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

VICH GL 49 (R2) (MRK) – METHOD USED IN RESIDUE DEPLETION STUDIES
October 2026
Revision at Step 9
For consultation at Step 4

STUDIES TO EVALUATE THE METABOLISM AND RESIDUE KINETICS OF VETERINARY DRUGS IN FOOD PRODUCING ANIMALS: VALIDATION OF ANALYTICAL METHODS USED IN RESIDUE DEPLETION STUDIES

Revision at step 9

Recommended for Consultation at Step 4 of the VICH Process
in October 2026
by the VICH Steering Committee

.

This Guideline has been developed and revised 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 HealthforAnimals, Rue d'Idalie 9-13, Box 5, B - 1050 Brussels (Belgium)
e-mail : sec@vichsec.org - Website : http://www.vichsec.org

1. INTRODUCTION 3
1.1. Objective of the guidance 3
1.2. Background 3
2. GUIDANCE 3
2.1 Purpose 3
2.2. Scope 3
3. PERFORMANCE CHARACTERISTICS 4
3.1. Linearity 4
3.2. Accuracy 5
3.3. Precision 5
3.4. Limit of Detection 6
3.5. Limit of Quantitation 6
3.6. Selectivity 6
3.7. Stability in Matrix 6
3.8. Processed Sample Stability 7
3.9. Robustness 7
4. GLOSSARY 8

1. INTRODUCTION

1.1. Objective of the guidance

This guidance document is intended to provide a general description of the criteria that have been found by the European Union (EU), Japan, United States of America (USA), Australia, New Zealand and Canada to be suitable for the validation of analytical methods used in veterinary drug residue depletion 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 (i.e., post approval control methods) and the validation requirements of the regulatory methods are usually well defined by various regulatory agencies worldwide and might even be defined by national or regional law. However, the residue depletion 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 depletion 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 VICH regions for use in the residue depletion studies. This validated method could 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 guidance provided here is intended to specifically address the validation of veterinary drug residue methods.

2. GUIDANCE

2.1 Purpose

The purpose of this document is to provide a general description of procedures that can be used for the validation of the analytical methods developed for the analysis of tissue samples obtained in residue depletion studies.

For purposes of this guidance, "acceptable" refers to the scientific evaluation of the analytical method in terms of the described validation criteria.

2.2. Scope

This guidance is only intended to apply to analytical procedures that have been developed for the evaluation of veterinary drug residue methods (assays developed to determine residues in marker residue depletion studies). It is not intended to define the criteria needed for validation of regulatory monitoring assay procedures.

This document provides performance characteristics of the residue assays that if followed would generally be considered acceptable by the regulatory agencies in the VICH regions. The intent is that methods validated according to this guidance will provide residue data that would generally be considered acceptable by the regulatory agencies in determining appropriate withdrawal periods.

While it is recognized that residue studies must be conducted under GLP using validated methodology, the actual method validation experiments do not fall within the scope of the GLP regulations. However, raw data generated as a result of a method validation should be archived as appropriate and be available for submission to regulatory authorities upon request.

3. PERFORMANCE CHARACTERISTICS

In general, 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 depletion studies.

3.1. 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, polynomial or other (as appropriate) regression plot of known concentration vs. response using a minimum of 5 different concentrations. Acceptability of weighting factors should be determined by evaluation of the residuals across three runs to determine if the residuals are 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 3.3. 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) could require log transformations to achieve linearity where other assays (e.g., ELISA, RIA) could 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.

3.2. 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. The accuracy should meet the range listed below:

Analyte Concentration*	Acceptable Range for Accuracy
< 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

3.3. 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 a within-run (repeatability) and between-run component.

The within- and between-run precision of the analytical method can be determined as part of the validation procedure. There is generally not a need to determine reproducibility (between-laboratory precision) in order to conduct a residue depletion study, because the laboratory that is developing the method is often the same laboratory assaying the samples from the residue study. Instead of establishing reproducibility of the assay, a between-run precision, can be determined. Within- and between-run 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. The precision should meet the range listed below.:

Analyte Concentration	Acceptable within-run precision (Repeatability), %CV	Acceptable between-run precision %CV*
< 1 µg/kg	30 %	45%
≥ 1 µg/kg < 10 µg/kg	25 %	32%
≥ 10 µg/kg < 100 µg/kg	15%	23%
≥ 100 µg/kg	10 %	16%

* as determined by the Horwitz equation $CV = 2^{(1-0.5 \log C)}$ where C = concentration expressed as a decimal fraction (e.g., 1 µg/kg is entered as 10^{-9}).

3.4. 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 could be used as long as a scientific justification is provided for their use. See Annexes 1, 2, and 3 for examples of acceptable methods for determining LOD.

3.5. 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 could be used as long as scientific justification is provided. See Annexes 1, 2, and 3 for examples of acceptable methods for determining LOQ.

3.6. Selectivity

Selectivity is the ability of a method to distinguish between the analyte being measured and other substances which might be present in the sample being analyzed. For the methods used in residue depletion studies, selectivity is primarily defined relative to endogenous substances in the samples being measured. Because the residue depletion studies are well controlled, exogenously administered components (i.e., other veterinary drugs or vaccines) could either be known or not be allowed during the study. If it is the intent to submit the validated method as a regulatory method, it might 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. That response should be no more than 20% of the response at the LOQ (see Section 3.5 above). See Annex 3 for a suggested protocol for determining accuracy, precision, LOD, LOQ and selectivity in a single study.

3.7. Stability in Matrix

Samples (tissue, milk, eggs or honey) collected from residue depletion studies are generally frozen and stored until assayed. It is important 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 should be periodically assayed at specified intervals (e.g., initially, 1 week, 1 month, 3 months). If the samples are frozen, freeze/thaw studies should 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 initial assay results or freshly fortified control sample assay results within the accuracy acceptance criteria established in Section 3.2.

3.8. 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. over a weekend. The stability of the analyte in the process sample extract might 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 might 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 accuracy acceptance criteria established in Section 3.2.

3.9. 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 can become apparent and variations in the ones most likely to affect the method performance should be evaluated.

4. 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 % bias.

Between-run Precision – Between-run precision expresses within-laboratory between-run variations.

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.

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.

Marker residue – The residue whose concentration is in a known relationship to the concentration of total residue in an edible tissue.

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.

Residue – Veterinary drug (parent) and/or its metabolite.

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 might be expected to be present.

Within-run Precision – Within-run precision expresses within-laboratory within-run variations.

Annex 1

Examples of Methods for Determining LOD and LOQ

One commonly used approach is referred to as the IUPAC definition.[1] 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 3.2 and 3.3), then the estimated LOQ is acceptable.

Annex 2

U.S. Environmental Protection Agency Method for Determining LOD and LOQ

The procedure described below is a slight modification of a procedure used by USDA's Interregional Project No. 4 program, which is published in 40 CFR Part 136, Appendix B.[2] This modified procedure can be found in Appendix 1 of the U.S. Environmental Protection Agencies document entitled "Assigning Values to Non-detected/Non-quantified Pesticide Residues in Human Health Food Exposure Assessments.[3] The procedure is provided below with minor modifications making it more representative of a tissue marker residue assay procedure example.

In this procedure, the estimation of the LOD and LOQ of a specific method for a specific analyte in a specific matrix can be done in the following two steps.

- The first step is to produce a preliminary estimate of the LOD and LOQ and to verify that a linear relationship between concentration and instrument response exists. These preliminary estimates correspond to what some term the IDL (Instrument Detection Limit) and IQL (Instrument Quantitation Limit), respectively. The matrix of interest will be fortified (spiked) at the estimate LOQ in the next step for the actual estimation of LOD and LOQ of the method.
- The second step is to use the initial estimate of the LOD and LOQ determined in Step 1 to estimate the method detection limit and the method quantitation limit in the matrix of interest.

An illustrative example follows:

Step 1. The analyst derives a standard curve for the method of interest. In this particular instance, the analyst prepares the standard solution in buffer or water with the following concentrations of the analyte of interest: 0.005, 0.010, 0.020, 0.050 and 0.100 µg/mL. For each concentration in the sample solution, the following instrument responses (measure peak height) are recorded:

Concentration (µg/mL)	Instrument Response (peak height)
0.100	206,493
0.050	125,162
0.020	58,748
0.010	32,668
0.005	17,552

In order to verify that a linear response is seen throughout the tested range, the instrument response is plotted as a function of injected concentration. The results (and associated statistics) are shown in Figure 1. Note from these results that the instrument response appears to be adequately linear throughout the range of tested concentrations (0.005 to 0.100 µg/mL) with a high R^2 value of 0.99 (see the "Summary of Fit" table in Figure 1). The residual standard deviation (root mean square error) is 8986.8. The equation which describes this relationship (provided in the "Parameter Estimates" table in Figure 1) is as follows:

$$Y = 15{,}120 + 1{,}973{,}099 * (\text{Concentration})$$

Where Y is the instrument response (peak height)

The estimated LOD and LOQ are calculated as follows (assuming these values are set to 3 and 10 standard deviations above the blank response, respectively):

1. The Peak Height at the LOD (Y_{LOD}) is calculated at 3 times the standard deviation while the Peak Height at the LOQ (Y_{LOQ}) is calculated at 10 times the standard deviation

   $$Y_{LOD} = 15120 + 3 * (8987) = 42{,}081$$

   $$Y_{LOQ} = 15120 + 10 * (8987) = 104{,}990$$

2. These values (peak height at LOD and peak height and LOQ) are then used to calculate the concentrations associated with these peak heights as follows:

   $$Y = 15{,}120 + 1{,}973{,}099 * (\text{Concentration})$$

   Rearranging,

   $$\text{Concentration} = (Y - 15{,}120) / 1{,}973{,}099$$

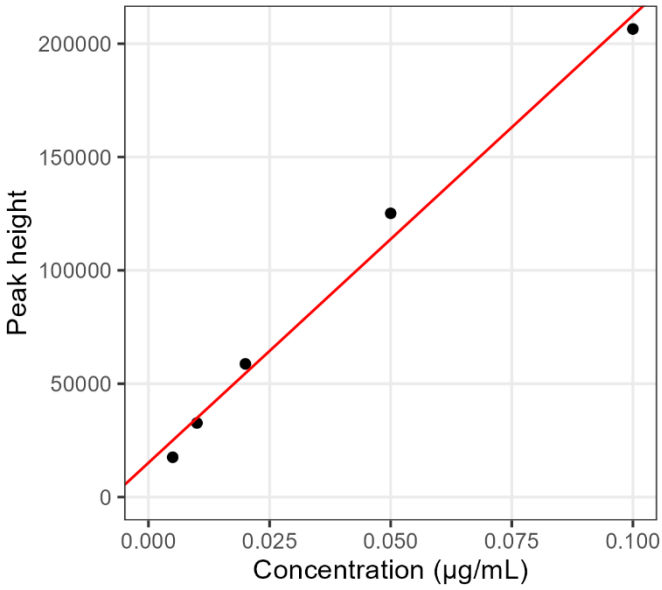
   Therefore,

   $$LOD = Y_{LOD} - 15{,}120 / 1{,}973{,}099 = (42{,}081 - 15{,}120) / 1{,}973{,}099 = 0.014 \text{ µg/mL}$$

   $$LOQ = Y_{LOQ} - 15{,}120 / 1{,}973{,}099 = (104{,}990 - 15{,}120) / 1{,}973{,}099 = 0.046 \text{ µg/mL}$$

   Thus, the initial estimated LOD and LOQ are 0.014 and 0.046 µg/mL, respectively which correspond to the IDL and IQL.

   These estimated LODs (or IDLs) and LOQs (or IQLs) are expressed in terms of the solution concentration and not in terms of the matrix concentration. At this stage, the solution concentration (µg/mL solution) should be converted to the effective concentration in the matrix (e.g., µg/g of matrix).



Summary of Fit	
R^2	0.99
Adjusted R^2	0.987
Root mean square error	8986.8
Mean of response	88125
Observations	5

Parameter Estimates				
Term	Estimate	Std. error	t value	p value
Intercept	15120	5835	2.59	0.0810
Slope	1973099	114317	17.26	0.0004

Figure 1. Statistical Results

Step 2. With the initial estimate of LOD (or IDL) and LOQ (or IQL) obtained and linearity verified, Step 2 involves estimating the LOQ and LOD in spiked matrix samples. This procedure uses the estimated instrumental LOQ and the procedure detailed in 40 CFR Part 136, Appendix B[2] to provide a better estimate of LOQ and verifies that method recoveries are acceptable.

The method calls for the analysis of 7 or more untreated control samples spiked at the estimated LOQ. The standard deviation of these samples is measured and the LOD and LOQ are determined as follows:

$LOD = t_{0.99} * S$

$LOQ = 3 * LOD$

where t = one-tailed t-statistic at the 99% confidence level for n-1 replicates

S = Standard Deviation of n sample spikes at the estimated LOQ

The following is a set of t-values for use in the above equation:

# of Replicates (n)	Degrees of Freedom (n-1)	$t_{0.99}$	# of Replicates (n)	Degrees of Freedom (n-1)	$t_{0.99}$
3	2	6.965	13	12	2.681
4	3	4.541	14	13	2.650
5	4	3.747	15	14	2.624
6	5	3.365	16	15	2.602
7	6	3.143	17	16	2.583
8	7	2.998	18	17	2.567
9	8	2.896	19	18	2.552
10	9	2.821	20	19	2.539
11	10	2.764	21	20	2.528
12	11	2.718	22	21	2.518

In this example, the analyst prepared 7 untreated control samples spiked at the above estimated LOQ of 0.05 µg/g. The following results were obtained:

Concentration detected (µg/g)	% Recovery
0.0397	79.4
0.0403	80.6
0.0400	80.0
0.0360	72.0
0.0498	99.6
0.0379	75.8
0.0388	77.6

Average Concentration: 0.0404 µg/g

Standard Deviation: 0.0044 µg/g

Average Recovery: 80.7%

Given that recoveries are adequate at the LOQ (average = 80.7%, range = 72.0% to 99.6%), the LOD and LOQ for the method are estimated as follows:

LOD = $t_{0.99}$ * S (for 7-1 = 6 degrees of freedom)

= 3.143 * 0.0044 µg/g

= 0.0138 µg/g

LOQ = 3 * LOD

= 3 * 0.0138 µg/g

= 0.0414 µg/g

Annex 3

Protocol for Residue Method Validation

Selectivity, LOD and LOQ are all interrelated and are affected by endogenous interferences that might be present in the matrix being assayed. LOD is often 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 (SD) 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 accuracy, precision, selectivity, LOD, and LOQ in one study.

1. Collect drug-free matrix from 6 separate sources (animals) and screen for any possible analyte contamination.
2. From the 6 sources, randomly select at least 3 samples to fortify (spike) at the estimated LOD (eLOD, determined during preliminary assay development), at the estimated LOQ (eLOQ, 3x the eLOD), and at 3 other concentrations that cover the expected concentration range. Repeat the fortification process for runs on Day 2 and Day 3 using a second and third set of 3 randomly selected samples per fortification level. Ensure that each source is represented at least once per fortification level across the 3 days. See Table 1 for an example.

Table 1 – Example of a minimal study design with six sources (animals) A through F

Fortification level	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	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

3. Assay the 18 samples each day and calculate the concentrations from the calibration standard curve.
4. Accuracy can be determined by comparison of the concentration found to the concentration added. Within- and between-run precision (i.e., intra- and inter-day precision) can be calculated as the %CV at each fortification level. Precision estimates incorporate variability due to the matrix, since the protocol uses 6 different sources of matrix. Acceptance criteria for accuracy and precision are provided in Sections 3.2 and 3.3, respectively.
5. If accuracy and precision are acceptable, plot the concentration found against the concentration added for all 3 days of assays. Whether variability remains constant across

fortification levels (homoscedasticity) or increases (heteroscedasticity) can be assessed visually from the plot or by examining the residuals of the regression.

6. With homoscedastic data, the LOD and LOQ can be calculated with the approach in Sections 6.3 and 7.3 of VICH GL2, respectively. A linear regression of concentration found vs concentration added is fitted, and then LOD is estimated as $3.3\sigma/s$ and LOQ as $10\sigma/s$, where σ is the regression's residual SD and s is the slope. The approach can also be used with heteroscedastic data, but σ and s must be estimated from a weighted regression in which the residuals are scaled by their weights prior to calculating σ. In this protocol, σ is the residual SD of the regression. The alternative VICH GL2 methods to estimate σ (SD of blanks or SD of y-intercepts) are not applicable because control samples may show zero or near-zero responses in sensitive methods, and only 3 runs are available for calculations in this scenario.
7. Alternatively, LOD and LOQ can be derived from the prediction intervals around the regression[4], as shown in Figure 1. With homoscedastic data, ordinary regression can be used, but with heteroscedastic data, weighted regression should be used.[5] The upper prediction interval is calculated to control the false positive rate (α), while the lower prediction interval is calculated to control the false negative rate (β). Typically, both α and β are 5%, so the prediction intervals are estimated to cover 90% of measurements, from the 5^{th} to 95^{th} percentile. The decision limit (Y_C) is where the upper interval crosses the y-axis. The corresponding point on the x-axis (L_C, also called the critical value or CC_α) represents the concentration where there is a 50% probability of correctly identifying the analyte as present. At the concentration where the decision limit intersects the lower prediction interval (also known as CC_β), the probability of detecting the analyte when it is present is $1-\beta$. This concentration is taken as the LOD. The LOQ is derived similarly. A determination limit (Y_Q) is established by multiplying the decision limit (Y_C) by 3, which is a commonly accepted ratio between LOD and LOQ (see e.g., VICH GL2). Where the determination limit intersects the lower prediction interval is taken as the LOQ.
8. The selectivity of the assay is considered sufficient if the response of the blank, control samples is less than 20% of the LOQ, as noted in Section 3.6. If this criterion is not met, the LOQ should be adjusted upward to ensure adequate selectivity.

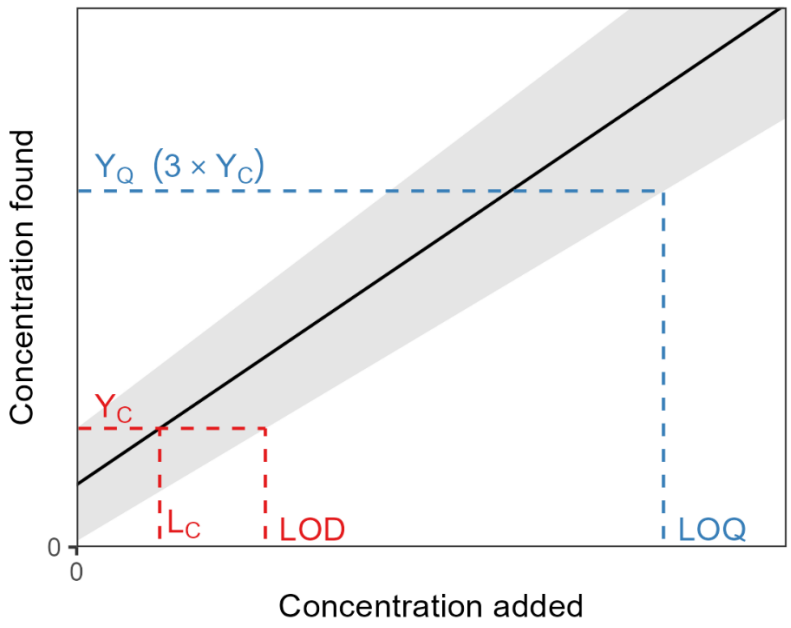


Figure 1 – estimating LOD and LOQ from the prediction intervals around the regression of concentration found vs concentration added. In this example, the intervals widen with increasing concentration, indicative of heteroscedasticity.

Example:

A validation procedure based on the above methodology was conducted on an LC-MS/MS milk assay procedure.

Control bovine milk samples obtained from 6 different animals were each fortified with the analyte at 0, 4.2 (eLOD), 14.0 (eLOQ), 35, 140 and 400 ng/mL, giving a total of 36 samples. For each run, milk samples from 3 of the 6 animals were randomly chosen for each fortification level, ensuring that each of the 6 animals was selected at least once per level. This results in 18 measurements per day and a total of 54 measurements across 3 runs on 3 different days.

The following parameters were evaluated: accuracy, within-run precision (repeatability), between-run precision, LOD and LOQ. The raw data are listed in Table 2.

Table 2 – Concentration of analyte (ng/mL) in fortified control milk across three days of analysis						
Conc. added	Run 1		Run 2		Run 3	
	Animal ID	Conc. found	Animal ID	Conc. found	Animal ID	Conc. found
0	B	0.494	A	0.233	B	0.154
	F	0.654	C	0.012	E	0.120
	D	0.588	C	0.117	F	0.313
4.2	B	4.38	D	4.97	A	3.80
	C	4.13	F	3.85	B	4.12
	E	4.33	F	4.41	E	3.67
14.0	C	13.2	A	11.1	D	11.8
	C	13.5	B	12.0	F	10.5
	E	11.9	E	12.8	D	11.7
35.0	A	31.5	A	31.0	B	27.3
	B	32.7	C	33.2	E	29.4
	E	34.4	D	32.9	F	25.5
140	B	131	C	137	A	118
	C	147	E	124	D	106
	E	127	F	131	F	118
400	A	396	D	396	A	335
	B	394	F	390	C	316
	B	384	F	373	E	344

At each fortification level, the mean concentration found was calculated and then the accuracy was expressed in percent recovery. To assess precision, the within- and between-run variances were estimated; these variances can be derived directly or with the aid of models, like mixed models or analysis of variance (ANOVA).[6] Here, a separate ANOVA was conducted for each fortification level, with "concentration found" as dependent variable and "run" as a discrete, independent variable. An ANOVA's mean square error represents the within-run variance at the given fortification level. Criteria for between-run precision are from the Horwitz equation (see Section 3.3) and are based on the total between-run variance, which is the sum of the within-run variance and the 'pure' between-run variance. The 'pure' between-run variance represents the variability due to run-to-run differences, beyond that within each run. It can be determined from a balanced ANOVA as the difference between the mean square for 'run' and the mean square error, divided by the number of measurements per group (3 per run in this example).[6] Equivalently, it is the between-run random effect variance in a mixed model. Precision estimates were expressed as CVs, calculated by dividing the square root of the variance by the mean and multiplying by 100.

Table 3 – Accuracy and precision estimates for each fortification level*							
Concentration added (ng/mL)	n	Mean concentration found (ng/mL)	Mean recovery (%)	Precision			
				Within-run		Between-run	
				Variance	CV (%)	Variance**	CV (%)
4.2	9	4.18	99.6	0.128	8.56	0.167	9.77
14	9	12.1	86.1	0.657	6.72	1.03	8.42
35	9	30.9	88.2	2.45	5.07	10.8	10.6
140	9	127	90.4	67.4	6.49	168	10.2
400	9	370	92.4	129	3.08	1180	9.30

*Recovery and CV calculated directly from the means and variances in the table may be slightly different than the stated values due to rounding. Deriving means and variances from ANOVA or mixed models, or conducting weighted regression to estimate prediction intervals (see below), can be done with statistical software (e.g., R or SAS).
**The sum of the within-run variance and the 'pure' between-run variance.

Method accuracy, within-run precision and between-run precision were all acceptable (Table 3).

Plotting concentration found vs concentration added showed that variability increased with concentration. Given this heteroscedasticity, weighted regression is appropriate. Commonly, each data point is weighted by the inverse of its variance (1/variance), although other weighting functions (e.g., $1/x$, $1/x^2$) may be suitable depending on the data. Here, weights were inverse variances, and to estimate them, the relationship between the empirical SD and concentration added was modelled. This relationship is often linear[5], and with the example dataset, it was SD = 0.297 + 0.076 * concentration. Squaring the SD yielded by this weighting function gave variance estimates, and inverting them yielded the weights 11.3, 2.64, 0.541, 0.115, 0.008, and 0.001 respectively for the 6 fortification levels that were then used in the weighted regression.

The weighted regression's slope was 0.893 and the intercept was 0.292 ng/mL. The SD of the weight-adjusted residuals was 0.906, so according to VICH GL2, LOD and LOQ were estimated as 3.35 ng/mL and 10.1 ng/mL, respectively.

The alternative approach of deriving LOD and LOQ from the weighted regression's prediction intervals is demonstrated in Figure 2. The prediction intervals were calculated according to Zorn et al.[4] with α and β set at 0.05 and using df = n - 4 to account for the parameters of both the weighted regression (2 df) and the weighting function (2 df). The decision limit was 0.763 ng/mL, the determination limit was 2.29 ng/mL, the critical value (Lc) was 0.528 ng/mL, the LOD was 1.21 ng/mL, and the LOQ was 3.16 ng/mL. Consistent with this estimated LOQ, accuracy and precision were acceptable at the lowest fortification level of 4.2 ng/mL (Table 3).

Regarding selectivity, the maximum response of blank, control samples (0.654 ng/mL) was less than 20% of the LOQ estimated with GL2, but it was about 20% of the LOQ estimated from

prediction intervals. So, in this case, taking a slightly higher LOQ would ensure that background noise is not considered quantifiable analyte.

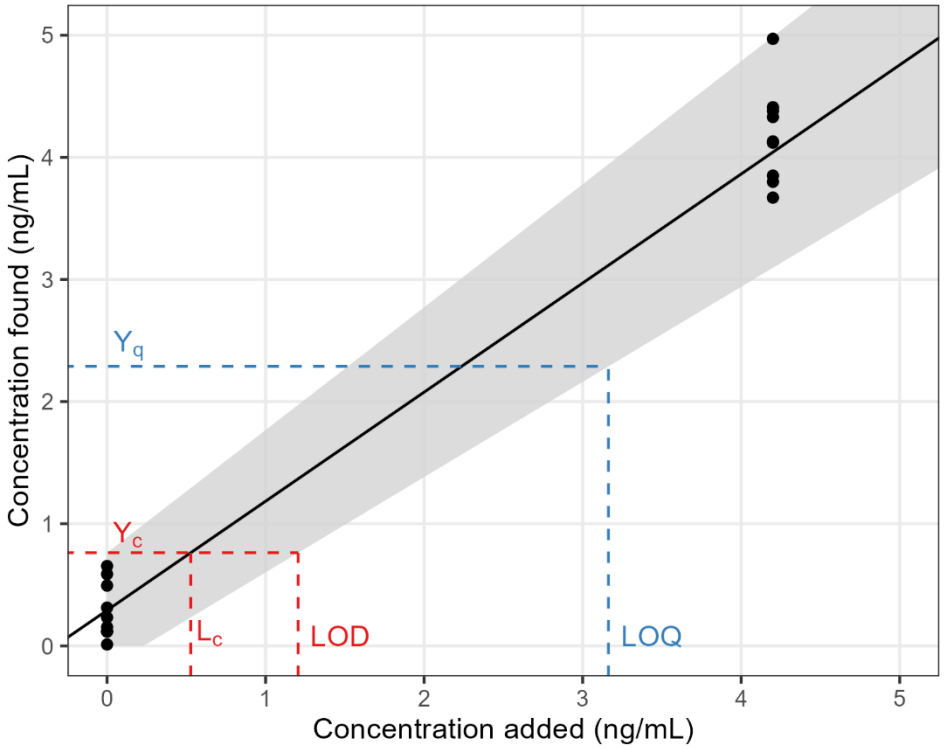


Figure 2 – derivation of LOD and LOQ from the example dataset.

---

[1] Codex Alimentarius Procedural Manual, 15th Ed., Twenty-eight Session of the Codex Alimentarius Commission, Rome, 2005, p 81.
[2] U.S. Code of Federal Regulations, Title 40: Protection of Environment , Part 136 – Guidelines Establishing Test Procedures for the Analysis of Pollutants, Appendix B to Part 136 – Definition and Procedure for the Determination of the Method Detection Limit – Revision 1.11.
[3] U.S. Environmental Protection Agency, Office of Pesticide Programs, March 23, 2000, "Assigning Values to Non-detected/Non-quantified Pesticide Residue in Human Health Food Exposure Assessments" Appendix 1, A-1 through A-8.
[4] 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.
[5] ISO 11843-2: 2000, Capability of detection – Part 2: Methodology in the linear calibration case.
[6] ISO 5725-2: 2019, 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.