



National Centre
for the Replacement
Refinement & Reduction
of Animals in Research

Welfare webinar: Minimising animal suffering with refined humane endpoints and early interventions

Question and answer session

This Q&A document accompanies the recording of an NC3Rs welfare webinar 'Minimising animal suffering with refined humane endpoints and early interventions' held over two parts in September 2025. Speakers from across sectors explored how refined, study-specific humane endpoints and early intervention strategies can be implemented in different research scenarios to minimise animal suffering. Written answers to the questions submitted during the webinar are provided by the speakers below.

Please [watch the webinar recording](#) for full context and further information.

Speakers and topics:

- [Penny Hawkins, RSPCA – Bone marrow ablation and reconstruction.](#)
- [Amy Marriott, NC3Rs Toxicology team and formerly Stuart Ainsworth's lab at the University of Liverpool – Snakebite envenoming studies.](#)
- [Romalie Alen, Labcorp – Managing weight loss in drug studies, and Mark Blee, NC3Rs Toxicology team – Upper body weight loss limits in short term toxicity studies.](#)
- [Klas Abelson, University of Copenhagen – Minimising potential suffering in rodent models for pain research.](#)
- [Claire Gibson, University of Nottingham – Improving *in vivo* modelling of ischaemic stroke and animal welfare.](#)

Penny Hawkins, RSPCA – Bone marrow ablation and reconstruction.

1. You say that a temperature drop of >0.5°C with non-recovery or a further decrease can be used as a humane endpoint indicator, how would you recommend measuring body temperature and which part of the body should be measured? What would the expected time frame for recovery be?

Body temperature in small animals can be measured using thermal cameras, non-contact thermometry or RFID (Radio Frequency Identification) temperature transponders. FLIR (Forward-Looking Infrared) thermal imaging cameras, which plug into mobile phones, are another (less accurate) option. It is best to scan the back, ear or abdomen of the mouse, as these areas provide reliable readings that correlate well with core body temperature. It may not always be possible to calibrate systems with absolute temperatures, but it is the temperature changes that are important. The [refining bone marrow ablation and reconstitution in mice] [report](#) doesn't give an absolute time frame for recovery, as there did not appear to be sufficient data on this – it is a judgement to be made in each case, taking other indicators into account.

2. Where you state a body weight loss of 20% should be the maximum considered as a humane endpoint, but ideally 15%, why isn't 15% the default recommended maximum? When would it not be possible to use 15%?

Setting humane endpoints relating to body weight loss is always a dilemma with respect to higher-severity protocols. Nobody wants individual animals to suffer avoidably, but if animals are euthanised when they could have recovered in a short period of time, naive animals will then have to be used and the total harms to animals will increase. Decisions like this, about humane endpoints, should always be discussed with the attending veterinarian, animal technologists and the relevant institutional ethics or animal care and use committee, such as the UK Animal Welfare and Ethical Review Body (AWERB). In this case, the [refining bone marrow ablation and reconstitution in mice] report sets out that the default maximum weight loss should be 15%, with 20% as an **absolute** maximum if there is special justification (e.g. weight loss is appearing to stabilise, animal is feeding and drinking, it is believed that the animal will recover and the licence severity limit will not be breached). If the rest of the guidance in the report is followed, animals should not lose more than 15% of body weight and they should recover.

3. Where can I find out more about the RSPCA Focus on Severe Suffering initiative and what resources and tools are available?

Please see our [Focus on Severe Suffering](#) website – this includes downloads of [RSPCA reports](#) on refining specific procedures, our [Roadmap refinement tool](#) to help reduce and avoid severe suffering (this is useful for refining other levels of suffering too), details and reports of our [RSPCA events](#) and much more. We're always keen to hear how helpful the site has been to users, so feedback would be welcome at animalsinscience@rspca.org.uk

Amy Marriott, NC3Rs Toxicology team and formerly Stuart Ainsworth's lab at the University of Liverpool – Snakebite envenoming studies.

1. How many mice would be considered a good representative sample for a venom study using this refined model of envenoming?

This would depend entirely on the research question being answered. We would recommend conducting pilot studies with a small number to then apply power calculations to ensure the study would be statistically powered.

2. Regarding tail microsampling – what method do you use to perform the bleeds? You stated that the actual severities are recorded as moderate due to messing about with the tails, I was interested to know if you've looked at refining the method (rather than just reducing the number of bleeds) to reduce this actual severity?

We used a microsampling approach, making a small incision in the tail vein to collect blood samples of less than 50 µL, with typical volumes of 10-20 µL. Samples were collected using either a pipette or a capillary tube set to a defined volume, thereby minimising the impact on the animals. Although the study was classified as moderate due to the repeated sampling regimen, both the sample volumes and frequency of sampling were within NC3Rs and Laboratory Animal Science Association recommended limits for microsampling procedures.

3. Can you elaborate on the use of biomarkers and point-of-care devices to refine studies – what are the potential barriers to using these and where could I source them from?

Biomarkers and point-of-care (PoC) devices can play an important role in refining studies, particularly by enabling earlier, less invasive and more sensitive detection of biological changes. This can reduce animal use, minimise suffering and improve data quality. They can be used as early indicators of efficacy or toxicity, and so can avoid the need to wait for overt clinical signs or humane endpoints to develop, thereby reducing suffering. The use of biomarkers and PoC devices can also reduce animal use by allowing for longitudinal monitoring of single animals, therefore having the potential to reduce group sizes. Laboratory tests for biomarkers usually take over 20 minutes or would require processing so are not appropriate for implementing prompt humane endpoints – the longer the processing time the longer the animal could continue to suffer. PoC devices provide results in real-time or near real-time, with reduced blood volumes – meaning endpoints can be implemented sooner, reducing suffering, and microsampling can be used as a refinement to other blood sampling methods, due to the reduced volume required.

PoC devices can be sourced commercially from life science companies, device manufacturers and from collaborative resources. The lab has published a list of available PoC devices, within the context of snakebite coagulopathy here: [Handheld Point-of-Care Devices for Snakebite Coagulopathy: A Scoping Review - PubMed](#).

Potential barriers to using PoC devices are that they may not be fully validated for use in rodent species. Several are developed for human patient use, so often require trial and validation in your model species.

Sometimes there may also be the risk of false negatives/positives which need to be fully evaluated and standardised before use, and the cost varies significantly between devices. All of these issues are discussed in the manuscript above.

Romalie Alen, Labcorp – Managing weight loss in drug studies, and Mark Blee, NC3Rs Toxicology team – Upper body weight loss limits in short term toxicity studies.

1. In the studies looking at body weight loss as a marker of maximum tolerated dose, when looking at the body weight percentage, was that percentage loss of body weight from baseline of the animal being dosed or from age/model matched controls?

Mark: It was percentage compared to Day 0 – so baseline. "Body weight loss was calculated for each individual animal as the percentage in weight lost from the pre-treatment weight." Please see the paper for further details:

- Chapman *et al.* (2013). A global pharmaceutical company initiative: An evidence-based approach to define the upper limit of body weight loss in short term toxicity studies. *Regulatory Toxicology and Pharmacology*. 67:27-38. doi: [10.1016/j.yrtph.2013.04.003](https://doi.org/10.1016/j.yrtph.2013.04.003)

Romalie: The Home Office have provided instruction that an individual should be considered from the previous highest weight at our facility but there are instances where the Day 1 at initiation of dosing body weight may be more appropriate and can be scientifically justified. The weight must be considered in context, for example, a rapidly growing animal (young rat in growth phase) that has not gained weight will have a big impact on multiple systems within the body such as reproduction, even though no weight loss is apparent but a fully grown mature animal with slight weight loss may not be as affected.

2. On the point of lean mass loss as part of body weight loss as measured by body composition scans – there has been a lot of misreporting that loss of lean mass equates to loss of muscle mass. They are not equivalent as scans such as DEXA when measuring lean mass is including all tissue that is not bone or fat so included organ tissues as well as skeletal muscle. Do you have any thoughts on this?

Mark: Loss of muscle mass, lean or total, reduces calorie capability, so there is impact on other systems and, therefore, an overall impact on the animal. Muscle has its own process for using calories.

Romalie: Most adult animals have a genetically determined range of muscle mass that is heavily influenced by factors such as age, species, sex, nutrition and exercise. While the number of muscle fibres is generally fixed at birth, the size of those fibres, and thus the total muscle mass, remains dynamic throughout an animal's life. Loss of lean muscle reduces metabolism, strength and mobility, and is true for all animals including humans. Loss of lean muscle directly reduces overall muscle mass because lean body mass consists primarily of skeletal muscle. Therefore, loss of lean muscle will impact on the individual's whole biological system. When testing for toxicological markers it may be that incorrect information is obtained as

the animal is already influenced by changes. As such, a loss of lean muscle mass is a consideration on the appropriate use of animals and the appropriate endpoints for that animal.

The DEXA scan measures "lean mass," which includes muscle, water and organs. Because it measures all fat-free, bone-free tissue, it is a highly reliable proxy for skeletal muscle mass. Therefore it would be difficult to differentiate the loss of "lean" skeletal muscle so a DEXA scan may not be appropriate for weight loss vs muscle loss investigations.

3. In terms of applying best humane endpoints for each drug individually, how would you implement this in a PPL protocol where you may not know exactly which drugs you will be testing over the course of the PPL? This is a particular issue faced by Pharma companies and CROs.

Romalie: A Project licence (PPL) is written to a specific framework outlined legally through ASPA and the Home Office requirements and are usually specific to an area of investigation. Those licences that are for toxicology are generally broader to encompass various materials under investigation. Each drug should be considered on an individual basis and the project licences may be adjusted according to the pharmacological changes that have no impact to the animal. Ultimately the goal is to ensure that no animal is under pain suffering distress or lasting harm.

Exaggerated pharmacology that provides animals with excessive changes such as reduced blood pressure, stopping eating or cellular changes which affect organs negatively, although not toxicological need to have an endpoint to ensure the animals do not suffer. Therefore, although a generalised Project licence is beneficial to screen toxicity, a more specific approach could be considered to allow more flexibility for pharmacological changes. However, specificity requires a longer time frame and therefore can be restrictive for a pharma company needing to move forward quickly with a compounds development.

4. Do you think it would ever be possible to add in biomarkers as generic early tox indicators, or some more bespoke to what you would expect with certain substances/chemicals?

Mark: This is done, usually if a target is suspected, e.g. KIM-1 for kidney and CRP for heart (usually longer term) – applicable to clinic. Thoughts on modern urine biomarkers are included in the recent publication led by the NC3Rs:

- Passini *et al.* (2025). Applying the 3Rs to urinalysis assessments in toxicity studies: refining procedures and adopting a case-by-case approach. *Comp Clin Pathol* 34, 311–322. [doi: 10.1007/s00580-025-03642-5](https://doi.org/10.1007/s00580-025-03642-5)

Klas Abelson, University of Copenhagen – Minimising potential suffering in rodent models for pain research.

1. What should my first steps be if I want to check whether a potential refinement (e.g. using analgesia) will affect model validity? Does this always depend on the model or are there ways to approach this that are more broadly applicable?

The first step would always be to find out whether the refinement (e.g. analgesia) theoretically is likely to have any effect. Using analgesics in a study, where for instance immunoparameters, inflammatory markers or neurochemical mechanisms are relevant, there is a potential risk that experimental read-out could be affected. If there is such a risk, one must study the literature and discuss with colleagues if any previous studies have been made on the subject. This could be of great help for designing the study in a proper way, in terms of refinement effort. However, since one experimental setup at one establishment rarely is identical to another, it may be necessary to evaluate the potential effects at some points, by setting up studies (pilot or full) comparing effects of various analgesic or other refinements.

2. You mention the importance of well-defined humane endpoints, what does this mean to you? Can you elaborate on the approach that you take to reviewing and refining humane endpoints?

The threshold for reaching a humane endpoint should be very clear and based on a strict scoring system, where the score of parameters used for welfare monitoring is harmonised across parameters. It is easy to say that the experiment will be stopped if, let's say, the body weight drops by 15 or 20 percent. But humane endpoints must also be detectable by a cumulative score, adding the score of several parameters together. This makes the detection more sensitive, which will make it possible to detect potential suffering at an earlier stage.

3. Do you see these refinements being implemented in regulatory or industry settings?

I cannot give a good answer to that, since I don't have insight in enough establishments of that kind. But where I do, in Denmark for instance, I would say that there is a clear understanding and acknowledgement of the importance of refinement, including many efforts of refinement being implemented.

Claire Gibson, University of Nottingham – Improving *in vivo* modelling of ischaemic stroke and animal welfare.

1. How do you ensure consistency in applying the traffic light system across different strains and sexes? Do these influence welfare outcomes or model performance?

The traffic light system is a useful reference system and starting point – users may need to modify locally to take account of observed differences according to strain and sex.

2. How would you objectively measure if the mice are showing a preference on one side of their body and is there specific guidance for this?

There isn't an objective measure but the 'expected' outcome is that animals will display some neglect on the side of their body which is opposite to the side of the brain in which the stroke has occurred. So users should assume the animals will show a preference for the side of the body which is on the same side of the brain in which the stroke occurred and ensure housing and monitoring is appropriate (e.g. animals may neglect or reduce grooming on the neglected side of their body).

3. Has work been done to validate the grimace scale in stroke models?

No – this is a piece of work that still needs to be done.