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 (21) 02 R3

ANNEX 2 – EVALUATION OF RESULTS FROM QUANTITATIVE TESTING

Full document title and reference	Annex 2 to OMCL GL Evaluation and Reporting of Results: Evaluation of Results from Quantitative Testing - PA/PH/OMCL (21) 02 R3
Document type	Guideline
Legislative basis	Council Directive 2001/83/EC and 2001/82/EC, as amended
Date of first adoption	December 2022
Date of original entry into force	March 2023
Date of entry into force of revised document	/
Previous titles/other references / last valid version	/
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.

EVALUATION OF RESULTS FROM QUANTITATIVE TESTING

Table of Contents

1. Introduction....2
2. Approach 1: Evaluation of results from testing performed based on known variability of the method...3
2.1 Evaluation of testing results....3
2.1.1.1. Calculation of T1 and T2 thresholds....3
2.1.1.2. Calculation of T1 and T2 using methods with well-known in-house precision....6
2.1.2. Quantitative impurity tests....6
2.1.3. Other quantitative testing....7
3. Approach 2: Evaluation of results based on statistical assay layouts (Ph. Eur. 5.3)....7
4. References....9

1. Introduction

This document is an Annex to the core document "Evaluation and reporting of results", *PA/PH/OMCL (13) 113* (in its current version), and it should be used in combination with it when planning, performing and documenting the evaluation process and reporting of results of quantitative tests (initial testing), as well as retesting as part of an OOS investigation.

The core document describes the steps for generation of testing results and treatment of data obtained from analytical testing which should be followed in order to obtain valid testing results. This Annex describes agreed decision rules for OMCLs when results are evaluated in relation to predefined validity criteria, where these criteria are not defined in the analytical method. This Annex is applicable for evaluation of the results (expressed as numerical values) obtained from quantitative analyses and includes several approaches which are not intended to be all-inclusive. Other approaches may be used for evaluation of the acceptability of test results (initial and retesting), as long as they are predefined, scientifically justified and documented.

When an OOS test result indicates that a sample is suspected to be non-conforming, i.e. it does not meet specification limits, it is necessary to decide whether a retest of that sample is appropriate.

This document should be considered as guidance for OMCLs and should not be taken as a list of compulsory requirements. It is left to the professional judgement and background experience of each OMCL to decide on the most appropriate procedures to be undertaken in order to demonstrate that the evaluation and reporting of results has been well managed.

2. Approach 1: Evaluation of results from testing performed based on known variability of the method

This approach is applicable for evaluation of results from quantitative testing of active pharmaceutical ingredients or finished products (e.g. assay and related substances/impurities) against predefined validity criteria (step 6 of the Initial testing decision tree in Section 4.1.1 of the core document). In this approach, the variability between individual testing results is evaluated against acceptable limits, based on the known variability of the method. The advantage of this approach is that the retest programme can be abandoned if it is evident after the first set of results that the observed repeatability is insufficient to reach a conclusion.

2.1 Evaluation of testing results

2.1.1. Assay

The relative standard deviation (RSD) of the testing results generated following the Initial testing decision tree (Section 4.1.1 of the core document) is compared to predefined validity criteria (acceptance criteria for the RSD values of testing results):

- T1 threshold: maximum RSD value above which the observed repeatability is insufficient to reach a conclusion and the testing can be stopped; an investigation is to be performed.
- T2 threshold: acceptable RSD value below which the average of the results can be calculated.

For RSD values between T1 and T2, additional testing should be performed to improve the reliability of estimated variability of results.

Step 6 of the Initial testing decision tree includes the following decision points:

- If the RSD is not greater than T1, the next step is to check whether the RSD obtained from testing results is greater than T2.
- If the RSD is not greater than T2, then the results are considered valid and the average can be calculated.
    a. If the mean result complies with the specification limit, it is reviewed and authorised for reporting (sample passes).
    b. If the mean result does not comply with the specification limit, an OOS investigation should be conducted (Annexes 3.1 and 3.2).
- If the RSD is greater than T2, additional determinations are performed, up to a maximum of six determinations, and the evaluation steps (described in the previous bullet points) are repeated after each set of results.

If at any stage the RSD is greater than T1, further determinations are useless, because it can be predicted that the RSD will not meet the threshold even after performing up to six determinations. The laboratory should investigate the reason for poor repeatability. As a consequence, the sample can be neither accepted nor rejected. The laboratory has to decide whether to report the testing result(s), provided that it is appropriately justified and documented.

2.1.1.1. Calculation of T1 and T2 thresholds

Examples of calculated values for T1 and T2 for compendial methods and methods without available validation data are given in Tables 1 and 2.

Table 1. Values for T1 (%) depending on the number of determinations for compendial methods/methods with no available validation data

B (%)	**Number of determinations (n)**				
	2	**3**	**4**	**5**	**6**
1.0	1.34	0.95	0.77	0.67	0.60
1.5	2.01	1.42	1.16	1.01	0.90
2.0	2.68	1.90	1.55	1.34	1.20
2.5	3.36	2.37	1.94	1.68	1.50
5.0	6.71	4.74	3.87	3.36	3.00
10.0	13.42	9.49	7.75	6.71	6.00

B = Upper specification limit (%)-100.

The maximum RSD value of the testing results (T1) can be calculated using the following formula:

$$T1 = B \cdot 1.342 \cdot \sqrt{\frac{1}{n-1}} \qquad [1]$$

- B = Upper specification limit (%)-100
- n = number of performed determinations

Table 2. Values for T2 (%) depending on the number of determinations for compendial methods/methods with no available validation data

B (%)	**Number of determinations (*n*)**				
	2	**3**	**4**	**5**	**6**
1.0	0.11	0.30	0.42	0.52	0.61
1.5	0.17	0.44	0.64	0.79	0.91
2.0	0.22	0.59	0.85	1.05	1.22
2.5	0.28	0.74	1.06	1.31	1.52
5.0	0.56	1.48	2.12	2.62	3.04
10.0	1.12	2.97	4.25	5.24	6.08

B = Upper specification limit (%)-100.

T2 value is calculated using quantiles of the t-Student distribution.

The t-value is calculated for a one-sided confidence level (95%) using the T.INV function in Excel:

$$\mathrm{t} = \mathrm{T.INV}(\text{probability}, \text{degrees of freedom}) \quad [2]$$

- For 95.0 level of confidence, probability is 0.95
- Degrees of freedom = n-1
- n = number of performed determinations

The T2 value is then calculated:

$$T2 = B \cdot \frac{\sqrt{n}}{2t} \qquad [3]$$

- B (%) = upper specification limit (%) -100
- t = t-value for the corresponding confidence level and degrees of freedom
- n = number of performed determinations

Examples of calculated values for T1 and T2 for the methods with available validation data are given in Tables 3 and 4.

Table 3. Values for T1 (%) depending on the number of determinations for MAH methods based on 5 degrees of freedom (99% degree of confidence)

Repeatability from validation (%) ν=5 (ICH[1])	df=1	df=2	df=3	df=4	df=5
0.4	1.61	1.46	1.39	1.35	1.32
0.6	2.42	2.19	2.08	2.03	1.99
0.8	3.23	2.91	2.78	2.70	2.65
1.0	4.03	3.64	3.47	3.38	3.31
1.2	4.84	4.37	4.17	4.05	3.97
1.4	5.65	5.10	4.86	4.73	4.64
1.6	6.45	5.83	5.56	5.40	5.30
1.8	7.26	6.56	6.25	6.08	5.96
2.0	8.06	7.29	6.95	6.75	6.62

df: degrees of freedom

Table 4. Values for T2 (%) depending on the number of determinations for MAH methods based on 5 degrees of freedom (95% degree of confidence)

Repeatability from validation (%) ν=5 (ICH[1])	df=1	df=2	df=3	df=4	df=5
0.4	1.03	0.96	0.93	0.91	0.90
0.6	1.54	1.44	1.40	1.37	1.35
0.8	2.06	1.92	1.86	1.82	1.80
1.0	2.57	2.41	2.33	2.28	2.25
1.2	3.08	2.89	2.79	2.73	2.70
1.4	3.60	3.37	3.26	3.19	3.15
1.6	4.11	3.85	3.72	3.65	3.60
1.8	4.63	4.33	4.19	4.10	4.05
2.0	5.14	4.81	4.65	4.56	4.49

df: degrees of freedom

T1 and T2 can be calculated using the critical F-values, under the assumption that the observed repeatability is not significantly worse than that reported in the dossier:

$$T = RSD_{method\ validation}\% \cdot \sqrt{F_{critical}} \qquad [4]$$

$F_{critical}$ is critical F-value when two variances are compared (from method validation and from obtained results) at a specified probability for the relevant degrees of freedom.

$$F_{critical} = \text{F.INV}(\text{probability}, \text{degrees of freedom1}, \text{degrees of freedom2}) \quad [5]$$

- for T1 values confidence level is set to 99%, probability is 0.99
- for T2 values confidence level is set to 95%, probability is 0.95
- degrees of freedom 1 = n_1-1, n_1 = number of performed determinations
- degrees of freedom 2 = n_2-1, n_2 = number of determinations performed for repeatability testing during method validation

2.1.1.2. Calculation of T1 and T2 using methods with well-known in-house precision

In cases where testing is performed using methods with well-known in-house precision (e.g. in OCABR routine analysis, monitored by control charts), precision estimates from in-house validation or from control charts can be used for calculation of T1 and T2.
T1 is calculated using formula [1]:

$$T1 = B \cdot 1.342 \cdot \sqrt{\frac{1}{n-1}}$$

- B = Upper specification limit (%)-100
- n = number of performed determinations

T2 can be calculated from the precision estimate:

$$T2 = 2 \cdot RSD$$

[6]

- RSD = relative standard deviation estimated during in-house method validation or from control charts

2.1.2. Quantitative impurity tests

The approach described above may be adapted to quantitative impurity tests where an analytical/instrumental response is obtained (e.g. peak area). If validation data are available, T1 and T2 are calculated using formula [4]; if no validation data are available the maximum value for RSD can be considered to be 10% when results are found near the upper specification limit or 15% when results are found near the limit of quantitation (LOQ)[1] or reporting threshold[2]. Responses below the LOQ or with a signal-to-noise ratio of less than 10 should be disregarded.

[1] Quantitation limit, as defined in "Validation of Analytical Procedures: Text and Methodology Q2 (R1)", ICH, 1994.

[2] Reporting threshold, as defined in the Ph. Eur. Chapter 5.10. Control of impurities in substances for pharmaceutical use".

2.1.3. Other quantitative testing

In case of validation data are not available and specification limits cannot be expressed as % of declared content, e.g. optical rotation, loss on drying, relative density, refractive index, viscosity, T1 and T2 are not applicable. Therefore, the acceptance criteria should be based on the measurement uncertainty estimated by e.g. equipment precision, precision based on PTS results or previous testing results. The RSD value of the result should not be greater than the expanded uncertainty of measurement, expressed as a relative value.

3. Approach 2: Evaluation of results based on statistical assay layouts (Ph. Eur. 5.3)

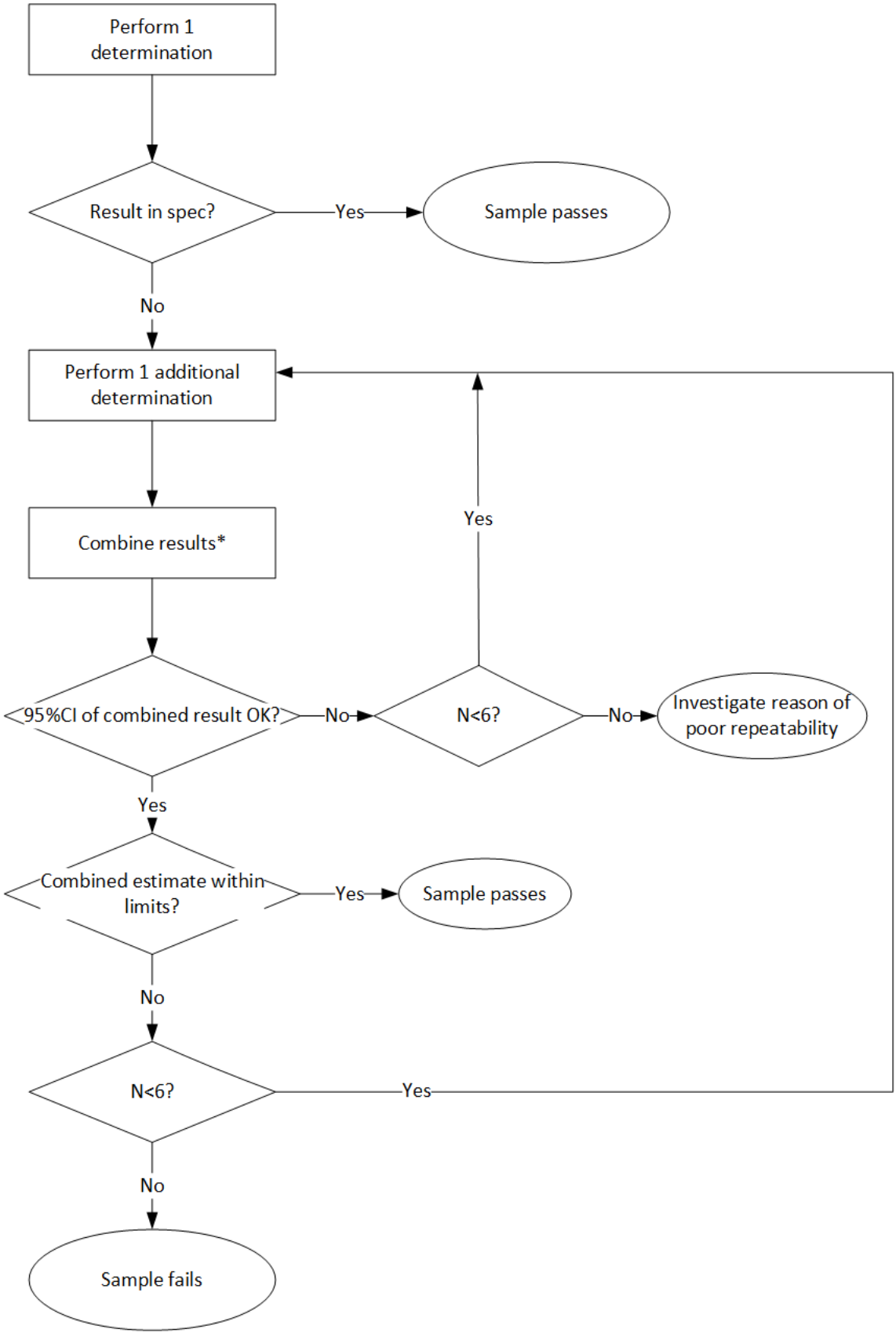
Results from bioassays in a statistical layout according to Ph. Eur. *5.3* (e.g. parallel lines, slope ratio, probit analyses) may be combined according to Ph. Eur. *5.3* (Section 6) or the relevant CombiStats functions. To combine assays, it is desirable that the individual results are obtained using identical or similar assay layouts.

In all cases, the 95% CI of the combined result should meet the validity criteria defined by the method.

CombiStats calculates three types of combinations. The combination used depends on whether the intervals are homogeneous. There are no strict rules as to which of the three should be used, but the following 'rule of thumb' can be of use:

- If the p-value for homogeneity is more than 0.100, the confidence intervals are sufficiently homogeneous to use the weighted combination.
- If the p-value is less than 0.100, the confidence intervals tend to be heterogeneous and it would be better to use the semi-weighted combination.
- The unweighted combination should only be used if there are enough assays, i.e. at least six.

Figure 1 - Decision tree for Approach 2:



* choose mode of combination considering guidance in text

4. References

1. Council of Europe, European Pharmacopoeia
2. Q3A Impurities in new drug substances, ICH Harmonised tripartite Guideline
3. Validation of Analytical Procedures: Text and Methodology Q2 (R1), ICH Harmonised tripartite Guideline