



The **APC**
Animal Procedures Committee

Final report of a LASA/APC Working Group to examine the feasibility of reporting data on the severity of scientific procedures on animals

October 2008

Chairman of the Working Group:

David Smith (President, LASA)

Current members of the Working Group:

Peter Collins (AstraZeneca)
David Everett (Covance)
Kathleen Mathers (MRC NIMR)
Kathy Murphy (Oxford University)
Peter Pearce (DSTL)
Clare Stanford (University College London)
Dominic Wells (Imperial College)
Sara Wells (MRC Harwell)
Alan White (GlaxoSmithKline)

Past members of the Working Group:

Ken Applebee (King's College London)
Paul Finnemore (then Oxford University)
Gill Fleetwood (GlaxoSmithKline)
Alex Harper (GlaxoSmithKline)
Ghislaine Poirier (GlaxoSmithKline)
Sally Robinson (AstraZeneca)

Observers:

Bob Kemp (APC Working Group on Suffering and Severity)
Peter Thornton (Animals Scientific Procedures Inspectorate)

Rapporteur:

Jane A. Smith

Invited participants:

Julie Lane (Central Science Laboratory)
Grace Panter (AstraZeneca)

Contents

	Page
Key points	i
Summary	ii-iv
1 Background	1
2 Method of working	1
3 Options for reporting severity	2
3.1 Current methods in the UK: severity bands of projects and licence abstracts	2
3.2 Other options considered by the WG	2
<u>Option 1</u> : Report animals by 'protocol severity limit'	2
<u>Option 2</u> : Report maximum actual severity using categories mild, moderate and substantial	2
<u>Option 3</u> : Report maximum actual severity using another extant classification system	3
<u>Option 4</u> : Report maximum actual severity using an intensity-duration grid	4
<u>Option 5</u> : Report maximum and remaining actual severity using two intensity-duration grids	5
4 Weighing benefits and burdens of Options for reporting 2 to 5	7
5 Integrating severity data with other information in the <i>Statistics</i>	10
5.1 Retrospective severity reporting requires 'retrospective' annual Returns	10
5.2 Adding severity data to the annual Returns form	11
5.3 Enhancing the value of published severity data	11
5.4 Balance of benefits over burdens in the Returns process	11
6 Severity reporting in practice	12
6.1 General need for guidance	12
6.2 Should the 'moderate' category be divided?	12
6.3 Quality assurance	13
6.4 Studies involving use of developmental stages	13
6.5 Unclassified procedures	13
6.6 Other questions concerning scope of severity reporting	13
7 Further considerations	14
7.1 Implementing a severity reporting scheme	14
7.2 Public presentation and use of severity data	14
7.4 Communication with the European Commission	14
Appendices	
1 Home Office descriptions of 'mild', 'moderate' and 'substantial'	15
2 Swiss severity reporting scheme	16
3 Examples considered by the Working Group	17
4 Why Europe should adopt retrospective reporting of scientific procedures on animals	28
5 FELASA statement on a LASA proposal for the European Community to adopt retrospective reporting of scientific procedures on animals	29
6 Suggestions for reducing the burden of reporting data on animal procedures	30
Tables	
1 Grid system for reporting maximum intensity and duration of adverse effect	4
2 Two-grid system for reporting maximum and remaining severity	5
3 Pros and cons of introducing retrospective severity reporting	8

Key points

1. The LASA/APC Working Group (WG) believes that the introduction of a process for retrospective reporting of the severity of scientific procedures on animals would be beneficial in terms of enhanced openness and public accountability, and could also bring animal welfare and scientific benefits (see Table 3 pp. 8-9 for details).
2. However, the WG is also mindful of the additional bureaucratic burdens that severity reporting would bring (see also Table 3, pp.8-9).
3. Taking into account the pros and cons of a range of different options for reporting, the WG recommends that severity be recorded and reported using a single code for each animal – mild, moderate or substantial – to indicate the maximum actual severity of adverse effects experienced. In the WG's view, reporting using this scheme would offer the most favourable balance of benefit over burden.
4. At present, data published in the Home Office *Statistics of Scientific Procedures on Living Animals* are collected when procedures are started. However, severity data have to be collected when procedures have finished. To avoid having to collect and return data at both the beginning and end of procedures, and to enable cross-referencing of severity data with other information in the *Statistics*, the WG recommends that all statistical data required in the annual Returns be collected when procedures have finished.
5. Any change in the method of reporting animal procedures will require a transition period, to ensure that there is no under- or over-counting of procedures, burdens on licensees reporting data are minimised through adequate training, and there is preservation of historical trends information.
6. To facilitate the Returns procedure, the WG recommends that severity data be added as additional rows in the Returns form – one row for each severity category, at the end of the existing list of 12 categories (i.e. between the current rows 12 and 13). By this means, no new columns will be necessary, and severity data can be reported at the end of each column already generated for a given procedure.
7. Any scheme for retrospective reporting of severity must be supported by detailed guidance, including a catalogue of worked examples covering the full range of species and a wide variety of regulated procedures and outcomes.

Summary

General remarks

1. This is the final report of a Working Group (WG) convened under the auspices of Laboratory Animal Science Association (LASA) and the Suffering and Severity Working Group of the Animal Procedures Committee (APC), to assess the feasibility of collecting and reporting data on the actual severity of adverse effects experienced by animals used in procedures regulated under the Animals (Scientific Procedures) Act 1986 [ASPA].
2. The WG believes that the introduction of a process for retrospectively recording and reporting the actual severity of effects experienced by animals used in regulated scientific procedures would be beneficial in terms of enhanced openness and public accountability, and could also bring animal welfare and scientific benefits (paragraph 4.1 and Table 3).
3. However, the WG is also mindful of the additional bureaucratic burdens that such recording and reporting would bring. This means that, in recommending any such severity reporting system, the benefits would have to be carefully weighed against the burdens of the various options (paragraph 4.2 and Table 3).

Options for reporting and the WG's recommended method

4. The WG identified or devised four possible options for reporting *actual* severity, and explored the advantages and disadvantages of each (these are set out in detail in paragraph 3.2 and summarised in paragraph 4.3)¹. The different options vary in their complexity, and therefore also in the detail in which they can capture actual severity and associated burdens of reporting.
5. Taking into account the pros and cons of each option, and the results of testing against a wide range of examples (Appendix 3), the WG recommends that severity be recorded and reported using a single code for each animal – mild, moderate or substantial – to indicate the maximum actual severity of adverse effects experienced. In the WG's view, recording and reporting severity using this scheme would offer the most favourable balance of benefit over burden (paragraph 4.4).
6. Accepting the above option for routine recording and reporting of severity data need not preclude the use of a more detailed system (such as the WG's two-grid system, described on pp. 5-7) for occasional *ad hoc* reporting of severity data for animal use that is of special concern.

Incorporating severity data in annual Returns

7. At present, data published in the Home Office *Statistics of Scientific Procedures on Living Animals* [the *Statistics*] are collected when procedures are started. However, severity data have to be collected when procedures have finished. In order to avoid having to collect and return data at both the beginning and end of procedures, and to enable cross-referencing of severity data with other information in the *Statistics*, the WG recommends that all statistical data required in the annual Returns be collected when procedures have finished (paragraph 5.1).
8. Any change in the method of reporting animal procedures will require a transition period, to ensure that there is no under- or over-counting of procedures, burdens on licensees reporting data are minimised through adequate training, and there is preservation of historical trends information.

¹ The WG rejected current methods of reporting severity (publishing information on severity bands of project licences in the annual Home Office *Statistics of Scientific Procedures on Living Animals* and narrative descriptions of potential adverse effects in project licence abstracts on the Home Office web-site), and also a fifth new option – to report animals used by protocol severity limit (Option 1, p.2) – because none of these would reflect actual adverse effects, and the last would greatly over-estimate actual severity.

9. Since the format for statistical reporting in Britain is influenced by European requirements, the WG has prepared a brief paper outlining the arguments for (and against) adopting such a system of retrospective reporting Europe-wide (Appendix 4). This has been submitted for consideration as part of current work on revision of EU Directive 86/609. However, implementation of severity reporting in Britain need not be conditional on an EU decision.
10. To facilitate the Returns procedure, the WG recommends that severity data be added as additional rows in the Returns form – one row for each severity category, at the end of the existing list of 12 categories (i.e. between the current rows 12 and 13). By this means, no new columns will be necessary, and severity data can be reported at the end of each column already generated for a given procedure – e.g. if the WG's recommendation in 5 above is accepted, severity data would be reported as 'number of animals' in three new rows labelled 'mild', 'moderate' and 'substantial' actual severity (paragraph 5.2).
11. When severity information is made available in the public domain, systems should be in place to enable wide cross correlation of these data with other information collected in the annual Returns, so as to set severity in context (e.g. with respect to species involved, genetic status, purpose and legislative reason for the procedure – paragraph 5.3).
12. In a purely statistical publication, there will be little information on the benefits sought through the use of scientific procedures that cause adverse effects on animals. Consideration should be given to means of increasing information on potential benefits – e.g. by providing illustrative examples in the *Statistics* publication itself, or via the Abstracts database² (paragraph 5.3).
13. The addition of severity data will increase the effort required to make annual Returns (see paragraph 5.4 for more detailed discussion). To provide some mitigation, the WG has put forward a range of suggestions for reducing and streamlining other aspects of the Returns (Appendix 6). These have been communicated to the relevant sections of the Government's Better Regulation taskforce.

Severity reporting in practice

14. Any scheme for retrospective reporting of severity must be supported by detailed guidance, including a catalogue of worked examples covering the full range of species and a wide variety of regulated procedures and outcomes. In relation to the WG's recommended option for reporting, guidance will need to provide clear answers to the question, "What kinds of effects 'count' as 'mild', 'moderate' and 'substantial'?", and will need to cover duration as well as intensity of adverse effects. Such guidance will also be valuable in explaining the data when they are presented in the public domain (paragraph 6.1).
15. The WG also notes that several recent reports have proposed that the 'moderate' category be sub-divided, and recommends that this possibility be re-visited once there is experience of retrospective reporting using the current three-category classification (paragraph 6.2).
16. The WG recommends that established processes for monitoring the working of ASPA be used to help provide quality assurance in the reporting of severity data. This would include Home Office inspections, Ethical Review Process reviews, and informal assessments by/input from Named persons (paragraph 6.3).
17. The WG recommends that protocols involving the use of developmental stages protected under ASPA and relevant EU legislation (i.e. embryonic, fetal and larval forms from the appropriate stage of development) be exempt from a requirement for retrospective reporting of severity (see paragraph 6.4 for explanation).
18. Studies under terminal anaesthesia are coded as such in the annual Returns and therefore there is no further need to record or report severity for protocols run under an 'unclassified' severity limit (6.5).

² <http://scienceandresearch.homeoffice.gov.uk/animal-research/publications-and-reference/001-abstracts/>

19. Further points concerning the scope of severity reporting in practice are laid out in paragraph 6.6.

Further considerations

20. The WG has obtained feedback on the feasibility of the various options for severity reporting from the researchers who provided examples (Appendix 3), and from workshops involving Named Animal Care and Welfare Officers and laboratory animal veterinarians. Before any such scheme goes 'live', there should be a pilot study of the chosen reporting method at one or more Designated Establishments (paragraph 7.1).
21. Similarly, it will be advisable to obtain feedback from a wider constituency of 'public' who will use the information reported, in order to gain further insight into how severity data will be used by sample 'publics', and, from this, to determine how the data collected can be presented to best effect (paragraph 7.2).
22. Statistical reporting, including reporting of severity, and also severity classification more generally are amongst the issues currently under consideration by the European Commission, as part of the revision of Directive 86/609. The WG intends to submit this report to the European Commission so that its conclusions can be taken into account in the revision process (paragraph 7.3).

Final report of a LASA/APC WG to examine the feasibility of reporting data on the severity of scientific procedures on animals

1 Background

- 1.1 In response to recommendations from a number of expert bodies³, in 2005 a Working Group (WG) was convened under the auspices of Laboratory Animal Science Association (LASA) and the Suffering and Severity Working Group of the Animal Procedures Committee (APC), to assess the feasibility of collecting and reporting data on the actual severity of adverse effects experienced by animals used in procedures regulated under the Animals (Scientific Procedures) Act 1986 (ASPA).
- 1.2 The WG represented nine establishments drawn from industry/pharmaceutical organisations (3), universities (3) and major government research institutes (3). The roles of project licence holder, personal licensee, Named Veterinary Surgeon and Home Office liaison officer were represented in the group. A representative of the APC and a Named Animal Care and Welfare Officer (NACWO) also participated. A Home Office inspector attended all meetings as an observer.
- 1.3 Working within the terms of reference provided by the APC, the group sought to devise a method of providing information about suffering and severity actually experienced by individual animals that could be published in an annual report – whilst at the same time seeking to minimise the additional burdens on individuals and establishments required to collect and/or report the data.

2 Method of working

- 2.1 Initially, in order to engage more widely with project licence holders, a questionnaire was used to establish current practice in the nine establishments represented in the WG. Analysis of 168 responses showed that 9 out of 10 licence holders whose work involves moderate and/or substantial protocols always or sometimes make records of adverse effects actually experienced by the animals. The biggest difficulty they foresee is the ability to record this information in a way that will facilitate annual Returns. Their major concern is the resource burden and its possible impact on time available to pursue the science itself, as well as animal welfare. A detailed report of the questionnaire findings can be found at: www.apc.gov.uk/reference/lasa-appendix1.pdf
- 2.2 Next, drawing on the questionnaire findings, the experience of working group members and consideration of reporting models used in other countries (e.g. Switzerland), a range of initial options for reporting severity were identified and discussed. These options were then tested against a range of protocols (involving a variety of species and adverse effects) including some from working group members' own establishments. An interim report of this work, which describes the background and method of working in more detail, is at: <http://www.apc.gov.uk/reference/lasa-report.pdf>.
- 2.3 This final report summarises the various options for reporting severity considered by the WG, reviews the pros and cons of each, and presents the WG's conclusions. It also records the WG's views on a number of issues likely to arise in reporting severity in practice, and notes important steps that will be required to facilitate such reporting.

³ Including, in the UK, the House of Lords Select Committee on Animals in Scientific Procedures (2002) <http://www.publications.parliament.uk/pa/ld/ldanimal.htm>; an APC review of cost-benefit assessment in the use of animals in research (2003) <http://www.apc.gov.uk>; Boyd Group/RSPCA report of discussions on categorising severity (2004) <http://www.boyd-group.demon.co.uk>; and the Nuffield Council on Bioethics working party report on the ethics of research involving animals (2005) <http://www.nuffieldbioethics.org>. The issue of retrospective reporting of severity is also under consideration across Europe, within a more general review of EU statistical reporting requirements, which is being conducted as part of the current process of revision of EU Directive 86/609.

3 Options for reporting severity

3.1 Current methods in the UK: severity bands of projects and licence abstracts

Currently, the only published sources of information pertaining to the severity of animal procedures in the UK are:

- (i) an unnumbered Table in the annual Home Office *Statistics of Scientific Procedures on Living Animals* [hereafter referred to as the *Statistics*], listing the number of project licences in force, granted and revoked during the relevant year by their overall 'severity band' (a prospective assessment that reflects the "overall level of cumulative suffering [likely] to be experienced by each animal" across a whole project⁴), categorised as mild, moderate, substantial and unclassified; and
- (ii) narrative descriptions of the severity of anticipated effects on animals used in licensed work, provided in some but not all abstracts of new project licences published on the Home Office web-site (since January 2005).

WG's conclusion:

Neither of the current methods is designed for reporting actual severity. Both are *predictions* of potential adverse effects, not retrospective assessments of the actual adverse effects caused in practice; and so neither indicates the degree of pain, suffering, distress or lasting harm actually experienced by the animals involved in the projects. Moreover, the former source (severity banding) gives no indication of the severity of effects likely to be caused to individual animals, nor of the anticipated effects of the different kinds of protocol used in a given project⁵.

3.2 Other options considered by the WG

Option 1: Report animals used by 'protocol severity limit'

In the UK, protocol severity limits (mild, moderate, substantial or unclassified⁶) are used to set an upper limit to the suffering that an animal can experience in a given protocol. They are set at the project licensing stage and are used for management of severity in practice as well as in ethical review.

Pros and cons / conclusion:

Reporting the number of animals used by protocol severity limit would provide more detailed information, at the protocol rather than project level, but is inappropriate because protocol severity limits are *upper limits* to the suffering that might be caused in a given protocol and not *actual* adverse effects. Reporting by this method would greatly over-estimate actual severity, and would therefore provide seriously misleading information.⁷ It would also imply that in a given protocol every animal experiences adverse effects of the same severity.

Option 2: Report maximum actual severity using categories mild, moderate & substantial

The maximum severity of adverse effects caused to an animal in a given protocol could be reported using a single code – mild, moderate or substantial. These are the categories currently used in delineating protocol severity limits, but in this case would be used to report *actual* severity. The categories are defined in Home Office guidance, as shown in Appendix 1.

Pros and cons

- The classification system mild, moderate and substantial is already understood and used (albeit for a different, but related, purpose) by licence holders, Named persons and others associated with ASPA; and many licence holders already routinely make records of actual

⁴ Home Office (2008) *Project licence application form* Note 1a. <http://scienceandresearch.homeoffice.gov.uk/animal-research/publications-and-reference/publications/licences/project-licences/>

⁵ See also Boyd Group/RSPCA (2004).

⁶ The 'unclassified' severity limit is used for procedures carried out wholly under general anaesthesia from which animals are not permitted to recover.

⁷ See also Boyd Group/RSPCA (2004)

adverse effects experienced by animals bearing this classification in mind (particularly for moderate and substantial severity protocols).

- The above features, combined with its relative simplicity, would make this option the least burdensome of the various means of reporting considered by the WG (see paragraph 5.4 for further discussion).
- Reporting just three categories of actual adverse effect would also enable severity data to be cross-correlated with other information in the *Statistics* (e.g. species and purpose) without making the publication unduly complex (see also section 5.3 on presentation of severity data).
- The system would be most meaningful for short-term protocols, but would have less descriptive power for more complex, longer-term procedures that generate variable severity profiles over time.
- For this – and any other – method of reporting to be meaningful, there would need to be much more detailed description of what the different categories comprise, covering duration as well as intensity of adverse effects (see also paragraph 6.1).
- The WG notes that the 'moderate' category is broad and encompasses a wide range of degrees of severity. See paragraph 6.2 for further discussion.

Option 3: Report maximum actual severity using another extant classification system

A number of European and other countries currently report data pertaining to the severity of animal procedures. The WG examined a number of these reporting schemes to see whether they offer advantages over Option 2 above.

Some countries, including Canada, New Zealand, and several states in Australia, classify the severity of *procedures* and report the number of animals used under each classification level. In these cases, the data reported do not necessarily reflect the actual experience of the animals; rather they are intended to capture what happens to the 'majority' of animals used in each procedure. They could therefore over- or under-represent the degree of suffering *actually* experienced by the animals.

The Netherlands and Switzerland collect and report data on the *actual* impact of scientific procedures on animals:

- in The Netherlands six categories (minor; minor-moderate; moderate; moderate-severe; severe; and very severe) are employed, but there is little or no written guidance on how these are to be assigned; rather, the system rests on the judgement of researchers, who are advised by local Animal Welfare Officers; and
- in Switzerland four severity groupings are used – the lowest of which (Degree of Severity 0) is essentially a 'no effect' category, not regulated in the UK. The descriptions of the categories are given in Appendix 2, and in practice are supported by numerous examples.

Of the above reporting schemes, the Swiss classification is the only one used to report actual severity that also provides clear descriptions of the categories. The WG therefore focused on the Swiss scheme when considering the pros and cons of reporting according to another, extant, classification system:

Pros and cons

- In terms of range of categories, the scheme used to collect and report data on actual severity in Switzerland does not differ significantly from that in Option 2, given above. This means that adopting such a scheme would not increase descriptive power, nor decrease the burdens on licensees and others who would be involved in collecting and reporting those data. Indeed, initially the burdens would most likely be greater than those imposed by Option 2, because the categories would be unfamiliar.

- The Swiss scheme does offer an illustration of category definitions that explicitly include *duration* as well as intensity of suffering (as it is noted would be needed should Option 2 above be adopted) – but these definitions highlight a potential problem:

With relatively few severity categories, the Swiss scheme combines more substantial but relatively short-term effects with milder but longer-term effects into single codes, as happens in the Swiss Degrees of Severity categories 2 and 3 (Appendix 2). Degree of Severity 2, for example, encompasses both moderate short-term adverse effects *and* mild (slight) medium to long-term adverse effects. Can these different effects really be considered 'equivalent' kinds of suffering, as the Swiss coding would suggest?

- With more codes, the need to combine more substantial short-term effects with longer-term but milder effects is reduced.

Option 4: Report maximum actual severity using an intensity-duration grid

In order to explicitly include duration as well as intensity of suffering in a reporting scheme, and avoid the problems alluded to above, the WG explored the possibility of reporting severity using a grid system, in which two parameters – maximum intensity of suffering and its duration – are considered independently (Table 1):

Table 1: Grid system for reporting maximum intensity & duration of adverse effects

Intensity of adverse effects	Duration of adverse effects		
	Short	Medium	Long
Mild	A	B	C
Moderate	D	E	F
Substantial	G	H	I

As in Option 2, both the intensity and duration categories would need careful and clear definition.

Pros and cons

- Increasing the number of categories in this way avoids having to combine into a single code more substantial but relatively short-term effects, and milder but longer-term, effects.
- The grid works well with procedures that have relatively simple severity profiles but, because it reports only maximum severity, is less successful in capturing the severity of more complex, longer-term procedures.
- This scheme requires those recording the data to make two judgements – duration and intensity of maximum suffering – and then to read the appropriate code from the grid. In theory, therefore, choosing codes for reporting should be no more burdensome than in Option 2 (which would require similar judgements if, as proposed, the category descriptions are modified to include duration as well as intensity of suffering). However, because there are more categories in Option 3, recording and reporting these data in the annual Returns would be more burdensome.
- In the resulting *Statistics* publication cross-correlating these severity data with other information in the *Statistics* (e.g. species, genetic status and purpose) would result in considerably more complex tables (unless some categories are combined).

Option 5: Report maximum & remaining actual severity using two intensity-duration grids

The WG also considered means of capturing severity in longer-term procedures involving more complex interventions that would result in better representation of the course of suffering over the whole procedure (i.e. what might be thought of as the 'area under the curve' in a graph of severity of adverse effect over the time).

One way of achieving this would be to report severity using two intensity-duration grids, to indicate (i) maximum severity and (ii) severity over the remainder of the procedure. Using this system, the actual severity experienced by an animal in a given procedure would be assigned a two-letter code – e.g. AA, EC (Table 2):

Table 2: Two-grid system for reporting maximum and remaining severity

Grid 1: *Maximum* severity within the procedure

Maximum intensity of adverse effects	Duration of maximum intensity of adverse effects within the procedure		
	Short ≤ one day	Medium >1 and ≤ 7 days	Long > 7 days
Mild	A	B	C
Moderate	D	E	F
Substantial	G	H	I

Grid 2: *Remaining* severity, over the rest of the procedure (*less* the maximum)

Overall intensity of adverse effects in rest of procedure	Duration of the remainder of the procedure		
	Short ≤ 7 days	Medium >7 and ≤ 28 days	Long >28 days
Mild	A	B	C
Moderate	D	E	F
Substantial	G	H	I

This two-grid scheme requires four independent decisions: two of which are subjective – intensity of the peak suffering and severity of any remaining suffering – and two of which are more objective: duration of the peak and duration of any remaining suffering. The durations given above are the WG's initial suggestions. Their appropriateness would need further consideration should such a reporting scheme be adopted.

Pros and cons

- Of the options considered by the WG, the two-grid system has the potential to provide the most representative picture of intensity and duration of severity over a wide range of procedures.
- A total of 39 codes are possible (not 81 as might first be thought, because ‘remaining’ severity cannot be greater than ‘maximum’ severity). It is likely that (a) for many but not all studies only a few of these codes will be needed, and (b) for many procedures, severity, and therefore codes for particular animals, can be predicted in advance, so that reporting can be done by a process of identifying and recording exceptions. But, nevertheless, this is still the most complex option considered by the WG and would therefore impose the greatest burdens on those recording and reporting the data.
- The codes would require amalgamation and explanation for public reporting, both for ease of understanding, and also to enable cross-correlation of data with other information in the *Statistics* (e.g. species and purpose) without overly increasing the complexity of the resulting publication. Since simplification would be required for reporting, it might be judged unnecessary to collect such complex data in the first place. However, it is also possible that the full data-set could be made available on-line (along with other information in the Returns), so that those with a particular interest could interrogate the detailed data for themselves (see also paragraph 5.3).⁸
- The WG devised a possible way of simplifying this system, so as to reduce the number of possible codes, but found that this revised reporting scheme was in fact less intuitive and more difficult to use than the original two-grid system.
- Although, in general, tests of the two-grid scheme suggest that, once explained, it is reasonably intuitive, the WG has identified some specific difficulties in applying this method of reporting in practice:
 - (i) A particular concern, revealed during testing against a wide range of examples, is that the scheme makes it difficult to record effects on animals that have the same severity throughout the duration of a procedure. Intuitively, one would expect that in such cases the maximum and remaining severity codes would be the same (e.g. AA, BB), but because the durations (of necessity) differ between the two grids, this is not always the case. For example, a procedure lasting more than 1 but less than 7 days that had mild effects on an animal throughout its duration would receive the code BA – not BB as might be expected.

It is possible that using different letters on Grids 1 and 2 (i.e. with Grid 2 codes running from J to R) could help solve this problem. However, when tested in a workshop of laboratory animal veterinary surgeons, 11% of all codes assigned using this modified two-grid system were wrong due to difficulties in coding adverse effects that had the same severity throughout the duration of a procedure. Some respondents also reported that they found the different letters on the two grids difficult and would prefer the same letters on each (i.e. as set out in Option 5 above).
 - (ii) Small differences in judgements on the duration of peak adverse effects can result in a wide range of different codes for procedures lasting less than a week. For example, it might be agreed that a procedure lasting less than 48 hours has maximum intensity of severity ‘moderate’. But disagreement about whether these moderate effects persist more than a day and/or persist for the full 48 hours would lead to four different codes (DA, EA, DD, and ED). In such a case, relatively minor differences in judgement lead to apparently wide variation in coding using the two-grid scheme as formulated above. See Appendix 3, example 6 and commentary in example 2.

⁸ An APC report proposes this strategy for the *Statistics* in general: www.apc.gov.uk/reference/stats-report270505.pdf

- (iii) The level of detail in this scheme also brings logistic difficulties in reporting certain types of animal use – e.g:
- some studies (such as some toxicity tests and use of certain disease models) in which animals are humanely killed at different times, having suffered different intensities of adverse effect, would require a large number of different severity codes (and associated judgements) – see especially example 9 in Appendix 3;
 - in studies involving wild animals, such as rodenticide studies, and similar ecological work, it is often difficult to discern duration and intensity of suffering to the level of detail required in this scheme – see last paragraph of example 12 in Appendix 3;
 - breeding, but not otherwise using, genetically altered animals, and particularly those in which the genetic alteration causes no overt adverse effect, will lead to a large number of different codes dependent simply on the time at which animals are killed – see example 7 in Appendix 3.
- (iv) Coding ‘maximum’ severity can be difficult when there is more than one ‘peak’ of adverse effects of the same intensity during a given procedure. To be consistent, in such instances duration of maximum intensity should be the sum of the peaks – but discerning this sum is difficult in practice and would increase burdens on those recording and reporting the information.

4 Weighing benefits and burdens of Options for reporting 2 to 5

4.1 The WG believes that the introduction of a process for retrospectively recording and reporting the actual severity of effects experienced by animals used in regulated scientific procedures would be beneficial in terms of enhanced openness and public accountability, and could also bring animal welfare and scientific benefits.

The latter benefits could include encouraging more consistent and detailed observation of adverse effects on animals, helping to promote and highlight priorities for refinement and benchmarking good practice, better informing the Ethical Review Process, and helping to ensure that accurate information on actual severity enters the public domain – in order, for example, to substantiate (or refute) claims by the research community that many or most procedures impose ‘mild’ effects on the animals involved. These and further benefits identified by the WG are set out in detail in Table 3 overleaf, along with some of the downsides of severity reporting.

4.2 However, the WG is also mindful of the additional bureaucratic burdens that such recording and reporting would bring. This means that, in recommending any such severity reporting system, the benefits will have to be carefully weighed against the burdens of the various options.

In particular, there is a balance to be struck in meeting the goals of:

- ensuring that this information is sufficiently representative of the actual adverse effects of scientific procedures on animals to help achieve the benefits listed above, and, importantly, neither under- nor over-estimates the severity of the effects concerned;
- enabling information on severity to be collected and reported in an intuitive, consistent and minimally bureaucratic way;
- providing easily understandable information for presentation in the public domain; and, as part of this, enabling cross-referencing of severity data with information on other aspects of animal use (e.g. species and purpose).

Table 3: Pros and cons of introducing retrospective severity reporting

	Pros	Cons
Animals	<p>Severity reporting should encourage more consistent and detailed observation of the animals, which in turn could stimulate more action to refine procedures and end-points – see also 'Named persons' below.</p> <p>Better recording and reporting of adverse effects should also help in highlighting priorities for refinement and benchmarking good practice.</p>	<p>The burden of severity reporting could divert resources (including time) that might have been used to enhance animal welfare</p> <p>Although it should be possible to track progress with refinement through published severity data, published information may not reveal all such progress – because reduction in severity can take place without crossing a severity category boundary. In this context, sub-dividing 'moderate' severity would help to provide a more accurate picture</p>
Animal care staff	<p>The process of classifying severity could trigger discussions involving animal care staff that might strengthen their interactions/relations with licence holders and Named persons.</p> <p>This might also help to reinforce the importance and value of their role, especially in observing and monitoring animals</p>	<p>There would be resource – mainly time – burdens on animal care staff, particularly if they are also personal licence holders</p> <p>Care would also be needed to ensure that licence holders do not pass the burdens and responsibility of severity reporting on to animal care staff, and thereby divert them from their primary role.</p>
Named persons⁹	<p>All of the following could assist Named persons in their statutory role:</p> <p>A requirement on licence holders to report severity retrospectively could lead to better definition of clinical signs by severity category and therefore more clarity - in advance - about the circumstances in which severity limits would be approached and considered breached</p> <p>The need to report severity could also provide a platform for discussion between licence holders and Named persons about severity limits <i>per se</i>, and this again could help to promote refinement</p> <p>Better guidance and examples to support severity classification could lead to more agreement on the significance of particular clinical signs and help in gaining agreement on the management of severity and end-points in practice</p>	<p>Named persons would undoubtedly be subject to pressure, with associated time and other resource burdens, to help licensees with grading and recording severity</p> <p>They might be subject to particular stress when faced with disagreement or inconsistency in grading particular uses of animals – but better guidance should help.</p>
Personal licensees (PILs)	<p>More systematic gathering and recording of severity information could assist PILs in meeting their statutory responsibilities – e.g. with respect to animal welfare (standard condition 12), severity limits (standard condition 13) and veterinary advice (standard condition 16)</p>	<p>PILs would need to keep more detailed records – in particular, of the severity of adverse effects experienced by animals used within moderate and substantial limit protocols. This would increase burdens on PILs</p> <p style="text-align: right;"><i>continued...</i></p>

⁹ That is Named Animal Care and Welfare Officers and Named Veterinary Surgeons. See: Home Office (2000). *Guidance on the operation of the Animals (Scientific Procedures) Act 1986*. TSO: London. Available at <http://www.archive.official-documents.co.uk/document/hoc/321/321.htm>

<p>Project licence holders (PPLs)</p>	<p>As with PILs, systematic gathering and recording of severity information could assist PPLs in meeting their statutory responsibilities – e.g. with respect to refinement of procedures (standard condition 6), and severity limits (standard conditions 6 and 9)</p> <p>The data collected could help PPLs to track their own progress with refinement, manage the project, and provide evidence to support the justification for any subsequent project licence application</p> <p>Taking on the burdens of severity reporting would bring an opportunity for PPLs to negotiate reduction in reporting requirements elsewhere in the annual Returns and perhaps other bureaucracy associated with ASPA</p> <p>The need to report severity might encourage closer engagement of more 'remote' PPLs with animal facility staff, which in turn could bring benefits for those staff and the animals</p>	<p>PPLs are responsible for keeping records on all animals on which procedures have been carried out (standard condition 9) and submitting these as part of the annual statistical return (standard condition 10). Retrospective reporting of severity would increase burdens in respect of both activities</p>
<p>Ethical review processes (ERPs) and the Home Office</p>	<p>Systematic recording and reporting of actual severity could enable ERPs and the Home Office to gain a better view of the adverse effects caused to animals in practice</p> <p>This could assist in retrospective and interim review of work in progress, especially in respect of application of the 3Rs</p> <p>There could, for example, be particular value in using these data to compare effects of the same protocol used in different projects and to track refinement over time (but on the latter, see the caveat under 'animals' above)</p>	<p>Increased burdens of collating, analysing and disseminating data; and possibly of dealing with increased requests under Freedom of Information legislation</p>
<p>Public</p>	<p>At present, useful data on the severity of scientific procedures on living animals are lacking – yet, arguably, these statistics are the most valued (alongside information about benefits) by wider 'publics'. Retrospective reporting of actual severity would help to fill this gap</p> <p>It would enable those with a particular interest (such as animal welfare and anti-vivisection organisations and patient and research defence groups) to gain a clearer picture of the relative severity of different areas of work.</p>	<p>It is difficult to link severity data with information on benefits sought – so the justification for the harms caused will be opaque (though there are some possibilities for consideration – see paragraph 5.3).</p>

- 4.3 In light of the criteria outlined in paragraph 4.2, and having tested the schemes against a wide range of examples (Appendix 3), the WG's conclusions on the pros and cons of the various options for reporting actual severity are as follows:

Option 2 – reporting maximum severity using a single code – mild, moderate or substantial: is the least complex of the reporting schemes. This means that although time and effort would be required to record and return this information, this option would impose the least additional burdens. The scheme also uses categories already familiar to those who would be required to record, report and use the severity data. Moreover, because there are relatively few categories, the scheme offers opportunities for wide cross-correlation of data within the *Statistics* without simplification – which would greatly improve the usefulness of the information collected.

Option 3 – reporting using the Swiss scheme: would offer no particular advantage over Option 2 – provided that existing Home Office descriptions of the categories used in Option 2 (see Appendix 1) are modified to include explicit reference to duration as well as intensity of severity.

Option 4 – single intensity-duration grid system, to report maximum severity: would offer a little more descriptive power than Option 2. Option 4 would require those reporting and recording data to make judgements similar to those required in Option 2, but, because there are more categories, would be more burdensome to report and more difficult to widely cross-correlate with other data in the *Statistics*. There will also be difficulties in respect of judgements about duration, similar to those outlined for Option 5 below, and in more detail in points (ii) and (iii) on pp. 6-7.

Option 5 – the two-grid system: is capable of providing the most representative picture of severity. But this scheme is also the most complex and therefore the most burdensome of the reporting systems explored by the WG. When published, data collected by this system would require additional interpretation (since the codes are complicated) and some amalgamation (since there are a large number of codes) both for ease of understanding and to enable wide cross-referencing of severity data with other information in the *Statistics*. There are also some important difficulties in applying Option 5 in practice. Of these, perhaps the most important (and most difficult to overcome) is that small differences in judgements about duration of adverse effects can lead to wide-scale variation in codes assigned.

- 4.4 **Taking into account the pros and cons of each option, the WG recommends adoption of Option 2 – i.e. retrospective recording and reporting of severity using a single code for each animal – mild, moderate or substantial – to indicate the maximum actual severity of adverse effects experienced. In the WG's view, recording and reporting severity using this scheme would offer the most favourable balance of benefit over burden.**
- 4.5 Of course, accepting this option for routine recording and reporting of severity data need not preclude the use of the more detailed two-grid system for occasional *ad hoc* recording and reporting of severity data for animal use that is of special concern (e.g. occasional use of two-grid scheme by ERPs or the Home Office).

5 Integrating severity data with other information in the *Statistics*

5.1 *Retrospective severity reporting requires 'retrospective' annual Returns*

- (i) Severity data, by their nature, have to be collected 'after the event' (i.e. when procedures are *finished*) and, in that sense, can be said to be 'retrospective'. However, at present, all (other) Home Office statistics on the use of laboratory animals are collected when procedures are *started*.

The WG recommends that all statistical data required in the annual Returns be collected when procedures have finished.

This will enable cross-referencing of data on severity with other information about animal use, and, importantly, will also avoid a need to collect two different sets of data – one at the start and one at the end of procedures – which the WG believes would be excessive and unnecessarily burdensome.

In making this recommendation, the WG is mindful that any such change in the method of reporting animal procedures will require a transition period, to ensure that there is no under- or over-counting of procedures, burdens on licensees reporting data are minimised through adequate training, and there is preservation of historical trends information.¹⁰

- (ii) Since the format for statistical reporting in Britain is influenced by European requirements, the WG has prepared a brief paper outlining the arguments for (and against) adopting such a system of retrospective reporting Europe-wide (Appendix 4). This has been submitted for consideration as part of current work on revision of EU Directive 86/609. The Federation of European Laboratory Animal Science Associations has recently issued a positive response to the WG's proposal that all such data should be collected retrospectively (Appendix 5). However, implementation of severity reporting in Britain need not be conditional on an EU decision.

5.2 Adding severity data to the annual Returns form

To facilitate the Returns procedure, the WG recommends that severity data be added as additional rows in the Returns form – one row for each severity category, at the end of the existing list of 12 categories (i.e. between the current rows 12 and 13).

For example, if the WG's recommendation in 5 above is accepted, severity data would be reported as 'number of animals' in each of three new rows labelled 'mild', 'moderate' and 'substantial' actual severity. By this means, no new columns will be necessary, and severity data can be reported at the end of each column already generated for a given procedure.

5.3 Enhancing the value of published severity data

Clearly, it is important that severity data are presented to best effect in the *Statistics* publication, so as to shed as much light as possible on the actual experience of the animals that are used in scientific procedures.

When severity information is made available in the public domain, systems should be in place to enable wide cross correlation of these data with other information collected in the annual Returns, so as to set severity in context (e.g. with respect to species involved, genetic status, purpose and legislative reason for the procedure).

This will also help, for example, in pin-pointing areas in which severity is changing over the years and to track progress with refinement. However, in a purely statistical publication, there will be little opportunity to indicate the benefits sought from procedures that cause adverse effects on animals.

Consideration should be given to means of increasing information on potential benefits – e.g. by providing illustrative examples in the *Statistics* publication itself, or via the Abstracts database.¹¹

5.4 Balance of benefits over burdens in the Returns process

The addition of severity data will increase the effort required to make annual Returns. To provide some mitigation, the WG has put forward a range of suggestions for reducing and streamlining other aspects of the Returns.

¹⁰ The possibility of moving to a system of retrospective reporting is also explored in an APC Statistics Working Group report, published in 2005 www.apc.gov.uk/reference/stats-report270505.pdf

¹¹ <http://scienceandresearch.homeoffice.gov.uk/animal-research/publications-and-reference/001-abstracts/>

These are listed in Appendix 6 and have been communicated to the relevant sections of the Government's Better Regulation taskforce.

It should also be noted that the WG's recommended reporting scheme is the least burdensome of the options considered:

For protocols run under a mild severity limit, Returns for all animals (bar a very few exceptions) would be under the 'mild' severity category; protocols run under a moderate severity limit would require project licence holders to choose whether to place animals in the 'mild' or 'moderate' categories; whilst 'substantial' severity limit protocols would require them to choose from all three severity categories. The burdens would therefore increase alongside the severity of the protocol, but there are likely to be fewer animals involved at the higher severity levels.

It should also be noted that for the majority of substantial protocols and for many moderate ones, licence holders already have systems in place for recording severity (see paragraph 2.1).

Moreover, responses to the WG's questionnaire to project licence holders¹² (paragraph 2.1) also suggest that for many, but not all, studies it is possible to predict severity in advance (e.g. whether animals will experience mild or moderate effects within a protocol that has a moderate severity limit), and so specific recording would be required for exceptions only.

6 Severity reporting in practice

6.1 General need for guidance

Any scheme for retrospective reporting of severity must be supported by detailed guidance, including a catalogue of worked examples covering the full range of species and a wide variety of regulated procedures and outcomes. Such guidance will also be valuable in explaining the data when they are presented in the public domain.

In particular, guidance will need to provide clear answers to the question, "What kinds of effects 'count' as 'mild', 'moderate' and 'substantial'?", and will need to cover duration as well as intensity of adverse effects.

The LASA/APC WG has undertaken an initial literature review of extant systems for assessing and categorising adverse effects into mild, moderate and substantial and how these are used in practice in the UK, and this has been passed to the APC's Suffering and Severity WG, which is also considering this question.

The cases considered by the WG (Appendix 3) provide a start in developing a catalogue of worked examples to support such guidance. The examples that accompany the Swiss scheme and the detailed descriptions in an augmented version of the New Zealand scheme (*pers comm.*) should also be helpful in assembling a UK catalogue.

6.2 Should the 'moderate' category be divided?

As noted in discussion of the pros and cons of the WG's recommended reporting scheme (Option 2, described on pages 2-3) the 'moderate' category is broad and encompasses a wide range of degrees of severity.

In this context, the WG notes that several recent reports have proposed that the 'moderate' category be sub-divided into upper and lower moderate¹³. The WG's case studies, presented in Appendix 3, lend some support to this view. They illustrate a broad range

¹² Report available at: www.apc.gov.uk/reference/lasa-appendix1.pdf

¹³ See APC review of cost-benefit assessment in the use of animals in research (2003) <http://www.apc.gov.uk>; Boyd Group/RSPCA report of discussions on categorising severity (2004) <http://www.boyd-group.demon.co.uk>; and report of a Nuffield Council on Bioethics on the ethics of research involving animals (2005) <http://www.nuffieldbioethics.org>.

effects that current guidance would place in the 'moderate' category, and, accordingly, for each example the WG has suggested how these might be divided into 'upper' and 'lower' moderate categories.

It is clear that sub-division of the moderate category would improve the descriptive power of the classification system. This would, in turn, assist in tracking progress in refining procedures. This is because refinement often takes place in small steps, and reductions in severity will not necessarily cross severity category boundaries (e.g. from moderate to mild), but might occur within categories (e.g. from the upper to the lower end of moderate).

Nevertheless, the WG also recognises that increasing the number of severity categories would increase the burdens of reporting, and that it might be difficult to clearly delineate the boundary between upper and lower moderate.

The WG recommends that the possibility of sub-dividing the moderate category be re-visited once there is experience of retrospective reporting using the current three-category classification.

6.3 *Quality assurance*

Good guidance and training, combined with good communication (e.g. between project licence holders, Named persons and ERPs), on assessment of severity should help to promote consistency in the classification of adverse effects. However, because judgements are involved, there will inevitably be some room for disagreement and no system will provide perfect consensus.

The WG recommends that established processes for monitoring the working of ASPA be used to help provide quality assurance in the reporting of severity data. This would include Home Office inspections, Ethical Review Process reviews, and informal assessments by/input from Named persons.

6.4 *Studies involving use of developmental stages*

The WG recommends that protocols involving the use of developmental stages protected under ASPA and relevant EU legislation (i.e. embryonic, fetal and larval forms from the appropriate stage of development) be exempt from a requirement for retrospective reporting of severity.

This is because of the difficulty (maybe impossibility) of assessing what these stages experience, and because including them would most likely artificially inflate the 'mild' severity categories – and in any case data on use of developmental stages are not currently reported in the *Statistics*. More generally, the WG is agreed that excluding areas in which severity judgements are, for similar reasons, difficult at the present time is likely to enhance public trust in the data that are reported.

6.5 *Unclassified procedures*

Studies under terminal anaesthesia are coded as such in the annual Returns and there is therefore no further need to record or report severity.

6.6 *Other questions concerning scope of severity reporting*

Should severity reporting include effects of 'contingent' factors, such as those pertaining to source of animals and quality of 'standard' housing and husbandry, as well as effects of the regulated procedures, as detailed in project licence 19b protocols?

It is the WG's view that any such contingent effects should not be reported, because the *Statistics* pertain to regulated procedures only. Therefore retrospective severity reporting should include only the severity of effects of regulated interventions. This would include effects of housing and husbandry when the methods used are covered in 19b protocols (e.g. individual housing of social animals – see Appendix 3, Case 2)

Should there be a 'no effect' category when no overt adverse effects have been caused to the animals?

The absence of overt signs does not necessarily mean that there are no adverse effects. The 'mild' intensity category should be regarded as 'up to and including mild'.

7 Further considerations

7.1 *Implementing a severity reporting scheme*

The WG has obtained feedback on the feasibility of the various options for severity reporting from the researchers who provided the examples in Appendix 3, and via a NACWO workshop organised by the Institute of Animal Technology, and also a Laboratory Animal Veterinary Association workshop for laboratory animal veterinarians.

The WG recommends that before any such scheme goes 'live', there should be a pilot study of the chosen reporting method at one or more Designated Establishments.

The feedback from the pilot will be particularly important in helping to ensure that, when implemented in practice, the balance of benefits over burdens of severity reporting is optimised, and that any specific difficulties not yet identified by the WG are addressed.

7.2 *Public presentation and use of severity data*

Similarly, it will be advisable to obtain feedback from a wider constituency of 'public' who will use the information reported, in order to gain further insight into how severity data will be used by sample 'publics', and, from this, to determine how the data collected can be presented to best effect.

7.3 *Communication with the European Commission*

Statistical reporting, including reporting of severity, and also severity classification more generally, are amongst the issues currently under consideration by the European Commission, as part of the revision of Directive 86/609.

The WG intends to submit this report to the European Commission so that its conclusions can be taken into account in the revision process.

APPENDIX 1: Home Office descriptions of 'mild', 'moderate', and 'substantial'¹⁴

Mild

"Protocols that, at worst, give rise to slight or transitory minor adverse effects. Examples include: small infrequent blood samples; skin irritation tests with substances expected to be non-irritant or only mildly irritant; minor surgical procedures under anaesthesia such as small superficial tissue biopsies or cannulation of peripheral blood vessels. However, if used in combination or repeated in the same animal, the cumulative severity may be increased beyond mild. Protocols may also be regarded as mild if they have the potential to cause greater suffering but contain effective safeguards to initiate effective symptomatic or specific treatment or terminate the protocol before the animal shows more than minor adverse effects."

Moderate

"Protocols regarded as moderate include toxicity tests (which do not involve lethal endpoints) and many surgical procedures (provided that suffering is controlled and minimised by effective post-operative analgesia and care). Protocols that have the potential to cause greater suffering but include controls which minimise severity, or terminate the protocol before the animal shows more than moderate adverse effects, may also be classed within the moderate severity limit."

Substantial

"Protocols that may result in a major departure from the animal's usual state of health or well-being. These include: acute toxicity procedures where significant morbidity or death is an endpoint; some efficacy tests of anti-microbial agents and vaccines; major surgery; and some models of disease, where welfare may be seriously compromised. If it is expected that even one animal would suffer substantial effects, the procedure would merit a 'substantial' severity limit."

¹⁴ Home Office (2000). *Guidance on the operation of the Animals (Scientific Procedures) Act 1986*. TSO: London
Available at <http://www.archive.official-documents.co.uk/document/hoc/321/321.htm>

APPENDIX 2: Swiss severity reporting scheme

In Switzerland, at the end of each experiment, researchers are required to report the number of animals that suffered (or not) each of four different 'degrees of severity' of adverse effect, so that annual statistics can be compiled. The definitions of the severity categories (numbered 0 to 3) encompass intensity as well as duration of adverse effect, and the *maximum* severity caused to the animals is reported. The basic definitions of the different 'degrees of severity' are listed below, and are elaborated (elsewhere) in detailed guidance containing numerous examples.¹⁵

Degree of Severity 0: "Interventions and manipulations in animals for experimental purposes as a result of which the animals experience no pain, suffering, injury, or extreme anxiety and no significant impairment of their general condition"

e.g. animals used for experiments which do not require authorisation (but have to be registered), or control animals on which no interventions with adverse effects were conducted

Degree of Severity 1: "Interventions and manipulations in animals for experimental purposes which subject the animals to a brief episode of mild stress (pain or injury)"

e.g. animals killed in pre-terminal narcosis or rabbits immunised without use of Freund's adjuvant

Degree of Severity 2: "Interventions and manipulations in animals for experimental purposes which subject the animals to a brief episode of moderate stress, or a moderately long to long-lasting episode of mild stress (pain, suffering, or injury, extreme anxiety, or significant impairment of general condition)"

e.g. animals in which electrodes were implanted in the brain or those which underwent adrenalectomy

Degree of Severity 3: "Interventions and manipulations in animals for experimental purposes which cause the animals severe to very severe stress, or subject them to a moderately long to long-lasting episode of moderate stress (severe pain, prolonged suffering or severe injury; extreme and persistent anxiety, or significant and persistent impairment of general condition)"

e.g. animals which underwent thoracotomy or ischaemia of the brain with subsequent substantial functional disorders. In long-term studies (e.g. lifetime toxicity study) spontaneously deceased animals must be allocated to severity degree 3, unless such animals are humanely killed at an earlier stage in clinical symptoms, when a correspondingly lower degree of severity will be allocated

¹⁵ Available at: http://www.tierversuch.ch/?show=AWLaw&nav_id=4104&lang=en

APPENDIX 3: Examples considered by the WG

1: Efficacy study of oncology compounds

Study design:

The aim of the study was to assess the potential of test compounds to inhibit tumour growth. Nude mice were sub-cutaneously implanted with tumour cells, so that a small, single tumour grew just under the skin on a flank where it would not cause discomfort, and after a set period initial tumour growth was measured. Animals with 'acceptable' levels of tumour growth (n=95) were then randomised into study groups, including control, sham-dosed control and test compound groups. Treatment was administered once daily by oral gavage. Reduction in tumour size compared with tumour growth in control animals indicated anti-tumour potential in the test compound.

Duration: Initial tumour growth took 14 days, at which point the animals were randomised into study groups. The treatment period post randomisation was 28 days. Total study duration was 42 days.

Monitoring:

Mice were observed at least twice daily throughout the study. Body weights and tumour size were recorded, along with any clinical signs. Mice were weighed daily for the first week and then 3 times weekly when tumours were also measured. If any mouse lost more than 8% initial body weight it was then weighed daily.

It was predicted that the majority of mice would suffer only 'mild' adverse effects because of the use of humane endpoints. Limits were set on tumour size and condition, weight loss and clinical condition beyond which the animals were humanely killed. However, the study was run under a 'moderate' severity limit because test compounds could have had unexpected adverse effects or there could have been unexpectedly rapid tumour growth or ulceration.

Severity classification

Tumours were no larger than 14mm in any direction. Observed effects fell into three categories:

Actual adverse effects on animals:	Severity coding	
	2-grid	Mild, Moderate, Substantial
91 mice, although bearing small tumours, appeared normal and had no overt adverse signs such as weight-loss	CC	Mild
3 mice suffered overt adverse effects (>10% body weight loss in one animal and ulceration of the tumours in two animals). They were humanely killed within 24 hours of effects appearing, on day 38	DC	Moderate (<i>lower</i>)
One animal had significant adverse effects including weight loss (>20%) and was humanely killed on day 40	GC	Substantial

2: Acute toxicity study

In pharmaceutical drug development acute toxicity studies in rodents are a regulatory requirement. The objective is to identify a dose causing major adverse effects after a single dose up to a limit of 2000 mg/kg. The animals receive a single dose of the test compound followed by up to 14 days observation post-dose. The following two examples have been selected at random:

Study design:

A dose-range (up to 2000 mg/kg) is initially investigated in 1 male and 1 female animal *per* dose level. A group of 5 + 5 animals is dosed at the anticipated maximum non-lethal dose. Occasionally it may be necessary to add an additional group of up to 5 + 5 animals. Typical study sizes range from 10 to 25 animals. In this case, study 1 involved 22 rats; and study 2, 10 rats

Monitoring:

These studies are run under a substantial severity limit, primarily because these will be the highest doses investigated in any toxicity study and major adverse effects are anticipated in a small number of animals. The animals are observed regularly every day (particularly on the day of dosing) for any signs that indicate a deviation from normal appearance or behaviour. In addition, body weight is recorded regularly. There are ERP approved site guidelines for translating clinical observations and weight-loss effects into severity category.

The vast majority of animals are expected to experience moderate effects in these studies, because steps are taken to mitigate potential adverse effects usually by implementation of humane endpoints. A smaller number of animals will experience substantial effects, as the purpose of the study is to identify major adverse effects up to a limit of 2000 mg/kg. However, the number of animals experiencing substantial effects is kept to a minimum because the study design involves the dose range being investigated in one male and one female. In studies where compounds of anticipated low toxicity are used and where the limit dose only is given (2000 mg/kg) then the majority of the animals may only experience mild effects.

Severity classification

Actual adverse effects and the WG's codes are shown in the table below. In study 1, six animals had a maximum intensity of adverse effects of mild and eleven of moderate. Five animals had a maximum intensity of adverse effects of substantial over a time period of less than 24h. In study 2, all animals appeared normal throughout the 14 days:

Actual adverse effects in Study 1:	Severity coding	
	2-grid	Mild, Moderate, Substantial
Transient increased heart rate and slight pilo-erection. Observed on the day of dosing only. Animals appeared normal for remainder of the study. 6 animals.	BB	Mild
Irregular heart rate, irregular respiration, pilo-erection. Seen intermittently over 24 h. Animals appeared normal for remainder of study. 11 animals.	EB	Moderate (<i>lower</i>)
One animal died on day 1. Previous observations included pilo-erection, decreased activity, pale, irregular respiration, eyes half shut, cold.	GG	Substantial
2 animals found dead on day 2, no previous observations recorded the evening before.	GG	Substantial
2 animals found dead on day 2. Previous observations included pilo-erection and cold extremities.	GG	Substantial
Actual adverse effects in Study 2:		
All 10 animals appeared normal throughout the study	AB	Mild

Using the 2-grid scheme to code the substantial effects – death of animals: In this case, death occurred due to the toxic effects of a compound, but, because of the short time-scale involved, without detection of any preceding clinical signs in two cases. Basing application of the retrospective reporting system solely on clinical signs *actually observed* would result in codes GA and GD for the first and second groups of two animals 'found dead' in the highest dose group of this acute toxicity study. Members of the Working Group would, however, prefer to code all of these deaths as GG. This coding would assume that there was short-term substantial suffering prior to death that was not evident from the last set of clinical observations, and so would give the benefit of the doubt to the animal. A further argument for a GG classification is that any other coding is unacceptable because this study type has the potential to cause significant adverse effects.

This is an example in which the 2-grid scheme adds complexity that provokes disagreement about details of coding, where there is in fact agreement on the main point – that, overall, the effects are substantial.

3: Long term study of effects following multiple vaccinations

Study design:

Marmosets were used to investigate potential long-term (18 month) effects following multiple vaccinations and other pre-treatments, administered over a one-month period (4 groups of n = 12). Parameters measured included cognitive behaviour, EEG, sleep, immunology, endocrinology and muscle strength. EEG (and thus sleep) were measured by means of pre-implanted radiotelemetry devices. Cognition and muscle strength were measured on a daily basis using home cage testing techniques, whereby animals had a choice of whether or not to engage. Banana milkshake and chopped nuts were used as positive reinforcement for each test respectively. These home cage tests could be conceived as an environmental enrichment because all animals were keen to 'engage'. Blood samples were taken on a monthly basis, and animals were trained to provide early morning urine samples when required.

Effects on the animals:

The animals experienced the effects of implantation surgery, anaesthesia and recovery, and monthly blood sampling as well as the administration of vaccines and pretreatments. There were no discernible short or long-term clinical signs or other adverse effects, centrally or locally, as a result of the vaccinations. Five animals (out of 48) exhibited a skin reaction above the small area of bone cement which held the EEG electrodes to the skull and further anaesthesia was required for treatment.

Severity classification

2-grid	Mild, Moderate, Substantial
EC	<p><u>Majority of animals:</u> Moderate (lower)</p> <p><u>Animals reacting to cement:</u> Moderate (probably higher moderate)</p>

4: Vasectomy of male mice for generation of pseudopregnant recipient mice

Procedure:

Embryo transfer is used in the generation of genetically altered mice and for rederivation of microbiologically clean stock from contaminated mice. Vasectomised mice are used to mate with females to produce pseudopregnant recipients that will accept transferred embryos.

Mice are anaesthetised with isoflurane and a 5-7mm scrotal incision is made under aseptic surgical conditions. This allows access to the vas deferens via small incisions in the peritoneal layers, the vas is ligated with cautery then the incisions are closed at the peritoneal and skin levels. Mice are given a dose of buprenorphine for post-op analgesia. Recovery and subsequent days post-op are usually uneventful.

Mice are then retained for 12 months, caged individually but with females introduced overnight at least two days per week.

Effects on the animals:

Mice usually show no overt clinical signs following such surgery and behave normally within minutes of recovery from the anaesthesia. The suggested classification gives the animals the benefit of the doubt.

Severity classification

2-grid	Mild, Moderate, Substantial
DC	Moderate (lower)

5: **Plasmid electrotransfer into skeletal muscle of mice**

Procedure:

Gene transfer into skeletal muscle is a potential treatment for muscle diseases. Muscle can also be used as a platform for expression of secreted proteins. Plasmid DNA can be used as an efficient gene vector provided it is delivered in combination with electrical pulses. Hyaluronidase increases the efficiency and reduces damage associated with electrotransfer.

The procedure involves injection of a dilute solution of hyaluronidase (25 microlitres) into the anterior tibial muscle of a 12 week old male mouse. This is carried out under Hypnorm/Hypnovel anaesthesia. 2 hours later the mouse is re-anaesthetised with isoflurane, plasmid DNA is injected into the same muscle and calliper electrodes are applied to the skin and 10 twenty millisecond electrical pulses are given (175V/cm at 1Hz). Depending on the experiment the mouse may then be given an injection of buprenorphine as the analgesic.

Effects on the animals:

Recovery is rapid with only transient (couple of hours) evidence of any lameness. There are no clinical signs of lameness 24 hours after the procedure. Animals are generally killed 5 days later and the muscles removed for post-mortem analysis. In a few cases there is clear evidence of significant muscle damage on histological evaluation.

Severity classification

2-grid	Mild, Moderate, Substantial
AA, or DD if there is clear evidence of muscle damage <i>post mortem</i>	Mild or Moderate (<i>lower</i>)

6: **Effect of pharmaceuticals that influence mood and behaviour on brain responses to environmental stimuli in mice and rats**

Study design:

Under general anaesthesia, a small burr hole is made in the animal's skull and a tiny micro-dialysis probe is implanted in the limbic region of the brain. The animals recover overnight and then are challenged with novel, naturalistic environmental stimuli over a period of 2 to 5 hours – e.g. being placed in a new cage, being exposed to a bright light, or another (new) animal of the same species and strain. The chemical effects of these stimuli on the limbic region of the brain are monitored by sampling via the micro-dialysis probe.

By these means, the changes in brain chemistry of animals that receive appropriate doses of drugs used to treat anxiety and depression can be compared with control animals that do not receive the drugs. In some studies, the responses of genetically altered (knock-out) mice are also compared. These knock-out mice show no overt phenotypic changes, except when challenged with novel environmental stimuli, when they behave like mice given anti-depressant drugs.

The pharmaceuticals employed are usually well established for clinical use in humans, and have the potential to alter the emotional state of the animals in a 'positive' way – e.g. reducing anxiety, enhancing feelings of well-being.

This technique is sensitive enough to detect small chemical changes in the brain of rodents following experimental challenges that cause no, or minimal, discomfort. Following the successful development of this technique in animals, it is being used increasingly in humans in a range of medical contexts (e.g. during surgery and for slow, localised infusion of medicines).

Effects on the animals:

Effects of surgery under general anaesthesia to implant microdialysis probe.

Minor discomfort through administration of test compounds – but the drugs used do not induce adverse effects at the doses used in this study. Effects of changes in naturalistic environmental stimuli for 2 to 5 hours. Total duration of study < 48 hours.

Severity classification

Maximum severity is moderate, and at 'lower' end – confirmed by effects of the same procedure in humans.

N.B. 2-grid codes depend on the view taken on how long adverse effects of this surgery persist.

If they are considered to last less than 48 hours, the remaining severity will be A, because the stimuli involved are mild (naturalistic) and brief, and the drugs administered do not induce adverse effects. Codes will be DA or EA depending on whether the surgery effects are considered to persist for more than 24 hours.

If surgery effects are thought to persist for the whole 48 hours, the codes could be DD or ED.

Severity coding	
2-grid	Mild, Moderate, Substantial
DA, DD, EA or ED	Moderate (lower)

Burdens of reporting

The researcher involved reports that the time taken to carry out individual assessment of severity is minimal for each study of this size: five minutes for the two studies. However, if large numbers of such studies are conducted over the course of a year, recording and reporting severity could bring considerable burdens.

7: Protocol to study gene function in limb development (involving breeding, but not otherwise using, genetically altered mice)

This protocol covers the breeding of 15 lines of genetically altered mice used to study the genetic mechanisms involved in limb development. The mice so bred are killed by a Schedule 1 method at different stages (e.g. embryo, fetus, neonate etc), so that the effects of the genetic alterations on the developing limbs can be assessed *post mortem*.

Effects on the animals:

In most of the lines that are maintained as breeding colonies under this protocol, the genetic alterations cause no discernible phenotypic effect (but the animals require DNA analysis from tissue taken from ear biopsy). Even if biopsy is not performed, such breeding is itself a regulated procedure under the Animals (Scientific Procedures) Act 1986 and the number of animals involved must be reported to the Home Office as part of the annual Returns. The severity of the procedure would be classified as 'mild'.

In one of the 15 lines, the genetic change results in a limb defect, and, in another line, the development of smaller than normal ear pinnae. In these lines, ear biopsy is not required because the obvious phenotypes indicate the animals' genotypes.

Severity classification

In both cases, the severity of the impact on the animals, in terms of pain, suffering, distress or lasting harm, is uncertain – but, giving the animals the benefit of the doubt, the limb defect would be classified as 'moderate'.

On this basis, the severity categorisation for animals humanely killed at 4 weeks would be as follows:

	Severity coding	
	2-grid	Mild, Moderate, Substantial
For the phenotypically 'normal' lines killed at 4 weeks	CC	Mild
For the 'limb defect' line killed at 4 weeks	FF	Moderate (lower)
For the 'small ear pinna' line killed at 4 weeks	CC	Mild

Using the two-grid system, other codes would be required, according to the stage of development at which the animals are killed, to reflect the differences in 'duration' of the 'breeding procedures' involved.

This example raises the question of what weight should be accorded to 'harm in the absence of suffering', which might apply to the transgenic mice that have smaller than normal pinnae, with no obvious evidence of hearing defect, nor other overt signs of the genetic alteration. The WG's view is that it would be appropriate to code such effects as 'up to and including mild'

Comment / burdens

In this case, it can be argued that the two-grid intensity-duration classification system brings unnecessary complications when used to report the severity of procedures involving breeding, but not otherwise using, genetically altered animals. When animals are killed at different times, a large number of code combinations are possible, because of the different 'durations' involved. But does it make sense to distinguish 'duration' of these procedures in this way, particularly for lines in which no overt adverse effect is caused by the genetic alteration?

The protocol described above is relatively simple, in that the genetically altered animals are 'merely' bred and not used in other regulated procedures. It is noted that the number of possible severity code combinations is further increased for animals that are bred and then subsequently used in other regulated studies.

8: Mouse mutagenesis study of a degenerative disease

Study design:

This study aims to dissect the genetic influences on time of onset of a specific degenerative disease.

1. Male mice are injected with a chemical mutagen, which produces random mutations in their sperm. Approximately 50 different genes carry changes in each sperm.
2. The mutagenized males are mated with female mice that have been genetically altered to carry a gene for a late onset degenerative disease.
3. The resulting offspring are tested for presence of the transgene, by ear biopsy at 3 weeks of age. Because the mothers carry the transgene in heterozygous form, around 50% of the offspring will not bear the transgene and are therefore humanely killed.
4. The remaining offspring, which carry both the transgene and random mutations caused by the chemical treatment, are then screened for signs of onset of the degenerative disease, using a scoring scheme specific to the disease. The aim is to ascertain how the gene changes caused by the chemical treatment influence the time of onset of the degeneration (i.e. to determine the role of gene modifiers in the onset of the disease).
5. Mice showing signs of degeneration are assessed using a specifically designed welfare assessment scoring scheme. Mice approaching humane end points (as set out in the Project licence) are humanely killed at about 20 weeks of age (17 weeks after biopsy).

Severity classification

(i) Experience of progeny that do not bear the transgene and are humanely killed

These animals are produced under a 'moderate' severity limit protocol, but do not bear the gene for the degenerative disease. It can therefore be argued that the maximum suffering they experience is that caused by the ear biopsy, which, under current HO guidance, would be classified as 'mild', or CC in the two-grid scheme (because even in the absence of clinical signs, genetic manipulations are regulated procedures, and are regarded as at least *up to* mild severity, throughout the animal's life).

However, it might also be argued that, although the effects are unknown, the mutation load borne by these animals by virtue of the ENU treatment of their fathers, should push the severity classification into the 'moderate' category, or FF in the two-grid scheme – see summary overleaf.

	Severity coding	
	2-grid	Mild, Moderate, Substantial
Depending on severity of mutation load, either:	CC (or FF)	Mild (or Moderate (<i>lower</i>))

(ii) Experience of mice screened for neurodegenerative disease

The Working Group was presented with adverse effect data for a sample of ten mice screened for signs of degeneration. Of these, 8 out of 10 experienced mild effects (classified as described above) during the days after onset of the condition, and two experienced more moderate effects – one of which was humanely killed prior to exceeding its severity limit due to considerable weight loss over the week.

	Severity coding	
	2-grid	Mild, Moderate, Substantial
8 mice	CC	Mild
2 mice	FF	Moderate (<i>upper</i>)

9: Scoring studies to assess treatments of arthritic symptoms in miceStudy design

Collagen together with adjuvant is injected subcutaneously on a single occasion. This results in a model of arthritis 3 to 4 weeks later. Animals are dosed daily with compound or vehicle and the effect on inhibiting joint swelling and limping monitored for a further 2 to 3 weeks.

Protocol severity limit

This is “moderate” because of development of swollen joints and limping and possible side effect of small skin ulcers at the site of initial injection. Severity is limited by treatment of ulcers and humanely killing the animals if they lose mobility or their general health begins to decline irreversibly.

Monitoring and assessment of severity of adverse effects

Animals are examined daily and effects scored and recorded locally.

A 5 point scoring system is used for both arthritis and ulcers.

For joint swelling, the first signs of swelling in digits or ankle are “mild”; any greater swelling “moderate”. For ulcers anything up to 10mm² scab at the base of the tail is scored as “mild”; a larger scab or open wound, which occurs in < 7% animals, is scored as “moderate”.

Severity classification

Possible codes are as follows:

Severity coding	
2-grid	Mild, Moderate, Substantial
Eight codes possible:	Two codes possible:
AA; CB; CC	Mild
EB; EC; FC; FE; FF	Moderate*

*If sub-divided could be *upper* or *lower*, depending on degree and duration of clinical effects.

Burdens of two-grid scheme

If the proposed two-grid system was used the possible combinations and resulting severity scores would be: AA; CB; CC; EB; EC; FC; FE; FF. Because of the lack of correlation between “arthritis” and development of ulcers each animal has to be evaluated individually for maximum severity (moderate arthritis or ulceration), and then assessed for remaining severity, looking at the overlap of effects and the time each lasted at each severity level.

In this example it is difficult to predict the severity category in advance for each grid and use of the scheme is a fairly lengthy process due to fact that wanted and unwanted adverse effects are recorded separately for each individual (as part of the protocol).

The researcher involved reports that, using the two-grid scheme, it took over an hour to assign codes for maximum severity in a study involving 164 mice; and estimating remaining severity took at least a further 30 minutes. About 10 such studies are performed per year, each involving approximately 150 mice. Therefore it is estimated that if the two-grid scheme was to be adopted the extra regulatory burden over and above the current intensive scoring would be 15 hours per year to assess, plus 2 hours a year to Return these two-grid data to the Home Office. It did not take long to understand the scheme and be able to use it, so it is anticipated there will be little further reduction in the time required to assign scores in future studies.

10: Two regulatory toxicity studies

In these examples, reporting was 'by group', because within each group individual animals experienced adverse effects of similar severity (or lack of them). Had there been significant differences in effects on individuals within a given group, reporting would have had to be 'by individual'.

Study 1: Long-term study involving dogs

Dogs were dosed intravenously with a pharmaceutical compound at daily intervals, for 39 weeks. Control animals were dosed with sterile saline. It was anticipated that the compound would produce only mild effects in the animals.

Effects on the animals:

The severity of any clinical signs actually observed was categorised using Company-wide guidelines (evolved over a period of time through discussion and practice).

No adverse effects were observed in the low dose and control groups. Some animals in the intermediate dose group occasionally passed loose faeces. In the high dose group, dogs occasionally showed subdued behaviour, increased salivation and loose faeces.

The study design and resulting severity classifications were as follows:

Group	Dose	Number of Animals	Severity coding	
			2-grid	Mild, moderate, substantial
1	Control	8	CC	Mild
2	Low	8	CC	Mild
3	Intermediate	8	CC	Mild
4	High	8	CC	Mild

Note that, in this study, none of the severity classification schemes distinguishes the different effects experienced by control *cf.* dosed animals, nor by low *cf.* higher dose groups.

Study 2: Short-term cyclical dosing study involving rats

This study had a more complicated design, involving dosing rats intravenously for five days with a pharmaceutical compound. In groups 1a and 4a (see table below), the dosing was followed by a 6-week observation period, and in groups 1, 2, 3 and 4 by a 5-day observation period. Controls were dosed with sterile saline. It was anticipated that the compound would produce effects up to moderate severity.

Effects on the animals:

The severity of clinical signs observed was classified using the guidance referred to above. Rats in the low and intermediate dose groups showed very mild body weight losses and associated reduction in food consumption. Clinical signs were more marked in the high dose groups, and included subdued behaviour, tremors and laboured breathing. In group 4a, animals were humanely killed on a case-by-case basis to avoid suffering, at between 12 and 45 days into the study.

Severity classification

The study design and resulting severity classifications are shown in the table.

Group	Dose	Regime	Number of animals	Severity coding	
				2-grid	Mild, moderate, substantial
1	Control	5 day IV + 5 days observation	30	CB	Mild
1a	Control	5 day IV + 6 weeks observation	20	CC	Mild
2	Low	5 day IV + 5 days observation	50	CB	Mild
3	Intermediate	5 day IV + 5 days observation	50	CB	Mild
4	High	5 day IV + 5 days observation	64	EB	Moderate (lower)
4a	High	5 day IV + 6 weeks observation	20	FF	Moderate (higher)

Burdens

A researcher working in this field estimates that, using the 2-grid system, it would take around 20 minutes to record actual severity in a relatively straightforward toxicity study involving 80 to 100 animals. With upwards of 240 such studies conducted under a single licence per year, the total time required to record severity using the 2-grid system for just this one kind of toxicity test would be 80 hours (over 11 working days) annually.

11: Fish pair-breeding studyStudy design:

- Breeding pairs of fathead minnows were exposed to a range of concentrations of a pharmaceutical, dosed via the water, to determine the effects of the compound on the animals' reproductive behaviour, when compared with control fish. There were 5 treatments, plus two control groups (dilution water control and solvent control), with 16 fish (8 pairs/replicates) per treatment. The breeding pairs were exposed to the compound for 42 days and then humanely killed.
- Embryos from two spawnings in each of the five treatments, and from the controls, were grown up until 28 days post hatch and then humanely killed (n= 50 eggs per treatment).

In addition, hatchability trials were conducted on two spawnings from each replicate. However, the number of eggs and hatched larvae were not reported as part of the annual statistical Returns, because the animals were humanely killed before they became capable of independent feeding, and so were not protected under the Animals (Scientific Procedures) Act.

Effects on the animals:

In this study, as in the majority of work with small fish, it is impossible to identify each individual within a treatment tank and so effects cannot be reliably followed in individual fish over the duration of a study. Rather, adverse effects observed at any given time are 'whole tank' assessments. At the end of a study, each fish is humanely killed and examined, and at this stage is given a unique identification number.

In this study the adverse effects were as follows:

1. Effects on adult breeding pairs (F0 generation)

At the end of the study, no effects were observed on mean wet weight, standard length and vitellogenin (egg yolk precursor) of fish exposed to the compound, compared with solvent control fish. In the top concentration there was a significant reduction in mean egg production when compared to the solvent control.

2. Effects on developing embryos (F1 generation)

As naturally occurs in fish breeding, some hatched embryos in both treated and control groups died of natural causes before reaching maturity. No significant differences were observed in hatching success, or mortality during the study in treated *cf.* solvent controls. At 28 days post hatch, mean wet weight and standard length were significantly reduced in fish which had been exposed to the top two concentrations – but there were no other effects or signs of suffering.

[Hatchability trials (not regulated under ASPA): No effect on hatchability was observed.]

Severity classification

Study	Severity coding	
	2-grid	Mild, moderate, substantial
1 – Adults (F0)	CC	Mild
2 – Developing embryos (F1)	CB	Mild

12: Evaluation of rodenticides, particularly in the context of meeting UK/EU/EPA legislative requirements

Study design:

Two sighting (dose ranging) tests were carried out in preparation for a full acute toxicity evaluation of a novel rodenticide. Two different formulations of the same active ingredient were tested. Each candidate formulation was administered to singly-caged Wistar rats by oral intubation and each animal was observed at least twice daily during the period when symptoms of poisoning were expected to appear. In total, 24 rats were tested.

Effects on the animals:

For formulation 1, 14 rats were dosed and observed for 17 days. Given the mode of action of the novel rodenticide, after 8 days it was deemed that this particular formulation was insufficiently toxic for practical rodent control. Symptoms (e.g. lethargy) for those animals given the highest doses appeared to be mild throughout the procedure, with rats on the lowest doses showing no symptoms at all.

For formulation 2, 10 rats were dosed and observed for 13 days. Two rats became severely ill after 6 days, but up to that point the symptoms had appeared to be mild. After another 2 days, one rat died and the other began to recover. The remaining 8 rats showed very mild, if any, symptoms of poisoning.

Severity classification

	Severity coding	
	2-grid	Mild, Moderate, Substantial
Formulation 1 – all animals / Formulation 2 – eight animals	AB	Mild
Formulation 2 – two animals	HB	Substantial

Researcher's comment on use of the 2-grid system

Using the two grid system in this example, it can be argued that the information contained in the code HB is potentially misleading, because the outcomes were substantially different for the two rats.

The rodenticide in this case failed to perform as expected due to bioavailability problems, which the manufacturer is hoping to correct. Further testing with a more effective formulation will likely lead to proportionally more animals being classified HB under test conditions, particularly, given the mode of action, if the dose-response curve is steep.

Rodenticide studies involving wild animals

This test was carried out on laboratory-strain rats. If a successful formulation emerges and an optimum concentration can be found, subsequent tests will be carried out on wild rats, in which it will be much more difficult to observe symptom intensity even in singly-caged animals.

The final stages of a rodenticide evaluation involve tests on colonies of wild rats under semi-natural conditions (e.g. arena trials) and on free-living populations. Arena trials do not allow close inspection of individual animals and in some circumstances particular individuals may be seen rarely once a trial starts (semi-natural conditions mean that plenty of cover is supplied for rodents to hide in). In field trials of rodenticides against rats or mice, efficacy is determined indirectly through activity measures (e.g. tracking plates) rather than by direct observation of animals. Indeed, sometimes, rats or mice are never seen at all.

APPENDIX 4:

Why Europe should adopt retrospective reporting of scientific procedures on animals:

A LASA proposal

1. Current methods of counting animal procedures in Europe

At present, there is no standard method of reporting statistical data on animal use across Europe:

- some countries, e.g. the UK, report animal procedures when they are started ('prospective' reporting);
- some, e.g. The Netherlands, report procedures when they are finished ('retrospective' reporting);
- whereas Switzerland reports animals 'in use' during a given year.

Prospective reporting means that animals are counted in the *Statistics* in the year in which they enter procedures, whereas retrospective reporting means that animals used in procedures that cross year-end(s) will not be counted until the subsequent (or later) year(s). However, if this thought to be a problem, it can be overcome – see 3iii below.

2. Disadvantages of prospective reporting

- (i) Reporting procedures started in a given year does not always reflect the eventual use of the animals.

This can be a particular problem with GA animals, which in the UK make up one third of the animals used. With prospective reporting all GA animals produced in a year must be recorded but often it is not certain whether they will subsequently be used for breeding, for samples for post-mortem analysis or subject to additional procedures. If used for a procedure in the subsequent year they are likely to end up being counted again, leading to inflated statistics.

- (ii) Reporting when procedures are started will not permit collection of data on the actual experience of animals.

Note that 'prospective' reporting of the use of animals according to the severity limits* of procedures into which they are entered would greatly over-estimate actual severity of any suffering. For example, animals may be entered into a procedure which has a substantial upper limit yet in practice very few may reach that limit and many may undergo no more than mild pain, suffering, distress or lasting harm.

3. Advantages of retrospective reporting

- (i) Reporting animal use when procedures are finished provides more accurate data on how animals are actually used (cf. (i) above – prospective reporting does not always reflect eventual use of animals).
- (ii) Retrospective reporting also enables data to be collected on the actual experience of the animal undergoing the procedure, including the severity of any suffering experienced and the duration of the procedure.

Reporting the actual effects of procedures on animals would bring a number of benefits, including:

- *Increased transparency* – providing public information on the severity of animal experiments;
- *Promotion of refinement* – helping to identify opportunities for refinement and, on a year-to-year basis, allowing for assessment of the impact of refinements, in terms of changes in the experience of animals undergoing the modified procedures – and so documenting good welfare practice.

(iii) Retrospective reporting allows identification of longer-term procedures lasting more than a year. Although, as noted above, the use of some animals will not be reported in the year in which the procedures are started, this difficulty can be overcome (if necessary) by also reporting the number of procedures started but not finished in a given year (as in the Swiss system).

4. A mixed prospective and retrospective reporting system is not viable

If there is to be retrospective reporting, it should be the sole method, used 'across the board' in order to:

- (i) enable cross-referencing of data on different aspects of animal use (e.g. purpose vs. severity); and
- (ii) avoid the additional bureaucratic burden of collecting two different sets of data – one at the start and one at the end of procedures.

Selecting a sub-set of procedures for retrospective reporting of severity in a given year (alongside an otherwise prospective system of reporting) would not reduce the burden of mixed reporting, because systems would need to be in place to collect severity information should it be required. Moreover, reporting the severity of only some categories of procedure would not enable any overall estimate of the welfare 'costs' to animals of experimentation, and could lead to public misinformation.

* That is, an upper limit to the suffering an animal can be caused in a given protocol – set at the licensing stage

APPENDIX 5:

FELASA statement on a LASA proposal for the European Community to adopt retrospective reporting of scientific procedures on animals

FELASA is composed of independent European national and regional laboratory animal science associations and was established by them in 1978. It can speak for laboratory animal scientists and technologists in more than twenty European countries and has Observer status at relevant EC fora.

As a science-based association, FELASA considers it has a responsibility to comment on any issue or change that may affect the use and welfare of animals in experimental or other scientific procedures. FELASA has therefore been actively involved in the ongoing discussions and consultations on the revision of the EU Directive EC86/609, which determines the minimum requirements for the regulation of animal experimentation by member states of the EU.

LASA proposal on retrospective reporting of scientific procedures on animals

The UK's Laboratory Animal Science Association (LASA), which is a constituent member association of FELASA, has proposed to the European Commission that the statistical reporting of scientific procedures carried out on animals should be performed retrospectively. LASA considers that this would provide more accurate data, increase transparency, and help promote refinement. A copy of LASA's full proposal is attached as Annex A (*here, Appendix 5, above*).

FELASA position on the proposal

1. FELASA has considered LASA's proposal and in principle supports the concept of retrospective reporting for the reasons detailed by LASA.
2. FELASA specifically supports the principle of retrospective reporting of severity and considers it can provide more meaningful data on the distress or pain that may be suffered by animals during scientific procedures.
3. Adoption of retrospective reporting should be in such a way that it does not significantly increase the overall administrative burden placed on research facilities and should be in place of, and not additional to, any other method of reporting, e.g. prospective.
4. Any system of retrospective reporting must be straightforward to implement and an appropriate lead-in/adaptation time should be allowed so as to enable those involved with reporting to become familiar with the new processes involved.
5. The data produced by retrospective reporting should be simple enough to be readily understandable by the general public.

FELASA, as a scientific association representing those involved with animal experimentation throughout Europe and with access to a wide range of expertise, will be happy to assist the European Commission by advising further on this or any other issue relevant to the revision of Directive EC86/609 and the conduct of animal experimentation within Europe.

FELASA October 2007

APPENDIX 6:

Suggestions for reducing the burden of reporting data on animal procedures

A LASA/APC Working Group (currently considering methods of reporting data on the severity of scientific procedures on animals) has recommended a series of changes that could help to streamline the Returns process as a whole.

The WG's suggestions take into account the continuing need to gather certain categories of data in order to meet European reporting requirements. They are offered below as a contribution to the current Home Office review of statistical reporting requirements within the Animals Scientific Procedures Better Regulation Action Plan, which will implement measures outlined in the recent Davidson Review.¹⁶

Row	Row Title	Recommendations
1	Species	<ul style="list-style-type: none"> • Retain
2	Species – CITES	<ul style="list-style-type: none"> • Remove – use practice similar to that currently used for rodenticide testing – e.g. state on the Returns form 'please give extra information if you have used a CITES species'
3	Stage of Development	<ul style="list-style-type: none"> • Remove and collect data on adults only, since data on developmental stages are not collated/reported in the <i>Statistics</i>
4	Genetic Status	<ul style="list-style-type: none"> • Retain – but simplify for GM/mutant animals. Genetic status categories are outdated and should be combined under the single heading GA; with the present arrangement it is easy to get the categories wrong – e.g. how to code the offspring of a nude mouse (mutant)-transgenic (GM) mouse cross?
5	Source of Animals	<ul style="list-style-type: none"> • Retain
6	Anaesthesia	<ul style="list-style-type: none"> • The WG considered possibility of removing this row because it might be explained via severity reporting. However, we note this category must be retained owing to its special status under the Animals (Scientific Procedures) Act, and in any case these data might usefully be linked to severity information
7	Neuromuscular Blocking Agents	<ul style="list-style-type: none"> • Remove – what is the value of these data?
8	Primary Purpose of the Procedure	<ul style="list-style-type: none"> • Retain – consider revising the list in a way that links primary purpose and body system
9	Body System	<ul style="list-style-type: none"> • Remove – difficult to categorise work in this way, therefore lots of 'bad' data collected; majority of procedures are 'system not relevant', therefore this row brings little added value; and Row 8 (primary purpose) is in any case retained<i>continued overleaf</i>

¹⁶ Davidson Review of the Implementation of EU Legislation (2006), page 31 point 3(d). http://www.hm-treasury.gov.uk/media/E/F/davidson_review281106.pdf

List A: Toxicology or other safety or efficacy evaluation		
A: Row 10	Purpose of evaluation	<ul style="list-style-type: none"> Retain – with minimal simplification and possible addition: delete A07, A08 and A22 (substances not tested); consider separating out vaccine testing under pharmaceutical evaluation – (an important cross reference for any severity data)
A: Row 11	Type of test/procedure	<ul style="list-style-type: none"> Retain but overhaul, simplify and consider adding. OECD categories over-stated and sometimes outdated; compare with EU requirements. Collect data on reasons for lethal tests that are performed (e.g. shell-fish toxin testing; botox)
A: Row 12	Legislative requirements	<ul style="list-style-type: none"> Retain. Delete A92 and A93 which seem redundant
List B: Procedures other than toxicology		
B: Row 10	Primary field of research	<ul style="list-style-type: none"> Simplify or remove – but may need to keep if Row 9 (body system) is removed
B: Row 11	Production and Breeding	<ul style="list-style-type: none"> Retain, but simplify for GM/mutant animals, as described for Row 3 (genetic status)
B: Row 12	Techniques of particular interest	<ul style="list-style-type: none"> Retain. Clarify inhalation. Fill in B00 by default
Row 13	Number of Procedures carried out on animals	<ul style="list-style-type: none"> Could delete. If retained, follow one of the two options listed below. Data on animals used plus data on re-use are sufficient to calculate numbers of procedures; therefore suggest removing this row and adding data on re-use as a table. This will avoid 'duplication' of many tables in the published <i>Statistics</i>. However, the WG is mindful that removing this row may not be possible because the aim of the <i>Statistics</i> is to record animal <i>procedures</i> – if so, one of the two options below could be followed
Row 14	Number of animals used for the first time	<ul style="list-style-type: none"> Retain; but delete 15 if 13 and 14 are retained
Row 15	Number of animals re-used for the first time in the current year	<ul style="list-style-type: none"> Retain; but delete 14 if 13 and 15 are retained

A further, more general recommendation is that the Returns be re-designed as an on-line flow chart, which would enable licence holders to select and show in the code list only those categories that relate to their own procedures – and so facilitate the reporting process.