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

P. Hawkins
*Royal Society for the Prevention of Cruelty to Animals*

D. B. Morton

O. Burman
*University of Lincoln*

N. Dennison
*Animals (Scientific Procedures) Inspectorate*

P. Honess
*University of Oxford*

See next page for additional authors

Follow this and additional works at: [https://animalstudiesrepository.org/resmlani](https://animalstudiesrepository.org/resmlani)

Part of the Animal Experimentation and Research Commons, Animal Studies Commons, and the Design of Experiments and Sample Surveys Commons

**Recommended Citation**


This Report is brought to you for free and open access by the Humane Society Institute for Science and Policy. It has been accepted for inclusion by an authorized administrator of the Animal Studies Repository. For more information, please contact eyahner@humanesociety.org.
Authors
P. Hawkins, D. B. Morton, O. Burman, N. Dennison, P. Honess, M. Jennings, S. Lane, V. Middleton, J. V. Roughan, S. Wells, and K. Westwood

This report is available at Animal Studies Repository: https://animalstudiesrepository.org/resmlani/1
A guide to defining and implementing protocols for the welfare assessment of laboratory animals: eleventh report of the BVAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement

P Hawkins (Editor), D B Morton (Chair), O Burman, N Dennison, P Honess, M Jennings, S Lane, V Middleton, J V Roughan, S Wells and K Westwood

The UK Joint Working Group on Refinement (JWGR) was established in 1989 by the British Veterinary Association Animal Welfare Foundation (BVAWF), the Fund for the Replacement of Animals in Medical Experiments (FRAME), the Royal Society for the Prevention of Cruelty to Animals (RSPCA) and the Universities Federation for Animal Welfare (UFAW). The aim is to provide up-to-date practical information on refinement. The JWGR prepares reports on specific topics, drawing together experts in each field to define contemporary best practice and ideals to aspire to. Professor David Morton chairs the Group and the secretariat is provided by the RSPCA. See http://www.rspca.org.uk/sciencegroup/researchanimals/implementing3rs/refinement (last checked 26 May 2010).

Corresponding author: P Hawkins. Email: phawkins@rspca.org.uk

Abstract

The refinement of husbandry and procedures to reduce animal suffering and improve welfare is an essential component of humane science. Successful refinement depends upon the ability to assess animal welfare effectively, and detect any signs of pain or distress as rapidly as possible, so that any suffering can be alleviated. This document provides practical guidance on setting up and operating effective protocols for the welfare assessment of animals used in research and testing. It sets out general principles for more objective observation of animals, recognizing and assessing indicators of pain or distress and tailoring these to individual projects. Systems for recording indicators, including score sheets, are reviewed and guidance is set out on determining practical monitoring regimes that are more likely to detect any signs of suffering. This guidance is intended for all staff required to assess or monitor animal welfare, including animal technologists and care staff, veterinarians and scientists. It will also be of use to members of ethics or animal care and use committees. A longer version of this document, with further background information and extra topics including training and information sharing, is available on the Laboratory Animals website.

Keywords: Refinement, welfare assessment, pain assessment, score sheets, humane endpoints

Laboratory Animals 2011; 45: 1–13. DOI: 10.1258/la.2010.010031

1 Introduction and aims

Reducing animal suffering through the refinement of husbandry and procedures is an important component of good science. It is also essential for humane reasons and is a specific requirement of legislation in some countries. If reducing animal suffering is to be effectively achieved, suffering must be detected as rapidly as possible so that appropriate action may be taken such as providing...
analgesia, applying a humane endpoint, reviewing husbandry and enrichment or euthanizing the animal.

Some signs of animal suffering are relatively easy to identify and assess, and many papers have been published on the assessment of welfare, both in general and following specific procedures. Despite this, there is still much reliance on subjective assessments and individual opinion.11

Discussion on these issues with veterinarians, animal care staff and scientists from a number of facilities in the UK established that it would be helpful to have further advice on objective methods for predicting and assessing welfare and animal suffering.11 The aim of this document, therefore, is to provide practical guidance on setting up and operating effective protocols for the welfare assessment of animals within individual projects. It should prove useful for all staff required to assess or monitor animal welfare, including animal technologists and care staff, veterinarians, scientists and members of ethics or animal care and use committees. Although it was produced in the UK, the issues and guidance apply worldwide.

There is also a longer version of this guide, which includes further background information and topics such as training and information sharing. The full version can also be used to assist in project design, as a discussion document for ethics or animal care and use committees, and to help funding bodies and regulators wishing to ensure that welfare will be properly assessed and suffering minimized in projects that they support or license [http://la.rsmjournals.com/cgi/content/full/la.2010.010031/DC2].

2 General principles for an effective welfare assessment scheme

The best approach to welfare assessment for each project depends on the type of establishment and its particular working practices, the nature of the research or testing, and the species and numbers of animals involved. However, there are some fundamental principles that should underpin all welfare assessment schemes. These are set out in Table 1 and explained more fully in the rest of the document.

2.1 A team approach

A team approach to welfare assessment is highly effective, because it allows input from people with different expertise, priorities and responsibilities. This should enable animal welfare to be given due priority, while also taking into account scientific requirements and the resources available for animal monitoring. Who is involved in welfare assessment, and how the process operates for each project, will depend on the nature of the individual project and establishment and the experience, expertise and resources available in-house. The end result could be a fixed group or a more fluid association of people, or the welfare assessment process may form part of the remit of an existing committee. It is most effective to take a flexible, tailored approach and consider the skills, knowledge, experience, motivation and authority that are required before identifying the team members who can bring them. Whatever the structure of the team, the competencies that are invaluable in establishing it are listed in Table 2.

2.2 Definition of good welfare

A baseline standard of good welfare should first be defined, to act as the point of reference for the species (and strain, where relevant) to be used in the study. This standard may apply establishment-wide for particular species or
strains. However, setting such a baseline is not always straightforward, particularly where behaviour is used as an indicator of welfare. For example, genetically altered (GA) mice with vestibular abnormalities spend much time circling in their cage. This is the normal behaviour for these animals, but is not necessarily desirable from a welfare point of view.\textsuperscript{1}

Similarly, understimulating housing can cause stereotypic behaviour,\textsuperscript{12} which is normal in such environments, but can indicate a serious welfare problem. The term ‘natural’ is sometimes used instead of normal, but this is no more descriptive – what is normal behaviour for a laboratory animal is not necessarily natural.\textsuperscript{5}

A more useful reference point for the welfare assessment protocol is to define a hypothetical ‘ideal’ level of welfare. This can be defined as: the state of being in animals when their nutritional, environmental, health, behavioural and mental needs are met.\textsuperscript{13} There are three key components to this ideal, set out in Table 3.

There is usually no need to measure all of these parameters to set the baseline. The ideal state can be assumed to exist if the team is confident that animal housing and care is consistent with best practice, the animals are healthy, behaving according to an appropriate time budget and they are fully habituated to their accommodation and husbandry routines.

\textsuperscript{1}In GA animals, use of the term ‘normal’ should be avoided in the context of welfare assessment because it refers to the wild type strain that has been used to generate the mutant line. The term ‘as described for wild type’ is thus more descriptive than ‘normal’ and ‘deviating from expected’ is preferable to ‘abnormal’.

There are many potential causes of suffering during the animals’ lives that may impact on current or future welfare assessments. Causes of deviation from the ideal welfare state may include early separation from the mother, transport, trapping, inappropriate housing, inadequate health care, scientific procedures and their after effects (expected and unexpected), husbandry procedures (such as cleaning out and identification) and euthanasia or release.\textsuperscript{2,14–16} These events can interact with one another. For example, stress due to early separation from the dam can influence nociception in rats.\textsuperscript{17}

Taking all this into account and assuming that the baseline standard is good, any deviation from this ideal state could indicate a welfare problem and should be investigated as such. Note, however, that some physiological parameters can alter in association with positive excitement, such as play, as well as with negative stimuli. Furthermore, many commonly used species, particularly rodents, do not always display behavioural signs of suffering that can easily be detected by human observers. These issues can be overcome by thoughtful selection and interpretation of welfare indicators.

### 2.3 Selection of appropriate welfare indicators

Key to the success of the welfare assessment scheme is the selection of welfare indicators that:

- Are readily and reliably recognizable;
- Are effective at providing good measures of welfare;

---

**Table 2** Competencies for a successful welfare assessment team

<table>
<thead>
<tr>
<th>The team should include members who:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can recognize a ‘normal’ animal and understand ‘normal’ behaviour for the species and strain, in the laboratory environment</td>
</tr>
<tr>
<td>Understand the science relating to the project</td>
</tr>
<tr>
<td>Can recognize all the potential welfare issues (both within the project and throughout the animals’ lives)</td>
</tr>
<tr>
<td>Can recognize and identify animals showing adverse effects as a result of scientific procedures</td>
</tr>
<tr>
<td>Can identify an animal who is not ‘normal’ but where this deviation from normality is not an adverse effect resulting directly from scientific procedures</td>
</tr>
<tr>
<td>Can assess and interpret welfare indicators</td>
</tr>
<tr>
<td>Are able to give advice on ameliorating (or avoiding) adverse effects</td>
</tr>
<tr>
<td>Bring in comparative knowledge across different species and institutes</td>
</tr>
<tr>
<td>Are able to address management and resource issues, such as staffing levels, either directly or by communicating with management</td>
</tr>
<tr>
<td>Will take ultimate responsibility for acquiring up-to-date information on welfare assessment</td>
</tr>
</tbody>
</table>

**Table 3** Components of an ‘ideal’ welfare state and examples of indicators associated with them

<table>
<thead>
<tr>
<th>Component</th>
<th>Characteristics</th>
<th>Examples of indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Physical state</td>
<td>Good level of physiological fitness, with no physical disabilities that either cause discomfort or pain, or that have an impact on physical function that could cause distress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Levels of stress and distress do not exceed those that would occur during the course of normal social interactions, for example. If parameters such as heart rate or blood pressure were measured; they would not be expected to indicate significant stress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The animal displays an ‘appropriate’ range of behaviours, according to what is known about the species and strain</td>
<td></td>
</tr>
<tr>
<td>(ii) Physiological/ biochemical state</td>
<td>Indicators relating to the observable physical condition of the animal, e.g. body weight, state of the coat, posture, lameness and excessive attention to surgical sites</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Physiological parameters such as heart rate, respiratory rate, levels of stress hormones such as corticosteroids</td>
<td></td>
</tr>
<tr>
<td>(iii) Psychological state</td>
<td>Changes in behaviour such as increased aggression to cage mates, withdrawal, stereotypies and changes in use of enrichment</td>
<td></td>
</tr>
</tbody>
</table>
• Are relevant to the project and species;
• Are practical to carry out and do not overly disturb the animal;
• Take the experimental design into account; and
• Lend themselves to consistent measurement, interpretation and analysis.4,18

It is preferable to use a combination of indicators from each of the categories listed in Table 3 to overcome difficulties with interpretation and to provide a more detailed and complete picture of an animal’s welfare.11,19

2.3.1 General indicators
A list of simple, objective welfare indicators, such as body weight and condition (physical state in Table 3), measured body temperature (physiological state) and food and water consumption (this may fall into any of the three categories), can be drawn up for use in most projects. These indicators can be directly and objectively measured, providing clear indicators that an animal’s welfare may be compromised. They are also useful for defining and implementing humane endpoints.11

2.3.2 Indicators specific to the project
The next stage is to predict the likely adverse effects, so that a list of indicators can be produced that is tailored to the study. For example, following vasectomy surgery, mice would be expected to experience a degree of pain associated with the wound site as an adverse effect. Indicators of this adverse effect would include body weight loss (a general indicator) and behavioural indicators specific to the project such as lifting a hind leg or pressing the abdomen to the cage floor.20

These behaviours can be used as indicators for welfare assessment post-vasectomy. Table 4 lists sources of information and guidance that can be used to predict adverse effects.

The list of potential adverse effects could include items such as discomfort or pain in specific areas of the body, nausea or other toxicological effects, anxiety, reduced physical ability and so on. The next step in the process is to consider what the behavioural or physiological indicators of these adverse effects might be. Some commonly used indicators that are associated with various adverse effects are arching the back, twitching and writhing and/or drawing in of the flank in rodents, slack muscle tone around the eyes in birds, skin colour changes in fish or reluctance to move in many species. The list of indicators for all species is continually expanding; a recent addition is the use of facial expressions in the mouse.22 Interactions with the environment, such as gnawing on chew blocks or nest building, can also provide extremely useful indicators, which is another reason for providing a structured environment.19

The Appendix [http://la.rsmjournals.com/cgi/content/full/la.2010.010031/DC1] lists sources of further examples of welfare indicators. For a more in-depth approach to estimating the probability of pain and distress, its consequences and associated risk levels, see National Health and Medical Research Council (2008)4 and for training in all of the above see Assessing the Health and Welfare of Laboratory Animals, http://ahwla.org.uk/ (last checked 26 May 2010). Table 5 sets out an approach for identifying indicators and

| Table 4 Sources of information and guidance on predicting potential adverse effects |
|-----------------------------|---------------------------------------------------------------|
| Information from care staff (and breeders, if possible and applicable) on sourcing, transport, identification methods, husbandry regimes, etc. and their welfare implications |
| Harm–benefit analyses that have already been carried out |
| Past experience, if other, similar studies have been carried out using the same species or strain |
| Results of in vitro and in silico studies, e.g. Quantitative Structure–Activity Relationship (QSAR) where chemical structure is correlated with biological activity |
| Searching the literature for publications of similar studies. Papers including adverse events can be extremely valuable |
| In a contract research setting, information obtained during the testing of other compounds (e.g. for agrochemicals, signs may be comparable between similar classes of compounds), chemical structure or information supplied with test compounds |
| Information from pilot studies, for example, data on absorption |
| In pharmaceutical research and development and safety assessment, information on the ‘target’, or predicted pharmacological action, not only of the parent compound, but also of any metabolites that may be formed* |
| For some models of animal diseases, e.g. using the same species as the target species, clinical signs of the disease in clinical cases |
| For models of human disease, clinical signs and symptoms of the disease in humans can provide pointers. This should be done with care, as clinical signs may differ in animals; research projects are usually only modelling certain aspects of a condition* |
| Contacting other researchers using the same or similar models, possibly using online discussion groups such as Compendia |
| Published resources, including those listed in the Appendix |

*Commercial computer databases (e.g. DEREK) are excellent resources for accessing known information on chemical structure and activity of parent compounds and their metabolites. They can also provide very good predictive activities for naïve compounds, although these predictive systems are not perfect

considering how each indicator might best be monitored and assessed.

Welfare assessment is an ongoing process and it is likely that the list of indicators will need to be reviewed and
updated as the work progresses, according to the validity of each indicator and the information it can provide. The aim is to include the minimum number of parameters necessary to detect adverse effects rapidly and effectively, yet not waste time gathering data that have no added value. It may be possible to reduce the number of indicators by using so-called ‘iceberg indicators’ that summarize other areas of welfare and are easy to understand. One such indicator is a lack of grooming after surgery, which could indicate an inability to coordinate grooming movements, postoperative pain or, if shortly after surgery, the side-effects of anaesthetic or analgesic agents.

Adverse effects and their behavioural indicators can sometimes be completely unpredictable, especially when testing novel compounds or in mutagenesis projects. In such cases, there will be some knowledge gaps and welfare assessment is approached blind, and thus there is a strong case for including as many potential indicators as is workable in the initial assessment. It may also be possible to acquire more information by consulting with outside experts and/or undertaking a small pilot study to help define appropriate indicators (see section 2.3.4). This is especially likely to be necessary when using new species or novel techniques.

2.3.3 Intervention points

It is essential for animal welfare, ethical and often legal reasons that clearly defined intervention points are set for each project. Suitable interventions when key signs appear, or reach a threshold level, should be defined at the time when indicators are discussed. For example, fluid therapy could be initiated as soon as signs of dehydration appear, or an intervention for weight loss of 10% in rats with a degenerative condition could be providing wet mash at floor level. Humane endpoints should also be defined at this stage, in which the animal is temporarily or permanently removed from the study.

Thresholds should be established that allow minimal animal suffering while still retaining the objective of the study, and so any intervention and its timing will need to be agreed with the researcher and regulator at an early stage. Some clinical signs may be expected as part of the model. One such case is pale extremities in animals used to study cardiovascular disease. The welfare assessment system should include clear guidance on those effects that are to be expected as part of the model, and what the endpoint is for each one.

2.3.4 Pilot studies

Pilot studies using a small number of animals can provide useful guidance on welfare indicators where these are difficult to predict, for example, for the testing of novel compounds or for newly developed experimental designs. The results of pilot studies can not only provide the indicators for the final project, but also help to guide refinements, including intervention points and humane endpoints.

The first animals in a pilot study should be monitored extremely carefully, using frequent sampling and a broad range of indicators, so as to gain as much information as possible about potential adverse effects and their progression. In the case of pilot studies where effects are highly unpredictable or potentially severe, it is advisable for an experienced animal technologist and/or the attending veterinarian to be present to assist with monitoring.

2.3.5 Indicators of positive welfare

Welfare assessment generally focuses on negative rather than positive welfare. However, it is also desirable to improve the animals’ quality of life as well as to minimize suffering, so the potential to define signs of positive welfare and the addition of these to the welfare assessment system should be considered. A general list of examples of behaviours to be considered is set out in Table 6.

Positive welfare signs will vary considerably with species, strain, life experience and individual temperament of animals. Indicators such as these therefore need careful interpretation, using the animal behaviour literature and advice from ethologists. For example, some strains of rodent are passive or have low activity levels, but this does not necessarily relate to individuals’ wellbeing. Some behaviours, such as tail chasing in rats, can mistakenly be believed to be play when the behaviour is in fact a self-directed activity in response to social isolation.

Knowledge about reliable signs of positive welfare is limited for many species at the time of writing. However, this is an evolving branch of animal welfare science and it is important to keep developments under review. For further information, see Boissy et al., Kirkwood et al., Wemelsfelder, and Yeates and Main.

2.4 Animal welfare indicator record systems

There are different systems in common use for recording welfare indicators. We have broadly categorized them as (i) relatively unstructured records, with a small number of objective signs and a reliance on written descriptions of adverse effects (free text), or (ii) more organized animal welfare assessment sheets with predetermined, but flexible, lists of indicators and minimal free text. The latter may be either numerical or binary score sheets. Numerical sheets aim to quantify the severity of adverse effects; binary systems simply note whether or not the adverse effects are present.

Relatively simple records with free text may be most appropriate in certain circumstances, for example during

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Examples of behaviours that can indicate positive welfare states</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good self-care, including grooming and ‘comfort’ behaviours</td>
<td>Normal activity levels and time budget, including sleep patterns</td>
</tr>
<tr>
<td>Normal activity levels and time budget, including sleep patterns</td>
<td>Seeking interactions with humans</td>
</tr>
<tr>
<td>Seeking interactions with humans</td>
<td>Curiosity and interest in exploring</td>
</tr>
<tr>
<td>Curiosity and interest in exploring</td>
<td>Appropriate social interactions with conspecifics, including allogrooming</td>
</tr>
<tr>
<td>Appropriate social interactions with conspecifics, including allogrooming</td>
<td>Mating ‘Anticipatory’ behaviour</td>
</tr>
<tr>
<td>Mating ‘Anticipatory’ behaviour</td>
<td>Using enrichment items, especially for ‘luxury’ behaviours</td>
</tr>
<tr>
<td>Using enrichment items, especially for ‘luxury’ behaviours</td>
<td>Interest in food treats</td>
</tr>
<tr>
<td>Interest in food treats</td>
<td>Play</td>
</tr>
<tr>
<td>Play</td>
<td>Vocalizations associated with positive welfare</td>
</tr>
<tr>
<td>Vocalizations associated with positive welfare</td>
<td>Normal learning and cognitive functions</td>
</tr>
</tbody>
</table>
of the assessor should protect the animal from avoidable action, although empathy and common sense on the part of the assessor may need to be weighed. It is also conceivable that an animal may be experiencing severe suffering and yet have a score that does not require action is needed (such as analgesia) according to a predetermined key attached to the sheet (Figure 1). Numerical scoring can also form the foundation of more complex assessments of welfare, for example in assessing cumulative suffering.33

However, numerical scores need careful interpretation. Simple addition may be justified in some cases but not others, and some parameters may need to be weighted. In the example shown in Figure 1, a score of 2 for staring coat and 2 for isolation from cage mates does not ‘equal’ a 4 for a hot, distended gut; an animal with a score of 4 is not suffering twice as much as an animal with a score of 2. It is also conceivable that an animal may be experiencing severe suffering and yet have a score that does not require action, although empathy and common sense on the part of the assessor should protect the animal from avoidable suffering in such cases.

In contrast, the binary system records either a ‘yes’ (present) or ‘no’ (absent) depending on whether the behaviour or effect is seen or not, with no description of its intensity. Core parameters such as body weight are also usually measured and recorded. The binary system is generally regarded as being more objective than the numerical system, as value judgements on severity are not required. However, objectivity should not be taken for granted; checks should still be made that people are using the system consistently.

The numerical and binary systems each have their own strengths and weaknesses and will be appropriate in different contexts (Table 8).

Whichever recording system is chosen, it should be adapted for specific studies and then regularly revised and modified with use if necessary. It may be that expected indicators occur infrequently, in which case they may be deleted, or if they are vital to the assessment, sampling frequency should be altered to ensure that they are picked up. Conversely, unexpected adverse effects may occur (as noted in the free text boxes) and indicators for these should be added to the sheets.

As a final note of caution, it should not be assumed that any welfare assessment system is infallible. There is always the potential for unpredicted adverse effects to occur, or for a particularly empathetic assessor to detect very subtle and previously unrecognized clinical signs. Balance is essential between striving for objectivity in welfare assessment on the one hand, and trusting the judgement of empathetic individuals on the other.

### 2.5 Assessment timing, duration and frequency

Once the recording system has been designed, a monitoring protocol can be set out that includes how often to assess animals, at what point in their activity phase and how long to spend on the assessment. The specific details of implementation will depend on the factors set out below. Clearly, the more animals that need to be assessed, the more time is needed. It is essential that adequate resources are in place to allow effective monitoring.

#### 2.5.1 Species and strain

Ideally, all species should be regarded as equally capable of suffering and should be given equal consideration. However, there are some genuine, practical issues that influence the level of complexity of the assessment and the time that it takes to conduct (and analyse) it. There are more measures of welfare available for some species than others, which can impact on the time required to carry out welfare assessment. Some species display behaviours that are comparatively easy for humans to detect and interpret. A depressed primate displays a characteristic hunched posture which is easy for most observers to recognize and empathize with.34 The signs of poor welfare may be more difficult and time consuming to recognize in other species such as small rodents and non-mammals; there are also likely to be differences between strains in both normal behaviour and responses to pain, suffering or distress.
Awareness of all of the above issues relating to genuine and perceived difficulties with assessing different species and strains, and consultation with ethologists and other relevant experts to resolve these, can help to facilitate consistent consideration for all animals.\textsuperscript{35}

With respect to GA animals, many gene manipulations have had both predictable and unpredictable characteristics that affect welfare, but which may not always be detected using standard phenotyping protocols.\textsuperscript{36} The welfare of founder animals should be carefully assessed over an

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|}
\hline
Component (see Table 3) & Score & Date/time & Date/time \\
\hline
Body weight & Normal or up to 5\% loss & 0 & \\
& 5--10\% loss & 1 & \\
& Over 10\% loss & 2 & \\
& General lack of grooming & 1 & \\
& Staring coat and/or ocular or nasal discharge & 2 & \\
& ‘Pinched’ features, ridge lines in skin & 4 & \\
& Faeces normal to slightly soft & 0 & \\
& Diarrhoea & 1 & \\
& Soft, distended gut, no faeces & 2 & \\
& Hard, hot, distended gut & 4 & \\
& Normal breathing & 0 & \\
& Slightly laboured breathing & 1 & \\
& Notably laboured breathing & 3 & \\
& Normal provoked behaviour & 0 & \\
& Slightly subdued & 1 & \\
& Moderate change in behaviour and/or apart from cage mates & 2 & \\
& Reacts violently/vocalization & 6 & \\
\hline
Other observations & & & \\
\hline
\textbf{TOTAL} & & & \\
\hline
\end{tabular}
\caption{Example taken from a numerical score sheet for rats used in inflammatory bowel disease studies (NB this is not a complete sheet)}
\end{table}
extended period, beginning at neonatal stages, so that any adverse effects can be identified and appropriate indicators used in the routine welfare assessment of the line. It is also good practice to extend this comprehensive inspection to genetic alterations being bred onto different background strains for the first time or being bred to homozygosity (this is especially relevant in mice).

2.5.2 The experimental design
The nature of the procedures will also influence the timing, duration and frequency of observations. Procedures at higher levels of severity require more frequent monitoring, as will protocols where a rapid onset of adverse effects is expected. The timescale of adverse effects can be more predictable if they have been scientifically evaluated (for example pertussis vaccine potency testing37) or there is a body of experience relating to a particular technique.

The timing of procedures should also be taken into account. There will obviously be a risk of discomfort or pain in the immediate postoperative period. For example, rats immediately following various types of abdominal surgery such as laparotomy, adrenalectomy and bladder manipulations have been found to display specific behaviours including twitching (usually observed as rapid fur movements on the back), back arching, belly pressing and writhing.9 Such behaviours have been found to occur at varying frequencies depending on the level of discomfort and it is necessary to observe each rat for at least 5 min to ensure that they are detected. Left untreated, these behaviours can be present for up to 24 h following the procedure.

Animals would need to be assessed more frequently post-surgery to assess whether analgesics were effective. Observations could then be made less frequently provided that there were no complications.

2.5.3 Housing environment
The time allowed for welfare assessment should take account of environmental stimulation such as nesting material, refuges and other structures provided. It may be necessary to move enrichment items or open nests to observe animals properly.

2.5.4 Husbandry practices
Disturbance caused by husbandry procedures such as cage cleaning can have a significant effect on animal behaviour and physiology. Following cage cleaning in the rat, exploratory behaviours, shelter use, heart rate and blood pressure all increase significantly.38–41 Behavioural and physiological parameters can take up to 2 h to return to pre-cage change levels in rodents,41 during which time these responses may mask important indicators and confound the welfare assessment. It may be advisable to conduct welfare assessment an hour or two after husbandry procedures such as cage change, as long as the potential severity of the procedure does not require more frequent monitoring.

2.5.5 The animals’ normal circadian rhythm
It is preferable to observe animals during the time when they would usually be most active, unless there is a potential for sleep disturbance as an adverse effect. Assessing awake, active animals will reduce the likelihood of missing essential signs. In the case of most rodents, this means conducting welfare assessment during the dark period, when they are predominantly active.9 There are obvious human resource issues associated with observing animals at night, but animals can be housed on an altered or a reversed light regime. Disruptions to the circadian cycle can also be used as a welfare indicator. For example, sleep disruption, reduced activity or responsiveness during periods of normally high activity, or increased activity when animals should be inactive, may indicate adverse effects.25,42 Care staff may notice changes in circadian behaviour patterns, or it may be appropriate to use sophisticated monitoring techniques such as video monitoring (which can be sped up or sampled for analysis), digital imaging or automated behaviour recognition systems such as Observer® (Noldus Information Technology, Wageningen, The Netherlands), Trafficage® (NewBehavior AG, Zurich, Switzerland) or HomeCageScan® (Clever Systems Inc, Reston, VA, USA).

3 Practical welfare assessment
This section provides generic guidance on welfare assessment in practice, which can be applied not only to domesticated mammals, but also to wild animals, fish, reptiles, birds and amphibia. Before implementing the welfare assessment protocol, it may be useful to go through the checklist in Table 9.

### Table 8 Comparison of numerical and binary observation systems

<table>
<thead>
<tr>
<th>Numerical system</th>
<th>Binary system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td><strong>Advantages</strong></td>
</tr>
<tr>
<td>Consistent, provided that guidance is clear and scoring options limited</td>
<td>Potentially more objective assessments – simply ‘present’ or ‘absent’</td>
</tr>
<tr>
<td>Considerable amount of data can be collected; data can be statistically analysed</td>
<td>Less time consuming as judgement on numerical scores is not required</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td><strong>Disadvantages</strong></td>
</tr>
<tr>
<td>Element of subjectivity in assigning scores</td>
<td>Less data collected</td>
</tr>
<tr>
<td>Can be time consuming</td>
<td>Need to consider threshold for marking as present, which still takes some time and can be subjective</td>
</tr>
<tr>
<td>Potentially not sensitive enough to detect any subtle changes</td>
<td>Cannot assign clear intervention points on the basis of a numerical score, although presence of some clinical signs can denote endpoints</td>
</tr>
<tr>
<td>Accumulated scores can over- or underestimate severity. Some changes may need to be weighted as more indicative than others, and the significance of certain criteria coinciding should be considered</td>
<td>No description of the intensity of each effect, unless this is built in (e.g. ‘pronounced piloerection’)</td>
</tr>
</tbody>
</table>
3.1 Making observations

Having the same person, or a very small number of people, observing the animals wherever possible can facilitate consistency and enable assessors to follow the progression of an animal’s condition more accurately.43 Animals may also consistently and enable assessors to follow the progression of an animal’s condition more accurately.43 Animals may also be able to tell the difference between different people (for example, on the basis of individual odours), so they may benefit from contact with familiar staff as opposed to strangers.44 Consistency of staff also enhances job satisfaction for animal technologists and carers, many of whom prefer to be responsible for the same animals throughout a study.

3.1.1 Observation from a distance

Assessing the general appearance, posture and behaviour without provoking any responses provides useful information, so animals should first be observed from a distance without moving, approaching or entering the enclosure or opening the cage. This enables the observer to see whether there are unprovoked behaviours that could indicate welfare problems, such as social animals isolated from conspecifics or nocturnal animals immobile but out of the nest. Alternatively, individuals may be playing, foraging or allogrooming, indicating that welfare is probably good. Animals should be observed for the predetermined time period to ensure that relevant indicators are more likely to be detected.

The enclosure should also be observed to see whether activities such as nest building, foraging or gnawing are reduced, or whether there is evidence of health problems including bleeding, vomiting or abnormal faeces (Figure 2). Clinical or behavioural indicators that can be observed without touching the animals or influencing their behaviour should be noted on to the assessment sheet at this stage. These include piloerection, postural changes, altered opercular beat frequency in fish, escape behaviour in amphibians, reduced mobility or favouring a surgical site in any species. These signs may be either specific to the project or unexpected, in which case they should be entered into the free text box if a structured sheet is used.

3.1.2 Opening the cage or entering the enclosure

The next stage is to examine the enclosure in more detail by removing the cage lid or entering the enclosure (or closely approaching it as appropriate). Enrichment items or nesting material should be removed or moved aside if necessary, and the animals’ reactions to this – and the observer – should be watched for a suitable period of time, as previously determined. Most species would normally respond with increased activity followed by a settling down period. Any specific or unexpected signs should be recorded, as above.

3.1.3 Handling the animals

After completing the above initial checks, terrestrial animals should be individually caught and handled to measure and score relevant core criteria such as body weight, body condition and temperature. This is also the time to assess some specific criteria that require handling, such as skin tenting, sensitive areas, tumour measurement or parameters that require blood sampling. Handling is stressful for many animals, which can affect both animal welfare and the scientific data.16 Observing animals following handling (for example, to record body weight) can facilitate observing some relevant behaviours, such as the postsurgical behaviours outlined in section 2.5.2, and reduce the time needed for the assessment.

It is not always appropriate to handle animals. Handling may cause discomfort or pain following certain procedures, or removing animals from their housing would cause excessive distress under some circumstances, as in aquatic animals or some breeding females. In such cases, extra time should be allowed for careful visual observations.

3.2 Highlighting potential welfare issues

Carrying out observations and noting them onto the assessment sheets, as set out above, may signify that there is a welfare problem. This may be due to the presence of a key indicator, because the animal’s score has reached a threshold level, or simply because someone feels that something may be wrong. Actions to be taken if animals are (or may be) suffering should have already been agreed and understood by all. There should be a clear line of reporting and everyone should know his or her responsibilities within it.

There should also be a failsafe system for flagging up enclosures containing animals that give cause for concern for any reason. All staff should know that animals housed in enclosures highlighted in this way require special attention and additional monitoring.17 Methods commonly used to draw attention to animals with welfare problems include message boards outside the room and coloured pegs on enclosures.

4 Reviewing welfare records

Timely reviews of welfare assessment records, during and after projects, are essential to ensure that welfare assessment systems are operating effectively; take account of any changes in the adverse effects noted; and ensure that any
changes in the nature of the project, knowledge about animal behaviour or new assessment techniques are taken into account.4

4.1 Reviewing adverse effects during projects

The key aim of welfare assessment review is to examine how well any adverse effects are being predicted, recognized and alleviated (Table 10). This could be done at one or more set points during a project, and/or in response to specific concerns about animal welfare or the effectiveness of the welfare assessment system that may arise. An advantage of setting appropriate interim review points during the life of a project is that initial, far-reaching systems of observations (for example, in the case of a pilot study or a founder GA animal) can be refined to include only the most relevant parameters or time points. This makes for more efficient use of resources as well as more effective welfare assessment.

Interim review should also help to detect any drift in the welfare consequences and the related clinical or behavioural indicators seen over time. This may be due to factors such as changes in the genotype or strain of animal used, changes in personnel performing procedures or in the duration of the study, including the aging of the animals in long-term studies. It may be appropriate to report the outcome of interim reviews to the ethics or animal care and use committee.

There should always be a retrospective review of welfare assessment records once a project has been completed, which should contribute to ethical and scientific reviews of the completed project.46 This should include the elements of interim review listed in Table 10.

Table 10 Points to include in interim reviews

<table>
<thead>
<tr>
<th>Assessment of the welfare indicators; that is, the frequency with which they have been observed, with the aim of removing any that are redundant or adjusting the observation protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review of observations recorded in free text boxes, to see whether any new indicators should be added</td>
</tr>
<tr>
<td>Review of interventions, including humane endpoints, and whether any animals were found dead; correlation of these with data from the sheets</td>
</tr>
<tr>
<td>Review of the timing of observations and interventions, to ensure that the frequency of assessment is appropriate throughout the project</td>
</tr>
<tr>
<td>Checks on consistency between observers, using external expertise where necessary. Statistical analysis can be used to confirm consistency in some circumstances43,45</td>
</tr>
<tr>
<td>Comparison between the predictions of severity made at the project planning stage and the level of severity observed in practice</td>
</tr>
</tbody>
</table>

Figure 2 Cage appearance in male HsdHan:NMRI mice with and without postlaparotomy pain. Mice in the upper row have built well-structured nests and are defaecating in a separate area (circles), as expected for the strain. Mice in the lower row are experiencing mild to moderate postlaparotomy pain; the cage area is unstructured without a separate area for defaecation and there are two nest-like resting places (arrows). Reproduced with permission from Arras et al.19
5 Liaison with ethics or animal care and use committees

It may be useful for the welfare assessment team to liaise with relevant ethics or animal care and use committees at the planning stage, to keep the committees informed about the welfare assessment protocol and to seek additional guidance on its effectiveness and acceptability with respect to the local culture of care at the establishment. Depending on the nature and role of the committee, topics for discussion may include:

- What will happen to each of the animals throughout the project, from sourcing to euthanasia, reuse, release or rehoming;
- What each animal will experience and where adverse effects on welfare are possible – including, but not only as a result of, experimental procedures;
- Which parameters will be monitored during the welfare assessment and how they were decided;
- How frequently animals will be assessed, when and why;
- How observations will be recorded and analysed;
- Explanation of the humane endpoints, how these were set and what will happen if they are exceeded.

This will facilitate discussion on the welfare assessment protocol and may enable it to be further refined before it is implemented for the first time.

6 Accessing further information

Welfare assessment is a rapidly developing field and it is vital to ensure that information on new scientific, technical and practical developments is brought to the welfare assessment team so that it can be incorporated into welfare assessment protocols. As well as the literature on welfare assessment, useful information is often gained from other fields. Journals that publish relevant papers together with examples of keywords relating to welfare assessment, useful information is often gained from other fields. Journals that publish relevant papers together with examples of keywords relating to welfare assessment are listed in the Appendix [http://la.rsmjournals.com/cgi/content/full/la.2010.010031/DC1]. Someone on the welfare assessment team could be made responsible for checking these journals regularly for relevant papers. Communicating with other establishments and research teams about experiences with different systems can also help to disseminate good practice and facilitate the exchange of ideas.

7 Summary recommendations

7.1 Setting up the welfare assessment system

Good practice in welfare assessment is recognized as being necessary for good science as well as animal welfare. A team approach to welfare assessment is most effective, where the team membership is considered in terms of fulfilling appropriate competencies. Sound team cohesion and constructive working relationships are essential, both internally and between the welfare assessment team and other ethics and animal care and use committees. Education and training should be provided for team members as necessary, to ensure their competence in welfare recognition and assessment.

The baseline standard of good welfare should be carefully defined for each project, using a general list of simple, objective indicators as the basis of the assessment system. A list of potential adverse effects should be set out and tailored to each specific project, using a wide range of information sources. The next stage is to predict those clinical and behavioural signs that may be associated with each adverse effect, taking into account how easy it will be to recognize and assess each one. Intervention points and humane endpoints are also defined at the project planning stage, as part of the process of setting up the assessment system. It may be necessary to conduct pilot studies to define welfare indicators. Current understanding of indicators of positive welfare should be researched and added to the protocol if appropriate.

A flexible approach should be taken to the choice of recording system. Organized animal welfare assessment sheets can reduce subjectivity, but care should be taken when interpreting numerical score sheets, as weighting and addition are not always straightforward. The risk of missing essential signs can be minimized by considering very carefully when animals will be assessed, for how long and how often. It is important to strike a balance between relying on objective assessment schemes and using experienced human judgement – both are necessary and complementary. The proposed welfare assessment protocol should be discussed with all relevant committees; to explain how welfare will be assessed and suffering reduced, and to obtain their input and advice if this is part of their remit.

7.2 Using the welfare assessment system

The welfare assessment team should be properly briefed so that everyone knows what they are supposed to be doing. Recording systems, such as assessment sheets, should have been updated if necessary. The animals and their enclosure are usually observed from a distance before opening the cage, or entering the enclosure and handling animals. Clear systems are necessary for highlighting welfare concerns if suffering is believed to be significant or reaches predefined limits. This includes ensuring that everyone is aware of requirements for interventions and humane endpoints and knows how to act on them. Excellent communication and teamwork should be maintained between all persons involved with the study and the welfare assessment team. This will ensure that the entire team works smoothly together to resolve issues rapidly if interventions are required. Animals should always be given the benefit of the doubt with respect to whether or not they are experiencing suffering.

7.3 Reviewing the welfare assessment system

Both planned and ad hoc interim reviews need to be conducted as appropriate. Part of this process involves relating observations made in practice to the predictions of the level and nature of adverse effects made during project planning.
RETROSPECTIVE REVIEWS OF WELFARE ASSESSMENT RECORDS CAN BE PERFORMED WHEN PROJECTS ARE COMPLETED, IN CONJUNCTION WITH ETHICAL AND SCIENTIFIC REVIEWS AND REPORTING REQUIREMENTS AS APPROPRIATE. RESULTS SHOULD BE COMMUNICATED TO ALL THOSE WHO WOULD BENEFIT. NEW KNOWLEDGE ABOUT BEHAVIOUR AND WELFARE ASSESSMENT, FOR EXAMPLE, GAINED FROM THE SCIENTIFIC LITERATURE OR FROM ATTENDING MEETINGS, SHOULD BE INCORPORATED INTO WELFARE ASSESSMENT PROTOCOLS.

ACKNOWLEDGEMENTS

THANKS TO MARY MOWAT AND BARRY PHILLIPS FOR THEIR INPUT AND USEFUL COMMENTS ON THE STRUCTURE OF THIS RESOURCE.

REFERENCES

7 Morton DB. Adverse effects in animals and their relevance to refining scientific procedures. ATLA 1990;18:29–39
8 Firth AM, Haldane SL. Development of a scale to evaluate postoperative pain in dogs. JAVMA 1999;214:651–9
17 Dickson AL, Leach MC, Flecknell PA. Influence of early neonatal experience on nociceptive responses and analgesic effects in rats. Lab Anim 2009;43:11–16
18 Leach MC, Thornton PD, Main DCJ. Identification of appropriate measures for the assessment of laboratory mouse welfare. Anim Welf 2008;17:161–70
25 Abou-Ismail UA, Burman OHP, Nicol CJ, Mendl M. Can sleep behaviour be used as an indicator of stress in group-housed rats (Rattus norvegicus)? Anim Welf 2007;16:185–8
26 Van der Harst JE, Spruit BM. Tools to measure and improve animal welfare: reward-related behaviour. Anim Welf 2007;16(Suppl. 1):87–73
32 Yeates JW, Main DCJ. Assessment of positive welfare: a review. Vet J 2008;179:293–300
40 Burn CC, Peters A, Mason GJ. Acute effects of cage cleaning at different frequencies on laboratory rat behaviour and welfare. Anim Welf 2010;15:161–71
Appendix

The Appendix sets out useful journals, web-based resources and discussion forums relating to welfare assessment. It forms part of the full report or it can be downloaded separately at [http://la.rsmjournals.com/cgi/content/full/la.2010.010031/DC1]