

The Flaws and Human Harms of Animal Experimentation

AYSHA AKHTAR

Abstract: Nonhuman animal (“animal”) experimentation is typically defended by arguments that it is reliable, that animals provide sufficiently good models of human biology and diseases to yield relevant information, and that, consequently, its use provides major human health benefits. I demonstrate that a growing body of scientific literature critically assessing the validity of animal experimentation generally (and animal modeling specifically) raises important concerns about its reliability and predictive value for human outcomes and for understanding human physiology. The unreliability of animal experimentation across a wide range of areas undermines scientific arguments in favor of the practice. Additionally, I show how animal experimentation often significantly harms humans through misleading safety studies, potential abandonment of effective therapeutics, and direction of resources away from more effective testing methods. The resulting evidence suggests that the collective harms and costs to humans from animal experimentation outweigh potential benefits and that resources would be better invested in developing human-based testing methods.

Keywords: animal research; medical testing; human health; human ethics; drug development; animal ethics

Introduction

Annually, more than 115 million animals are used worldwide in experimentation or to supply the biomedical industry.¹ Nonhuman animal (hereafter “animal”) experimentation falls under two categories: basic (i.e., investigation of basic biology and human disease) and applied (i.e., drug research and development and toxicity and safety testing). Regardless of its categorization, animal experimentation is intended to inform human biology and health sciences and to promote the safety and efficacy of potential treatments. Despite its use of immense resources, the animal suffering involved, and its impact on human health, the question of animal experimentation’s efficacy has been subjected to little systematic scrutiny.²

Although it is widely accepted that medicine should be *evidence based*, animal experimentation as a means of informing human health has generally not been held, in practice, to this standard. This fact makes it surprising that animal experimentation is typically viewed as the default and gold standard of preclinical testing and is generally supported without critical examination of its validity. A survey published in 2008 of anecdotal cases and statements given in support of animal experimentation demonstrates how it has not and could not be validated as a necessary step in biomedical research, and the survey casts doubt on its predictive value.³

I am deeply indebted to David DeGrazia, Tom Beauchamp, and John Pippin for their careful review and helpful comments. The opinions expressed here are those of the author and do not represent the official position of the U.S. Food and Drug Administration or the U.S. government.

Aysha Akhtar, M.D., M.P.H., is a neurologist and preventive medicine specialist and Fellow at the Oxford Centre for Animal Ethics, Oxford, United Kingdom.

Cambridge Quarterly of Healthcare Ethics (2015), 24, 407–419.

© Cambridge University Press 2015. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.
doi:10.1017/S0963180115000079

I show that animal experimentation is poorly predictive of human outcomes,⁴ that it is unreliable across a wide category of disease areas,⁵ and that existing literature demonstrates the unreliability of animal experimentation, thereby undermining scientific arguments in its favor. I further show that the collective harms that result from an unreliable practice tip the ethical scale of harms and benefits against continuation in much, if not all, of experimentation involving animals.⁶

Problems of Successful Translation to Humans of Data from Animal Experimentation

Although the unreliability and limitations of animal experimentation have increasingly been acknowledged, there remains a general confidence within much of the biomedical community that they can be overcome.⁷ However, three major conditions undermine this confidence and explain why animal experimentation, regardless of the disease category studied, fails to reliably inform human health: (1) the effects of the laboratory environment and other variables on study outcomes, (2) disparities between animal models of disease and human diseases, and (3) species differences in physiology and genetics. I argue for the critical importance of each of these conditions.

The Influence of Laboratory Procedures and Environments on Experimental Results

Laboratory procedures and conditions exert influences on animals' physiology and behaviors that are difficult to control and that can ultimately impact research outcomes. Animals in laboratories are involuntarily placed in artificial environments, usually in windowless rooms, for the duration of their lives. Captivity and the common features of biomedical laboratories—such as artificial lighting, human-produced noises, and restricted housing environments—can prevent species-typical behaviors, causing distress and abnormal behaviors among animals.⁸ Among the types of laboratory-generated distress is the phenomenon of contagious anxiety.⁹ Cortisone levels rise in monkeys watching other monkeys being restrained for blood collection.¹⁰ Blood pressure and heart rates elevate in rats watching other rats being decapitated.¹¹ Routine laboratory procedures, such as catching an animal and removing him or her from the cage, in addition to the experimental procedures, cause significant and prolonged elevations in animals' stress markers.¹² These stress-related changes in physiological parameters caused by the laboratory procedures and environments can have significant effects on test results.¹³ Stressed rats, for example, develop chronic inflammatory conditions and intestinal leakage, which add variables that can confound data.¹⁴

A variety of conditions in the laboratory cause changes in neurochemistry, genetic expression, and nerve regeneration.¹⁵ In one study, for example, mice were genetically altered to develop aortic defects. Yet, when the mice were housed in larger cages, those defects almost completely disappeared.¹⁶ Providing further examples, typical noise levels in laboratories can damage blood vessels in animals, and even the type of flooring on which animals are tested in spinal cord injury experiments can affect whether a drug shows a benefit.¹⁷

In order to control for potential confounders, some investigators have called for standardization of laboratory settings and procedures.¹⁸ One notable effort was

made by Crabbe et al. in their investigation of the potential confounding influences of the laboratory environment on six mouse behaviors that are commonly studied in neurobehavioral experiments. Despite their “extraordinary lengths to equate test apparatus, testing protocols, and all possible features of animal husbandry” across three laboratories, there were systematic differences in test results in these labs.¹⁹ Additionally, different mouse strains varied markedly in all behavioral tests, and for some tests the magnitude of genetic differences depended on the specific testing laboratory. The results suggest that there are important influences of environmental conditions and procedures specific to individual laboratories that can be difficult—perhaps even impossible—to eliminate. These influences can confound research results and impede extrapolation to humans.

The Discordance between Human Diseases and Animal Models of Diseases

The lack of sufficient congruence between animal models and human diseases is another significant obstacle to translational reliability. Human diseases are typically artificially induced in animals, but the enormous difficulty of reproducing anything approaching the complexity of human diseases in animal models limits their usefulness.²⁰ Even if the design and conduct of an animal experiment are sound and standardized, the translation of its results to the clinic may fail because of disparities between the animal experimental model and the human condition.²¹

Stroke research presents one salient example of the difficulties in modeling human diseases in animals. Stroke is relatively well understood in its underlying pathology. Yet accurately modeling the disease in animals has proven to be an exercise in futility. To address the inability to replicate human stroke in animals, many assert the need to use more standardized animal study design protocols. This includes the use of animals who represent both genders and wide age ranges, who have comorbidities and preexisting conditions that occur naturally in humans, and who are consequently given medications that are indicated for human patients.²² In fact, a set of guidelines, named STAIR, was implemented by a stroke roundtable in 1999 (and updated in 2009) to standardize protocols, limit the discrepancies, and improve the applicability of animal stroke experiments to humans.²³ One of the most promising stroke treatments later to emerge was NXY-059, which proved effective in animal experiments. However, the drug failed in clinical trials, despite the fact that the set of animal experiments on this drug was considered the poster child for the new experimental standards.²⁴ Despite such vigorous efforts, the development of STAIR and other criteria has yet to make a recognizable impact in clinical translation.²⁵

Under closer scrutiny, it is not difficult to surmise why animal stroke experiments fail to successfully translate to humans even with new guidelines. Standard stroke medications will likely affect different species differently. There is little evidence to suggest that a female rat, dog, or monkey sufficiently reproduces the physiology of a human female. Perhaps most importantly, reproducing the preexisting conditions of stroke in animals proves just as difficult as reproducing stroke pathology and outcomes. For example, most animals don't naturally develop significant atherosclerosis, a leading contributor to ischemic stroke. In order to reproduce the effects of atherosclerosis in animals, researchers clamp their blood vessels or artificially insert blood clots. These interventions, however, do not replicate the elaborate pathology of atherosclerosis and its underlying causes. Reproducing human

diseases in animals requires reproducing the *predisposing* diseases, also a formidable challenge. The inability to reproduce the disease in animals so that it is congruent in relevant respects with human stroke has contributed to a high failure rate in drug development. More than 114 potential therapies initially tested in animals failed in human trials.²⁶

Further examples of repeated failures based on animal models include drug development in cancer, amyotrophic lateral sclerosis (ALS), traumatic brain injury (TBI), Alzheimer's disease (AD), and inflammatory conditions. Animal cancer models in which tumors are artificially induced have been the basic translational model used to study key physiological and biochemical properties in cancer onset and propagation and to evaluate novel treatments. Nevertheless, significant limitations exist in the models' ability to faithfully mirror the complex process of human carcinogenesis.²⁷ These limitations are evidenced by the high (among the highest of any disease category) clinical failure rate of cancer drugs.²⁸ Analyses of common mice ALS models demonstrate significant differences from human ALS.²⁹ The inability of animal ALS models to predict beneficial effects in humans with ALS is recognized.³⁰ More than twenty drugs have failed in clinical trials, and the only U.S. Food and Drug Administration (FDA)-approved drug to treat ALS is Riluzole, which shows notably marginal benefit on patient survival.³¹ Animal models have also been unable to reproduce the complexities of human TBI.³² In 2010, Maas et al. reported on 27 large Phase 3 clinical trials and 6 unpublished trials in TBI that all failed to show human benefit after showing benefit in animals.³³ Additionally, even after success in animals, around 172 and 150 drug development failures have been identified in the treatment of human AD³⁴ and inflammatory diseases,³⁵ respectively.

The high clinical failure rate in drug development across all disease categories is based, at least in part, on the inability to adequately model human diseases in animals and the poor predictability of animal models.³⁶ A notable systematic review, published in 2007, compared animal experimentation results with clinical trial findings across interventions aimed at the treatment of head injury, respiratory distress syndrome, osteoporosis, stroke, and hemorrhage.³⁷ The study found that the human and animal results were in accordance only half of the time. In other words, the animal experiments were no more likely than a flip of the coin to predict whether those interventions would benefit humans.

In 2004, the FDA estimated that 92 percent of drugs that pass preclinical tests, including "pivotal" animal tests, fail to proceed to the market.³⁸ More recent analysis suggests that, despite efforts to improve the predictability of animal testing, the failure rate has actually increased and is now closer to 96 percent.³⁹ The main causes of failure are lack of effectiveness and safety problems that were not predicted by animal tests.⁴⁰

Usually, when an animal model is found wanting, various reasons are proffered to explain what went wrong—poor methodology, publication bias, lack of pre-existing disease and medications, wrong gender or age, and so on. These factors certainly require consideration, and recognition of each potential difference between the animal model and the human disease motivates renewed efforts to eliminate these differences. As a result, scientific progress is sometimes made by such efforts. However, the high failure rate in drug testing and development, despite attempts to improve animal testing, suggests that these efforts remain insufficient to overcome the obstacles to successful translation that are inherent to the use of animals.

Too often ignored is the well-substantiated idea that these models are, for reasons summarized here, intrinsically lacking in relevance to, and thus highly unlikely to yield useful information about, human diseases.⁴¹

Interspecies Differences in Physiology and Genetics

Ultimately, even if considerable congruence were shown between an animal model and its corresponding human disease, interspecies differences in physiology, behavior, pharmacokinetics, and genetics would significantly limit the reliability of animal studies, even after a substantial investment to improve such studies. In spinal cord injury, for example, drug testing results vary according to which species and even which strain within a species is used, because of numerous interspecies and interstrain differences in neurophysiology, anatomy, and behavior.⁴² The micropathology of spinal cord injury, injury repair mechanisms, and recovery from injury varies greatly among different strains of rats and mice. A systematic review found that even among the most standardized and methodologically superior animal experiments, testing results assessing the effectiveness of methylprednisolone for spinal cord injury treatment varied considerably among species.⁴³ This suggests that factors inherent to the use of animals account for some of the major differences in results.

Even rats from the same strain but purchased from different suppliers produce different test results.⁴⁴ In one study, responses to 12 different behavioral measures of pain sensitivity, which are important markers of spinal cord injury, varied among 11 strains of mice, with no clear-cut patterns that allowed prediction of how each strain would respond.⁴⁵ These differences influenced how the animals responded to the injury and to experimental therapies. A drug might be shown to help one strain of mice recover but not another. Despite decades of using animal models, not a single neuroprotective agent that ameliorated spinal cord injury in animal tests has proven efficacious in clinical trials to date.⁴⁶

Further exemplifying the importance of physiological differences among species, a 2013 study reported that the mouse models used extensively to study human inflammatory diseases (in sepsis, burns, infection, and trauma) have been misleading. The study found that mice differ greatly from humans in their responses to inflammatory conditions. Mice differed from humans in what genes were turned on and off and in the timing and duration of gene expression. The mouse models even differed from one another in their responses. The investigators concluded that “our study supports higher priority to focus on the more complex human conditions rather than relying on mouse models to study human inflammatory disease.”⁴⁷ The different genetic responses between mice and humans are likely responsible, at least in part, for the high drug failure rate. The authors stated that every one of almost 150 clinical trials that tested candidate agents’ ability to block inflammatory responses in critically ill patients failed.

Wide differences have also become apparent in the regulation of the same genes, a point that is readily seen when observing differences between human and mouse livers.⁴⁸ Consistent phenotypes (observable physical or biochemical characteristics) are rarely obtained by modification of the same gene, even among different strains of mice.⁴⁹ Gene regulation can substantially differ among species and may be as important as the presence or absence of a specific gene. Despite the high degree of genome conservation, there are critical differences in the order and function of

genes among species. To use an analogy: as pianos have the same keys, humans and other animals share (largely) the same genes. Where we mostly differ is in the way the genes or keys are expressed. For example, if we play the keys in a certain order, we hear Chopin; in a different order, we hear Ray Charles; and in yet a different order, it's Jerry Lee Lewis. In other words, the same keys or genes are expressed, but their different orders result in markedly different outcomes.

Recognizing the inherent genetic differences among species as a barrier to translation, researchers have expressed considerable enthusiasm for genetically modified (GM) animals, including transgenic mice models, wherein human genes are inserted into the mouse genome. However, if a human gene is expressed in mice, it will likely function differently from the way it functions in humans, being affected by physiological mechanisms that are unique in mice. For example, a crucial protein that controls blood sugar in humans is missing in mice.⁵⁰ When the human gene that makes this protein was expressed in genetically altered mice, it had the opposite effect from that in humans: it caused *loss* of blood sugar control in mice. Use of GM mice has failed to successfully model human diseases and to translate into clinical benefit across many disease categories.⁵¹ Perhaps the primary reason why GM animals are unlikely to be much more successful than other animal models in translational medicine is the fact that the "humanized" or altered genes are still in nonhuman animals.

In many instances, nonhuman primates (NHPs) are used instead of mice or other animals, with the expectation that NHPs will better mimic human results. However, there have been sufficient failures in translation to undermine this optimism. For example, NHP models have failed to reproduce key features of Parkinson's disease, both in function and in pathology.⁵² Several therapies that appeared promising in both NHPs and rat models of Parkinson's disease showed disappointing results in humans.⁵³ The campaign to prescribe hormone replacement therapy (HRT) in millions of women to prevent cardiovascular disease was based in large part on experiments on NHPs. HRT is now known to *increase* the risk of these diseases in women.⁵⁴

HIV/AIDS vaccine research using NHPs represents one of the most notable failures in animal experimentation translation. Immense resources and decades of time have been devoted to creating NHP (including chimpanzee) models of HIV. Yet all of about 90 HIV vaccines that succeeded in animals failed in humans.⁵⁵ After HIV vaccine gp120 failed in clinical trials, despite positive outcomes in chimpanzees, a *BMJ* article commented that important differences between NHPs and humans with HIV misled researchers, taking them down unproductive experimental paths.⁵⁶ Gp120 failed to neutralize HIV grown and tested in cell culture. However, because the serum protected chimpanzees from HIV infection, two Phase 3 clinical trials were undertaken⁵⁷—a clear example of how expectations that NHP data are more predictive than data from other (in this case, cell culture) testing methods are unproductive and harmful. Despite the repeated failures, NHPs (though not chimpanzees or other great apes) remain widely used for HIV research.

The implicit assumption that NHP (and indeed any animal) data are reliable has also led to significant and unjustifiable human suffering. For example, clinical trial volunteers for gp120 were placed at unnecessary risk of harm because of unfounded confidence in NHP experiments. Two landmark studies involving thousands of menopausal women being treated with HRT were terminated early because of

increased stroke and breast cancer risk.⁵⁸ In 2003, Elan Pharmaceuticals was forced to prematurely terminate a Phase 2 clinical trial when an investigational AD vaccine was found to cause brain swelling in human subjects. No significant adverse effects were detected in GM mice or NHPs.⁵⁹

In another example of human suffering resulting from animal experimentation, six human volunteers were injected with an immunomodulatory drug, TGN 1412, in 2006.⁶⁰ Within minutes of receiving the experimental drug, all volunteers suffered a severe adverse reaction resulting from a life-threatening cytokine storm that led to catastrophic systemic organ failure. The compound was designed to dampen the immune system, but it had the *opposite* effect in humans. Prior to this first human trial, TGN 1412 was tested in mice, rabbits, rats, and NHPs with no ill effects. NHPs also underwent repeat-dose toxicity studies and were given 500 times the human dose for at least four consecutive weeks.⁶¹ None of the NHPs manifested the ill effects that humans showed almost immediately after receiving minute amounts of the test drug. Cynomolgus and rhesus monkeys were specifically chosen because their CD28 receptors demonstrated similar affinity to TGN 1412 as human CD28 receptors. Based on such data as these, it was confidently concluded that results obtained from these NHPs would most reliably predict drug responses in humans—a conclusion that proved devastatingly wrong.

As exemplified by the study of HIV/AIDS, TGN 1412, and other experiences,⁶² experiments with NHPs are not necessarily any more predictive of human responses than experiments with other animals. The repeated failures in translation from studies with NHPs belie arguments favoring use of *any* nonhuman species to study human physiology and diseases and to test potential treatments. If experimentation using chimpanzees and other NHPs, our closest genetic cousins, are unreliable, how can we expect research using other animals to be reliable? The bottom line is that animal experiments, no matter the species used or the type of disease research undertaken, are highly unreliable—and they have too little predictive value to justify the resultant risks of harms for humans, for reasons I now explain.

The Collective Harms That Result from Misleading Animal Experiments

As medical research has explored the complexities and subtle nuances of biological systems, problems have arisen because the *differences* among species along these subtler biological dimensions far outweigh the *similarities*, as a growing body of evidence attests. These profoundly important—and often undetected—differences are likely one of the main reasons human clinical trials fail.⁶³

“Appreciation of differences” and “caution” about extrapolating results from animals to humans are now almost universally recommended. But, in practice, how does one take into account differences in drug metabolism, genetics, expression of diseases, anatomy, influences of laboratory environments, and species- and strain-specific physiologic mechanisms—and, in view of these differences, discern what is applicable to humans and what is not? If we cannot determine which physiological mechanisms in which species and strains of species are applicable to humans (even setting aside the complicating factors of different caging systems and types of flooring), the usefulness of the experiments must be questioned.

It has been argued that some information obtained from animal experiments is better than no information.⁶⁴ This thesis neglects how misleading information can be worse than no information from animal tests. The use of nonpredictive animal

experiments can cause human suffering in at least two ways: (1) by producing misleading safety and efficacy data and (2) by causing potential abandonment of useful medical treatments and misdirecting resources away from more effective testing methods.

Humans are harmed because of misleading animal testing results. Imprecise results from animal experiments may result in clinical trials of biologically faulty or even harmful substances, thereby exposing patients to unnecessary risk and wasting scarce research resources.⁶⁵ Animal toxicity studies are poor predictors of toxic effects of drugs in humans.⁶⁶ As seen in some of the preceding examples (in particular, stroke, HRT, and TGN1412), humans have been significantly harmed because investigators were misled by the safety and efficacy profile of a new drug based on animal experiments.⁶⁷ Clinical trial volunteers are thus provided with raised hopes and a false sense of security because of a misguided confidence in efficacy and safety testing using animals.

An equal if indirect source of human suffering is the opportunity cost of abandoning promising drugs because of misleading animal tests.⁶⁸ As candidate drugs generally proceed down the development pipeline and to human testing based largely on successful results in animals⁶⁹ (i.e., positive efficacy and negative adverse effects), drugs are sometimes not further developed due to unsuccessful results in animals (i.e., negative efficacy and/or positive adverse effects). Because much pharmaceutical company preclinical data are proprietary and thus publicly unavailable, it is difficult to know the number of missed opportunities due to misleading animal experiments. However, of every 5,000–10,000 potential drugs investigated, only about 5 proceed to Phase 1 clinical trials.⁷⁰ Potential therapeutics may be abandoned because of results in animal tests that do not apply to humans.⁷¹ Treatments that fail to work or show some adverse effect in animals because of species-specific influences may be abandoned in preclinical testing even if they may have proved effective and safe in humans if allowed to continue through the drug development pipeline.

An editorial in *Nature Reviews Drug Discovery* describes cases involving two drugs in which animal test results from species-specific influences could have derailed their development. In particular, it describes how tamoxifen, one of the most effective drugs for certain types of breast cancer, “would most certainly have been withdrawn from the pipeline” if its propensity to cause liver tumor in rats had been discovered in preclinical testing rather than after the drug had been on the market for years.⁷² Gleevec provides another example of effective drugs that could have been abandoned based on misleading animal tests: this drug, which is used to treat chronic myelogenous leukemia (CML), showed serious adverse effects in at least five species tested, including severe liver damage in dogs. However, liver toxicity was not detected in human cell assays, and clinical trials proceeded, which confirmed the absence of significant liver toxicity in humans.⁷³ Fortunately for CML patients, Gleevec is a success story of predictive human-based testing. Many useful drugs that have safely been used by humans for decades, such as aspirin and penicillin, may not have been available today if the current animal testing regulatory requirements were in practice during their development.⁷⁴

A further example of near-missed opportunities is provided by experiments on animals that delayed the acceptance of cyclosporine, a drug widely and successfully used to treat autoimmune disorders and prevent organ transplant rejection.⁷⁵ Its immunosuppressive effects differed so markedly among species that researchers

judged that the animal results limited any direct inferences that could be made to humans. Providing further examples, PharmaInformatic released a report describing how several blockbuster drugs, including aripiprazole (Abilify) and esomeprazole (Nexium), showed low oral bioavailability in animals. They would likely not be available on the market today if animal tests were solely relied on. Understanding the implications of its findings for drug development in general, PharmaInformatic asked, “Which other blockbuster drugs would be on the market today, if animal trials would have not been used to preselect compounds and drug-candidates for further development?”⁷⁶ These near-missed opportunities and the overall 96 percent failure rate in clinical drug testing strongly suggest the unsoundness of animal testing as a precondition of human clinical trials and provide powerful evidence for the need for a new, human-based paradigm in medical research and drug development.

In addition to potentially causing abandonment of useful treatments, use of an invalid animal disease model can lead researchers and the industry in the wrong research direction, wasting time and significant investment.⁷⁷ Repeatedly, researchers have been lured down the wrong line of investigation because of information gleaned from animal experiments that later proved to be inaccurate, irrelevant, or discordant with human biology. Some claim that we do not know which benefits animal experiments, particularly in basic research, may provide down the road. Yet human lives remain in the balance, waiting for effective therapies. Funding must be strategically invested in the research areas that offer the most promise.

The opportunity costs of continuing to fund unreliable animal tests may impede development of more accurate testing methods. Human organs grown in the lab, human organs on a chip, cognitive computing technologies, 3D printing of human living tissues, and the Human Toxome Project are examples of new human-based technologies that are garnering widespread enthusiasm. The benefit of using these testing methods in the preclinical setting over animal experiments is that they are based on *human* biology. Thus their use eliminates much of the guesswork required when attempting to extrapolate physiological data from other species to humans. Additionally, these tests offer whole-systems biology, in contrast to traditional *in vitro* techniques. Although they are gaining momentum, these human-based tests are still in their relative infancy, and funding must be prioritized for their further development. The recent advancements made in the development of more predictive, human-based systems and biological approaches in chemical toxicological testing are an example of how newer and improved tests have been developed because of a shift in prioritization.⁷⁸ Apart from toxicology, though, financial investment in the development of human-based technologies generally falls far short of investment in animal experimentation.⁷⁹

Conclusion

The unreliability of applying animal experimental results to human biology and diseases is increasingly recognized. Animals are in many respects biologically and psychologically similar to humans, perhaps most notably in the shared characteristics of pain, fear, and suffering.⁸⁰ In contrast, evidence demonstrates that critically important physiological and genetic differences between humans and other animals can invalidate the use of animals to study human diseases, treatments, pharmaceuticals, and the like. In significant measure, animal models specifically,

and animal experimentation generally, are inadequate bases for predicting clinical outcomes in human beings in the great bulk of biomedical science. As a result, humans can be subject to significant and avoidable harm.

The data showing the unreliability of animal experimentation and the resultant harms to humans (and nonhumans) undermine long-standing claims that animal experimentation is necessary to enhance human health and therefore ethically justified. Rather, they demonstrate that animal experimentation poses significant costs and harms to human beings. It is possible—as I have argued elsewhere—that animal research is more costly and harmful, on the whole, than it is beneficial to human health.⁸¹ When considering the ethical justifiability of animal experiments, we should ask if it is ethically acceptable to deprive humans of resources, opportunity, hope, and even their lives by seeking answers in what may be the wrong place. In my view, it would be better to direct resources away from animal experimentation and into developing more accurate, human-based technologies.

Notes

1. Taylor K, Gordon N, Langley G, Higgins W. Estimates for worldwide laboratory animal use in 2005. *Alternatives to Laboratory Animals* 2008;36:327–42.
2. Systematic reviews that have been conducted generally reveal the unreliability and poor predictability of animal tests. See Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, et al. Comparison of treatment effects between animal experiments and clinical trials: Systematic review. *BMJ* 2007;334:197. See also Pound P, Bracken MB. Is animal research sufficiently evidence based to be a cornerstone of biomedical research? *BMJ* 2014;348:g3387. See Godlee F. How predictive and productive is animal research? *BMJ* 2014;348:g3719. See Benatar M. Lost in translation: Treatment trials in the SOD 1 mouse and in human ALS. *Neurobiology Disease* 2007;26:1–13. And see Akhtar AZ, Pippin JJ, Sandusky CB. Animal studies in spinal cord injury: A systematic review of methylprednisolone. *Alternatives to Laboratory Animals* 2009;37:43–62.
3. Mathews RAJ. Medical progress depends on animal models—doesn't it? *Journal of the Royal Society of Medicine* 2008;101:95–8.
4. See Shanks N, Greek R, Greek J. Are animal models predictive for humans? *Philosophy, Ethics, and Humanities in Medicine* 2009;4:2. See also Wall RJ, Shani M. Are animal models as good as we think? *Theriogenology* 2008;69:2–9.
5. See note 3, Mathews 2008. See also Hartung T, Zurlo J. Food for thought... alternative approaches for medical countermeasures to biological and chemical terrorism and warfare. *ALTEX* 2012;29:251–60. See Leist M, Hartung T. Inflammatory findings on species extrapolations: Humans are definitely no 70-kg mice. *Archives in Toxicology* 2013;87:563–7. See Mak IWY, Evaniew N, Ghert M. Lost in translation: Animal models and clinical trials in cancer treatment. *American Journal in Translational Research* 2014;6:114–18. And see Pippin J. Animal research in medical sciences: Seeking a convergence of science, medicine, and animal law. *South Texas Law Review* 2013;54:469–511.
6. For an overview of the harms-versus-benefits argument, see LaFollette H. Animal experimentation in biomedical research. In: Beauchamp TL, Frey RG, eds. *The Oxford Handbook of Animal Ethics*. Oxford: Oxford University Press; 2011:812–18.
7. See Jucker M. The benefits and limitations of animal models for translational research in neurodegenerative diseases. *Nature Medicine* 2010;16:1210–14. See Institute of Medicine. *Improving the Utility and Translation of Animal Models for Nervous System Disorders: Workshop Summary*. Washington, DC: The National Academies Press; 2013. And see Degryse AL, Lawson WE. Progress towards improving animal models for IPF. *American Journal of Medical Science* 2011;341:444–9.
8. See Morgan KN, Tromborg CT. Sources of stress in captivity. *Applied Animal Behaviour Science* 2007;102:262–302. See Hart PC, Bergner CL, Dufour BD, Smolinsky AN, Egan RJ, LaPorte L, et al. Analysis of abnormal repetitive behaviors in experimental animal models. In Warrick JE, Kauleff AV, eds. *Translational Neuroscience and Its Advancement of Animal Research Ethics*. New York: Nova Science; 2009:71–82. See Lutz C, Well A, Novak M. Stereotypic and self-injurious behavior in rhesus macaques: A survey and retrospective analysis of environment and early experience. *American Journal of Primatology* 2003;60:1–15. And see Balcombe JP, Barnard ND, Sandusky C. Laboratory routines cause animal stress. *Contemporary Topics in Laboratory Animal Science* 2004;43:42–51.

The Flaws and Human Harms of Animal Experimentation

9. Suckow MA, Weisbroth SH, Franklin CL. *The Laboratory Rat*. 2nd ed. Burlington, MA: Elsevier Academic Press; 2006, at 323.
10. Flow BL, Jaques JT. Effect of room arrangement and blood sample collection sequence on serum thyroid hormone and cortisol concentrations in cynomolgus macaques (*Macaca fascicularis*). *Contemporary Topics in Laboratory Animal Science* 1997;36:65–8.
11. See note 8, Balcombe et al. 2004.
12. See note 8, Balcombe et al. 2004.
13. Baldwin A, Bekoff M. Too stressed to work. *New Scientist* 2007;194:24.
14. See note 13, Baldwin, Bekoff 2007.
15. Akhtar A, Pippin JJ, Sandusky CB. Animal models in spinal cord injury: A review. *Reviews in the Neurosciences* 2008;19:47–60.
16. See note 13, Baldwin, Bekoff 2007.
17. See note 15, Akhtar et al. 2008.
18. See Macleod MR, O'Collins T, Howells DW, Donnan GA. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke* 2004;35:1203–8. See also O'Neil BJ, Kline JA, Burkhardt K, Younger J. Research fundamentals: V. The use of laboratory animal models in research. *Academic Emergency Medicine* 1999;6:75–82.
19. Crabbe JC, Wahlsten D, Dudek BC. Genetics of mouse behavior: Interactions with laboratory environment. *Science* 1999;284:1670–2, at 1670.
20. See Curry SH. Why have so many drugs with stellar results in laboratory stroke models failed in clinical trials? A theory based on allometric relationships. *Annals of the New York Academy of Sciences* 2003;993:69–74. See also Dirnagl U. Bench to bedside: The quest for quality in experimental stroke research. *Journal of Cerebral Blood Flow & Metabolism* 2006;26:1465–78.
21. van der Worp HB, Howells DW, Sena ES, Poritt MJ, Rewell S, O'Collins V, et al. Can animal models of disease reliably inform human studies? *PLoS Medicine* 2010;7:e1000245.
22. See note 20, Dirnagl 2006. See also Sena E, van der Worp B, Howells D, Macleod M. How can we improve the pre-clinical development of drugs for stroke? *Trends in Neurosciences* 2007;30:433–9.
23. See Gawrylewski A. The trouble with animal models: Why did human trials fail? *The Scientist* 2007;21:44. See also Fisher M, Feuerstein G, Howells DW, Hurn PD, Kent TA, Savitz SI, et al. Update of the stroke therapy academic industry roundtable preclinical recommendations. *Stroke* 2009;40:2244–50.
24. See note 23, Gawrylewski 2007. There is some dispute as to how vigorously investigators adhered to the suggested criteria. Nevertheless, NXY-059 animal studies were considered an example of preclinical studies that most faithfully adhered to the STAIR criteria. For further discussion see also Wang MM, Guohua X, Keep RF. Should the STAIR criteria be modified for preconditioning studies? *Translational Stroke Research* 2013;4:3–14.
25. See note 24, Wang et al. 2013.
26. O'Collins VE, Macleod MR, Donnan GA, Horky LL, van der Worp BH, Howells DW. 1,026 experimental treatments in acute stroke. *Annals of Neurology* 2006;59:467–7.
27. See note 5, Mak et al. 2014.
28. See note 5, Mak et al. 2014.
29. See Perrin S. Preclinical research: Make mouse studies work. *Nature* 2014;507:423–5. See also, generally, Wilkins HM, Bouchard RJ, Lorenzon NM, Linseman DA. Poor correlation between drug efficacies in the mutant SOD1 mouse model versus clinical trials of ALS necessitates the development of novel animal models for sporadic motor neuron disease. In: Costa A, Villalba E, eds. *Horizons in Neuroscience Research*. Vol. 5. Hauppauge, NY: Nova Science; 2011:1–39.
30. Traynor BJ, Bruijn L, Conwit R, Beal F, O'Neill G, Fagan SC, et al. Neuroprotective agents for clinical trials in ALS: A systematic assessment. *Neurology* 2006;67:20–7.
31. Sinha G. Another blow for ALS. *Nature Biotechnology* 2013;31:185. See also note 30, Traynor et al. 2006.
32. See Morales DM, Marklund N, Lebold D, Thompson HJ, Pitkanen A, Maxwell WL, et al. Experimental models of traumatic brain injury: Do we really need a better mousetrap? *Neuroscience* 2005;136:971–89. See also Xiong YE, Mahmood A, Chopp M. Animal models of traumatic brain injury. *Nature Reviews Neuroscience* 2013;14:128–42. And see commentary by Farber: Farber N. Drug development in brain injury. *International Brain Injury Association*; available at <http://www.internationalbrain.org/articles/drug-development-in-traumatic-brain-injury/> (last accessed 7 Dec 2014).
33. Maas AI, Roozenbeek B, Manley GT. Clinical trials in traumatic brain injury: Past experience and current developments. *Neurotherapeutics* 2010;7:115–26.

34. Schneider LS, Mangialasche F, Andreasen N, Feldman H, Giacobini E, Jones R, et al. Clinical trials and late-stage drug development in Alzheimer's disease: An appraisal from 1984 to 2014. *Journal of Internal Medicine* 2014;275:251–83.
35. Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, et al. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences USA* 2013;110:3507–12.
36. Palfreyman MG, Charles V, Blander J. The importance of using human-based models in gene and drug discovery. *Drug Discovery World* 2002 Fall:33–40.
37. See note 2, Perel et al. 2007.
38. Harding A. More compounds failing phase I. *The Scientist* 2004 Sept 13; available at <http://www.the-scientist.com/?articles.view/articleNo/23003/title/More-compounds-failing-Phase-I/> (last accessed 2 June 2014).
39. See note 5, Pippin 2013.
40. See note 5, Hartung, Zurlo 2012.
41. Wiebers DO, Adams HP, Whisnant JP. Animal models of stroke: Are they relevant to human disease? *Stroke* 1990;21:1–3.
42. See note 15, Akhtar et al. 2008.
43. See note 2, Akhtar et al. 2009.
44. Lonjon N, Prieto M, Haton H, Bröchner CB, Bauchet L, Costalat V, et al. Minimum information about animal experiments: Supplier is also important. *Journal of Neuroscience Research* 2009;87:403–7.
45. Mogil JS, Wilson SG, Bon K, Lee SE, Chung K, Raber P, et al. Heritability of nociception I: Responses of 11 inbred mouse strains on 12 measures of nociception. *Pain* 1999;80:67–82.
46. Tator H, Hashimoto R, Raich A, Norvell D, Fehling MG, Harrop JS, et al. Translational potential of preclinical trials of neuroprotection through pharmacotherapy for spinal cord injury. *Journal of Neurosurgery: Spine* 2012;17:157–229.
47. See note 35, Seok et al. 2013, at 3507.
48. Odom DT, Dowell RD, Jacobsen ES, Gordon W, Danford TW, MacIsaac KD, et al. Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nature Genetics* 2007;39:730–2.
49. Horrobin DF. Modern biomedical research: An internally self-consistent universe with little contact with medical reality? *Nature Reviews Drug Discovery* 2003;2:151–4.
50. Vassilopoulos S, Esk C, Hoshino S, Funke BH, Chen CY, Plocik AM, et al. A role for the CHC22 clathrin heavy-chain isoform in human glucose metabolism. *Science* 2009;324:1192–6.
51. See Guttman-Yassky E, Krueger JG. Psoriasis: Evolution of pathogenic concepts and new therapies through phases of translational research. *British Journal of Dermatology* 2007;157:1103–15. See also The mouse model: Less than perfect, still invaluable. *Johns Hopkins Medicine*; available at http://www.hopkinsmedicine.org/institute_basic_biomedical_sciences/news_events/articles_and_stories/model_organisms/201010_mouse_model.html (last accessed 10 Dec 2014). See note 23, Gawrylewski 2007. See note 2, Benatar 2007. See note 29, Perrin 2014 and Wilkins et al. 2011. See Cavanaugh S, Pippin J, Barnard N. Animal models of Alzheimer disease: Historical pitfalls and a path forward. *ALTEX* online first; 2014 Apr 10. And see Woodrooffe A, Coleman RA. ServiceNote: Human tissue research for drug discovery. *Genetic Engineering and Biotechnology News* 2007;27:18.
52. Lane E, Dunnett S. Animal models of Parkinson's disease and L-dopa induced dyskinesia: How close are we to the clinic? *Psychopharmacology* 2008;199:303–12.
53. See note 52, Lane, Dunnett 2008.
54. See note 5, Pippin 2013.
55. Bailey J. An assessment of the role of chimpanzees in AIDS vaccine research. *Alternatives to Laboratory Animals* 2008;36:381–428.
56. Tonks A. The quest for the AIDs vaccine. *BMJ* 2007;334:1346–8.
57. Johnston MI, Fauci AS. An HIV vaccine—evolving concepts. *New England Journal of Medicine* 2007;356:2073–81.
58. See Rossouw JE, Andersen GL, Prentice RL, LaCroix AZ, Kooperberf C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy menopausal women: Principle results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321–33. See also Andersen GL, Limacher A, Assaf AR, Bassford T, Beresford SA, Black H, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: The Women's Health Initiative randomized controlled trial. *JAMA* 2004;291:1701–12.
59. Lemere CA. Developing novel immunogens for a safe and effective Alzheimer's disease vaccine. *Progress in Brain Research* 2009;175:83.

The Flaws and Human Harms of Animal Experimentation

60. Allen A. Of mice and men: The problems with animal testing. *Slate* 2006 June 1; available at http://www.slate.com/articles/health_and_science/medical_examiner/2006/06/of_mice_or_men.html (last accessed 10 Dec 2014).
61. Attarwala H. TGN1412: From discovery to disaster. *Journal of Young Pharmacists* 2010;2:332–6.
62. See Hogan RJ. Are nonhuman primates good models for SARS? *PLoS Medicine* 2006;3:1656–7. See also Bailey J. Non-human primates in medical research and drug development: A critical review. *Biogenic Amines* 2005;19:235–55.
63. See note 4, Wall, Shani 2008.
64. Lemon R, Dunnett SB. Surveying the literature from animal experiments: Critical reviews may be helpful—not systematic ones. *BMJ* 2005;330:977–8.
65. Roberts I, Kwan I, Evans P, Haig S. Does animal experimentation inform human health care? Observations from a systematic review of international animal experiments on fluid resuscitation. *BMJ* 2002;324:474–6.
66. See note 60, Allen 2006. See also Heywood R. Target organ toxicity. *Toxicology Letters* 1981;8:349–58. See Fletcher AP. Drug safety tests and subsequent clinical experience. *Journal of the Royal Society of Medicine* 1978;71:693–6.
67. See note 60, Allen 2006. See note 5, Pippin 2013. See also Greek R, Greek J. Animal research and human disease. *JAMA* 2000;283:743–4.
68. See note 60, Allen 2006. See also note 5, Leist, Hartung 2013.
69. Food and Drug Administration. Development & approval process (drugs); available at <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/> (last accessed 7 Dec 2014). See also <http://www.fda.gov/drugs/resourcesforyou/consumers/ucm143534.htm> (last accessed 7 Dec 2014).
70. Drug discovery pipeline. IRSF; available at <http://www.rettssyndrome.org/research-programs/programmatic-overview/drug-discovery-pipeline> (last accessed 24 Sept 2014).
71. See note 60, Allen 2006.
72. Follow the yellow brick road. *Nature Reviews Drug Discovery* 2003;2:167, at 167.
73. See note 5, Pippin 2013.
74. For data on aspirin, see Hartung T. Per aspirin as astra ... *Alternatives to Laboratory Animals* 2009;37(Suppl 2):45–7. See also note 5, Pippin 2013. For data on penicillin, see Koppányi T, Avery MA. Species differences and the clinical trial of new drugs: A review. *Clinical Pharmacology and Therapeutics* 1966;7:250–70. See also Schneiersson SS, Perlman E. Toxicity of penicillin for the Syrian hamster. *Proceedings of the Society for Experimental Biology and Medicine* 1956;91:229–30.
75. See note 67, Greek, Greek 2000.
76. Oral bioavailability of blockbuster drugs in humans and animals. *Pharmainformatic*. available at http://www.pharmainformatic.com/html/blockbuster_drugs.html (last accessed 19 Sept 2014).
77. Sams-Dodd F. Strategies to optimize the validity of disease models in the drug discovery process. *Drug Discovery Today* 2006;11:355–63.
78. Zurlo J. No animals harmed: Toward a paradigm shift in toxicity testing. *Hastings Center Report* 2012;42 Suppl:s23–6.
79. There is no direct analysis of the amount of money spent on animal testing versus alternatives across all categories; however, in 2008 the *Chronicle of Higher Education* reported that funding of research involving animals (under basic research) of the National Institute of Health (NIH) remained steady at about 42 percent since 1990. See Monastersky R. Protesters fail to slow animal research. *Chronicle of Higher Education* 2008:54. In 2012, NIH director Francis Collins noted that the NIH's support for basic research has held steady at 54 percent of the agency's budget for decades. The remainder of the NIH's budget is heavily funded toward clinical research, suggesting that *preclinical* human-based testing methods are much less funded. See also Wadman M. NIH director grilled over translational research centre. *Nature News Blog* 2012 Mar 20. Available at <http://blogs.nature.com/news/2012/03/nih-director-grilled-over-translational-research-center.html> (last accessed 5 Mar 2015). There is no data that suggests that the NIH's funding of animal experimentation has decreased. A 2010 analysis estimates that at least 50 percent of the NIH's extramural funding is directed into animal research; see Greek R, Greek J. Is the use of sentient animals in basic research justifiable? *Philosophy, Ethics, and Humanities in Medicine* 2010;5:14.
80. For a helpful discussion on animal pain, fear, and suffering, see DeGrazia D. *Taking Animals Seriously: Mental Lives and Moral Status*. New York: Cambridge University Press; 1996:116–23.
81. See Akhtar A. *Animals and Public Health: Why Treating Animals Better Is Critical to Human Welfare*. Hampshire, UK: Palgrave Macmillan; 2012:chap. 5.