Comment

A 'Road Map' Toward Ending Severe Suffering of Animals Used in Research and Testing

Elliot Lilley, Penny Hawkins and Maggy Jennings

Ending severe suffering is a desirable goal for both ethical and scientific reasons. The RSPCA has pledged to work toward the end of such suffering for laboratory animals, and in this article we outline a practical approach that establishments can follow to achieve this aim

Introduction

The introduction of EU *Directive 2010/63/EU*[1] controlling experiments on animals, and the associated updating of the *UK Animals (Scientific Procedures) Act 1986 (ASPA)*,[2] have focused attention on the need to reduce animal suffering in scientific procedures. Classification of levels of suffering into mild, moderate and severe, and the need to report actual levels of severity, have provided added impetus to the drive to refine the most severe models and procedures, and have resulted in greater recognition that high levels of suffering impact on an animal's physiological responses, increasing the variability of experimental data. So, ending severe suffering is a desirable goal for scientific reasons, as well as ethical and legal ones.

There are currently no data on the proportion of the approximately four million procedures carried out on animals in 2013 in the UK that caused severe suffering — or indeed, for any previous year. This will change from 2014, because the EU Directive and the *ASPA* will require research establishments to assess and report the actual level of suffering experienced by individual animals undergoing regulated procedures. This is good news for three reasons. Firstly, it will give the public a clear indication of the levels of suffering that laboratory animals experience, which will be an important step toward open and honest reporting of animal use in research. Secondly, it will give establishments an opportunity to evaluate how successful their *refinement* programmes are in reducing suffering, and highlight areas where more work is needed. Finally, these new data will allow regulators, welfare organisations and research funding bodies to focus resources on areas of research where suffering is the most severe and/or *refinement* is lacking, and to track progress of Three Rs-related programmes targeted at reducing suffering.

Clearly, the responsibility to end severe suffering falls on the whole scientific community, who need to accept this as a worthwhile goal and commit to achieving it. There needs to be a coordinated effort from researchers, industry and academia, regulatory authorities, funding bodies and scientific journals. But animal welfare organisations also have a key role, and even before revision of the *ASPA*, the RSPCA had pledged to work toward ending severe suffering. Since 2011, the Society has been developing a programme of work with the scientific community, aimed at producing innovative, challenging and feasible approaches to the achievement of this goal. This article focuses on the role of research establishments, and outlines some practical steps that can be taken to create a 'road map' to end severe suffering.

Stages on the Road

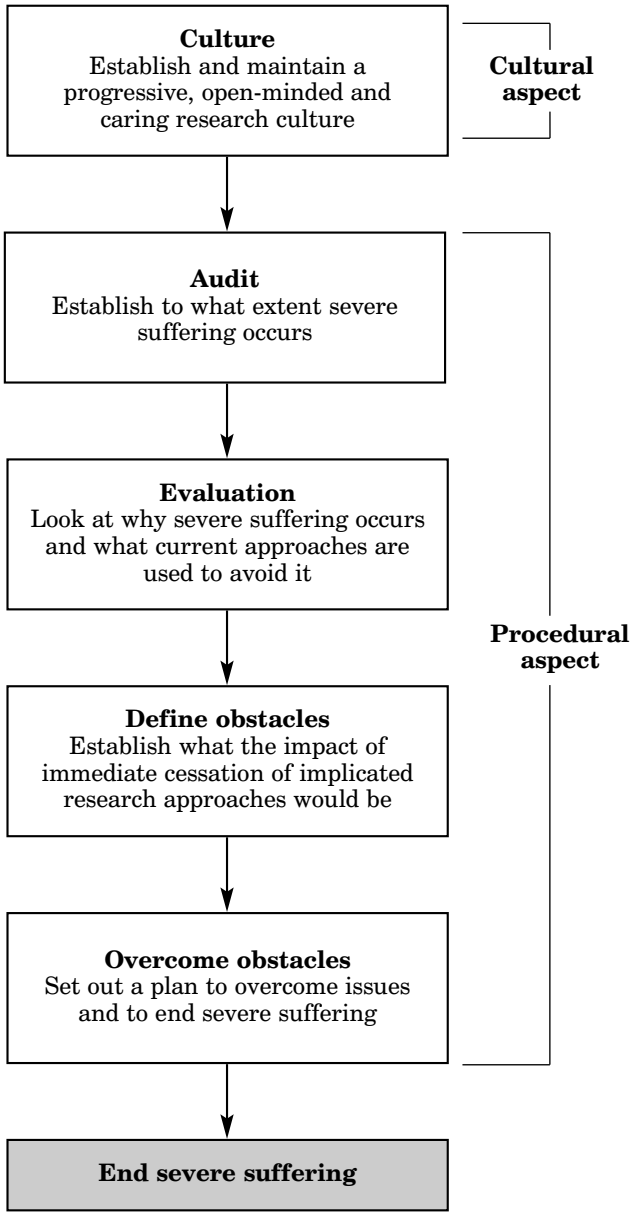
There are two complementary aspects to the 'road map' approach (Figure 1):

1. The *cultural* aspect — developing an environment that will support and encourage positive attitudes toward change; and
2. The *procedural* aspect — developing the activities and initiatives that will end severe suffering.

The cultural aspect — establishing the culture of care

Fundamental to ending severe suffering is the belief that this is both desirable and possible —

Figure 1: Stages on the road to ending severe suffering of animals used in research and testing



and is worthy of the necessary time and resources. A progressive, open-minded and caring research culture should be willing to embrace these concepts. The need for a culture (or climate) of care to be developed and maintained in designated establishments is now included as an integral part of the EU Directive (Recital 31) and the *ASPA*. The term 'culture of care' is much used, but is not formally defined; however, it is generally accepted that such a culture would have several components that collectively provide a framework which delivers high standards with respect to the legal, welfare, Three Rs and ethical aspects of animal use — and within which concerns such as severe suffering can be constructively addressed. Components of a culture of care include:

— A collective responsibility and accountability for the welfare of animals, shared by all staff.

— Demonstrable support at corporate level from senior management.

— Internal openness, including the ability to raise, share and resolve concerns without negative repercussions for the individual.

— Support for 'Named Persons' (such as Animal Care and Welfare Officers, Veterinary Surgeons, Information and Training and Competency Officers).

— A robust framework for the training, assessment of competence and continued professional development of all staff.

— Effective and well-supported institutional ethical reviews of scientific work.[3,4]

— A good ethics or animal care and use committee, e.g. the Animal Welfare and Ethical Review Body (AWERB) in the UK.

The ethics or animal care and use committee should be a major contributor in developing and maintaining the overall culture of care, as well as providing an ideal forum for initiating and driving activities that aim to reduce animal suffering.

The procedural aspect — developing the 'road map'

Assuming that the environment is supportive and the decision has been made to tackle severe suffering, a stepwise process can be initiated toward achieving positive change, as judged by genuine progress toward ending severe suffering.

The first step on the road is to establish the extent of the issue in-house, beginning with a 'severity audit' of all protocols, procedures and models (hereinafter referred to as 'procedures') to determine which are classified as having the potential to cause severe suffering. These procedures can then be reviewed, by using records of day-to-day observations of the animals, to identify those where the *actual* severity was severe.

Step two is to take these and carry out a detailed evaluation of:

a) why each procedure was used and what factors resulted in it being severe;

b) whether that level of suffering was really necessary to achieve the scientific objective;

c) what proportion of animals suffered severely;

d) what refinements were already in place, and

whether there is potential for further implementation of all Three Rs;

e) whether there are any scientific obstacles to refining the procedure or avoiding it altogether; and

f) if so, how these obstacles might be overcome.

Each of these concepts is discussed in more detail below.

Why the procedure was used and what factors made it severe

Most 'severe' procedures are used: a) to 'model' diseases that cause severe suffering in humans or other animals (e.g. multiple sclerosis, rheumatoid arthritis, chronic pain, sepsis); or b) in procedures where death or morbidity is the endpoint for regulatory purposes (e.g. acute toxicity tests, vaccine testing, batch testing of toxins used for medical purposes). In the former, a severe, end-stage disease model may be deemed essential to fully validate a particular hypothesis or therapeutic intervention, or it may be perceived (rightly or wrongly) that scientific journal editors will demand to see data from a 'gold standard' disease model which is severe. In the case of regulatory procedures, there may be a severe safety or efficacy test that is considered essential for successful registration (for example, in toxicity testing with significant morbidity or death as the endpoint). The ultimate scientific and regulatory reasons for conducting the severe procedure *per se*, and the contribution that it makes to the overall project, should be discussed and clearly understood. What made the actual severity of the procedure 'severe' may be immediately obvious — for example, highly invasive orthopaedic surgery or the after-effects of MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine; a neurotoxin) administration to a marmoset. Where there is no single, obvious cause of severe suffering, it will be necessary to conduct a thoughtful review of the animal's lifetime experience, to see whether cumulative causes can be identified. These might take the form of a series of procedures or 'technical acts' without adequate time for the animal to recover between them, or the combined impact of procedures and husbandry-related factors such as singly housing social animals or long-term housing.

Establishing the reason(s) for the procedure and the putative cause of severe suffering are fundamental to the next steps along the road.

Is a severe level of suffering really necessary?

It is widely accepted that severe suffering should only be permitted if there is a clear and justifiable scientific benefit. A rigorous analysis of the *scientific* justification for a procedure should be part of the harm–benefit evaluation within the project authorisation process, and this should thoroughly scrutinise the necessity for severe suffering, if it is involved. So, under what circumstances might there be a scientific requirement to cause or permit severe suffering?

One justification often given for severe animal models of disease is that they help scientists to understand and treat diseases that cause severe suffering in patients. However, it may not always be necessary to allow disease 'models' to progress to a severe state in order to answer the scientific question being asked. The prudent use of biomarkers (as surrogates for disease progression) in mechanistic, pathway-focused 'models' allows potential therapeutics to be assessed, without the need to create a disease model. This approach is increasingly being used within the pharmaceutical industry.[5,6]

In addition, human patients who experience severe pain or distress are usually given analgesics and/or sedatives, and other means of palliative care will always be used in order to provide further relief. It is possible to make the case that animals used to 'model' human diseases should receive the same level and nature of care — on both ethical and scientific grounds. In fact, given that therapeutic agents for diseases are rarely taken in isolation by human patients, it could be argued that combining candidate compounds with analgesics might even improve translatability in some studies.

The translational, or predictive, value of the procedure is, of course, an essential consideration. A significant number of recent publications have highlighted the poor predictive value of animal models for both drug safety and clinical efficacy. Reasons for this are likely to include poor study design and data analysis, poor reporting, publication bias, and fundamental differences in the underlying biology of humans and other animals.[7,8] A full appreciation of the limitations of animal models is particularly important when there is the potential for severe suffering.

For instance, a severe model may be used to assess the potential therapeutic value of a new drug, but this would have little benefit if the model had poor predictive value.[9–12] As an example, the dextran sodium sulphate (DSS) colitis model in mice may be useful for understanding the underlying mechanisms and pathology of intestinal inflammation, much of which may be relevant to human colitis, but candidate drug molecules that ameliorate DSS-induced inflammation in the mouse may not be effective in clinical trials and *vice versa* (13). It would be very difficult to justify causing severe suffering for such a tenuous benefit.

Regulatory authorities that demand data from severe procedures should be challenged (by the sci-

entific community, animal welfare organisations and the public) to explain empirically why this is necessary and why an alternative, less severe approach, is not acceptable. This is especially important if death as an endpoint is required, as in some toxicity and vaccine efficacy tests.[14] The EU Directive acknowledges that death as an endpoint should be avoided, and its guidance on severity assessment states that death should generally be regarded as causing severe suffering. Humane endpoints[15] should therefore always be defined and implemented (e.g. by humanely killing the animal, removing them from the study and/or administering pain relief), for both ethical and legal reasons.

What proportion of animals suffered severely?

This step is included for two reasons. First, it may be necessary for establishments conducting more than one severe procedure to prioritise them for action. Second, keeping track of the proportion of animals that experience severe suffering, in accordance with the legal requirement of the EU Directive to assess and report actual severity, is a good way to monitor progress on the issue. For example, a targeted effort to refine a severe protocol may result in the proportion of animals with an actual severity of 'moderate' to increase from 60% to 80%. On the one hand, this shows that the refinements are effective for many of the animals — and on the other hand, it provides an incentive to understand why the remaining 20% still experience severe suffering.

Reviewing refinement and the Three Rs

The simplest way to eliminate severe suffering in a procedure, of course, is to replace it with a humane alternative (even if this means asking a slightly different scientific question) — and this should be the ultimate goal. *Replacement* can often be thought to involve a direct 1:1 exchange of an animal model with a non-animal assay system. In reality, alternatives will often require researchers to use a battery of *in silico*, *in vitro* and human study approaches, in order to address a particular research question. The key principle is that alternative thinking is needed; the presumption against the use of animals must be the first step.

If this cannot be done, then greater efforts to reduce and refine the procedure clearly need to come into play. With respect to the principle of *reduction*, this obviously involves seeking further statistical advice to see whether more could be done to minimise the numbers of animals used. A creative approach to reducing suffering *via* experimental design may also be feasible; two examples are, the use of small numbers of animals in a stepwise approach, stopping when sufficient data have been obtained, rather than using a larger cohort of animals all at once, and assessing whether the duration of the experiment can be reduced without affecting data quality.

For example, if systemic infection causes severe suffering in a sepsis study, it may be possible to characterise a biomarker (e.g. core temperature or circulating levels of an inflammatory cytokine) to indicate disease progression, and to use this to terminate the study at an earlier time-point. Refining humane endpoints (for more information, see http://www.humane-endpoints.info/) in this way is particularly important in areas of research where mortality studies are common (e.g. vaccine development, sepsis research), and can give rise to both *reduction* and scientific benefits, because more data can be obtained from animals which are humanely killed, rather than 'found dead'.

Refinement is often the easiest of the Three Rs to achieve, and is clearly critically important with respect to severe suffering. Breaking the animal's full life experience (not only the experimental procedure) down into its component steps, analysing the potential for suffering, and refining each aspect — the 'aggregation of marginal gains' principle[16] — is an extremely useful tool.

If there is an obvious, single cause of severe suffering, it may be possible to take a targeted approach and tackle this directly, e.g. by reviewing the provision of analgesia in the case of pain, or even conducting the entire procedure under terminal anaesthesia (for example, in LPS-induced endotoxaemia). If the severe suffering is thought to be as a result of a series of technical acts or procedures, the potential for each of them to be further refined could be explored. This might involve supporting the animal during the procedure (e.g. with additional nesting material, liquid nutrition, hand feeding, heat pads), reducing restraint periods, and using positive reinforcement training to avoid restraint altogether. The rationale for the experimental design should also be reviewed; could animals be allowed more time to recover between interventions (for example, following surgical implantation of a radio telemetry device), or could the number of technical acts per animal be reduced?

Whether or not one or more causes of severe suffering can be identified, reviewing and refining housing, husbandry and care should always be a step along the road. For example, providing a good quality and quantity of space, group housing social animals, providing animal refuges, nesting materials and chew blocks, are all well-established means of reducing suffering and improving welfare.

It is also very important to use records of day-to-day observations and welfare indicators to evaluate whether refinements to husbandry or procedures really are effectively reducing suffering and improving welfare. Key to the whole

process is a comprehensive understanding of how to identify, assess and alleviate suffering in the species (and strain and individuals, where appropriate) used in the study.[17] It is important to bring people together who have sufficient knowledge, motivation, expertise and experience, who are prepared to 'think outside the box', and who have the time to do so — in effect, some form of well-informed, resourced and motivated '*refinement* team'. This team, if properly resourced, can ensure that *refinement* is properly considered, fully implemented and robustly reviewed. A well-constituted ethics or animal care and use committee can help to ensure that the 'right' staff are involved, and that additional external knowledge, expertise and perspectives are sought, where necessary.

Notwithstanding the above, most researchers consider that they have adequately implemented the Three Rs, particularly *refinement*, within their projects. The question is whether they have been sufficiently challenging — have they sought out the right expertise, covered all the issues, and asked the 'right' questions? Having a '*refinement* team', as suggested above, can help ensure that all relevant points are addressed, while paying special attention to projects involving procedures that cause severe suffering.

Whether there are scientific obstacles to reducing severity and how these might be overcome

A critical review of the justification for, and necessity of, severe procedures should reveal any scientific obstacles that need to be overcome to end severe suffering, taking an honest and realistic view of the impact (either negative or positive) on the scientific objectives of avoiding or refining the severe procedures. It may be useful to imagine that, due to a change in the legislative framework, research that causes severe suffering was banned with immediate effect. *Based on this hypothetical scenario*, what would be the impact of such a ban on the research that is currently being undertaken, and, more importantly, how would any resulting problems be dealt with?

Any obstacles to ending severe suffering should be clearly set out, and the genuine impact should be evaluated. It might be the case that a particular model may not be as essential as was first thought, and that by re-framing the research question, its use could be avoided.

In an ideal world, it would be possible to replace severe models by thinking in mechanistic terms about the research question, rather than trying to replicate the human disease, or to refine all severe procedures so that the level of suffering was significantly reduced — all without negatively impacting the science.

However, the situation becomes more challenging where there is a regulatory requirement to use a severe endpoint, or where a severe disease model is regarded as the 'gold standard' for a particular area of research. Here, it may be argued that ending severe suffering would have a significant impact on research progress or product registration in the short term. Either way, the consequences of immediate cessation of the use of all severe procedures should be evaluated, and a plan to mitigate them should be established.

In fact, what qualifies a model as being regarded as the 'gold standard' in a particular area of research should be assessed and challenged. It may be that there has been a thorough evaluation of the validity of a particular model and that data from that model had high translational value with respect to animal or human clinical benefit. Alternatively, it may be that a model is considered to be the 'gold standard', merely because it is used by the majority of researchers in a particular field, because of a particularly influential publication, ease of use, cost consideration, or a combination of these and other additional factors. Unless a model has very high predictive value, it should not be considered the 'gold standard'.

There are, unfortunately, some cases where a strong scientific justification for causing severe suffering can be made; for example, the induction of some disease models such as rheumatoid arthritis, experimental autoimmune encephalomyelitis, and stroke. Where this is the case, there should be a coordinated effort from researchers both to develop and validate humane endpoints in order to limit suffering experienced by animals in such studies, and to develop and adopt alternative, non-severe, approaches.

It may be useful to apply a 'stretch objective' to ending severe suffering by imposing a challenging, but achievable, time-point, after which no further severe studies would be undertaken. This target may be different, depending on the model or procedure being used and the specific obstacles that are identified. However, having a fixed point in time to work toward can be a powerful motivating force for achieving challenging goals.

Summary

Ending severe suffering is desirable and achievable. In this article we have outlined a 'road map', or series of steps, that can be used to evaluate the extent to which severe suffering occurs within establishments, and some practical principles that can be employed in order to work toward ending it. We urge researchers to act upon these principles, and to begin a journey toward a future where causing severe suffering to animals used in scientific research is a thing of the past.

Author for correspondence:
Elliot Lilley
RSPCA Research Animals Department
Wilberforce Way
Southwater
Horsham RH13 9RS
UK
E-mail: elliot.lilley@rspca.org.uk

References

[1] Anon. (2010). *Directive 2010/63/EU* of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official Journal of the European Union* **L276**, 20.10.2010, 33–79.

[2] Home Office (2014). *Guidance on the Operation of the Animals (Scientific Procedures) Act 1986*, 164pp. London, UK: The Stationery Office.

[3] Smith, J., van den Broek, F., Martorell, J., Hackbarth, H., Ruksenas, O. & Zeller, W.; FELASA Working Group on Ethical Evaluation of Animal Experiments (2007). Principles and practice in ethical review of animal experiments across Europe: Summary of the report of a FELASA working group on ethical evaluation of animal experiments. *Laboratory Animals* **41**, 143–160.

[4] Franco, N. & Olsson, A. (2013). Is the ethical appraisal of protocols enough to ensure best practice in animal research. *ATLA* **41**, P5–P7.

[5] Hunter, J. (2011). Have animal models of disease helped or hindered the drug discovery process? *Annals of the New York Academy of Sciences* **1245**, 1–2.

[6] Fishman, M. (2013). Power of rare diseases: Found in translation. *Science Translational Medicine* **5**, 1–3.

[7] Landis, S.C., Amara, S.G., Asadullah, K., Austin, C.P., Blumenstein, R., Bradley, E.W., Crystal, R.G., Darnell, R.B., Ferrante, R.J., Fillit, H., Finkelstein, R., Fisher, M., Gendelman, H.E., Golub, R.M., Goudreau, J.L., Gross, R.A., Gubitz, A.K., Hesterlee, S.E., Howells, D.W., Huguenard, J., Kelner, K., Koroshetz, W., Krainc, D., Lazic, S.E., Levine, M.S., Macleod, M.R., McCall, J.M., Moxley, R.T., 3rd, Narasimhan, K., Noble, L.J., Perrin, S., Porter, J.D., Steward, O., Unger, E., Utz, U. & Silberberg, S.D. (2012). A call for transparent reporting to optimize the predictive value of preclinical research. *Nature, London* **490**, 187–191.

[8] Ioannidis, J., Greenland, S., Hlatky, M., Khoury, M., Macleod, M., Moher, D., Schulz, K. & Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* **383**, 166–175.

[9] Hackam, D. (2007). Translating animal research into clinical benefit. *BMJ* **334**, 163–164.

[10] Levin, L.A. & Danesh-Meyer, H.V. (2010). Lost in translation: Bumps in the road between bench and bedside. *JAMA* **303**, 1533–1534.

[11] Ioannidis, J. (2012). Extrapolating from animals to humans. *Science Translational Medicine* **151**, 151–154.

[12] McGonigle, P. & Ruggeri, B. (2014). Animal models of human disease: Challenges in enabling translation. *Biochemical Pharmacology* **87**, 162–171.

[13] Petersen, Y., Björkenberg, B., Christjansen, K., Stahlhut, M., Pedersen, H. & Thorkildsen, C. (2009). Is the murine dextran sodium sulfate-induced colitis model valid for predicting drug efficacy in IBD? *Gastroenterology* **136**, Suppl. 1, A48.

[14] Jennings, M., Morton, D., Charton, E., Cooper, J., Hendriksen, C., Martin, S., Pearce, M.C., Price, S., Redhead, K., Reed, N., Simmons, H., Spencer, S. & Willingale, H. (2010). Application of the Three Rs to challenge assays used in vaccine testing: Tenth report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement. *Biologicals* **38**, 684–695.

[15] Ashall, V. & Millar, K. (2013). An opportunity to focus on the 'humane' in experimental endpoints: Moving beyond *Directive 2010/63/EU*. *ATLA* **41**, 307–312.

[16] Lilley, E. & Jennings, M. (2013). Refinement: Lessons from the 2012 Olympics. *ATLA* **41**, P28–P29.

[17] Hawkins, P., Morton, D., Burman, O., Dennison, N., Honess, P., Jennings, M., Lane, S., Middleton, V., Roughan, J., Wells, S. & Westwood, K. (2011). A guide to defining and implementing protocols for the welfare assessment of laboratory animals. *Laboratory Animals* **45**, 1–13.