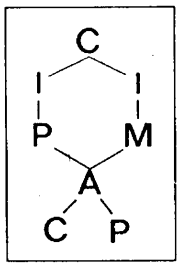
CIPAC

COLLABORATIVE INTERNATIONAL PESTICIDES ANALYTICAL COUNCIL LIMITED

Commission Internationale des Méthodes d'Analyse des Pesticides (CIMAP)

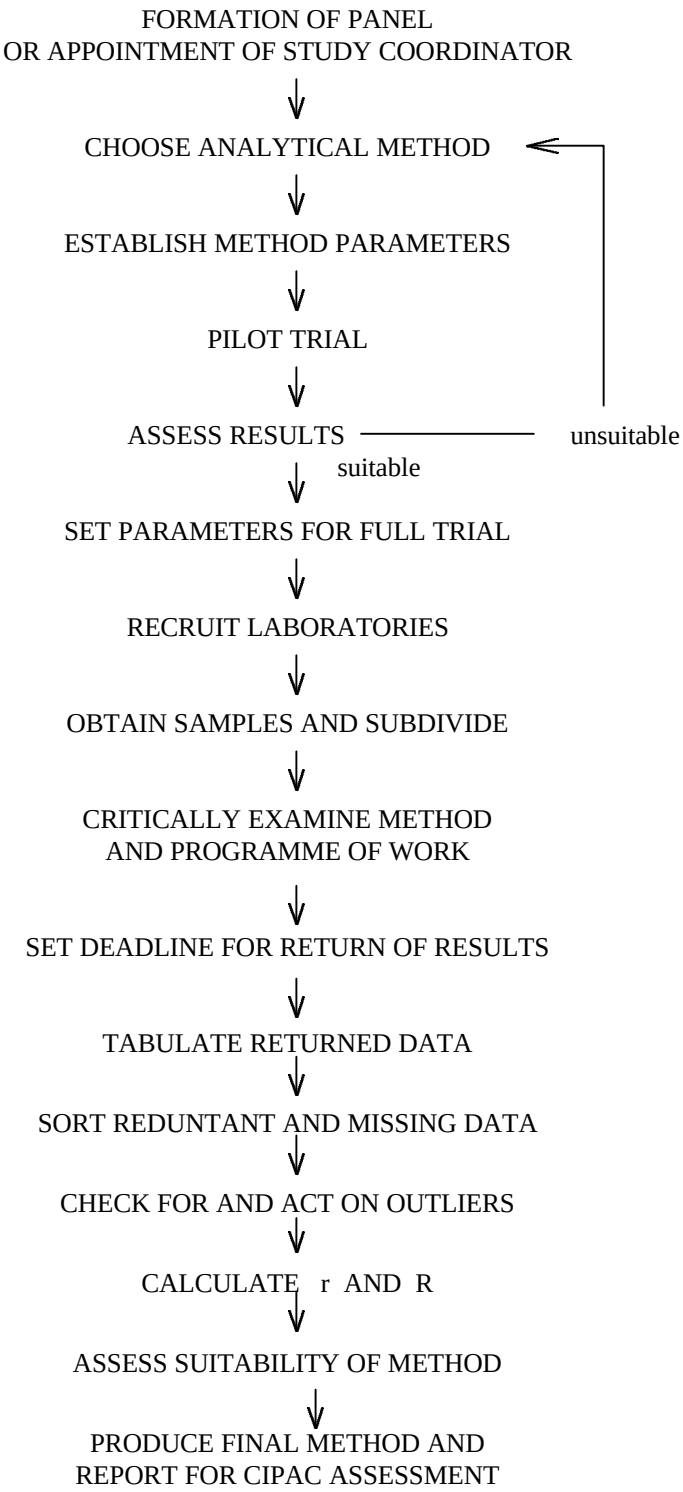


CIPAC Guidelines for Collaborative Study Procedures for Assessment of Performance of Analytical Methods

This document was originally formulated as CIPAC 3092 by Mr. D.S. Farrington, MAFF, Starcross, Devon, England, and revised as CIPAC 3426 by Mr. P. Baker, ICI, Yalding, Kent, England.
Mr. P. Baker, ICI, Yalding, Kent, England ' Dr. N.C. Franklin and Dr. K. Pavel, Bayer AG, Leverkusen, West Germany, prepared the final version submitted to CIPAC in May 1989.

CONTENTS

CIPAC COLLABORATIVE STUDY PROTOCOL



1.AIM OF THE DOCUMENT

This document outlines the procedure which should be followed when performing CIPAC collaborative studies on pesticidetechnical materials and formulated products. It is based on information from existing CIPAC practice, and includes the recom-mendations of the IUPAC Workshop of Harmonization of Collaborative Analytical Studies,' Geneva, Switzerland, May 4-5, 1987 [1]. It takes into consideration ISO 5725-(1986) [2] according to which the statistical analysis is carried out.

2. INTRODUCTION

The aim of each CIPAC Committee is to produce, for each of its allocated pesticides, a satisfactory metho of analysis which has been verified by international collaborative study. Such methods are required in case of dispute in trade, and to support the use of national and international specifications.

In addition to recommending procedures for active ingredient determination, the national PAC or the pane may be required to undertake development and collaborative testing of methods of analysis for impurities.

3. CHOICE OF METHODS FOR COLLABORATIVE STUDY

A collaborative study is defined as: An inter-laboratory study in which each laboratory uses the defined method of analysis to analyse identical portions of homogeneous materials to assess the performance characteristics obtained using that method of analysis.

It should be stressed that a collaborative study is primarily a test of the method and not of the laboratory. The method must be followed as closely as practicable, and any deviations from the method as described, no matter how trivial they may seem, must be noted on the report form.

The choice of method may be limited, as there may be only one available, or a literature search may reveal one method with particular advantages. often "in housell methods can be drawn on from members, and in particular from manufacturers; these have the advantage that at least one member is able to advise on possible areas of difficulty.

Analytical methods for the determination of concentrations of chemical components should possess the following features:

a) Applicability; the method should generally be applicable to a range of products and concentration levels, yet retain selectivity and sensitivity.
b) Reliability; the accuracy and precision (both repeatability and reproducibility)should be acceptable
c) Practicability; speed and low cost are an advantage and the method should use readily available reagents and equipment and be safe in the hands of trained operators.

A single satisfactory method may not always be found for the analysis of all available technical and formulated samples of a particular pesticide. In this case problem formulations should be isolated and deal with using another analytical procedure.

If at this stage one method of analysis appears to be suitabie for consideration the panel or study coordinator should obtain further specific information befote embarking on any inter-laboratory work.

Collaborative studies require substantial effort and should only be applied to methods-that have receive adequate prior testing. This should include estimates of the total withinlaboratory standard deviation of th analytical results over the concentration range of interest. In addition information on systematic erro should be obtained.

The prior testing should include the following if possible:

1. The ability of the method to measure the physical and chemical forms of the analyte likely t be present in the Materials
2. The effect of other substances that are likely to be present in appreciable concentrations i production Materials and which may interfere with the determination.
3. The results obtained by applying the method to reference Materials
4. The recovery from synthetic material specially prepared with known amounts of analyte.
5. The results of a comparison of the method with existing tested methods intended for similar Materials
6. The procedures specified for calibration and blank correction should not cause a signifikar bias.
7. The ruggedness of the method should be checked by the originating laboratory; for thi purpose Youden's procedure [3] is recommended.

The method then should be clearly and unambiguously written up taking into account ISO Guide 18-1978 (E) [4].

Normally the Provision of all information would be the responsibility of the originating laboratory or study coordinator. In the absence of any clearly defined originating laboratory the panel or national PAC chairman should endeavour to obtain a study coordinator to sponsor the method at this preliminary stage and undertake the work outlined above.

4.PILOT TRIAL

Before initiating even a small scale trial, it is advisable to consult with other interested parties to ensure the mutual acceptability of the proposed method.

Having established a method which can be reasonably expected to work, small scale pilot trials should take place involving only three or four laboratories who, though not all experienced in the practical work, are familiar with the setting up of the method.

These laboratories should follow the agreed analytical procedure, but have the freedom to suggest modifications to the inethod if they seem to be a valuable improvement. These should be reported to the study coordinator. Restricting the trial to a small number of participants speeds the flow of information if any difficulties are encountered. Thus the method actually adopted for full collaborative trial may differ slightly from the original method.

Results of this study may then be used in calculation of repeatability and reproducibility, without preliminary outlier tests. They also give an indication of the agreement between duplicate measurements of the same sample which collaborators should aim to achieve in the full trial.

From the results of the,pilot trial the panel or national PAC should decide if the method is suitable to go forward to a full collaborative trial. If it is considered to be unsuitable, the method must either be modified or a different method chosen. The pilot trial would then have to be repeated with a "new" method.

5.FULL COLLABORATIVE TRIAL

5.1 Recruitment of Laboratories

Laboratories should be invited to participate by means of a notice (CIPAC Information Sheet) sent out via the secretary of CIPAC. This notice should not only cite the principle of the method, but also, in the case of instrumental methods, give the important parameters which need to'be followed. An assurance should be obtained from prospective collaborators, in advance, that they can essentially follow the method.

Ideally, laboratories participating in the collaborative trial should be chosen at random from those likely to use the test method under consideration, but in practice only a limited number of laboratories are likely to be willing to take part. If the study is intended for international use, laboratories from different countries'should participate. Laboratories invited to participate should have personnel experienced in the analytical procedures employed, though not necessarily with the method itself.

The recommended minimum number of participating laboratories in a collaborativestudy is 8. When it is impossible to obtain this number, the study may be conducted with less, but with an absolute minimum of 5 laboratories, with the accompanying expansion in the confidence limits of the precision. Although it is desirable to have more than 8 laboratories, studies containing more than about 15 laboratories become difficult to administer.

5.2 Preparation

The following points should be considered when drawing up the programme:-

5.2.1 An analytical reference material must be available (Note 1).

5.2.2 The range of levels encountered in practice should be selected for the trial (e.g. technical, 500 g/kg EC, etc).

5.2.3 Having chosen the levels, a number of Materials within these levels should be analysed. If several manufacturers produce material at similar levels, more than one sample may be chosen. Additionally, in order to test the robustness of the method, efforts may be made to select a difficult sample. The minimum number of samples used in a collaborative study is 5. This may be reduced t an absolute minimum of 3 when only a single level of a single matrix, e.g. solely one concentration of solely one type of formulation or technical material is available.

5.2.4 Duplicate analyses should be performed by each laboratory (Note 2).

5.2.5 The weight of material required by each labotatory should be determined.

Note 1 Validated reference Materials should be available either from a manufacturer of the material or fror a specialist supplier of analytical reference Materials (cf CIPAC D, p.186 ff).

5.2.6 One laboratory, usually the originating laboratory or study coordinator, should be responsible for sub-sampling and dispatch of the samples. It is the responsibility of the laboratory organizing distribution to ensure that reference Materials and samples are thoroughly homogenized prior to dispatch. If samples need to be melted before use, the instructions provided should have been validated to ensure that the chemical stability is not affected. For difficult samples, e.g. suspension concentrates,'it may be necessary to provide a procedure for sub-sampling.

5.2.7 Laboratories may be sent additional material and undergo a small scale preliminary trial with which to familiarize themselves before starting on the full scale exercise.

5.2.8 A deadline for the return of results should be indicated.

5.3 Instructions for the Performance of the Analysis

The method should be written up in CIPAC format.

The study coordinator should ensure the following isprovided:

5.3.1 the necessary instructions to collaborators on sub-sampling and/or homogenization of samples, sub samples or reference Materials

5.3.2 any particular steps which should be taken to equilibrate the equipment before use, (e.g. the numbe of injections of samples or standards before starting GLC analysis, and what repeatability of duplicate injections should be attained prior to starting the analysis);

5.3.3 the injection sequence of samples, calibration solutions and the number of calibration solutions required in chromatographic methods, (Note 3);

5.3. any questions which participating laboratories need to be asked in order to ensure the collection of full details of their operating conditions, particularly any Departure from the method;

5.3.5 the equation and or calculations to be used to convert the data obtained into the results to be reported, including the meaning of symbols used in the equation;

5.3.6 the number of signifikant figures required in the reported -results;

5.3.7 special information concerning storage of reference Materials samples or prepared solutions and th estimated she-lf life of solutions and reagents;

Note 2 The procedure which should be adopted is to analyse two samples from each level, with each sample weighed and analysed once, and the procedure repeated at a later date. For chromatographic analyses each individual sample is injected twice, thus giving two sets of results. These duplicate injections are not required for the subsequent statistical analysis, which uses their averages, but they do give a good indication of the efficiency of the chromatography, and provide the panel or national PAC with useful additional information.

Note 3 After the required equilibration programme, the normal injection,sequence adopted for chromatographic methods should be:
Calibration (1), Sample (A), Sample (A), Calibration (2), Sample (B), Sample (B), Calibration(1),where Calibration (1) and Calibration (2) are separate weighings of reference Materials. Thus two results are obtained for each sample, both of which rely on the average values from the calibration injections which bracket them. A preliminary calibration exercise using 0.5 w, 1 w, and 2 w grams of reference material, (w = wt. used in the study) should also be carried out. This will check the linearity of response; the middle calibration solution can be used to check the accuracy of the collaborative trial calibration solutions, provided that these are not subject to signifikant decomposition.

5.3.8 if appropriate, at what time breaks may be taken in the analysis, or where immediate progress to th next stage is essential;

5.3.9 safety hazards and precautions - adequately described;

5.3.10 in a chromatographic, spectroscopic, titrimetric or similar method, the minimum acceptable conditions, (e.g. plate counts, resolution, sensitivity), and a typical chromatogram, spectrum or titration curve etc. should be included with the method of analysis;

5.3.11 critical parameters of the method must be adhered to. If other parameters can be varied, this may be noted, (Note 4);

5.3.12 preprinted report sheets, which should be so designed that when completed they give enough data to Anable a complete check on the calculations to be carried out.

6 DISPATCH OF SAMPLES

When everything for the study has been prepared, the samples and documentation should be sent out in separate packages. It is essential that all international regulations for the tra'nsportation and labelling of pesticides be complied with.

7 STATISTICAL TREATMENT OF RESULTS

Statistical treatment of results is based on the procedure described in ISO 5725 - 1986 (E) [21, although some modifications have been made in line with the recommendations of the IUPAC Workshop.

When all the results have been returned to the panel or study coordinator, they should be tabulated and statistically assessed.

7.1 Preliminarv calculations

Before starting the calculation of repeatability and reproducibility, the data should be examined for the following points:

7.1.1 Redundant Data

If a laboratory carries out more than the required number of replicate analyses, all the results should be reported. An explanation from the originating laboratory of why this was done, and which results should be considered the most accurate, should be provided. If they are all valid, this may be taken into account by a statistical procedure, or the required number of results should be selected using a strictly random method.

7.1.2 Missing Data

Test results may be missing, due to loss of sample or a slip in the experiment. Completely empty cells can be ignored. when rejection of partly empty cells reduces the number of participating laboratories below the minimum number acceptable for the study, insert the average cell spread where no cell exists. In this context a cell is defined as the result obtained from measurement 1 of the sample and the result obtained from measurement 2 of the sample.

Note 4 Such ranges should have been investigated as far as possible by the laboratory which is responsible for initiating the study.

7.1.3 Outliers

Outliers are entries among the original test results, or in tables derived from them, that deviate so much from other entries that they are considered irreconcilable with them.

A graphical presentation of the results (cf Appendix 1) is often useful in recognizing such deviating data. A cursory examination will show whether highly significant outliers are likely to be present. In case of doubt the tests according to Cochran [2, paragraphs 11.2.3 and 12] and/or Grubbs [5], (cf Appendix 2) should be applied before any further calculation.

Cochran's maximum variance test and Grubbs (and/or double Grubbs outlier tests (Note 5) are used in combination with the procedure outlined below:

$p > 5\%$, i.e. Cochran's and/or Grubbs test statistic is less than its 5% critical value; the item is accepted.

$5\% > p > 1\%$, i.e. the test statistic lies between 5% and l% critical value, it is a straggler and should be marked with an asterisk; the test is statisticaly significant.

$p < 1\%$, i.e. the test statistic is greater than its l% critical value; this is a statistical outlier, and the test is highly significant.

p is the probability of the observed value of the test statistic. The 5% and l% critical values of Grubbs and Cochran's test are available in tabulated form (cf Appendix 2).

In the first instance the precision data, mean (x), repeatability (r) and reproducibility (R), are calculated with no outlying results removed, but using only valid data. Then those laboratories or data are removed when they turn out to be stragglers or outliers according to the above mentioned tests. Removal of outlier should be stopped if more than 22% i.e. 2 of 9 laboratories are removed (Note 5).

The prime decision as to whether unexplained outliers or stragglers may be retained as correct items lies with the study coordinator.

Determine whether the stragglers or outliers can be explained by some technical or computational error etc. When there is a reasonable explanation, the item is considered a real outlier, and may be corrected or discarded in keeping with the explanation obtained. When several unexplained stragglers or outliers occur at different levels within the same laboratory, that laboratory may be considered an outlier. It may be reasonable here to discard some or all of its data. The prime decision for this rests with the study coordinator, though advice should be sought from a statistician, who should be asked to draw attention to such features as a bimodal distribution within the results.

After the outlying data has been eliminated, the precision parameters are then recalculated.

7.2 Computation of r and R

The computation of r and R is carried out according to ISO 5725 - 1986, paragraph 14 [2].

Note 5 The Grubbs tests should be applied only to laboratory means, not to individual values of replicated design, since these values are not independent.

In making statistical calculations from the reported data, the full power of the calculator or computer should be used with no rounding or treating until the final reported mean and standard deviations are achieved. If the calculation of standard deviations must be conducted in steps, with the transfer of intermediate results, the number'of significant figures to be retained in squares calculations should be at least 1 plus 2 times the number of figures of the data.

Repeatability (r) - the value below which the absolute difference between two single test results obtained with the same method on identical test material under the same conditions (same operator, same apparatus same laboratory, short interval of time), may, with a specific probability (95% unless otherwise stated), be expected to lie.

Reproducibility (R) - the value below which the absolute difference between two single test results obtained with the same method on identical test m'aterial under different conditions (different operators, different apparatus, different laboratories and/or different time), may, with a specific probability (95% unless otherwise stated), be expected to lie.

A single test result is the value obtained by applying the test method fully once to a single specimen, and may be the mean of two or more observations.

The mean ($\bar{x}$), together with r and R should be calculated for each sample.

7.3 Functional Relationship between r, (or R) and $\bar{x}$

ISO 5725 describes a number of procedures for determining any functional relationship between r, R and $\bar{x}$ However, in the case of CIPAC methods, the Materials under investigation are all of different composition and from different production processes, and therefore such a reiationship is in no way certain. The alternative to determining this relationship is to compute separate values of r and R for each material investigated. The later approach, in keeping with current practice, is recommended.

8 EVALUATION OF THE STATISTICAL RESULTS

8.1 Presentation of Precision Data

Having obtained a full report on the statistical treatment of the results, the panel or national PAC should take decisions concerning the following questions:

8.1.1 Are any discordant results due to any defect in the description of the method?
8.1.2 What actions should be taken with respect to rejecting outlying laboratories?
8.1.3 Do results justify the establishment of final values or r and R?
8.1.4 If so, what are these final values? What is the region in which the precision data apply?

ISO recommends publication in tabular form:

Range or level	r	R
From .....to.....	.....	.....
From .....to.....	.....	.....
From .....to.....	.....	.....

and recommends a footnote such as:

"The precison data were determined from an experinent conducted in (year) involving (p) laboratöries and (q) levels".

8.2 Utilization of Precision Data

Initially r and R are intended to be used as criteria to determine whether a difference between two single test results can be ascribed to random fluctuations. A difference larger than r or R is suspect, and may justify the conclusion that there exists a systematic difference betwean the two test results, or justify additional investigation.

Thus r and R may be called "critical differences", to be applied to a pair of test results respectively obtaine under repeatability and reproducibility conditions.

It is sometimes necessary to compare the averages of two or more tests, or to compare the averages of a series with a specific value, and in such cases the "critical differences" can be derived from r and R as explained in ISO 5725 - 1986, Paragraph 19.2.1 to 19.2.4 [21.

The parameters r and R are calculated with a probability level of 95%. If these parameters need to be determined at other probability levels the method quoted in ISO 5725-1986 Paragraph 19.1.1 [2] may be used. In such cases the probability level should be denoted by a subscript, e.g. r_{99} or R_{90}.

The "critical differences" r and R may be used in a variety of ways, e.g.

-to compare test results from a batch of product with a product specification
-to compare test results obtained by a supplier and a customer on the same batch of product
-to design quality control procedures.

Specifically, if a single determination is carried out in a customer laboratory, then this value should not differ from' the manufacturerls value by more than 0.707 R once in twenty cases on average, [2, Paragraph 19.2.3].

If the laboratory performs the whole analysis in duplicate, then the difference obtained between this value and the manufacturerls value should not exceed

$$0.707 \quad \sqrt{R^2 - 0.5r^2}$$

once in twenty cases on average, [2, Paragraph 19.2.3]

R and r may also be used in estimating the minimum tolerances which can be placed on declared contents i FAO specifications.

9 ACCEPTABILITY OF r AND R

Whether or not it is felt by the panel or national PAC that final figures for reproducibility and repeatability are satisfactory depends on individual circumstances. If no other method is available, and the panel or national PAC feels that there is little chance of significantly reducing r or R by a further collaborative trial, then the results should stand as a true statement of the .methodls capabilities. These should be incorporated into any derived documents, e.g. reports or specifications.

However an alternative approach to determine whether the tested method is acceptable or not, is to compare the reproducibility relative standard deviation of the study RSD_R(Exp.) with the empirical linear relationship between the"logarithm to the base 10 of the reproducibility relative standard deviation RSD_R(Calc.) and the logarithm of the concentration of the analyte, expressed as a decimal fraction [6][7].

In its exponential form this relationship is given as the function of the so-called Horwitz curve:

$$RSD_R(\text{Calc.})\% = 2^{(1 - 0.5\log c)}$$

where c is the concentration of the analyte as a decimal fraction (e.g. for loo% concentration c = 1).

Thus if RSD_R(Exp.) as determined from the collaborative study is not larger than RSD_R.(Calc.) calculated for the relevant concentration, the gethod should be acceptable. The reproducibility relative standard deviation RSD_R (Exp.) may be calculated from R using the equation:

$$RSD_R(\text{Exp.}) = \frac{R * 100}{2.8 * \bullet}$$

provided that R itself has been calculated according to ISO 5725 - 1986 [2], • is the mean of the concentration of the analyte in %

10 FINAL REPORT

The analytical method and any associated documents should be written up in their final form, and a final report should be compiled giving the results of the study and the recommendations of the panel or national PAC. If these are approved than they should be submitted to CIPAC for final assessment.

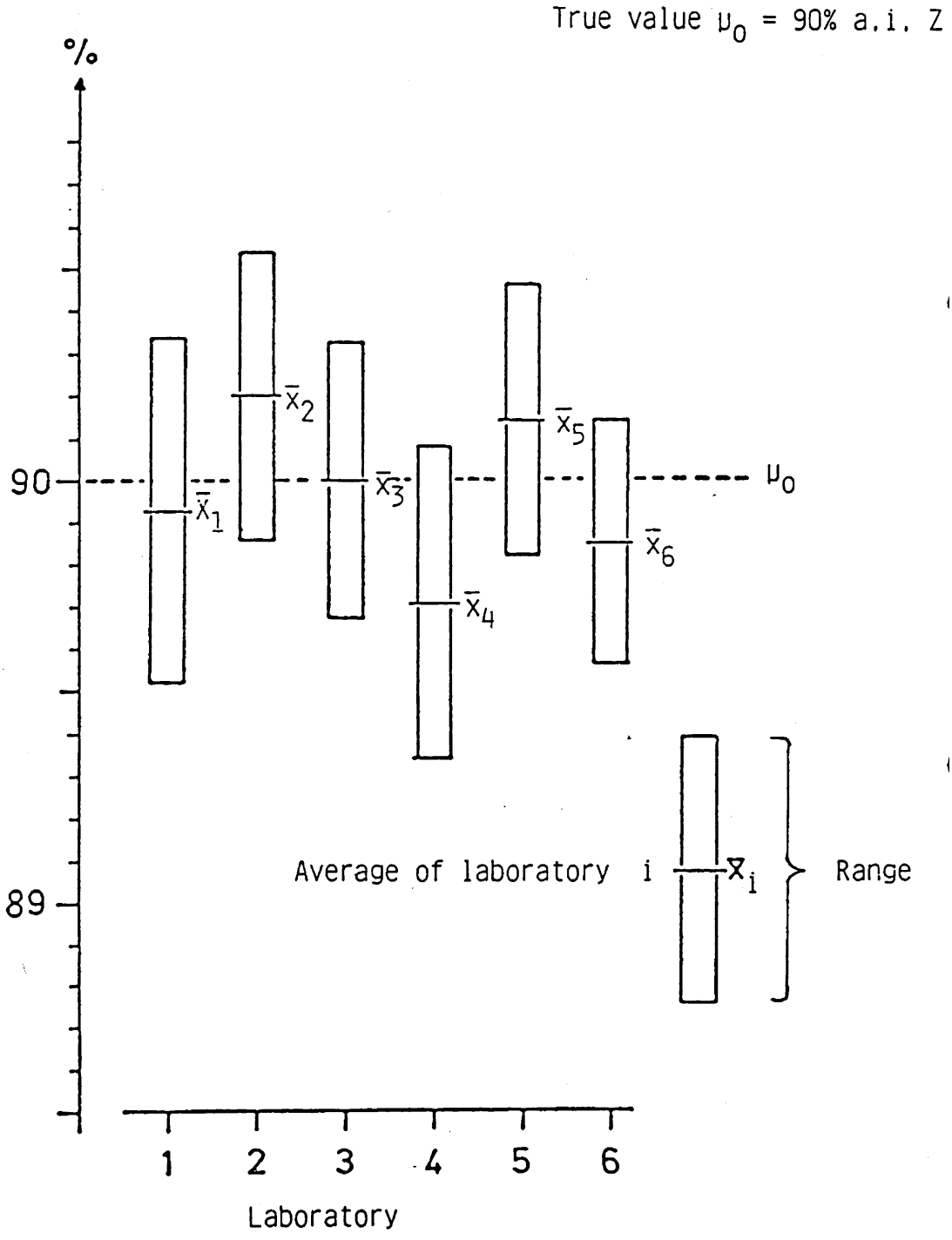
Appendices 1. Graphical presentation of results
2. The Outlier Tests according to Grubbs [5]

11 REFERENCES

[l] IUPAC Recommendations on the Harmonization of Collaborative Analytical Studies, Geneva, Switzerland, (1987).
[2] International Standard ISO 5725 - 1986 (E). Precision of test methods - Determination of repeatability and reproducibility by inter-laboratory tests. Second edition.
[3] JAOAC. 71, 161, (1988), 4th Final Draft.
[4] ISO - Guide 18 - 1978 (E) Layout for a standard meth'od of chemical analysis.
[5] F.E.Grubbs and G.Beck, Technometrics, 14, (1972), 847.
[6] W. Horwitz, Anal. Chem., 54, (1982), 67A.
[7] K. W. Boyer, W. Horwitz, R. Albert, Anal. Chem., 57, (1985), 454.

Appendix 1

GRAPHICAL PRESENTATION OF RESULTS



APPENDIX **2**

The Outlier Tests according to Grubbs [5]

1. Calculation of the Test Value r''_{lower} and r''_{upper}

Arrange the individual laboratory averages $\bullet_i$ in ascending order:

$\bullet_{(1)}$ = smallest laboratory value

$\bullet_{(n)}$ = largest laboratory value.

Calculate the total mean, $\bullet$ and standard variation, s.

Calculate the difference between:
the total mean and the smallest laboratory average value

$$= \bullet - \bullet_{(1)} \quad \text{and}$$

the total mean and the largest laboratory average value

$$= \bullet_{(n)} - \bullet$$

Compare the two differences, and with the largest difference calculate:

$$r''_{lower} = \frac{\bullet - \bullet_{(1)}}{s}$$

$$r''_{upper} = \frac{\bullet_{(n)} - \bullet}{s}$$

These test values r''_{lower} and r''_{upper} are compared with the corresponding tabulated critical values $r_{\bullet,n}$ (see Tables)

2. Evaluation

If

$r''_{lower} > r_{\bullet,n}$

or

$r''_{upper} > r_{\bullet,n}$

then the extreme value checked, $\bullet_{(1)}$ or $\bullet_{(n)}$ is an outlier at the probability value of $\bullet$. It is recommended that a probability value of $\bullet = 0.01$ should be taken for the testing of the total values.

Table

Critical values $r_{\alpha;n}$ for the Grubbs outlier test

Two-sided test

n \ α	0.10	0.05	0.02	0.01
3	1.153	1.155	1.155	1.555
4	1.463	1.481	1.492	1.496
5	1.672	1.715	1.749	1.764
6	1.822	1.887	1.944	1.973
7	1.938	2.020	2.097	2.139
8	2.032	2.12'6	2.221	2.274
9	2.110	2.215	2.323	2.387
10	2.176	2.290	2.410	2.482
11	2.234	2.355	2.485	2.564
12	2.285	2.412	2.550	2.636
13	2.331	2.462	2.607	2.699
14	2.371	2.507	2.659	2.755
15	2.409	2.549	2.705	2.806
16	2.443	2.585	2.747	2.852
17	2.475	2.620	2.785	2.894
18	2.504	2.651	2.821	2.932
19	2.532	2.681	2.854	2.968
20	2.557	2.709	2.884	3.001
21	2.580	2.733	2.912	3.031
22	2.603	2.758	2.939-	3.060
23	2.624	2.781	2.963	3.087
24	2.644	2.802	2.987	3.112
25	2.663	2.822	3.009	3.135
26	2.681	2.841	3.029	3.157
27	2.698	2.859	3.049	3.178
28	2.714	2.876	3.068	3.199
29	2.730	2.893	3.085	3.218
30	2.745	2.908	3.103	3.236

n \ α/2	0.05	0.025	0.01	0.005

One-sided test

Modified table according to F.E. Grubbs and G. Beck
Technometrics Vol. 14 (1972) pp. 847 et seq.

Table continued

Critical values $r_{\alpha;n}$ for the GRUBBS outlier test

Two-sided test

n \ α	0.10	0.05	0.02	0.01
31	2.759	2.924	3.119	3.253
32	2.773	2.938	3.135	3.270
33	2.786	2.952	3.150	3.286
34	2.799	2.965	3.164	3.301
35	2.811	2.979	3.178	3.316
36	2.823	2.991	3.191	3.330
37	2.835	3.003	3.204	3.343
38	2.846	3.014	3.216	3.356
39	2.857	3.025	3.228	3.369
40	2.866	3.036	3.240	3.381
42	2.887	3.057	3.261	3.404
44	2.905	3.075	3.282	3.425
46	2.923	3.094	3.302	3.445
48	2.940	3.111	3.319	3.464
50	2.956	3.128	3.336	3.483
52	2.971	3.143	3.353	3.500
54	2.986	3.158	3.368	3.516
56	3.000	3.172	3.383	3.531
58	3.013	3.186	3.397	3.546
60	3.025	3.199	3.411	3.560
65	3.055	3.230	3.442	3.592
70	3.082	3.257	3.471	3.622
75	3.107	3.282	3.396	3.648
80	3.130	3.305	3.521	3.673
85	3.151	3.327	3.543	3.695
90	3.171	3.347	3.563	3.716
95	3.189	3.365	3.582	3.736
100	3.207	3.383	3.600	3.754

n \ α/2	0.05	0.025	0.01	0.005

One-sided test

Modified table according to F.E. Grubbs and G. Beck
Technometrics Vol. 14 (1972) pp. 847 et seq.