Working Group Report

Application of the Three Rs to challenge assays used in vaccine testing: Tenth report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement

Maggy Jenningsa,*, David B. Mortonb, Emmanuelle Chartonc, Jane Coopera, Coenraad Hendriksen d, Stella Martin e, Michael C. Pearce f, Scott Price g, Keith Redhead e, Nick Reed h, Hugh Simmons h, Stephen Spencer e, Hilary Willingale

a RSPCA, Research Animals Department, Wilberforce Way, Southwater, Horsham, West Sussex RH13 9RS, UK
b School of Biosciences, Emeritus Professor of Biomedical Science & Ethics, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
c Council of Europe, European Directorate for the Quality of Medicines and HealthCare (EDQM), 7 Allée Kastner, 67081 Strasbourg, France
dNetherlands Vaccin Institute (NVI), P.O. Box 457, 3720 AL Bilthoven, The Netherlands
ePfizer Animal Health, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK
f Veterinary Medicines Directorate, Woodham Lane, New Haw, Surrey KT15, 3LS, UK
gVeterinary Laboratories Agency, Woodham Lane, New Haw, Addlestone, Surrey KT15 3NB, UK
hHome Office, ASPD, Mailpoint 21B, 4th Floor Seacole SW, 2 Marsham Street, London SW1P 4DF, UK

ARTICLE INFO

Article history:
Received 8 April 2010
Accepted 8 April 2010

Keywords:
Vaccines
Three Rs
Humane endpoints
Guideline harmonisation

ABSTRACT

This report aims to facilitate the implementation of the Three Rs (reduction, refinement and replacement) in the testing of vaccines for regulatory and other purposes. The focus is predominantly on identification of reduction and refinement opportunities in batch potency testing but the principles described are widely applicable to other situations that involve experimental infections of animals. The report should also help to interpret the requirements of the European Pharmacopoeia with regard to the use of alternative tests, humane endpoints and other refinements. Two specific worked examples, for batch potency testing of Clostridium chauvoei and canine leptospira, with recommendations for harmonisation of international test requirements for these and other vaccines, are provided as appendices online.

1. Introduction and aims of the report

Testing of veterinary vaccines is a significant area of experimental animal use within European and other countries with a vaccine manufacturing industry. The need to apply the Three Rs of reduction, refinement and replacement [1] to the testing of vaccines for both veterinary and human use, and opportunities to do so, have been discussed by Castle [2], Metz et al. [3], Halder et al. [4] and Hendriksen [5], amongst others, and application of the Three Rs to veterinary vaccines specifically has been reviewed most recently by Cooper and Jennings [6]. The tests that use most animals and cause most suffering, and in which animals may die, are the challenge assays used to assess batch potency of certain vaccines. These involve the induction of disease by infecting animals with pathogens or exposing them to associated toxins. For the tests to be valid, some animals have to show typical signs of disease. This inevitably causes considerable suffering to unvaccinated control animals and to those vaccinated with low vaccine doses which do not protect against disease.

A batch potency test is performed routinely on every batch of a vaccine. The purpose is to demonstrate that the batch to be marketed will be at least as potent as the batch of minimum potency/titre shown to give satisfactory results in key efficacy studies. The test can also provide information on the consistency of the

E-mail address: mjennings@rspca.org.uk (M. Jennings).

* Corresponding author.

doi:10.1016/j.biologicals.2010.04.004
manufacturing process. Alternatives to replace batch potency tests involving challenge are urgently needed on animal welfare grounds. Research has already produced accepted replacements for several batch potency tests including those for tetanus toxoid, inactivated Newcastle disease and swine erysipelas vaccines [7–9]. Test development, validation and acceptance is a very lengthy process, typically taking more than 10 years [10], and in many cases it presents a considerable scientific challenge. A more immediate positive impact on animal welfare could be achieved by refinement of batch potency assays to reduce the suffering involved. Reduction in the numbers of animals that have to be used should also be possible.

Refinement has a beneficial impact on science as well as welfare, since the welfare of experimental animals affects the quality of scientific data produced. For example, when animals are stressed they may appear outwardly ‘normal’, but are likely to experience subtle, yet uncontrolled physiological and biochemical changes that can influence the variability, reliability and reproducibility of data collected [11,12]. Reducing animal suffering can improve the reliability and reproducibility of data and this can reduce the likelihood that tests will have to be repeated. The variability of results is also reduced, allowing smaller group sizes to be used, saving time, resources and animals. Such savings can offset the cost of developing, validating and implementing refinement, and the cost of obtaining a variation to the standard test monograph.

There is considerable scope for applying the Three Rs to batch potency testing and the report focuses on this. However, the principles described can be applied in all situations that require experimental infections including model development, proof-of-concept, challenge validation, challenge passage, and full scale efficacy studies.

2. Regulatory requirements and the Three Rs

There are two types of national and international legislation that impact on the use of animals for vaccine testing:

(i) regulation of the use of animals in scientific procedures;
(ii) regulation of the production and marketing of vaccines.

The first requires animal use and suffering in scientific procedures to be minimised. The second defines the necessary quality, safety and efficacy test requirements for vaccine products, many of which necessitate the use of animals in such procedures. To reduce the conflict between these different regulations, the principles of humane science, enshrined in the regulation of scientific procedures on animals, need to be carried through into all relevant vaccine testing regulations. However, the latter are not always straightforward to interpret, particularly with respect to the flexibility within the animal test requirements. A clearer understanding of the legislation, and of the roles and responsibilities of the various regulatory and other bodies involved, can help identify where and how implementation of the Three Rs can be achieved. To assist with this, the European situation is described below, with reference to other countries where particularly relevant.

2.1. Regulation of scientific procedures on animals

Scientific procedures that may cause animals pain, suffering or distress are currently regulated in the European Union (EU) by Directive 86/609/EEC [12]. This states that experiments “shall not be performed if another scientifically satisfactory method of obtaining the result sought, not entailing the use of an animal, is reasonably and practically available.” Furthermore, when choosing test methods, “those which use the minimum number of animals, involve animals with the lowest degree of neurophysiological sensitivity, cause the least pain, suffering, distress or lasting harm... shall be selected.” Similar statements are made within Council of Europe Convention ETS 123 [14] on experimental animals, which applies across a wider range of countries. These principles are translated into national legislation, for example in the UK by the requirements of the Animals (Scientific Procedures) Act 1986 [15]. In practice, this means that in all EU countries there is a legal requirement to apply the Three Rs to any regulated work involving the use of animals including the testing of vaccines.

Countries outside Europe have different legislation, implemented in different ways, but the Three Rs are internationally supported and integral to the animal protection legislation of most countries that have legislation controlling animal use, for example, Canada, Australia and the USA. Some countries also require the likely benefits of research proposals to be weighed against the potential harms to the animals used. Manufacturers therefore need to weigh the need for, and benefits of, carrying out tests to assure that their vaccines are safe and efficacious, against the harms to any animals used to provide those assurances.

2.2. Regulation of the production and marketing of vaccines

For a vaccine to be accepted for marketing in Europe, manufacturers must submit an application for a Marketing Authorisation (MA). The requirements for, and guidance on, data that should be generated and presented in the dossier that supports the MA are set out in a number of documents (see Section 2.2.1 and Appendix 3) including EU directives and the European Pharmacopoeia (Ph. Eur.) [16].

The requirements have to be interpreted according to their legal force, and whether they are relevant and can be logically applied to a particular product. In so doing, it is necessary to consider the nature of the method of manufacture and the starting materials used, and to consider carefully what data or routine testing requirements should be applied to the product to ensure batch consistency. This means that, depending on how the product is manufactured, additional tests (animal and/or non-animal) may have to be performed to confirm consistency of production. Equally, some tests (animal or non-animal) may not need to be performed.

2.2.1. The European Pharmacopoeia

European Pharmacopoeia monographs set mandatory requirements for products that are on the market in countries that are signatories to the European Pharmacopoeia Convention.3 The Ph. Eur. comprises:

- General notices which clarify how to interpret various sections of the pharmacopoeia and individual monographs;
- General chapters which give methods that are used for multiple monographs;
- General monographs which are applicable to a wide range of products; and
- Specific monographs which are applicable to products of a particular type (as stated in the Ph. Eur. Definition section).

Requirements that are common across a range of products are not usually repeated in each product-specific monograph. The specific monographs must therefore be read in conjunction with

---

3 To date, there are 36 signatories to the Convention, in addition to the European Union itself.
the relevant general monographs, for example monograph 0062, Vaccines for Veterinary Use, and with an understanding of the contents of the General Notices of the Ph. Eur. (Chapter 1.1).

If there is no specific monograph for a product, it will need to comply with the general pharmacopoeial standards. In the case of a veterinary vaccine, it will therefore have to comply with monograph 0062, Vaccines for Veterinary Use.

2.2.1.1. The Three Rs in the European Pharmacopoeia. Since the signing in 1986 of the European Convention on the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes, the Ph. Eur. has carried out a programme of work to reduce, refine and remove the use of animals in its texts [17,18]. The Ph. Eur. actively encourages the implementation of the Three Rs through the following statements.

General notices:

“......This does not imply that performance of all the tests in a monograph is necessarily a prerequisite for a manufacturer in assessing compliance with the Pharmacopoeia before release of a product. The manufacturer may obtain assurance that a product is of Pharmacopoeia quality from data derived, for example, from validation studies of the manufacturing process and from in-process controls.”

“......The tests and assays described are the official methods upon which the standards of the Pharmacopoeia are based. With the agreement of the competent authority, alternative methods of analysis may be used for control purposes, provided that the methods used enable an unequivocal decision to be made as to whether compliance with the standards of the monographs would be achieved if the official methods were used. In the event of doubt or dispute, the methods of analysis of the Pharmacopoeia are alone authoritative.”

General Monograph on Vaccines for Veterinary Use (0062):

“......In accordance with the General Notices, alternative test methods may be used to demonstrate compliance with the monograph and the use of such tests is particularly encouraged when this leads to replacement or reduction of animal use or reduction of suffering”

This means that there is flexibility in demonstrating compliance with the requirements of the Ph. Eur.; manufacturers do not have to perform all test methods, but should demonstrate that the product would comply if subjected to these tests. Since manufacturers and the Competent Authorities have a duty to ensure that animal usage is kept to a minimum and animal health and welfare legislation is upheld, both need to critically assess whether a test is necessary, and whether reduction and refinement options could be applied. Competent Authorities should also encourage the development of alternative methods which lead to a reduction, refinement or replacement of animal use, and data from these should be accepted when suitably validated.

2.2.1.1.1. Humane endpoints. The general monograph 0062 Vaccines for Veterinary Use, also makes it clear that the principle of humane endpoints (see Section 3.6 of this report) must be applied in the tests conducted:

“......In accordance with the provisions of the European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes, tests must be carried out in such a way as to use the minimum number of animals and to cause the least pain, suffering, distress or lasting harm. The criteria for judging tests in monographs must be applied in light of this. For example, if it is indicated that an animal is considered to be positive, infected etc. when typical clinical signs occur then as soon as it is clear that the result will not be affected the animal in question shall be either humanely killed or given suitable treatment to prevent unnecessary suffering.

Specific definitions of humane endpoints are not usually included in test descriptions within specific monographs. However, an example of where this has been done is supplement 6.4 of the Ph. Eur., which contains revisions of the monographs Rabies vaccine for human use prepared in cell cultures (0216) and Rabies vaccine (inactivated) for veterinary use (0451). These include a section on alternative endpoints describing typical clinical signs to be noted and a typical score chart. The analyst is expected to ‘validate’ the endpoint for a sufficient number of batches by scoring the test in the usual way, but also using the alternative endpoint. Since the test is carried out routinely for release of batches of vaccine, manufacturers have the opportunity to do the alternative scoring without having to do additional tests for validation. This approach is the one that is most likely to lead to the use of alternative endpoints for other vaccines.

2.2.1.2. The Three Rs and batch potency testing. It is made clear in the General Notices, General Monographs and Specific Monographs of the Ph. Eur. that alternative batch potency tests from the examples given in the Specific Monographs are acceptable. Any such test has to be validated and give assurance to the manufacturers and the Competent Authority that the batch is of suitable quality and meets the requirements of the Ph. Eur. and the standards agreed through the MA. Only in rare cases of doubt or dispute over whether the batch of product is of pharmacopoeial quality, does the test have to be performed exactly as described in the monograph.

2.2.2. Replacement alternatives in batch potency testing

Currently, a promising approach to the replacement of individual batch potency tests is in vitro antigen quality and quantification assays as part of the consistency approach [19]. This requires proof of consistency of production, and aims to demonstrate that each new batch of vaccine produced is of a similar quality to a vaccine batch of the same provenance, and is of proven efficacy and safety. The strategy involves demonstrating consistency using a battery of physico-chemical, immuno-chemical and in vitro methods [20].

An example of where this has been successful is the antigen quantification assay as a replacement for the rabies (NIH) potency test. This is accepted by the regulatory authorities on condition of demonstrated validity. Generally, antigen quantification tests are based on the use of monoclonal antibodies quantifying relevant antigen epitopes; they have limited value when vaccines are adjuvanted. However, even then it can be possible, as there is now an accepted in vitro potency test for oil-adjuvanted Newcastle disease vaccines [8].

2.2.3. Options for validation of alternative methods

The most appropriate approach to validation of alternative potency test methods will depend on how broad the applicability is for the alternative assay. Two possible options are:

(i) For a manufacturer-specific alternative method which is to form part of the dossier for a manufacturer-specific vaccine, a suitable in-house validation study will probably be acceptable to the relevant regulatory authorities. To enhance the chances of such methods being accepted by regulators within Europe,
scientific advice can be sought from the European Medicines Agency (EMEA) or from National Competent Authorities, something that is regularly encouraged in the UK.

(ii) For a vaccine-specific alternative method where it is intended that the alternative test be accepted into a regulatory monograph or guideline, more formal validation rather than just a successful in-house study is likely to be required. Organisations such as the European Centre for the Validation of Alternative Methods (ECVAM) and the Biological Standardsisation Programme of the European Directorate for the Quality of Medicines and HealthCare (EDQM) work towards validating alternative methods to be included in regulatory guidelines and monographs. There is a defined procedure for such validation, which is outlined in Appendix 3 (Section A3.2.).

There is no guarantee that an alternative method will be accepted and manufacturers have to pay a fee for each licence variation for the introduction of any change in the assays they submit. Together, these factors create a strong disincentive for industry to explore and validate alternative methods. To facilitate the acceptance of alternatives and avoid conflicts of opinion between different Competent Authorities, manufacturers should communicate with all relevant parties at an early stage. (The roles, responsibilities and interactions of the various institutions and regulatory authorities involved in the regulation of vaccines and development of test guidelines within Europe are illustrated in Appendix 3, Fig. 1) It would also help if Competent Authorities would waive the fee for licence variations that implement the Three Rs, as done by the UK Veterinary Medicines Directorate (VMD).

2.2.4. Recent three Rs initiatives

There are two recent Three Rs initiatives within Europe. The Committee for Medicinal Products for Veterinary Use (CVMP) is exploring ways to establish stronger EU ties with EDQM and ECVAM and to consider how advice concerning Three Rs issues for the development, manufacture and testing of veterinary medicinal products could be provided. Phased assessment of the registration dossier is also being considered. This would help to minimise the number of animals used in clinical safety and efficacy studies, and also avoid animals being used to validate a method if there are indications at an early stage that the method may not be accepted for regulatory purposes. This initiative is at an early stage and any significant change to the way in which veterinary pharmaceuticals are authorised may have to wait for the next review of the legislation, due in 2014.

The European Partnership for Alternative Approaches to Animal Testing (EPAA) is a joint Three Rs initiative from the European Commission, which was formed in 2005. It consists of individual companies and trade associations representing the chemicals, soaps and detergents, cosmetics, crop protection, pharmaceutical, bio-industry and animal health sectors, together with the following Commission services: Directorate Generals Enterprise and Industry, Research and Development, SANCO, Environment, and the Joint Research Centre (ECVAM). In 2010, the EPAA in collaboration with ECVAM organised a workshop on the consistency approach to vaccine quality control and the JWGR encourages them to pursue further vaccine-related Three Rs activities as well.

2.3. Regulations in other countries

Other countries and market areas where vaccines are manufactured either have their own pharmacopoeias and/or regulatory requirements, for example, the US Code of Federal Regulations Title 9 (9CFR) and Japanese Pharmacopoeia (Japanese Veterinary Biologicals Product Standards), or they base their requirements on those of other countries. As these regulations have developed independently, there are some differences in the detail of test methods, test requirements, animal numbers and endpoints. Harmonisation of international guidelines, taking the most refined methods as the standard, would have a very positive impact on animal welfare and obviate the need for repetition of tests for different markets. Appendices 1 and 2 illustrate how this could be done for Clostridium chauvoei and canine leptospira vaccines and provide a template for applying the principles to other vaccines.

Although full international harmonisation may take a long time, harmonisation even between some of the major markets, for example US, EU and Japan, would facilitate mutual acceptance of data and bring real Three Rs benefits.

3. Practical application of refinement and reduction

To maximise application of refinement and reduction in batch potency and other tests, it is important to critically review every aspect of experimental design and test procedures, and every aspect of animals’ life-time experiences (see Table 1). This includes factors such as housing and care, transport, and handling in addition to the potentially adverse effects of vaccination and challenge.

3.1. Materials and equipment

Good preparation and storage of materials will maximise efficiency and reproducibility of tests and minimise the number of animals that have to be used. Minimising the need to passage challenge materials and to check virulence in animals also helps to reduce animal use.

The same principle applies to equipment and data management systems. Adequate maintenance of these increases the chance that test results are reliable, accurate and precise, and reduces the likelihood of repeat testing being needed, again minimising animal use. Requirements are stipulated in good manufacturing practice (GMP) standards and these should be applied by vaccine manufacturers and others involved indirectly, such as animal suppliers and contract research organisations.

Where possible, equipment, materials and services should be obtained from sources which operate to recognised quality assured standards such as the ISO standards. Alternatively, some small scale suppliers may be willing to participate in specific quality assurance (QA) audits.

3.1.1. Preparation, maintenance and storage of vaccine challenge materials and reagents

Whenever possible, bacterial and viral challenge strains should not be maintained by routine in vivo passage. As well as the additional use of animals, this can lead to increased variability in the biological nature and virulence of the strain, and each passage increases the likelihood of contamination with extraneous agents which could influence the outcome of subsequent challenges. Ideally, a master seed/working seed system should be used to maintain challenge stocks to minimise variation.

The best storage conditions for each toxin, or bacterial or viral strain, should be ascertained from the available literature and storage viability tested. Factors to be considered include: the form of storage (for example, chilled liquid, frozen or freeze-dried); storage temperature; toxin concentration; whether (for some bacteria) they should be stored as vegetative cells or spores, and the cell density; types of excipients and percentage inclusion; and the method of freezing and thawing.
Table 1
Factors to consider in relation to refinement and reduction.

| Materials and equipment                      | Use of appropriate well-maintained and, where relevant, sterile equipment. |
| C ventures for selection of animals          | Careful preparation, maintenance and storage of materials; and consideration of their nature (e.g., irritancy, tissue compatibility, sterility, temperature) when administered. |
| Animal husbandry and care                    | Selection of an appropriate species and strain of animals with consideration of other factors such as age, weight and sex. |
| Numbers of animals and statistical design    | Use of a consistent source of high health status animals. |
| Administration of substances                | Implementation of animal housing and care that takes into account the physical and behavioural needs of the animals as well as the need to be able to monitor them without too much disturbance. |
| Humane endpoints                             | Use of sympathetic handling and restraint procedures. |
| Monitoring animals                           | Application of an appropriate experimental and statistical design with well justified numbers of animals. |
| Staff                                        | Timing of challenge to facilitate monitoring (in relation to the animals’ time budget® and staff availability). |
|                                               | Use of the most refined methods including: |
|                                               | - use of an appropriate gauge needle (i.e. the smallest gauge appropriate to the species, route and substance administered); |
|                                               | - selection of the least invasive route likely to cause least trauma and pain to the animals; |
|                                               | - selection of an appropriate and least harmful site/s for administration and suitable preparation of the site to facilitate accurate administration first time; |
|                                               | - use of aseptic technique; |
|                                               | - exploration of opportunities to administer reduced volumes. |
|                                               | - Description and implementation of humane endpoints to minimise level and duration of suffering. |
|                                               | - Careful, regular and timely monitoring of animals for adverse effects including those associated with the administration procedure itself. |
|                                               | - Use of anaesthetic and analgesics to reduce pain. |
|                                               | - Sufficient, appropriately trained and competent staff who can implement all of the above. |

* The relative amounts of time that an animal spends performing different behaviours.

The effects of storage on the activity of toxins or the viability/virulence of bacterial or viral challenges should be checked; frequently initially, then less so if they appear to be stable in storage. The results should be recorded and monitored to identify any trends.

### 3.1.2. Challenge validation

All challenge assays involve the use of unvaccinated animals to confirm the toxicity or virulence of the challenge material used. If challenge tests are performed relatively frequently, i.e. once or more each year, no other validation of the material should be needed. If challenge tests are less frequent, it may be necessary to validate the challenge annually. For most viral challenges, plaque assay counts on a suitable cell line can be used for validation of test material. For bacterial challenges, viable counts on agar will have to be supported by less frequent in vivo challenges. For toxins, if possible, an in vitro method such as a cell line assay should be used for routine validation and animals should only be used when part of the challenge test.

The performance of the challenge material during testing should be recorded and checked to detect any changes in toxicity or virulence. Where a reduction in virulence of the master seed is detected, it may be necessary to passage the bacterium or virus through the relevant animal species to retain the virulence of the challenge. The way this is done should be optimised to keep the numbers of animals and level of suffering to a minimum.

### 3.1.3. Calibration and maintenance of equipment

Test equipment and data recording and management systems need to be validated and maintained properly, so that data are measured and recorded accurately and precisely. High quality records of such maintenance need to be kept and audited regularly. Adequate records should also be kept for each batch test run, for inspection in the event of aberrant test results.

Facilities should validate new equipment and carry out regular calibrations and conformity checks on existing equipment. Ideally, this should be linked to national or international standards. Subcontractors used to maintain equipment should also have some form of quality control system that gives confidence that they are operating to suitable standards. Ideally both subcontractors and suppliers should be audited to ensure they comply with the requirements of GMP and QA.

Examples of equipment that should be regularly inspected and calibrated include: measuring devices such as micro-pipettes, thermometers and telemetric implants; storage equipment such as incubators, freezers and refrigerators; and environmental monitoring equipment. Building management systems, including temperature, humidity and lighting are also important as they can have a profound effect on the welfare of animals.

Staff should be adequately trained, experienced and competent to conduct measurements and undertake data recording and management.

These comments apply equally to test equipment used in the development and use of alternative in vitro tests, for example ELISA readers and associated software programmes and data storage systems.

### 3.2. Selection of animals

Species, strain, source: selection of the appropriate species and strain of animal, suitably susceptible to the pathogen or toxin and responsive to the test vaccine, is important for the generation of reliable and reproducible scientific data. The strain and source of animals should be consistent to reduce any possibility that variability of results could be related to genetic factors. The use of a regular and reliable source will avoid any need to assess new strains and use more animals.

Age and weight: experimental groups should comprise animals of similar age and weight at critical time points, for example at vaccination and challenge. This allows consistency in the monitoring of progression of clinical signs, which can make it easier to define earlier endpoints and minimise variability.

Sex: the sex of animals is unlikely to affect the data obtained, but may have an indirect affect on animal welfare if it affects the ability...
to group-house social animals. Thus, for species where it is easier to group-house females than males there is an animal welfare benefit to using females. However, the overall benefit is lost, if as a consequence, males are killed and wasted. In short-term studies, or in studies with juveniles, it should be possible to house animals in mixed sex groups.

Health status: animals should be of high health status and, in the case of the common laboratory species, preferably specific pathogen free (SPF), to ensure that experimental results are not influenced or complicated by sub-clinical background pathology, and to assure consistency.

Good suppliers of the common laboratory species monitor the health status of their breeding colonies. This involves periodic evaluation of a representative sample of the colony using serological, microbiological and other pathological tests to ensure that the colony is free of relevant specified pathogens and to claim SPF status. Specified pathogens, methods of detection and frequency of sampling are described in recommendations made by the Federation of European Laboratory Animal Science Associations (FELASA) [21–23].

Farm animal species from different agricultural sources are likely to vary in health status. It is therefore good practice for all animals in each test to be bought from the same source, preferably one that is closed to entry of new stock. This will ensure consistency in test results and reduce the potential for spread of endemic disease. Disease can be especially severe and problematic in young animals (for example, the occurrence of diarrhoea and pneumonia in calves), particularly if animals from different sources are mixed. Vaccination and preventive treatments on farm prior to delivery may therefore be necessary.

It is important to check the veterinary health plans in place at the source, the health monitoring strategies, and whether both the source as a whole and the individual animals are free from disease. Inspection by the testing establishment veterinarian is advisable as is investigation by a veterinary pathologist of any culls or losses.

Any use of specially prepared animals, such as colostrum-deprived or gnotobiotic, requires special husbandry measures such as irradiated food or antibiotics in drinking water. The health and welfare of such animals will require close monitoring. Their use should be considered carefully to assess whether data obtained are relevant and valid with respect to the expected vaccine efficacy in normal animals, particularly in the case of ruminants where much of the development of the immune system occurs after birth, and the role of the gut associated lymphoid tissue is very important.

Fish are increasingly used in vaccine studies. Generally, these animals are not reared in SPF conditions. Consequently, indicators of abnormal behaviour and/or clinical signs of disease need to be defined and their health and welfare status should be monitored on arrival and during housing in the laboratory.

Acclimatisation: a period of acclimatisation to the new environment is recommended for all animals as this allows them to recover fully from any adverse effects of transport, mixing of individuals, and changes in diet, environment and handlers. Acclimatisation also allows for the incubation of diseases, most of which are likely to show within 3 to 4 days of arrival, and are often complicated by a change or variation in diet. This is particularly important where animals are to join an established group.

An acclimatisation period of around one week allows animals that are unsuitable for any reason to be identified before being used in a study. This period is generally considered within industry to be the minimum necessary for animals used for regulatory testing (although it may be precluded in some instances, for example in the evaluation of vaccine safety in day-old chicks). It is also the time recommended by Obernier and Baldwin [24] in their review of the literature on the physiological acclimatisation period for laboratory animals after transport.

Allocation: animals should be randomly assigned to test groups to minimise selection bias.

3.3. Animal husbandry and care

The many factors to be addressed in the husbandry and care of animals are well documented in various guidelines [14,25]. Species dependent requirements for facilities, environment and care described in such guidelines should be set as the minimum standard, with opportunities for improvement and further refinement encouraged, explored and implemented.

The principles of optimal care and husbandry apply to all uses of animals for scientific purpose and are not specific to testing vaccines. However, there are points that require particular attention in vaccine challenge tests, especially where features of housing and care and the animals’ response to these (for example, levels of activity, social interactions, interaction with objects in the environment) are used to help observe and implement humane endpoints. Factors to consider include:

Social housing: animals should be group-housed unless there are compelling scientific or veterinary reasons not to do so. If individual animals in a test group might die or reach a humane endpoint at different times, for example in control groups in challenge tests, care should be taken to ensure that a single remaining animal is not left alone for any longer than absolutely essential.

Housing: housing should satisfy animals’ physiological, behavioural and psychological needs and provide comfort and security. Disturbance should be kept to a minimum to avoid affecting the animals’ behaviour, particularly when this is used as a determinant of humane endpoints. However, this has to be balanced with the need to facilitate the regular and detailed observation of animals that is essential in challenge tests.

Flooring and bedding: solid flooring together with appropriate bedding material is particularly important for animals who may experience pain or ill-health during the course of challenge infection. Bedding material should be selected that improves welfare, but does not make animals too difficult to observe. It may even be possible to use materials that help to reveal an endpoint, for example white shredded paper to show haematuria.

Physical environment: modifying environmental conditions, such as ambient temperature, may help improve the welfare of sick animals. Modification of lighting patterns, i.e. use of reverse-lighting for species that are nocturnally active, enables animals to be observed when at their most active during normal staff working hours. This allows more meaningful observation of their behaviour and removes the need to disturb them during their inactive period.

Diet and water: diets should be consistent and of assured quality as changes may have an impact within or between tests; water supply should be similarly monitored and its quality maintained. Any recommended supplementation or medication of feeds should be assessed for potential impact on test outcomes.

Animals suffering from certain harmful effects of a challenge may show reduced food or water intake. It may be possible to counter this by the addition of ingredients such as fresh vegetables, such as sweet-corn or carrot, fruit, such as apples or grapes, and grains, and to increase water intake by moistening foodstuffs.

Enrichment: environmental enrichment encourages expression of a wider range of behaviours than is possible in a barren cage. This increases the scope for scoring relevant clinical signs because animals often change the level of interaction with enrichment items when they experience adverse effects, and this may occur before other detectable clinical signs are present. For example,
Environmental enrichment must be appropriate for the species, strain, type of accommodation, and nature of the test protocol. Enrichment should be designed and monitored to ensure it has a beneficial effect for the animals. Experienced staff should be able to judge improvements in the condition or demeanour of animals, or any reduction in adverse behaviour such as fighting, hair plucking or tail chewing. It is also important to show that there are no adverse effects on the outcome of tests.

3.4. Numbers of animals and statistical design

The design of regulatory vaccine tests should ensure that results provide an accurate assessment of test material using the lowest number of animals possible and causing the minimum pain, distress and suffering to animals.

3.4.1. Test numbers in regulatory requirements

There is considerable disparity regarding group sizes for particular species for individual vaccines specified in the Ph. Eur. and its equivalent in other countries. Tests that use species such as birds, mice or fish typically specify much larger numbers of animals than those that use dogs, cats or cattle. For example, in the 9CFR requirements for Pasteurella multocida vaccines, the potency of bovine vaccine in calves requires 10 vaccinates and 5 controls, whereas for the avian vaccine in chickens, 20 vaccinates and 10 controls are specified. There is also disparity between the numbers of animals required in European and American regulations for tests on the same vaccine (see examples in Appendices 1 and 2).

It is difficult to see how these disparities can be scientifically justified and it is important to explore opportunities to reduce numbers and ensure group sizes have a good scientific and statistical basis. It is quite probable that the populations of rodents, birds and fish from which test groups are derived are more homogenous than those of cats, dogs or farm animals, so it may be argued that the group sizes required for the former should be no larger than for the latter. If the numbers of birds, mice and fish required could be reduced to the numbers required for tests using dogs, cats, or farm animals for example, a considerable reduction in the total numbers of animals used would be obtained. Further reductions in the numbers are possible if more than one batch of vaccine is tested simultaneously with each test sharing a single control group.

A re-evaluation of the current test requirements and harmonisation to achieve a consistent minimum number of animals both within and between individual pharmacopoeias is therefore recommended by the JWGR, together with a simple fast-track process to get revised test requirements implemented quickly.

3.4.2. Statistical design

Aside from the regulatory requirements, it is good practice to re-examine the statistical design of all tests carried out in-house at regular intervals as this provides the historical data to allow more accurate statistical analysis and assessment of the need for controls. Determination of the number of animals required for a test system is dependent on many factors, including the predictive value of test systems and the type of measurements made i.e. whether these are continuous or binomial. It requires specialist knowledge of the relevant statistical and epidemiological principles involved and it is recommended that an expert in statistics is consulted.

3.4.2.1. Predictive value of test systems. The number of animals to be used in a test system is driven by the required accuracy of the test result, expressed as the predictive value. This concept applies at the level of the individual animal in the test, and to the test system overall. There are two types of predictive value:

- **Positive predictive value:** the probability that if a test result is positive, the material tested is truly positive.
- **Negative predictive value:** the probability that if a test result is negative, the material tested is truly negative.

Predictive values are partly dependent on the sensitivity and specificity of the individual test or test system:

- **Sensitivity:** the probability that, if a test material is positive, a test or test system will return a positive result.
- **Specificity:** the probability that, if a test material is negative, a test or test system will return a negative result.

The positive predictive value of a test or test system can be increased by increasing the test or test system specificity, or by increasing the probability of the test material being truly positive prior to testing. The negative predictive value of a test or test system result can be increased by increasing the test or test system sensitivity or increasing the probability of the test material being truly negative prior to testing.

The sensitivity and specificity of an individual test or test system can often be modified by changing cut-off points or endpoints. Predictive values are also dependent on the probability of the test material being truly negative or positive known as the priori. This value can often be estimated based on previous experience with the test material, or on information about the test material.

3.4.2.2. Type of measurements. Continuous data: when the test outcome of interest is continuous data, appropriate sample sizes can be determined using well established methods and software requiring input of:

- the probability of making type 1 and type 2 errors;
- the minimum difference of interest between a treatment group and control group;
- the variability of the data.

**Binomial data:** the situation with binomial data for example, yes/no or positive/negative results for individual animals, is different and determination of the number of test animals needed requires several steps:

**Step 1** Determine the sensitivity and specificity of the individual animal test.
**Step 2** Define the threshold result for the test system to pass or fail the test article. The threshold result can be set so that the proportion of treated animals and the proportion of control animals that reach the test endpoint are fixed and do not vary. Alternatively, the threshold result can be the ratio of

---

4 Cut-off point: an arbitrary dividing line between $+$ and $-$, or between responder and non-responder.
5 Type 1 error: rejection of null hypothesis when it is actually true. Type 2 error: acceptance of null hypothesis when it is false.
the proportion of treated animals reaching the test endpoint, divided by the proportion of control animals which reach the test endpoint, known as the relative risk or risk ratio.

Step 3 With the sensitivity and specificity of a test for the individual animal known, calculate the probability of a truly positive animal returning a positive result and a truly negative animal returning a negative result. The binomial distribution can then be used to partition treated and control animals into positive and negative results.

Step 4 Check compliance of the results with the outcome specifications declared in Step 2. If compliance is not achieved, either the number of animals can be increased, or the sensitivity and specificity of the test for the individual animal may need to be reviewed. This step will need to be repeated until compliance can be achieved.

3.5. Administration of test material

Administration procedures themselves have the potential to cause adverse effects in addition to any effects from the challenge material. Opportunities for refining such procedures are described in an earlier JWC Report [26] and other useful references are Diehl et al. [27] and Wolfensohn and Lloyd [28].

Aseptic practice should be observed to limit the potential for local infection at injection sites and contamination of test materials, which may interfere with the immune response. Other factors including needle size, differing tissue sensitivity in different areas of the body, and temperature and composition of the inoculum, can all influence the degree of pain and discomfort experienced by an animal and should all be considered and optimised (see Table 1).

Some vaccines, for example certain clostridial vaccines, can cause transitory discomfort and possible longer-term irritation, self-mutilation and abscessation. In such cases, the deposition of smaller volumes over more than one site should be considered. This reduces the likelihood of discomfort and tissue pressure necrosis, and will allow the presentation of a greater surface area of animal/vaccine interface to elicit an immune response. However, these advantages need to be weighed against the discomfort of additional needle sticks.

3.6. Humane endpoints

Endpoints in challenge assays are traditionally some of the most severe; the performance of a test vaccine is assessed against claims for that vaccine and if it is claimed to protect against a lethal challenge, the endpoint of the test system may be death in challenged animals. Such endpoints are likely to be of the most substantial severity and result in the greatest pain and distress. They are clearly not humane, and are also unpleasant for animal care staff to observe.

Death as an endpoint can be strongly questioned on scientific as well as animal welfare grounds. It is an indirect scientific measure, because animals in vaccine tests frequently die from secondary and tertiary effects rather than the infection challenge itself. Deaths of animals challenged with microorganisms are usually the result of secondary or tertiary effects of infection. For example, animals given a neurotropic microorganism such as rabies virus might show secondary neurotoxic effects such as convulsions, muscle weakness and incoordination. This may prevent them from reaching water, in which case they may die from the tertiary effects of dehydration and heart failure due to increased viscosity of the blood, neither of which is directly related to the infectious agent.

An alternative approach, preferable for both animal welfare and scientific reasons, is to define an accurate scientific endpoint that is a surrogate measure of the claim for the test material, using the occurrence of one or more clinical signs known to precede death. This results in far less suffering and is the principle underlying the development of humane endpoints, defined by the OECD [29] as “the earliest indicator in an animal experiment of substantial pain, distress, suffering, or impending death”. The aim is to stop an experiment at the earliest point at which the scientific objectives are achieved in order to avoid animals suffering unnecessarily.

The need to develop more humane endpoints in vaccine potency testing has been argued for many years (e.g. [2,31]). Valid humane endpoints need to be defined for every challenge potency test in the Ph. Eur., and any requirement to exceed a humane endpoint and cause significant suffering, or death, should be stringently questioned.

3.6.1. Defining humane endpoints for batch potency tests

The objective of a challenge assay used for batch potency testing is to demonstrate with a reasonable level of confidence that (i) the challenge is sufficiently to cause disease, and (ii) the vaccine protects against disease. The standard tests require that unvaccinated animals and/or animals immunised with unprotective vaccines show specific signs of challenge or die within the observation period. More humane endpoints can be developed by identifying reliable, reproducible, predictive clinical signs, or non-clinical physiological markers such as levels of protective antibodies, that provide sufficient information about the potency of the vaccine and its ability to protect animals at specific doses. These indicators should then be validated against the traditional endpoint. The endpoints selected need to have the highest reduction in suffering with respect to both its duration and intensity [31].

3.6.1.1. Identifying suitable signs. The first step is to carefully observe animals through all stages in the development of disease following infection, recording any clinical signs that occur. These may be physiological, such as weight loss or diarrhoea, or behavioural, including changes in the animals’ social behaviour, use of enrichment items such as nesting material or chew blocks, or food and water consumption.

Some signs, for example weight loss and body temperature, are easily measurable and can be objectively assessed. Others, such as convulsions, dyspnoea, piloerection and social behaviours may not be so easily measurable, but they can still provide reliable information and be related to the scientific outcome measure. It may be possible to assign numerical ‘scores’ to these (see below). Where measurement of physiological parameters requires sampling of body fluids such as blood or urine (for example, for assay of haematological and biochemical indicators of organ function or failure, or monitoring blood leukocyte count for decline as an indication of infection) there are additional factors to consider. Firstly, if such measures are used to help predict and refine endpoints, then the speed of analysis is important. Secondly, invasive monitoring techniques can in themselves cause suffering or stress. The benefits that they provide in terms of improved animal monitoring and implementation of more humane endpoints therefore needs to be weighed against their potential for causing additional stress and discomfort, particularly in smaller species. This also applies to the measurement of physiological parameters using implanted telemetry devices, transponders or chips. The implantation procedures can cause suffering, so the harms and benefits of implanting devices should be carefully thought through and the most refined approach used [32].

This approach has also been reported for the rodent protection test used to confirm in vivo efficacy of novel antiviral, antibacterial and antifungal agents [30].
The reliability of each sign needs to be assessed in relation to its utility for reliably predicting the required scientific outcomes and whether it might give false positive or false negative information. Using a combination of signs may improve predictability, especially if behavioural and physiological data can be correlated. For example, when testing the potency of a whole cell pertussis vaccine, scoring clinical signs together with weight loss and reduced body temperature signal an inevitable deterioration more reliably than the individual signs alone. The selected humane endpoints are then validated by observing whether or not the animals continue to morbidity or death and the data are analysed to ensure they are sufficiently robust from a statistical point of view.

The success of a humane endpoint trial can be measured in animal welfare terms by calculating the reduction in the number of days over which the animals suffer. For example, the whole cell pertussis potency assay above uses a mouse body temperature of 34.5 °C as the humane endpoint. This occurs at a mean of 2 days (range 1–7 days) earlier than the normal experimental endpoints of death or severe suffering requiring euthanasia [31].

3.6.1.2. Monitoring animals and avoiding observer variation. The next step is to consider how animals should be monitored and how variation between human observers can be minimised. This is important because differences in observational skills and interpretation can lead to relevant clinical signs being discarded and endpoints being applied inconsistently.

Animals should be assessed at times and frequencies that will best help to identify the early onset of harmful effects. The timing of challenge and subsequent observations therefore needs to take into account the following factors.

3.6.1.2.1. Normal circadian activity patterns of the animals. Important clinical signs are more likely to be observed if animals are monitored when they are likely to be most active, or performing specific behaviours such as feeding. The optimal time for assessing animals will depend on the species and strain-specific behaviour patterns and time budgets, in conjunction with husbandry routines. It is best to observe them at approximately the same times each day.

The frequency and duration of direct cage-side observations should be carefully considered, with input from animal care staff. The use of behaviour recognition software can improve measurement of animal behaviour in a variety of settings, and can make 24 h monitoring and analysis a possibility.

3.6.1.2.2. Predicted onset and duration of clinical signs of disease. The observations made when defining humane endpoints should enable the time between the challenge and the onset of adverse effects to be predicted. If animals display clinical signs and then recover, it should also be possible to determine approximately how long animals are likely to experience adverse effects. Monitoring protocols should then be set up that ensure animals are observed when they are especially likely to be suffering. This information can also be used to time the challenge so that these critical periods coincide with maximum staff availability.

3.6.1.2.3. Staff availability. There needs to be sufficient staff available to monitor all animals individually. A challenge should be administered so that the critical endpoint phase is expected to occur during the working day and not during a weekend or holiday. The person responsible for the test should be present or rapidly available during this period. Ideally, the study director should be contactable, but if this is not possible, a trained member of staff must be available to make decisions to help ensure that humane endpoints are effectively implemented (see also Section 3.7).

3.6.1.2.4. Avoiding observer variation. Observer variation can be reduced by good training and teamwork, and by ensuring that clinical signs are clearly described and accurately recorded in welfare assessment sheets that are tailored to the type of test and the type of product. The use of video to record clinical signs and aspects of scoring is an invaluable aid, both for helping uniformity of scoring between personnel and for training.

Structured welfare assessment sheets that include a list of agreed clinical signs [33] should be available to all relevant staff and preferably be included in standard operating procedures (SOPs). They should be discussed with all concerned, including the principle investigator and veterinary and animal care staff, before starting work, and be reviewed at regular intervals to see whether signs have been effectively predicted and whether any humane endpoint might be brought forward even further. They should also be considered by ethical review processes or animal care and use committees, so that others with additional expertise may contribute.

One approach to reducing variation when staff use welfare assessment sheets is to simply record signs as being either present or absent, with no quantitative judgements. If necessary, an allowance can also be made for uncertainty (a record such as ‘possibly present’), that would highlight the need for closer observation, for example where loose stools precede diarrhoea, or slight coughing precedes pneumonia. Signs can also be given numerical scores, but this involves making value judgements and needs careful definition and agreement to ensure consistency between observers.

The chosen method of data recording needs to take into account the clinical signs to be recorded, observer consistency, ease of use, clarity and effectiveness. Examples of commonly used recording schemes include: paper records, computer collation of data and direct entry into hand-held devices such as palm tops. All systems for data recording should be accurate, contemporaneous and attributable. Results should be reproducible and data should be stored in a form suitable for archiving and which is traceable.

3.6.1.2.5. Avoiding observer bias. Observer bias occurs where observers are aware of different treatments and may, unwittingly, make biased decisions based upon that knowledge. It is important therefore, wherever possible, to randomise the distribution of cages in a study so that the observer is unaware of the treatment group in each cage. This ‘blinded’ approach has the added benefit that any environmental impact associated with cage location, for example high or low light levels, should be neutralised across the different groups.

3.6.2. Use of analgesia and anaesthesia

The use of suitable analgesics and, where appropriate, anaesthetics, should be considered when it is expected and predicted that the challenge process could be painful at any stage. It is theoretically possible to anaesthetise an animal throughout an experiment of short duration, but many infections have incubation periods of several days and take time to produce clinical signs. In any case, anaesthesia may not be appropriate as some of the more common clinical signs are related more to distress than pain, for example through inability to reach the water bottle due to muscle weakness.

Analgesics may be useful for some types of challenge that cause pain at certain stages (for example, with Cl. chauvoei or rabies). Their value will depend on the time course of the infection and analgesic regimes should take account of the expected period of discomfort and pain after the challenge has been administered, with doses repeated as necessary. If analgesia is used, the type of analgesic, dose and treatment regime need to be empirically examined to assess not only whether it is effective and beneficial, but also to confirm that it does not materially interfere with the course of the disease or identification of humane endpoints. Similarly, analgesia should not have a significant effect on the robustness of the test’s ability to discriminate between effective and non-effective batches of vaccine.
In summary, the decision to use analgesics and anaesthetics, and even to provide palliative care, should depend on validated studies and critical observation of whether this actually reduces suffering for the animal, either in duration or intensity.

### 3.7. Staff issues

A team of well trained, highly competent staff is integral to implementation of the Three Rs. All relevant staff should have the opportunity to be involved in the planning, development and review of tests, procedures and systems and be encouraged to contribute ideas on the Three Rs and animal welfare. Input and feedback between all those involved will create an environment where opportunities for implementing the Three Rs are actively sought throughout the process of vaccine research, development and testing.

#### 3.7.1. Training

Training is fundamental to the provision of a high standard of animal welfare, implementation of the Three Rs, delivery of high quality, valid test results, and compliance with legal and product registration requirements. In all aspects of vaccine testing it is therefore essential that staff should be well trained and competent in the procedures and activities they perform. This will be well worth the resources required i.e. financial costs, time, and the expertise of trainers and supervisors.

All staff should have a personal development plan tailored to their job which should include a training record, and regular review of training needs. Training should cover both knowledge and practical skills. Staff may not be allowed to carry out certain regulated procedures before they are fully authorised, so much of their initial training will be theoretical. Theory is important, but can only provide an introduction to the set of practical skills required, so this needs to be followed by a period of supervision or apprenticeship until competence in all necessary practical skills is attained. Competency will require formal assessment, allowing that different individuals will not all require the same sets of skills, and will take different times to attain them.

The learning process should not be considered to end once initial training and formal assessment has occurred, a personal licence or authorisation has been gained, or upon the completion of the supervision period. Ongoing performance and results should be monitored for consistency and for conformity with expected standards and results. Continuing staff development is also important and personal development requirements should be considered for all individuals. The provision of continuous professional development (CPD) or refresher training will encourage the maintenance of current competencies, and the addition of new ones [34].

There are several useful publications that describe the competencies, and associated training and supervision needs of staff involved in the care and use of animals generally [35–37] and many countries have a system of training provision linked to assured quality standards.

#### 3.7.1.1. Vaccine testing—specific training

As with the husbandry and care of animals, there are certain topics that are particularly relevant to vaccine challenge assays and these are likely to require more specific in-house training. These topics are listed in Table 2 alongside the categories of staff for which they are particularly important.

#### 3.7.2. Additional staff issues

There are additional staff related issues that can play a part in refinement which should be considered when planning and conducting a programme of vaccine testing.

Staff numbers should be maintained at levels that allow adequate time for test programmes to be conducted correctly. This is essential for the maintenance of scientific quality, animal welfare, health and safety of staff, and compliance with legal requirements. It will also avoid the unnecessary animal use that results from repetition of unsatisfactory tests. Consistent staffing will give similar benefits, leading to less variability in observations, procedures and results, and possibly less stress to test animals.

When scheduling challenge tests, critical time points should coincide with the availability of staff best able to assess health, determine humane endpoints, and take decisions on the course of the test. For example, the time at which a challenge is administered should be planned such that any predictable adverse effects occur during normal working hours. If this is not possible, additional staffing must be provided to ensure frequent

<table>
<thead>
<tr>
<th>Training requirement/topics</th>
<th>Relevant Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope, purpose and requirements of relevant regulations, guidelines and monographs and their application to the conduct of vaccine testing challenge assays</td>
<td>Animal technicians</td>
</tr>
<tr>
<td>Husbandry and care of the species of animals used, including appropriate environmental enrichment</td>
<td>✓</td>
</tr>
<tr>
<td>Normal species specific behaviour and likely strain differences</td>
<td>✓</td>
</tr>
<tr>
<td>Correct application of most refined techniques for administration of test substances, challenge organisms and analgesics by different routes</td>
<td>✓</td>
</tr>
<tr>
<td>Possible adverse effects associated with administration procedures</td>
<td>✓</td>
</tr>
<tr>
<td>Expected outcomes of tests - effects on behaviour and health caused by the test substances and challenge organisms used</td>
<td>✓</td>
</tr>
<tr>
<td>The nature and level of pain, suffering or distress in the relevant species and methods of assessment</td>
<td>✓</td>
</tr>
<tr>
<td>Use of relevant systems (e.g. score sheets) for recording observations</td>
<td>✓</td>
</tr>
<tr>
<td>Actions to be taken in the event of adverse reactions, suffering or ill-health</td>
<td>✓</td>
</tr>
<tr>
<td>Definition, determination and implementation of humane end-points</td>
<td>✓</td>
</tr>
<tr>
<td>Appropriate methods of euthanasia and of confirmation of death</td>
<td>✓</td>
</tr>
<tr>
<td>Collection and storage of data in accordance with GMP, GLP, QA</td>
<td>✓</td>
</tr>
<tr>
<td>Handling, storing and validating test and challenge materials</td>
<td>✓</td>
</tr>
<tr>
<td>Hazards associated with handling test and challenge materials</td>
<td>✓</td>
</tr>
</tbody>
</table>
observation of animals at these critical points, thereby mini-
mising any suffering. A responsible person such as the study
director, veterinarian or senior animal technician must always be
easily contactable.

4. Summary and recommendations

There is considerable scope for applying the Three Rs to vaccine
batch potency testing. Although the test requirements are driven by
international regulations such as the Ph. Eur., so too is the need to
implement the Three Rs with respect to the tests specified.

Everyone involved with designing and implementing such regula-
tions including regulators within Competent Authorities, manu-
facturers, study directors, test facility managers and animal care
staff, can all contribute in some way using the principles and ideas
within this report. Although this focuses specifically on applying
the Three Rs in vaccine batch potency testing, the principles can be
applied to vaccine studies more widely, including to model devel-
opment, proof-of-concept, challenge validation, challenge passage,
and efficacy studies.

The practical recommendations below are separated into three
categories aimed at companies and individual staff, Competent

4.1. Practical refinement that companies and individual staff can do
now

- All staff involved in the planning, development, conduct and
  review of tests, procedures and systems should be encouraged to
  read this report as a ‘thought starter’ to facilitate imple-
  mentation of reduction, refinement and, where feasible,
  replacement. Input and feedback between all those involved
  will create an environment where opportunities for imple-
  menting the Three Rs are sought throughout the challenge
  assay process and the development of novel vaccines.
- As part of this process, relevant staff should periodically come
together to consider how the Three Rs could be applied in
practice to every aspect of experimental design and test
procedures, and every aspect of the animals’ life-time experi-
ence for the vaccines they work with. Section 3 of this report
provides the background to facilitate a review of:
  - materials and equipment;
  - selection criteria for animals;
  - animal housing, care, handling and transport;
  - acclimatisation of animals;
  - numbers of animals, statistics and experimental design;
  - administration of test materials;
  - reduction of the adverse effects of challenge;
  - development of humane endpoints;
  - monitoring animals, use of anaesthesia and analgesia; and
  - staff training.
- High standards, for example GMP or equivalent, should be
  applied to all aspects of vaccine testing, since production of
good quality data reduces the likelihood of having to repeat
tests.
- Specific examples of reduction and refinement that can
  immediately be applied to two vaccines, Cl. chauvoei and canine
  leptospira, are given in Appendices 1 and 2, respectively. These
examples also provide a template for similar review of test
requirements and methods for other vaccines. The JWGR
recommends that research teams run the vaccines they are
working with through the template to explore the reduction
and refinement opportunities.
- One significant and immediate reduction in the level and
duration of pain and suffering to which a test animal is exposed
can be made by the identification of humane endpoints for test
systems. A valid humane endpoint needs to be defined for
every batch potency vaccine test involving challenge. Any
requirement in the Ph. Eur., and its equivalents to exceed this,
in particular to require death as an endpoint, should then be
challenged by those involved with the tests.
- To facilitate the acceptance of alternative methods, whether
  refinement, reduction or replacement, and avoid conflicts of
  opinion between different Competent Authorities, manufac-
turers should communicate with all relevant parties (see,
Appendix 3, Fig. 1) at an early stage during the development
and validation of alternative methods.
- The EPAA should include more vaccine-related Three Rs
activities in its remit.

4.2. Actions for Competent Authorities

- Competent Authorities should encourage the development of
  alternative methods and data from these alternatives should be
  accepted when suitably validated, and where this leads to a
  reduction, refinement or replacement of animal use.
- Competent Authorities that regulate animal experiments have
  a duty to ensure that animal usage is kept to a minimum and
  animal health and welfare legislation is upheld, so they must
  critically assess whether the tests they are asked to authorise
  are necessary, and whether reduction and refinement options
  could be applied.
- Competent Authorities should waive the fee for licence varia-
tion when satisfactory alternative methods that demonstrably
improve animal welfare are proposed and/or used, and provide
a simple, fast, harmonised system for approval of such
methods.

4.3. Actions for the European Pharmacopoeia, the 9CFR and
equivalent bodies

- Improving awareness of the opportunities for applying the
  Three Rs within the regulatory requirements for testing
vaccines is key to the implementation of each ‘R’ in practice,
but it is currently not easy to interpret test requirements and
determine how flexible these are. Greater clarity of the Ph. Eur.
text is needed, together with provision of additional user-
friendly guidance.
- Wider knowledge of the test requirements in countries outside
  of the EU and USA would also be helpful with greater inter-
action between the regulatory bodies concerned. This would
determine where harmonisation or mutual acceptance is
possible and identify where further reduction, refinement or
replacement opportunities could be developed and
implemented.
- Harmonisation between the Ph. Eur., the 9CFR and other
national equivalents is necessary. In particular, the JWGR
recommends that current test requirements are re-evaluated
to define consistent, statistically justifiable minimum numbers
of animals in test and control groups. Where a need for har-
monisation is identified, a request should be sent to interested
parties for example, Group 15 V of the Ph. Eur.
- Specific recommendations for revision of the Cl. chauvoei and
canine leptospira vaccine monographs are included in
Appendices 1 and 2.

Acknowledgements

The JWGR would like to thank Gavin Thomson for his contribu-
tions and advice with regard to the potency test for canine leptospira.
The authors would like to express their gratitude to the late Secretary to the European Pharmacopoeia Commission, Mr. Peter Castle, who was an inspiration and guide at the start of this work.

Appendix. Supplementary material

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.biologicals.2010.04.004.

References

[34] FELASA. Guidelines for the continuing education for persons involved in animal experiments: recommendations of a Federation of European Laboratory Animal Science Associations (FELASA) working group. Unpublished.