PD CEN/TR 10345:2013

BSI Standards Publication

Guideline for statistical data treatment of inter laboratory tests for validation of analytical methods

bsi.

...making excellence a habit.™

Document Sharing Hub

National foreword

This Published Document is the UK implementation of CEN/TR 10345:2013.

The UK participation in its preparation was entrusted to Technical Committee ISE/102, Methods of Chemical Analysis for Iron and Steel.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2013. Published by BSI Standards Limited 2013

ISBN 978 0 580 81663 5

ICS 03.120.30; 17.020

Compliance with a British Standard cannot confer immunity from legal obligations.

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 December 2013.

Amendments issued since publication

Date	Text affected

Documents
Document Sharing Hub

Full document details: PD CEN/TR 10345:2013

TECHNICAL REPORT

RAPPORT TECHNIQUE

TECHNISCHER BERICHT

CEN/TR 10345

December 2013

ICS 17.020; 03.120.30

Supersedes CEN/TR 10345:2008

English Version

Guideline for statistical data treatment of inter laboratory tests for validation of analytical methods

Guide pour le traitement statistique des données de validation de méthodes d'analyse, issues d'essais interlaboratoires

Richtlinien für die Behandlung von statistischen Daten von verschiedenen Laboren für die Validierung von Analysenverfahren

This Technical Report was approved by CEN on 29 July 2013. It has been drawn up by the Technical Committee ECISS/TC 102.

CEN members are the national standards bodies of Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom.

EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

CEN-CENELEC Management Centre: Avenue Marnix 17, B-1000 Brussels

© 2013 CEN All rights of exploitation in any form and by any means reserved worldwide for CEN national Members.

Ref. No. CEN/TR 10345:2013 E

Documents
Document Sharing Hub

Full document details: PD CEN/TR 10345:2013

Contents

Page

Foreword 3
1 Scope 4
2 Normative references 4
3 Principle 4
4 Preliminary rules 4
5 Procedure 6
6 Report 9
7 General remarks 9
Annex A (normative) **Steps for the validation of a draft Standard** 10
Annex B (normative) **Synoptic of the operations described in Annex A** 13
Annex C (informative) **Examples** 18
Bibliography 44

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

Foreword

This document (CEN/TR 10345:2013) has been prepared by Technical Committee ECISS/TC 102 “Methods of chemical analysis of iron and steel”, the secretariat of which is held by SIS.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. CEN [and/or CENELEC] shall not be held responsible for identifying any or all such patent rights.

This document supersedes CEN/TR 10345:2008.

In comparison with the previous version of CEN/TR 10345, the following significant technical change was made in Annex A: correction of the error in the last sentence of A.2 concerning the appropriate number of significant figures.

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

1 Scope

This Technical Report is a guideline to carry out the statistical evaluation of data from an inter laboratory test for method validation.

Its purpose is to detail the methodology of ISO 5725-1:1994, ISO 5725-2:1994 and ISO 5725-3:1994 for the treatment of the data collected under the conditions used within the ECISS/TC 102 working groups.

NOTE The present document is not a simplification of the ISO 5725 standard, which is the only reference document.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 5725-1:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

ISO 5725-2:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*

ISO 5725-3:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate measures of the precision of a standard measurement method*

3 Principle

An inter laboratory test for method validation is organized at each stage of the development of a standard draft. Changing economic conditions have led to the optimization of the work of the participating laboratories. The principle retained by ECISS/TC 102 is to have three values by participant laboratory: two values obtained in repeatability conditions (day 1) and a third obtained in intra laboratory reproducibility conditions (day 2). The data evaluation requires a complex statistical analysis, which may be very confusing for a non-specialist, even if it is widely detailed in the ISO 5725 standard. Consequently, it seems useful to clarify the methodology of this standard for the above purpose and to underline that difficulties found should be discussed and solved with statisticians.

Values that are identified as statistically abnormal at 99 % (outliers) using numerical Cochran's and Grubbs' tests lead to the elimination of the laboratory that produced them, at the stage at which they are detected: this principle is adopted even though we risk wrongly eliminating one result in one hundred. Nevertheless, it is essential to advise the laboratory concerned about the reasons for these eliminations and to pay particularly attention to this laboratory's results.

Furthermore, in the case of a laboratory which produces values that are determined as statistically significant at 95 % (stragglers) by numerical Cochran's and Grubbs' tests, particular attention should be paid to all the other values produced by this laboratory.

4 Preliminary rules

4.1 First rule ('to be clear')

The inter laboratory test should be adapted in order to meet the following requirements:

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

— to estimate the variances linked to the tested method (repeatability, intra laboratory reproducibility, inter laboratories reproducibility);

— to check that inter laboratories variance is compatible with defined criteria for referee or routine methods (Aim CVR or Max CVR) within their full range of application.

The following shall be imposed on the working groups:

a) minimal number of participating laboratories (8 / 10 / 15… - see ISO 5725-1:1994, Annex B);

b) rigorous implementation of the working programme;

c) appropriate number of significant figures to be given for each transmitted value, in order to allow an optimal statistical data evaluation.

4.2 Second rule ('to be modest')

The statistical treatment shall be performed by application of the ISO 5725 standard, and by following the procedure described hereafter. In order to solve the statistical problems encountered, the help of an expert should be sought. It should be noted that ISO 5725-2 and ISO 5725-3 don't give an exhaustive description of all situations and that they clearly indicate that statistical data treatment should be performed by a person experienced in work planning and in statistical analysis (ISO 5725-2:1994, 6.2). Various situations may arise in practice which require the application of the variance analysis general modes (ANOVA), and these ISO 5725 standards only specify simplified procedures.

4.3 Minimal characteristics of data population

At least two samples should be tested for each concentration range to be determined, (for example between 0,010 and 0,099 we shall have two samples, between 0,10 and 0,99 we shall have two samples and so on), and should never be less than 5 for the full range of values. It is useful, when possible, that the inter laboratory test be performed using certified reference materials (CRMs) representing at least 50 % of the total number of samples to be tested. The remaining samples can be internal reference materials provided by laboratories on condition that their homogeneity has been tested and found to be compatible with inter laboratory test requirements.

In the present economic situation, an inter laboratory test should be planned with at least 8 laboratories from at least 5 different countries.

The values provided by the participating laboratories should have a sufficient number of significant figures in order to enable correct statistical data treatment; although the number of significant figures does not influence the precision of the result, the transmission of rounded values containing fewer figures means that the work cannot be correctly evaluated. Expressed as a w/w percentage, values should generally be written under the form listed below, i.e. containing 3 or 4 significant figures:

xx,xx
x,xxx
0,xxx x
0,0xx x
0,00x xx
0,000 xxx

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

5 Procedure

5.1 Bases

Here we look at the only case where the data population is strictly obtained by the methods defined in Clause 3, that is to say, 3 determinations for each sample and from each laboratory:

— two determinations under repeatability conditions called ‘Day 1,1’ and ‘Day 1,2’;

— the third determination under reproducibility conditions called ‘Day 2’.

Statistical data evaluation is performed for each content level, that is to say in the present case, sample by sample.

5.2 Raw data examination

Raw data shall be typed into a table and then printed. The raw data table should be studied in order to detect potential typing errors in the data supplied by each participating laboratory and/or in the final input stage (ISO 5725-2:1994, 7.2.6).

Performing a normality test of the data population to be tested is recommended.

5.3 Intra laboratory repeatability variance (Cochran’s test)

A first graphical evaluation of the raw data may be performed in order to test the intra laboratory repeatability consistency by using Mandel’s k test (all data included): its only purpose is to get an overview of the data population.

Further statistical treatments are carried out under the hypothesis that the intra laboratory repeatability variances belong to the same normal population. Cochran’s test should therefore be performed (ISO 5725-2:1994, 7.3.3) in order to detect unexpected values of intra laboratory variances, which shall then be discarded so that statistical analysis may be pursued.

Strictly speaking, Cochran’s test should only be used to evaluate a population of measurements obtained in repeatability conditions.

In practice, it is advisable to proceed as follows:

— perform Cochran’s test with ‘Day 1,1’ and ‘Day 1,2’ values, after a normality test for this data population;

— discard laboratories having an unexpected variance.

It is advisable not to perform an iteration when using Cochran’s test, except in the case of a large laboratory population (i.e. greater than 15), or when there is a particular statistical reason to justify it. A common rule sometimes used is that the outliers shall not represent more than 10 % of the whole data.

In practice, in order to have a clear view of the data after the Cochran’s test has been performed, it is advisable either to print a new table of the remaining data or to clearly identify discarded data.

5.4 Intra laboratory reproducibility variance (Grubbs’ test)

Further statistical evaluations are carried out under the hypothesis that intermediate variances (intra laboratory reproducibility) belong to the same normal population. Grubbs’ test should be applied (ISO 5725-2:1994, 7.3.4) in order to detect unexpected means that should then be discarded so that statistical analysis may be pursued.

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

In practice, it is advisable to proceed in the following order:

— apply Grubbs' test to the 'daily means' data, that is to say to the pair of values '(Day 1,1 + Day 1,2)/2' and 'Day 2' for all of the laboratories;

— firstly, test the highest 'mean' after having confirmed that the data population is normal;

— if the test is positive, discard the laboratory concerned;

— secondly, test the lowest 'mean';

— if the test is positive, discard the laboratory concerned;

— if neither of the two former tests detect an unexpected 'mean', perform the test with the two highest 'mean' values;

— if the test is positive discard the laboratory(ies) concerned;

— perform the test with the two lowest 'mean' values;

— if the test is positive discard the laboratory(ies) concerned.

NOTE 1 It is advisable not to perform iterations when using Grubbs' test.

NOTE 2 It is not necessary to perform the test whenever the value or the two values to be tested come from a population containing respectively a second or a third identical value.

5.5 Inter laboratory reproducibility variance (Grubbs' test)

A first graphical evaluation of the raw data may be performed in order to test the inter laboratory consistency using Mandel's h test (all retained data remaining after the application of Cochran and Grubbs' tests are included).

Further statistical treatments are carried out under the hypothesis that the laboratories means belong to the same normal population. Grubbs' test should be performed (ISO 5725-2:1994, 7.3.4) in order to detect unexpected laboratories mean values which shall then be discarded so that statistical analysis may be pursued.

In practice, it is advisable to proceed in the following order:

— firstly, test the highest 'laboratory mean', '(Day 1,1 + Day 1,2 + Day 2)/3', after having confirmed that the data population is normal;

— if the test is positive discard the laboratory concerned;

— secondly, test the lowest 'laboratory mean';

— if the test is positive discard the laboratory concerned;

— if neither of the two former tests detect an unexpected 'laboratory mean', perform the test with the two highest 'laboratory mean' values;

— if the test is positive discard the laboratories concerned;

— perform the test with the two lowest 'laboratory mean' values;

— if the test is positive discard the laboratories concerned.

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

NOTE 1 It is advisable not to perform iterations when using Grubbs' test.

NOTE 2 It is not necessary to perform the test whenever the value or the two values to be tested come from a population containing respectively a second or a third identical value.

5.6 Retained data examination

The application of Cochran and Grubbs' tests may lead to the elimination of raw data from some laboratories. These eliminations have been performed for each content level tested: therefore it is necessary to carry out a critical inter level examination that may lead to discard all data from one or more laboratories if it is established that multiple outlier or straggler values come from these laboratories. These eliminations have to be justified.

In practice, in order to have a clear view of the remaining data, it is advisable either to print a new table of the retained data or to clearly identify discarded data.

5.7 Calculation of the variances associated to the tested method

Calculations are performed with a strict application of ISO 5725-3:1994, (Annex C); it generally results in three variances labelled Vr, VRw and VR.

It is the result of a variance analysis evaluation assessment (ANOVA) which assumes that all the former stages were performed, and which presupposes that intra laboratory repeatability (residual, Vr), intra laboratory reproducibility (VRw) and inter laboratory reproducibility variances (VR) are discernible and quantifiable within the test experiment conditions.

5.8 Treatment for unexpected calculated variances

Despite the implementation of all of the above specifications there could still exist some anomalies:

— Variances are normally graded in the order Vr < VRw < VR; if this is not the case it could be that the results of the calculation are not realistic because an hypothesis was not verified. In particular this may occur when the conditions to detect a difference between the intra laboratory reproducibility variance and the residual variance (intra laboratory repeatability variance) don't exist; then strictly speaking, the ISO 5725 standard formulae are not applicable and consequently only one global residual intra laboratory variance can be calculated. The same approach should be taken concerning VRw and VR.

— Variances are of course positive numbers (they are the sum of square numbers); nevertheless, the ISO 5725 standard formulae can lead to negative values. In such cases these results should not be taken into account because some hypotheses may not have been satisfied or because of the lack in variability of the transmitted data (rounded values or an insufficient number of significant figures).

— etc.

These anomalies should be dealt with by an expert and the resulting position taken by the committee could be that standard deviation values for repeatability and/or reproducibility don't fit the criteria for publication (ISO 5725-2:1994, 7.7.2).

5.9 Estimation of a function linking variance and level

Calculations are performed in strict agreement with the specifications of ISO 5725-2:1994, 7.5.

Particular attention should be paid to the correlation coefficient of the functions between:

— $\lg m = a + b \cdot \lg r$

— $\lg m = a + b \cdot \lg Rw$

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

— $\lg m = a + b \cdot \lg R$

Typically, the value of each correlation coefficient should be at least greater than 0,9.

Nevertheless, correlation coefficient with values from 0,7 to 0,9 can be admitted, after consensus.

Values lower than 0,7 should be rejected as they show a lack of correlation. In such cases, only individual r, Rw and R should be edited in the validated method (the edition of the smoothed r, Rw and R values is not allowed).

6 Report

The report of the statistical evaluation shall be submitted to all of the participating laboratories, so that they can verify that there was no error in the transcription of their results and in order to get their opinion concerning the evaluation performed.

The report of the working group convenor shall include:

— complete statistical report;

— participating laboratories corresponding remarks;

— comments and answers of the convenor concerning these remarks.

This report is then sent to the technical committee together with the method accompanied by the remarks and technical comments of the working group members.

7 General remarks

Most of the inter laboratory tests for the validation of standard methods have not been evaluated under the rules of the present document. Indeed there is no standardized method which mentions the impossibility of the evaluation of one of the three variances, based on the data produced by an inter laboratory test (variances and the corresponding standard deviations calculations are systematically performed and published). It is important to verify if the present document is mentioned in a particular standard for the section concerning the statistical evaluation of the data issued from the corresponding validation inter laboratory test.

It is possible to perform the statistical evaluation of the data in accordance with ISO 5725 using software. Nevertheless, it should be kept in mind that there is no software able to make decisions concerning abnormal situations.

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

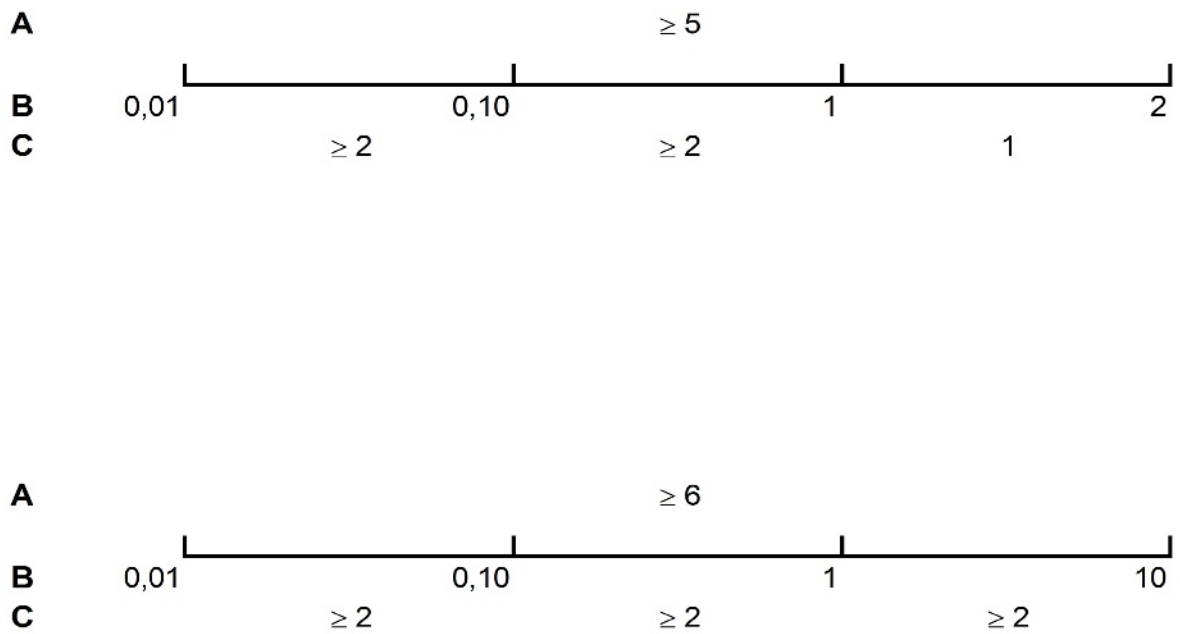
Annex A
(normative)

Steps for the validation of a draft Standard

A.1 Decision from a TC to create a working group

Selection of N laboratories: minimum 8 laboratories from not less than 5 countries.

Selection of P samples: a minimum of 5 samples to cover the entire range. If the range is greater than a factor of 10 then for each sub-range of a factor of 10 there shall be at least 2 samples. See example below:



Key

A minimum number of samples

B content range (in %)

C minimum number of samples for each sub-range of a factor of 10

Figure A.1 — Example of selection of P samples

Selection of samples: minimum 50 % CRM (if possible).

A.2 Laboratories performance on the specified tests

Rigorous application of the draft standard.

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

Rigorous application of the previous scheme in order to produce: two results under repeatability conditions “Day 1,1” and “Day 1,2” and one under reproducibility conditions “Day 2”.

Transmission of results (%) with the appropriate number of figures: 3 significant figures for values < 0,100 % and 4 significant figures for values ≥ 0,100 %.

A.3 Statistical work to be done at each level (for each of the P samples)

A.3.1 General

General table including all values produced.

Examination of the raw data, for typing errors detection.

Normality test.

A.3.2 Intra laboratory repeatability variance (Cochran's test: ISO 5725-2:1994, 7.3.3)

Apply only to the set of N couples of values “Day 1,1” and “Day 1,2”.

Discard the laboratory(ies) which shows unexpected variance (outliers).

If the outliers represent more than 10 % of the total number of the laboratories, ask for the advice of a statistician.

Edit a new table containing the set of laboratories having the same “intra laboratory repeatability variance” for further tests.

Normality test.

A.3.3 Intra laboratory reproducibility variance (Grubbs' test: ISO 5725-2:1994, 7.3.4)

Apply only to the 'daily means': set of 2N values “(A+B)/2” and “C”.

Test the highest 'daily mean'.

Discard the laboratory that shows the unexpected variance (outlier).

Test the lowest 'daily mean'.

Discard the laboratory that shows the unexpected variance (outlier).

If neither of the two tests above detect an unexpected 'mean':

Test the couple of highest 'daily mean'.

Discard the laboratory(ies) that show the unexpected variances (outliers).

Test the couple of lowest 'daily mean'.

Discard the laboratory(ies) that show the unexpected variances (outliers).

Edit a new table containing the set of laboratories having the same “intra laboratory reproducibility variance” for further tests.

Normality test.

AVA Documents
Document Sharing Hub

Full document details: PD CEN/TR 10345:2013

A.3.4 Inter laboratories reproducibility variance (Grubbs' test: ISO 5725-2:1994, 7.3.4)

Apply only to the “laboratory means”: set of N values “(A+B+C)/3”.

Test the highest “mean value”.

Discard the laboratory that shows the unexpected variance (outlier).

Test the lowest “mean value”.

Discard the laboratory that shows the unexpected variance (outlier).

If neither of the two tests above detect an unexpected 'mean':

Test the couple of highest “mean values”.

Discard the laboratories that show the unexpected variances (outliers).

Test the couple of lowest “mean values”.

Discard the laboratories that show the unexpected variances (outliers).

Edit a new table containing the set of laboratories having the same inter laboratory reproducibility variance.

A.3.5 Inter level laboratory performance

If a laboratory is a straggler or an outlier on several levels the convenor should consider the possibility of rejecting all results from that laboratory.

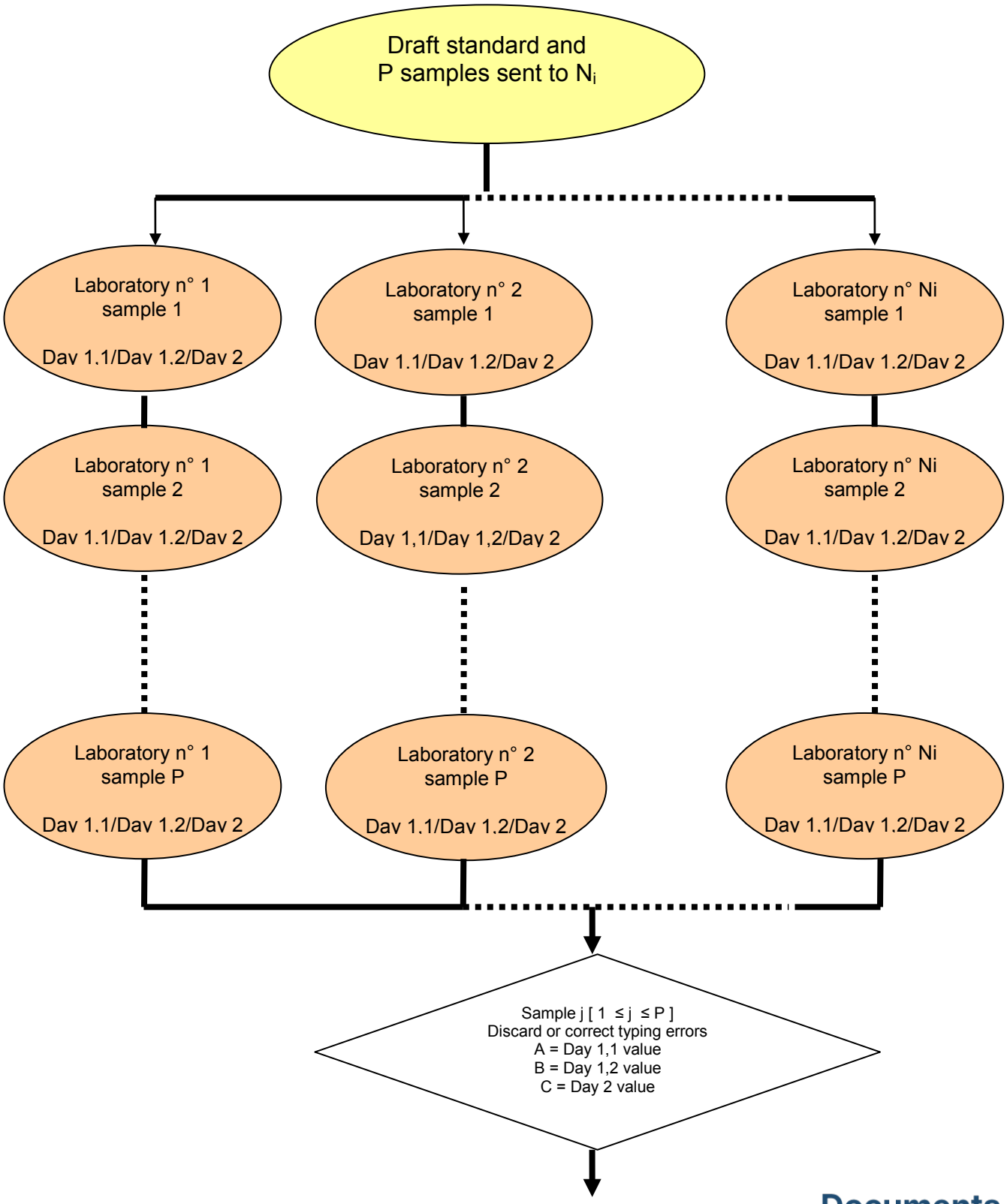
Consider, if possible, the levels of the h and k Mendel's consistency statistics for the laboratory(ies) to be discarded as a tool to support the decision.

Restart the entire processes of statistical evaluation, without this (or these) laboratory(ies).

Documents
Document Sharing Hub
Full document details: PD CEN/TR 10345:2013

Annex B
(normative)

Synoptic of the operations described in Annex A



Documents
Document Sharing Hub

Full document details: PD CEN/TR 10345:2013