmann DA, Lynn F, Schiffer JM, Stark GV, Kohberger R, Quinn CP, Nuzum EO.** 2012. Anthrax vaccine-induced antibodies provide cross-species prediction of survival to aerosol challenge. Sci Transl Med **4:**1–26. https://doi.org/10.1126/scitranslmed.3004073.

23. **Festing MF.** 2018. On determining sample size in experiments involving laboratory animals. Lab Anim **52:**341–350. https://doi.org/10.1177/0023677217738268.

24. **Flórez-Vargas O, Brass A, Karystianis G, Bramhall M, Stevens R, Cruickshank S, Nenadic G.** 2016. Bias in the reporting of sex and age in biomedical research on mouse models. eLife **5:**1–14. https://doi.org/10.7554/eLife.13615.

25. **Fox JG, Bennett BT.** 2015. Laboratory animal medicine: historical perspectives. p 1–21. In: Fox JG, Anderson LC, Otto GM, Pritchett-Corning KR, Whary MT, editors. Laboratory animal medicine, Boston (MA): Academic Press. https://doi.org/10.1016/B978-0-12-409527-4.00001-8.

26. **Franco NH.** 2013. Animal experiments in biomedical research: a historical perspective. Animals (Basel) **3:**238–273. https://doi.org/10.3390/ani3010238.

27. **Goodman J, Chandna A, Roe K.** 2015. Trends in animal use at US research facilities. J Med Ethics **41:**567–569. https://doi.org/10.1136/medethics-2014-102404.

28. **Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW Jr, Atkins D, Meerpohl J, Schünemann HJ.** 2011. GRADE guidelines: 4. Rating the quality of evidence–study limitations (risk of bias). J Clin Epidemiol **64:**407–415. https://doi.org/10.1016/j.jclinepi.2010.07.017.

29. **Hackam DG.** 2007. Translating animal research into clinical benefit. BMJ **334:**163–164. https://doi.org/10.1136/bmj.39104.362951.80.

30. **Hackam DG, Redelmeier DA.** 2006. Translation of research evidence from animals to humans. JAMA **296:**1731–1732. https://doi.org/10.1001/jama.296.14.1731.

31. **Haidich AB.** 2010. Meta-analysis in medical research. Hippokratia **14 (Suppl 1):**29–37.

32. **Hajar R.** 2011. Animal testing and medicine. Heart views **12:**42–42. https://doi.org/10.4103/1995-705X.81548 PubMed

33. **Hall ED, Traystman RJ.** 2009. Role of animal studies in the design of clinical trials. Front Neurol Neurosci **25:**10–33. https://doi.org/10.1159/000209470.

34. **Higgins JPT, Green S, editors.** 2008. Cochrane handbook for systematic reviews of interventions. 1st ed. Chichester (United Kingdom): John Wiley and Sons.

35. **Hirst JA, Howick J, Aronson JK, Roberts N, Perera R, Koshiaris C, Heneghan C.** 2014. The need for randomization in animal trials: an overview of systematic reviews. PLoS One **9:**1–11. https://doi.org/10.1371/journal.pone.0098856.

36. **Hooijmans CR, IntHout J, Ritskes-Hoitinga M, Rovers MM.** 2014. Meta-analyses of animal studies: an introduction of a valuable instrument to further improve healthcare. ILAR J **55:**418–426. https://doi.org/10.1093/ilar/ilu042.

37. **Hooijmans CR, Ritskes-Hoitinga M.** 2013. Progress in using systematic reviews of animal studies to improve translational research. PLoS Med **10:**1–4. https://doi.org/10.1371/journal.pmed.1001482.

38. **Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW.** 2014. SYRCLE's risk of bias tool for animal studies. BMC Med Res Methodol **14:**1–9. https://doi.org/10.1186/1471-2288-14-43.

39. **Horn J, de Haan RJ, Vermeulen M, Luiten PG, Limburg M.** 2001. Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. Stroke **32:**2433–2438. https://doi.org/10.1161/hs1001.096009.

40. **Hudson-Shore M.** 2016. Statistics of scientific procedures on living animals Great Britain 2015— highlighting an ongoing upward trend in animal use and missed opportunities. Altern Lab Anim **44:**569–580. https://doi.org/10.1177/026119291604400606.

41. **Hutton JL, Williamson PR.** 2000. Bias in meta-analysis due to outcome variable selection within studies. J R Stat Soc Ser C Appl Stat **49:**359–370. https://doi.org/10.1111/1467-9876.00197.

42. **Ioannidis JPA.** 2008. Why most discovered true associations are inflated. Epidemiology **19:**640–648. https://doi.org/10.1097/EDE.0b013e31818131e7. Erratum in: Epidemiology 2009 **20:**629.

43. **Ioannidis JP.** 2012. Extrapolating from animals to humans. Sci Transl Med **4:**1–5. https://doi.org/10.1126/scitranslmed.3004631.

44. **Ioannidis JPA.** 2016. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q **94:**485–514. https://doi.org/10.1111/1468-0009.12210.

45. **Khan K, Kunz R, Kleijnen J, Antes G.** 2011. Systematic reviews to support evidence-based medicine 2nd ed. London. CRC Press.

46. **Khan KS, Kunz R, Kleijnen J, Antes G.** 2003. Five steps to conducting a systematic review. J R Soc Med **96:**118–121. https://doi.org/10.1177/014107680309600304.

47. **Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG.** 2010. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol **8:**1–5. https://doi.org/10.1371/journal.pbio.1000412.

48. **Kilkenny C, Parsons N, Kadyszewski E, Festing MF, Cuthill IC, Fry D, Hutton J, Altman DG.** 2009. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. PLoS One **4:**1–11. https://doi.org/10.1371/journal.pone.0007824.

49. **Knight A.** 2007. Systematic reviews of animal experiments demonstrate poor human clinical and toxicological utility. Altern Lab Anim **35:**641–659. https://doi.org/10.1177/026119290703500610.

50. **Knight A.** 2008. Systematic reviews of animal experiments demonstrate poor contributions toward human healthcare. Rev Recent Clin Trials **3:**89–96. https://doi.org/10.2174/157488708784223844.

51. **Kolar R.** 2006. Animal experimentation. Sci Eng Ethics **12:**111–122. https://doi.org/10.1007/s11948-006-0011-1.

52. **Landis SC, Amara SG, Asadullah K, Austin CP, Blumenstein R, Bradley EW, Crystal RG, Darnell RB, Ferrante RJ, Fillit H, Finkelstein R, Fisher M, Gendelman HE, Golub RM, Goudreau JL, Gross RA, Gubitz AK, Hesterlee SE, Howells DW, Huguenard J, Kelner K, Koroshetz W, Krainc D, Lazic SE, Levine MS, Macleod MR, McCall JM, Moxley RT 3rd, Narasimhan K, Noble LJ, Perrin S, Porter JD, Steward O, Unger E, Utz U, Silberberg SD.** 2012. A call for transparent reporting to optimize the predictive value of preclinical research. Nature **490:**187–191. https://doi.org/10.1038/nature11556.

53. **Leenaars CHC, Kouwenaar C, Stafleu FR, Bleich A, Ritskes-Hoitinga M, De Vries RBM, Meijboom FLB.** 2019. Animal to human translation: a systematic scoping review of reported concordance rates. J Transl Med **17:**1–22. https://doi.org/10.1186/s12967-019-1976-2.

54. **Leenaars M, Hooijmans CR, van Veggel N, ter Riet G, Leeflang M, Hooft L, van der Wilt GJ, Tillema A, Ritskes-Hoitinga M.** 2012. A step-by-step guide to systematically identify all relevant animal studies. Lab Anim **46:**24–31. https://doi.org/10.1258/la.2011.011087.

55. **Leenaars M, Savenije B, Nagtegaal A, van der Vaart L, Ritskes-Hoitinga M.** 2009. Assessing the search for and implementation of the Three Rs: a survey among scientists. Altern Lab Anim **37:**297–303. https://doi.org/10.1177/026119290903700312.

56. **Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D.** 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ **339:**b2700. https://doi.org/10.1136/bmj.b2700.

57. **Macleod MR, van der Worp HB, Sena ES, Howells DW, Dirnagl U, Donnan GA.** 2008. Evidence for the efficacy of NXY-059 in experimental focal cerebral ischaemia is confounded by study quality. Stroke **39:**2824–2829. https://doi.org/10.1161/STROKEAHA.108.515957.

58. **Mak IW, Evaniew N, Ghert M.** 2014. Lost in translation: animal models and clinical trials in cancer treatment. Am J Transl Res **6:**114–118.

59. **Matthews RAJ.** 2008. Medical progress depends on animal models—doesn't it? J R Soc Med **101:**95–98. https://doi.org/10.1258/jrsm.2007.070164.

60. **Maurer KJ, Quimby FW.** 2015**.** Animal models in biomedical research, p 1497–1534. In: Fox JG, Anderson LC, Otto GM, Pritchett-Corning KR, Whary MT, editors. Laboratory animal medicine, 3rd ed. Boston (MA): Academic Press.

61. **McCullough LD, de Vries GJ, Miller VM, Becker JB, Sandberg K, McCarthy MM.** 2014. NIH initiative to balance sex of animals in preclinical studies: generative questions to guide policy, implementation, and metrics. Biol Sex Differ **5:**1–7. https://doi.org/10.1186/s13293-014-0015-5.

62. **Mignini LE, Khan KS.** 2006. Methodological quality of systematic reviews of animal studies: a survey of reviews of basic research. BMC Med Res Methodol **6:**1–6. https://doi.org/10.1186/1471-2288-6-10.

63. **Moja L, Pecoraro V, Ciccolallo L, Dall'Olmo L, Virgili G, Garattini S.** 2014. Flaws in animal studies exploring statins and impact on meta-analysis. Eur J Clin Invest **44:**597–612. https://doi.org/10.1111/eci.12264.

64. **Møller AM, Myles PS.** 2016. What makes a good systematic review and meta-analysis? Br J Anaesth **117:**428–430. https://doi.org/10.1093/bja/aew264.

65. **Morgan SJ, Elangbam CS, Berens S, Janovitz E, Vitsky A, Zabka T, Conour L.** 2013. Use of animal models of human disease for nonclinical safety assessment of novel pharmaceuticals. Toxicol Pathol **41:**508–518. https://doi.org/10.1177/0192623312457273.

66. **Mueller KF, Briel M, Strech D, Meerpohl JJ, Lang B, Motschall E, Gloy V, Lamontagne F, Bassler D.** 2014. Dissemination bias in systematic reviews of animal research: a systematic review. PLoS One **9:**1–15. https://doi.org/10.1371/journal.pone.0116016.

67. **Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z.** 2018. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol **18:**1–9. https://doi.org/10.1186/s12874-017-0468-4.

68. **Nikolakopoulou A, Mavridis D, Salanti G.** 2014. How to interpret meta-analysis models: fixed effect and random effects meta-analyses. Evid Based Ment Health **17:**64. https://doi.org/10.1136/eb-2014-101794.

69. **O'Collins VE, Macleod MR, Donnan GA, Horky LL, van der Worp BH, Howells DW.** 2006. 1,026 experimental treatments in acute stroke. Ann Neurol **59:**467–477. https://doi.org/10.1002/ana.20741.

70. **O'Rourke K.** 2007. An historical perspective on meta-analysis: dealing quantitatively with varying study results. J R Soc Med **100:**579–582. https://doi.org/10.1177/0141076807100012020.

71. **Osborne N, Avey MT, Anestidou L, Ritskes-Hoitinga M**. 2018. Improving animal research reporting standards: HARRP, the first step of a unified approach by ICLAS to improve animal research reporting standards worldwide. EMBO Rep **19:**1–5. https://doi.org/10.15252/embr.201846069.

72. **Pasquali P.** 2018. The importance of animal models in research. Res Vet Sci **118:**144–145. https://doi.org/10.1016/j.rvsc.2018.02.001.

73. **Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, Macleod M, Mignini LE, Jayaram P, Khan KS.** 2007. Comparison of treatment effects between animal experiments and clinical trials: systematic review. BMJ **334:**1–6. https://doi.org/10.1136/bmj.39048.407928.BE.

74. **Pound P, Bracken MB.** 2014. Is animal research sufficiently evidence based to be a cornerstone of biomedical research? BMJ **348:**g3387. https://doi.org/10.1136/bmj.g3387.

75. **Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I.** 2004. Where is the evidence that animal research benefits humans? BMJ **328:**514–517. https://doi.org/10.1136/bmj.328.7438.514.

76. **Prankel S.** 2011. The costs and benefits of animal experiments: an evaluation with bias. By Andrew Knight. Palgrave Macmillan: Basingstoke, Hampshire, UK,; Hardcover, 272 pp; ISBN 978-0-230-57686-5; Paperback, ISBN: 978-0-230-57687-2. Molecular Diversity Preservation International. Animals (Basel) **2:**25–26. https://doi.org/10.3390/ani2010025.

77. **Ritskes-Hoitinga M, Leenaars M, Avey M, Rovers M, Scholten R.** 2014. Systematic reviews of preclinical animal studies can make significant contributions to health care and more transparent translational medicine. Cochrane Database Syst Rev. https://doi.org/10.1002/14651858.ED000078.

78. **Sandercock P, Roberts I.** 2002. Systematic reviews of animal experiments. Lancet **360:**586. https://doi.org/10.1016/S0140-6736(02)09812-4.

79. **Sena ES, Currie GL, McCann SK, Macleod MR, Howells DW.** 2014. Systematic reviews and meta-analysis of preclinical studies: why perform them and how to appraise them critically. J Cereb Blood Flow Metab **34:**737–742. https://doi.org/10.1038/jcbfm.2014.28.

80. **Sena ES, van der Worp HB, Bath PMW, Howells DW, Macleod MR.** 2010. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. PLoS Biol **8:**1–8. https://doi.org/10.1371/journal.pbio.1000344.

81. **Smith AJ, Allen T.** 2008. Use of databases, information centers and guidelines when planning research that may involve animals. Anim Welf **14:**347–359.

82. **Smith R.** 1987. Comroe and Dripps revisited. Br Med J (Clin Res Ed) **295:**1404–1407. https://doi.org/10.1136/bmj.295.6610.1404.

83. **Sweet M, Moynihan R.** [Internet]. 2007. Improving population health; the uses of systematic reviews. [Cited 25 January 2020]. Available at: https://www.milbank.org/publications/improving-population-health-the-uses-of-systematic-reviews/.

84. **Taylor K, Gordon N, Langley G, Higgins W.** 2008. Estimates for worldwide laboratory animal use in 2005. Altern Lab Anim **36:**327–342. https://doi.org/10.1177/026119290803600310.

85. **Tsilidis KK, Panagiotou OA, Sena ES, Aretouli E, Evangelou E, Howells DW, Al-Shahi Salman R, Macleod MR, Ioannidis JPA.** 2013. Evaluation of excess significance bias in animal studies of neurological diseases. PLoS Biol **11:**1–10. https://doi.org/10.1371/journal.pbio.1001609.

86. **van der Worp HB, Howells DW, Sena ES, Porritt MJ, Rewell S, O'Collins V, Macleod MR.** 2010. Can animal models of disease reliably inform human studies? PLoS Med **7:**1–8. https://doi.org/10.1371/journal.pmed.1000245.

87. **van der Worp HB, van Gijn J.** 2007. Clinical practice. Acute ischemic stroke. N Engl J Med **357:**572–579. https://doi.org/10.1056/NEJMcp072057.

88. **Van Norman GA.** 2019. Limitations of animal studies for predicting toxicity in clinical trials: Is it time to rethink our current approach? JACC Basic Transl Sci **4:**845–854. https://doi.org/10.1016/j.jacbts.2019.10.008.

89. **Vesterinen HM, Sena ES, Egan KJ, Hirst TC, Churolov L, Currie GL, Antonic A, Howells DW, Macleod MR.** 2014. Meta-analysis of data from animal studies: A practical guide. J Neurosci Methods **221:**92–102. https://doi.org/10.1016/j.jneumeth.2013.09.010. Erratum in: J Neurosci Methods 2016 **259:**156.

90. **Wang Q, Liao J, Hair K, Bannach-Brown A, Bahor Z, Currie GL, McCann SK, Howells DW, Sena ES, Macleod MR.** 2018. Estimating the statistical performance of different approaches to meta-analysis of data from animal studies in identifying the impact of aspects of study design. bioRxiv**:**256776. https://doi.org/10.1101/256776.

91. **Whiting P, Wolff R, Mallett S, Simera I, Savović J.** 2017. A proposed framework for developing quality assessment tools. Syst Rev **6:**1–9. https://doi.org/10.1186/s13643-017-0604-6.

92. **Will TR, Proaño SB, Thomas AM, Kunz LM, Thompson KC, Ginnari LA, Jones CH, Lucas S-C, Reavis EM, Dorris DM, Meitzen J.** 2017. Problems and progress regarding sex bias and omission in neuroscience research. eNeuro **4:**1–10.