

Relating criteria for humane endpoints to objectives

D. J. Fry

Home Office, PO Box 31, Shrewsbury SY3 7WN, UK

Summary

An important determinant for terminating a live animal experiment is the achievement of the scientific objective. When this has been met any further animal suffering produces no extra scientific value and so should be avoided. To do so requires that objectives are clearly specified, and that there is a way of identifying when they have been reached. Good use can be made of pilot studies using a few animals to find not only the likely progress of an experimental approach, but also pointers to when the objective has been attained. For the same animal model, the criteria for the humane endpoint can differ according to the objective: thus for a rodent arthritis model the endpoint can be at a lower level of suffering when the objective is to test an agent for efficacy in prophylaxis than when effectiveness against established disease is to be determined. These arguments will be illustrated by data provided by experimenters who have been persuaded by them and, using different approaches, have identified stages at which endpoints can be drawn for animal work of various types.

It has been recognized for many years that when the scientific objective of an experiment is attained, or for some reason the objective is clearly unattainable, any further risk to the animals involved of pain, distress or harm is not offset by the obtaining of further useful information, and the experiment should end. However, the extent to which this can be put into practice depends crucially on setting clear objectives and reliable determinants of when they are met. In the UK this approach has been promoted for at least 10 years by Home Office inspectors, and as they are required by law to provide advice on all programmes of scientific work involving live vertebrates with potential to cause pain, suffering, distress or lasting harm, this small group is in a good position to assess how widespread is the approach. It seems, in some areas, to have become normal practice. For example, in testing the efficacy of potential rodenticides, where there is a clear objective—namely to identify compounds that will kill rodent pests—the scientific end is achieved when death can be reliably predicted, and the experimental animals are

routinely humanely killed when such predictive signs are seen.

However, it is less evident in academic science, and in order to see how prevalent this approach is in the culture of academic institutions, the 28 programmes of work submitted for project licences in a 6-month period from one area of the country were assessed. These were from different scientists and six different academic institutions. Only two gave either clear objectives from which objective-achieved endpoints could be set, or a design strategy which would enable such endpoints to be identified. Less than a third had aims sufficiently clear for such objectives to be stated when individual experiments were being designed.

These submissions cover programmes of work and the picture might be different at the individual experiment level. However, some indication that it is not comes from reports requested on experiments undertaken: not one of the last 18 reports submitted gave the purpose of the experiment(s) or considered the extent to which the effects on the animals, described in the report, were

necessary in order to meet the objective of the work. The tendency is to use a standard experimental set-up to generate results, with general humane endpoints based on the level of suffering considered acceptable for the type of work, rather than to clarify the purpose of a particular experiment and to tailor the endpoints accordingly.

The purpose of this presentation is to highlight this problem, which is unlikely to be unique to British scientists, and discuss some types of work where the endpoints set for particular animal models could vary according to the objectives of the work.

Timed endpoints—serial measurement

Where the time course of an effect has been determined, there is scope for setting as a humane endpoint the time by which a biologically significant change should have occurred. Fig 1 illustrates this for tumour growth studies. It shows a typical time course and magnitude of response in experiments where a set tumour mass is implanted subcutaneously in an immunodeficient mouse and subsequent growth determined from external measurements. The prospective anti-cancer agent is given with the tumour implantation, the controls receiving vehicle only. The graph shows mean values at each time point for 8 animals, and is representative of the way several groups present their results. After day 5, control and experimental values are recorded as significantly different. For reasons of confidentiality, Fig 1 does not show an actual study, but there are many that conform to this pattern.

As with other experimental neoplasia studies, general limiting endpoints—tumour size, haemorrhage, ulceration, interference with mobility, etc.—can be set. However, in this type of experiment the experimental measurement is taken in the living animal and the results can be easily calculated before the next measurement is taken, giving scope to set an endpoint based on the results obtained.

Thus if the objective is simply to determine whether the untried agent can reduce tumour growth *in vivo*, as when compounds showing promise *in vitro* are screened, then

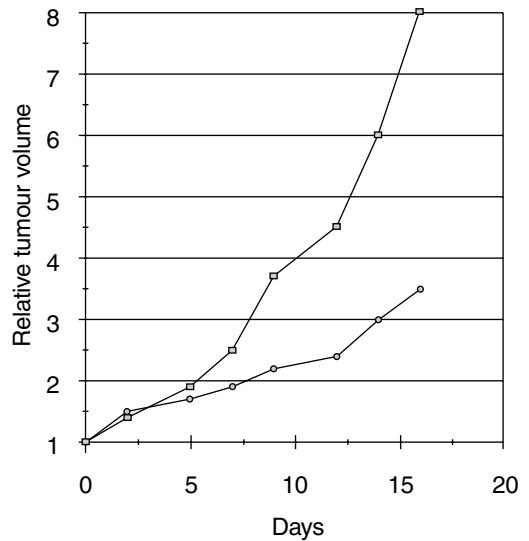


Fig 1 Effect of a potential therapeutic on xenograft tumour growth. —□— Control; —○— Anti-cancer agent

the endpoint could be: 'when mean control tumour volume has increased over four times or after two successive measurements have shown significant difference between experimental and control groups, whichever is the sooner'. On that basis, the study shown in Fig 1 went on many days too long.

If, however, the objective is to see how the agent affects the time course of tumour growth, the endpoint would be when sufficient measurements have been taken to establish the pattern of tumour growth in the experimental group, or the general severity endpoint is reached by either group, whichever is the sooner. With that objective a study of the type shown in Fig 1 might need to continue for some days longer in the experimental group, e.g. if there are indications that the effect is not persistent and valuable additional information is likely to be obtained from later time points, so that a higher level of overall severity can be justified.

Similar arguments would apply where measurements are obtained postmortem from animals killed at successive time points if the results are readily calculated, but careful design would be needed to avoid wastage of animals when the timed endpoint was reached.

Fixed time endpoints

Where set times must be specified beforehand, as in serial studies depending on post-mortem or *ex vivo* material when results are not quickly calculated, a pilot study using only a few animals, or the first experiment in a series, can be used to set future timed endpoints. Fig 2 shows results from a study using 1000 chickens both to provide commercial levels of virus antigen from a new strain used for the first time (where there was a persuasive case that *in vivo* production was necessary), and to determine the time to humanely kill the birds in later runs.

Figures for 750 birds are shown in Fig 2. The upper graph shows how many birds were slightly sick or worse at each time point. When birds were judged more than slightly sick they were humanely killed, and the lower graph shows the cumulative total of birds thus culled (about 10% of the total sick at each time point). A further 250 birds (hence the 1000) were used to provide estimates of virus antigen produced at various time points.

The objective was to obtain sufficient antigen for commercial production of an efficacious vaccine. The estimates showed that that level of antigen was present by 50 h after infection. After this, higher levels could be obtained but the additional material did not make sufficient difference to efficacy to warrant the extra suffering, so in future all birds were humanely killed at or before 50 h (or if birds were getting sick more quickly than usual, at the time when over 10% of those infected showed signs of sickness).

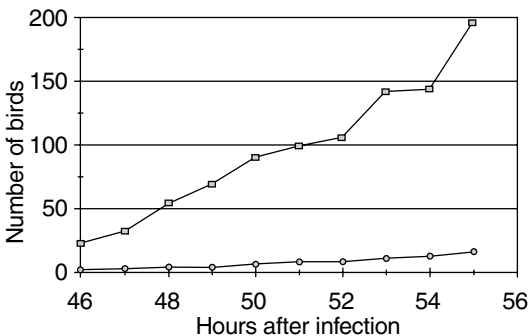


Fig 2 Time course of birds becoming sick
 —□— Slightly sick and culled; —○— Culled

This preliminary experiment leading to a simple timed endpoint, cut the number of birds falling ill in a large-scale production by half, and a key element was obtaining from the company a clear objective, and agreement on when it was achieved.

The use of pilot or preliminary experiments to determine suitable humane endpoints is applicable to a wide range of studies, but how often is this approach actually employed?

Different objectives—different endpoints

Although it is helpful to define suitable endpoints for an animal model, these have to be placed in the context of the studies on that model. This has become evident in discussion of appropriate endpoints for an antigen-induced arthritis model, in which, following a sensitization protocol, the arthritic changes are induced by injection of antigen into one knee joint, with the other receiving saline injection and used as the control. The scientific measure used is the difference in diameter between the antigen-injected and the other knee joint. The time course of the response for one set of rats and one batch of antigen is shown in Fig 3. The pattern is typical but the peak effect can be a day or so later and the size of the residual swelling can also vary between sets of rats and batches of antigen.

When the objective is to determine whether an agent has a prophylactic effect

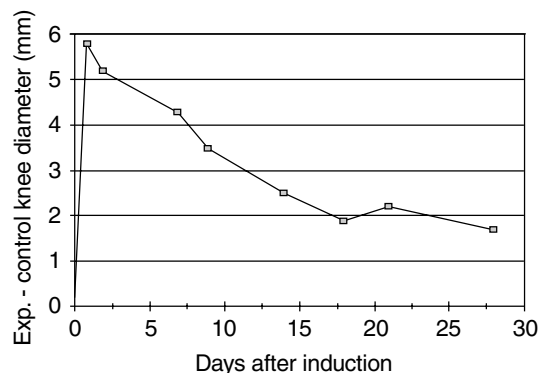


Fig 3 Time course of knee swelling in antigen-induced monoarthritis (mBSA induction, rat)

against the acute swelling, an endpoint of say 3 mm difference between the two knee diameters can be set at one day after induction, since within such endpoints the efficacy of an agent should be clear, even if the onset is at the slow end of the range.

However, this group is principally interested in developing preparations that will be effective against established arthritis, and for that would start testing on day 7, when previous histological studies have shown that the acute inflammatory response has changed to a chronic inflammation. In order to minimize severity, one objective would be to see whether the use of centrally acting analgesics, which would alleviate the pain of the acute phase, would affect the results. For this an endpoint has to be set which allows for the effects on the group not receiving analgesic, but ensures that animals showing excessive swelling, or vocalization on contact, or restricted mobility, for example, would be humanely killed. This more severe endpoint would apply only to sufficient animals for there to be confidence that analgesia would not prejudice the tests.

With analgesia used thereafter, and the animals not normally showing any limitation in mobility or signs of pain or deteriorating condition, these can be used as endpoint criteria, since any animal showing them is likely to have some abnormality which could bias the test results.

As with the tumour growth illustration, the different objectives of determining whether a preparation is effective, and determining the time course of an effect, can be met by different periods of observation. For the former, a cut-off at day 15 should be sufficient to pick out effective agents, but for

showing a lasting effect the full period of 28 days is likely to be required.

Conclusion

In these examples there was scope for reducing cumulative severity by tailoring the endpoints to the particular purpose of the studies, but it required pressure from an outside authority (in this case a Home Office inspector) before the scientists concerned thought of doing so. The surveys above indicate that experimenters do not routinely think along these lines, and that this is an area for reducing avoidable suffering that needs to be more widely considered. The concept of setting objective-achieved endpoints is not novel, and it is usually readily appreciated when the points are raised, but some stimulus is needed to bring it to mind when a scientist is planning a set of animal experiments. Some key questions are as follows:

- What is the precise objective of the experiment/set of experiments?
- Are there reliable indicators of when it is achieved, or will clearly not be achieved?
- If not, could pilot experiments provide such indicators?

There are some experiments where the purpose is to generate data from which hypotheses can be formulated, and retrospective analysis is appropriate, but most are to test ideas that can be clearly stated and are amenable to setting objective-achieved endpoints. So greater promotion of this approach among groups of scientists, those involved in regulating or advising on animal work, and journal editors could give considerable benefit in refinement of animal experiments.