European Directorate for the Quality of Medicines & HealthCare | Direction européenne de la qualité du médicament & soins de santé

Co-funded by the European Union

EUROPEAN UNION

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

Co-funded and implemented by the Council of Europe

General European OMCL Network (GEON) QUALITY MANAGEMENT DOCUMENT

PA/PH/OMCL (14) 92 R3

ANNEX 3 – VERIFICATION OF OOS RESULTS
ANNEX 3.4 - VERIFICATION OF OUT-OF-SPECIFICATION RESULTS IN ANIMAL TESTING

Full document title and reference	Annex 3.4 to OMCL GL Evaluation and Reporting of Results: Verification of Out-Of-Specification Results in Animal Testing - *PA/PH/OMCL (14) 92 R3*
Document type	Guideline
Legislative basis	Council Directive 2001/83/EC and 2001/82/EC, as amended
Date of first adoption	August 2014
Date of original entry into force	October 2014
Date of entry into force of revised document	10 April 2025
Previous titles/other references / last valid version	Last valid version: *PA/PH/OMCL (14) 92 R2*
Custodian Organisation	The present document was elaborated by the OMCL Network / EDQM of the Council of Europe
Concerned Network	GEON

N.B. This OMCL Quality Management System document is applicable to members of the European OMCL Network only. Other laboratories might use the document on a voluntary basis. However, please note that the EDQM cannot treat any questions related to the application of the documents submitted by laboratories other than the OMCLs of the Network.

VERIFICATION OF OUT-OF-SPECIFICATION RESULTS IN ANIMAL TESTING

Contents

1. INTRODUCTION 2
2. ASSAY VALIDITY AND CONTROLS 2
3. RELEVANCE OF THE FORMAT OF SPECIFICATION LIMITS 3
4. EXAMPLES OF RETEST PROGRAMMES 4
4.1. POTENCY ASSAY (LOWER LIMIT SPECIFICATION) 4
4.2. POTENCY ASSAY (SINGLE DILUTION COMPARISON WITH STANDARD) 5
5. FLOW CHART - OOS STRATEGY 6
6. EXAMPLE OF A RETEST STRATEGY FOR SEROLOGY TEST USING ELISA 7
7. OTHER CONSIDERATIONS 9

1. INTRODUCTION

Retest programmes that involve animal testing should be designed to minimise the use and the suffering of the tested animals as much as possible, in line with Directive 2010/63/EU on the protection of animals used for scientific purposes.

This Annex focuses on reduction of animal use through the retest strategy, embracing the 3R principles. Other 3R approaches should be considered (e.g. testing the bulk instead of several final filling batches, or application of the humane endpoint for symptomatic animals), but are not described in this document. In this respect, in the case of an OOS result, the strategy should be to repeat the *in vitro* test first, provided that enough material is available.

In contrast to the requirements of the core document, the retest strategy described in this document introduces further tolerances to be applied to the specification limit for the potency.

The retest programme depends on the type of product and assay (e.g. human vaccines and veterinary vaccines). Another element to consider for the retest programme is how the authorised specifications were set in consideration of clinical relevance: in some cases, the specification limits for vaccines give the lower/upper values or only lower specification limits.

In general, as the *in vitro* test is well characterised and validated, its contribution to the global variability is lower than the *in vivo* part.

2. ASSAY VALIDITY AND CONTROLS

Adequate internal controls are required for biological assays conducted in vivo or in a combined in vivo and in vitro test system. Internal controls shall be adequately established and qualified during assay verification or validation. Moreover, these internal controls shall be able to control the testing process, detect any deviations occurring during testing and ensure the validity of the results obtained throughout the whole process. Reliable and properly validated controls provide confidence in the results obtained during product testing.
It is important that in vivo and in vitro parts of the test are monitored by appropriate internal controls. Depending on the context of the test, it is important to take into account false positive and false negative results when choosing internal controls.

In the interest of the 3R recommendations, OMCLs should continuously evaluate the need for control groups. For instance, negative control animal groups may be dispensable, once it has been established that there is no significant background reactivity. For the tests combining *in vivo* and *in vitro* assays, are the negative controls used solely to validate the *in vitro* assay considered sufficient to validate the entire test (e.g. control of vaccine, when the potency of the vaccine to be tested is compared only to a reference vaccine)?

If applicable, negative control sera for use in the *in vitro* assay may be collected in sufficient amounts during the test development and validation, for future use in the *in vitro* part of the assay.

For *in vitro* assays, positive controls are mandatory as part of system suitability criteria, particularly when the testing is based on a comparison with a reference standard. However, positive controls may not be required for the *in vivo* assay; this remains at the OMCL's discretion and, above all, must be properly justified.

For combined *in vivo* and *in vitro* tests, the *in vitro* test should ideally employ its own system suitability controls (negative and positive controls are mandatory) rather than relying solely on control material derived from the *in vivo* part of the combined test. This is to permit the identification of invalid or trending *in vitro* tests and therefore to take targeted measures rather than having to repeat both the *in vitro* and *in vivo* parts of the test.

For the *in vivo* assay, examples of positive controls are a BRP vaccine, a reference vaccine provided by the manufacturer or an internal control that has been well characterised and validated.

For the *in vitro* assay, for example in serological assays, negative and positive controls should be reference sera, sera from the manufacturer or internal samples that have been well characterised and validated. In general, acceptance criteria for all controls and control materials used in the *in vivo* or combined *in vivo* and *in vitro* test should be established during assay implementation and validation, to confirm assay validity. Additionally, test performance (established controls) and trends should be monitored using appropriately designed control charts for both the *in vivo* and *in vitro* parts of the test, as applicable.

3. RELEVANCE OF THE FORMAT OF SPECIFICATION LIMITS

If an OMCL uses a different method than the manufacturer's method, the manufacturer's specifications cannot be used and the OMCL must establish its own specification limits. In this case, care must be taken to ensure that the *in vivo* variability of the method is taken into account when establishing specification limits. When obtaining an OOS result, it is also important to ensure that the OOS result is not linked to *in vivo* limits that are too strict and not representative of the variability of the *in vivo* response.

In this context, the relative potency method could be preferred above reporting in absolute terms i.e. geometric mean titre (GMT) and comparison with a lower GMT acceptance limit for the following reasons. Any variability between experiments (each experiment always includes both analysis of the sample(s) to be tested and of the reference standard) that is not related to the individual variability of each animal (age/weight) is more likely to be normalised/corrected when using the ratio method, as any systemic factor(s) that may contribute to variability will most likely impact both the result of the test sample and the reference standard to a similar extent. Such factors will most likely result in much higher variability of potency results when reporting in absolute GMT. As such, the ratio method does not usually require revision of the acceptance limit in case of changes/modifications of the experimental assay. However, random factors may exist that affect either the test

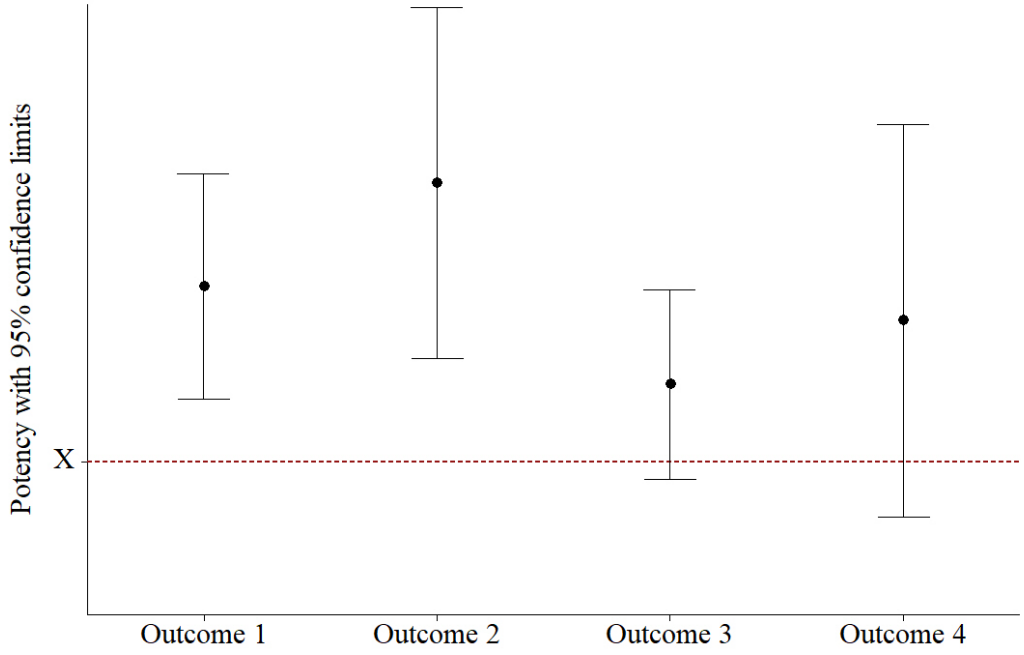
vaccine or reference vaccine group, which could result in a random OOS. Thus, it is important to identify the best method at the OMCL level.

4. EXAMPLES OF RETEST PROGRAMMES

4.1. POTENCY ASSAY (LOWER LIMIT SPECIFICATION)

The following example is for the potency assay of a vaccine with a one-sided potency specification for the lower confidence limit on the estimate: the vaccine complies with the test if the lower confidence limit (*p*-value = 0.95) is not less than X, but no upper limit is set. As for many *in vivo* tests, variability is high (Ph. Eur. validity criterion for 95% CI of this assay is 50-200% of estimate).

Figure 1 - Possible outcomes of initial OMCL test



In outcome 1, all assay validity criteria have been met and the vaccine potency estimate is within specification.

In outcomes 3 and 4, the criterion on the lower confidence limit is not met (values shown are
< X). Moreover, in outcome 4, the confidence limit validity criterion was exceeded. The OMCL would need to repeat the test in order to confirm the OOS and appropriately evaluate the suitability of the batch.

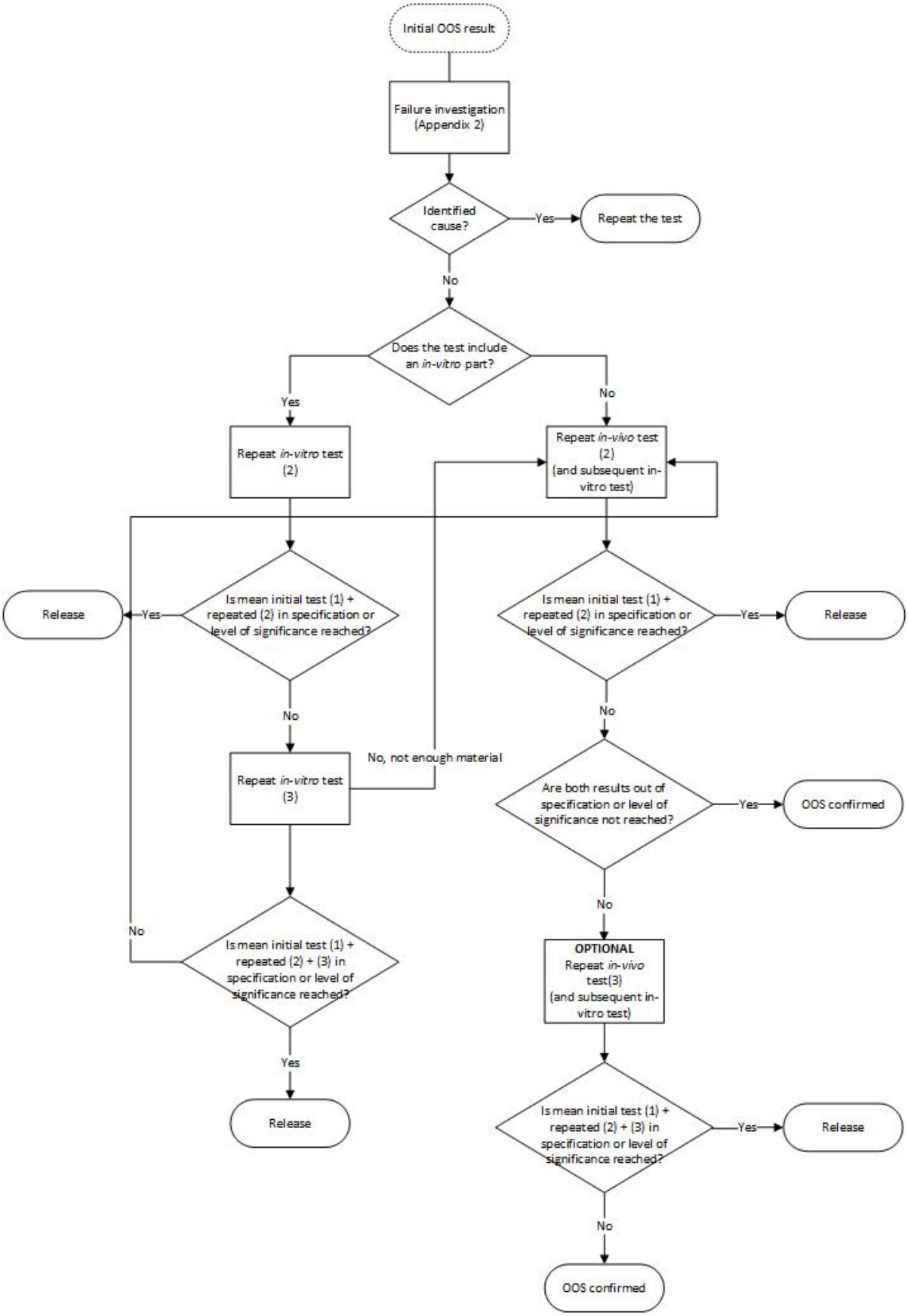
In outcome 2, the criterion on the lower confidence limit is met (value shown is ≥ X), but the width of the confidence interval exceeds the Ph. Eur. validity criterion (confidence interval shown is 45% to 222% of the estimate). In this case, a deeper review of the data (validity criteria) from the MAH and from the OMCL is recommended. As a next step, a decision to accept the results should be justified – even if the confidence interval validity criteria were not met – or a retest should be considered, which should also be justified in view of the objectives of the above-mentioned EU directive.

4.2. POTENCY ASSAY (SINGLE DILUTION COMPARISON WITH STANDARD)

In this example, the immunisation capacity of a vaccine is tested by a direct comparison of assay responses with those for the BRP, using two groups of 10 animals. The vaccine complies with the minimum requirement specification if the observed responses in the vaccine group significantly exceed those in the BRP group (p-value ≤ 0.05). In a case where the level of significance is not met in the first run, a retest must be performed. If it is considered appropriate to combine the 20 values obtained for each of the test vaccine and BRP groups, and this combination leads to an acceptable p-value in line with the level of significance, the vaccine can be released. An example with analysis performed using CombiStats is displayed in Appendix 1.

5. FLOW CHART - OOS STRATEGY

Figure 2 – OOS Strategy



After an initial OOS result is discovered, the first step is to start a failure investigation. The aim is to investigate possible sources of error(s) in the testing process. The evaluation is carried out in collaboration between the technical staff and the supervisor by means of a documented investigation. A Guide is provided in Appendix 2, to be considered as an example of possible questions to be investigated. It is left to the OMCL to select the most appropriate questions/topics to be investigated, depending on the type of test used. The initial OOS result is invalidated if the assigned cause is identified, then the test is repeated. If the cause is not identified, the initial OOS result is still considered valid and confirmed by repeating the in vitro test (reanalysis) or repeating the in vivo test (retesting) and used in the final calculation of the assay. Repeating the in vivo test implies performance of the in vitro test (e.g. in the case of re-immunisation, see Fig. 2).

If the testing includes in vitro and in vivo parts, the retest strategy should be designed to give priority to the repetition of the in vitro test first (reanalysing the same sample). The maximum number of repetitions for the in vitro test should be two, for a total of three determinations including the initial result.

In general, the test sample passes the test if the mean value from initial test and repeat test(s) meets the specifications, or the required significance level is achieved in the repeat test(s). Alternatively, a combined evaluation of the first and repeat test (runs) with CombiStats is also possible. Repetition of *in vitro* tests (reanalysis) requires that enough material (for example serum) is available. If not, the *in vivo* part has to be retested as a next step.
If necessary, new samples should be ordered for retesting.

A maximum of three runs is allowed for the *in vivo* testing but should be restricted to situations where a third test is required to ultimately confirm the result, when the previous results are discrepant.

Individual results can be combined to report a mean value (and confidence limits). According to Ph. Eur. Chapter *5.3*, which describes how to evaluate the homogeneity of results generated in more runs, three approaches can be considered:

- If the individual assay results are not independent, a simple mean (i.e. unweighted mean) can be calculated;
- If results are independent and homogeneous (homogeneity testing, p-value > 0.10), a weighted mean can be calculated taking into account the within-assay variability;
- If results are independent and heterogeneous (homogeneity testing, p-value ≤ 0.10), a semi-weighted mean can be calculated taking into account the within-assay variability as well as the between-assay variability.

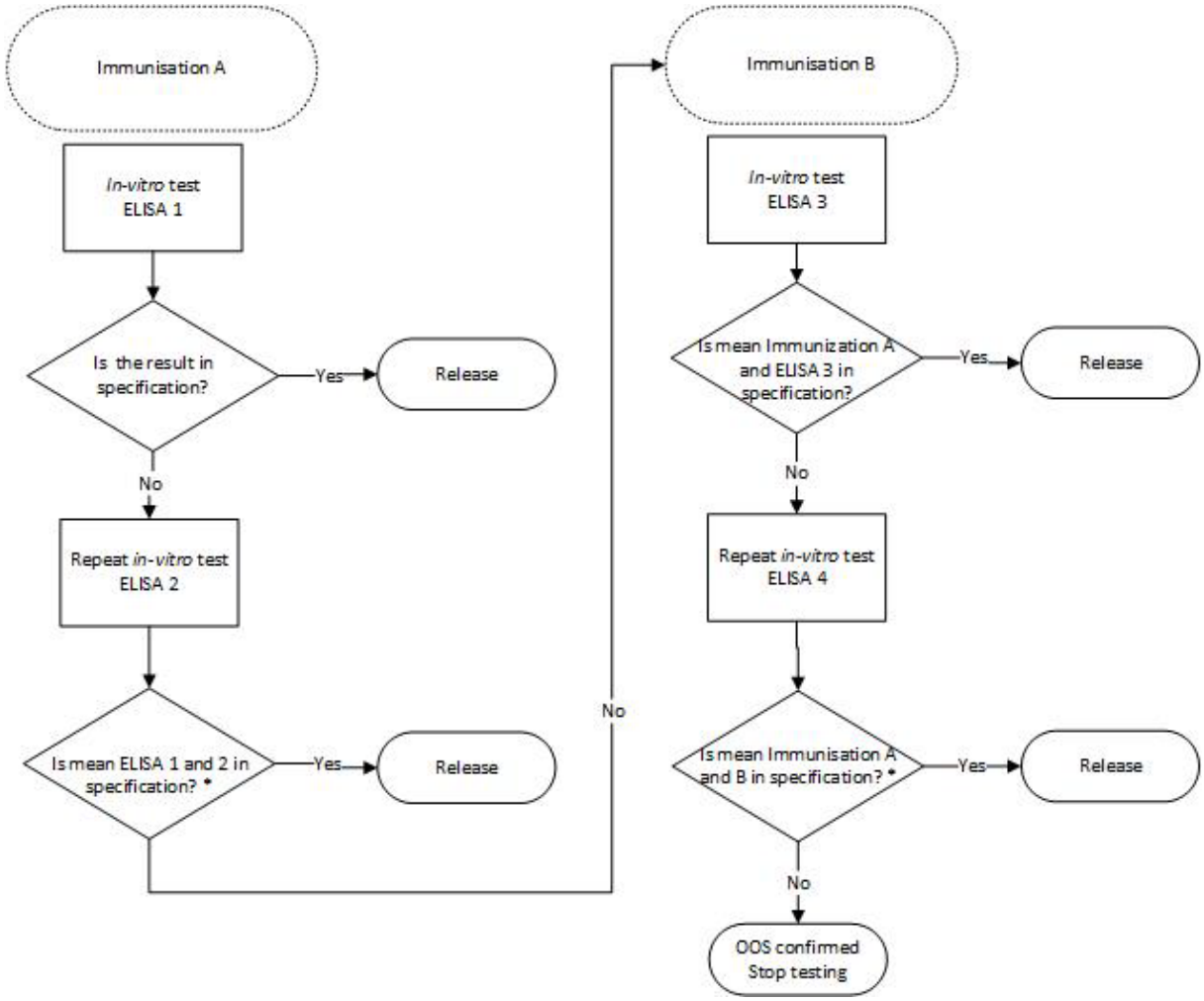
All valid results, including the initial OOS result, are included in the calculation.

6. EXAMPLE OF A RETEST STRATEGY FOR SEROLOGY TEST USING ELISA

Serological assays are composed of an *in vivo* part (animal immunisation and collection of sera) and an *in vitro* part (determination of antibody titre by ELISA). In the event of an OOS result after a first assay (Immunisation A - ELISA 1) it is recommended to carry out a second ELISA on the sera from the first immunisation (Immunisation A - ELISA 2). This must be seen in a 3R context, in order to reduce the number of animals required, compared to directly performing a second immunisation. The average of the results of the two ELISAs carried out on the first immunisation will then be calculated. If the mean result is in

specification, the result will be considered compliant. If the mean result is still OOS, a second immunisation will be performed (immunisation B) and tested by ELISA (Immunisation B - ELISA 3). At this stage, given that two ELISAs have been carried out for immunisation A, while only one ELISA result has been obtained for immunisation B, it is important not to give greater weight to immunisation A when calculating the average result. The mean result for the two immunisations will therefore be assessed on the basis of the average result for immunisation A and the result for immunisation B (ELISA 3). If the mean result is still non-compliant, a second ELISA will be carried out on the B immunisation (Immunisation B - ELISA 4). The final mean result will then be based on the average result of the two immunisations. * For each immunisation, if the results of the two ELISAs are discrepant, a third ELISA can be carried out. The average of the three ELISAs will then be taken into account to determine the status of the batch.

Fig. 3. Example of flowchart for serology retesting strategy



7. OTHER CONSIDERATIONS

Where specifications are not available (e.g. from the MAH), they can be set up by the OMCL during the method validation (no additional run needed). Moreover, test validity criteria could/should be derived from the validation process.

If there is any deviation in the method or in the interpretation/reporting of results (e.g. *p*-value versus relative potency) compared to the manufacturer, the OMCL should define its own specifications.

In cases of potential OOS results, it is recommended to already inform the MAH at this stage, especially if the time for confirmation of the OOS is long (e.g. due to repetition of *in vivo* testing).

In cases of confirmed OOS results, the results previously generated in the OMCL (control charts, trend analysis) should be reviewed and checked for consistency alongside the MAH results, when available.

After confirming OOS results, an investigation should be opened and discussion with the manufacturer started. This is a case-by-case situation.

Appendix 1

Assay

Information

Product	Rabies vaccine inactivated
Method	SNT mouse sera

Remark

Immunisation capacity of test vaccine compared to standard (BRP adjusted to 1 IU/dose). Direct comparison is based on 'antigenic units' that are assigned to the BRP. CombiStats analyses whether the test vaccine is more immunogenic than the BRP.

Preparations

		Information	Potency	
Table	Preparation	ID	Potency	Value
1	Standard	BRP	Assigned	1 IU/dose
2	Sample 1	Test	Assumed	? IU/dose

Raw data

Table 1										
Preparation	Standard									
ID	BRP									
Potency	Assigned									
Potency value	1 IU/dose									
Dose	**Rep.1**	**Rep.2**	**Rep.3**	**Rep.4**	**Rep.5**	**Rep.6**	**Rep.7**	**Rep.8**	**Rep.9**	**Rep.10**
1 dose	10.1	8.5	5.0	6.0	7.1	14.3	1.3	8.5	8.5	14.3

Table 2										
Preparation	Sample 1									
ID	Test									
Potency	Assumed									
Potency value	? IU/dose									
Dose	**Rep.1**	**Rep.2**	**Rep.3**	**Rep.4**	**Rep.5**	**Rep.6**	**Rep.7**	**Rep.8**	**Rep.9**	**Rep.10**
1 dose	0.1	17.0	12.0	10.1	3.0	24.0	20.2	14.3	8.5	12.0

Analysis options

Assay: Single-dose
Model: Wilcoxon-Mann-Whitney test
Response: Quantitative
Design: Completely randomised

Limit test

		Limit tested		
Preparation	**Units**	**Value**	**Probability**	**Level of significance**
Sample 1: Test	IU/dose	1	0.079023	non-significant

Specification: *p*-value (probability) ≤ 0.05
p-value = 0.0790 (FAIL)

Assay

Information

Product	Rabies vaccine inactivated
Method	SNT mouse sera

Remark

Immunisation capacity of test vaccine compared to standard (BRP adjusted to 1 IU/dose). Direct comparison is based on 'antigenic units' that are assigned to the BRP. CombiStats analyses whether the test vaccine is more immunogenic than the BRP.

Preparations

		Information	Potency	
Table	Preparation	ID	Potency	Value
1	Standard	BRP	Assigned	1 IU/dose
2	Sample 1	Test	Assumed	? IU/dose

Raw data

Table 1										
Preparation	Standard									
ID	BRP									
Potency	Assigned									
Potency value	1 IU/dose									
Dose	**Rep.1**	**Rep.2**	**Rep.3**	**Rep.4**	**Rep.5**	**Rep.6**	**Rep.7**	**Rep.8**	**Rep.9**	**Rep.10**
1 dose	4.4	12.2	8.5	8.5	1.8	7.2	7.2	8.5	3.1	10.2

Table 2										
Preparation	Sample 1									
ID	Test									
Potency	Assumed									
Potency value	? IU/dose									
Dose	**Rep.1**	**Rep.2**	**Rep.3**	**Rep.4**	**Rep.5**	**Rep.6**	**Rep.7**	**Rep.8**	**Rep.9**	**Rep.10**
1 dose	10.2	16.9	11.7	10.2	10.2	10.2	8.4	0.1	0.1	8.4

Analysis options

Assay: Single-dose
Model: Wilcoxon-Mann-Whitney test
Response: Quantitative
Design: Completely randomised

Limit test

		Limit tested		
Preparation	**Units**	**Value**	**Probability**	**Level of significance**
Sample 1: Test	IU/dose	1	0.168985	non-significant

Specification: *p*-value (probability) ≤ 0.05
p-value = 0.1690 (FAIL)

Assay

Information

Product	Rabies vaccine inactivated
Method	SNT mouse sera

Remark

Immunisation capacity of test vaccine compared to standard (BRP adjusted to 1 IU/dose). Direct comparison is based on 'antigenic units' that are assigned to the BRP. CombiStats analyses whether the test vaccine is more immunogenic than the BRP.

Preparations

		Information	Potency	
Table	Preparation	ID	Potency	Value
1	Standard	BRP	Assigned	1 IU/dose
2	Sample 1	Test	Assumed	? IU/dose

Raw data

Table 1										
Preparation	Standard									
ID	BRP									
Potency	Assigned									
Potency value	1 IU/dose									
Dose	**Rep.1**	**Rep.2**	**Rep.3**	**Rep.4**	**Rep.5**	**Rep.6**	**Rep.7**	**Rep.8**	**Rep.9**	**Rep.10**
1 dose	10.1	8.5	5.0	6.0	7.1	14.3	1.3	8.5	8.5	14.3
1 dose	4.4	12.2	8.5	8.5	1.8	7.2	7.2	8.5	3.1	10.2

Table 2										
Preparation	Sample 1									
ID	Test									
Potency	Assumed									
Potency value	? IU/dose									
Dose	**Rep.1**	**Rep.2**	**Rep.3**	**Rep.4**	**Rep.5**	**Rep.6**	**Rep.7**	**Rep.8**	**Rep.9**	**Rep.10**
1 dose	0.1	17.0	12.0	10.1	3.0	24.0	20.2	14.3	8.5	12.0
1 dose	10.2	16.9	11.7	10.2	10.2	10.2	8.4	0.1	0.1	8.4

Analysis options

Assay: Single-dose — Model: Wilcoxon-Mann-Whitney test

Response: Quantitative — Design: Completely randomised

Limit test

		Limit tested		
Preparation	**Units**	**Value**	**Probability**	**Level of significance**
Sample 1: Test	IU/dose	1	0.043519	*

Specification: *p*-value (probability) ≤ 0.05
p-value = 0.0435 (PASS)

Appendix 2 – Guide for failure investigation of *in vivo* testing of initial OOS results

Sample information
-Medicinal product: -Batch number: -Expiry date: -Other:
Analytical procedure
-Type of test: -Standard Operating Procedures/References (Pharmacopoeia, guidelines, etc.): -Other:
Result(s)
-Release specification value(s): -Assay validity criteria: -Suspected OOS result(s):
OOS investigation - Investigation on errors in test performance (CAPA)
-Animal housing and environmental issues: ○ Insufficient/inadequate space (such as overcrowding, inaccessible areas, absence of isolated environments to manipulate animals, etc.) ○ Inappropriate levels of temperature and humidity ○ Inadequate air conditioning systems ○ Inadequate individually ventilated cages ○ Inappropriate conditions of lighting ○ External disturbances (such as noise and vibration) ○ Wrong cleaning and maintenance ○ Inadequate feeding, watering and care ○ Incorrect security measures ○ Violations of ethical standards and animal welfare guidelines ○ Other: -Animal handling issues: ○ Wrong shipping, transportation and delivery ○ Inadequate rest periods ○ Animal identification errors ○ Inadequate grouping ○ Effects of variability within animal groups ○ Inadequate animal training/handling by different personnel/staff turnover ○ Operator errors (such as weighing errors, using improper techniques during procedures, errors in preparation and use of samples, errors in the collection of biological samples, inadequate manipulation of animals, incorrect equipment used, etc.) ○ Inadequate documentation and data collection (such as recording of weight, counts of deceased and surviving animals, etc.) ○ Mistakes in following standard operating procedures during the testing ○ Other:

- Animal health issues:
  - ○ Transportation stress
  - ○ Inadequate environmental enrichment
  - ○ Infectious diseases (such as microbial contamination, food and water contamination)
  - ○ Inadequate veterinary and staff care
  - ○ Interference from cage mate
  - ○ Mishandling during housing or experimentation
  - ○ Other:

Other possible causes:

OOS final assessment and Decision on the retest programme

○ **<u>The cause(s) of an OOS has been identified</u>**

➢ The initial OOS result is invalidated and the test is repeated

NB. Define corrective and/or implemented action(s):

○ **<u>The cause(s) of an OOS has not been identified</u>**

➢ The initial OOS result is still considered valid and is to be confirmed:

(1) <u>if the test includes an *in vitro* part</u>

○ repeat *in vitro* test (reanalysis of the same sample)

NB. Define strategy (methods/parameters, number of repetitions, approach for data analysis, etc.):

○ repeat *in vivo* test (re-immunisation) and retest in *in vitro* test

NB. Define strategy (methods/parameters for re-immunisation, methods/parameters for reanalysis of the *in vitro* test, *number of repetitions, approach for data analysis, etc.):*

(2) <u>repeat *in vivo* test</u>

○ *NB. Define strategy (methods/parameters, number of repetitions, approach for data analysis, etc.):*

Technician(s) *(Signature, Date)*

Supervisor *(Signature, Date)*