Practical use of distress scoring systems in the application of humane endpoints

M. H. Lloyd & S. E. Wolfensohn
Veterinary Services, University of Oxford, c/o University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK

Summary

This paper deals with the use of pain and distress scoring systems for two experimental models: elastase-induced emphysema in rats and allergic encephalomyelitis in mice. The selected protocols showed that even very simple systems can be used successfully, giving consistent results and permitting humane endpoints to be defined.

The assessment of pain and distress in animals is problematic since it can be subjective and based upon anthropomorphic assumptions that things which are distressing for humans will also be distressing for animals, which is not necessarily the case (Flecknell 1994). This may lead to a tendency to overestimate the pain and suffering experienced by animals in some situations, and underestimate it in others. Objective methods of assessing pain and distress in animals are valuable for several reasons: they allow potential refinements and new research techniques to be evaluated critically; they allow real judgments to be made about the need for, and the efficacy of, analgesics; and they allow for the implementation of humane endpoints. Several distress scoring systems have been used in many of our research projects, but they are hard to validate, and there are a number of difficulties which we and others have found (Beynen *et al*. 1987, 1988).

Selection and sensitivity of parameters

Many research projects produce changes in just one organ system, and a general scoring system is not sensitive enough to pick up specific indicators of ill health in these cases. To overcome this, there are two choices:

(1) to modify the system to include a very large checklist of all the possible clinical parameters which may change. This produces a system which checks many parameters, but since in many cases only a few parameters will be discriminatory the score may still be low overall even when the severity of signs in a particular category is high; and
(2) to modify the list to include only those specific changes which are anticipated in the particular project, i.e. adapting the sheet for each project.

Between assessor variations

Since scoring is based on the subjective judgments of individuals, it is possible that two people with different backgrounds and experiences will score the same animal differently. This can be avoided by having several observers score the animal at each point and average the scores, or by having all assessments performed by just one or a few experienced persons. Veterinarians, investigators and animal care staff can, with experience, become consistent in their judgments (Orlans 1996) and thus reduce this variation between assessors.

In practice, scoring systems can be valuable tools in keeping the costs to animals in biomedical research to a minimum, and this paper describes how such systems have been used and modified in two different research projects to overcome these problems. Both projects were authorized under the UK Animals (Scientific Procedures) Act 1986, and work was carried out at Bayer PLC, Pharmaceutical Division, Research Department, Stoke Court, Stoke Poges, Slough SL2 4LY, UK and at Celltech Therapeutics Ltd, Bath Road, Slough SL1 4EN, UK.

Use of a distress scoring sheet in an elastase-induced emphysema model in the rat

Barrier-reared Sprague Dawley rats from Harlan UK (Shaws Farm, Bicester, Oxon) were group housed on solid floors in a part barrier facility, screened frequently for microbiological contaminants. Rats had emphysema induced by the intranasal or intratracheal instillation of up to 600 i.u. porcine pancreatic elastase under general anaesthesia (Finlay *et al.* 1996). Control animals received vehicle only. Emphysema was allowed to develop for a period of up to 8 weeks and various treatment regimens and controls were used to alter the development of the disease. At the end of the period all remaining animals were humanely killed and their lungs examined.

During an initial pilot study, the animals were scored every 4 h for 48 h following the instillation of porcine pancreatic elastase or vehicle, using a distress scoring system based on that proposed by Morton and Griffiths (1985). This system assesses pain and distress in animals objectively using a number of physiological and behavioural parameters: body weight (usually reflecting food and water intake), appearance, clinical signs, natural behaviour, and provoked behaviour. A numerical score is allocated to each group, 0 for normal, 3 for grossly abnormal. If 3 is scored more than once, an extra point is awarded for each three, giving a maximum score of 20. This ensures that severe signs in one category are not cancelled out by mild signs in another and gives extra weighting for more severe clinical signs, but reduces the impact of a 'false-positive' result in a single category: it is unlikely that such a false-positive would occur twice. We prepared a simple distress scoring sheet using this system, including criteria to assist with scoring, which could be used easily by investigators or animal care staff (Wolfensohn & Lloyd 1998) (Fig 1). Scoring was carried out by only two or three researchers, who initially scored the animals together, so they became very consistent in their assessments.

During the study, the researchers found that occasionally animals would die peracutely, usually within one hour of the procedure, due to massive lung haemorrhage. Such animals showed few clinical signs prior to death and even careful observation and scoring did not help identify these animals. Some animals would become moderately distressed and then recover, others continued deteriorating slowly until death. Animals which reached a score of 14 did not recover and invariably died. No animals were found to deteriorate after 24 h, and most animals began to recover after 18 h. Control animals rarely scored more than 4. For subsequent studies the period of scoring every 4 h was reduced to 24 h, although animals continued to be monitored closely for the duration of the study, and any animal reaching a score of 14 was humanely killed.

Although the animals exhibited many general signs of distress, the main effects of the experimental treatment were respiratory, and it was felt that the score sheet lacked sensitivity. In an effort to refine the humane endpoint, the distress score sheet was modified for use during the critical 24 h period, to include more details of respiratory parameters (Fig 2). Clinical signs which were felt to be particularly severe, e.g. dehydration, were given extra weighting, and animals were awarded a score out of 19. This specific score sheet was used alongside the general one initially to make sure it worked, then it was used alone. Animals consistently scored higher using the specific sheet despite the maximum possible score being lower (Fig 3), and the range of scores recorded by individuals was greater. The

Parameter	Animal ID	Score	Date/Time	Date/Time
Appearance	Normal	0		
	General lack of grooming	1		
	Coat staring, ocular and nasal discharges	2		
	Piloerection, hunched up	3		
Food and water intake	Normal	0		
	Uncertain: body weight $<5\%$	1		
	Intake: body weight 0–15%	2		
	No food or water intake	3		
Clinical signs	Normal T, cardiac and respiratory rates	0		
	Slight changes	1		
	T $\pm$ 1°C, C/R rates $\uparrow$ 30%	2		
	T $\pm$ 2°C, C/R rates $\uparrow$ 50% or very	3		
Natural behaviour	Normal	0		
	Minor changes	1		
	Less mobile and alert, isolated	2		
	Vocalization, self mutilation, restless or still	3		
Provoked behaviour	Normal	0		
	Minor depression or exaggerated response	1		
	Moderate change in expected behaviour	2		
	Reacts violently, or very weak and precomatose	3		
Score	If you have scored a 3 more than once, score an extra point for each 3	2–5		
	Total	0–20		

0–4 = Normal
5–9 = Monitor carefully, consider analgesics
10–14 = Suffering, provide relief, observe regularly. Seek second opinion from named animal care and welfare officer and/or named veterinary surgeon. Consider termination
15–20 = Severe pain; does your experimental protocol need rethinking?

Fig 1 General distress scoring sheet

Parameter	Animal ID	Score	Date/Time	Date/Time
Appearance	Normal	0		
	General lack of grooming	1		
	Piloerection, fresh ocular and nasal discharges	2		
	Piloerection, hunched up	3		
	Above and eyes half closed	4		
Natural behaviour	Normal	0		
	Minor changes	1		
	Less mobile and isolated, but alert	2		
	Restless or very still, not alert	3		
Hydration status	Normal	0		
	Abnormal skin pinch test	5		
Clinical signs	Normal respiratory rate and pattern	0		
	Slight changes, increased rate only	1		
	Increased rate with abdominal breathing	2		
	Decreased rate with abdominal breathing	3		
	Marked abdominal breathing and cyanosis	4		
Provoked behaviour	Normal	0		
	Minor depression or exaggerated response	1		
	Moderate change in expected behaviour	2		
	Very weak and precomatose	3		
	Total	0–19		

Fig 2 Modified distress scoring sheet to include respiratory parameters

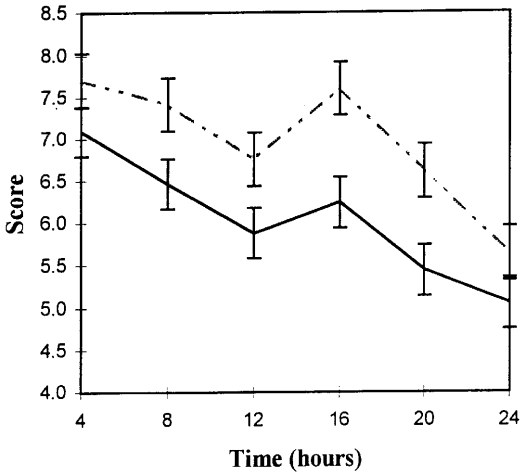

Fig 3 Mean distress scores of rats using old and new systems together. Distress scores determined using both sheets concurrently in a group of 30 rats given PPE. The modified sheet produced consistently higher scores and a slightly greater range indicating increased sensitivity.
(— old; – · · – New)

Table 1 Clinical signs of disease in the experimental allergic encephalomyelitis model

Grade 0	Normal
Grade 1	Loss of righting reflex/tail paralysis
Grade 2	Incomplete hind limb paralysis
Grade 3	Complete hind or front limb paralysis
Grade 4	Moribund

After Smith *et al*. (1994)

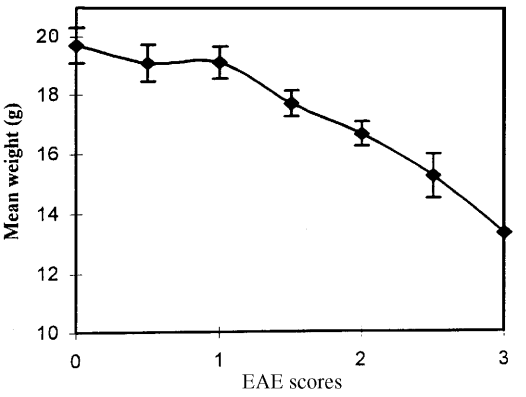

Fig 4 Weights of animals developing clinical signs of EAE. EAE scores greater than 1 were associated with a significant fall in body weight; $n = 14$, $P < 0.0001$

specific sheet certainly appeared to be more sensitive, and was felt by the researchers to be easier to use than the general one, but neither sheet was able to identify the animals which died per-acutely, so some further modification is needed. However, both sheets were used successfully to fix a humane endpoint.

Use of distress scoring in a mouse model of experimental allergic encephalomyelitis

Barrier-reared Biozzi mice from Harlan UK (Shaws Farm, Bicester, Oxon) were housed in shoebox cages in a conventional unit. In a pilot study, 16 mice were immunized with spinal cord homogenate in Freund's Complete Adjuvant injected intradermally into the base of the tail under general anaesthesia to induce experimental allergic encephalomyelitis (EAE) (Smith *et al*. 1994). Animals were weighed daily, their clinical condition monitored, and the progression of the neurological disease monitored. After 24 days all animals were euthanized. The neurological deficits exhibited by the animals were examined to determine an EAE score, according to the grading system shown in Table 1.

Experimental allergic encephalomyelitis is an autoimmune paralytic and inflammatory disease of the central nervous system (CNS) used as a model for demyelinating disorders such as multiple sclerosis. The disease follows a fairly predictable course, and may be chronic and relapsing allowing long-term benefits to be studied. Affected animals that do not die usually recover. These features make defining a humane endpoint difficult, since animals may sometimes develop quite severe disease and then recover fully. Several different scoring systems exist for assessing the neurological effects of EAE and defining an experimental endpoint (King *et al*. 1983, Brosnan *et al*. 1985, Sedgwick *et al*. 1987, Smith *et al*. 1994) but they do not take into account the effects of the disease on the general health and well-being of the animal. In this study, the relationship between body weights and EAE scores was analysed using

the GLM procedure on SAS (SAS Institute Inc., (1989) SAS/STAT User's Guide, Version 6, 4th edn, Cary, NC, USA). The normality of all residuals was tested. Neurological deficits appeared from day 10–15, and once animals reached an EAE score of more than 1, this was associated with a significant loss in body weight ($P < 0.0001$, see Fig 4), although animals often recovered neurological function and regained body weight subsequently. Since even minor neurological deficits had adverse effects on the health of the animals, this suggested that the neurological scoring system was too specific, and needed to be combined with an assessment of general distress in order to determine a humane endpoint for subsequent studies. A humane endpoint was fixed using both general and specific parameters as follows. Animals showing one or more of the following were humanely killed.

- Weight loss: greater than 30%.
- Clinical signs: not eating or drinking for more that 24 h.
- EAE score: greater than 4 for more than 24 h, complete hind limb paralysis or quadriparesis.

Conclusion

Reliable methods for scoring animal distress provide experimenters with the ability to make quantitative assessments of animal pain and suffering, and to assess the efficacy of any proposed refinement. If any system is to work, it has to be consistent and easy to use, and be specific and sensitive enough to detect subtle changes in an animal's well-being. General systems may have to be made more specific to increase sensitivity, and specific systems used for determining experimental endpoints combined with assessment of general parameters for determining humane endpoints. These protocols show that even very simple systems can be used successfully, giving consistent results and permitting humane endpoints to be defined.

Acknowledgments Thanks are due to Manuel Berdoy for help with data analysis, Barbara Briggs for her work on emphysema, and Sarah Lane for information about EAE.

References

Beynen AC, Baumans V, Bertens APMG, Havenaar R, Hesp APM, van Zutphen LFM (1987) Assessment of discomfort in gallstone bearing mice: a practical example of the problems encountered in an attempt to recognize discomfort in laboratory animals. *Laboratory Animals* **21**, 35–42

Beynen AC, Baumans V, Bertens APMG, *et al.* (1988) Assessment of discomfort in rats with hepatomegaly. *Laboratory Animals* **22**, 320–5

Brosnan JV, Fellowes R, Craggs RI, King RHM, Bowley TJ, Thomas PK (1985) Changes in lymphocyte subsets during the course of experimental allergic neuritis. *Brain* **108**, 315–34

Finlay GA, O Donnell MD, O Connor CM, Hayes JP, FitzGerald MX (1996) Elastin and collagen remodelling in emphysema; a scanning electron microscope study. *American Journal of Pathology* **49**, 1405–15

Flecknell PA (1994) Refinement of animal use–assessment and alleviation of pain and distress. *Laboratory Animals* **28**, 222–31

King RHM, Craggs RI, Gross MLP, Tompkins C, Thomas PK (1983) Suppression of experimental allergic neuritis by Cyclosporin-A. *Acta Neuropathologica, Berlin* **59**, 262–8

Morton DB, Griffiths PHM (1985) Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Veterinary Record* **116**, 431–6

Orlans FB (1996) Invasiveness scales for animal pain and distress. *Lab Animal* **25**, 23–5

Sedgwick J, Brostoff S, Mason D (1987) Experimental allergic encephalomyelitis in the absence of a classical delayed-type hypersensitivity reaction. *Journal of Experimental Medicine* **165**, 1058–75

Smith RM, Morgan A, Wraith DC (1994) Anti-class II MHC antibodies prevent and treat EAE without APC depletion. *Immunology* **83**, 1–8

Wolfensohn SE, Lloyd MH (1998) *A Handbook of Laboratory Animal Management and Welfare*, 2nd edn. Oxford: Blackwell Science Publications