Avoiding mortality in animal research and testing

Report of two workshops held by the RSPCA, LASA, LAVA and the IAT
University of Cambridge, 19 September 2017 and 1 October 2018

Penny Hawkins, Sharon Brookes,
James Bussell, Ngaire Dennison,
Helmut Ehall, Anne-Marie Farmer,
Theresa Langford, Chris Lelliott,
Elliot Lilley, Ian Ragan,
Kathy Ryder, Sara Wells

Summary

This report aims to provide guidance, and stimulate discussion, on ways of avoiding mortality in animals in research and testing. It is concerned with mortality that could feasibly be avoided, by reducing the number of animals 'found dead', reducing unpredicted mortality of animals being used in procedures, and challenging perceived requirements for death as an endpoint. Suggested approaches to avoiding mortality include: reviewing welfare assessment; undertaking pilot studies; improving staff training; data or record mining; and reviewing and challenging regulatory requirements. The report also sets out some issues that may need careful consideration, e.g. interpreting indicators of impending mortality in aged animals, and considering harms and benefits of increasing surveillance when this might cause additional distress. It concludes with a 'wish list' of developments that would further help to avoid mortality, and action points for scientists, animal technologists, veterinarians, regulators and members of ethics and/or animal care and use committees.

Authors

Penny Hawkins
RSPCA Research Animals Department

Sharon Brookes
Animal and Plant Health Agency

James Bussell
Wellcome Sanger Institute

Ngaire Dennison
University of Dundee

Helmut Ehall
Envigo

Anne-Marie Farmer
University of Cambridge

Theresa Langford
Institute of Animal Technology

Chris Lelliott
Wellcome Sanger Institute

Elliot Lilley
RSPCA Research Animals Department

Ian Ragan
NC3Rs

Kathy Ryder
Home Office Animals in Science Regulation Unit

Sara Wells
Medical Research Council Harwell Institute

1. Background

This workshop report is part of a joint initiative to reduce and avoid severe suffering for animals used in research and testing, organised by the RSPCA, LASA, LAVA and the IAT. Avoiding mortality was chosen as a topic to address within the severe suffering initiative because, under European Union and UK law, actual severity is assumed to be severe if an animal is 'found dead' when undergoing a scientific procedure, and the death is likely to be due to the procedure (unless an informed decision can be made that severe suffering did not occur before death). The scope of this report includes both reducing the number of animals 'found dead', and challenging perceived scientific or regulatory requirements for death as an endpoint, as there is immediate potential to avoid these causes of mortality*.

The group divided these into three main categories:

- **unpredicted mortality in stock animals held for future breeding or experimental use, both wild type and genetically altered (GA); e.g. some (background) strains may have a higher mortality rate than others, or some GA animals may develop an unanticipated lethal phenotype;**
- **unpredicted mortality in animals undergoing procedures (e.g. disease 'models'); and**
- **predictable mortality of animals, for example in studies to fulfil those regulatory requirements where, currently, death is explicitly required.**

Two workshops were initiated to discuss good practice approaches for predicting and avoiding death, and to share these more widely. We ensured that relevant expertise, and areas of animal use, were represented at the workshops, to maximise opportunities for cross-disciplinary thinking. This included representatives from all four convening organisations and other expert participants from areas including regulatory toxicology, fundamental, preclinical and veterinary research, and those managing large colonies of GA and aged mice (the agenda for the first meeting is attached at Appendix 1).

This report summarises the working group's discussion and conclusions. It is intended for people involved in laboratory animal sciences, including scientists, animal technologists, veterinarians, facility managers, regulators and members of ethics and/or animal care and use committees, such as UK Animal Welfare and Ethical Review Bodies (AWERBs), European Union (EU) Animal Welfare Bodies (AWBs), Animal Care and Use Committees (ACUCs) or Animal Ethics Committees (AECs) worldwide. Although produced in the UK, the conclusions apply globally. The authors hope that the report will be useful both when responding to issues of avoidable mortality, and for regular review of establishment practice and protocols in general.

2. Reasons for avoiding mortality

There are legal, ethical, animal welfare, and scientific reasons to avoid mortality. For example, in European and UK legislation death as an endpoint of a procedure must be avoided as far as possible and replaced by early and humane endpoints. Except in very specific situations (e.g. some regulatory studies in which mortality is currently considered a necessary endpoint), when an animal dies in a research environment, not only is an animal's life lost, but data and resources are also often lost. For example, in sepsis research, mortality is often reported as an endpoint in scientific papers. However, a recent expert working group considered this issue and found no justification for death as an endpoint. Whether or not animals survive is simply a binary readout, and more valuable data can be obtained by using discrete biomarkers and clinical signs to define earlier humane endpoints [1].

Mortality should therefore be closely monitored, challenged and avoided wherever possible (see also reference 2). A better understanding of the causes of death, identification of early signs of potential mortality, and implementation of intervention strategies are all required to effectively prevent animal deaths. The working group agreed that zero avoidable mortality should be the goal; there should never be an 'acceptable' level because that removes the incentive to challenge the status quo and make further efforts to reduce mortality.

3. Approaches to avoiding mortality

Potential ways to avoid mortality are set out in sections 3.1 to 3.7 below; some are generic, and others apply to specific research fields. The actions and approaches are listed broadly according to the amount of resource required.

3.1 Challenge endpoints requested by journal reviewers or editors

Mortality as an experimental endpoint is sometimes requested by journal editors, or peer reviewers, when papers are submitted reporting research into diseases that can lead to human or animal mortality (e.g. sepsis research). This may be merely because previous publications in the field have done so. Such requests should always be robustly challenged, because death as an endpoint is rarely justifiable.

* Although all animal mortality raises ethical concerns, this report does not cover circumstances in which there is a scientific requirement to humanely kill animals (e.g. as part of a research project, or for tissue collection), or when animals are humanely killed when severity limits are approached, or when an authorised clinical or scientific endpoint is reached.

3.2 Review welfare assessment

Reviewing welfare assessment protocols, including indicators and timing and frequency of observations, is an obvious course of action when seeking to avoid deaths.

Questions to ask include:

- **Is an indicator of mortality being missed?**
- **Could animal observation and monitoring be made more effective and timely, to help refine and implement humane endpoints and other interventions?**

If there is a risk of mortality, or deaths have occurred, reviewing welfare assessment should be allocated adequate time, resource and input from people with relevant expertise. A review of welfare assessment should include:

(i) investigating the current state of knowledge regarding the assessment of pain, suffering and distress, both in general and for the specific model, or area of investigation; and

(ii) reviewing welfare assessment records.

The primary aim should be to ensure that the assessment protocols and recording systems are sufficiently tailored to the species, protocol and circumstances. For example; could new, potentially useful indicators be added, or are any of the current indicators not being observed in practice? The box below lists some examples of resources to help review welfare assessment.

General resources include

- National Three Rs centres
- The European Commission Working Document on a severity assessment framework [3], and the accompanying worked examples [4]. The EC working document is especially useful because it sets out a structured approach for observing animals, which helps to reduce the risk of missing important indicators
- A guide to defining and implementing protocols for the welfare assessment of laboratory animals: eleventh report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement [5]
- Classification and reporting of severity experienced by animals used in scientific procedures: a FELASA/ECLAM/ESLAV Working Group report [6]
- The UK National Centre for the Three Rs (NC3Rs) website includes resources related to welfare assessment. They include the Grimace Scale for recognition of pain (**nc3rs.org.uk/grimacescales**), the Procedures with Care resource (procedureswithcare.org.uk) and the general welfare hubs for rodents (**nc3rs.org.uk/rodent-welfare-hub**) and non-human primates (**nc3rs.org.uk/welfare-non-human-primates**)
- Assessing the Health and Welfare of Laboratory Animals (AHWLA) (**ahwla.org.uk**)
- The University of Cambridge 3Rs search tool (**ubs.admin.cam.ac.uk/3rs/3rs-search-tool**)

Assessing mouse welfare:

- The Mouse Welfare Terms website (**mousewelfareterms.org**), with standardised terms for describing characteristics of laboratory mice
- The Mouse Genome Informatics website (**informatics.jax.org**) contains information on the phenotypes of mouse lines and links to references
- The International Mouse Phenotyping Consortium database (**mousephenotype.org**) also contains information on mouse phenotypes including viability

Assessing fish welfare:

- A Zebrafish Health and Welfare Glossary is being developed (**wiki.zfin.org/display/ZHWG/Zebrafish+Health+and+Welfare+Glossary+Home**)
- The Fish Indicators of Stress and Health (FISH) website [7]
- Behavioural indicators of welfare in farmed fish –a review paper [8]

Assessing the welfare of farmed animals:

- European Food Safety Authority (EFSA) opinions (**ec.europa.eu/food/animals/welfare/efsa_opinions_en**)
- Advancing Animal Welfare Assurance (AssureWel) (**assurewel.org**)

In addition, the RSPCA severe suffering web pages* list examples of model-specific refinements including welfare assessment, and other relevant reports include Percie du Sert et al. 2017 (stroke) [9] and Workman et al. 2010 (cancer) [10]. Databases to search for other specific examples include Norecopa (**norecopa.no/**) and Google Scholar; others are listed within the European Commission website**. Examples of useful search terms include: 'animal welfare assessment', 'animal pain assessment' and 'animal distress assessment', also 'pain behav*' or 'distress behav*' when used with 'animal' or the species in question. It would be helpful if authors could include keywords like these in publications that describe innovative and effective protocols for assessing animals, within any in vivo research field.

Consulting widely on welfare assessment can help to obtain useful insights from other fields, e.g. a scoring system might be transferred from one model, or research field, to another. As an example, indicators of toxicity following the injection of lipopolysaccharide (LPS) [11] might form a basis for preliminary endpoints in a study where animals could be at risk of septicaemia. Such consultations could include discussions with people who have expertise in the specific or similar models, or of assessing animal welfare; professional networks and mailing lists are important sources of expertise and contacts. The authors frequently consult colleagues from their own and other establishments, including in other countries, to benefit from expertise and information that may not be cited in the scientific literature. They also undertake formal training opportunities in other facilities, sometimes facilitated by actively seeking travel bursaries.

Following the review and consultation, it may be necessary to make changes to elements of the welfare assessment protocol such as the indicators used, times of day at which animals are assessed, frequency of observations, or implementation of the protocol at different disease stages or times of day. For example, if nocturnal rodents die during the night (because they are most active then, and/or less frequently observed), what steps could be taken to address this? The welfare assessment protocol should thus become further tailored to the specific model and circumstances (see the two case studies). This is standard good practice; robust review, including all necessary expertise, is vital when there is a risk of death so it is essential to allocate time and resources to this.

Going forward, the success of the welfare assessment protocol, and any changes to this, should ideally be evaluated at intervals during the project – and always evaluated as part of the post-study review, with the results reported to the local AWERB, AWB or relevant committee.

* science.rspca.org.uk/sciencegroup/researchanimals/severesuffering/resources/reports

** ec.europa.eu/environment/chemicals/lab_animals/3r/key_resources/databases_en.htm

Case study I:
Disease models that include a risk of sudden death, e.g. avian influenza A virus (AIV)

The Animal and Plant Health Agency (APHA) strives to develop and implement humane alternatives to in vivo tests wherever possible, including in vitro models (continuous and primary cell lines), in ovo models (embryonated eggs used before the embryos became protected animals) and ex vivo models (animal tissues). However, some regulatory tests require in vivo data.

AIV is a notifiable avian disease*, and the World Organisation for Animal Health (OIE) testing requirements for AIV include an intra-venous pathogenicity index (IVPI) test of 'severe' severity. Virus strains are considered to be highly pathogenic (HPAIV) if they cause over 75% mortality within 10 days. APHA, an OIE reference laboratory, has successfully refined this test to reduce the number of birds found dead by defining a humane end point more clearly as follows; any bird scored as having persistent (>24 hours) moderate signs is observed more frequently (from two observations per day to three or more), and humanely killed if any severe signs are observed such as paralysis or torticollis (twisted neck), or if the bird is unresponsive or cannot eat or drink. Results obtained using these endpoints are accepted by all regulators.

Research work is also undertaken to study AIV pathogenicity and transmission in chickens, turkeys and ducks. Most protocols for ducks have been reduced from severe to moderate, or mild in many cases, by increasing the frequency of welfare inspections (from one to three per day). Staff also sought to reduce the number of times each bird is handled for sampling, because this is stressful. The APHA internal ethical review of protocols before in vivo studies begin includes balancing the need to minimise the number of samples with ensuring that sufficient data will be obtained to achieve the scientific goal. This highlighted the importance of carrying out research and literature reviews to understand the pathogen – in this case APHA reviewed LowPAIV vs. HPAIV, enabling virulence to be predicted based on molecular characterisation, as data showed that certain viral characteristics result in higher virulence causing greater morbidity and/ or a higher risk of sudden death.

Focussed refinement of humane endpoints, as above, has enabled APHA to reduce severity to mild or moderate for most birds; 97% of ducks, 85% of chickens and 75% of turkeys. It has also enabled numbers found dead to be substantially reduced over the last 5 to 7 years; chickens from 17% to 5% and ducks from 7% to 0%. However, turkeys remain problematic at around 20% mortality, despite husbandry and monitoring protocols that include extra care and clinical monitoring, with reduced sampling stress. APHA has an ongoing goal to address this, including understanding the differences in species response to AIV infection and minimising the impact on turkeys and the numbers used in study protocols.

* oie.int/animal-health-in-the-world/oie-listed-diseases-2017

Case study II:
Conditional lung tumour models in mice

KRAS mutant mice are used to model human non-small cell lung cancer (NSCLC). 'Standard' KRAS knock-in mice develop tumours at around 6 weeks after viral infection, with adenocarcinomas from around 16 weeks. Median survival is 185 days (range around 110 to 210 days). Genetic modifications are investigated to test whether loss or gain of these genes sensitises or protects against KRAS driven tumour development. If protectant, mice should have extended survival times. The desired scientific outcomes are tissues to provide tumour cells for culture and ex-vivo studies and data to generate a survival curve. Both require animals to develop advanced tumours. Death as an endpoint is unsuitable when harvesting tumour tissue, so that both ethically and scientifically there was an imperative to define a surrogate endpoint for the survival curve data.

Initial endpoints for euthanasia were animals being 'noticeably ill' or losing 20% body weight, but some were still found dead and others did not develop large enough tumours. The monitoring system was updated so that treated mice were weighed 2 to 3 times per week until their weight fluctuated, or they displayed a noticeable increase in respiratory rate, indicating effects of lung cancer. They were then subjected to a 'scruff test' (analogous to 'stressing' humans with cardiovascular compromise on a treadmill) and their recovery assessed. Each mouse was tightly restrained by the scruff for 10 to 15 seconds, initially once a week, then at least three times a week, and finally daily as clinical signs progressed. Scores were allocated from 1 (normal, breathing as before the scruff test) to 5 (struggling to breathe afterwards, staggering when moving), with an endpoint of 5+ (seizure, lies gasping).

Animals scoring 4 and above often did not show obvious clinical signs other than increased respiratory rate, although changes in gait and forelimb position have been noted. Weight loss is still a poor predictor of endpoint, so body condition scoring [12] has been used to make weight data more meaningful.

Some challenges remain:

- **the monitoring method is time consuming and requires training for consistency;**
- **scruffing the animals is stressful, but reducing or eliminating this would necessitate other endpoints, such as basal respiratory rate, which might be harder to apply consistently. Activity levels, if automated behavioural monitoring equipment were available, might be a solution; and**
- **reducing the score for euthanasia would reduce suffering but risk some animals having tumours of insufficient size.**

However, no animals have been found dead since the scruff test was initiated. The endpoint permits the growth of sufficiently large tumours, and the data is acceptable to the scientific community as a surrogate for death for survival studies (and has been included in publications).

3.3 Undertake pilot studies (or studies in parallel) to identify approaches to predicting and avoiding mortality

Where there is a higher risk of death (or severe suffering) associated with an experiment, evaluation studies should be considered to identify ways of predicting mortality, refining procedures and otherwise minimising suffering. These may be conducted:

1. as separate pilot studies, before the main study; or
2. in parallel with the main study, conducted at the same time and using an extra group of animals. This will increase the number of animals, but will use fewer animals than a separate pilot study. This is because the same control group as the 'main' experiment can act as the control for the refinement, so a separate control group will not be necessary; or
3. also in parallel with the main study, but using a group within this study to evaluate a refinement, such that no additional animals are required (sometimes referred to as a 'piggyback' approach).

Ideally, pilot studies conducted before projects begin (or done in parallel with an extra group) will be designed so that the data they generate could be incorporated into data from the main experiment, to avoid additional animal use. Carefully planned and executed studies like these can not only reduce suffering, but also improve the validity of the scientific data obtained from the main experiments.

The following text refers to 'pilot' studies for clarity, but all the factors and indicators we describe can also be evaluated in parallel with ongoing projects as in points (2) and (3) above.

Like all experiments, pilot studies need clear objectives, for example 'to establish or redefine the earliest point at which scientifically valid data can be obtained'. Clear objectives will also help to identify logistical, scientific and animal welfare issues that may arise in subsequent experiments. A single, well-designed pilot study should be able to address all the issues that may arise in the definitive study, but two or more pilots might be needed for more complex studies. Investing resource in a pilot study is particularly worthwhile where a new model is being introduced that is planned for long term use and could involve many animals. The PREPARE checklist for designing animal experiments (**norecopa.no/PREPARE**) and the Experimental Design Assistant (**nc3rs.org.uk/experimental-design-assistant-eda**) can be helpful with respect to planning pilot studies.

Examples of factors that can be addressed within pilot studies as ways of reducing and avoiding mortality are listed below.

- **Whether the proposed clinical signs and indicators of mortality are sufficiently robust and predictive. For example, evaluations could include asking another person 'blinded' to the treatment to score videos of animals, and/or perform statistical analyses of clinical sign 'scores' against actual outcomes. The aim is to ensure that indicators are used because they are genuinely predictive, rather than intuitive or 'traditional'.**
- **Whether the procedure could be associated with a significant level of mortality, e.g. because it is technically challenging, especially invasive or severe. A pilot study could compare procedures and test whether an alternative, less technically or physiologically demanding model might yield equivalent results*.**
- **End-points, both scientific and welfare (humane), and their effectiveness. Is it possible to identify new biomarkers or clinical signs which could be used to reduce the risk of severe suffering or mortality? Could any parameters that are being monitored for scientific purposes (e.g. heart rate or blood pressure) also be used to predict mortality and define endpoints? Pilot studies can also be used to evaluate whether sufficient data can be obtained about the mechanisms of a disease process before animals develop advanced disease.**
- **Times of day when there is a higher risk of mortality, for example in relation to circadian rhythms or time elapsed after dosing or surgery, or time points during disease progression at which animals are especially susceptible.**
- **Ways of reducing severity and mortality via refinements (e.g. supportive husbandry measures such as supplying highly palatable food) or resolving technical issues (e.g. revising the way a technically difficult procedure is undertaken or equipment is used).**
- **Treatment variables that may increase the risk of mortality, e.g. dose rates and duration of interventions. Undertaking 'bottom up' rather than 'top down' dose ranging studies in regulatory testing will reduce suffering and mortality risk; see also the Response Surface Pathway (RSP) approach [14].**

With respect to pilot studies aiming to define robust indicators of future mortality, those that occur most frequently in the literature are body temperature, body weight and difficulty in rising or locomotion [2]. For example, a combination of temperature and body weight has been used to effectively predict death in mice used in infectious disease research [15], ocular herpes studies [16], and in aged mice [17, 18], and body temperature has been used with physical activity, and food and water consumption, to predict death in mouse models of staphylococcal enterotoxic shock [19] and lymphoma [20] respectively (see section 3.5 on animal monitoring).

Some publications describe studies which take animals to death as an endpoint, then calculate the number of 'false positives' and 'false negatives' to get a measure of the robustness of their indicators of mortality. It may be possible to make a case for such studies, but only when there is absolutely no alternative way to achieve international regulatory adoption of non-lethal endpoints, or acceptance within a research discipline of refined disease models. Also, rather than undertaking pilot studies, this kind of evaluation should only be done via a parallel, 'piggyback' approach, using a protocol with death as an endpoint that would be run in any case, so that no additional animals are allowed to die.

* Kamp et al. describe a systematic and meta-analysis of mortality in different mouse models of Delayed Cerebral Ischaemia, identifying significant differences in mortality rates according to the experimental approaches and protocols [13].

3.4 Improve staff training

The risk of mortality will be reduced if all relevant staff are adequately trained and competent in three areas; conducting procedures, welfare assessment, and humane killing. This section refers to the modular training set out within the common education and training framework to fulfil Directive 2010/63/EU [21]. Trainees must complete this modular programme before they begin a period of work under supervision, after which they must be assessed before they can be signed off as competent.

3.4.1 Training in conducting procedures

European Union training modules 7, 8, 20, 21 and 22 all address various aspects of conducting procedures, including restraint, dosing and sampling, sedation, local and general anaesthesia, and surgery. For procedures with an additional risk of mortality, or where the impact is difficult to predict, special training needs should be identified by personnel such as supervisors, veterinarians, senior animal technologists or those appointed to a role that specifically oversees training needs (such as the Named Training and Competence Officers (NTCOs) in the UK). Discussion and observation in an environment of coaching, or mentoring, with more experienced members of the team (such as an animal technologist, scientist and/or veterinarian) is also useful.

3.4.2 Training in welfare assessment

EU Module 5, on recognising pain, suffering and distress, provides an introduction to the topic, but is not tailored to achieving competence in monitoring models where there is the risk of mortality or significant suffering. It is especially important for the trainer to devote adequate time to ensuring appropriate competence in welfare assessment when procedures carry a significant risk of mortality, or severe suffering, or if adverse effects are unknown.

3.4.3 Training in humane methods of killing

This topic is the subject of EU module 6 (humane methods of killing). For all levels of severity and types of procedure, humanely killing animals when humane endpoints are approached should be straightforward, and undertaken with minimal delay. However, where there is an increased risk of mortality, and animals may rapidly become moribund and die, it is essential to ensure that processes are in place to identify at any time, within minutes, a person who can competently kill a suffering animal.

3.4.4 Sourcing training materials and expertise

Many training videos and other resources are available online, and it may be possible to receive training in monitoring impacts and endpoints either online or from distant sites via video links. The UK Named Information Officer, or equivalent, should be able to help identify relevant materials and contacts. However, the quality of any videos sourced from sites that were not produced by welfare refinement experts should be reviewed and discussed with colleagues possessing relevant expertise, before being used as training materials with respect to reducing mortality.

Training materials that group members have found useful are listed below.

- **The general resources to help review welfare assessment (see box on page 4), which are also useful training aids.**
- **Training materials for defining and monitoring humane endpoints can often be found within documents on refining models, e.g. the examples in the RSPCA/LASA/LAVA/IAT severe suffering web pages* and NC3Rs impact page (nc3rs.org.uk/our-impacts). Other useful sources are guidance on assigning severity, such as the UK Home Office advisory notes on recording and reporting the actual severity of regulated procedures [22] and severity classification of GA animals [23].**
- **Humane Endpoints in Laboratory Animal Experimentation website (humane-endpoints.info/en) – provides guidance on applying humane endpoints, with a secure database of videos and photographs to help with training.**
- **The e-learning resource 'Recognition and prevention of pain, suffering and distress in laboratory animals' aims to deliver the objectives set out for EU Module 5, and the principles and examples also apply internationally (by Newcastle University and NC3Rs) (cbctraining.ncl.ac.uk/eM-EU5/story_html5.html).**
- **The Laboratory Animal Science Association (LASA) (2nd edition, 2016) Guiding Principles for Supervision and Assessment of Competence as required under EU and UK legislation [24] aims to help establishments set up a robust framework for training, supervision and assessment of competence.**
- **The e-learning resource 'Euthanasia in laboratory animals' aims to deliver EU Module 6 (by Newcastle University and NC3Rs) (cbctraining.ncl.ac.uk/eM-EU6/story_html5.html).**

Even experienced personnel may struggle to stay up to date with current methods and thinking – so using new materials to question current practice should not be seen as 'conflict', but viewed positively as an integral part of the Culture of Care [25]. Discussions should include the designated veterinarian and the persons responsible for overseeing animal care and welfare, ensuring that species-specific information is available, and that staff are trained and competent (as specified in Directive Articles 24 and 25; named persons in the UK). Including staff with different roles will help to ensure that issues can be fully discussed and resolved, and the results can be reported and disseminated if appropriate.

Where relevant experience does not exist within the team, it may be possible to draw on the skills of another group within the facility or externally, including overseas if necessary. The staff listed above, and local committees such as the AWB, are likely to be able to assist with identifying groups from which high quality scientific and welfare-friendly theory and skills can be acquired. Keeping staff informed of technical and scientific developments regarding the Three Rs (replacement, reduction and refinement) is also a task of the AWB, and the UK AWERB should support staff with respect to provision of appropriate training.

3.5 Invest in animal monitoring

Animal monitoring software and hardware technologies are rapidly developing, providing new opportunities to increase the frequency and level of observations, and to seek earlier indicators of ill health (such as temperature changes, activity anomalies or decreased food intake). A web search using terms such as '(lab) animal behaviour/behavior' and 'monitoring' or 'analysis' will yield a number of companies, and internal and external colleagues should be able to advise on those whose equipment they have used. These systems can deliver a very high volume of data, so it is a necessity to have bioinformatics experience within the team establishing such equipment.

Additional harms should not be imposed on animals when using these systems, e.g. any products that require single housing of social animals (because they cannot recognise individuals) or do not permit

* science.rspca.org.uk/sciencegroup/researchanimals/severesuffering/resources/reports

enrichment (because it will obscure animals for video recording) should be avoided. State-of-the-art systems are now available that use RFID (radio frequency identification) chips to identify individuals and/or sophisticated software that can track animals even if they pass behind other individuals or objects.

Harms caused by any invasive elements will obviously need to be considered against the benefits of increased ability to implement humane endpoints. For example, body temperature can be monitored using an implanted telemetry device or RFID chip. If the device is implanted solely to help define or implement humane endpoints, rather than as part of the scientific protocol, a harm-benefit assessment will be required. Device implantation is an invasive procedure which may require sedation or general anaesthesia (depending on the size of the device and the animal), but this may be justifiable if accurate and timely body temperature data can be used to prevent severe suffering and mortality. Developments in non-contact, infrared thermometry have made it possible to avoid using implanted devices. For example, this technique has been used to achieve a 41 % reduction in mortality in a murine model of septic shock [26].

While body temperature is clearly a very useful potential indicator of survival/mortality, authors emphasise that it is essential to understand, and account for, the many factors that influence this. Such factors include ambient temperature, the presence and amount of nesting material, circadian rhythms, age, strain, and so on.

3.6 Develop in-house data and/or record mining

Information, data and records that can help to avoid mortality may already be available, and if they are not, it is helpful to consider whether these could be feasibly obtained. The workshop identified three ways of reviewing such information, with the aim of going beyond the reviews of welfare assessment set out in section 3.2 and taking a more strategic, top-level approach.

3.6.1 Review by the local AWB, AWERB, ACUC or ethics committee

Regularly discussing fates of animals with the local committee, including animals 'found dead', can help to identify causes of mortality and possible approaches to preventing these. For example, one author sat on an AWERB that reviewed the fate of all animals at the establishment, including unexpected mortality and animals humanely killed because an endpoint had been reached, or because they were surplus to requirements. The unit manager would present tables with lists of fates for each project and explain any areas of concern to the AWERB. This was a standing agenda item, which helped to significantly reduce avoidable wastage as well as so-called 'found deads'.

3.6.2 Structured review of assessment records

Regular, structured reviews of welfare assessment records and outcomes can help to identify reliable predictors of mortality, as referred to in section 3.2. For example, animal technologists can be tasked with reviewing records, in conjunction with expert statistical advice, to see whether there are significant correlations between indicators noted on welfare assessment sheets and subsequent mortality. Conversely, it may be possible to identify indicators that are not good predictors of mortality or welfare in general, and remove these from welfare assessment protocols if they are not adding value.

3.6.3 Data mining approaches

Informatics databases used to record data from animals can be a substantial resource for understanding mortality, for example when evaluating how mortality rates can be affected by changes in breeding and husbandry practices or environmental conditions. This is especially helpful where very large numbers and/or GA strains are involved.

It is understood that GA strains are more likely to have unpredicted effects than non-GA strains, and a small proportion of GA strains will potentially have higher mortality rates. Quantifying this in a meaningful way requires calculations of current or 'baseline' mortality, and a definition of what constitutes elevated mortality. In mutant colonies, the appropriate baseline 'control' group could be the wild-type mice bred within the mutant colony (e.g. littermates or other wild types from the same matings as the mutants). Alternatively, wild-type colonies of the same genetic background can provide a good reference for mutant strains. Note that the current mortality rate should not necessarily be regarded as 'acceptable' (see section 2), but it is necessary to calculate this to identify any welfare problems and detect any changes.

Some factors to include when considering calculation of the appropriate baseline mortality are set out in box 1. While this is mostly applicable for larger mouse facilities, equivalent record-keeping will be important for other species.

Box 1: Examples of factors to include when calculating baseline mortality

- Age range of the animals
- Background strain of GA animals
- Sex of animals
- Temperature in the room and/or cage
- Genetic status of animals
- Litter history of the dam
- Age of the dam, for neonates/juveniles
- Presence of the sire, for neonates/juveniles
- Animal's diet
- Whether animals are breeding stock or being used in procedures
- The room in the facility in which the animals are held

It is good practice to include as many parameters in the database as possible, to establish a representative baseline rate and ensure sufficient data to inform opportunities for intervention if necessary. For example, the room in which the animals are held can turn out to be critically important if data analysis reveals that mortality is higher in some rooms than others, all other recorded factors being equal. In addition, where routines may vary between different areas of a facility or over time (such a change in the diet supplier), these should be accurately logged in order to interpret potential differences. A suitable recording system should include easily interpretable terms to annotate animal conditions and fates. Box 2 overleaf sets out a practical example using informatics in a large facility housing many strains of GA mice.

Box 2: A database informatics approach to understanding neonate mortality

A facility housing large numbers of mice (over 95,000 individuals, in 475 colonies) set up an in-house informatics database which enabled the age at death for individual mice to be annotated.

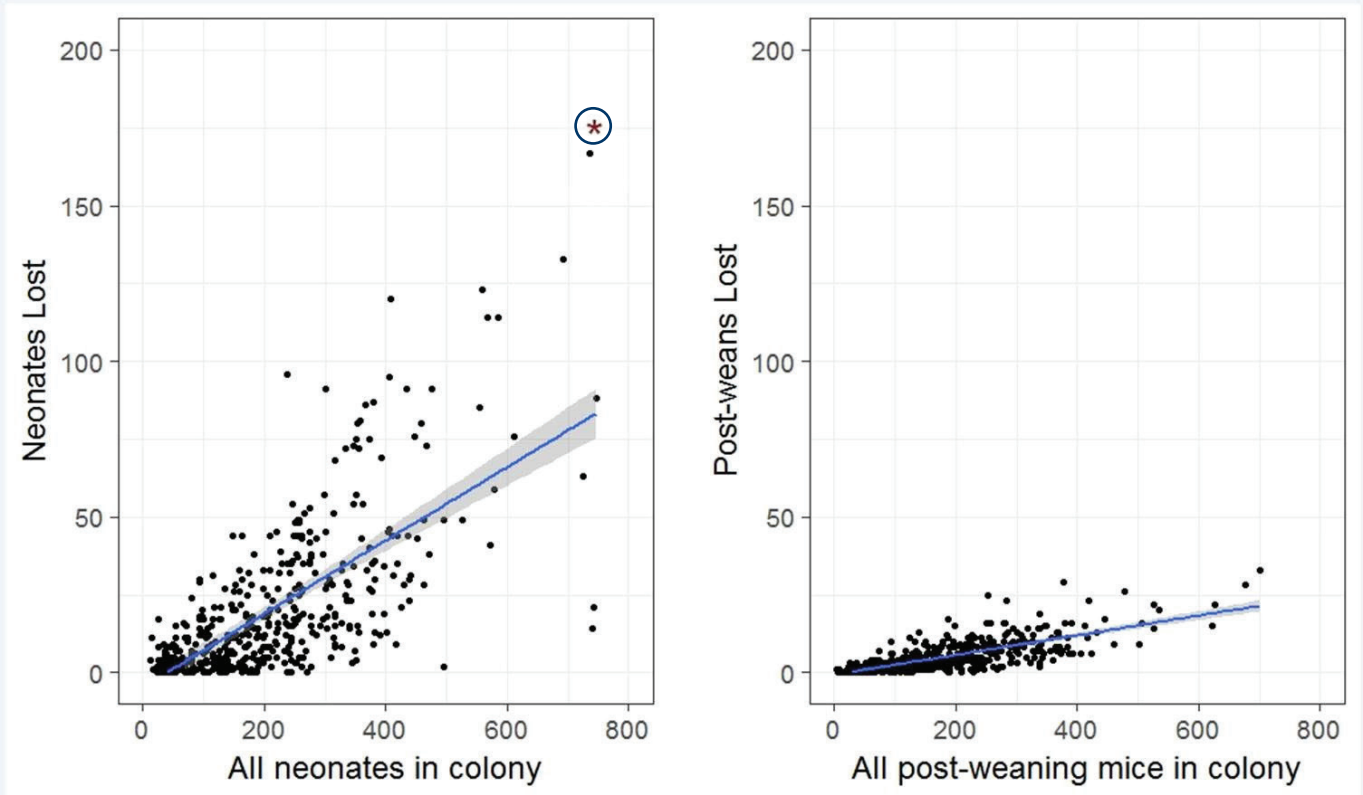
The goal was to establish the baseline level of mortality for GA-colonies in the facility. This was done by using the database to count the numbers of mice in each colony that were recorded in the database as having died for one of the following reasons: found dead; missing from the cage (in the case of neonate litter losses); euthanased because they were sick; or humanely killed because they reached an welfare endpoint unrelated to the scientific study. These mice are termed Mice Lost. The next step was to count the total number of mice generated in each colony. Using these two pieces of information, graphs were made comparing the numbers of Mice Lost to the total number of mice in the colony. By calculating and plotting the mortality level for each colony and assessing the gradient of the best fit line, an overall mortality rate of 19 % was determined.

It was suspected that very young mice may have higher mortality, so the database was used to analyse how mouse age affected mortality rates. To do this, the database was used to separate mice into one of three time frames according to their age at death; neonate (up to 6 days), pre-weaning (between 6 and 21 days) and post-weaning (over 21 days). The data were plotted on a graph to determine the rate of mortality. This analysis indicated that birth to 21 days of age was the critical period for mice with respect to mortality. Further analysis found that mortality was 12 % for neonates (up to 6 days), so most mortality was within this neonate phase (see fig. 1).

This is just a summary of a more complex analysis. Plotting graphs in this way highlighted colonies which deviated substantially from the baseline rate, so required investigation on animal welfare grounds, such as the colony marked with the circled asterisk in figure 1a. When prioritising which colonies to investigate, it is important to note the large variability in the mortality rate between different GA colonies. Nevertheless, a comparison of the graphs indicates that mortality rates in neonate mice were almost four times that of post-weaning mice. Neonate mortality in mice is highly dependent on multiple factors (e.g. age of the dam, trio vs. duo matings, presence of the sire, the genotypes involved in the mating). Further investigations to pinpoint and improve mortality rates need to compare specific circumstances across all colonies.

However, in absolute numbers, this database analysis shows that efforts in preventing mortality at an institutional scale may be best directed towards examining routines and procedures involving mouse neonates (see section 4.2).

Figure 1. Using database information to calculate overall mortality rate in a colony of GA mice



Legend: Scatterplot showing number of neonate (up to 6 days, 1a) or post-weaning mice (over 21 days, 1b) lost due to early mortality compared to the total number of mice in the colony. Each point represents a single colony from 475 GA colonies. The slope shows the best fit line determined by linear regression. The slope of the line gives the mortality rate, 11.8% for neonates and 3.1% for post-weaning mice. The shaded area shows the 99% confidence interval for the best fit line. The circled asterisk shows an example of a colony with a particularly high mortality rate that required investigation.

Once the baseline mortality rate has been established, and compared to the mortality rates of the animals of interest, the next step is to decide what degree of increased mortality is worthy of further investigation, balancing the risks of excessive false alarms against missed opportunities for interventions. The example in box 2 demonstrates two things: (i) efforts to prevent mortality ought to be focused according to the ages of the animals, and (ii) there is large variability in the mortality rate between different GA colonies. This approach should be applicable in many other settings and species.

Animal management databases thus provide valuable opportunities to calculate mortality rates, but analyses will be limited by the quality and the richness of the data input. Some facilities use bespoke, in-house monitoring software and data analysis programmes, whereas others use off-the-peg, commercially available monitoring software and spreadsheets. Whatever the level of surveillance and analysis, adequate commitment and resources are essential in order to collect sufficient good-quality data; post-hoc analysis using incomplete or poor quality data sets will not be as useful. Finally, complete and accurate use by all relevant staff is essential to derive maximum value and minimal bias from the data. Ideally, this data can be formatted for reporting to investigators or bodies such as the AWERB, AWB or ACUC to help implement both reduction and refinement.

3.7 Review regulatory requirements and their interpretation

There is a widespread and increasing recognition that dying animals are not good models of any specific aspect of toxicity, apart from maybe death itself. Now that more subtle indicators of toxicity can be detected, and the cause of toxicity is more important than the outcome, data from dying animals is becoming even less relevant. In the interest of good science, committees and bodies responsible for the design of test protocols and/or guidelines (e.g. the OECD*) are making a concerted effort to eliminate mortality as an endpoint wherever possible.

For example, in recent updates to guidelines in which high doses are likely, such as acute and genetox studies, death as an endpoint has been considered (and for the most part avoided), either by improving the dose response assays to determine the Maximum Tolerated Dose (MTD) or by limiting the top dose required. Any guidance that still requires death as an outcome should be updated as a matter of urgency, not only in the interest of animal welfare, but also to provide robust scientific data. For setting the MTD, development should continue of in vitro assays to predict aspects of toxicity (non, low, medium and high), so that the start of the MTD testing is done at a sensible point.

Researchers should be mindful that regulators are of course interested in the outcome of a test, and are also interested in achieving, and proving, adequate exposure to the test article. The latter may be wrongly interpreted as a requirement for death as an outcome. However, as already discussed, the overall premise and interpretation of toxicity studies and their endpoints are becoming far subtler, with biomarkers becoming more commonplace as sensitive indicators of both toxicity and efficacy. Consequently, the drive to develop and improve in vitro methods to identify potential biomarkers is a driver for the refinement and ultimately the replacement of in vivo tests. Furthermore, sharing information and approaches to refining and replacing regulatory tests will greatly improve the general acceptance of alternative approaches by regulators.

3.8 Challenge regulators' requirements

There is definitely an impetus to remove the use of death as an endpoint in regulatory studies. However, there are still a number of extant OECD guidelines of this type, and often a residual mindset that prefers death to other measures because of its lack of ambiguity.

Most studies that require death as an endpoint are those for the determination of acute toxicity. The demonstration that single dose acute oral toxicity testing added no significant scientific value to information from other studies led to its removal from the international guideline, ICH M3 [27], for pharmaceutical drug development. However, other sectors, for example the chemicals and pesticides industries, still adhere to practices such as the US Environmental Protection Agency's 'six pack' of tests [28]. There are moves to replace this, but meanwhile adherence to death as an end point will depend on the intended purpose of the testing.

There are two potential approaches to avoiding death in those regulatory tests that currently require this – replace the test with a humane alternative, or challenge the regulatory requirement and make a case for a humane endpoint.

Attempts to replace acute toxicity tests have not been easy nor resulted in as much change to practice as one might wish. In some sectors and regions, alternative approaches are accepted (e.g. cosmetics). In contrast, taking fish acute toxicity tests as an example, whilst the 'threshold approach' [29] (which could provide a 40% reduction in the number of fish used) is included in guidance on risk assessment of agrochemicals from the European Food Safety Authority (EFSA) [30], there is no recommendation included as to how this approach should be integrated into test packages. Similarly, use of the Fish Embryo Acute Toxicity Test (OECD TG236) [31] to fulfil REACH requirements for acute fish toxicity data has been set back by a report commissioned by the European Chemicals Agency (ECHA) which states this test is only suitable as part of a weight-of-evidence package due to a lack of quality data generated using the test guideline [32]. It is likely that registrants will therefore continue to submit data generated using adult fish, hampering the collection of further data that could support the use of TG236.

Another approach to reducing mortality is to replace death by an alternative endpoint, and in this connection, the use of 'evident toxicity' has had some success. Evident toxicity relies on the detection by animal care staff of behavioural and other signs which predict that death would occur at the next highest dose or concentration of the test chemical. Note, this is a step before moribundity – evident toxicity comprises clear signs that predict death or severe toxicity at the next highest concentration, and does not warrant euthanasia. The test procedure is essentially similar to the Acute Toxic Class method already approved for oral, dermal and inhalation toxicity testing. The use of evident toxicity as an end point was accepted in 2002 in the Acute Oral Fixed Dose Procedure (OECD TG420) [33] but in the absence of guidance on what constitutes evident toxicity, the test has not gained widespread use.

This concern surfaced repeatedly in efforts led by the UK NC3Rs to gain approval for the inhalation toxicity equivalent, the Fixed Concentration Procedure (FCP). Only extensive retrospective analysis of a large historical data set, and detailed comparison of the classifications made by the FCP with other methods succeeded in overcoming mistrust of evident toxicity and led to acceptance of

* Organisation for Economic Co-operation and Development, a regulatory body: oecd.org/about/

the draft guideline in 2017 (OECD TG433) [34, 35], 13 years after its first publication [36]. Curiously, revision of the dermal toxicity test guideline (OECD TG402) to incorporate evident toxicity instead of death as an endpoint (the dermal equivalent of the FCP) has also now been approved [37], but without a similar body of supporting evidence. It is a concern that the OECD approval process seems so variable not only in time required but also in the evidence base demanded. However, the slow adoption into use of TG420 may indicate that where a method is approved without extensive and clear guidance on its use, it may struggle to gain wide acceptance.

One clear positive outcome from the experiences with gaining approval for TG433 is that the supporting evidence was based on analysis of historical data alone, an encouraging precedent for other changes to current regulatory practice.

Similar activities are currently underway to provide a definition of moribundity in order to update and refine acute fish toxicity studies (TG203). This will involve identifying clinical signs displayed in individual fish that can indicate impending death within the same study, and thus allow the animal to be euthanased, avoiding death as an endpoint.

4. Issues that may need additional consideration

The group identified some additional issues that may need to be taken into account when implementing some approaches to avoiding mortality, or that could require discussion with the local ethics or animal care and use committee. These are listed below, with some suggested actions.

4.1 Monitoring animals can cause harm

Monitoring animals can involve disturbing them, which can risk causing discomfort or distress if the animals are used in severe procedures or are becoming moribund. Assessing some clinical signs used as indicators that animals are likely to die can thus have a negative welfare impact. For example, the 'scruff test' in the lung tumour model (see case study II) is at best uncomfortable and, at worst, distressing for the animal. Similarly, body temperature measurements, blood sampling for biomarkers and other invasive measures, or monitoring that requires disturbing animals from their routine behaviours, will all cause a degree of harm. In some instances, the stress caused by the additional interference could increase morbidity or mortality, which raises animal welfare and ethical issues and may interfere with the scientific goals of the study. Also, where a model is under development, closer monitoring than usual may be required in order to identify definitive clinical endpoints. This may cause additional harm to the individuals being monitored, in order to reduce harm to those animals subsequently used.

What to do:

It is important to recognise that some monitoring techniques, in some situations, can cause a degree of discomfort, pain or distress. This obviously presents a dilemma, if the imperative to avoid mortality (or to avoid approaching an endpoint) could result in causing more suffering in the immediate short-term. Allocating time to focus on refining a monitoring technique is an obvious first step. For example, it is possible to train animals to climb onto scales for weighing without being handled, and if a technique requires anaesthesia the least distressing agents and delivery methods can be researched and used.

The harms and benefits associated with using the monitoring technique can also be considered by relevant staff and/or the local committee, and outcomes can be monitored to see whether the technique adds value. This could include defining indicators to identify when excessive stress has been caused by the monitoring technique, reviewing whether these indicators have been observed, and seeing whether mortality is reduced in practice. Correlating more 'invasive' indicators with non-invasive observations may also help to assess how much value is added by the former. In some cases, it may be justifiable to use implanted telemetry devices to monitor relevant parameters remotely, but surgical implantation is an invasive procedure and a careful harm-benefit assessment will be necessary [38].

4.2 Observing neonatal rodents

There is a common belief that disturbing a dam with young pups, particularly in mice, will induce the dam to kill and eat some or all of the pups. In many cases this has led to facilities having a policy of not inspecting female mice in the first few days post partum. However, the working group was unable to find any published studies that have evaluated whether human disturbance directly causes cannibalism. Alternative explanations for dams killing and/or eating pups could be that the dam may have killed her pups regardless of being disturbed by a human, or may have eaten pups that were already dead from other causes. Unless neonatal pups are checked, and at least counted, perinatal mortality cannot be properly evaluated or understood. This is a critically important animal welfare and ethical issue; for example, the Alive Pup Project has estimated that an additional 1.7 million mice would need to be bred every year within the European Union to compensate for an early litter loss of 25% [39]. New analyses that involve monitoring neonates earlier than previously may lead to an apparent increase in mortality, however this is likely to be because the losses were previously undetected.

What to do:

The benefits of checking and monitoring neonatal rodents should be objectively considered against any potential risks to the dam and pups. Anecdotal accounts of pup loss caused by disturbance due to monitoring should be taken seriously, but also carefully and critically reviewed, rather than taken as an indicator that these animals should not be monitored. In practice, pup mortality and litter loss may be due to many risk factors other than disturbance, including mis-mothering, problems with lactation, early life stress, issues with maternal nutrition, genetic status of the dam and pups, general welfare status of the dam and whether or not the sire is present (see boxes 1 and 2). Neonatal mice are routinely checked at one author's facility with no negative impacts, as demonstrated by comparing live birth and viability data.

There may be a case for refining observation practices; for example consideration could be given to the time of day at which litters are checked and whether this is during the animals' active or inactive periods; and whether nests are opened and animals handled, or it is possible to slowly raise the closed cage and look up through the floor. It is essential to gather adequate data to enable robust analysis of the impact (positive or negative) of any changes in practice.

If there is an apparent increase in mortality, this is may well be an artefact of better detection rather than an actual increase. Although increased mortality may seem negative or demoralising, it indicates enhanced monitoring, which should lead to better animal care and welfare because it provides a more effective benchmark for improvements.

4.3 Monitoring animals more effectively can lead to difficult decisions

Monitoring animals more effectively is an essential goal from animal welfare, scientific and ethical perspectives. However, using more effective indicators of mortality, and the data mining approaches set out in 3.6.3, may yield new information that leads to difficult decisions. For example, when using more sensitive indicators it might become apparent that a severity limit (or humane endpoint) has been reached earlier than anticipated, at a time when data essential to the project have not yet been obtained.

What to do:

Legal requirements should always be met when severity limits are approached or reached, and the agreed criteria for intervention (usually euthanasia) should be applied. For example, the UK project licence includes a standard condition (18) to the effect that the licence holder must inform the regulator if a severity limit has been (or is likely to be) breached [40]. If it appears that the project is of a higher severity than was previously thought, one option is to refine the endpoint. If this would mean that the objectives were no longer achievable, an alternative option is to discuss this with the local committee, regulator and others with expertise in experimental design and refinement – it may be feasible to redesign the experiment and implement other refinements. If none of the above approaches are feasible, then there may be justification for applying to increase the harms permitted in the license, if the likely benifits exeed these. This should be critically considered and should not involve causing prolonged, severe suffering.

4.4 Practicalities relating to resources

It is easy to assume that a greater level of observation, and the use of earlier indicators, is a straightforward way of reducing mortality. However, in practice it is often necessary to allocate resources pragmatically, focussing efforts where needs are greatest. For example, the resources required for continuous observations of animals by video monitoring and analysis would currently be outside the reach of many research grants and animal facilities.

What to do:

The fact that resources are finite does not preclude the opportunities (or responsibility) to look for easily measurable indicators, or to identify times when heightened surveillance is possible and appropriate, e.g. in the case of a novel GA line or a new treatment. Furthermore, lower-cost alternatives to expensive and complex home-cage monitoring techniques are increasingly available. These can help to set up a triage system, in which animals showing abnormal levels of activity could be monitored in more detail, allowing efforts to be focussed on times when additional care and attention is required. Furthermore, many funding bodies emphasise the importance of implementing the Three Rs, so there is a strong case for including equipment (or resources, e.g. additional staff hours) that would help to implement refinement in grant applications.

Lower-tech (and therefore lower-cost) alternatives can also be considered. For example, one group working on a bacterial challenge experiment habituated their mice to receiving a chocolate muffin or cucumber treat at set times of the day, well before procedures began. Throughout the experiment, they could use the first time point when a mouse did not leave the nest for their treat as the indicator that the animal would not recover, so could be humanely killed. This method effectively reduced mortality and improved data output, because animals were not found dead.

4.5 Interpreting indicators in aged animals

Ageing and longevity studies pose particular issues with respect to reducing mortality, especially if the objective of a project is to assess the causes and modifiers of death. Further challenges arise when monitoring animals on long-term studies undergoing interventions such as diet restriction, drug treatment or genetic alterations. In mice, some inbred lines accumulate pathologies such as tumours, behavioural anomalies and skin disorders, and it is likely that any control or wild type groups will also be subject to suffering as they age. Furthermore, many indicators of ill-health in juvenile animals may not be applicable in ageing cohorts. Taking mice as an example, signs such as hair thinning (see picture below) or a reduction in daily activity would be more worrying in young versus much older animals.

This raises a number of ethical and animal welfare issues. Defining humane endpoints for aged animals obviously requires a balance between avoiding age-related suffering or death on the one hand, and not euthanasing animals before sufficient data have been gathered on the other. Although ageing is a 'natural' process, it can be argued that mice would not normally live as long as they do in the laboratory, so humans are artificially prolonging their lives and exposing them to age-related pathologies. The group's consensus view is that pathologies associated with ageing should be regarded as harms, when considering 'cumulative severity' in the context of the animal's life experiences [41] and when implementing refinements and humane endpoints.

What to do:

Ageing animals (whether they are controls or animals where an intervention has taken place) should obviously be afforded the same level of care and unbiased assessment of pain, suffering, distress and lasting harm as any other animal. This is especially important because any harms can be potentially very long-lasting. In mice, there is an increasing literature on the use of 'frailty indexes', which include parameters such as grip strength, walking speed and physical activity [42]. These can be useful tools to help assess declining condition and define age-related humane endpoints.

Sharing information

At the time of writing, the level of avoidable mortality within animal research and testing as a whole is unfortunately unknown. Greater openness and communication are required with respect to mortality rates and causes, and approaches to reducing these, some of which are listed below.

- **Authors can include relevant information in publications, such as protocols for assessing and monitoring animals, and predicting impending mortality, using suitable keywords to help ensure that others can find the information. The increase in online journals and supplementary information means that word limits are no longer an obstacle to providing more detail. Authors should also be open about mortality during a project; e.g. the ARRIVE guidelines require an explanation if any animals are not included in the analysis.**
- **Journal editors and reviewers can also insist on the above.**
- **Posters and presentations can include information about mortality and how this was avoided; many poster presenters also provide A4 flyers of their posters, which are a helpful way of disseminating information and contact details.**
- **The non-technical project summaries, required as part of project authorisation applications within the UK and EU, can also mention how mortality will be avoided in high-risk projects, and these can be made publicly available via searchable databases.**

5. Role of the local committee: AWB, AWERB, ACUC or ethics committee

Institutional bodies such as ethics committees, AWERBs, AWBs and ACUCs can play an important role in avoiding and reducing mortality, since many of their tasks relate to this aim.

For example, relevant EU AWB and UK AWERB tasks include:

- **advising staff on matters related to animal welfare, in relation to their acquisition, accommodation, care and use;**
- **advising on the Three Rs;**
- **establishing and reviewing internal operational processes regarding monitoring, reporting and follow-up in relation to animal welfare; and**
- **following the development and outcome of projects, taking into account the effect on the animals used, and identifying and advising on elements that further contribute to the Three Rs.**

It is helpful for individual committees and bodies to review these tasks, and reflect on how well they are fulfilling these and contributing to avoiding deaths. Further explanation and practical guidance on addressing the tasks can be found in the relevant EU Working Document [43] and sections 6 and 10 of the RSPCA/LASA Guiding principles on good practice for AWERBS [44]. For retrospective review*, it is important to develop an effective process (perhaps by holding a workshop to help develop the approach locally, with input from animal technologists, the veterinarian and scientific staff), and to develop user-friendly documentation and channels for feedback. The PREPARE guidelines also include some useful criteria and topics that are relevant to retrospective review (**norecopa.no/prepare/prepare-checklist**).

Examples of specific issues to consider include:

- **how effectively actual severity was predicted;**
- **whether anything has changed which might alter the original harm-benefit analysis;**
- **how well welfare assessment sheets and monitoring procedures are working; and**
- **whether there were unexpected harms.**

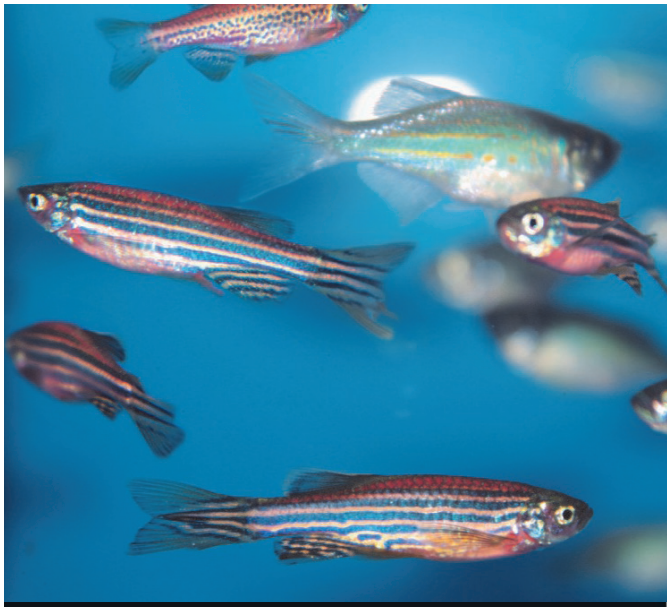
* Note that retrospective review is applied to all projects, whereas formal retrospective assessment is required for all projects using non-human primates, cats, dogs or Equidae, and those involving severe suffering.

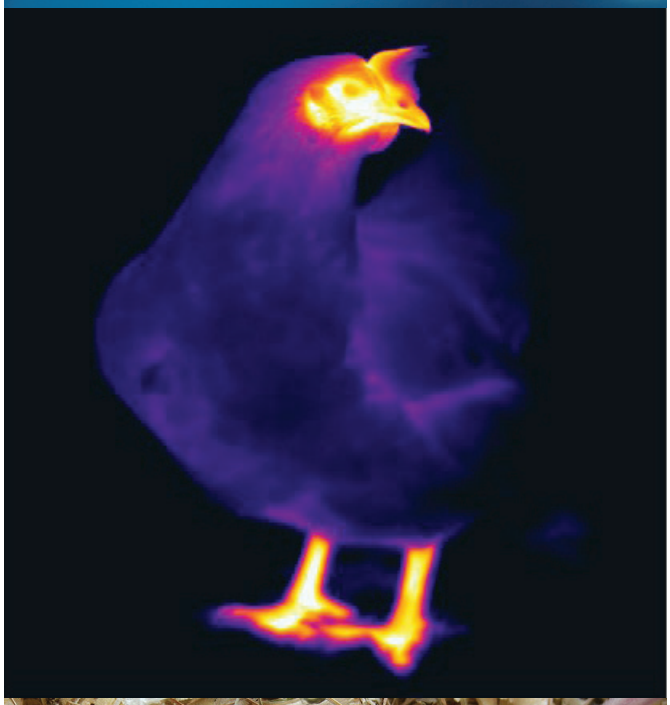
6. Wish list

The Working Group's discussions were based on currently available resources, practice and technologies, but the authors also identified a 'wish list' of developments they believe would further contribute to avoiding mortality.

These are listed below.

- Species-specific and disease-model specific databases of indicators to help predict deaths; ideally, these would be driven by controlled lexicons, such as Mouse Welfare Terms, Fish Welfare Terms or OBO MPath (obofoundry.org/ontology/mpath.html).
- Objective validation of species-specific 'frailty indexes', or equivalent, for predicting impending mortality.
- Guidance to help facilities to implement data mining and set up informatics databases.
- Cheap, user-friendly software to set up and analyse informatics databases.
- Increased use (e.g. within drug discovery) of mechanism-based model approaches with less severe endpoints and less risk of mortality, moving away from animal 'models' of human disease that recapitulate disease symptoms.
- Active, strategic efforts to replace strains (conventional and GA) with inherent, significant mortality.
- Development and validation of multi-factorial, species- and model-specific assessment and monitoring systems to help predict impending mortality in high-risk models and situations.
- Research to empirically evaluate whether assessing neonatal rodents leads to mortality, and if so, which factors contribute to this and how to address any problems and make decisions regarding monitoring protocols.
- Further research and development into more sophisticated, cheaper, automated systems for continuous, real-time animal monitoring (e.g. nc3rs.org.uk/rodent-big-brother).
- Mechanisms to share results of the kinds of pilot studies, and studies in parallel, described in this report (section 3.3).
- A library of technologies, techniques and approaches to avoiding mortality.
- Regulatory bodies jointly committing to end requirements for death as an endpoint within tests that currently require this, e.g. in some toxicity testing and in vaccine potency tests.
- Training resources specifically tailored to help researchers and animal care staff develop strategies designed to avoid death, e.g. relating to welfare assessment, monitoring and humane endpoints.





7. Action points

These action points, taken from the report, set out some principles and initiatives that can be put forward and acted on by scientists, animal technologists and care staff, veterinarians and members of bodies such as ethics committees, AWERBs, AWBs and ACUCs.

- **Make a commitment to reviewing mortality and seeing whether action is needed, regardless of whether there is currently a perceived problem.**
- **Challenge the concept of an 'acceptable rate' of mortality.**

and

- Regularly review literature relating to welfare assessment, including noting descriptions in methods sections of 'mainstream' publications.
- Discuss welfare assessment with internal and external colleagues, including the local ethics and/or animal care and use committee.
- Instigate or conduct reviews of current in-house welfare assessment protocols, to evaluate how effectively these are detecting early indicators of mortality.
- If mortality rates are, or are likely to be, a problem, explore the potential for pilot studies, or studies in parallel, to address this.
- Explore the potential for 'data mining' approaches to review causes and times of mortality.
- Instigate or conduct a review of in-house training in conducting procedures, welfare assessment and humane killing, including the syllabus, assessing and maintaining competency, and the potential to use new training aids.
- Actively share approaches to avoiding mortality between research groups and establishments, for example via user groups and communication networks such as the UK AWERB Hub system and AWB platforms within the EU.
- Vigorously challenge requests from editors or peer reviewers to include mortality data in the absence of an explicit regulatory requirement.

If working in a regulatory environment:

- Critically review regulatory requirements and how these are interpreted – can endpoints be refined?
- If there is a current regulatory requirement for death as an endpoint, explore ways of challenging this, in the short and long term.
- Contact regulatory bodies and request that they delete tests requiring death as an endpoint as soon as refined protocols, or humane alternatives, are validated and accepted.
- Share information on approaches to refining regulatory tests and their endpoints, e.g. by identifying biomarkers.

Actions relating to members of local committees:

- When working to avoid mortality, ensure that any difficult issues or dilemmas that arise are thoughtfully considered, e.g. by relevant local committees that include members with different expertise and viewpoints.
- Help to ensure that the local ethics and/or animal care and use committee is fulfilling all of its tasks that relate to refining endpoints and reducing mortality, and that it is efectively communicated with and supported.
- Suggest that the local ethics or animal care and use committee reviews all fates of animals
- Make (or support) a case for sufficient resource to support any reviews or refinements that will reduce mortality.

For journal editors and peer reviewers:

- Insist on full details of animal numbers and mortality throughout studies in papers submitted for publication, so that animals do not 'go missing' within publications because they have died.

For competent authorities:

- Ensure that non-technical summaries are transparent with respect to mortality and make them easily available to the public in searchable databases.

For all:

- Look at the 'wish list' in section 6 and see whether there is anything you can do to implement or support any of the suggested actions.

Acknowledgements:

We thank the University of Cambridge for providing the meeting venue and catering for both workshops. We are also grateful to the following people for reviewing the workshop report: Maggy Jennings, Patrick Sinnett-Smith, Adrian Smith and Jane Smith. Thanks to Belinda Farnfield for participating in the first workshop.

References:

1. Lilley E, Armstrong R, Clark N et al. (2015) Refinement of animal models of sepsis and septic shock. Shock 43: 304-316
2. Toth LA (2000) Defining the moribund condition as an experimental endpoint for animal research. ILAR Journal 41: 72–79 **doi.org/10.1093/ilar.41.2.72**
3. European Commission (2012) Working Document on A Severity Assessment Framework. EC, Brussels. **ec.europa.eu/environment/chemicals/lab_animals/pdf/Endorsed_Severity_Assessment.pdf**
4. European Commission (2013) Examples to Illustrate the Process Of Severity Classification, Day-To-Day Assessment and Actual Severity Assessment. EC, Brussels. **ec.europa.eu/environment/chemicals/lab_animals/pdf/examples.pdf**
5. Hawkins P, Morton DB, Burman O et al. (2011) A guide to defining and implementing protocols for the welfare assessment of laboratory animals: Eleventh report of the BVA(AWF)/FRAME/RSPCA/UFAW Joint Working Group on Refinement. Laboratory Animals 45: 1-13 **journals.sagepub.com/doi/pdf/10.1258/la.2010.010031**
6. Smith D, Anderson D, Degryse A-D et al. (2018) Classification and reporting of severity experienced by animals used in scientific procedures: FELASA/ECLAM/ESLAV Working Group report. Laboratory Animals 52(1S): 5–57 **journals.sagepub.com/doi/pdf/10.1177/0023677217744587**
7. F.I.S.H. – Fish Indicators of Stress and Health website (University of Liverpool) **liverpool.ac.uk/integrative-biology/research/adaptation-to-environmental-change/fish-indicators-stress-health/about/**
8. Martins CIM, Galhardo L, Noble C et al. (2012) Behavioural indicators of welfare in farmed fish. Fish Physiol. Biochem. 38: 17–41 **ncbi.nlm.nih.gov/pmc/articles/PMC3276765/**
9. Percie du Sert N, Alfieri A, Allan SM et al. (2017) The IMPROVE Guidelines (Ischaemia Models: Procedural Refinements Of in Vivo Experiments). Journal of Cerebral Blood Flow & Metabolism 37(11): 3488–3517 **journals.sagepub.com/doi/full/10.1177/0271678X17709185**
10. Workman P, Aboagye EO, Balkwill F et al. (2010) Guidelines for the welfare and use of animals in cancer research. Br. J. Cancer. 102: 1555–1577 **ncbi.nlm.nih.gov/pmc/articles/PMC2883160/**
11. Kadl A, Pontiller J, Exner M et al. (2007) Single bolus injection of bilirubin improves the clinical outcome in a mouse model of endotoxemia. Shock 28: 582-8
12. Ullman-Culleré MH & Foltz CJ (1999) Body condition scoring: a rapid and accurate method for assessing health status in mice. Lab. Anim. Sci. 49: 319-23
13. Kamp MA, van Lieshout JH & Dibué-Adjei M (2017) A systematic and meta-analysis of mortality in experimental mouse models analyzing delayed cerebral ischemia after subarachnoid hemorrhage. Translational Stroke Research 8(3): 206–219 **link.springer.com/article/10.1007/s12975-016-0513-3**
14. Dewi S, Aune T, Aasen Bunæs JA et al. (2014) The development of response surface pathway design to reduce animal numbers in toxicity studies. BMC Pharmacology and Toxicology 15: 18 **doi.org/10.1186/2050-6511-15-18**
15. Trammell RA & Toth LA (2011) Markers for predicting death as an outcome for mice used in infectious disease research. Comp. Med. 61: 492-8 **ncbi.nlm.nih.gov/pmc/articles/PMC3236690/**
16. Hankenson FC, Ruskoski N & van Saun M (2013) Weight loss and reduced body temperature determine humane endpoints in a mouse model of ocular herpesvirus infection. J. Am. Assoc. Lab. Anim. Sci. 52: 277–285 **ncbi.nlm.nih.gov/pmc/articles/PMC3690449/**
17. Trammell RA, Cox L & Toth LA (2012) Markers for heightened monitoring, imminent death, and euthanasia in aged inbred mice. Comp. Med. 62: 172–178 **ncbi.nlm.nih.gov/pmc/articles/PMC3364702/**
18. Ray MA, Johnston NA, Verhulst S et al. (2010) Identification of markers for imminent death in mice used in longevity and aging research. J. Am. Assoc. Lab. Anim. Sci. 49: 282-8 **ncbi.nlm.nih.gov/pmc/articles/PMC2877298/**
19. Vlach KD, Boles JW & Stiles BG (2000) Telemetric evaluation of body temperature and physical activity as predictors of mortality in a murine model of staphylococcal enterotoxic shock. Comp. Med. 50: 160-6 **ncbi.nlm.nih.gov/pubmed/10857007**
20. Hunter JE, Butterworth J, Perkins ND et al. (2014) Using body temperature, food and water consumption as biomarkers of disease progression in mice with Eμ-myc lymphoma. Br. J. Cancer. 110: 928-34 **doi.org/10.1038/bjc.2013.818**
21. European Commission (2014) A Working Document on the Development of a Common Education and Training Framework to Fulfil The Requirements Under the Directive. EC, Brussels ec.europa.eu/environment/chemicals/lab_animals/pdf/Endorsed_E-T.pdf
22. Home Office (2014) Advisory Notes on Recording and Reporting the Actual Severity of Regulated Procedures. Home Office, London **assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/662489/NotesActualSeverityReporting.pdf**
23. Home Office (2014) Severity Classification of Genetically Altered Animals Under the Animals (Scientific Procedures) Act 1986. Home Office, London **assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/662491/AdviceSeverityAssessmentGA.pdf**
24. Laboratory Animal Science Association (2016) Guiding Principles for Supervision and Assessment of Competence as required under EU and UK legislation, 2nd Edition. A report by the LASA Education, Training and Ethics Section (M Jennings and M Berdoy eds). **lasa.co.uk/wp-content/uploads/2016/09/LASA_supervision_and_competence_2016.pdf**
25. See: norecopa.no/more-resources/culture-of-care
26. Laitano O, Van Steenbergen D, Mattingly AJ et al. (2018) Xiphoid surface temperature predicts mortality in a murine model of septic shock. Shock 50(2): 226-232 **doi:10.1097/SHK.0000000000001007**
27. European Medicines Agency (2009) ICH Guideline M3(R2) On Non-Clinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorisation for Pharmaceuticals. EMA, London **ema.europa.eu/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-2.pdf**
28. US Environmental Protection Agency Data (2018) Requirements for Pesticide Registration. **epa.gov/pesticide-registration/data-requirements-pesticide-registration**
29. Creton S, Clook M & Wheeler JR (2014) Application of the threshold approach for acute fish toxicity testing to plant protection products: A proposed framework. Chemosphere 96: 195-200 **doi.org/10.1016/j.chemosphere.2013.10.015**
30. European Food Safety Authority (2013) Guidance on Tiered Risk Assessment for Plant Protection Products for Aquatic Organisms in Edge-Of-Field Surface Waters. **efsa.europa.eu/en/efsajournal/pub/3290**
31. OECD Test No. 236: Fish Embryo Acute Toxicity (FET) Test. **doi.org/10.1787/9789264203709-en**

References (cont):

32. Scholz S, Klüver N & Kühne R (2016) Analysis of the Relevance and Adequateness of Using Fish Embryo Acute Toxicity (FET) Test Guidance (OECD 236) To Fulfil the Information Requirements and Addressing Concerns Under REACH. Report ECHA-UFZ contract ECHA/2014/341. **echa.europa.eu/documents/10162/13639/fet_report_en.pdf**
33. OECD Test No. 420: Acute Oral Toxicity - Fixed Dose Procedure. **doi.org/10.1787/20745788**
34. OECD Test No. 433: Acute Inhalation Toxicity: Fixed Concentration Procedure. **doi.org/10.1787/20745788**
35. Sewell F, Ragan I & Marczylo T (2015) A global initiative to refine acute inhalation studies through the use of 'evident toxicity' as an endpoint: Towards adoption of the fixed concentration procedure. Regul. Toxicol. Pharmacol. 73(3): 770-9 **doi.org/10.1016/j.yrtph.2015.10.018**
36. Stallard N, Whitehead A & Indans I (2003) Statistical evaluation of the fixed concentration procedure for acute inhalation toxicity assessment. Human & Experimental Toxicology 22: 575-585
37. OECD Test No. 402: Acute Dermal Toxicity. **doi.org/10.1787/20745788**
38. Hawkins P (2014) Refining housing, husbandry and care for animals used in studies involving biotelemetry. Animals 4: 361-373 **mdpi.com/2076-2615/4/2/361/pdf**
39. See **approjectweb.wordpress.com/**
40. Home Office (2018) Advice Note: Animals (Scientific Procedures) Act 1986 Project Licence Standard Condition 18 Notification. Home Office, London **assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/705833/Standard_Condition_18_Advice_Note.pdf**
41. Animals in Science Committee (2017) Review of Harm-Benefit Analysis in The Use Of Animals In Research. **assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/675002/Review_of_harm_benefit_analysis_in_use_of_animals_18Jan18.pdf**
42. Whitehead JC, Hildebrand BA & Sun M (2013) A clinical frailty index in aging mice: comparisons with frailty index data in humans. The Journals of Gerontology: Series A 69(6): 621–632 **doi.org/10.1093/gerona/glt136**
43. European Commission (2014) A Working Document on Animal Welfare Bodies and National Committees to Fulfil the Requirements Under the Directive. EC, Brussels **ec.europa.eu/environment/chemicals/lab_animals/pdf/endorsed_awb-nc.pdf**
44. RSPCA and LASA (2015) Guiding Principles on Good Practice for Animal Welfare and Ethical Review Bodies. A report by the RSPCA Research Animals Department and LASA Education, Training and Ethics Section (M Jennings ed). **lasa.co.uk/PDF/AWERB_Guiding_Principles_2015_final.pdf**

Appendix 1: Agenda for the first workshop, held at the University of Cambridge on 19 September 2017

9.45 **Welcome and plan for the day – Penny Hawkins, RSPCA**

9.50 **ASRU presentation to set the scene – Kathy Ryder, ASRU**
Four talks on avoiding mortality and refining endpoints – success stories. Each talk followed by discussion to identify what people feel are the most generally applicable/transferable aspects of each, that they may be able to apply in their own field; are there some 'universals'?

10.10 **Predicting mortality and avoiding 'found deads', either in GA lines or from the perspective of having very large numbers of animals – Sara Wells, MRC Harwell; Chris Lelliott and James Bussell, Wellcome Trust Sanger Institute**

10.40 **Challenging death as an endpoint within a 'traditional' model that currently includes this. Lung tumour – Ngaire Dennison, University of Dundee**

11.10 **Predicting mortality in a disease model that includes a risk of sudden death – avian influenza. Sharon Brookes, APHA**

11.40 **A new OECD guideline for inhalation toxicity based on evident toxicity: trials and tribulations – Ian Ragan, NC3Rs**

12.10 **General discussion from morning, identify some elements for the report**

12.40 **Break**
Three brief presentations on models or procedures in which unpredicted mortality is still a problem. Group discussion as to whether these could be refined, considering (a) what might be done in the here and now, and (b) if money and technology (e.g. remote monitoring, behavioural recognition software) were no object.

1.10 **Predicting mortality in aged animals – Sara Wells**

1.30 **Fish in regulatory toxicology – death as an endpoint – Helmut Ehall, Envigo**

1.50 **Analysing previous data to help predict mortality – Belinda Farnfield, LAVA**

2.10 **Summing up key points and agreeing broadly what to put in report and how to publish it, tasks and timelines**

3.30 **End**

The second workshop was held on 1 October 2018 in order to finalise the report.

Published by:
RSPCA Research Animals Department,
Wilberforce Way, Southwater, Horsham, West Sussex RH13 9RS
Tel: 0300 123 0329 www.rspca.org.uk/researchanimals

For further copies of this booklet, please email: **research.animals@rspca.org.uk**
ISBN 978-0-901098-17-7