European Medicines Agency
Veterinary Medicines and Inspections

London, 14 December 2009
Doc. Ref. EMEA/CVMP/VICH/463202/2009-CONSULTATION

VICH Topic GL49

at step 4

GUIDELINES FOR THE VALIDATION OF ANALYTICAL METHODS USED IN RESIDUE DEPLETION STUDIES

DRAFT AGREED BY VICH STEERING COMMITTEE	November 2009
ADOPTION BY CVMP FOR RELEASE FOR CONSULTATION	10 December 2009
END OF CONSULTATION (DEADLINE FOR COMMENTS)	20 May 2010

Comments should be provided using this template to vet-guidelines@emea.europa.eu
Or by fax +44 20 7418 8447

7 Westferry Circus, Canary Wharf, London, E14 4HB, UK
Tel. (44-20) 74 18 84 00 Fax (44-20) 74 18 86 60
E-mail: mail@emea.europa.eu http://www.emea.europa.eu

VICH International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products

VICH GL 49 (MRK) – METHOD USED IN RESIDUE DEPLETION STUDIES

November 2009

For consultation at Step 4 - Draft 1

GUIDELINES FOR THE VALIDATION OF ANALYTICAL METHODS USED IN RESIDUE DEPLETION STUDIES

Recommended for Consultation

at Step 4 of the VICH Process

on 6 November 2009

by the VICH Steering Committee

This Guideline has been developed by the appropriate VICH Expert Working Group and is subject to consultation by the parties, in accordance with the VICH Process. At Step 7 of the Process the final draft will be recommended for adoption to the regulatory bodies of the European Union, Japan and USA.

Secretariat : C/O IFAH, rue Defacqz, 1 - B - 1000 Bruxelles (Belgium) - Tel. +32-2-543.75.72, Fax +32-2-543.75.85
e-mail : sec@vichsec.org - Website : http://www.vichsec.org

Table of Contents

1. INTRODUCTION

1.1. Objective

This guidance document is intended to provide a general description of the criteria that has been found to be acceptable to the European Union (EU), Japan, United States of America (USA), Australia, New Zealand and Canada for the validation of analytical methods used in veterinary drug residue studies.

1.2. Background

During the veterinary drug development process, residue depletion studies are conducted to determine the concentration of the residue or residues present in the edible products (tissues, milk, eggs or honey) of animals treated with veterinary drugs. This information is used in regulatory submissions around the world. Submission of regulatory methods (post approval control methods) and the validation requirements of the regulatory methods are usually well defined by various regulatory agencies worldwide and may even be defined by law. Consequently, it has been difficult to harmonize the procedures used for validation of these methods. However, the residue studies are generally conducted before the regulatory methods have been completed. Often times the in-house validated residue methods provide the framework for the methods submitted for regulatory monitoring. Harmonization of the validation requirements for methodology used during residue studies and submitted to the regulatory agencies in support of the maximum residue limits (MRLs) and withdrawal periods should be achievable. It is the intent of this document to describe a validation procedure that is acceptable to the regulatory bodies of the EU, Japan, USA, Australia, New Zealand and Canada for use in the residue depletion studies. This validated method may continue on to become the "regulatory method" but that phase of the process will not be addressed in any detail in these guidelines.

A variety of validation guidelines exist for analytical methodology and many of the aspects of those validation procedures are incorporated in this document (VICH GL1 (Validation Definition), October 1998 and VICH GL2 (Validation Methodology), October 1998). However, there are aspects of residue validation procedures that are addressed in this guidance document that are not addressed in previous documents. The guidelines provided here are intended to specifically address the validation of veterinary drug residue methods.

1.3. Scope

These guidelines are only intended to apply to analytical procedures that have been developed for the evaluation of residue assays. These are not intended to define the criteria needed for validation of regulatory monitoring assay procedures.

The format of this document is to provide performance characteristics of the residue assays that if followed will be acceptable to the regulatory agencies of the EU, Japan, USA, Australia, New Zealand and Canada. The intent is that methods validated according to this guideline will provide residue data that will be acceptable to the regulatory agencies in determining appropriate withdrawal periods.

2. PERFORMANCE CHARACTERISTICS

There are specific performance characteristics of a method validation. Those performance characteristics are defined as follows:

Linearity

Accuracy

Precision

Limit of Detection

Limit of Quantitation

Selectivity

Stability in Matrix

Process Sample Stability

Robustness

Each of the characteristics will be described below as they apply to the validation of methods intended for use in veterinary drug residue studies.

2.1. Linearity

A calibration curve should be generated in which the linear relationship is evaluated across the range of the expected matrix (tissue, milk, egg or honey) concentrations. Calibration standard curves can be generated in three formats depending upon the methodology: standards in solvent/buffer, standards fortified into control matrix extract and standards fortified into control matrix and processed through the extraction procedure. Linearity should be described by a linear regression plot of known concentration vs. response using a minimum of 5 different concentrations. The linear relationship in general is best described by unweighted linear regression, but may be fit to a weighted linear regression with weighting factors of 1/concentration (1/X) or 1/concentration^2 ($1/X^2$), if justified. Acceptability of the weighting factors should be determined by evaluation of the residuals across three runs (Are they randomly distributed?). Evaluation of the residuals should be carried out across at least three separate runs.

The recommended acceptance criterion for a standard curve is dependent upon the format of the standard curve. Calibration standard curves generated by fortification of control matrix and processed through the procedure are subject to the same acceptance criteria as the samples (see Section 0 Precision). Calibration standard curves generated by standards in solvent/buffer or

by fortification of control matrix extract would require more stringent acceptance criteria (Repeatability ≤ 15% at all concentrations except at or below LOQ where it can be ≤ 20%).

Some assays (e.g. microbiological assays) may require log transformations to achieve linearity where other assays (e.g., ELISA, RIA) may require a more complicated mathematical function to establish the relationship between concentration and response. Again, acceptability of the function selected should be verified by evaluation of the residuals generated when that function is used.

2.2. Accuracy

Accuracy refers to the closeness of agreement between the true value of the analyte concentration and the mean result that is obtained by applying the experimental procedure. Accuracy is closely related to systematic error (analytical method bias) and analyte recovery (measured as percent recovery). Recommended accuracy for residue methods will vary depending upon the concentration of the analyte. Recommended mean accuracies based on the concentration of the analyte as provided by the CODEX Guidelines[i] are listed below:

Analyte Concentration*	Acceptable Range
< 1 µg/kg	-50 % to +20 %
≥ 1 µg/kg < 10 µg/kg	-40 % to +20 %
≥ 10 µg/kg < 100 µg/kg	-30 % to +10 %
≥ 100 µg/kg	-20 % to +10 %

* µg/kg =ng/g = ppb

2.3. Precision

Precision of a method is the closeness of agreement between independent test results obtained from homogenous test material under stipulated conditions of use. Analytical variability between different laboratories is defined as reproducibility, and variability from repeated analyses within a laboratory is repeatability. Single-laboratory validation precision should include an intra-day (repeatability) and inter-day component.

It is considered adequate to determine the intra- and inter-day precision of the analytical method as part of the validation procedure. There is generally not a need to determine reproducibility (inter-laboratory precision) in order to conduct a residue depletion study, because the laboratory that is often developing the method is the same laboratory assaying the samples from the residue study. Instead of establishing reproducibility of the assay an inter-day precision can be determined. Inter-day precision can also be referred to as between-day precision whereas repeatability is defined as within-day (intra-day) precision. Intra- and inter-day precision should be determined by the evaluation of a minimum of three replicates at three

different concentrations representative of the intended validation range (which should include the LOQ) across three days of analysis.

For the purposes of the residue method validation, acceptable variability is dependent upon the concentration of the analyte. Recommended acceptable precision as provided by CODEX Guidelines[ii] are listed in the table below:

Analyte Concentration	Repeatability (intra-laboratory/inter-day precision), %CV
< 1 µg/kg	35 %
≥ 1 µg/kg < 10 µg/kg	30 %
≥ 10 µg/kg < 100 µg/kg	20 %
≥ 100 µg/kg	15 %

2.4. Limit of Detection

The limit of detection (LOD) is the smallest measured concentration of an analyte from which it is possible to deduce the presence of the analyte in the test sample with acceptable certainty. There are several scientifically valid ways to determine LOD and any of these may be used as long as a scientific justification is provided for their use. See Annex 1 and Annex 2 for examples of acceptable methods for determining LOD and Annex 3 for a suggested protocol for determining accuracy, precision, LOD, LOQ and selectivity in a single study.

2.5. Limit of Quantitation

The LOQ is the smallest measured content of an analyte above which the determination can be made with the specified degree of accuracy and precision. As with the LOD, there are several scientifically valid ways to determine LOQ and any of these may be used as long as scientific justification is provided. See Annex 1 and Annex 2 for examples of acceptable methods for determining LOQ and Annex 3 for a suggested protocol for determining accuracy, precision, LOD, LOQ and selectivity in a single study.

2.6. Selectivity

Selectivity is the ability of a method to distinguish between the analyte being measured and other substances which may be present in the sample being analyzed. For the methods used in residue studies, selectivity is primarily defined relative to endogenous substances in the samples being measured. Because the residue studies are well controlled, exogenously administered components (i.e., other veterinary drugs or vaccines) are either known or not allowed during the study. If it is the intent to submit the validated method as a regulatory method, it may be prudent for the investigator to test known products used in the animals being tested for possible interference.

A good measure of the selectivity of an assay is the determination of the response of control samples (see section 2.5 above). That response should be no more than 20% of the response at the LOQ. See Annex 3 for a suggested protocol for determining accuracy, precision, LOD, LOQ and selectivity in a single study.

2.7. Stability in Matrix

Samples (tissue, milk, eggs or honey) collected from residue studies are generally frozen and stored until assayed. It is necessary to determine how long these samples can be stored under the proposed storage conditions without excessive degradation prior to analysis. As part of the validation procedure or as a separate study, a stability study needs to be conducted to determine the appropriate storage conditions (e.g., 4°C, -20°C, or -70°C) and length of time the samples can be stored prior to analysis.

Samples should be fortified with known quantities of analyte and stored under the appropriate conditions. Samples will be periodically assayed at specified intervals (e.g. initially, 1 week, 1 month, 3 months). If the samples are frozen, freeze/thaw studies need to be conducted (3 freeze/thaw cycles – one cycle per day at a minimum). Alternatively, incurred samples can be used with initial assays conducted to determine the starting concentrations. The recommended protocol for assessing stability in matrix is the analysis of two different concentrations in triplicate near the high and low end of the validation range. Stability in matrix is considered acceptable if the mean concentration obtained at the specified stability time point agrees with the freshly fortified control sample assay results (initial assay results if incurred samples are used) within ± 15%.

2.8. Processed Sample Stability

Often, the samples are processed one day and assayed on a second day or because of an instrument failure are stored additional days, e.g. weekend. The stability of the analyte in the process sample extract may be examined as necessary to determine stability under processed sample storage conditions. Examples of storage conditions would be 4 to 24 hours at room temperature and 48 hours at 4°C. Other storage conditions may be investigated consistent with the method requirements. The recommended protocol for assessing processed sample stability is the analysis of two different concentrations in triplicate near the high and low end of the validation range. Processed sample stability is considered adequate if the mean concentration obtained at the specified stability time point agrees with the initial assay results or with freshly fortified and processed control sample assay results within the acceptance criteria (± 15%) of the assay.

2.9. Robustness

Evaluation of the robustness of regulatory methods is of major importance. Evaluation of robustness for residue methodology is less of a concern for residue methods as these are usually conducted within a single laboratory using the same instrument. However, robustness should still be evaluated particularly for areas of the method that could undergo changes or modifications over time. These might include reagent lots, incubation temperatures, extraction solvent composition and volume, extraction time and number of extractions, solid phase extraction (SPE) cartridge brand and lots, analytical column

brand and lots and HPLC elution solvent composition. During the development, validation or use of the assay, method sensitivity to any or all of these conditions may become apparent and variations in the ones most likely to affect the method performance should be evaluated.

3. GLOSSARY

Accuracy – The accuracy of an analytical procedure expresses the closeness of agreement between the true value of the analyte concentration and the mean result that is obtained by applying the analytical procedure. This is generally expressed as % recovery or % error.

Control sample – Tissue, milk, egg or honey from an animal that has not been treated with the veterinary drug under investigation.

Incurred sample – Tissue, milk, egg or honey from an animal treated with the veterinary drug under investigation that has a residue concentration of the analyte of interest.

Inter-Day Precision – Inter-day precision expresses within-laboratory across-day variations.

Intra-Day Precision – Intra-day precision expressed within-laboratory within-day variations.

Limit of Detection – The limit of detection of an individual analytical procedure is the lowest amount of analyte in a sample that can be detected with acceptable certainty but not quantitated as an exact value.

Limit of Quantitation – The limit of quantitation of an individual analytical procedure is the lowest amount of analyte in a sample that can be quantitatively determined with acceptable precision and accuracy.

Linearity – The linearity of an analytical procedure is its ability (within a given range) to obtain test results that are directly proportional to the concentration (amount) of analyte in the sample

Matrix – The matrix is basic edible animal products (tissue, egg, milk or honey) that contains or could contain the residue of interest.

Precision – The precision of an analytical procedure expresses the closeness of agreement between a series of measurements obtained from multiple sampling of the same homogenous sample under prescribed conditions. The precision of an analytical procedure is usually expressed as the variance, standard deviation or coefficient of variation of a series of measurements.

Processed Sample – A processed sample is a sample that has been extracted or otherwise processed to remove the analyte from much of the original sample matrix.

Repeatability – Repeatability expresses the precision under the same operating conditions over a short interval of time.

Reproducibility – Reproducibility expresses the precision between laboratories.

Robustness – The robustness of an analytical procedure is a measure of its capacity to remain unaffected by small variations in method parameters and provides an indication of its reliability during normal usage**.**

Selectivity – Selectivity is the ability to assess the analyte in the presence of components (endogenous materials, degradation products, other veterinary drugs) that may be expected to be present.

Veterinary drug residues – all pharmacologically active substances, whether principles, excipients or degradation products and their metabolites that remain in foodstuffs obtained from animals to which the veterinary drug product in question has been administered. In practice, a specific drug residue (principle, excipients, or metabolite) is referred to in the analytical procedure as the analyte of interest.

Annex 1

Examples of Methods for Determining LOD and LOQ

One commonly used approach is referred to as the IUPAC definition.[iii] In that procedure the LOD is estimated as mean of 20 control sample (from at least 6 separate sources) assay results plus 3 times the standard deviation of the mean. The LOQ then becomes the mean of the same results plus 6 or 10 times the standard deviation of the mean. Testing of the accuracy and precision at the estimated LOQ will provide the final evidence for determination of the LOQ. If the %CV for the repeatability measurement at that concentration is less than or equal to the accuracy and precision acceptance criteria (Section 2.2 and 2.3), then the estimated LOQ is acceptable.

Annex 2

Codex Alternative Methods for Determining LOD and LOQ

An alternative method for determining LOD and LOQ has been recommended by Codex Alimentarius[iii]. The method is said to overcome the problems associated with the IUPAC defined method (i.e. the high variability at the limit of measurement can never be overcome) in Annex 1. In this approach, the LOD is determined by a rounded value of the reproducibility relative standard deviation (RSD) when it goes out of control (i.e. where 3 X RSD = 100%; RSD = 33%, rounded to 50% because of the high variability). This method is then directly related to the analyte in matrix and not just the analyte.

The Limit of Quantitiation (LOQ) then corresponds to the LOD and becomes defined as where the RSD = 25%. This is consistent with where the upper limit of detection merges with the lower limit of quantitation. As in the IUPAC method defined in Annex 1, testing of the accuracy and precision at the estimated LOQ will provide the final evidence for determination of the LOQ. If the %CV for the repeatability measurement at that concentration is less than or equal to the accuracy and precision acceptance criteria (Section 2.2 and 2.3), then the estimated LOQ is acceptable.

Annex 3

Protocol for Residue Method Validation

Selectivity, LOD and LOQ are all interrelated and are affected by endogenous interferences that may be present in the matrix being assayed. LOD is often time difficult to determine particularly in LC/MS assays where control samples actually provide zero response at the retention time of the analyte. Without a response, it is impossible to calculate a standard deviation and therefore impossible to determine the LOD based on the mean plus 3 times the SD of the mean. Even if a mean plus 3 times the SD of the mean can be determined, it is often related to the instrument limit of detection rather than the method limit of detection. The following protocol is designed to determine specificity, LOD, LOQ, precision and accuracy in one study.

1. Collect drug free matrix from 6 separate sources (animals) and screen for any possible analyte contamination.
2. Fortify (spike) 1 each of a minimum of 3 samples (each source randomly selected such that each source is represented at least once at each concentration) of the 6 control samples at 0, at the estimated LOD (determined during assay development), at 3 times the estimated LOD (estimated LOQ), and 3 other concentrations that will encompass the expected concentration range (Table 1). Repeat the fortification process for Day 2 and Day 3 using a second and third set of 3 each (each source randomly selected such that each is represented at least once at each concentration) of the 6 control samples.

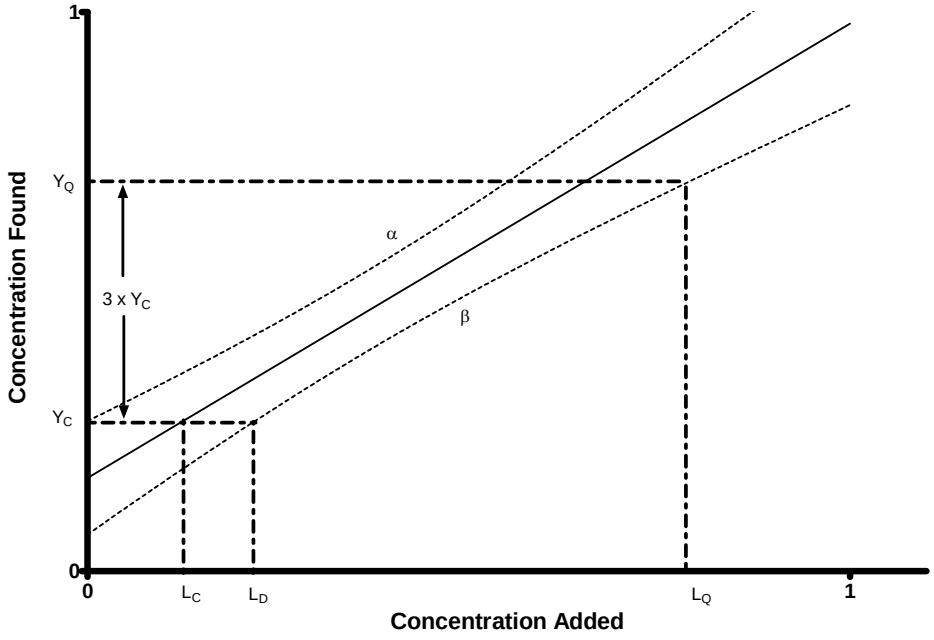
Table 1. Example of Minimum Study Design to Allow Determination of LOD, LOQ, Accuracy and Precision (Six Sources/Animals: A, B, C, D, E, and F) Within One Study

Fortification Concentration	Animal/Source ID†		
	Day/Run 1	Day/Run 2	Day/Run 3
0 (Control)	B, F, D	A, C, C	B, E, F
eLOD*	B, C, E	D, F, F	A, B, E
eLOQ (3 X eLOD)*	C, C, E	A, B, E	D, F, D
Lower part of Validation Range	A, B, E	A, C, D	B, E, F
Middle of Validation Range	B, C, E	C, E, F	A, D, F
Upper Part of Validation Range	A, B, B	D, F, F	A, C, E

* eLOD (estimated LOD) is generally determined from preliminary studies conducted during method development. eLOQ (estimated LOQ) is determined as 3 times eLOD.

† each source randomly selected such that each source is represented at least once at each concentration across the 3 validation runs.

3. Assay the 18 samples each day and evaluate the results against a calibration standard curve.
4. Plot the results of concentration found against concentration added across all three days of assays. This will normalize the data results across days and allow all the data from the 3 runs to be used in the determination of the LOD and LOQ.
5. Establish a decision limit by calculating prediction intervals around the weighted regression line with the upper confidence interval line based upon the probability α (false positive) and the lower confidence interval line based upon the probability β (false negative)[iv]. The decision limit (Y_C) then becomes the point at which the upper confidence limit crosses the Y-axis and can be converted to concentration by estimating from the regression line to the x-axis (L_C). This is the critical point where 50% of the responses are real. The L_D or LOD can be determined by estimating concentration from the lower confidence limit β that reduces the false negative rate to what level is assigned to β. Typically, both α and β are set equal to 5%.
6. Establish a determination limit (Y_Q) by multiplying the detection limit (Y_C) by 3 (commonly accepted ratio between LOD and LOQ is 3). The LOQ (L_Q) can then be determined by estimating where the line Y_Q crosses the lower confidence limit β that reduces the false negative rate for the determination of LOQ to what level is assigned to β (typically 5%).
7. Inter-day precision can be determined by calculating the %CV at each concentration evaluated. Accuracy can be determined by comparison of the results obtained to the fortification levels. Acceptance criteria for accuracy and precision are provided in Sections 2.2 and 2.3, respectively.



This approach takes into consideration the interrelationship between specificity, LOD and LOQ. By determining LOD and LOQ using 6 different sources of matrix, the variability due to the matrix as well as the variability of the assay is taken into account. Since specificity for residue

methods is dependent upon the possible interference of matrix components this approach also addresses specificity and insures that specificity is acceptable at the LOD and LOQ determined. This approach is consistent with the determination of the detection limit and quantitation limit specified in VICH GL2 (Validation Methodology) Guideline.

Data Set Example:

A validation procedure based on the above methodology was conducted on an ELISA assay.

Control swine serum obtained from six different animals were each fortified with the analyte at 0, 50, 150, 300, 600 and 1200 ng/mL giving a total of 36 samples. Because this was a serum assay and it was relatively easy to run, all six fortification levels were run on each of three days. Had this been tissue samples, we would have randomly chosen 3 of the 6 animals (insuring that each of the 6 animals were run at least once) at each of the fortification levels to run on each of the 3 days of assay for a total of 18 samples per day.

Based on these three days of analyses which consisted of 108 assays total (for tissue assays it would have been 54 assays total) the following determinations were done: repeatability (intra-day precision), inter-day precision, LOD and LOQ. The raw data and the results of the statistical analyses are listed below:

Run	**Fortification Level, ng/mL**	**Results, ng/mL**					
		Animal A	**Animal B**	**Animal C**	**Animal D**	**Animal E**	**Animal F**
1	0	nr	nr	nr	nr	nr	nr
	50	9	32	59	18	18	25
	150	162	160	148	145	133	128
	300	251	303	331	295	270	260
	600	508	514	592	513	568	609
	1200	907	1186	1162	1037	1050	1097
2	0	nr	nr	nr	nr	nr	nr
	50	26	41	40	36	37	27
	150	155	168	130	144	143	177
	300	234	251	335	307	251	247
	600	504	522	553	516	650	580
	1200	999	1030	1037	1020	985	996
3	0	1	nr	8	nr	nr	1

	50	39	60	71	50	68	48
	150	157	179	159	167	172	148
	300	290	277	336	319	299	278
	600	565	572	611	586	648	579
	1200	1071	1190	1218	1262	1246	1160

nr = no response

The statistical evaluation of the above data was conducted as follows: The percentage recovery was calculated for each sample using the concentration obtained and the fortification concentrations prior to analysis. A model which included the fixed effect of treatment (fortification level) and the random effects of run (day), sample preparation within the run, run by treatment interaction and residual was used to obtain the least squares means and estimates of variation.

In order to assess within-day variability, the residual variance was used in calculating the CV for each treatment and across treatments. The CVs were calculated by dividing the square root of the residual variance by the mean and multiplying by 100.

In order to assess across-day variability, the sum of the residual variance, the variance due to run, sample within run and run by treatment was used as the estimate of variance when calculating CVs for each treatment and overall treatments.

The results of the analysis were as follows:

Fortification* Level, ng/mL	**n**	**Recovery, %**	**95% Confidence Interval**	**Intra-day Precision, CV%**
150	36	102.7	93.1 – 112.2	11.8
300	36	95.0	86.1 – 104.6	9.2
600	36	94.3	85.6 – 103.0	7.8
1200	36	91.3	93.1 – 103.7	3.3
Overall	144	95.8	87.9 – 103.7	17.7

* 50 ng/mL fortification level was below the LOD and was not used to determine precision

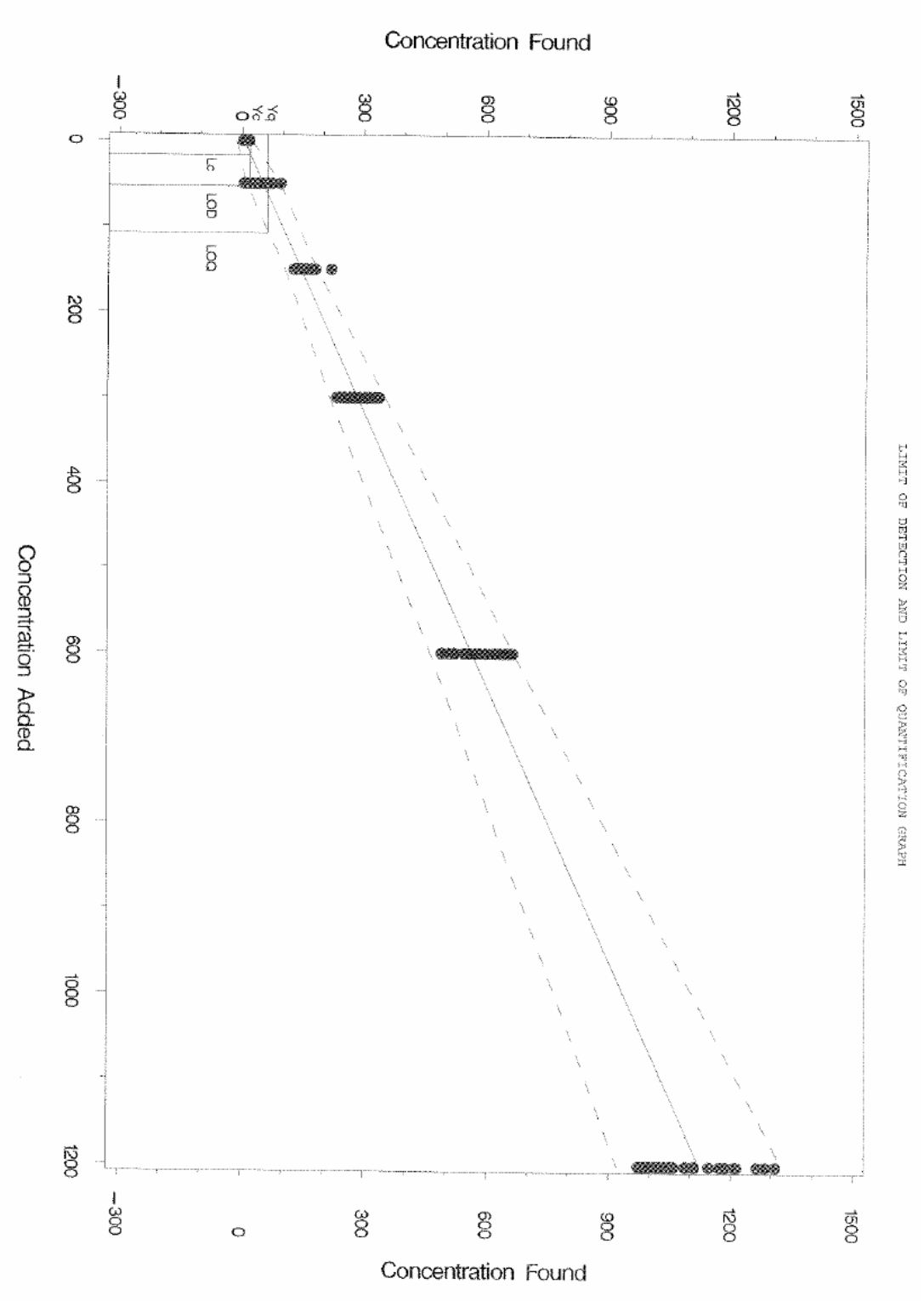
Repeatability (Overall Intra-Day Precision) = 17.7%

Inter-Day Precision = 19.1%

LOD = 62 ng/mL

LOQ = 112 ng/mL

A graphical representation of the determination of LOD and LOQ is provided below:



This is a straightforward way to accurately determine precision, accuracy, LOD and LOQ within one study across three days of validation. The LOD and LOQ are consistent with what one would expect from a subjective evaluation of the data. The precision is a bit high but considering this is an ELISA procedure it is not unexpected but yet still is acceptable based on the precision criteria outlined in this document.

---

[i] Codex Guidelines for the Establishment of a Regulatory Programme for Control of Veterinary Drug Residues in Foods, Part III Attributes of analytical Methods for Residue of Veterinary Drugs in Foods, p. 41, CAC/GL 16-1993.
[ii] Codex Guidelines for the Establishment of a Regulatory Programme for Control of Veterinary Drug Residues in Foods, Part III Attributes of analytical Methods for Residue of Veterinary Drugs in Foods, p. 42, CAC/GL 16-1993.
[iii] Codex Alimentarius Procedural Manual, 15th Ed., Twenty-eight Session of the Codex Alimentarius Commission, Rome, 2005, p 81.
[iv] Zorn ME, Gibbons RD, Sonzogni WC. Weighted Least-Squares Approach to Calculating Limits of Detection and Quantification by Modeling Variability as a Function of Concentration, *Anal Chem* **1997,** 69, 3069-3075.