codex alimentarius commission E

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

WORLD HEALTH ORGANIZATION

JOINT OFFICE: Viale delle Terme di Caracalla 00153 ROME Tel: 39 06 57051 www.codexalimentarius.net Email: codex@fao.org Facsimile: 39 06 5705 4593

Agenda Item 7(a) **CX/PR 09/41/5-Add.2**

April 2009

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON PESTICIDE RESIDUES

Forty-first Session

Beijing, China, 20-25 April 2009

PROPOSED DRAFT REVISION OF THE GUIDELINES ON THE ESTIMATION OF UNCERTAINTY OF RESULTS FOR THE DETERMINATION OF PESTICIDE RESIDUES (CAC/GL 59-2006)

Prepared by IAEA

This document is being circulated at Step 3 of the Procedure. Governments and interested international organizations are invited to prepare their comments and be ready to consider it at the forthcoming 41st Session of the Codex Committee on Pesticide Residues.

PART 1 – BACKGROUND INFORMATION AND JUSTIFICATION

INTRODUCTION

1. The determination of residues at trace levels (0.001 - 10 mg/kg) is subject to considerable analytical variability. To have a sound estimate of the quality of results the expression of measurement uncertainty (MU) is used. Due to the large scope of analytes and commodities in this field validation and consequently the estimation of MU can be extremely demanding. However, in order to minimise disputes resulting from questionable exceedance of regulatory limits, the estimation and reporting of MU is essential for demonstrating the boundaries and equivalence of analytical results generated in different places. When the compared results are obtained by analytical methods affected by significantly different bias this difference should be considered in the comparison of the performed evaluations (e.g., through an adequate estimation of the MU). More over, MU estimation is an essential prerequisite for accreditation and a must for laboratories.

2. Institutes have often limited financial, personnel and time resources which were necessary for thorough MU calculation as laid out in numerous guidance papers on MU. Therefore, and in the interest of rationalizing laboratory work it is impractical to calculate individual values for countless commodity/pesticide combinations[1], especially when using the rigorous bottom-up approach. Consequently, it was proposed in ALINORM 07/30/24, paras. 156-160 to develop a simplified guidance document for the estimation of MU, e.g., based on proficiency testing (PT) results, method validation and quality assurance data.

3. At the 39th CCPR Meeting a discussion paper was agreed, which would form the basis of a guidance document to be prepared and tabled at the 40th CCPR Meeting. The Committee decided to undertake new work in that regard. Accordingly a revised draft paper to support this attempt was prepared by IAEA taking account of contributions from responses to CX/PR 09/41/5 towards the revision of CAC/GL 54-2004 at the 41st CCPR Meeting.

[1] More than 1,000 pesticides are known worldwide; more than 220 pesticides have a Codex Reference Number

4. A revision of the MU guidance document using this discussion paper and the proposal for an extension of CAC/GL 54-2004 takes into consideration respective inputs by the EWG. The target output would be a straightforward guideline including a hands-on practical supplement based on empirical top-down concepts taking into account the results of proficiency testing (PT) schemes. MU evaluation based on the Horwitz approach is deemed less suitable.

5. The revision of the guidance document intends to support the practical adoption of the MU concept in food laboratories dealing with pesticide residue analysis. The key objectives are:

(a) adaptation of MU estimation by taking into consideration the complexity of pesticide residue analysis (i.e. several working steps eventually involving instrument calibration, chemical changes of the target analytes; limited laboratory resources; large number of combinations of commodities and pesticides);

(b) elaboration of a practically oriented and straightforward guidance (e.g. based on empirical top-down concepts);

(c) allowing for simplified MU estimation for the ease of compliance with ISO Standard 17025. Uncertainty related to sampling is not subject of this guideline.

BACKGROUND

6. The need to control analytical procedures – and consequently the necessity for quantitative expression of MU – is widely recognized. The technical part of ISO Standard 17025 requires the estimation of MU as an essential parameter which laboratories must have in place[2].

7. A number of guidance documents describe different approaches towards estimating MU. With regard to pesticide residue analysis, bottom-up calculations in particular are perceived as overly complicated and extremely laborious. This may in part be due to the fact that the MU concept had originally been developed for physical measurements where influencing factors and analytical parameters are limited and rather straightforward to define and to calculate. The concept is not easily transferable to complicated and multi-factorial chemical residue analysis procedures.

8. Pesticide residue methods involve several independent processes: (a) sample preparation, processing and storage, (b) extraction of analyte(s), (c) clean-up, (d) quantitation of analyte(s). Each sub-procedure can involve several steps including sample comminution, weighing, pipetting, calibration, and so on. Each procedural and/or working step will influence MU values, eventually different from analyte to analyte, from commodity to commodity, and mostly it is concentration dependent.

9. As a consequence, there is still incomplete understanding and limited adoption of the uncertainty concept. Commonly applicable procedures are still missing as adapted for the particular purpose of pesticide residue analysis in food, with its diversity of influencing factors. Therefore, specific guidance applicable to pesticide residue analysis of food would be useful towards simplification and wider acceptance of the uncertainty concept.

10. Taking into consideration relevant guidelines and publications, particularly in terms of top-down approaches on MU, this paper intends to outline specific pathways applicable to pesticide residues in food, including practical examples.

MU CONCEPTS FOR PESTICIDE RESIDUE ANALYSIS

11. Difficulties related to MU were discussed by CCMAS in 2007 (see ALINORM 07/30/23, paras 6-10). Although pesticide residue analysis in its complexity was not of particular concern, the matter is perceived similarly in CX/MAS 07/28/2-Add.2. The guidance document on MU summarizes the situation and draws together various developments in that area. In parts A to L the main relevant approaches as outlined in different publications are summarized and discussed. However, there is no specific guidance for particular analytical procedures as to which approach would be applicable for which purpose.

[2] See ISO/IEC Standard 17025, Para. 5.10.3.1: In addition … test reports shall … include the following: "c) where applicable, a statement on the **estimated uncertainty** of measurement; information on uncertainty is needed in test reports when it is relevant to the validity of application of the test results, when a client's instruction so requires, or when the uncertainty affects compliance to a specification limit."

12. ISO/TS 21748:2004[3] provides additional mathematical concepts especially for estimating zones of acceptance and rejection around analytical values; straightforward top-down approaches are also discussed. One important statement relevant in this context is that the reproducibility standard deviation obtained from collaborative studies is considered as a valid basis for MU evaluation. If accuracy (or trueness) data can be utilized, e.g., with respect to an established reference value based on (certified) reference material, then uncertainty associated with the estimated bias should be included in the MU budget. Evaluating uncertainty according to ISO/TS 21748 comprises the following elements:

(a) repeatability, reproducibility and bias estimates from collaborative study;

(b) laboratory bias and precision within that expected on the basis of collaborative studies;

(c) laboratory bias and precision under control and effects appropriately combined to form a combined uncertainty estimate.

Most of the interlaboratory trials performed in this field cannot allow the estimation of the uncertainty through this approach since participants use different analytical methods and the reference value is the consensus mean.

13. In guideline EA-4/16[4] it is recognized that "laboratories cannot in general be expected to initiate scientific research to assess the uncertainties associated with their measurements and tests". The guideline, among others, describes the use of validation and method performance data for uncertainty evaluation. Data accumulated during validation and verification of test methods, interlaboratory studies according to ISO 5725, accumulated quality control data, and proficiency testing schemes typically characterize test method performance.

14. SANCO ACQ Guidelines[5] support this line of action towards evaluating MU associated with proficiency test results. Eurolab Technical Report[6] and NORDTEST Report[7] TR 537 outline in greater detail, among others, the use of method validation and PT data for estimating MU.

ALTERNATIVE MU APPROACHES

15. A comprehensive and easily applicable MU concept is not provided by existing guidelines in terms of the practical application to pesticide residue analysis in foodstuffs. Calculating uncertainty budgets for thousands of relevant pesticide/crop combinations and dozens of analytical methods used in pesticide residue analysis is not practical in routine laboratory operation.

16. Empirical approaches proposed show alternatives also for pesticide residue analysis of foodstuffs. Practical and straightforward guidance for application in the determination of pesticide residues in foodstuffs could be made available through top-down MU concepts. Validation data, repeatability, reproducibility, outcomes of PT schemes can be utilized for simplified MU estimation applicable in food control laboratories.

[3] Technical Specification ISO/TS 21748:2004: Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation, First edition 2004-03-15

[4] EA-4/16 EA guidelines on the expression of uncertainty in quantitative testing, December 2003 rev00

[5] Document N° SANCO/2007/3131 - METHOD VALIDATION AND QUALITY CONTROL FOR PESTICIDE RESIDUE ANALYSIS IN FOOD AND FEED (www.crl-pesticides.eu)

[6] Eurolab Technical Report No. 1/2007, March 2007, Measurement uncertainty revisited: Alternative approaches to uncertainty evaluation (www.eurolab.org)

[7] NORDTEST Report TR 537, HANDBOOK FOR CALCULATION OF MEASUREMENT UNCERTAINTY IN ENVIRONMENTAL LABORATORIES, EDITION 2

17. Based on a series of PT schemes, the ACQ Guidelines of the EC indicate that actual and target values according to different performance and quality criteria were well within the same order of magnitude. For instance, values derived from Fitness-for-Purpose (FFP), the Horwitz equation (see annex) and standard deviation calculated from EC PT schemes, after rejection of outliers (Qn), expressed in (%), were very similar. Accordingly, the evaluation of the recent EC PT schemes demonstrates that a FFP variability of 25% can be accepted as a sound representation of performance under these circumstances. As a consequence, accepting 25% variability as a standard deviation would lead to a generalized assumption of ±50% MU. The laboratories must prove that this uncertainty value is adequate to describe the performance of their analytical method, for instance through satisfactory participations in proficiency tests (see para 21.). The extrapolation of good performance for the analysis of one analyte/matrix combination to a broader scope of analysis should be based on the knowledge of the equivalence of the performance of the method within the considered scope of analysis. When the proficiency tests involve the reporting of average values of replicated measurements collected within the participant laboratory, the relative expanded uncertainty of 50 % (2 x 25%; 25 % is the reference RSD for some PT) underestimates the uncertainty associated with single measurements.

18. Accepting such a generalized approximation for pesticide multi-residue analysis methods, a generalized top-down approach might result in larger MU values than such derived for each individual pesticide/commodity combination by systematic bottom-up calculations. However, the application of generic MU is considerably more practical and easier to obtain. Generalized values, like ±50% MU, mostly would expand safety margins around MRLs. It could make a difference especially when getting close to MRLs/ trigger values. On the other hand, for laboratories it would mean a considerable rationalization in terms of time, resources and workload that otherwise had to be devoted to systematic bottom-up MU evaluation.

MU ESTIMATION BASED ON HORWITZ FORMULAS

19. Similarly to the PT based approach MU may be estimated using empirical Horwitz formula. These generalized expressions are based on countless empirical interlaboratory comparison data. This approach takes into account that expected MU values are dependent on the residue level, i.e., the higher the residue concentration, the lower the anticipated relative MU. The Horwitz approach is expressed by the following equation (equivalent approximations exist):

$$RSD_R = 2^{1-0.5\log c} = 2 * c^{-0.1505}$$

with:

RSD_R = expected relative interlaboratory standard deviation (%)

c = concentration of the analyte (expressed as kg/kg, i.e., 0.01 mg/kg = 0.00000001 kg/kg)

Accordingly putting real figures into the above formulas concentration dependent RSD_R values are obtained, i.e.:

0.01 mg/kg ⇒ 32.0 %
0.1 mg/kg ⇒ 22.6 %
1 mg/kg ⇒ 16.0 %

20. RSDR values depending on the respective concentration levels can be transformed into MU by multiplying with an appropriate coverage factor, normally k = 2. Advantages of this concept[8] include the incorporation of laboratory bias because laboratory variability is also randomized. Deviations generated by different laboratories have been included. The Horwitz equation was found to be widely applicable to all concentration, methods and analytes.

21. Drawbacks associated with the approach are that appropriate and sufficient data are needed as the basis for the estimation of a valid relation between concentration and uncertainty since the Horwitz data came from a highly diverse range of collaborative trials with concentrations ranging from 0.05 µg/kg to 60%, involving a large number of other compounds than pesticides. Prescribed methods were used, and PT data were not included. The resulting estimates of uncertainty accordingly are based on the distribution of between-laboratory standard deviations.

[8] L. Alder et al.: Estimation of Measurement Uncertainty in Pesticide Residue Analysis. JAOAC International. Vol. 84, No 5, 2001, 1569-1577.

MU VALUES IN PESTICIDE RESIDUE ANALYSIS BASED ON EMPIRICAL DATA

22. Data derived from systematic method validation for verifying recovery values and associated standard deviations characterizing the use of analytical methods can be utilized. A step by step practical guidance should incorporate representative examples of commonly used analytical methods.

23. In practical terms a guidance document would incorporate empirical data and outcomes of PT schemes. In particular the following information and data could be utilized:

(a) Concentration dependent RSDs according to Horwitz for estimating MU, e.g., for fatty matrices, whereas approximated RSDs of 25% would be applicable for non-fatty matrices over the range of relevant trace level according to existing data derived from PT schemes conducted in the European Union. However, laboratory have to prove their proficiency for reporting measurement with this uncertainty value.

(b) Method validation data, including recovery, repeatability and intermediate precision.

(c) Quality assurance data such as control charts derived from the routine application of methods.

(d) Results from the participation in PT schemes.

Note: There is a relation between the following precision values:

s_{ip} – intermediate precision standard deviation;
s_{bl} – between laboratory precision standard deviation;
s_R – reproducibility standard deviation.

$$S_R = \sqrt{S_{ip}^2 + S_{bl}^2}$$

When each laboratory report the mean of n replicated measurements obtained within the laboratory in intermediate precision conditions (e.g. different days), the reproducibility of the method is reduced from s_R to $s_{R;n}$:

$$S_{R;n} = \sqrt{\frac{S_{ip}^2}{n} + S_{bl}^2}$$

Application example:

Considering the reproducibility, $S_{R;2}$, of an analytical method estimated in a interlaboratory trial where each participant reports the mean of two replicates obtained within a laboratory in intermediate precision conditions, equal to 0.2325 mg/kg (mean concentration of 0.93 mg/kg and RSD of 25 %):

$$S_{R;2} = 0.2325 = \sqrt{\frac{S_{ip}^2}{2} + S_{bl}^2}$$

If the relative intermediate precision standard deviation is 17 % (i.e. 0.1581 mg/kg), then:

$$S_{R;2} = 0.2325 = \sqrt{\frac{0.1581^2}{2} + S_{bl}^2}$$

Therefore:

$$S_{bl} = \sqrt{0.2325^2 - \frac{0.1581^2}{2}} = 0.2038\ mg/kg$$

Therefore, the reproducibility, s_R, associated with single measurements is:

$$S_R = \sqrt{S_{ip}^2 + S_{bl}^2} = \sqrt{0.17^2 + 0.2038^2} = 0.2654\ mg/kg$$

Therefore, it should be reported an expanded relative uncertainty associated with single measurements of 57 %.

24. Implementing a PT-based simplified ±50% MU approach should only be used by individual laboratories if the following analytical performance and quality criteria can be demonstrated:

(a) Within-laboratory SD smaller than the between-laboratories SD.

(b) Successful participation in PT schemes (z-score ≤|2| for 95%, z-score ≤|3| for not more than 5% of the values).

(c) Small method and/or laboratory bias for recovery tests.

(d) Verification of analytical performance by regularly analysing suitable reference material, if available.

RECOMMENDATION / PROPOSAL

25. As is an emerging practice in the EC and elsewhere already, empirical top-down estimation of ±50% MU could complement a mathematically stringent bottom-up calculation model if the respective empirical quality criteria are met. Alternatively the Horwitz formula approach of estimating concentration-dependent MU based on the evaluation of results of interlaboratory collaborative tests could be applied as well. However, the laboratory must prove the applicability of this uncertainty value to their measurements.

26. It is proposed to further develop a specific guidance for the application of empirical MU concepts applicable particularly in the field of pesticide residue analysis of foodstuffs.

Note: Laboratories uncomfortable with these empirical approaches or where such is not deemed applicable may wish to apply step-by-step bottom-up calculation to specifically generate distinct individual uncertainty estimates as given elsewhere[9] including guidance on the treatment of concentration levels eventually conflicting with trigger values[10].

[9] primarily: EURACHEM/CITAC Guide CG 4, Quantifying Uncertainty in Analytical Measurement, Second Edition, QUAM 2000.1

[10] EURACHEM/CITAC Guide, Use of uncertainty information in compliance assessment, First Edition 2007

PART 2 – PROPOSED EXTENSION OF CAC-GL 59-2006:

PRACTICAL AND SIMPLIFIED MU ESTIMATION BASED ON TOP-DOWN APPROACHES

Underlying principles, formulas and statistics for PT based estimation of MU

Within-laboratory reproducibility standard deviation is combined with estimates of the method and laboratory bias using PT data:

$$U' = k * u' = \sqrt{u'(R_W)^2 + u'(bias)^2}$$

where:

$$u'(bias) = \sqrt{RMS'_{bias}{}^2 + u'(C_{ref})^2}$$

and:

$$RMS'_{bias} = \sqrt{\frac{\sum (bias'_i)^2}{m}}$$

and:

$$u'(C_{ref}) = \frac{\sum_i \frac{S'_{Ri}}{\sqrt{n_i}}}{m}$$

with:

U'	=	expanded relative uncertainty
k	=	coverage factor
u'	=	combined relative standard uncertainty
$u'(R_W)$	=	intermediate precision relative standard uncertainty
u'(bias)	=	relative standard uncertainty component from method and laboratory bias, based on PT data
RMS'_{bias}	=	root mean square of relative bias values
$bias'_i$	=	relative bias of PT i [obtained $result_i$ – assigned $value_i$)/assigned $value_i$]
$u'(C_{ref})$	=	average relative standard uncertainty of assigned values
S'_{Ri}	=	interlaboratory relative standard deviation of PT i
n_i	=	number of participants in PT i
m		total number of PT schemes

Practical application

(1) Prerequisites for using an expanded relative uncertainty of 50 %:

- The laboratory has demonstrated its technical capability to generate reliable results at the required level of quality, i.e. by:
  - sound validation data for the respective analytical method;
  - acceptable quality control data, e.g., control charts for respective methods and compounds;
  - successful participation in PT schemes which fulfil PT quality criteria according to the Harmonized Protocol[11], ISO Guide 43-1 etc.;
  - the laboratory has been rated as well-performing (e.g. Category A, Sufficient Scope at 90%, e.g., according to PT evaluation applicable within the EC);
  - Evidence of the equivalence of the performance of the analytical method for each analyte/commodity combination from the scope of analysis.

(2) Uncertainty evaluation using laboratory evaluation data:

- identification of the main sources of uncertainty (weighing, calibration, purity, temperature, volumetric glassware, …);

[11] M Thompson, S L R Ellison, R Wood; The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report); Pure Appl. Chem. 78(1) 145-196 (2006)

- evaluation of the order of magnitude of the uncertainty of basic laboratory operations in relation to the overall uncertainty of the procedure;
- expected result:
  - uncertainty of basic laboratory operations almost negligible;
  - random run-to-run variability as the principal source of MU.
- estimation of overall bias and recoveries from in-house validation experiments (fortification, spiking, reference materials, …):
  - the mean of the resulting relative standard deviation taken as relative uncertainty is associated with random variation;
  - analyte mean recovery within 70-120 %..

(3) Comparison with PT results:

- series of PT rounds with slightly varying concentrations and matrices;
- the relative standard deviation of valid data is comparable to the expected relative standard deviation (comparing PT results with real laboratory data).

(4) Verification of uncertainty estimates:

- checks using observed within-laboratory precision;
- checks using certified reference materials or suitable test materials;
- checks using reference methods;
- checks based on the results of PT (including external QA data or measurement audits);
- checks based on comparison of results with other laboratories,
- comparison with other uncertainty estimates based on different approaches or different data (some approaches are expected to produce significantly different MU estimations).

(5) Conclusion:

- PT data can provide strong support for the laboratory estimate of MU based on validation data;
- PT data can form the basis for estimating MU, using the dispersion of relative differences.

Evaluation of uncertainty estimates against PT results

Checking the quality of uncertainty estimates may apply the zeta (ζ) score formula laid out in the Eurolab Report[6]:

$$\zeta = \frac{x - x_a}{\sqrt{u(x)^2 + u(x_a)^2}}$$

with:

x	=	laboratory result
x_a	=	assigned value
u(x)	=	standard uncertainty of laboratory results
$u(x_a)$	=	standard uncertainty of assigned values

Uncertainties are considered correct if $|\zeta|$ is in the range 0 to 2; underestimated if $|\zeta|$ is frequently over 2. Equivalent to the zeta score, the E_n number can be calculated by replacing the expanded uncertainties U(x) and $U(x_a)$ by u(x) and $u(x_a)$ in the above formula if both coverage factors are 2.

Limitations

In general, proficiency tests are not carried out frequently enough to provide good estimates of the performance of an individual laboratory's implementation of a test method. However, in the special case where:

- the types of test items used in the scheme are appropriate to the types tested routinely,
- the assigned values in each round are traceable to appropriate reference values, and,

- the uncertainty associated with the assigned value is small compared with the observed spread of results,

the dispersion of the differences between the reported values and the assigned values obtained in repeated rounds provides a basis for an evaluation of the uncertainty (see Eurolab and NORDTEST references).

A PT-based top-down approach is therefore applicable where PT data support this. Referring to EC-PT schemes this approach could be different for various matrices and pesticide/ matrix combinations.

Certain matrix/pesticide combinations would need separate MU evaluation following the guidelines and approaches given elsewhere.

Summary

With the assumptions and prerequisites outlined for conducting and evaluating PT schemes and classification of laboratory performance, based on top-down approaches, an estimate MU of ±50% as a generalized value would provide an acceptable and practical approximation of pesticide residue analysis of foodstuffs to daily laboratory reality.

ANNEX: PRACTICAL APPLICATION OF THE BOTTOM-UP APPROACH ON MU ESTIMATION BY UTILIZING PT AND INTERNAL VALIDATION AND QUALITY ASSURANCE DATA

Requirements

Method Validation and Quality Control Procedures for Pesticide Residues Analysis in Food and Feed. Par. 64: Acceptability of analytical performance for routine analysis.

- Individual recovery result should normally be in the range of the mean recovery ±2 x %RSD.
- Addition of a spiked sample to each batch of analysis.
- Results may be used for quality control charts.

Long-term quality control

Suitable samples for long-term quality control are:

- Certified reference materials.
- Remaining materials from proficiency tests.
- Other materials with suitable (and stable) concentrations of pesticides.
- Spiked samples.
- Matrices and analytes have to be stable.

Example

In a rolling programme, each technician analyses twice a year one QC sample. These sample are treated as routine samples

Evaluation of PT results by Sum of Weighted z-Scores

The sum of weighted z-scores (SWZ) is used in the PT evaluation as an overall laboratory assessment. SWZ may also be used to evaluate all PT results of a laboratory over a long period of time.

Each analyte is evaluated as follows:

- Acceptable z-scores: The sum of all z-scores between 0 and 2, multiplied by 1.
- Questionable z-scores: The sum of all z-scores greater than 2 but less than or equal to 3, multiplied by 3.
- Unacceptable z-scores: The sum of all their z-scores greater than three, multiplied by 5.

The total sum of these three groups is then divided by the number of results (n) for each analyte.

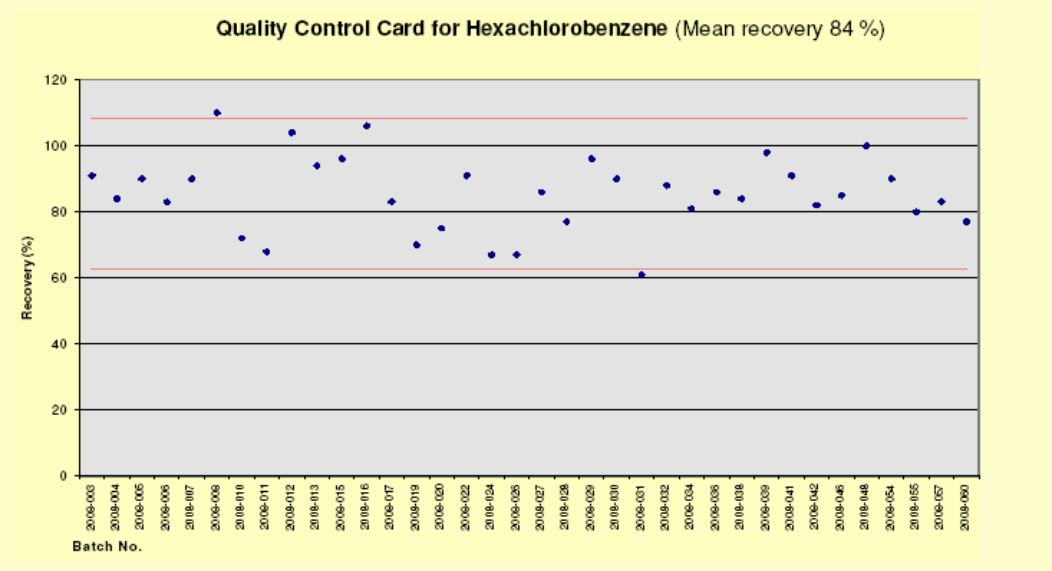

Figure: Example of a control chart of mean recovery values