Validation of Alternative Methods for the Potency Testing of Vaccines

The Report and Recommendations of ECVAM Workshop 31[1,2]

Coenraad Hendriksen,[3] Jean-Marc Spieser,[4] Arnoud Akkermans,[3] Michael Balls,[5] Lukas Bruckner,[6] Klaus Cussler,[7] Arnold Daas,[4] Johan Descamps,[8] Roland Dobbelaer,[9] Julia Fentem,[5] Marlies Halder,[5] Margot van der Kamp,[10] Roger Lucken,[11] Julie Milstien,[12] Dorothea Sesardic,[13] Donald Straughan[14] and Antonio Valadares[15]

[3]*National Institute of Public Health and the Environment (RIVM), 3720 BA Bilthoven, The Netherlands;* [4]*European Department for the Quality of Medicines, Council of Europe, 67029 Strasbourg Cedex 1, France;* [5]*ECVAM, JRC Institute for Health & Consumer Protection, 21020 Ispra (VA), Italy;* [6]*Institute of Virology and Immunoprophylaxis, 3147 Mittelhäusern, Switzerland;* [7]*Paul-Ehrlich-Institute, Paul-Ehrlich-Strasse 51–59, 63207 Langen, Germany;* [8]*SmithKline Beecham Biologicals SA, 89 Rue de l'Institut, 1330-Rixensart, Belgium;* [9]*Department of Biological Standardisation, Scientific Institute of Public Health, Ministry of Social Affairs, Health and Environment, Juliëtte Wydsmanstraat 14, 1050 Brussels, Belgium;* [10]*Netherlands Centre Alternatives to Animal Use, Yalelaan 17, 3584 CL Utrecht, The Netherlands;* [11]*Trawsnant, Cross Inn, Llandysul, Dyfed SA44 6LX, UK;* [12]*World Health Organization — Vaccine Supply and Quality Unit, 20 Avenue Appia, 1211 Geneva 27, Switzerland;* [13]*National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3QG, UK;* [14]*9 Beaufort West, Bath, Avon BA1 6QB, UK;* [15]*Laborotório National de Investigação Veterinaria, Estrada de Benfica 701, 1500 Lisbon, Portugal*

Preface

This is the report of the thirty-first of a series of workshops organised by the European Centre for the Validation of Alternative Methods (ECVAM). ECVAM's main goal, as defined in 1993 by its Scientific Advisory Committee, is to promote the scientific and regulatory acceptance of alternative methods which are of importance to the biosciences and which reduce, refine or replace the use of laboratory animals. One of the first priorities set by ECVAM was the implementation of procedures which would enable it to become well-informed about the state-of-the-art of non-animal test development and validation, and the potential for the possible incorporation of alternative tests into regulatory procedures. It was decided that this would be best achieved by the organisation of ECVAM workshops on specific topics, at which small groups of invited experts would review the current status of various types of *in vitro* tests and their potential uses, and make recommendations about the best ways forward (1).

The joint ECVAM/AGAATI (Advisory Group on Alternatives to Animal Testing in

Immunobiologicals) workshop on Validation of Alternative Methods for the Potency Testing of Vaccines was held in Angera, Italy, on 14–16 November 1997, under the co-chairmanship of Coenraad Hendriksen (RIVM, The Netherlands) and Jean-Marc Spieser (European Department for the Quality of Medicines [EDQM], France). The participants, all experts in vaccine quality control and/or validation procedures, came from international regulatory or government organisations, national control laboratories and vaccine manufacturers. The principle objective of the workshop was to discuss, in an informal atmosphere, the complex issue of guidelines for the validation of alternative methods for the potency testing of vaccines. The specific aims of the workshop were: a) to discuss existing guidelines for validation; b) to explore specific guidelines for the validation of alternative methods for the potency testing of vaccines, and to agree on strategies for the preparation of such guidelines; and c) to discuss the implementation of validated alternative methods in a regulatory context. The outcome of the discussions and the recommendations agreed by the workshop participants are summarised in this report.

Introduction

Vaccines are considered to be the most cost-effective tools in the prevention of infectious diseases. In the coming years, their importance will continue to increase because of the priority given by the World Health Organization (WHO) and the European Union (EU) to the eradication of a number of diseases, the emergence of antibiotic-resistant strains of bacteria, the impact of viral infections, the high incidence of infections in large livestock industries, and various other factors. Traditionally, laboratory animals have played a major role in the quality control of vaccine batches, and, consequently, the use of animals for this purpose is extensive. It is estimated that about 10% of the total number of laboratory animals used in Europe are required for the safety and potency testing of vaccines for veterinary and human application.

However, although vaccines are becoming more important in human and veterinary health care, the total number of animals used for quality control is likely to decrease in the near future, for a number of reasons, namely: a) the evolution of the concept of vaccine quality control, as there is a growing feeling that emphasis should be put on ensuring the consistency of a product which has been shown to be safe and efficacious; b) the priority which is given to the development and implementation of alternative methods to the classical animal tests; and c) the development of a new generation of vaccines which are better defined, thereby reducing the need for extensive quality control and thus changing the role of final product testing in the control of product quality.

Although the workshop focused in particular on the development and validation of alternative methods for testing the potency of vaccines, it was recognised that the relevance of these methods in routine testing can only be seen in relation to the evolution of the concept of quality control. Therefore, this issue was discussed simultaneously. Furthermore, recommendations were made in relation to the further implementation of these methods in test guidelines.

In this report, the term "alternatives" is used in the context of the Three Rs, referring to methods which can replace, reduce and/or refine the use of laboratory animals (2). The term *"in vitro"* is sometimes used when referring specifically to replacement alternative methods.

Concepts of the Quality Control of Vaccines

Vaccines are derived from living organisms in a batch-wise procedure. This implies that their characteristics can vary from batch to batch. Each batch produced in one production run is considered unique. Strict controls must be in place to ensure production consistency, so that the safety and efficacy of each batch can be assured.

While the quality of vaccines is the prime responsibility of the manufacturer, the Competent Authority or National Control Authority approves procedures to ensure that biological products intended for use in humans or animals are of adequate safety and efficacy. Quality is a goal which the manufacturer approaches throughout the steps of product development and by defining in detail the processes by which the vaccine is

produced and tested, including the in-process and final product testing to be carried out. Attainment of quality depends largely on the quality control tests performed at various critical steps during the production process, and on the application of Good Manufacturing Practice. Quality assurance systems control the reliability and reproducibility of each of the production and testing steps (3, 4). Quality control should be seen as an instrument for monitoring consistency in the production and testing of vaccines (5).

Consistency of production means that each batch of a product is of the same quality and is within the same specifications as a batch which has been shown to be safe and efficacious in human trials or in the target animal species. The formerly established concepts of quality control were based on the uniqueness of each individual batch. A shift in emphasis away from reliance on final product testing will require development of a control scheme for each product or product class. It is within this context that the development and validation of alternative methods for potency testing were considered by the workshop participants. The need to establish systems for monitoring consistency is crucial in this respect.

## Monitoring Systems to Demonstrate Consistency in Vaccine Production and Testing

The principle of monitoring, which is a very important issue in quality assurance programmes, is to test whether data can be considered as coming from one and the same population. If so, this would indicate that the process of production and testing is consistent. Monitoring can be performed by using control charts. A minimum of 10–20 assay results have to be available before a reliable control chart can be established. When a new method is being introduced, these results can be taken from the validation study and used to calculate fixed limits (for example, two standard deviations). Another approach can be a continuous adaptation of the limits, based on all valid results or on a set of the most recent assay results (rolling average of, for example, 30 consecutive assays). Two types of control charts can be used — graphical or numeri-

cal. The merits of a graphical control chart are that: a) trends can be seen easily; and b) by using pre-printed forms, which can be filled in by hand, graphical charts are user-friendly. The advantages of a numerical control chart are that: a) computer analysis of the data is possible; b) it enables the use of cumulative sum and recalculation of overall variance and mean; and c) graphical representations of the numerical control charts can easily be produced. Examples of control charts are the Shewhart Control Chart (6) and the CUSUM (Cumulative Sum) Control Chart (7, 8), which is more sensitive to shifts (9).

A variety of parameters can be used for monitoring the performance of potency assays, including: a) ED50 (dose of vaccine needed to have an effect in 50% of the animals) and PD50 (dose of vaccine protecting 50% of the animals against the effects of a challenge); b) in the case of serological tests, the average antibody response (in multi-dilution assays of one or all groups of animals) on the condition that test doses in the different tests are the same; c) variance in response in *in vitro* or *in vivo* tests; d) slopes of dose–response curves in multi-dilution assays; and e) the response of the positive control and/or internal standard.

Before being used in a control chart, data shoud be checked for their normal distribution. If this is not the case, a transformation (for example, logarithmic) can be applied. Figure 1 shows the effect of log transformation on the distribution of the data (in this case ED50 values of the reference preparation were used). After log transformation, outliers are on both sides of the distribution curve, whereas before log transformation, outliers are only seen at the higher end of the distribution curve.

Generally, alternative tests such as serological assays are better able to monitor consistency than classical (challenge) tests. This is because of the nature of the parameters measured (for example, quantitative antibody responses versus lethality) and the additional inherent variability of classical challenge models. The lack of inherent variability can be considered an additional advantage of alternative methods, although it should be stressed that the acceptance of alternative methods should not only be based on this advantage, but also on the relevance or validity of the model.

Figure 1: Distribution of ED50 values
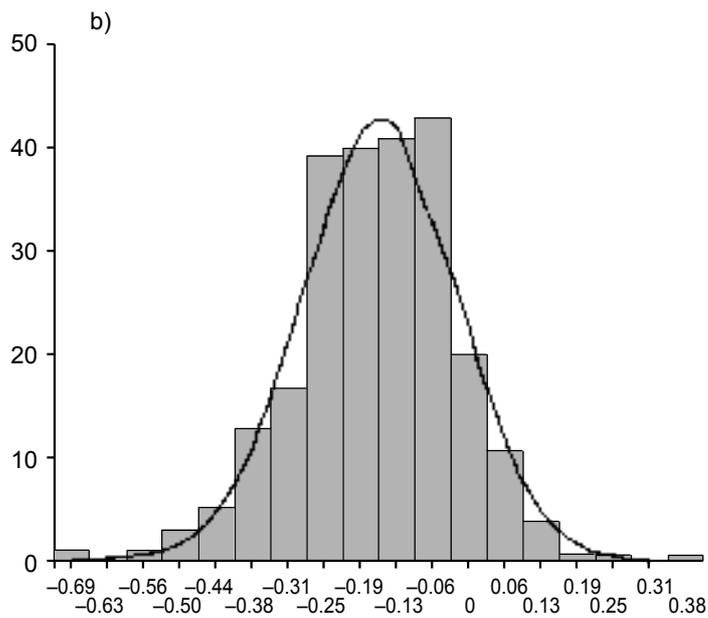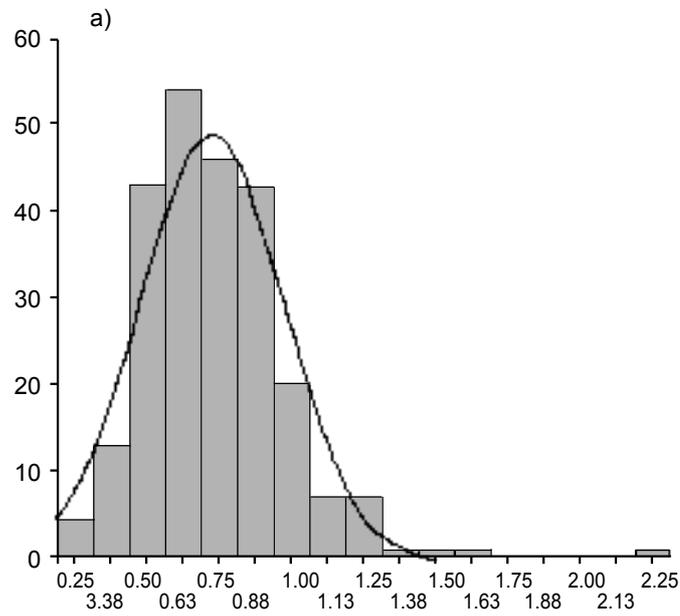
a) *Before log transformation SD = 0.25, mean = 0.73, n = 241.*

b) *After log transformation SD = 0.14, mean = –0.16, n = 241.*

Scientific Basis for the Introduction of Alternative Methods

As a general principle, potency tests have been designed to measure the ability of the vaccine to protect against subsequent challenge with the active component responsible for pathogenicity. Retrospectively, it can be concluded that the routine use of these methods has not resulted in the release of ineffective vaccines. A number of alternative methods (both serological assays and *in vitro* tests) have now been developed and proposed as replacements to current challenge-based potency tests (10).

It is impossible to discuss the development of an alternative method without a clear understanding of the purpose for which the assay is to be used. It is essential to discriminate between testing procedures (vaccine research and development, evaluation of novel vaccines, or existing vaccines in novel presentations and delivery systems during the developmental, pre-clinical or pre-licensing stages), and the batch quality control of licensed vaccines and fully established products. It is generally accepted that, for batch release of established products, it should be possible to use tests which are different from those used to determine the efficacy of novel formulations. However, these batch release tests are a subset of those used to characterise the product going into clinical trials. Even so, it is important to remember that, for a vaccine, the overriding objective of any testing procedure for whatever purpose is the evaluation of the ability of the vaccine to protect against the pathogenic entity against which it is targeted. Hence, in the ideal situation, to develop effective and credible alternative methods, it is essential to understand the mechanism of protection (11). For vaccines, this generally means both the mechanism of induction of a protective immune response and the mode of action of the pathogenic entity in causing disease. Also, development of mechanistically based assays would require an understanding of how virulence factors exert their pathogenic effects. Such assays will eventually lead to the complete replacement of animal models which test for protection. This must be the ultimate goal in the development of such alternatives.

However, the scientific knowledge presently available limits the development of such mechanistically based assays. The use of targeted or monoclonal antibodies could improve the specificity and reliability of immunoassays, but these antibodies are generally not able to discriminate between all the domains of an antigen required for its biological function. Similarly, immunoassays cannot always differentiate between protective and non-protective antibodies and, as such, are limited in their prediction of protection. An additional difficulty, especially for the human vaccines, is our limited understanding of the relevance to humans of the *in vivo* models currently used to test for protection. Due to the lack of information that could lead to mechanistic approaches being undertaken, we are forced to consider possible "empirical" approaches that might provide a pragmatic solution to the problem, at least in the short term. The use of quality control tests to monitor consistency, rather than to establish the "true" effectiveness of a vaccine, might be such a solution.

For some vaccine products, *in vitro* assays will not be able to completely replace animal models for the demonstration of efficacy for the foreseeable future. Therefore, it will be difficult to adopt *in vitro* methods or serological-based approaches where protection is not dependent on antibody response. Nevertheless, it is likely that *in vitro* assays can be validated for certain specific purposes, for instance to monitor certain aspects of consistency in production.

Finally, it remains to be mentioned that the need for potency testing in animals is likely to disappear in the future when current vaccines are replaced by highly purified and well-characterised products from genetic engineering technologies.

Alternatives to the Potency Testing of Human and Veterinary Vaccines: Current Status

Potency tests in animals are currently required for each single vaccine batch. Ideally, a potency test (*in vivo* or *in vitro*) should be able to ensure that vaccine lots meet at least the potency of the lots for which efficacy in the target species was demonstrated and accepted during the licensing procedure. In the ideal situation, the experimental model (*in vivo* or *in vitro*) should reflect the target species, in all relevant aspects, or should be performed for veterinary vaccines

in the target species, as is the case for many poultry vaccines. In the classical concept of quality control (the uniqueness of the vaccine batch produced), *in vitro* antigenicity models are generally considered not to be valid replacements for *in vivo* assays because they quantify the amount of antigen and do not reflect the immunogenicity of an antigen, and they do not reflect the rather complex cascade of responses after *in vivo* immunisation.

However, with the experimental animal models, the ideal situation is seldom met and, in most cases, potency tests are performed by using experimental *in vivo* models which differ substantially from the target species (Table I). In addition, the primary goal of potency testing should be monitoring of consistency after demonstration of efficacy in the target species.

With regard to the experimental *in vivo* models, the following comments can be made.

1. In general, there are differences between the immune responses of laboratory animals and the target species, including humans.

2. In challenge tests, or other "functional" potency tests, the challenge organism is a defined laboratory strain. Its characteristics might differ from the strains in natural infection and the route of infection might be different. For example, with whole cell pertussis vaccine potency tests, the challenge strain (strain 18323) used differs antigenically from the wild strains of micro-organism (12). Also, the chal-

lenge route (intracerebral) differs substantially from the route of natural infection (respiratory tract).

3. Serological tests might measure non-protective antibodies as well as protective antibodies.

This clearly shows that concessions to the ideal situation have already been made, and that judging replacement of an *in vivo* model by an *in vitro* method on a case-by-case basis might not be such a big step as it first seems. In fact, *in vitro* methods are already being used routinely for a number of vaccines (Table II).

If, for a given vaccine, replacement is considered inappropriate and *in vivo* testing appears to be indispensable, several rational approaches to reducing the use of laboratory animals should be considered.

1. Potency tests in animals could be conducted at an earlier stage of production (up-stream), and final lots produced from the bulk could be released on the basis of *in vitro* data, thereby reducing the number of tests.

2. Potency tests in animals should be required only if the composition of the final product is changed (for example, new components, new seed lot, etc.).

3. For a strictly limited period of time, or for a strictly limited number of batches, potency could be estimated *in vitro* and *in vivo* in order to gain additional data. For instance, potency could be determined *in vivo* and *in vitro* during the development

Table I: Human and veterinary vaccines for which animal testing is required on every final batch

Vaccine type	Human vaccines	Veterinary vaccines
Toxoid vaccines	Diphtheria, tetanus	Tetanus, *Clostridium* spp.
Inactivated bacterial vaccines	Haemophilus B, pertussis	Leptospira, swine erysipelas, *Escherichia coli*, etc.
Inactivated viral vaccines	Polio, rabies	Rabies, parvoviruses, Newcastle disease, etc.

Table II: Vaccines for which no routine animal potency test is required

Vaccine type	Human vaccines	Veterinary vaccines
Live bacterial vaccines	Oral typhoid, BCG	Brucellosis
Polysaccharide vaccines	Meningococcal and pneumococcal	—
Live viral vaccines	Measles, mumps	Distemper, parvoviruses, Marek's disease, etc.
Inactivated viral vaccines	Influenza, hepatitis A, hepatitis B	Feline leucosis

phase of a vaccine. Once a product is licensed, *in vivo* testing should be discontinued.

4. Validated single-dilution assays and other simplified animal models could be used.

Of course, all these approaches imply that consistency of production has been adequately demonstrated and is maximally maintained by a programme of quality assurance. Furthermore, the validity of the alternative method (*in vitro* or *in vivo*) used should be reassessed when significant changes in production occur, or when the component is included in a new combination vaccine. Significant reduction, or even complete replacement, of routine potency testing in animals should be possible without jeopardising the quality of information on the product.

With respect to inactivated veterinary vaccines, a completely different situation is evident, since the vaccines can be tested in the target species and, in fact, batch potency testing of poultry vaccines is regularly performed in the target species. Vaccines for mammals are usually tested in laboratory animal models which are based on vaccination-challenge tests (for example, swine erysipelas, leptospirosis, blackleg) or serological procedures (for example, *Escherichia coli*, parvovirus vaccines). Performance of the potency test completely *in vitro* is not yet common; however, the monograph on feline leucosis vaccine is the first monograph which does not stipulate an animal test for batch potency testing (13).

Another promising approach has been undertaken with the monograph on inactivated rabies vaccine for veterinary use which, since its last revision (14), includes an *in vitro* immunochemical method for the determination of the glycoprotein content of the test vaccine as an alternative to the classical mouse challenge test for potency testing. Currently, both tests are performed in parallel, and it is hoped that the data collected will show a good correlation between the two test systems and result in the replacement of the mouse potency test in the near future.

With the veterinary vaccines, challenge tests in laboratory animals should not be used as gold standards in validation studies of (future) alternative tests; the necessity of efficacy testing of vaccines in the target species during licensing offers the unique possibility of comparing the results obtained in the target species with the results obtained by using alternative methods. So far, there has been limited progress in implementing the Three Rs concept in the regulatory requirements for potency testing of veterinary vaccines; however, alternative methods for potency testing of various clostridial toxoids (15, 16) and swine erysipelas vaccines (17) are available and have been incorporated into the drafts for revision of the monographs concerned. These models have already been published for consultation in *Pharmeuropa*, the official journal of the European Pharmacopoeia Commission.

Some of the alternative methods developed to replace challenge tests in the quality control of vaccines reached the validation stage, as outlined below.

The Validation of Alternative Test Systems for the Potency Testing of Vaccines: Case Studies

Recently, the use of an *in vitro* method (an ELISA procedure based on the quantification of genetically engineered hepatitis B surface antigen) was validated in an in-house study. This method was intended to replace the *in vivo* ED50 mouse potency test. The in-house validation study was focused on aspects such as the accuracy of the model, intra-assay and inter-assay precision, specificity, and quantification limits. Based on the results of this study, it was concluded that the ELISA could serve as a means of routinely monitoring the consistency of production of hepatitis B vaccine. Nevertheless, it was still felt to be necessary to perform the *in vivo* assay for stability studies and for the validation of major process changes, such as scaling-up of production.

A second example is the validation of serological methods as alternatives to the potency test for tetanus vaccine. For the veterinary tetanus vaccine, where the standard method assesses potency by toxin neutralisation in mice (18), the conclusion of an inter-laboratory study, performed in seven laboratories in six European countries, was that both ELISA and the toxin binding inhibition (ToBI) test were suitable alternative test systems for assessing potency (15). Currently, another interlaboratory study is in progress to evaluate the use of ELISA and the ToBI test for the challenge-based potency test for human tetanus vaccines. This study, commissioned by the EDQM and ECVAM, is divided into several phases, giving the opportunity to allow appropriate monitoring and any changes to the study design which are deemed necessary. A report of this study will be available in 1999.

One of the problems identified in these validation studies was that only a few specific guidelines on the validation of alternative methods for vaccine potency testing were available.

Overview of Guidelines on the Validation of Alternative Methods

In the last decade, efforts have been made to establish guidelines on the validation of alternative methods for the toxicity testing of chemicals. In 1995, based on experience gained during several large-scale validation studies on toxicological tests, ECVAM published recommendations concerning the practical and logistical aspects of validating alternative methods for the toxicity testing of chemicals (19, 20). Five main stages in the evaluation of new test methods were identified: test development; prevalidation; validation (involving a formal interlaboratory study); independent assessment; and progression toward regulatory acceptance.

Before a test can be considered for validation or regulatory use, it must have been properly developed, in terms of defining its purpose, the need for the test, the derivation of the method, its applicability for the intended purpose, the case for its inclusion in a prevalidation/validation study, the provision of a protocol/Standard Operating Procedures, and the definition of a prediction model for the interpretation and application of the results (19–22). Prevalidation includes three main phases: protocol refinement; protocol transfer; and protocol performance (22). The informal interlaboratory study carried out at the prevalidation stage involves assessing the interlaboratory transferability of the method, undertaking any optimisation and standardisation of the protocol which may be considered necessary, and identifying any unexpected problems with the test procedures. The objective of the prevalidation stage is to ensure that any method included in a formal validation study adequately fulfils the criteria defined for inclusion in such a study. In the validation stage, the main objective is to conduct an interlaboratory blind trial as a basis for evaluating whether a test can be shown to be relevant and reliable for its specific purpose, according to defined performance criteria (19, 23). The design of the study should reflect its objectives, which need to be clearly stated and realistic. The following main stages are involved: selection of tests; selection of laboratories; selection and distribution of test reagents; data collection and analysis; assessment of the outcome according to the prediction model; and proposal of the next steps to be taken (19, 23, 24). The published results of a validation should be considered in detail by one or more independent assessment panels, under the auspices of appropriate national or international organisations, before any regulatory authorities, or their

equivalent, are asked to consider the formal acceptance of any satisfactorily validated method for incorporation into the regulatory framework (21).

Apart from general guidelines, little has been published on the validation of alternative methods in vaccine quality control. Practical information on the validation of *in vitro* methods for toxoid vaccine potency testing was provided by the WHO (25). These guidelines included specifications for the number of assays (for each formulation of vaccine, not less than ten valid parallel line assays using both the classical test and the alternative model), the criteria for acceptance and the cases for reassessment of the validity of the assay. General guidelines on validation procedures have been given by the WHO (4), the European Pharmacopoeia (26), and the International Conference on Harmonisation (ICH; 27).

Prevalidation and Validation of Alternative Methods for Vaccine Potency Testing: Background Information

Those tests which involve the most animals and the largest amount of suffering should be given the highest priority for the validation of alternative methods. However, this is moderated by whether the state-of-the-art seems likely to make replacement a reasonable possibility for any particular product test. In this context, for veterinary vaccines, the validation of alternative methods for potency testing of clostridial, erysipelas and rabies veterinary vaccines has generated considerable interest. For human vaccines, much activity has occurred in the testing of tetanus and diphtheria toxoid and hepatitis B vaccines. There are still a very substantial number of products for which there are presently no useful alternative tests.

There are very many products, the tests for which are very product-specific; however, for each product there are usually very few manufacturers. A consequence of this is that each validation study is of concern to only a small group of individuals. Although the manufacturers should have the prime responsibility for initiating validation studies, they may have neither the human nor the financial resources to do this. Hence, the responsibility for initiating these studies has

become an uneasy compromise among all the parties concerned (for example, manufacturers, national and international control bodies, pharmacopoeial authorities). Significant progress has been made with a small group of widely used human and veterinary vaccines. The use of alternative methods for tetanus potency testing has been allowed by the WHO since 1990 (28), but no international collaborative study was performed until recently. This study, based on a collaboration between the European Pharmacopoeia, the RIVM, ECVAM and several national control laboratories, is a good example of the cooperation between various parties which is essential to make validation studies successful.

Guidelines on Various Aspects of Validation of Potency Tests for Vaccines

There have been a number of attempts to produce general guidelines for the validation of potency tests for veterinary and human vaccines. General acceptance of these guidelines has been very difficult to achieve, largely because the views of regulators in different countries and those of industry vary greatly. These differing views essentially stem from entrenched ideas about what constitutes validation, so that methods may be considered valid by one authority but invalid by another. This in turn makes manufacturers wary of introducing new techniques, since the risk of rejection by one authority is far greater than if established methods are employed, even if those methods are themselves not fully validated. Without doubt, this situation has limited, and continues to limit, progress, particularly in the development and introduction of alternative tests.

It is evident that clear, practical, understandable and widely accepted guidelines are needed. However, decisions must be taken about the purpose of these guidelines. They need to provide guidance not only to those intending to develop and validate methods, but also to assessors and experts who must evaluate the results of the studies as part of the registration process. Secondly, consensus is needed to decide the broad contents of the guidelines. It was agreed at this workshop that guidelines need to include definitions of terminology, particularly a

clear definition of what constitutes a potency test, and the terms applicable to validation studies themselves. As far as possible, these definitions should be consistent with those already promulgated in existing documents. The definitions given in Appendix 1 are derived from documents of the WHO (3, 4), the European Pharmacopoeia (26), the ICH (27) and ECVAM (19, 23). In addition, it was considered that guidelines should include advice on required levels of laboratory competence (accreditation), experimental design, statistical tools, establishment of criteria, use of reference preparations, and techniques for the maintenance and continuity of validity.

The ultimate objective is to provide guidelines which will facilitate the development, validation, implementation and regulatory acceptance of alternative methods, but it was agreed that this cannot be done in isolation. To be acceptable, alternatives have to be at least as good as existing *in vivo* methods for the purpose for which they are designed.

The first requirement is to define what a potency test is supposed to do. Potency must be correlated with efficacy but a potency test does not necessarily need to measure efficacy directly. The test must, however, be capable of detecting batches which possess an activity different to that of a batch, or batches, for which efficacy has been demonstrated. The test should be capable of detecting batches which have either a lower or higher activity, and should normally be capable of quantifying that difference, although for routine batch release purposes it might be sufficient that the test indicates only that the test batch has a potency at least equal to the reference preparation or to an efficacious batch produced previously.

There are therefore two basic aspects of validation which need to be considered: validation of the correlation with efficacy; and validation of the method itself.

Since the methods used vary greatly, it follows that the approaches for validation are likely to vary. Consequently, it would seem reasonable to structure a guideline so that a general section on principles is followed by more-specific guidance for different types of tests applicable to different types of product.

As a starting point, the following groups of tests might be considered.

1. Test for live vaccines — viral titrations, bacterial counts, viability assays.

2. Challenge models — laboratory animals/target species.

3. Serological methods — mainly for inactivated products.

4. Cell-mediated immunity assays.

5. Totally *in vitro* approaches — antigen quantitation.

In addition, the relationship between validation of potency methods and process validation (maximising quality assurance) should be considered.

Based on this general outline, it was agreed that a first draft of a guideline would be prepared. ECVAM will be asked to establish a task force to work on this. It was agreed that the guideline would be discussed in relation to its application for both human and veterinary products; a decision will need to be made as to whether separate but parallel guidelines should be developed for the two areas.

Apart from the lack of clear guidelines, some general difficulties in developing, validating and implementing alternatives were identified. These difficulties stem in part from: a) the "good experience" and "reassurance" provided to both manufacturers and regulators by the continued use of familiar, but sometimes imperfect, animal tests; b) cost and resource considerations of developing and validating alternative methods; and c) the inevitably slow process of multinational agreement to revise pharmacopoeial monographs. An indication of the timetable usually required for the whole process of test development to regulatory acceptance is shown in Table III.

Biostatistical Methods in the Validation of Alternative Methods in Vaccine Potency Testing

Biostatistical methods are an essential tool in the analysis of data generated during a validation study. Providing recommendations on the application of these methods is a difficult task. However, there was general agreement that the appropriate statistical tools should be selected and applied by a biostatistician familiar with the conduct and requirements of validation studies. The biostatistician should endeavour to formulate the choice of methods during the test devel-

Table III: Approximate time-scale from test development and implementation in test guidelines

Stage	Required time (years)	Total time (years)
Development of test	2–3	2–3
In-house validation	1	3–4
Prevalidation	1	4–5
Validation study	2–3	6–8
Acceptance of test	1.5	7.5–9.5
Preparation of reference preparations	1.5	9–11
Implementation of test in regulations	0.5	9.5–11.5

opment phase, but the results obtained during prevalidation could feasibly give rise to the adoption of another, more appropriate, statistical approach than was intended initially. To prevent, as far as possible, the selection of an inappropriate design and/or statistical method, it is strongly recommended that the biostatistician is involved at all stages of the development and validation of the test, and also that the analyses are carried out by two independent biostatisticians. This could prevent misinterpretation of the data, or biased conclusions due to the choice of methods.

It should be borne in mind that the answer to whether the response measured in an assay is directly related to the efficacy of the product cannot be provided by statistical methods, since these can only support or reject the validity. In order to select an optimal design, it is necessary to formulate the expected ultimate goal of the alternative method at an early stage, and statistical analyses should be performed. The following factors play an important role:

— whether the alternative is a single-dilution or a multiple-dilution assay;

— whether the method has to be able to assay products from different manufacturers against the same reference;

— whether the existing method is known to reflect reliably the efficacy of the product;

— whether this alternative method has to give quantitative results (i.e. a potency estimate) or only has to answer the question of whether or not the product meets a given minimum (and perhaps a maximum) requirement; and

— what the key variables are that could possibly jeopardise the study (if they were not prescribed in the Standard Operating Procedure), and which variables can be ignored.

Selection, Preparation, Validation and Maintenance of Reference Materials

In many cases, the application of alternative methods in vaccine potency testing will require the selection of reference preparations, both for international standardisation purposes and quality assurance. An example is the introduction of reference sera when challenge potency tests are being replaced by a serological-based method. These preparations have to be calibrated before introduction into routine testing.

Many current principles used in the standardisation of reference materials will apply equally to human and veterinary vaccines. Traditionally, International Standards and reference reagents have been established and calibrated in interlaboratory validation studies so that national standards and in-house reference preparations can be calibrated and used for potency and other testing of biological products. International Standards, established by the WHO, are known as primary standards. Secondary standards, calibrated where appropriate against the WHO standard, can be regional

(for example, European Pharmacopoeia working standard), national, or specific to a particular company (manufacturer's working standard).

As the alternative assays are being developed for testing vaccines, the need for standardisation will continue. However, while it will be desirable to use the same reference standard for both the classical method and any alternative test, it may be necessary to develop method-specific reference materials. For example, it may sometimes be appropriate for a reference material to be a vaccine, but in other cases a (homologous) serum derived from an animal or group of animals vaccinated with the product may be more appropriate. In other cases (for example, rabies vaccine potency test), specific monoclonal antibodies will be of use. For some tests, highly purified antigens might be the best material. Generally, there needs to be some clearly defined relationship (for example, homologous reference serum samples) between the reference material and the product of demonstrated efficacy and potency. Furthermore, it is important to demonstrate in the test systems involved that the test samples behave in the same way as the reference samples, though they need not necessarily be identical.

One point which is particularly worth noting is that establishing reference and standard materials requires a very substantial resource input. Consequently, it is always advisable to consider whether the material can be made to serve multiple purposes. The preparation of reference materials is the ultimate responsibility of the regulatory bodies or, for international standards, the WHO.

Reference preparations must be stable. Therefore, particular attention needs to be given to the method of preparation and the formulation of the reference preparation. This may mean that the material will be stored under different conditions than the commercial product itself, even when the reference is the actual vaccine. Documentation should be provided of the preparation, including the characterisation and storage conditions. These should relate to the stability of the reference preparation and not interfere with intended use. Presentation of appropriate filling volume will ensure optimisation of its use in the most economical way. The validity of the reference preparation depends on it being maintained prop-

erly. In order to do so, the stability of the material must be understood and monitored. In principle, trend analysis of data generated during routine testing might provide useful information on the reference preparation.

The key piece of information that is required in respect of any standard or reference preparation is its activity, which will be defined in terms of units used in the test method(s). Where an International Standard exists, calibration is in International Units (IU). However, we need to be very sure of this assigned value. For calibration, an officially recognised method should be used. The method used must therefore be fully validated with respect to its characteristics, its variability, its robustness and its accuracy. This will allow definition of the activity of the reference material in those terms. In the case of a standard, definition of the assigned value should be the result of an interlaboratory study, preferably involving both manufacturers and regulatory authorities.

Long-term maintenance of standard and reference preparations also requires that defined procedures are in place for replacement of the reference when this is necessary. This can usually be achieved with multiple assays, comparing the new and old materials, but it is obviously essential that both preparations are produced by identical means and that results obtained are comparable in terms of repeatability, robustness and specificity, and expressed with reference to the International Standard wherever appropriate.

Harmonisation

Positive steps are being taken with respect to the adoption of alternative potency test methods, which can promote harmonisation of requirements for biological products. It is to be expected that the shift to reliance on production consistency rather than solely on final product testing presents an opportunity to promote harmonisation in this area.

Guidelines for the validation of alternative tests can contribute to this effort. There will be a need to ensure that there is a strong scientific basis in assessing the credibility of a validation study. Involving control authorities (working according to guidelines from different regional and/or international bodies [WHO, European Pharmacopoeia, etc.])

in collaborative studies for test validation should be strongly considered as a mechanism to promote harmonisation. The existing channels through the WHO, the European Pharmacopoeia, ECVAM, and other multilateral groupings, ICH and national authorities should be used to their full extent in this effort.

Conclusions and Recommendations

1. Information on the potency of vaccines should ideally be obtained in the target species. However, in most cases, potency tests are performed with experimental *in vivo* models which differ from the natural infection in the target species. Thus, concessions to the ideal situation have already been made and replacement of these *in vivo* models by *in vitro* methods may not necessarily imply loss of valid information.

2. Vaccine quality control should be seen as providing confirmatory evidence of the consistency in production. This places emphasis on *in vitro* methods which may usefully replace or supplement the *in vivo* tests currently used in batch testing.

3. The change in emphasis to reliance on production consistency in addition to final product testing presents an opportunity to reduce the numbers of animals being used and to promote harmonisation in this area.

4. When it is considered impossible to introduce *in vitro* potency tests, and *in vivo* testing is still considered indispensable for a given vaccine, rational approaches, such as shifting *in vivo* potency testing to an earlier stage in the production, or the use of single dilution assays, should be considered, in order to minimise the use of laboratory animals. The use of endpoints other than lethality or severe clinical signs should be evaluated.

5. During test development and validation, attention should be given to the identification of parameters which can be used to monitor consistency in production and testing.

6. Challenge tests in laboratory animals should not be considered as the gold standard for the validation of (future) alternative tests, although they may be. For veterinary vaccines, the necessity of efficacy testing of vaccines in the target species during licensing offers the unique possibility to compare the results obtained in the target species with the results obtained by using alternative methods.

7. In order to gain experience with an *in vitro* assay in routine testing, such an assay could be conducted in parallel with *in vivo* potency tests in a small number of animals for a strictly limited period of time (for example, 2 years) or with a strictly limited number of batches (for example, the first 5–10 batches).

8. For human rabies vaccines, data on parallel studies conducted with both the *in vivo* and the *in vitro* models to estimate the potency of batches should be made available to the regulatory bodies. On the basis of valid data, deletion of the *in vivo* test from the European Pharmacopoeia monograph for human rabies vaccines should be considered.

9. Those tests which involve the most animals and the most suffering should be afforded the highest priority for the development and validation of alternative methods. Obviously, consideration must also be given to whether the state-of-the-art is such that replacement is a reasonable possibility.

10. In the validation of alternative methods in vaccine potency testing, appropriate statistical methods should be selected and applied by a biostatistician.

11. To prevent the selection of an inappropriate design, and/or statistical method, it is strongly recommended that a biostatistician, familiar with validation studies, is involved at all stages of the test development, validation and data analysis.

12. Involving control authorities (working according to guidelines from different regional and/or international bodies, such as the WHO and the European Pharmacopoeia, in collaborative validation studies should be strongly considered as a mechanism to promote harmonisation.

13. A task force should be established to draft guidelines for the validation of alternative methods for the potency testing of vaccines.

References

1. ECVAM (1994). ECVAM News & Views. *ATLA* **22**, 7–11.
2. Russell, W.M.S. & Burch R.L. (1959). *The Principles of Humane Experimental Technique*. 238pp. London, UK: Methuen.
3. WHO (1992). *WHO Technical Report Series No. 882: Guidelines for national authorities on quality assurance for biological products,* pp. 31–44. Geneva, Switzerland: WHO.
4. WHO (1997). General introduction. In *Manual of Laboratory Methods for Potency Testing of Vaccines Used in the WHO Expanded Programme on Immunisation*. *WHO/VSQ/97.04, pp. 3–12.* Geneva, Switzerland: WHO.
5. Griffiths, E. (1996). Assuring the quality of vaccines: regulatory requirements for licensing and batch release. In *Methods in Molecular Medicine: Vaccine Protocols* (ed. A. Robinson, G. Farrar & C. Wiblin) pp. 269–288. Totowa, NJ, USA: Humana Press.
6. Shewhart, W.A. (1931). *Economic Control of Quality of Manufactured Products, 495pp.* New York, USA: Nostrand.
7. Anon (1991). Minitab Statistical Software Release 8, *QC Supplement: Quality Control and Improvement*, pp. 21.15–21.20. Rosemont, PA, USA: Quickset.
8. Rowlands, R.J., Wilson, D.W., Nix, A.B., Kemp, K.W. & Griffiths, K. (1980). Advantages of CUSUM techniques for quality control in clinical chemistry. *Clinica Chimica Acta* **108**, 393–397.
9. Finney, D.J. (1964). Control charts. In *Statistical Methods in Biological Assay*, pp. 161–162. London, UK: Charles Griffin & Company Limited.
10. Hendriksen, C.F.M., Garthoff, B., Aggerbeck, H., Bruckner, L., Castle, P., Cussler, K., Dobbelaer, R., van de Donk, H., van der Gun, J.W., Lefrancois, S., Milstien, J., Minor, P.D., Mougeot, H., Rombaut, B., Ronneberger, H.D., Spieser, J-M., Stolp, R., Straughan, D.W., Tollis, M. & Zigtermans, G. (1994). Alternatives to animal testing in the quality control of immunobiologicals: current status and future prospects. The report and recommendations of ECVAM workshop 4. *ATLA* **22**, 420–434.
11. Sesardic, D. (1996). Requirements for valid alternative assays for testing of biological therapeutic agents. *Developments in Biological Standardisation* **86**, 311–318.
12. Arico, A., Gross, R., Smida, J. & Rappuoli, R. (1987). Evolutionary relationship in the genus *Bordetella*. *Molecular Microbiology* **1**, 301–308.
13. Council of Europe (1998). Feline leukaemia vaccine (inactivated). In *European Pharmacopoeia, 3rd Edition, Suppl. 2*. Strasbourg, France: Council of Europe.
14. Council of Europe (1997). Rabies vaccine (inactivated) for veterinary use. In *European Pharmacopoeia, 3rd Edition*. Strasbourg, France: Council of Europe.
15. Hendriksen, C.F.M., Woltjes, J., Akkermans, A.M., van der Gun, J.W., Marsman, F.R., Verschure M.H. & Veldman, K. (1994). Interlaboratory validation of *in vitro* serological assay systems to assess the potency of tetanus toxoid in vaccines for veterinary use. *Biologicals* **22**, 257–268.
16. EDQM (1997). Alternative potency testing and other possible related quality issues for veterinary clostridial vaccines. *Pharmeuropa*, Special Issue, Bio 97-1. Strasbourg, France: European Department for the Quality of Medicines.
17. Beckmann, R. & Cussler, K. (1994). Potency testing of erysipelas vaccines in mice. ELISA versus challenge. *ALTEX* **11**, Suppl., 39–45.
18. Council of Europe. (1997). Tetanus vaccine for veterinary use. In *European Pharmacopoeia, 3rd Edition*. Strasbourg, France: Council of Europe.
19. Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H. & Zucco, F. (1995). Practical aspects of the validation of toxicity test procedures. The report and recommendations of ECVAM workshop 5. *ATLA* **23**, 129–147.
20. Balls, M. & Karcher, W. (1995). The validation of alternative test methods. *ATLA* **23**, 884–886.
21. Fentem, J. & Balls, M. (1997). The ECVAM approach to validation. In *Animal Alternatives, Welfare and Ethics* (ed. L.F.M. van Zutphen & M. Balls), pp. 1083–1089. Amsterdam, The Netherlands: Elsevier.
22. Curren, R.D., Southee J.A., Spielmann, H., Liebsch M., Fentem, J.H. & Balls, M. (1995). The role of prevalidation in the development, validation and acceptance of alternative methods. *ATLA* **23**, 211–217.
23. Balls, M., Blaauboer, B., Brusick, D., Frazier, J., Lamb, D., Pemberton, M., Reinhardt, C., Roberfroid, M., Rosenkranz, H., Schmid, B., Spielmann, H., Stammati, A-L. & Walum, E. (1990). Report and recommendations of the CAAT/ERGATT workshop on the validation of toxicity test procedures. *ATLA* **18**, 313–337.
24. Fentem, J.H., Prinsen, M.K., Spielmann, H., Walum, E. & Botham, P.A. (1995). Validation: lessons learned from practical experience. *Toxicology in Vitro* **9**, 857–862.
25. WHO (1997). Validation of potency assays for toxoid vaccines based on *in vitro* methods for measuring antitoxin in the sera of vaccinated animals. In *Manual of Laboratory Methods for Potency Testing of Vaccines Used in the WHO Expanded Programme on Immunization*, pp. 175–178. WHO/VSQ/97.04. Geneva, Switzerland: WHO.
26. Council of Europe (1996). Analytical Validation. In *Pharmeuropa Technical Guide*, pp. 28–40. Strasbourg, France: Council of Europe.
27. ICH (1995). *Validation of Analytical Methods: Definitions and Technology. ICH Topic Q 2A (CPMP/ICH/381/95)*. London, UK: ICH.
28. WHO Expert Committee on Biological Standardization (1990). *Technical Report Series, Vol. 800, Requirements for Diphtheria, Tetanus, Pertussis and Combined Vaccines*, pp. 87–179. Geneva, Switzerland: WHO.

Appendix 1

Terminology

Accuracy is the closeness of the agreement between the accepted reference value and the mean of the repeated values found.

Antigenicity is the extent to which an antigen will react with the immune response elicited by the immunogen.

Efficacy is the extent to which a presumed useful effect is achieved. Efficacy must necessarily be measured in the target species.

Immunogenicity is the extent to which an antigen is capable of eliciting a specific type of immune response in the host animal.

Limit of detection is the lowest amount of the biologically active compound in a sample which can be detected, but not necessarily quantitated, as an exact value.

Limit of quantitation is the lowest amount of the biologically active compound in a sample which can be quantitatively determined with appropriate precision and accuracy.

Linearity refers to results which are either directly or, by a well defined mathematical transformation, proportional to, the concentration of the biologically active compound in the sample.

Potency is a measure of some parameter which has been shown to be related directly or indirectly with efficacy. The parameter may be measured by a challenge or non-challenge method in the target species, or in a laboratory animal model or in an *in vitro* assay.

Precision is the closeness of agreement between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions:

— *repeatability (= intra-assay precision)* expresses the precision under the same operating conditions over a short interval of time;

— *intermediate precision (= inter-assay precision)* expresses intralaboratory variations, for example, different days, different analysts, or different equipment;

— *reproducibility* expresses the variance between laboratories (collaborative studies).

Range is the interval between the upper and lower concentrations of the biologically active compound in the sample for which it has been demonstrated that the analytical procedure has a suitable level of precision and accuracy.

Reference is in-house preparation, the activity of which may be expressed relative to a standard preparation or in appropriate units derived from the test method.

Robustness is a measure of its capacity to remain unaffected by small, but deliberate, variations in the method parameters and provides an indication of its reliability during normal usage.

Specificity is the ability to assess unequivocally the biologically active compound in the presence of compounds which may be expected to be present.

Standard is the preparation of defined activity and composition available to any manufacturer, normally through a national or international authority.

Validation is the process by which the reliability and relevance of a procedure are established for a specific purpose.