The Old School, Brewhouse Hill, Wheathampstead,
Hertfordshire AL4 8AN, UK

Animal Welfare 2011, 20: 281-301
ISSN 0962-7286

Classifying the severity of scientific animal use: a review of international systems

N Fenwick[†], E Ormandy[†‡], C Gauthier[†] and G Griffin*[†]

[†] Canadian Council on Animal Care (CCAC), Suite 1510-130 Albert St, Ottawa ON, K1P 5G4, Canada
[‡] University of British Columbia, Animal Welfare Program, 2357 Main Mall, Vancouver, BC, Canada V6T 1Z4
* Contact for correspondence and requests for reprints: ggriffin@ccac.ca

Abstract

Severity classification systems (ie pain scales, categories of invasiveness, degrees of severity etc) are used to classify the adverse effects experienced by animals used for scientific purposes. Currently, eleven countries use severity classification systems. These systems have developed in various ways, depending on each country's process for overseeing the use of animals in science, as well as the particular aspects emphasised by those individuals who have championed their implementation. Severity classification serves four main purposes: as a tool to assist animal ethics committees in ethical review; education of animal users about concepts for humane animal experimentation; provision of data to inform the public about scientific animal use; and provision of data to inform national policies. At a time when the newly accepted European Union Directive will make the reporting of severity data mandatory, we review the characteristics of international severity classification systems and how they have evolved; analyse the effectiveness of some systems; and identify emerging challenges for severity classification.

Keywords: *animal ethics committees, animal use statistics, animal welfare, ethical review, severity classification, Three Rs*

Introduction

Severity classification systems — typically going from least invasive, or harmful, to most invasive — are used to describe and sort the various adverse effects that animals may experience when they are used for scientific purposes. These may include (but are not limited to): discomfort, pain, distress, fear, nutritional deprivation, and behavioural deprivation. Classification systems are referred to internationally by various terms, including: pain scales, categories of invasiveness, degrees of severity, harm scales, and domains of animal welfare compromise. Here, we have chosen to use the term 'severity classification' as inclusive of the characteristics of various systems.

Various uses for severity classification have been put forward, but four main purposes have been identified for nationally legislated (or mandated) systems (eg Orlans 1990; Griffin *et al* 2007): i) a practical tool to assist animal ethics committees (AECs; this will be used to refer to animal ethics review committees, animal care committees, institutional animal care and use committees, etc) in ethical review of animal use protocols; ii) the education of animal users about the concepts of humane animal experimentation (ie the Three Rs tenet of Replacement, Reduction and Refinement; Russell & Burch 1959); iii) the provision of data to inform the public on the numbers of animals that may experience each level of severity; and iv) the provision of data that will inform the development of national policies on the use of animals in science.

Additional uses for severity classification systems have been proposed, including: measurement of refinement progress (Orlans 2000); prioritisation of the development of Three R's alternatives (Smyth 1978; Orlans 2000); screening of research funding proposals (Shapiro & Field 1988); improvement of communication between investigators and animal care staff with "'rules-of-thumb' upper limits for the impacts that can be caused in particular procedures" (Smith *et al* 2005; p 19); and provision of a tool for monitoring compliance (EU 2010). However, this review will focus on whether mandated, national severity classification is successfully fulfilling the four 'main' purposes outlined above, and how this affects animals used in science. To achieve this, we report on the characteristics of international severity classification systems and how they have evolved, analyse the effectiveness of some systems, and identify emerging challenges for severity classification.

Characteristics of international severity classification systems and how they evolved

To our knowledge, eleven countries currently have mandated, national severity classification: Australia, Canada, Finland, Germany, The Republic of Ireland, The Netherlands, New Zealand, Poland, Sweden, Switzerland, and the United Kingdom (UK). This number will increase

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

with the adoption of the new European Union (EU) Directive 2010/63 which requires severity classification for its 27 member states (EU 2010). The new Directive will take full effect from January 1, 2013. The United States (US) also has a mandated national scientific animal use classification system, but it is markedly different from the rest of the countries that we report on. It does not classify severity directly, but instead reports on whether pain-relieving drugs were required and/or used (USDA 2000). Reviews of this system are provided in Stephens *et al* (1998) and Orlans (2000).

There are some general similarities in the way severity classification is used by each of the eleven countries, even though processes for ethical review of animal-based studies vary. When preparing animal-based protocols, investigators typically forecast potential adverse effects that animals may experience during their experiment. This method of pre-assigning severity, known as 'prospective assignment', is based on a precautionary approach that considers the maximum level of pain and distress animals might experience. The potential adverse effects are classified by reviewing the types of experimental procedures that will be performed; in some cases, outcomes to animals are also considered. Investigators then submit protocols to an AEC for approval to proceed with the proposed research. Although led by precedents (ie previous classification of similar procedures) and guidance documents (ie criteria for a particular category), assignment to a severity class typically requires professional and ethical judgments to be made by investigators, veterinarians and AEC members. In all eleven countries, it is implicit that stronger justification is required for scientific use of animals at higher levels of severity. In some jurisdictions, once the experiment is completed, investigators are additionally required to review the accuracy of the severity level, a process termed 'retrospective assignment'. Whether collected prospectively or retrospectively, the numbers of animals assigned to each severity level are then reported to the regulating body, usually on an annual basis. Currently, of the eleven countries, only six publish the numbers of animals that are assigned annually to each category (Table 1); however, with the approval of the EU Directive, the number of countries publishing severity data will greatly increase. For comparison of the characteristics of the severity classification systems of the eleven countries and the EU Directive, refer to Table 1. For complete descriptions of each system refer to Appendix 1.

The origin of severity classification

Where did the idea for severity classification come from? The catalyst appears to have been the 1978 publication *Alternatives to Animal Experiments* by David H Smyth, then president of the UK Research Defence Society. He proposed categorising experimental procedures and collecting national statistics on animal experiments as a way to prioritise efforts for the development of Three R's alternatives. Smyth argued that attention should be focused on procedures which cause the most pain and distress and devised a system to classify uses of animals based on types of experimental procedures.

Smyth's novel classification system introduced the idea of using classification to capture implementation of the Replacement principle and originated the idea of classifying non-recovery use of an animal separately (Table 1). Only two categories in Smyth's system included procedures that cause pain to conscious animals, and they grouped all degrees of pain, from minor to very severe, together. This was a limitation and most subsequent severity classification systems have made attempts to tease apart the degree and/or the duration of pain and distress experienced by animals used in science (Table 1).

Although Smyth's scale appears to have been the first of its kind in the English-language scientific literature, a similar idea was not only proposed but also legislated by both The Netherlands and Sweden. It has been suggested that these developments were linked to Smyth (eg Orlans 1993), however, there were other factors. In Sweden, the 1979 creation of a five-category system arose out of the desire to make the work of newly legislated AECs "reasonably uniform" (Obrink 1982; p 56) rather than to prioritise development of Three R's alternatives. The Netherlands also mandated severity classification in 1979 with three categories of 'discomfort' and sub-categories to account for duration of harm (Orlans 2000). In addition, neither The Netherlands nor Sweden implemented a category for non-recovery procedures (as Smyth had proposed). The Swedish and Dutch systems introduced two additional features that are now fundamental to severity classification: they were designed for use as part of the ethical review process, and they divide animal use by pain levels. These features are reflected in current international systems, although the number of severity levels range from three (Finland and Sweden) to nine (Australia) (Table 1), and the names and/or descriptions provided for each category are unique for each system (refer to Appendix 1).

In North America, Orlans built on Smyth's ideas: in the early 1980s she began writing about classification of severity to assist in regulating animal use for teaching purposes (Orlans 1980). Orlans suggested that while all levels of use may be appropriate for professional scientific research, only lower levels (ie non-pain-inflicting) are appropriate for pre-college student use. This novel application of severity classification was not incorporated into a national severity system until 2005, when Poland mandated that educational use of animals could only occur at the lowest severity category (Smith *et al* 2005).

The US began establishing institution-based AECs in the late 1970s (Orlans 1993). Then, in 1982, Karl Obrink, a Swedish professor, presented an overview of Swedish law on laboratory animals to the Scientists' Center for Animal Welfare (SCAW), an association that aims to improve animal welfare outcomes for animals used in science (Orlans 1993). As a result, SCAW created its own severity classification system with the view that it could be used (voluntarily) as a tool for reviewing protocols by US AECs

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Table 1 Characteristics of international severity classification systems.

Country (year of adoption)	Number of categories	Accounts for duration of severity	Contains category for non-recovery use	Contains category for unacceptable procedures	Contains category for relative replacement use	Data reported prospectively (P) or retrospectively (R)	Publishes numbers of animals per impact category
Australia (1994)	9	No	Yes	No	No	P	Yes
Canada (1987)	5	Yes	No	Yes	Yes	P	Yes
Finland (1986)	3	No	No	No	No	P	Yes
Germany (1995)	4	Yes	No	No	No	P	No
Ireland (2006)	4	Yes	No	No	No	P	No
The Netherlands (1979)	6	Yes	No	No	No	R	Yes
New Zealand (1988)	5	Yes	No	No	No	R	Yes
Poland (2006)	5	Yes	No	Yes	No	P	No
Sweden (2002)	3	Yes	No	No	No	P	No
Switzerland (1994)	4	Yes	No	No	No	R	Yes
UK (1986)	4	No	Yes	No	No	P	No
EU (2010)	4	Yes	Yes	No	No	R	Yes

Table references: Australia (New South Wales 2009); Canada (CCAC 1991); Finland (Orlans 2000; Purves 2000); Germany (Purves 2000; Smith *et al* 2005); Ireland (Department of Health and Children 2006); The Netherlands (Orlans 2000; Smith *et al* 2005); New Zealand (Bayvel *et al* 2007; Animal Welfare Directorate 2008); Poland (Ministry of Science and Higher Education 2006); Sweden (Rony Kalman *et al* in preparation); Switzerland (Swiss Federal Veterinary Office 2006, undated); United Kingdom (Home Office 2000; APC 2008); and European Union (EC 2010).

(Orlans 1987). Similar to Smyth's scale, the SCAW system included a level for animal use that may be described as 'relative replacement': replacing more sentient animals, such as vertebrates, with animals that current scientific evidence indicates have a significantly lower potential for pain perception, such as some invertebrates (CCAC 2009). This system also included a category to describe procedures that are typically unacceptable regardless of their scientific merit (Table 1). The United States Department of Agriculture (USDA) proposed including the SCAW severity classification system in national scientific animal use policy but this was rejected by US legislators (Orlans 2000).

Further developments in the use of severity classification were occurring worldwide. In 1984, The Netherlands became the first country to publicly report the data collected from its severity classification system (Orlans 2000). This important development marked the beginning of using severity data to contribute to public accountability in science. By the end of the 1980s, five additional countries had adopted some form of severity classification: Switzerland, Finland, UK, Canada and New Zealand (Table 1).

In 1986, the UK legislated the use of a two-tiered, four-level system to classify severity. The descriptive names of the categories came from the Verbal Rating Scale (VRS) used in human medicine (LASA 1990) which describes increasing pain intensities as 'no pain, mild pain, moderate pain and severe pain' (Williamson & Hoggart 2005). The UK severity classification labels are 'unclassified', 'mild severity', 'moderate severity' and 'substantial severity' (Home Office 2000). The two tiers of severity classification — protocols and projects — was (and is) unique. The first tier (protocols) assigns 'severity limits' to individual scientific protocols, reflecting the upper limit of negative effects from procedures in that protocol. The second tier (projects) assigns 'severity bands' to scientific projects, typically composed of a group of related protocols (Home Office 2000). The severity bands are intended to reflect the severity experienced by the 'average' animal being used in a project (ie a suite of

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

protocols); however, this approach has been criticised because the number of protocols within each severity band is published but the number of animals experiencing the differing levels of severity is not (APC 2003).

Outcomes to animals in severity classification

The late 1980s marked the start of including specific outcomes to animals, in addition to types of procedures, in the classification of severity. This idea had appeared in previous literature (eg Ross 1981) but occurred first in national policy in 1987 when Canada developed a peer-reviewed severity classification system based on the SCAW system. The Canadian system differed from SCAW in two main ways. First, it included a wider category for experiments that involve moderate to severe pain and distress (referred to as category of invasiveness [CI]-D) due to criticism from Canadian investigators that the SCAW categories jumped too quickly from minor stress to significant but unavoidable stress (CCAC 1987). Second, the Canadian system also included (and still includes) examples of outcomes to animals that should not occur. For example, animals undergoing CI-C procedures (mild to moderate pain and distress) "must not show self-mutilation, anorexia, dehydration, hyperactivity, increased recumbency or dormancy, increased vocalization, aggressive-defensive behavior or demonstrate social withdrawal and self-isolation" (CCAC 1991).

By 1988, New Zealand had incorporated elements of both the Swedish and the SCAW severity classification systems into guidelines for AECs (Orlans 1993). Then, in 1997, New Zealand became the second country to include outcomes to animals when it introduced severity classification for use in conjunction with the existing guidelines (Williams *et al* 2006). Based on Mellor and Reid's (1994) novel 'domains of animal welfare compromise', the New Zealand system requires prediction of the severity of experiments with reference to the domains of welfare compromise. The five domains are presently described as follows: i) Nutrition — water deprivation, food deprivation, malnutrition; ii) Environment — challenging outdoor and indoor conditions; iii) Health — disease, injury, functional impairment; iv) Behaviour — individual, group or interactive restrictions; and v) Mental state — unpleasant or noxious experiences (Animal Welfare Directorate 2008). Investigators use a non-numerical grade (A, B, C, D or E) to assess the degrees of compromise within each domain. For example, "grades A and B represent no welfare compromise to low-level tolerable compromise, [while] grade E represents compromise likely to cause very severe suffering" (Mellor *et al* 2009; p 82). The use of a non-numerical score is an important distinction from severity classification that uses numerical scores which are added together to determine overall severity (eg Porter 1992) and more in keeping with the widely accepted use of professional and ethical judgment in classifying severity.

Severity classification today

Severity classification systems have continued to evolve. Australia began mandating severity classification in 1994 with Germany following in 1995. In 2002, Sweden, which had ceased to use its original system in 1989 (Orlans 2000), reintroduced a new classification system. In 2005, Poland first introduced a five-level system modified from SCAW and Mellor and Reid (1994), and, as mentioned above, uniquely specified that only the lowest level of severity is permissible for education purposes. The present version of the Polish system (revised in 2006) describes each severity level with reference to outcomes to animals, similar to the Canadian and New Zealand systems.

In 2006, The Republic of Ireland began using severity classification, and from January 1, 2013, the new EU Directive will require the inclusion of severity classification in the animal statistics report required from member states. Finally, it is important to note that although not all countries mandate severity classification, in some it is used voluntarily by institutions to educate animal users and/or as part of the ethical review process. For example, many US institutions voluntarily use the SCAW system, and SCAW is also recommended to Japanese institutions by the Science Council of Japan (Kuhara 2008). Similarly, some AECs in Brazil have begun to use three-level severity classification (mild, moderate and substantial) (ATP Filipecki, personal communication 2010) and one research institution in Brazil voluntarily uses the Canadian severity classification system (Silla *et al* 2009).

In spite of the many differences between systems, no large discrepancies in the classification of procedures have developed. For example, withdrawal of blood is consistently considered low severity and experiments with death as an endpoint are consistently considered high severity (refer to Appendix 1). However, the variety of severity classification systems invites the question of which systems, or which characteristics, are most effective.

Effectiveness of severity classification

As stated, severity classification serves four main purposes: i) a practical tool to assist AECs in ethical review; ii) the education of animal users about concepts for humane animal experimentation; iii) the provision of data to inform the public about scientific animal use; and iv) the provision of data to inform national policies. Are these goals achieved? The following sections examine the effectiveness of severity classification in Canada, New Zealand and the UK.

Does severity classification provide a practical tool to assist AECs in ethical review?

Many severity classification systems appear to have originated as practical tools to assist in ethical review of protocols (eg Sweden). Severity classification is used by Canadian AECs as an administrative tool to pre-sort protocols and to ensure that the most invasive ones receive additional scrutiny (Griffin *et al* 2007). A study of the role of Canadian AECs in humane animal experimentation found that AEC

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

members use severity classification: i) to determine whether the benefits need to be greater (Schuppli 2004); ii) to determine how closely animal numbers should be scrutinised (eg more effort is directed at reducing animal numbers in protocols at higher levels of invasiveness) (Schuppli & Fraser 2005); and iii) to 'test' the accuracy of the protocols (eg do the procedures outlined match the severity classification selected by the investigator?) (Schuppli 2004). In addition, Schuppli (2004; p 90) found that:

> …some [AEC] members viewed inconsistencies in the way that the application form was filled out as reason for a negative recommendation because investigators were not portraying information accurately. Inconsistencies included whether the listed drugs matched the drugs described in procedures or whether the category of invasiveness [severity level] matched what was expected from procedures.

There is ambiguity surrounding the acceptability of procedures at the highest severity level (known as CI-E): some Canadian institutions prohibit any CI-E procedures, while others permit highly invasive tests in order to fulfill legislated requirements (ie safety and toxicology studies). In these instances, the CI-E severity classification identifies animals that may experience severe adverse effects and signals the need for extra vigilance from AECs and animal care personnel.

A UK study found that severity classification assisted AECs in defining the upper limits of suffering, identifying humane endpoints, and prioritising targets for Three R's interventions (ie identifying procedures that have the greatest adverse effect on animals) (Smith & Jennings 2004). However, this and another study found the effectiveness to be limited by several factors. These included: the lack of adequate descriptions to define each category; the lack of sufficient examples of procedures per category; and the absence of worked examples to explain the process of severity classification in supporting guidance documents (Smith & Jennings 2004; Smith *et al* 2005). Smith and Jennings (2004) also noted that classifying severity was made more difficult when the outcomes from procedures are uncertain. In addition, UK severity bands were identified as having 'very limited value' because it is difficult to provide a single assessment of a project that includes protocols of varying severities (Smith & Jennings 2004). These studies concluded that severity classification could be improved by increasing the focus on individual animals and by assessing outcomes to animals (Smith & Jennings 2004; Smith *et al* 2005).

These Canadian and UK findings suggest that AECs do use prospectively assigned severity classification as a tool in ethical review. Greater clarity in descriptions of severity levels would remove some ambiguities and assist AECs in determining outcomes to animals and opportunities for refinement. No published study on inter-rater reliability of severity classification using a national system is available; however, a theoretical severity classification system reported an inter-rater reliability score of 0.80 with trained raters (ie individuals using the system to classify severity agreed 80% of the time) (Shapiro & Field 1988).

Are severity classification systems educating animal users about concepts for humane animal experimentation?

Obtaining empirical evidence that severity classification is educating animal users is complicated by the difficulty of separating the educational effect of severity classification from the entire ethical review process. However, there is some information available that may help assess whether the process of classifying severity educates animal users about humane experimentation. In Canada, classifying severity is perceived by regulators to help educate, sensitise and alert investigators and AECs to the degree of pain and distress that experimental procedures cause animals. In addition, the existence of a category for procedures that do not use vertebrate animals or cephalopods is thought to convey the concept of relative replacement (as per Orlans 1990). Similarly, the highest severity level is used to convey the concept of procedures that are considered highly questionable (and thus requiring greater justification) or that are unacceptable regardless of scientific merit (CCAC 1997a).

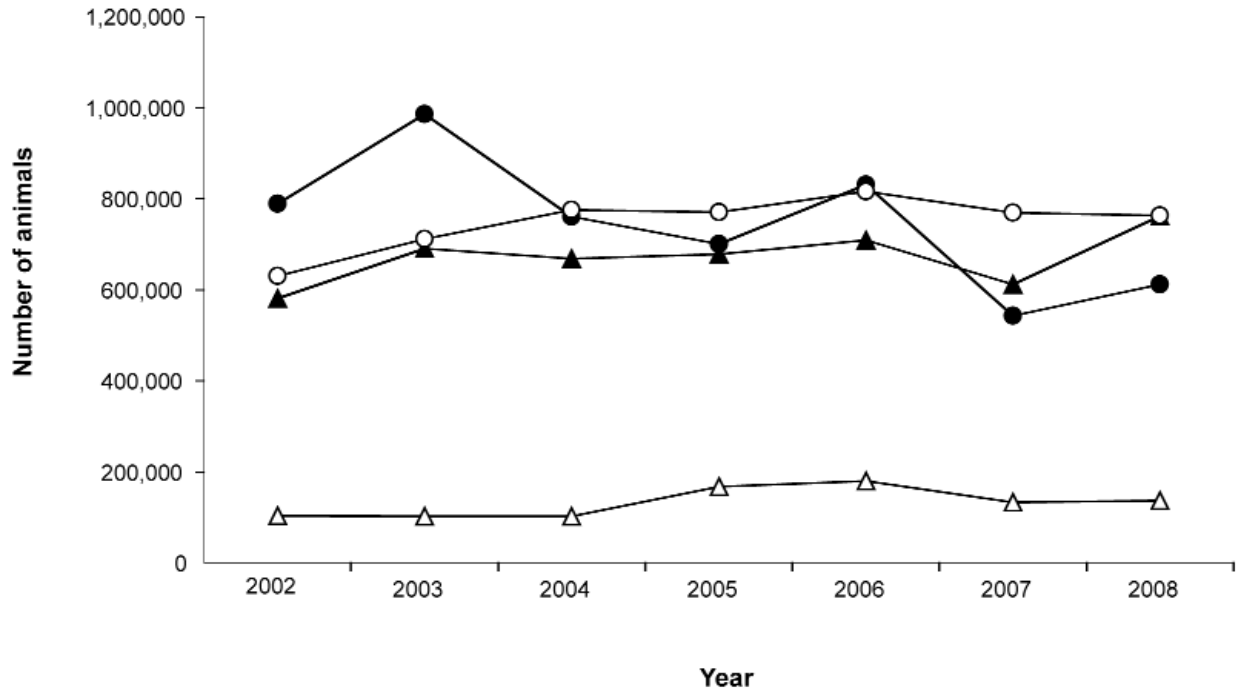
Bowd (1997) studied the effectiveness of one Canadian AEC in educating investigators and identified several inconsistencies between the beliefs of the investigators and nationally endorsed ethical positions and policies. Although not specifically analysing the role of severity classification, the study found that there was a need for greater education of investigators on the ethical principles behind guidelines and policies. This was also identified by an internal survey of the recommendations made to Canadian institutions by peer-assessment teams and resulted in publication of a policy training guide for investigators (CCAC 1999).

In New Zealand, severity classification is perceived to assist investigators in evaluating the harm that may be done to animals during an experimental procedure and to ensure that benefits outweigh potential adverse effects to animals (Williams *et al* 2006). However, Williams *et al*'s (2006) review of the operation and effectiveness of New Zealand severity classification suggested that appropriate and balanced terminology was needed for each category, and that a sixth category should be added for procedures that are deemed unacceptable. The additional category for unacceptable procedures was proposed primarily to communicate to the general public which procedures are not permitted; however, it would also serve to educate animal users.

The severity classification system in the UK is believed to "[e]ncourage deeper thought" by animal users about the effect their experiments will have on animals (APC 2003; p 115). Soon after it was mandated, the Laboratory Animal Science Association (LASA) identified the need for a consistent method for assigning severity (LASA 1990). LASA surveyed animal users and created a 'Severity index' (SI) that itemised the components of severity common to many procedures and gave each a numerical scoring range. The scores could then be summed to assign an SI that, in turn, would be used along with professional judgment to place procedures and protocols into a severity category. This attempt to standardise severity classification suggested that the existing system was not sufficiently educational on

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Figure 1



Trends in number of animals per severity level (category of invasiveness [CI]) in Canada from 2002–2008. Closed circles = CI-B; closed triangles = CI-C; open circles = CI-D; open triangles = CI-E.

its own, and that more detailed description and guidance was needed. Smith and Jennings' (2004) study of the UK system found that adequately detailed descriptions and examples of severity levels and examples of how to classify procedures and protocols were essential for clear communication about animal suffering and cost-benefit assessment. Participants in this study (including investigators, veterinarians and animal welfare advocates) also agreed that the process of assigning severity assists investigators in thinking about animal suffering and refinement options. Elements identified as counterproductive to raising awareness of severity included classification based on the 'average' rather than the individual animal, and the label 'moderate', which may serve to downplay adverse effects and/or become a default category. Similarly, category labels that are overly pain-related were believed to inadequately reflect adverse effects, such as stress, anxiety and other effects, such as nausea (Smith & Jennings 2004).

The studies cited above suggest that the process of severity classification does serve an educative purpose and that the greatest educational benefit is likely to arise from systems with detailed names and descriptions for each category, and with accompanying examples of procedures and protocols that represent each severity level. Practical experience in Canada corroborates these findings as the Canadian severity classification system has been revised twice (in 1989 and 1991) to add detailed examples and to adapt severity classification to non-biomedical use of animals. In addition, requests from investigators for specific directions resulted in inclusion of worked examples for classifying the severity of procedures used for genetically engineered animals (GEAs) (CCAC 1997b) and in field studies (CCAC 2003).

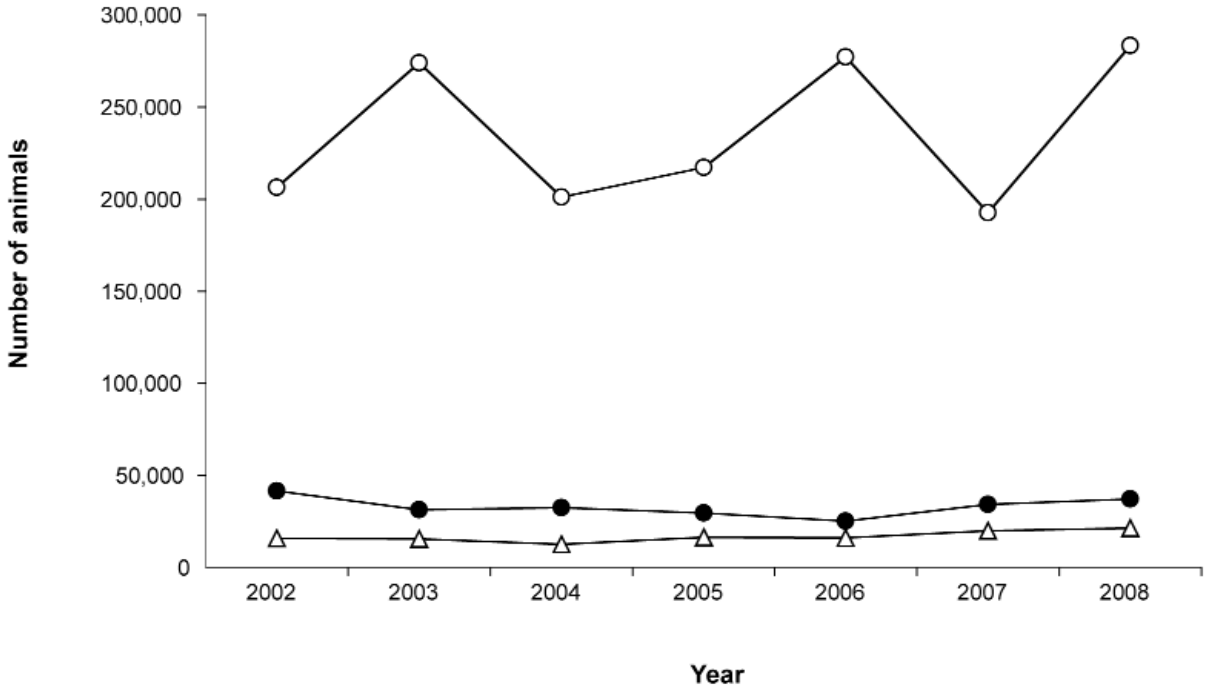
Is severity classification contributing to public accountability?

To our knowledge, of the eleven countries discussed in this paper, only six publicly report severity classification data by number of animals per category (Australia, Canada, Finland, The Netherlands, New Zealand and Switzerland) (Table 1). Publishing the data provides members of the public with the opportunity to review the proportion of animals exposed to each severity level. For example, the majority of scientific animal use occurs at the lower severity levels with less invasive consequences to animals (Williams *et al* 2006). In addition, the public (and policy-makers) can compare the use of animals in science between countries, although direct comparison is difficult as countries vary in the number of categories used and the definitions of 'animal' and 'procedure'. In the future, more uniform international data are expected to be publicly available as the EU directive (which requires severity classification) is fully implemented (EU 2010).

In Canada, animal use statistics are published annually by species, purpose of animal use and severity level (category of invasiveness) (CCAC 2008). Figure 1 shows the trends in severity from 2002–2008 for Canada and shows that the

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Figure 2



Trends in number of animals per severity level in New Zealand in 2002–2008. Open circles = Little/no suffering; filled circles = Moderate Suffering; open triangles = Severe suffering.

numbers of animals in the most severe category (CI-E) fluctuate between 101,655 at the lowest point (in 2003) and 179,781 at the highest point (in 2006). There has also been an increase in the numbers of animals in CI-D since 2002 (attributed to the fact that all newly created GEAs are put into CI-D until a stable phenotype has been established). Following publication, questions from the public and media are received and these typically focus on animal use in the highest severity level. These requests for information are used as an indication of areas of public concern and also highlight a key challenge to prospectively assigning severity: ensuring that the public understands that data only reflects the potential level of pain and distress that a specified number of animals may have experienced. This remains an ongoing challenge in ensuring accountability and transparency to the Canadian public.

New Zealand publishes retrospectively assigned severity data and is therefore reporting actual rather than predicted severity. Figure 2 shows the severity trends from 2002–2008 for New Zealand (New Zealand Government 2009). Although New Zealand collects data from five levels of severity, it publishes three in which levels 'A' and 'B' are combined, and 'D'; and 'E' are combined. For the most part, the number of animals in the combined 'D' and 'E' category ('severe') has remained constant, and includes the fewest numbers of animals. Since retrospective assessment and reporting provides more accurate statistics, in turn this could be expected to increase public confidence in the reported annual animal numbers (Williams *et al* 2006).

The UK publishes the number of animals used per scientific procedure and per purpose of use, but keeps much of its severity data confidential (APC 2009). This has been criticised because the published statistics do not allow the public to determine the nature, level and duration of pain, suffering and distress actually experienced by animals (Smith & Jennings 2004; Nuffield 2005). In addition to more detailed reporting, the Nuffield (2005) review suggested that examples of procedures and protocols for each category should be publicised to improve public understanding of severity classification. These should include examples of animals used over extended periods of time, and describe both the research and factors such as the conditions of breeding, housing and handling.

The examples of Canada, New Zealand and the UK suggest that publishing severity statistics provides, at least, the opportunity for members of the public to assess the effects of scientific use on animals and this, by extension, contributes to public accountability.

Is severity classification informing national policies on the use of animals in science?

In Canada, analysing severity data trends has proven useful to establish whether the number of animals in the most severe categories is changing. It has also provided valuable information about how the Three Rs are being implemented and where to target future Three R's efforts. For example, Canada expects the number of animals in the highest severity categories to decrease or at least remain steady in

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

the presence of increasing funding for animal-based research as the number of animals used in research is reduced and techniques are refined. In New Zealand, information from severity classification is used to track "Three R's developments" (Bayvel *et al* 2007; p 712), to provide justification for Three R's research (Bayvel *et al* 2007), and to provide confirmation to regulators that cost-benefit assessments are being undertaken (Williams *et al* 2006).

Canada can provide three specific examples of how severity classification has been able to inform policy. First, considerable numbers of animals reported in the highest severity levels (CI-D and CI-E) prompted the development of the *CCAC Guidelines on: Choosing an Appropriate Endpoint in Experiments Using Animals for Research, Teaching and Testing* (CCAC 1998). The years following the publication of this guideline saw a reduction in the number of animals reported in CI-E, and reflected the increased level of monitoring of animals and use of clinical signs to determine endpoints (Gauthier 2004).

Second, Canada's reporting of prospectively assigned severity data has had the unfortunate consequence of skewing animal use data to suggest that the worst-case scenario has actually been experienced by all animals. In particular, protocols involving the generation of a GEA line are required to be assigned to category CI-D (CCAC 1997b) because of the unpredictable impact of genetic engineering techniques, and the associated potential for pain and distress. Although most GEAs do not experience pain and distress, if the protocol is not reclassified, this results in the inflation of numbers of animals in that severity category. Canadian animal use data thus convey the false impression that there is a substantial increase in the number of animals experiencing considerable levels of pain and distress. Efforts are currently underway to address this inaccuracy through the revision of the *CCAC Guidelines on: Transgenic Animals* (CCAC 1997b) as *CCAC Guidelines on: Genetically Engineered Animals Used in Science* (CCAC, in preparation).

Third, in the years 2004–2006, Canada saw an increase in numbers of animals in the highest severity category (CI-E). Examination of the data showed that most of the animals in this category were being used for regulatory testing, in particular, considerable numbers of mice were being used for shellfish toxin testing. This finding prompted Guy and Griffin's (2009) study of the Three Rs in shellfish toxin testing, and identification of several areas where federal government policy (including the implementation of validated non-animal methods) could have a direct influence on the numbers of animals in the highest severity category. Using severity classification data to guide policy follows the intention first expressed by Smyth (1978): that categorisation of animal experiments could be used as an indication of where the greatest efforts should be placed to implement the Three Rs. These examples also show that there is a clear role for severity classification in national animal use policy development.

Discussion: emerging challenges for severity classification

Severity classification has been used for 30 years and has provided: an important tool for AECs in evaluating the ethical acceptability of animal-based science; an educative process for investigators; and meaningful animal use statistics to inform both the public and animal use policy. However, developments in science continue to present new challenges to old policies and several key challenges for severity classification have emerged.

The first challenge is to accurately reflect the number of animals that experience the various levels of severity. Reporting retrospectively assigned severity classifications would ensure that published information more accurately reflects the pain and distress experienced by animals and thereby increase public accountability. Retrospective reporting may also assist in further understanding the outcomes to animals from particular procedures. The use of retrospective assignment and reporting must be weighed against the possible administrative burden placed on institutions, investigators and AECs. However, it may prove to be acceptable: a recent UK study suggested that the administrative burden could be acceptably minimised if all statistics on animal use were reported retrospectively and if appropriate training initiatives were in place to aid in transitioning to a new system (APC 2008). Use of retrospective assignment and reporting is also supported by the Europe-based Federation of Laboratory Animal Science Associations (FELASA 2007), and the new EU directive includes a component of retrospective assignment and reporting for the highest severity level (EU 2010).

Another challenge relates to classifying the welfare status of animals maintained in breeding colonies (ie when not being used in experimental protocols). Growing concern for the quality of life of otherwise healthy animals housed in laboratory facilities has triggered consideration of how to acknowledge these effects in severity classification, which is mostly based on the impact of experimental procedures on the animals. Effects on quality of life become even more apparent in situations where animals are bred specifically to have a disease condition, for example diabetes. In addition, there are instances where the need to maintain a defined health status dictates housing the animal in an impoverished environment, for example immune-compromised mice must be maintained in ventilated cages, and provision of nesting material is difficult because it can lead to skin irritations (Baumans *et al* 2006). How might these varying levels of severity be captured?

Severity classification systems are also challenged by the need to ensure that outcomes to animals are considered along with experimental procedures that are used. This would increase the focus on the animals, an approach that aligns with recent international farm animal welfare standards (eg NFACC 2009). With this approach, outcome-based rather than prescriptive guidelines are used. For example, the effect of a particular aspect of housing (ie flooring material) is measured by a parameter

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

relevant to animal welfare (ie incidence of lameness). This approach requires that animal facilities have a good animal welfare assessment system in place. While checklists for monitoring animal welfare have been advocated since the pivotal work of Morton and Griffiths (1985), considerable work is needed to establish reliable indicators of pain and distress, and to fully understand and classify the severity of scientific procedures on animals.

Finally, classifying harms to GEAs at different stages of the production of a new animal line is emerging as a particularly complex challenge. Attempts to accurately identify the pain and distress experienced by an animal have led to discussions about whether severity classification should more appropriately reflect the lifetime experience, and nowhere is this of greater relevance than for GEAs. In the UK, being born a GEA is considered to be a procedure, and animals are thereby assigned to a particular severity level. This recognises that the animal may experience negative welfare due to having a harmful genetic alteration. If the animal is then used in a protocol — for example, to determine the effect of the genetic modification on its phenotype, these additional procedures may impose an added burden on the animal, over and above what might be experienced by a non-modified or 'conventional' animal. In addition, for GEAs, it remains to be seen how severity classification can address animal experiences where there has been a change to the animals' telos (ie the manner in which it expresses its interactions with its environment [Rollin 1998; Gauthier & Griffin 2005]). For now, in cases where the outcome of genetic modification is unknown, investigators and AECs might use the inability to accurately assign severity as an indication that a pilot study is needed in order to determine the likely effect on the animal.

Conclusion and animal welfare implications

Our analysis has shown that to maximise the effectiveness of severity classification and sharpen the focus on improving animal welfare and minimising pain and distress, severity classification systems should have the following characteristics:

• Some form of retrospective assessment of the severity assignments should be carried out and reported in order to validate initial classifications. This would also serve an educative role. Retrospective assessment should increase the accuracy of severity classifications and hence increase both investigators' and public confidence in the data. More accurate data would also be valuable for sound policy development.

• Severity levels should be defined in such a way that the quality of life of the animal and the outcomes for the animal are considered in the classification process (ie not just the scientific procedures that are performed). This is of particular importance when considering GEAs or animals bred to have a disease that will experience the effects throughout their lifespan.

• Classification should be focused on the experience of individual animals, not on the experience of a group of animals or an average experience.

• Descriptive supporting guidance should be provided, including detailed examples of what procedures and outcomes to animals are appropriate for each level of severity.

Acknowledgements

The authors gratefully acknowledge the following people for their generous assistance in the research and preparation of this paper: Wendy Clarence, Julie Dale, Ana Tereza Pinto Filipecki, Vanessa Carli Bones Silla, Anne Thibaudeau, Elin Weber and Drs Daniel Butzke, Eila Kaliste, Marilyn Keaney, Anna Olssen, Margaret Rose, Catherine Schuppli, Patrick Sheridan, Krystyna Vocadlo, Marina von Keyserlingk and Francoise Wemelsfelder. We especially thank Paulina Jawor for considerable help with translation of Polish documents, and the anonymous reviewers of this paper for their insightful comments.

References

Animal Procedures Committee (APC) 2003 *Report of the Cost-Benefit Working Group of the Animal Procedures Committee*. http://apc.homeoffice.gov.uk/reference/costbenefit.pdf

Animal Procedures Committee (APC) 2008 *Final Report of a LASA/APA Working Group to Examine the Feasibility of Reporting Data on the Severity of Scientific Procedures on Animals*. http://www.lasa.co.uk/synopsis_LASA_APC_severity%20project.html

Animal Procedures Committee (APC) 2009 *Report of the Suffering and Severity Working Group on the Strengths and Weaknesses of the Current System of Severity Limits as a way of Prospectively Assessing Suffering and Severity*. http://apc.homeoffice.gov.uk/reference/Suffering%20and%20Severity%20paper.pdf

Animal Welfare Directorate 2008 *Animal Use Statistics*. http://www.biosecurity.govt.nz/files/regs/animal-welfare/pubs/naeac/animal-use-statistics.pdf

Baumans V, Clausing P, Hubrecht R, Reber A, Vitale A, Wyffels E and Gyger M 2006 *FELASA Working Group Standardization of Enrichment: Working Group Report*. Laboratory Animals Ltd: London, UK. http://www.lal.org.uk/pdffiles/FELASA_Enrichment_2006.pdf

Bayvel ACD, Carsons LA and Littin KE 2007 Severity assessment — The New Zealand experience and perspective. *Proceedings of the Sixth World Congress on Alternatives. Alternatives to Animal Testing and Experimentation 14*: 711-714. http://altweb.jhsph.edu/bin/q/t/paper711.pdf

Bowd AD 1997 The educative role of an animal care committee in Canada: a case study. *Animal Welfare 6*: 309-315

Canadian Council on Animal Care (CCAC) 1987 *Confidential memo to Dr H Rowsell*. CCAC: Ottowa, Canada

Canadian Council on Animal Care (CCAC) 1991 *CCAC Policy on: Categories of Invasiveness in Animal Experiments*. http://www.ccac.ca/en/CCAC_Programs/Guidelines_Policies/POLICIES/CATEG.HTM

Canadian Council on Animal Care (CCAC) 1997a *CCAC Guidelines on: Animal Use Protocol Review*. http://www.ccac.ca/en/CCAC_Programs/Guidelines_Policies/GDLINES/PROTOCOL/g_protocol.pdf

Canadian Council on Animal Care (CCAC) 1997b *CCAC Guidelines on: Transgenic Animals*. http://www.ccac.ca/en/CCAC_Programs/Guidelines_Policies/GDLINES/TRANSGEN/g_transgenic.pdf

Canadian Council on Animal Care (CCAC) 1998 CCAC *Guidelines on: Choosing an Appropriate Endpoint in Experiments Using Animals for Research, Teaching and Testing*. http://www.ccac.ca/en/CCAC_Programs/Guidelines_Policies/GDLINES/ENDPTS/g_endpoints.pdf

Canadian Council on Animal Care (CCAC) 1999 *CCAC Guidelines on: Institutional Animal User Training*. http://www.ccac.ca/en/CCAC_Programs/Guidelines_Policies/GDLINES/NIAUT/g_niaut.pdf

Canadian Council on Animal Care (CCAC) 2003 CCAC *Guidelines on: the Care and Use of Wildife*. CACC: Ottowa, Canada http://www.ccac.ca/en/CCAC_Programs/Guidelines_Policies/GDLINES/Wildlife/Wildlife.pdf

Canadian Council on Animal Care (CCAC) 2008 *Trends in Animal Use*. http://www.ccac.ca/en/Publications/New_Facts_Figures/trends/figure6.htm

Canadian Council on Animal Care (CCAC) 2009 *Three Rs Microsite*. http://www.ccac.ca/en/alternatives/index.html

Canadian Council on Animal Care (CCAC) (in preparation) *CCAC Guidelines on: Genetically Engineered Animals*. CACC: Ottowa, Canada

Department of Health and Children 2006 *Application For A License*. http://www.dohc.ie/other_health_issues/uaeosp/Application%20Form.pdf?direct=1

European Union (EC) 2010 *Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the Protection of Animals Used for Scientific Purposes*. European Commission: Brussels, Belgium. http://eur-lex.europa.eu/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:EN:PDF

Federation of European Laboratory Animal Scientists (FELASA) 2007 *FELASA Statement on a Proposal from the Laboratory Animal Science Association (LASA) for the European Community to Adopt Retrospective Reporting of Scientific Procedures on Animals*. http://www.felasa.eu/docs/Statement_RetrospectiveReporting_2007-11-13.pdf

Gauthier C 2004 Overview and analysis of animal use in North America. *Alternatives to Laboratory Animals 32(1A)*: 275-285

Gauthier C and Griffin G 2005 Using animals in research, teaching and testing. *Revue Scientifique et Technique (International Office of Epizootics) 24*: 735-745

Griffin G, Dansereau M and Gauthier C 2007 Categories of invasiveness: a precautionary approach. *Proceedings of the Sixth World Congress on Alternatives. Alternatives to Animal Testing and Experimentation 14*: 715-720. http://altweb.jhsph.edu/wc6/paper715.pdf

Guy A and Griffin G 2009 Adopting alternatives for the regulatory monitoring of shellfish for paralytic shellfish poisoning in Canada: interface between federal regulators, science and ethics. *Regulatory Toxicology and Pharmacology 54*: 256-263

Home Office 2000 *Guidance on the Operation of the Animals (Scientific Procedures) Act 1986*. http://www.archive.official-documents.co.uk/document/hoc/321/321-05.htm#gen113

Kuhara T 2008 Attitudes to severity assessment in Japan. *Proceedings of the Sixth World Congress on Alternatives. Alternatives to Animal Testing and Experimentation 14*: 721-722. http://altweb.jhsph.edu/wc6/paper721.pdf

Laboratory Animal Scientists Association (LASA) 1990 LASA Working Party Report: The assessment and control of the severity of scientific procedures on laboratory animals. *Laboratory Animals 24*: 97-130. http://la.rsmjournals.com/cgi/reprint/24/2/97.pdf

Mellor DJ and Reid CSW 1994 Concepts of animal well-being and predicting the impact of procedures on experimental animals. In: Jenkin G, Mellor DJ and Baker RM (eds) *Improving the Well-being of Animals in the Research Environment* pp 3-18. Glen Osmond: SA, Australia

Mellor DJ, Patterson-Kane E and Stafford KJ 2009 *The Sciences of Animal Welfare*. Wiley-Blackwell: Oxford, UK

Ministry of Science and Higher Education 2006 *Skala inwazyjności doświadczeń przeprowadzania na zwierzętach*. http://www.nauka.gov.pl/nauka/krajowa-komisja-etyczna-ds-doswiadczen-na-zwierzetach/zwierzeta-labolatoryjne/zwierzeta-labolatoryjne/artykul/skala-inwazyjnosci-badan-na-zywych-kregowcach/. [Title translation: Scale of invasiveness of experiments conducted on animals]

Morton DB and Griffiths PHM 1985 Guidelines on the recognition of pain, distress and discomfort in experimental animals and a hypothesis for assessment. *The Veterinary Record 116(16)*: 431- 436

National Farm Animal Care Council (NFACC) 2009 *Code of Practice for the Care and Handling of Dairy Cattle*. Dairy Farmers of Canada and National Farm Animal Care Council: Ottawa, Canada. http://www.dairyfarmers.ca/what-we-do/animal-health-and-welfare/code-of-practice-for-the-care-and-handling-of-dairy-cattle

New South Wales 2009 *Form L: Animal Use Statistics 2009 Explanatory Notes*. http://www.animalethics.org.au/__data/assets/word_doc/0009/249993/form-l2009.doc

New Zealand Government 2009 *NAEAC Annual Reports 2002-2008*. http://www.biosecurity.govt.nz/regs/animal-welfare/naeac/annual-reports

Nuffield Council on Bioethics 2005 *The Ethics of Research Involving Animals*. Nuffield Council on Bioethics: London, UK. http://www.nuffieldbioethics.org/fileLibrary/pdf/RIA_Report_FINAL-opt.pdf

Obrink KJ 1982 Swedish law on laboratory animals. In: Dodds WJ and Orlans FB (eds) *Scientific Perspectives on Animal Welfare*. Academic Press: New York, USA

Orlans FB 1980 Animal welfare. *BioScience 30*: 144-145

Orlans FB 1987 Research protocol review for animal welfare. *Investigative Radiology 22*: 253-258

Orlans FB 1990 Animal pain scales in public policy. *Alternatives to Laboratory Animals 18*: 41-50

Orlans FB 1993 *Animal pain scales in public policy: An international perspective. In the Name of Science: Issues in Responsible Animal Experimentation*. Oxford University Press: New York, USA

Orlans FB 2000 Public policies on assessing and reporting degrees of animal harm; international perspectives. In: Balls M, van Zeller A-M and Halder ME (eds) *Progress in the Reduction, Refinement and Replacement of Animal Experimentation*. Elsevier Science: London, UK

Porter DG 1992 Ethical scores for animal experiments. *Nature 356*: 101-102

Purves KE 2000 Global harmonization of pain and distress classification systems: current analysis and opportunity for change. *Journal of Applied Animal Welfare Science 3*: 49-61

Rollin BE 1998 *The Unheeded Cry: Animal Consciousness, Animal Pain and Science*, Iowa State University Press: Iowa, USA

Rony Kalman I, Olsson AS, Bernardi C, van den Broek F, Bronstad A, Gyertyan I, Lang A, Marinou KA and Zeller W 2010 Ethical evaluation of procedures for experiments involving animals: Recommendations for ethics committees. In: Howard B, Nevalainen T and Perretta G (eds) *COST Manual of Laboratory Animal Care and Use: Refinement, Reduction, and Research*. CRC Press: USA

Ross MW 1981 The ethics of experiments on higher animals. *Social Science & Medicine 15F*: 51-60

Russell WMS and Burch RL 1959 *The Principles of Humane Experimental Technique*. Methuen & Co Ltd: London, UK

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Schuppli CA 2004 *The role of the animal ethics committee in achieving humane animal experimentation.* PhD Thesis, The University of British Columbia, Vancouver, Canada. http://www.landfood.ubc.ca/animalwelfare/publications/pdfs/theses/phdthesisSchuppli.pdf

Schuppli CA and Fraser D 2005 The interpretation and application of the Three Rs by AEC members. *Alternatives to Laboratory Animals 33*: 487-500

Shapiro KJ and Field PB 1988 A new invasiveness scale: its role in reducing animal distress. *PsyETA Bulletin 2*: 43-46

Silla VCB, Marthos SM and Molento CFM 2009 Two years description of the Animal Use Ethic Committee of the Agrarian Science Sector of the Federal University of Paraná. *Ciência Rural 39*: 2093-2098

Smith JA and Jennings M 2004 *Categorising the severity of scientific procedures on animals: Summary and reports from three round-table discussions on the use of severity limits and bands in the UK.* RSPCA Research Animals Department: Horsham, UK. http://www.boyd-group.demon.co.uk/severity_report.pdf

Smith JA, van den Broek FAR, Martorell JC, Hackbarth H, Ruksenas O and Zeller W 2005 Principles and practice in ethical review of animals experiments across Europe. *A report prepared by the FELASA working group on ethical evaluation of animal experiments.* http://www.felasa.eu/docs/WG_ethical_review_final.pdf

Smyth DH 1978 *Alternatives to Animal Experiments.* Scolar Press: London, UK

Stephens ML, Mendoza P, Weaver A and Hamilton T 1998 Unrelieved pain and distress in animals: An analysis of USDA data on experimental procedures. *Journal of Applied Animal Welfare Science 1*: 15-26

Swiss Federal Veterinary Office 2006 *Interim and final reports on animal experiments: Explanatory notes for Form C (800.116-1.03).* http://www.bvet.admin.ch/themen/tierschutz/00777/00779/index.html?lang=en

Swiss Federal Veterinary Office (undated). *Retrospective classification of animal experiments according to degrees of severity prior to the experiment (stress categories).* http://www.tierversuch.ch/?show=AWLaw&nav_id=4105&lang=en

United Stated Department of Agriculture (**USDA**) 2000 Information resources for institutional animal care and use committees [online database]. *Animal Welfare Information Center* Resource Series No 7. http://www.nal.usda.gov/awic/pubs/IACUC/iacuc.htm

Williams VM, Mellor DJ and Marbrook J 2006 Revision of a scale for assessing the severity of live animal manipulations. *Alternatives to Animal Testing and Experimentation 23*: 163-169

Williamson A and Hoggart B 2005 Pain: a review of three commonly used pain rating scales. *Journal of Clinical Nursing 14*: 798-804

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 Current descriptions of international severity classification systems.

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
Australia *Categorisation of procedures*	1. Observation involving minor interference	Animals are not interacted with or, where there is interaction, it would not be expected to compromise the animal's welfare any more than normal handling, feeding etc. There is no pain or suffering involved. Examples: • Observational study only • Breeding animals for supply, where only normal husbandry procedures are used • Breeding or reproductive study with no detriment to the animal • Feeding trial, such as Digestible Energy determination of feed in a balanced diet • Behavioural study with minor environmental manipulation • Teaching of normal, non-invasive husbandry, such as handling, grooming etc
	2. Animal unconscious without recovery	Animal is rendered unconscious under controlled circumstances with little or no pain or distress. Capture methods are not required. Any pain is minor and brief and does not require analgesia. Procedures are carried out on the unconscious animal which is then killed without regaining consciousness. Examples: • Laboratory animals killed painlessly for dissection, biochemical analysis, etc • Teaching surgical techniques on live, anaesthetised patients which are not allowed to recover following the procedure
	3. Minor conscious intervention	Animal is subjected to minor procedures which would normally not require anaesthesia or analgesia. Any pain is minor and analgesia usually unnecessary, although some distress may occur as a result of trapping or handling. Examples: • Injections, blood sampling in conscious animal • Minor dietary or environmental deprivation or manipulation, such as feeding nutrient-deficient diets for short periods • Trapping and release as used in species impact studies, etc • Trapping and humane euthanasia for collections of specimens • Stomach tubing, shearing
	4. Minor surgery with recovery	Animal is rendered unconscious with as little pain or distress as possible. A minor procedure such as cannulation or skin biopsy is carried out and the animal allowed to recover. Depending on the procedure, pain may be minor or moderate and post-operative analgesia may be appropriate. Field capture using chemical restraint methods is also included here. Examples: • Biopsies • Cannulations • Sedation/anaesthesia for relocation, examination or injections/blood sampling
	5. Major surgery with recovery	Animal is rendered unconsciousness with as little pain or distress as possible. A major procedure such as abdominal or orthopaedic surgery is carried out and the animal allowed to recover. Post-operative pain is usually considerable and at a level requiring analgesia. Examples: • Orthopaedic surgery • Abdomimal or thoracic surgery • Transplant surgery • Mulesing, castration without anaesthesia

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
Australia (cont)	6. Minor physiological challenge	Animal remains conscious for some or all of the procedure. There is interference with the animal's physiological or psychological processes. The challenge may cause only a small degree of pain/distress or any pain/distress is quickly and effectively alleviated. Examples: • Minor infection, minor or moderate phenotypic modification, early oncogenesis • Arthritis studies with pain alleviation • Prolonged deficient diets, induction of metabolic disease • Polyclonal antibody production • Antiserum production
	7. Major physiological challenge	Animal remains conscious for some or all of the procedure. There is interference with the animal's physiological or psychological processes. The challenge causes a moderate or large degree of pain/distress which is not quickly or effectively alleviated. Examples: • Major infection, major phenotypic modification, oncogenesis without pain alleviation • Arthritis studies with no pain alleviation, uncontrolled metabolic disease • Isolation or environmental deprivation for extended periods • Monoclonal antibody raising in mice
	8. Death as an endpoint	This category only applies in those rare cases where the death of the animal is a planned part of the procedures. Where predictive signs of death have been determined *and* euthanasia is carried out before significant suffering occurs, they may be placed in category 6 or 7. Examples: • Lethality testing (including LD_{50}, LC_{50}). It does not include: death by natural causes; animals which are euthanised as part of the project; animals which are euthanised if something goes wrong; animals euthanised for dissection or for use as museum specimens; or accidental deaths
	9. Production of genetically modified animals	This category is intended to allow for the variety of procedures which occur during the production of genetically modified animals. As animals in this category may be subjected to both minor *and* major physiological challenges *and* surgical procedures, this category reflects the varied nature of the procedures carried out. It effectively includes ALL animals used in GM production other than the final progeny which are used in a different category of procedure. Examples: • Initial breeding animals for GM production • Animals culled as part of the GM production process
Canada *Categories of invasiveness*	A. Experiments on most invertebrates, or on live isolates	Possible examples: the use of tissue culture and tissues obtained at necropsy or from the slaughterhouse; the use of eggs, protozoa or other single-celled organisms; experiments involving containment, incision or other invasive procedures on metazoan
	B. Experiments which cause little or no discomfort or stress	Possible examples: domestic flocks or herds being maintained in simulated or actual commercial production management systems; the short-term and skilful restraint of animals for purposes of observation or physical examination; blood sampling; injection of material in amounts that will not cause adverse reactions by the following routes: intravenous, subcutaneous, intramuscular, intraperitoneal or oral but not intrathoracic or intracardiac (C); acute non-survival studies in which the animals are completely anaesthetised following rapid unconsciousness, such as anaesthetic overdose, or decapitation preceded by sedation or light anaesthesia; short periods of food and/or water deprivation equivalent to periods of abstinence in nature

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
Canada (cont)	C. Experiments which cause minor stress or pain of short duration	Possible examples: cannulation or catheterisation of blood vessels or body cavities under anaesthesia; minor surgical procedures under anaesthesia, such as biopsies, laparoscopy; short periods of restraint beyond that for simple observation or examination, but consistent with minimal distress; short periods of food and/or water deprivation which exceed periods of abstinence in nature; behavioural experiments on conscious animals that involve short-term, stressful restraint; exposure to non-lethal levels of drugs or chemicals. Such procedures should not cause significant changes in the animal's appearance, in physiological parameters, such as respiratory or cardiac rate, or faecal or urinary output, or in social responses *Note: during or after Category C studies, animals must not show self-mutilation, anorexia, dehydration, hyperactivity, increased recumbency or dormancy, increased vocalisation, aggressive-defensive behaviour or demonstrate social withdrawal and self-isolation*
	D. Experiments which cause moderate to severe distress or discomfort	Possible examples: major surgical procedures conducted under general anaesthesia, with subsequent recovery; prolonged (several hours or more) periods of physical restraint; induction of behavioural stresses such as maternal deprivation, aggression, predator-prey interactions; procedures which cause severe, persistent or irreversible disruption of sensorimotor organization; the use of Freund's Complete Adjuvant (see *CCAC Guidelines on Acceptable Immunological Procedures*). Other examples include induction of anatomical and physiological abnormalities that will result in pain or distress; the exposure of an animal to noxious stimuli from which escape is impossible; the production of radiation sickness; exposure to drugs or chemicals at levels that impair physiological systems. *Note: Procedures used in Category D studies should not cause prolonged or severe clinical distress as may be exhibited by a wide range of clinical signs, such as marked abnormalities in behavioural patterns or attitudes, the absence of grooming, dehydration, abnormal vocalization, prolonged anorexia, circulatory collapse, extreme lethargy or disinclination to move, and clinical signs of severe or advanced local or systemic infection, etc*
	E. Procedures which cause severe pain near, at, or above the pain tolerance threshold of unanaesthetised conscious animals	This category of invasiveness is not necessarily confined to surgical procedures, but may include exposure to noxious stimuli or agents whose effects are unknown; exposure to drugs or chemicals at levels that (may) markedly impair physiological systems and which cause death, severe pain, or extreme distress; completely new biochemical experiments which have a high degree of invasiveness; behavioural studies about which the effects of the degree of distress are not known; use of muscle relaxants or paralytic drugs without anaesthetics, burn or trauma infliction on unanaesthetised aniamls; a euthanasia method not approved by the CCAS; any procedures (eg the injection of noxious agents or the induction of severe stress or shock) that will result in pain which approaches the pain tolerance threshold and cannot be relieved by analgesia (eg when toxicity testing and experimentally induced infectious disease studies have death as the endpoint
Finland	Class 2, minor or short duration pain	Information not readily available
Harm assessment	Class 1, procedures that might cause severe pain, suffering or distress or serious illness to the animal	Information not readily available

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
Finland (cont)	Class 0, no pain, suffering or distress	Information not readily available
Germany	Level 0 None	No pain, suffering, harm or fear
Severity of harm	Level 1 Mild	Light, short-term pain or harm. Duration of harm taken into account as well: < 1 day, 1–7 days, 7–30 days, > 30 days
	Level 2 Moderate	Middle-grade, short-term pain or harm. Duration of harm taken into account as well: < 1 day, 1–7 days, 7–30 days, > 30 days
	Level 3 Severe	Significant or middle to long-lasting, middle-grade pain; suffering; harm; or fear. Duration of harm taken into account as well: < 1 day, 1–7 days, 7–30 days, > 30 days
Ireland	Mild	Information not readily available
Categories of severity	Moderate	Information not readily available
	Substantial	Information not readily available
	Severe and likely to be prolonged	Information not readily available
The Netherlands	Minor	Information not readily available
Categories of pain or discomfort	Minor/moderate	Information not readily available
	Moderate	Information not readily available
	Moderate/severe	Information not readily available
	Severe	Information not readily available
	Very severe	Information not readily available
New Zealand *Grading of manipulations*	Grade A, no impact or virtually no impact	**Mental state**: Field observations of grazing behaviour on farms, or benign handling of tame and trained animals which are familiar with all personnel and procedures and with the place where the procedures are conducted **Food/water**: Animals kept outdoors eating their usual food in appropriate amounts; grazing trials on treated pastures; offering supplements to naturally available food; provision of complete, balanced rations to meet all nutritional requirements of animals maintained indoors **Environmental challenge**: Exposure to ambient conditions which are within the thermoneutral range; reduced barometric pressures which do not cause increases in red blood cell production. **Disease/injury/functional impairment**: Studies of healthy uninjured animals which are kept in physical conditions which do not themselves lead to injuries such as lameness or compression sores; studies to establish normal characteristics of healthy animals **Behaviour:** Studies of wild or undomesticated animals in their natural habitats; field studies of domesticated animals
	Grade B, little impact	**Mental state**: Experiments on completely anaesthetised animals which do not regain consciousness; simple venipuncture or venisection; injection of non-toxic substances; skin tests which cause low-level irritation without ulceration/erosion; feeding trained animals by orogastric tube; movement of free-range domesticated animals to unfamiliar housing; minor restrictions of water and/or feed intake beyond the normal period of satiation **Food/water**: Water priming for kidney function tests; short-term overall food intake restrictions or excesses which are within usual tolerance levels for the species; short-term changes in dietary composition which cause no clinical signs of deficiency or toxicity, but which would cause such symptoms in the longer term **Environmental challenge**: Exposure to levels of cold or heat which are outside the thermoneutral range, or barometric pressures which increase red blood cell production, but which remain within the capacity of the animals to adapt and do not lead to debility in the long term

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
New Zealand (cont)	Grade B, little impact (cont)	**Disease/injury/functional impairment**: Studies of vaccines using killed pathogens; tuberculosis tests; induction of mild fever without other debilitating effects; induction of subclinical parasitism; healing of minor superficial incisions, cuts or wounds; minor surgical and/or pharmacological modification of homeostatic capacity (eg creation of non-obstructive gut fistulae; splenectomy; endocrine gland removal with complete and permanent hormone replacement therapy); physical conditions which cause transient lameness of low intensity, mild compression sores or abrasions **Behaviour**: Mild and short-term physical restraint; keeping free-range domesticated animals in a yard; movement of free-range domesticated livestock to unfamiliar housing; operant conditioning with positive reinforcement in barren laboratory environments; benign preference tests in unnatural surroundings
	Grade C, moderate impact	Includes manipulations of minor impact and long duration or moderate impact and short duration **Mental state**: Recovery from major surgeries like thoracotomy, orthopaedic procedures, hysterectomy or gall bladder removal with effective use of analgesics; surgical procedures on conscious animals but with the use of local anaesthesia and systemic analgesic; movement of excitable free-range domesticated livestock to unfamiliar housing; short term capture, handling and restraint of wild or semi-domesticated animals that exhibit marked flight responses; moderate restrictions of water and/or feed intake beyond the normal period of satiation **Food/water**: Simulation of usual overall intake restrictions often experienced by pregnant/lactating ruminants during cold winters or drought; dietary induction of milk fever in cattle; induction of mild deficiency or toxicity signs by feeding diets containing inadequate or excessive amounts of essential nutrients **Environmental challenge**: Short-term exposure to severe extremes of cold or heat which would lead to collapse if prolonged **Disease/injury/functional impairment**: Studies of live vaccines; induction of clinical parasitism; induction of mild reversible infectious diarrhoea; moderate surgical and/or pharmacological modification to homeostatic capacity (eg limited gut resection; endocrine gland removal with delayed or incomplete hormone replacement therapy); physical conditions which cause minor chronic lameness or other injuries; studies of the effects of infectious or toxic agents that cause rapid death without distress **Behaviour**: Medium-term restrictions of instinctive behaviour; medium-term holding of ruminants in a metabolism crate; long-term restraint leading to the development of reversible stereotypies; changing social group composition
	Grade D, high impact	Includes manipulations of moderate impact and long duration or high impact and short duration **Mental state**: Recovery from major surgery under anaesthesia without the use of postoperative analgesics; marked social or environmental deprivation; longer term capture, handling, restraint or housing, without the use of tranquilisers, of wild or semi-domesticated animals that exhibit marked flight responses **Food/water**: Dietary induction of advanced pregnancy toxaemia in sheep or ketosis in dairy cattle; dietary induction of advanced signs of nutrient deficiency or excess; severe deleterious effects of dietary toxins; severe restrictions of water and/or feed intake beyond the normal period of satiation

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
New Zealand (cont)	Grade D, high impact (continued)	**Environmental challenge**: Prolonged exposure to severe cold or heat which would lead to failure of thermoregulation and collapse, but the exposure is terminated just before those outcomes **Disease/injury/functional impairment**: Studies of severe facial eczema; induction of severe diarrhoea or severe infectious pneumonia; protracted or irreversible pharmacological modification of homeostatic capacity (eg chemical induction of diabetes mellitus without replacement therapy); marked surgical modification of homeostatic capacity (eg extensive gut resection; cutting of sensory or motor nerves serving large areas of the body from which no self-mutilation injury results; precise lesioning of limited areas of the brain but with intervention before collapse); physical conditions which cause moderate chronic lameness or other injuries; studies of the effects of infectious and toxic agents which cause either a protracted death with minor distress or a rapid death with moderate distress **Behaviour**: Application of marked and repeated noxious stimuli from which escape is impossible; prolonged periods (several hours or more) of close physical restraint; marked alterations to the perceptual or motor functions of animals to test consequent behaviour
	Grade E, very high impact	Manipulations of high impact and long duration **Mental state**: Conducting major surgeries without the use of anaesthesia on control animals in assessing efficacy of analgesics; testing the efficacy of analgesics in animals with severe induced pain **Food/wate**r: Experiments which cause animals to die from poisoning by toxins in the diet; protracted and severe restrictions on water and/or feed intake **Environmental challenge**: Purposeful exposure of conscious animals to lethal extremes of cold, heat or barometric pressure which duplicate naturally occurring conditions **Disease/injury/functional impairment**: Studies of methods for killing pest animals; cutting of sensory or motor nerves serving large areas of the body from which self-mutilation injury results; evaluation of vaccines where death is the measure of failure to protect; studies of the effects of infectious or toxic agents which cause either a protracted death with marked distress or a rapid death with severe distress **Behaviour**: Application of marked and repeated extremely noxious stimuli from which escape is impossible; prolonged periods (several hours or more) of close physical restraint
Poland Scale of invasiveness of experiments conducted in animals	Grade 1, non-invasive procedures in which none of the animals in the experiment is exposed to any suffering or injury	Examples: • Behavioural observations on small animals in captivity including the use of behavioural tests in animals not undergoing any treatment
	Grade 2, procedures resulting in slight momentary gain, stress or prolonged mild discomfort	Examples: • momentary restraint (several minutes) in order to make a simple observation or treatment • blood collection • administration of substances intravenously, intramuscularly, subcutaneously, peritoneally or orally in an amount/dose not resulting in adverse effects • killing (euthanasia) using acceptable methods which cause an immediate almost stressless loss of consciousness • terminal experiments under deep anaesthesia, after which the animal does not awaken from anaesthesia but is euthanised using methods recommended by National Ethical Committee • short duration food deprivation which is similar to food deprivation periods naturally occurring for different animals in nature

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
Poland (cont)	Grade 3, procedures resulting in moderate pain or stress	Examples: • cannulation or catheterisation of vessels or body cavity with general or local anaesthesia • minor surgical procedures with anaesthesia, in which there is no removal or change in tissue essential for life and in an amount to significantly influence function of organism (biopsy, laparoscopy, opening peritoneal cavity, the removal of the gonads) • short lasting restraint, beyond grade 2, without sedatives or anaesthesia • short duration water/food deprivation longer than periods occurring in nature • administration of a non lethal dose of pharmacological or chemical substances • injection into the heart and thorax • exposure to harmful, stressful stimuli, but with the possibility of escape These procedures cannot lead to significant changes in animal's behaviour, or changes in physical parameters (eg breathing, pulse, feed/water intake, defaecation). These procedures cannot provoke animal self-mutilation, dehydration, anorexia, increase in motor activity, increase in vocalisation, as well changes in animal behaviour (eg increase in aggression level, isolation from other individuals or other pathological social behaviour)
	Grade 4, procedures resulting in severe pain/stress and usually irreversible damage to the body and physiological functions	Experiments causing moderate or severe suffering (distress) and ailment, without long lasting or severe suffering (clinical symptoms: abnormal behaviour, lack of self-cleaning, dehydration, vocalisation, stillness, long lasting anorexia, symptoms of infection) Examples: • major surgery in general anaesthesia ie involving opening main body cavities (skull, thorax, spinal canal, abdomen cavity, pelvis) or exposure of major structures including blood vessel, lymphatic vessel, musculature, bone or gland systems, or removal and/or changes in tissue essential for life in amount that significantly influence the function of organism • behavioural stress (deprivation of maternal care, aggression, predator-prey interaction) • long-term (several and hours and longer) restraint • administration of chemical/pharmacological substances resulting in disturbed physiological state (toxicological tests) • inducing radiation sickness • the use of Freunds complete adjuvant • causing anatomical or physiological disturbances resulting in pain and suffering • exposure to harmful, unavoidable stimuli • recovery and maintaining the life of mutilated animals after invasive surgery (performed under anaesthesia); • induction of fatal disease (eg radiation sickness, infections, inherited and cancerous diseases) with a humane endpoint; • causing disruption of sensorimotor organisation and severe psychological disorders (eg by permanent isolation of the young from their mothers) • the exposure to harmful, highly stressful stimuli without any possibility of escape

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
Poland (cont)	Grade X, proceedings causing extreme suffering	Experiments causing acute pain, close to or beyond the tolerance threshold in unanaesthetised animals, involving procedures causing severe suffering and pain or leading to death preceded by a long-lasting period of suffering due to surgery, exposure to harmful chemical substances or other procedures which impair physiological systems Examples: • using muscle relaxants without anaesthesia • killing methods not recommended by National Ethical Committee • exposure to severe stress or shock • major surgery causing substantial changes in the body • causing wounds and burns without anaesthesia • causing death by poisoning (eg strychnine), by dehydration or starvation or by the effect of temperature and pressure • induction of acute psychosis (eg exposing restrained animals to acute stress or replacement of the mother by a punishing phantom) and agonistic behaviour leading to injury and death Grade X procedures of invasiveness may be carried out only in exceptional cases if it is allowed or required by special regulations
Sweden *Severity classification*	Minor, experiments where animals are not at risk of being exposed to more than minor pain and/or other discomfort	Examples of experiments of mild severity: • Restraint for physical examination • Gavage feeding • Skin tests with non-irritating substances • Injections with non-irritating substances • Blood sampling from peripheral blood vessels (artery and vein) • Sedation or anaesthesia to facilitate handling • Experiments under anaesthesia in animals that are euthanised without recovery • Anaesthesia for minor, superficial surgical interventions with recovery • Species-specific mild food or water deprivation • Euthanasia using accepted methods/techniques • Tail tissue sampling from rodents
	Moderate, experiments where animals are not at risk of being exposed to more than moderate pain and/or other discomfort which under normal circumstances can be fully handled from the animal protection viewpoint by users with appropriate knowledge and techniques	Examples of experiments with moderate severity: • Permanent catheterisation of peripheral or central blood vessels (artery and vein) • Larger surgical interventions under anaesthesia in the abdominal cavity, thorax, skeleton or central nervous system with recovery and appropriate postoperative care and analgesia • Blood sampling through retro-orbital puncture in small rodents • Injections with irritating substances • Housing in metabolic cages • Disease models where animals are subject to pain or suffering which is minimised with appropriate care • Immunisation using Freund's complete adjuvant • Behaviour studies with harmful stimuli and the possibility to escape • Toxicity tests without lethal endpoint • Combined, repeated interventions or interventions of long duration, each of which of minor severity

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
Sweden (cont)	Considerable: Experiments where animals are at risk of being subject to considerable pain and/or other discomfort which cannot always be eliminated even with appropriate knowledge and techniques.	Examples of experiments of considerable severity: • Larger surgical interventions under anaesthesia without adequate postoperative analgesia • Shock, burn and radiation experiments where animals may be subject to considerable pain or suffering • Tumour biology experiments where the tumour growth must be followed until advanced stages • Infectious biology experiments, including experiments for development, testing and control of vaccines, where animals can be expected to be seriously ill or with lethal endpoint • Behaviour experiments with harmful stimuli without the possibility to escape or with considerable restraint • Toxicity tests with lethal endpoint • Antibody production using the ascites method • Serious hypoxia to induce central nervous system injury • Induction of serious disease conditions without alleviating treatment • Combined, repeated interventions or interventions of long duration, each of which of moderate severity
Switzerland *Degrees of severity (stress categories)*	No stress: severity grade 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. Examples in veterinary practice: withdrawal of blood samples for diagnostic purposes; subcutaneous injection of a drug Animals which were not exposed to any pain, suffering, injury or severe fear through interventions and procedures for experimental purposes and whose general well being was not significantly impaired. As a general rule, animals used for experiments which do not require authorisation or control animals on which no interventions with adverse effects were conducted, will be counted under this heading
	Mild stress: severity grade 1	Interventions and manipulations in animals for experimental purposes which subject the animals to a brief episode of mild stress (pain or injury) Examples in veterinary practice: injection of a drug requiring the use of restraint; castration of male animals under anaesthesia Animals which suffered a minor, temporary adverse effect (pain or injury) caused by intervention and procedures for experimental purposes. For example, animals which were killed in pre-terminal narcosis or rabbits which were immunised without the use of Freund's adjuvant will be counted here
	Moderate stress: severity grade 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) Examples in veterinary practice: surgical treatment of a single leg-bone fracture; castration of female animals Animals which suffered a medium-severe, short-term or slight but medium to long-term adverse effect (pain, suffering or injury, severe fear or substantial impairment of general well being) caused by intervention and procedures for experimental purposes. For example, animals in which electrodes were implanted in the brain or those which underwent adrenalecto-my will be counted here

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press

Appendix 1 (cont)

Country and name of system	Description of severity level	Examples, interpretation and/or guidance
United Kingdom *Protocol severity limits and Severity banding of projects*	Mild severity	(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 severity	(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 postoperative 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 severity	(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 antimicrobial 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
	Unclassified	(Protocols) performed entirely under general anaesthesia, from which the animals does not recover consciousness. This includes the preparation and use of decerebated animals.

https://doi.org/10.1017/S0962728600002761 Published online by Cambridge University Press