



1 15 October 2021
2 EMA/CVMP/SWP/735325/2012 Rev.2*
3 Committee for Medicinal Products for Veterinary Use (CVMP)

4 Guideline on determination of withdrawal periods for 5 edible tissues

6 Draft

Draft agreed by Safety Working Party (SWP-V)	30 September 2021
Adopted by CVMP for release for consultation	7 October 2021
Start of public consultation	15 October 2021
End of consultation (deadline for comments)	17 December 2021

7
8 Comments should be provided using this [template](#). The completed comments form should be sent to Vet-guidelines@ema.europa.eu

9
10

Keywords	<i>Withdrawal periods, tissues, statistical approach, alternative approach</i>
-----------------	---

11
12 *The current revision consists of administrative changes made in order to align the guideline to the
13 new definitions and terminology provided by Article 4 of Regulation (EU) 2019/6. The references to the
14 legislation applicable and other scientific guidelines have also been updated.



16 Guideline on determination of withdrawal periods for
17 edible tissues

18 **Table of contents**

19 **Executive summary..... 3**

20 **1. Introduction (background) 3**

21 **2. Scope..... 4**

22 **3. Legal basis..... 4**

23 **Statistical approach to the establishment of withdrawal periods..... 5**

24 **4. General considerations..... 5**

25 4.1. Statistical approach 5

26 4.1.1. Calculation model 5

27 4.1.2. Data base 5

28 4.1.3. Linear regression analysis assumptions 5

29 4.1.3.1. *Homogeneity of variances* 6

30 4.1.3.2. *Log-linearity* 6

31 4.1.3.3. *Normality of errors* 6

32 4.1.4. Estimation of withdrawal periods by regression analysis..... 7

33 4.2. Possible alternative approach 8

34 4.3. Injection site residues 9

35 4.4. MRL or other reference value..... 9

36 **4.5 Generic products 9**

37 **5. Example for the statistical analysis of residue data..... 11**

38 **6. Discussion on the regression analysis..... 19**

39 6.1. To what extent a departure from the regression assumptions may be acceptable?..... 19

40 6.2. Withdrawal periods should be set by interpolation and not by extrapolation..... 19

41 6.3. Should the 95% or the 99% tolerance limit be applied?..... 19

42 6.4. Dealing with 'less than' values..... 20

43 6.5. Dealing with obvious outliers..... 20

44 6.6. Combining data sets 20

45 6.7. The possibility of overriding one study with another 21

46 **7. References..... 22**

47 **Annex A 24**

48 **Annex B 26**

49 **Annex C 30**

50 **Annex D: Comparisons of different approaches for dealing with values**

51 **below the LOQ 32**

52

53 **Executive summary**

54 This document, originally published in 1997 as the CVMP Note for guidance: approach towards
55 harmonisation of withdrawal periods, provides detailed guidance on how to establish withdrawal
56 periods and was developed by the CVMP in order to provide a standardised approach for derivation of
57 withdrawal periods within the European Union. Much of the document is focused on the statistical
58 approach used by CVMP, but an alternative, for use in those cases where the data do not allow use of
59 the statistical approach, is also described. The issue of withdrawal periods for substances with a 'No
60 MRL required' classification is also addressed.

61 **1. Introduction (background)**

- 62 1. Even where Community MRLs have been established, similar products in various Member States
63 may differ greatly with respect to the withdrawal periods established by national authorities.
- 64 2. The 1997 note for guidance enabled applicants and assessors from all member states to use the
65 same approach for determining withdrawal periods (WPs), leading to fewer discrepancies between
66 authorised WPs for the same product in different member states (MS). The same approach is also
67 used in centralised and decentralised procedures.
- 68 3. The Committee considers that the statistical approach offers the greatest opportunity for
69 harmonisation.
- 70 4. Consumer safety needs to be assessed for all pharmacologically active substances which are
71 intended for use in food producing animals in accordance with Regulation (EC) No 470/2009 [21].
72 Pre-slaughter withdrawal periods are determined in order to ensure that the residues deplete to
73 permissible concentrations. The withdrawal period is set individually for each veterinary product as
74 part of the marketing authorisation process.
- 75 5. New residue depletion studies for the establishment of withdrawal periods should be conducted in
76 accordance with VICH GLs 48 (14) and 49 (15) and VICH GL 57 for aquatic species (20), and with
77 Directive 2004/10/EC (GLP)(17). Data from these studies should be adequate in most cases to use
78 a statistical method. On these occasions, applicants should use the statistical software provided by
79 the CVMP (found on the EMA website) in order to determine a suitable WP for their product(s). The
80 underlying statistics for this software are described in Annexes A-C of this guideline.
- 81 6. Occasionally, data from studies are insufficient to evaluate the withdrawal period using the
82 statistical method. This could be the case for old studies that were conducted before the
83 publication of the requirements described in Volume 8 of The rules governing medicinal products in
84 the EU (replaced by Regulation (EU) 2018/782 (13)), or VICH GLs 48 and 49 or VICH GL 57 for
85 aquatic species. This may also be the case for recent studies where most of the residue
86 concentrations are below the LOQ of the analytical method, or where inappropriate timepoints have
87 been chosen.
- 88 7. Under these circumstances, a more pragmatic approach is necessary. For this reason, an
89 alternative method, which has been used successfully throughout the union for many years, has
90 also been included; however, it should only be used where the statistical method(s) cannot be
91 used.

92 The objective of the present paper is to provide guidance on how to establish withdrawal periods for
93 edible tissues of food producing animals. This also includes consideration of the injection sites.
94 However, injection site residues are more specifically addressed in the guideline on injection site
95 residues (EMA/CVMP/542/03-FINAL) (12). This guideline does not address withdrawal periods in milk;

96 guidance is provided in the CVMP Guideline on determination of withdrawal periods for milk
97 (EMA/CVMP/SWP/735418/2012) (18).

98 Emphasis has been put on a statistical approach. As the method of first choice, a linear regression
99 technique is recommended. Data from an actual residue study were used to demonstrate the
100 applicability of this recognised statistical technique. A step by step procedure is described which has
101 been drawn up with the FDA guideline (1,2) as a basis. It is recommended in this paper to determine
102 withdrawal periods at the time when the upper one-sided 95% tolerance limit for the residue is below
103 the MRL (or other reference value¹) with 95% confidence. However, for comparison of approaches (cf.
104 FDA), 99% tolerance limits with 95% confidence are also calculated.

105 **2. Scope**

106 This guideline describes a standardised approach for the determination of withdrawal periods within the
107 European Union, focusing particularly on use of a statistical method, but providing additional guidance
108 on an alternative approach, for use in those cases where the data do not allow use of the statistical
109 approach (e.g., where the statistical assumptions are not met).

110 In addition, the guideline discusses the possible need for withdrawal periods for products containing
111 substances for which a 'No MRL required' status has been established, as well as considerations for
112 generic products.

113 **3. Legal basis**

114 In line with article 10, 14, 16 and 35 of Regulation (EU) 2019/6 of the European Parliament and of the
115 Council of 11 December 2018 on veterinary medicinal products as repealing Directive 2001/82/EC
116 entered into force on 28 January 2019 and applicable from 28 January 2022 onwards, the product
117 information for veterinary medicinal products for use in food producing species shall contain
118 information on the withdrawal period. The dossier requirements in order to enable the determination of
119 withdrawal periods are mentioned in part II. 3.B of the Commission Delegated Regulation (EU)
120 2021/805. Article 4 of the Regulation (EU) 2019/6 defines the following definition for the withdrawal
121 period:

122 "*(34) 'withdrawal period' means the minimum period between the last administration of a veterinary*
123 *medicinal product to an animal and the production of foodstuffs from that animal which under normal*
124 *conditions of use is necessary to ensure that such foodstuffs do not contain residues in quantities*
125 *harmful to public health"*

¹ In the absence of a numerical MRL, another reference value may be used as the relevant limit for residue concentrations. This is applicable whenever "MRL" is mentioned further below.

126 **Statistical approach to the establishment of withdrawal** 127 **periods**

128 **4. General considerations**

129 **4.1. Statistical approach**

130 **4.1.1. Calculation model**

131 The calculation model for the statistical determination of withdrawal periods is based on accepted
132 pharmacokinetic principles. According to the pharmacokinetic compartment model, the relationship
133 between drug concentration and time through all phases of absorption, distribution and elimination is
134 usually described by multi-exponential mathematical terms. However, the terminal elimination of a
135 drug from tissues, the residue depletion, in most cases follows a one-compartment model and is
136 sufficiently described by one exponential term. The first order kinetic equation for this terminal
137 elimination is:

$$138 C_t = C_0' e^{-kt}$$

139 C_t is the concentration at time t , C_0' is a pre-exponential term (fictitious concentration at $t = 0$) and k
140 is the elimination rate constant.

141 Linearity of the plot $\log_e C$ versus time indicates that the model for residue depletion is applicable and
142 linear regression analysis of the logarithmic transformed data can be considered for the calculation of
143 withdrawal periods.

144 **4.1.2. Data base**

145 Regression analysis requires data which are independent from each other. Normally, residue depletion
146 data meet this assumption because they originate from individual animals. In cases of duplicate or
147 triplicate measurements of samples the mean value of each sample has to be used for the calculation.
148 To avoid biasing slope and intercept, each data point of the regression line should originate from the
149 same number of repeated sample measurements.

150 When all, or most of, the reported data from a slaughter time-point are 'less than' values (data which
151 are below the LOQ), excluding the whole time point from the analysis should be considered. However,
152 it should be borne in mind that 3 time points are necessary to allow a meaningful regression analysis.
153 See section 6.4 and Annex D for further discussion on 'less than' values.

154 The numbers of animals to be used for residue depletion studies is specified in guideline VICH GL48
155 (14) and VICH GL 57 for aquatic species (20). There, depending on the animal species and type of
156 depletion study, 4-10 animals per time point are recommended.

157 In some cases, depending on the validation of the analytical method and how this has been conducted,
158 corrections for recovery and/or stability of the analyte(s) in relevant matrices may be required.

159 **4.1.3. Linear regression analysis assumptions**

160 It is necessary for linear regression analysis that the following regression assumptions are valid:

- 161 • assumption of homogeneity of variances of the \log_e -transformed data on each slaughter day,
- 162 • assumption of linearity of the \log_e -transformed data versus time,

163 • assumption of a normal distribution of the errors.

164 **4.1.3.1. Homogeneity of variances**

165 It should be confirmed that the variances of the log_e-transformed concentrations of the different
166 slaughter days are homogeneous.

167 Several tests are available. The FDA (1,2) recommends Bartlett's test. Bartlett's test is said to be the
168 most powerful test, but it is extremely sensitive to deviations from normality. Furthermore, the test
169 should only be used, when each group numbers 5 or more. Equal sample sizes are not required (3).

170 Other commonly used tests for homogeneity of variances are Hartley's test and Cochran's test.
171 Hartley's test can only be used if all groups are of the same size (3).

172 In the Committee's view, Cochran's test is the best choice. It is easier to perform than Bartlett's test,
173 and it uses more information than Hartley's test. Furthermore, it is not as sensitive to departures from
174 normality as Bartlett's test. Cochran's test may be used for data whose group sizes do not differ
175 substantially by calculating the harmonic mean of the group sizes.

176 **4.1.3.2. Log-linearity**

177 Visual inspection of a plot of the data is often sufficient to assure that there is a useful linear
178 relationship. Obvious deviations from linearity at early time points may indicate that the drug
179 distribution processes have not yet ended. These time points should therefore be excluded. Deviations
180 from linearity at late time points may be due to concentrations below the limit of detection. Depletion
181 kinetics cannot be observed at these time points, and it is justified to exclude these data. It should,
182 however, be borne in mind that all other time points have to be kept, unless there is a clear
183 justification for their omission.

184 For statistical assurance of the linearity of the regression line an analysis of variances has to be
185 performed (lack-of-fit test). The usual procedure is to compare the variation between group means and
186 the regression line with the variation between animals within groups (see Section 5, Step 5; F-test
187 (3)).

188 An appropriate supplementation to the lack-of-fit test is the test of the significance of the quadratic
189 time effect according to Mandel (9). The question is, whether a quadratic fit is better than the linear fit.
190 The calculation procedure is described in Annex C of this guideline.

191 **4.1.3.3. Normality of errors**

192 A good visual test is to plot the ordered residuals versus their cumulative frequency distribution on a
193 normal probability scale. Residuals are the differences between the observed values and their
194 expectations (i.e., the difference between the observed log_e-transformed concentration and the value
195 predicted by the regression line).

196 A straight line indicates that the observed distribution of residuals is consistent with the assumption of
197 a normal distribution. In order to verify the results of the residual plot, the Shapiro-Wilk test can be
198 applied. This test has been shown to be effective even if sample sizes are small (4).

199 The plot of the cumulative frequency distribution of the residuals can be used as a very sensitive test.
200 Deviations from a straight line, indicating non-normality of the residuals, may be due to:

201 • deviations from normality of the log_e-transformed residue concentrations within one or more
202 slaughter groups,

- 203 • deviations from log_e-linearity of the regression line,
- 204 • non-homogeneity of variances,
- 205 • outliers.

206 In the selected presentation of the data using standardised residuals (standardised by dividing by the
 207 residual error $s_{y.x}$), an outlier would have a value < -4 or $> +4$, indicating that the residual is 4
 208 standard deviations off the regression line (see Fig. 1, 2).

209 **4.1.4. Estimation of withdrawal periods by regression analysis**

210 The withdrawal period should be estimated using the results of linear regression calculations.
 211 Withdrawal periods are determined as the time when the upper one-sided tolerance limit with a given
 212 confidence is below the MRL¹. If this time point does not make up a full day, the withdrawal period is
 213 to be rounded up to the next day.

214 The FDA (1,2) recommends calculating the 99th percentile of the population with a 95% confidence
 215 level by a procedure which requires the non-central t-distribution.

216 The calculation of the one-sided upper tolerance limit (95% or 99%) with a 95% confidence according
 217 to K. Stange (5) is recommended in this guideline. This method of calculation has comparable results
 218 (see Annex B) and is easier to perform since only the percentage points of the standardised normal
 219 distribution are required.

220 With the Stange equation, one estimates (with a confidence of $1-\alpha$) the proportion of $1-\gamma$ of the
 221 population which at least is to be expected to be below the one-sided upper tolerance limit. The
 222 respective percentage points of the standardized normal distribution are $u_{1-\alpha}$ and $u_{1-\gamma}$ (e.g., for
 223 $1-\alpha = 0.95$ is $u_{1-\alpha} = 1.6449$, for $1-\gamma = 0.95$ is $u_{1-\gamma} = 1.6449$, and for $1-\gamma = 0.99$ is $u_{1-\gamma} = 2.32635$).

224 The equation published by K. Stange (5) is:

225
$$y = a + bx + k_T s_{y.x}$$

with

226
$$k_T = \frac{\sqrt{(2n - 4)}}{(2n - 4)^* - u_{1-\alpha}^2} \left[\sqrt{(2n - 4)^*} u_{1-\gamma} + u_{1-\alpha} W_n \right]$$

227
$$W_n = \sqrt{u_{1-\gamma}^2 + [(2n - 4)^* - u_{1-\alpha}^2] \left[\frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}} \right]}$$

228
$$S_{xx} = \sum x_i^2 - \frac{1}{n} (\sum x_i)^2$$

229
$$s_{y.x} = \text{residual error } ()^* = (2n - 5), \text{ according to Graf et al. (6)}$$

230 A revised version of the Stange equation (using the term $(2n-5)$ instead of $(2n-4)$ in the three
 231 parentheses marked above by an asterisk) was published by Graf et al. in 1987 (6). The use of this
 232 equation results in slightly higher tolerance limits. According to Stange (5) the equation is valid for
 233 $n \approx 10$, whereas Graf et al. (6) restrict validity to $n \approx 20$.

234 A listing of data comparing the results of both equations to the results of the FDA procedure can be
 235 found in Annex B1 of this paper.

236 Remark: For reasons discussed below (see Section 6.3) the selection of the 95% tolerance limit with
 237 95% confidence is preferred.

238 **4.2. Possible alternative approach**

239 The statistical approach should be used whenever a data set fulfils the minimum requirements for a
240 statistical analysis. The statistical significance levels given in this guideline should be considered as
241 recommendations, not as strict rules, in that any violation of regression assumptions would not
242 automatically trigger use of an alternative approach. A decision to not use a statistical approach should
243 always be scientifically justified and based on statistical expert judgement.

244 The following statistical tests are referred to: F-test, Cochran test, Bartlett test, Shapiro-Wilk test, the
245 most critical of which is the lack-of-fit test (F-test). Significant deviations from a straight line cannot be
246 accepted for the model recommended in the guideline.

247 In many cases, the question of whether the statistical method can be used or not is dependent on the
248 number of time points with a sufficient number of observations above the LOQ; the validation of the
249 LOQ is therefore pivotal in this regard. The statistical method could probably be used in more
250 situations where a lower LOQ is demonstrated.

251 Whenever data available do not permit the use of the statistical model, an alternative approach has to
252 be considered in order to determine appropriate withdrawal periods.

253 The approach depends on many parameters such as sample size, number, frequency and choice of
254 slaughter timepoints, variability of the data, and analytical factors (e.g. Limit of Detection (LOD)).

255 One concept is the establishment of the withdrawal period at the time point where the concentrations
256 of residues in all tissues for all animals are at or below the respective MRLs¹. However, when one has
257 determined that time point, the estimation of a safety span should be considered in order to
258 compensate for the uncertainties mentioned above.

259 The value of a safety span depends on various, not easy to specify, factors which are decided by the
260 study design, the quality of the data and finally by the pharmacokinetic properties of the active
261 substance(s). As a result, an overall recommendation cannot be provided. An approximate guide for a
262 safety span is likely to be a value of 10 - 30% of the time point when all marker (relevant) residues
263 are at or below the MRL¹. Alternatively, a safety span might be calculated from the terminal tissue
264 depletion half-life, possibly a value of 1-3 times $t_{1/2}$.

265 **Examples of how certain factors might influence the size of the safety span:**

- 266 • If, at the first time point at which residues are below the MRL¹, all values are below the LOQ, then
267 a safety span of 10% may be acceptable.
- 268 • If there are long gaps between time points and if residue levels are already close to the MRL at the
269 timepoint before the one at which they actually fall below the MRL¹, then a safety span of 10%
270 may be appropriate.
- 271 • If there is high variability between animals at each timepoint then a safety span of 30% may be
272 appropriate.
- 273 • The proximity of the residue value to the MRL¹ should be taken into account and a safety span at
274 the higher end of the standard range (i.e., a safety span of 30%) considered in those cases where
275 the residue finding is at the MRL.
- 276 • If the first timepoint at which all residues are <MRL¹ is less than 10 days, and the underlying data
277 show a high degree of variability, then a longer safety span (>30%) should be considered (16).

278 **4.3. Injection site residues**

279 When considering the establishment of withdrawal periods for parenterally administered drugs, it is
280 important to take into account the residues of the intramuscular (IM) or subcutaneous (SC) injection
281 site. As a standard approach, the injection site residues are compared to the muscle MRL or Injection
282 Site Residue Reference Value (ISRRV) if available. The guideline on injection site residues
283 (EMA/CVMP/542/03-FINAL) (12) specifically addresses this point.

284 **4.4. MRL or other reference value**

285 The withdrawal period ensures that the foodstuffs from food producing animals do not contain residues
286 in quantities harmful to public health. As a default the establishment of the withdrawal period is at the
287 time point where the concentrations of residues in all tissues for all animals are at or below the
288 respective MRLs as laid down in Commission Regulation (EU) No 37/2010. If no (numerical) MRLs are
289 available other reference values may be used.

290 For compounds for which it was not necessary to establish a MRL (substances with are classified as 'No
291 MRL required' in Table 1 of the Annex to Commission Regulation (EU) No 37/2010) it **should not** be
292 interpreted as automatically implying that no withdrawal period is necessary.

293 If there is any indication that the amount of drug derived residues in an edible portion may exceed the
294 ADI or other reference value², a withdrawal period has to be set. The respective edible portion should
295 include the injection site muscle for substances to be injected intramuscularly or subcutaneously.

296 Since no MRLs are set for such compounds, the withdrawal period has to be estimated on the basis of
297 the ADI. An ADI based assessment of residues should cover the most relevant endpoint. Depending on
298 the type of ADI, residues of concern may be either the total drug related residues or the
299 toxicologically, pharmacologically and/or microbiologically active fraction of the total residues (the
300 general principles can be found in the guideline on injection site residues (12)).

301 For compounds which may cause injection site residues with potential harmful (e.g. pharmacological)
302 effects, it may be necessary to establish a precautionary withdrawal period even when an ADI has not
303 been set (e.g. in the case of hormones the naturally occurring levels in tissues should be used as the
304 starting point for the determination of a withdrawal period). In addition, other reference values may be
305 used, such as daily intake values for vitamins or other food-additives, set by EFSA.

306 **4.5 Generic products**

307 When the formulation (active and inactive ingredients), the dose schedule, the route(s) of
308 administration and the target species of a specific generic product, are identical to a currently
309 approved product (i.e., the reference product), or it has been adequately justified that any differences
310 in formulation are so minor such that they will not impact on residue depletion, then the withdrawal
311 period of the latter can be used for the former.

312 When a generic product is accepted as bioequivalent to a reference product, then it can be assumed
313 that residue depletion from edible tissues (muscle, fat, liver and kidney) will be comparable for both
314 products. In this case, assuming that the generic product will be used under the same conditions as
315 the reference product (target population, posology, route of administration, etc), the withdrawal period
316 of the reference product can typically be applied to the generic product.

² The ADI should preferably be used. However, depending on the particular substance this could be another health based guidance value like the 'normal intake via food' in case of substances being part of the human diet.

317 However, in the case of products administered subcutaneously or intramuscularly, small differences in
318 composition may have significant effects on injection site depletion which may not be detected in the
319 standard blood level bioequivalence studies. See also 'Bioequivalence GL, section 4.4' (7). Therefore,
320 for such formulations, in addition to the demonstration of bioequivalence equivalent depletion of
321 residues from the injection site should be demonstrated, in order that the withdrawal period
322 established for the reference product can be adopted. However, if it is demonstrated that the depletion
323 is slower at the injection site, resulting in a longer withdrawal period than that established for the
324 reference product, this longer withdrawal period should be taken as the overall withdrawal period. If it
325 is demonstrated that the depletion of the generic product is faster at the injection site when compared
326 to the reference product, then the established withdrawal period for the reference product will still be
327 adopted. This is because it is not known whether the other edible tissues may be withdrawal period
328 determining. See also the GL on injection site residues (12).

329 For further dosage forms that are able to leave local residues (e.g., topically applied products) residue
330 data from the site of administration would also be required. Samples of relevant tissues (e.g. muscle,
331 subcutaneous fat or skin/fat from the application site) should be analysed.

332 For applications under Article 19 (generic 'hybrid'), where a change of the target species and/or the
333 route of administration is requested, information on tissue residue depletion is considered to be
334 necessary. Changes in the dose (or frequency of dosing) may also require residue depletion data. Any
335 justification for the absence of data should be based on scientific argumentation.

336 Remark: For experimental design of blood level bioequivalence studies the bioequivalence guideline
337 provided by the CVMP (7) should be taken into account.

338 For specific problems concerning milk, see the CVMP Guideline on determination of withdrawal periods
339 for milk (EMA/CVMP/SWP/735418/2012) (18).

340 5. Example for the statistical analysis of residue data

341 Data constructed from an empirical residue depletion study on cattle treated subcutaneously with a
 342 veterinary drug were used to demonstrate the applicability of the statistical model for the estimation of
 343 withdrawal periods. The residue data for the marker residue in the target tissues liver and fat are listed
 344 in Table 1 (see Annex A). An ADI of 35 µg per day for a 60 kg person has been assumed for the total
 345 residue. The MRLs for the marker residue have been set at 30 µg/kg and 20 µg/kg for liver and fat,
 346 respectively (no MRL was set for muscle).³

347 **Calculation procedure**

348 **Step 1:** Inspection of the data (listed in Table 1, Annex A)

349 Data below the LOQ (i.e., 2 µg/kg) were set to one-half of the LOQ (i.e., 1.0 µg/kg). See Annex D.

350 For fat, the day 35 samples were excluded from calculation because of too many values below the LOQ
 351 (10 of 12 observations). Data for liver on day 35 were not available.

352 **Step 2:** Calculation of the linear regression parameters of the log_e-transformed data

353 **Table 2:** Linear regression parameters

Parameter	Liver	Fat
Number of values *	n = 48	n = 48
Intercept	a = 5.64 ± 0.35	a = 5.84 ± 0.36
Slope	b = - 0.16 ± 0.02	b = - 0.17 ± 0.02
Correlation coefficient	r = - 0.7927	r = - 0.8026
Residual error	s _{y.x} = 0.9930	s _{y.x} = 1.0258

354 * excluded data: day 35 for fat (day 35 for liver: not assayed)

355 **Step 3:** Visual inspection of the regression line

356 Both the regression line for liver and the regression line for fat passed through all slaughter groups. No
 357 time points have to be excluded at the end or at the beginning of the line (see Figs. 3 and 4).

358 **Step 4:** Homogeneity of variances

359 Due to the amount of data given per group and due to the equal group sizes, it was possible to use all
 360 three tests discussed above. The equations and percentage points have been published in L. Sachs (3).
 361 The results of the tests are summarized in the Tables 3-5.

362 **Table 3:** Bartlett's test

Tissue	Test value	Degrees of freedom	Probability	Significance
liver	$\hat{\chi}^2 = 4.24$	df = 3	P > 0.05	n.s.
fat	$\hat{\chi}^2 = 5.95$	df = 3	P > 0.05	n.s.

363 n.s.: differences are not significant

³ When this guideline was initially developed there was greater flexibility on whether MRLs should be established in all four standard tissues. It is now standard practice to always establish an MRL in muscle.

364 **Table 4:** Cochran's test

Tissue	Test value	Degrees of freedom	Probability	Significance
liver	\hat{G} max= 0.343	df ₁ = 11 df ₂ = 4	P > 0.05	n.s.
fat	\hat{G} max= 0.442	df ₁ = 11 df ₂ = 4	P > 0.05	n.s.

365 n.s.: differences are not significant

366 **Table 5:** Hartley's test

Tissue	Test value	Degrees of freedom	Probability	Significance
liver	\hat{F} max=3.46	df ₁ = 4 df ₂ = 11	P>0.05	n.s.
fat	\hat{F} max=4.68	df ₁ = 4 df ₂ = 11	P>0.05	n.s.

367 n.s.: differences are not significant

368 Conclusion: The variances of the log_e-transformed data at each time point are homogeneous.

369 **Step 5:** Analysis of variances (showing lack of fit) according to L. Sachs (3)

370 The ratio

$$\hat{F} = \frac{\text{MS between group means and the regression line}}{\text{MS within groups}}$$

371 was calculated and compared to the 5% percentage point of the F-distribution. Generally, a significant
372 ratio indicates that the log_e-linear model appears to be inadequate.

373 **Table 6:** ANOVA table for liver

Source of variation	Degrees of freedom	Sum of square (SS)	Mean square (MS=SS/df)
Between group means and the regression line	2	0.784	0.3919
Within groups (departure of y-values from their group mean)	44	44.573	1.0130
\hat{F} (test) = 0.3869 (df ₁ = 2, df ₂ = 44) P>0.05 n.s.			

374 n.s.: no significant deviation from linearity

375 **Table 7:** ANOVA table for fat

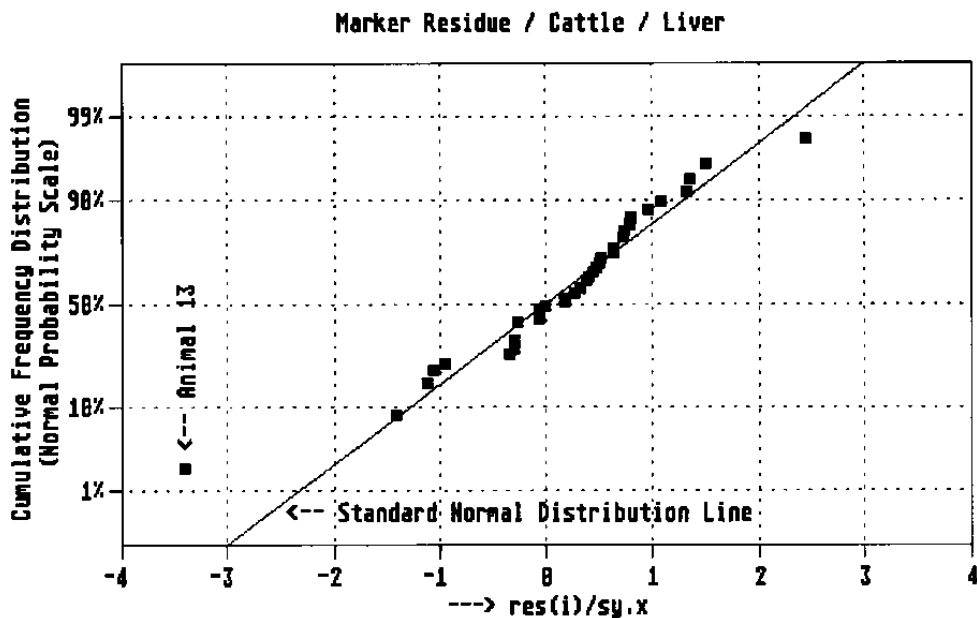
Source of variation	Degrees of freedom	Sum of square (SS)	Mean square (MS= SS/df)
Between group means and the regression line	2	6.240	3.1199
Within groups (departure of y-values from their group mean)	44	42.165	0.9583
\hat{F} (test) = 3.2557 (df ₁ = 2, df ₂ = 44) 0.05 > P > 0.025 n.s. *			

376 * Potential deviation from linearity emerges.

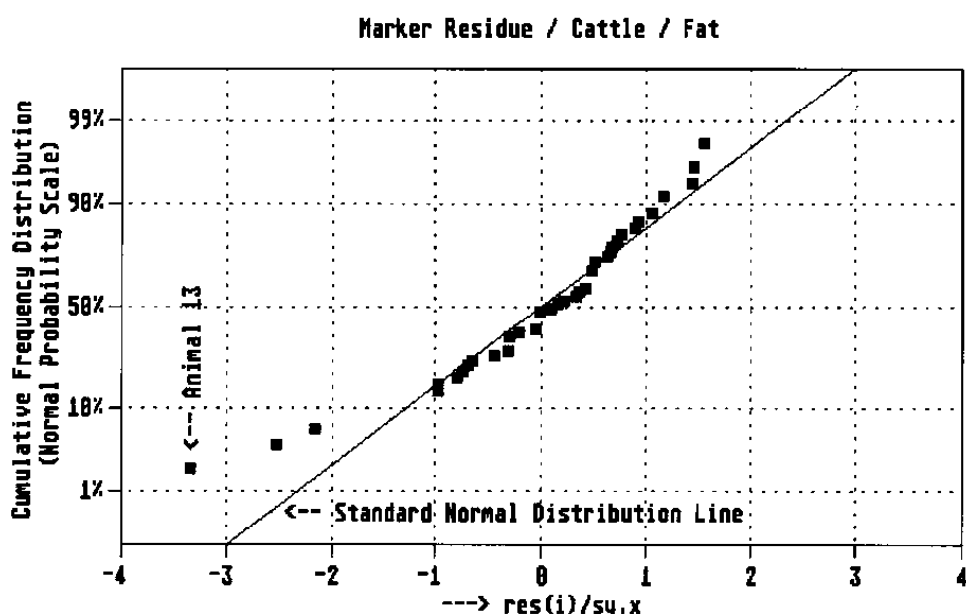
377 Conclusion: The assumption of linearity of the log_e-transformed data versus time can be upheld for
 378 liver. In the case of fat, a potential deviation from linearity emerges. A critical re-inspection of the
 379 plotted data (Fig. 4) suggests that day 7 may possibly belong to an earlier phase of residue depletion.
 380 Excluding day 7 from calculation might therefore be taken into account. This approach was not
 381 followed up here because the linearity assumption was not seriously violated.

382 **Step 6:** Calculation of residuals and plot of cumulative frequency distribution according to the
 383 recommendation of the FDA 1983 (2).

384 The plots for the ordered residuals (standardised by the residual error s_{y.x}) versus their cumulative
 385 frequency on a normal probability scale are shown in Figure 1 (liver) and Figure 2 (fat).



386 **Figure 1:** Cumulative frequency distribution of residuals for liver
 387



388
389 **Figure 2:** Cumulative frequency distribution of residuals for fat

390 Conclusion: Fat shows a marked departure from the straight line at the negative end of this line. The
391 value which deviates most belongs to the animal numbered 13. The plot for liver also shows that the
392 sample of animal 13 deviates from the standard normal distribution line. This is a possible indication
393 that the residue data of animal 13 tend to be outliers.

394 In order to verify the results of the residual plot, the Shapiro-Wilk test for normality was performed
395 according to G. B. Wetherill (4). The coefficients required for calculation of the test value \hat{W} were
396 taken from Table C7 (see (4), pp. 378 - 379) and compared to the percentage points for the
397 Shapiro-Wilk test, published in Table C8 (see (4) p. 380). The assumption of a normal distribution (in
398 this case a normal distribution of the errors) holds as long as the test value \hat{W} exceeds the 10%
399 percentage point for the given sample size.

400 **Table 8:** Shapiro-Wilk test

Tissue	Test value	n	Probability	Significance
Liver	$\hat{W} = 0.960$	48	$P > 0.10$	n.s.
Fat	$\hat{W} = 0.922$	48	$P < 0.01$	*
Fat (animal 13 excl.)	$\hat{W} = 0.955$	47	$P > 0.10$	n.s.

401 n.s.: No significant deviation from normality; * Significant deviation from normality

402 Conclusion: No deviation from normality could be observed for liver. For fat, there was a significant
403 deviation of the errors from normality when testing all fat samples. As discussed above, sample 13
404 may possibly be seen as outlier. Excluding animal 13 from calculation for fat, the distribution returned
405 to normality.

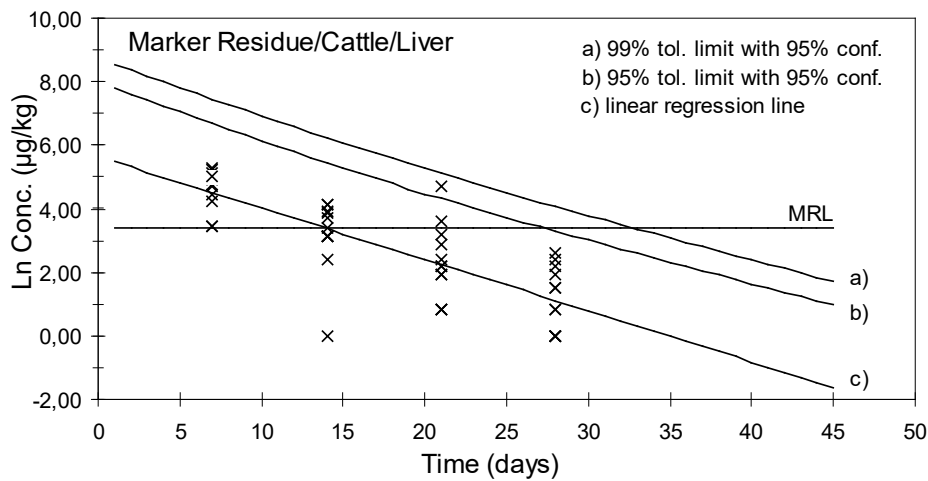
406 **Step 7:** Calculation of the one-sided 95% and 99% upper tolerance limits (both with a 95% confidence
407 level) according to K. Stange (5):

408 The numerical values are summarised in Table 9 and 10. Plots of withdrawal period calculations for
409 liver and fat are shown in Figures 3 and 4.

410 **Table 9:** Results for liver (full data set, including animal 13)

Days post dose	Statistical tolerance limits with 95% confidence	
	95% Tolerance limit (µg/kg)	99% Tolerance limit (µg/kg)
26	35.7	77.9
27	30.9	67.4
28	26.8*	58.3
29	23.3	50.5
30	20.3	43.7
31	17.6	38.0
32	15.3	33.0
33	13.4	28.7*

411 * first concentration below the MRL (30 µg/kg) for liver

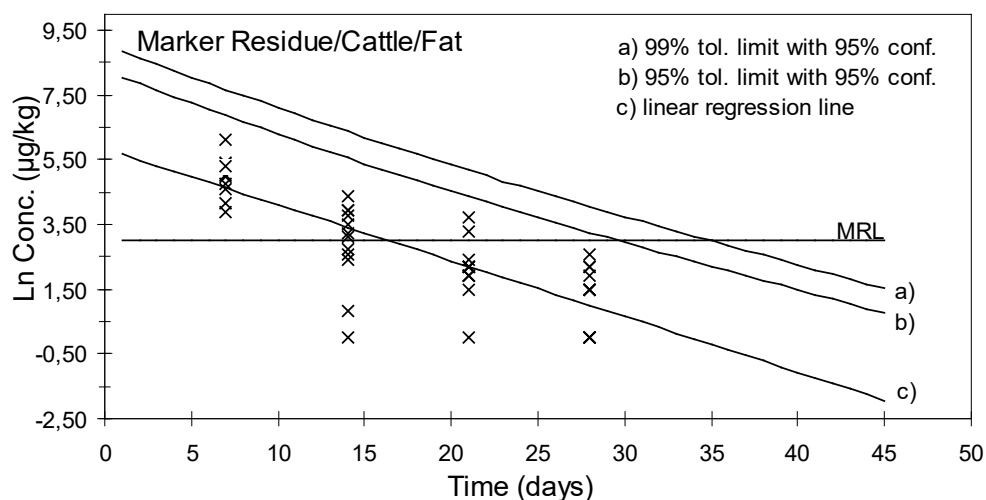


412 **Figure 3:** Plot of withdrawal period calculation for liver

414 **Table 10:** Results for fat (full data set, including animal 13)

Days post dose	Statistical tolerance limits with 95% confidence	
	95% Tolerance limit (µg/kg)	99% Tolerance limit (µg/kg)
26	35.1	78.6
27	30.1	67.2
28	25.8	57.5
29	22.2	49.3
30	19.1*	42.3
31	16.4	36.3
32	14.2	31.2
33	12.2	26.8
34	10.5	23.1
35	9.1	19.9*
36		17.2

415 * first concentration below the MRL (20 µg/kg) for fat



416
417 **Figure 4:** Plot of withdrawal period calculation for fat

418 The MRLs for the target tissues liver and fat are 30 µg/kg and 20 µg/kg, respectively. The time points
419 when the residues in fat and liver dropped below their MRLs are summarised in Table 11.

420 **Table 11:** Withdrawal periods obtained for the full data set, including animal 13:

Withdrawal times obtained from	Liver	Fat
95% tolerance limit (95% conf.)	28 days	30 days
99% tolerance limit (95% conf.)	33 days	35 days

421 **Re-evaluation of data excluding animal 13**

422 **Table 12:** Test results (excluding animal 13)

	Liver	Fat
Bartlett's test	0.05 > P > 0.025	P > 0.05
Cochran's test	P > 0.05	P > 0.05
Lack of fit test	P > 0.05	P > 0.05
Shapiro-Wilk test	P > 0.10	P > 0.10

423 The regression assumptions are not seriously violated.

424 Taking into account MRLs of 30 µg/kg and 20 µg/kg for liver and fat, respectively, the withdrawal
425 times listed below were estimated:

426 **Table 13:** Withdrawal periods obtained (excluding animal 13)

Withdrawal times obtained from	Liver	Fat
95% tolerance limit (95% conf.)	26 days	29 days
99% tolerance limit (95% conf.)	31 days	33 days

427 **Step 8:** Estimation of the withdrawal period for the injection site (using an alternative approach)

428 In the example discussed here, the withdrawal periods estimated in Step 7 were based on the MRLs for
429 the target tissues fat and liver. An MRL for muscle was not established for the drug under review, and
430 so the residues at the injection site cannot be compared to the muscle MRL. Therefore, the withdrawal

431 period for injection site residues has to be calculated on the basis of the ADI being 35 µg (per day for a
 432 60 kg person) for the total residue (listed in Table 1, Annex A).

433 It has to be shown that the ADI is not exceeded when the usual food package (0.5 kg) includes 0.3 kg
 434 injection site (instead of 'normal' muscle). In some cases, the CVMP will have set an Injection Site
 435 Residue Reference Value (ISRRV), which can be used as a surrogate for the muscle MRL for injection
 436 sites only (13).

437 For this purpose, marker residue concentrations from Table 1 were converted to total residues
 438 according to the average ratios marker/total (0.3 for liver, fat and kidney, and 0.6 for injection site
 439 muscle), determined in a total residue depletion study. The daily intake of the total residue from each
 440 tissue type was calculated using the standard food consumption figures (300 g injection site, 100 g
 441 liver, 50 g kidney and 50 g fat). In other words, the total residue in the 0.5 kg food package was
 442 determined for each slaughter day by using the following equation:

$$RI = (cL \times FL / RL) + (cK \times FK / RK) + (cF \times FF / RF) + (cM \times FM / RM)$$

RI = residue intake (µg)
 c = concentration of the marker residue (µg/kg)
 F = food consumption figures (0.3 kg muscle, 0.1 kg liver, 0.05 kg kidney, 0.05 kg fat)
 R = ratio marker residue vs. total residue
 (to be applied when the ADI refers to the total residues)

Indices L, K, F, M = liver, kidney, fat and muscle (here injection site)

443 Day 28 was not excluded from calculation even though there were only 2 values (out of 12) above the
 444 LOQ for the injection site. However, day 35 was excluded because data for liver and kidney were not
 445 available. Data below the LOQ were set to one-half of the LOQ. The results of this calculation are listed
 446 in the last column of Table 1 (Annex A).

447 As residue depletion from the injection site was rather erratic (high animal to animal variation) the
 448 statistical requirements for regression analysis were not met by these data for the daily dietary residue
 449 intake. The data revealed a significant deviation from normality and the homogeneity of variances was
 450 slightly violated.

451 **Table 14:** Test results

Edible portion		
Bartlett's test	0.05 > P > 0.025	*
Lack-of-fit test	P > 0.05	n.s.
Shapiro-Wilk test	0.05 > P > 0.02	**

- 452 n.s.: no significant deviation from linearity
 453 * potential non-homogeneity of variances
 454 ** significant deviation from normality

455 Furthermore, the tolerance limits crossed the ADI-line far after the time range when data for the total
 456 residue intake were available (95% tolerance limit: day 35 using Stange, 99% tolerance limit: day 42
 457 (Table 15).

458 **Table 15:** Data set of 48 animals, 12 per slaughter day, ADI = 35 µg/person, Upper 95 and 99%
 459 tolerance limits with 95% confidence

Days post dose	Upper 95% tolerance limit (µg)	Upper 99% tolerance limit (µg)
33	54.47	
34	43.18	
35	34.27*	155.46
36	27.22	122.97
37	21.64	97.37
38	17.22	77.18
39	13.71	61.23
40	10.92	48.62
41	8.71	38.63
42	6.95	30.71*
43	5.54	24.43
44		19.45

460 * first total residue level below the ADI (35 µg/person)

461 Since the time period between day 28 and day 35/42 was not covered by data and since the regression
 462 assumptions were not met, the statistical approach of setting a withdrawal period seemed to be
 463 inadequate.

464 Therefore, an alternative approach was applied:

465 Inspection of the data for the daily dietary residue intake (Table 1) showed that on day 28 the highest
 466 individual residue amount (calculated as 32.3 µg) was just below the ADI (35 µg/day). In order to
 467 account for the high variability of the residue data, especially the variability of the injection site data, a
 468 safety span has to be added to the depletion time of 28 days. A safety span of 7 days can be seen as
 469 appropriate. This safety span corresponds to 25% of the 28 day depletion time. The alternative
 470 approach would then result in a withdrawal period of 35 days.

471 On the whole, it should be noted here that any alternative approach is of course rather subjective and
 472 depends on the significance given to specific aspects of the information available.

473 Remark: The final withdrawal period has to be set in a way that the residues in all target tissues drop
 474 below their specific MRLs and ISRRVs, and, in addition, that the amount of residues in the edible
 475 portion drops below the ADI. This means that the longest withdrawal period has to be selected in order
 476 to be in full compliance with the MRLs, ISRRV and the ADI. In the example discussed here, the
 477 withdrawal times obtained from the statistical 95% tolerance limits for fat and liver residues were 30
 478 and 28 days, respectively. However, the withdrawal period of 35 days derived when comparing the
 479 calculated daily total residue intake including the injection site to the ADI would determine the
 480 conclusive withdrawal period.

481 **6. Discussion on the regression analysis**

482 Data on residues in cattle liver and fat (constructed from real empirical data) were analysed by using a
483 set of basic statistical tests in order to prove that linear regression analysis is an appropriate model for
484 estimation of withdrawal periods. It was shown that assumptions on which the regression analysis is
485 based could in principle be upheld when tested on these data. Only in the case of fat was the normality
486 assumption violated (Shapiro-Wilk test). However, excluding one sample (which was suspected to be
487 an outlier) the distribution of the fat data returned to a normal distribution.

488 The statistical procedure applied to these data revealed a number of problems associated with
489 estimating withdrawal periods:

490 ***6.1. To what extent a departure from the regression assumptions may be*** 491 ***acceptable?***

492 The first general question is where to set the significance levels of the tests and to what extent a
493 departure from the regression assumptions may be acceptable. Second, should these assumptions
494 absolutely dictate whether the calculation model can be used or not?

495 In other words, one could be faced with a situation in which the data do not sufficiently satisfy the
496 statistical assumptions. In this situation one has to decide whether the calculation procedure should be
497 stopped, strictly according to the rules of statistics, or whether the calculation procedure may be
498 continued under more investigative considerations. As long as the regression assumptions are not
499 seriously violated, the tolerance limits might be used as a reference for an appropriate safety span. In
500 our view, this pragmatic approach will at least provide rough orientation for a potential withdrawal
501 period.

502 ***6.2. Withdrawal periods should be set by interpolation and not by*** 503 ***extrapolation***

504 The tolerance limits may cross the MRL (or ADI-line or other reference value-line) far after the time
505 range when data were available, which may not be acceptable.

506 Again, it has to be considered whether the treatment of the data should be done strictly according to
507 the rules of statistics, or whether an extrapolation can be allowed. In our view, a slight extrapolation
508 may be possible because the depletion kinetic is assumed to be linear with time (log_e-linearity).
509 Furthermore, tolerance limits are described by hyperbolic curves. Accordingly, the withdrawal period is
510 unlikely to be underestimated when derived by slight extrapolation.

511 Extrapolation has to be considered with care, when there is indication (e.g. from pharmacokinetic
512 parameters) of a slower final depletion kinetic. Extrapolation far removed from the range of observed
513 data should be avoided. In cases when a withdrawal period can only be derived by a significant
514 extrapolation, further residues data must be provided to confirm the suitability of the derived
515 withdrawal period.

516 ***6.3. Should the 95% or the 99% tolerance limit be applied?***

517 Calculations were performed with both the 95% and the 99% one-sided upper tolerance limits (each
518 with a 95% confidence level). Taking into account the MRLs proposed (in the example above) for the
519 target tissues liver and fat, and using the full data set (including animal 13), withdrawal periods of 28
520 and 30 days (95% tolerance limit) and 33 and 35 days (99% tolerance limit) were calculated for
521 respectively liver and fat. These withdrawal periods were derived by a minimal extrapolation at the

522 95% tolerance limit for fat and by increased extrapolation at the 99% tolerance limit for both fat and
523 liver.

524 When applying the 99% tolerance limit, one is often confronted with the problem of extreme
525 extrapolation which may result in inadequate withdrawal periods. The 95% tolerance limit in some
526 cases may diminish the extrapolation problem and is therefore expected to provide more realistic
527 withdrawal periods.

528 For the reasons above the more pragmatic approach - the selection of the 95% tolerance limit for
529 setting withdrawal periods - is preferred.

530 **6.4. Dealing with 'less than' values**

531 Generally, these data cannot be excluded from calculation a priori, since they are due to real
532 observations concerning the depletion kinetics. Setting these data to one-half of the LOQ should be
533 considered. 'Less than' values may also be estimated by special procedures (10,11). See also Annex D.

534 If, however, the majority of data from one slaughter day are below the LOQ, the whole time point
535 could be excluded. This should be the case in particular when the time point in question is a late one
536 which is well off the regression line defined by the other data.

537 **6.5. Dealing with obvious outliers**

538 For example, could there be any justification to reject the residue data measured for animal 13 of the
539 present data set?

540 Inspection of the residue data indicated that animal 13 may possibly be an outlier. The residues in all
541 the tissues of this animal (including the injection site) were at or below the LOQ at a relatively early
542 time point post dose (day 14, see Table 1). As discussed earlier, the regression assumptions were
543 violated for fat when the full data set was evaluated. Exclusion of animal 13 gave a more reliable basis
544 for the statistical estimation of the withdrawal period.

545 Usually, due to the limited number of animals and due to the biological animal-to-animal variability,
546 exclusion of values has to be considered with great care. A formal test for outliers has not been
547 recommended in this guideline. It may occur, however, that there is a clear reasoning for an exclusion,
548 but removal of data points defined as statistical outliers should only be accepted if there is a causal
549 justification (e.g. dosing error, sampling/analytical error, any of which should be properly
550 documented).

551 **6.6. Combining data sets**

552 The benefits and drawbacks of combining studies are discussed in a general section of the 'Guideline
553 on statistical principles for clinical trials for veterinary medicinal products (pharmaceuticals)'
554 (EMA/CVMP/EWP/81976/2010)(19). Generally, such a meta-analysis could have advantages as well as
555 disadvantages: On the one hand, there could be an increase in precision and reliability of results, and
556 sacrificing animals could be reduced. On the other hand, problems might arise if the study
557 characteristics are too different, and if low-quality data are combined with high-quality data, the
558 results might be less reliable than those of an analysis of the high-quality data alone. Thus,
559 combination of data sets might be considered appropriate when the underlying studies are 'similar' and
560 of 'similar quality' (e.g., similar study design, same breeds, animal weight range, dosing, comparable
561 analytical methods etc.). It would only be appropriate to derive withdrawal periods using the statistical
562 approach, analysing the combined data sets, if the results of the two (or more) studies had been
563 shown to be statistically comparable (for example not statistically different from each other in respect

564 to key parameters such as residual errors of the populations; slope and starting concentrations (C0) of
565 residues. Differences in these and other parameters might indicate differences due to subtle (i.e. not
566 easy to notice) differences in the study designs or other influencing factors.

567 **6.7. The possibility of overriding one study with another**

568 Whether to use or discount a study should depend solely on the quality and validity of the data and
569 not, for example, on the age of the study. Expert judgement is needed, however, to determine
570 whether an 'old' study still reflects contemporary good veterinary and analytical practice (are the
571 animal breeds, treatment and housing conditions and analytical techniques still 'state of the art' and
572 representative of current practices, can these differences have any significant impact on the results?).
573 If old data are considered valid in respect to relevant study design and quality criteria then they should
574 not be discounted in favour of more recently generated residue data.

575 **7. References**

- 576 1. FDA, General Principles for Evaluating the Safety of Compounds Used in Food-Producing Animals,
577 1994
- 578 2. FDA, General Principles for Evaluating the Safety of Compounds Used in Food-Producing Animals,
579 1983
- 580 3. Lothar Sachs, Angewandte Statistik, 7th Ed., Springer Verlag Berlin, Heidelberg, New York,
581 London, Paris, Tokio, 1992
- 582 4. G. Barrie Wetherill, Intermediate Statistical Methods, Chapman and Hall, London, New York, 1981
- 583 5. Kurt Stange, Angewandte Statistik, Vol. II, pp. 141-143, Springer Verlag, Berlin, Heidelberg, New
584 York, 1971
- 585 6. U. Graf, H.J. Henning, K. Stange, P.T. Wilrich, Formeln und Tabellen der angewandten
586 mathematischen Statistik, 3rd ed., Springer Verlag, Berlin, Heidelberg, New York, London, Paris,
587 Tokio, 1987
- 588 7. CVMP, Guideline on the conduct of bioequivalence studies for veterinary medicinal products,
589 EMA/CVMP/016/2000-Rev.4, July 2021
- 590 8. D.B. Owen, Handbook of Statistical Tables, Addison-Wesley Publishing Company, Reading,
591 Massachusetts, 1962
- 592 9. John Mandel, The Statistical Analysis of Experimental Data, Interscience Publ., J. Wiley & Sons,
593 New York, London, Sydney 1964
- 594 10. Helsel, D.R., Less than obvious, *Environ. Sci. Technol.*, Vol 24, No. 12, pp. 1766-1774, 1990
- 595 11. Newman, M.C., Dixon, P.M., Looney, B.B., Pinder, J.E., Estimating mean and variance for
596 environmental samples with below detection limit observations, *Water Resources Bulletin*, Vol 25,
597 No. 4, pp. 905-916, 1989.
- 598 12. CVMP, Guideline on injection site residues, EMEA/CVMP/542/03-FINAL, April 2005
- 599 13. COMMISSION REGULATION (EU) 2018/782 of 29 May 2018 establishing the methodological
600 principles for the risk assessment and risk management recommendations referred to in Regulation
601 (EC) No 470/2009
- 602 14. VICH GL48: Studies to evaluate the metabolism and residue kinetics of veterinary drugs in food-
603 producing animals: marker residue depletion studies to establish product withdrawal periods
604 (EMA/CVMP/VICH/463199/2009, 12 February 2015)
- 605 15. VICH GL49: Studies to evaluate the metabolism and residue kinetics of veterinary drugs in
606 food-producing animals: validation of analytical methods used in residue depletion studies,
607 (EMA/CVMP/VICH/463202/2009, 12 February 2015)
- 608 16. Schefferlie & Hekman, The size of the safety span for pre-slaughter withdrawal periods; *Journal of*
609 *Veterinary Pharmacology and Therapeutics* Volume 32, Issue Supplement s1, 17 JUL 2009
- 610 17. Directive 2004/10/EC of the European Parliament and of the Council of 11 February 2004 on the
611 harmonisation of laws, regulations and administrative provisions relating to the application of the
612 principles of good laboratory practice and the verification of their applications for tests on chemical
613 substances
- 614 18. Guideline on Determination of Withdrawal Periods for Milk (EMA/CVMP/SWP/735418/2012) (2000)

- 615 19. Guideline on statistical principles for veterinary clinical trials (EMA/CVMP/EWP/81976/2010)
- 616 20. VICH GL57 on Studies to evaluate the metabolism and residue kinetics of veterinary drugs in food-
617 producing species: marker residue depletion studies to establish product withdrawal periods in
618 aquatic species (EMA/CVMP/VICH/517152/2013) (2018)
- 619 21. Regulation (EC) No 470/2009 of the European Parliament and of the Council of 6 May 2009 laying
620 down Community procedures for the establishment of residue limits of pharmacologically active
621 substances in foodstuffs of animal origin, repealing Council Regulation (EEC) No 2377/90 and
622 amending Directive 2001/82/EC of the European Parliament and of the Council and Regulation (EC)
623 No 726/2004 of the European Parliament and of the Council.
- 624 22. Regulation (EU) 2019/6 of the European Parliament and of the Council of 11 December 2018 on
625 veterinary medicinal products and repealing Directive 2001/82/EC
- 626 23. Commission Delegated Regulation (EU) 2021/805 of 8 March 2021 amending Annex II to
627 Regulation (EU) 2019/6 of the European Parliament and of the Council
- 628

629 **Annex A**

630 **Table 1:** Individual results for the marker residue in cattle and calculated daily total residue intake
 631 (Data constructed from a real empirical data set)

Animal number	Days post dose	Liver	Fat	Kidney	Muscle	Inj. site	Daily intake*
		(µg/kg)					(µg)
1	7	85.5	96.8	27.0	11.3	123.8	111.0
2	7	141.8	225.0	29.3	11.3	74250.0	37214.7
3	7	198.0	213.8	47.3	15.8	6750.0	3484.5
4	7	31.5	48.3	18.0	4.5	n.a.	-
5	7	119.3	119.3	38.3	9.0	18000.0	9066.0
6	7	108.0	204.8	38.3	18.0	922.5	537.8
7	7	171.0	157.5	6.8	15.8	19125.0	9646.9
8	7	31.5	450.0	11.3	2.3	24.8	99.8
9	7	189.0	65.3	13.5	20.3	4050.0	2101.1
10	7	67.5	195.8	18.0	6.8	495.0	305.6
11	7	135.0	148.5	49.5	20.3	65.3	110.7
12	7	150.8	202.5	60.8	20.3	4500.0	2344.2
13	14	<2.0	<2.0	<2.0	<2.0	2.3	1.8
14	14	22.5	11.3	6.8	2.3	180.0	100.5
15	14	60.8	78.8	20.3	11.3	85.5	79.5
16	14	60.8	51.8	9.0	4.5	2025.0	1042.9
17	14	47.3	33.8	13.5	4.5	121.5	84.4
18	14	22.5	24.8	2.3	2.3	13.5	18.8
19	14	11.3	2.3	2.3	<2.0	<2.0	5.0
20	14	22.5	15.8	13.5	4.5	585.0	304.9
21	14	49.5	51.8	4.5	6.8	49500.0	24775.9
22	14	22.5	13.5	4.5	2.3	105.8	63.6
23	14	40.5	22.5	9.0	4.5	20.3	28.9
24	14	29.3	42.8	18.0	6.8	31.5	35.7
25	21	36.0	27.0	11.3	6.8	33.8	35.3
26	21	9.0	9.0	2.3	2.3	4.5	7.1
27	21	9.0	6.8	2.3	<2.0	<2.0	5.0
28	21	6.8	6.8	2.3	<2.0	<2.0	4.3
29	21	18.0	6.8	2.3	<2.0	<2.0	8.0
30	21	6.8	11.3	2.3	<2.0	<2.0	5.0
31	21	108.0	40.5	11.3	9.0	14850.0	7469.6
32	21	11.3	9.0	4.5	<2.0	11.3	11.7
33	21	2.3	4.5	2.3	<2.0	31.5	17.7
34	21	2.3	9.0	6.8	<2.0	<2.0	3.9
35	21	24.8	9.0	4.5	4.5	11.3	16.2
36	21	2.3	<2.0	<2.0	<2.0	<2.0	1.6
37	28	4.5	4.5	<2.0	<2.0	4.5	4.7
38	28	2.3	4.5	<2.0	<2.0	<2.0	2.2
39	28	11.3	9.0	2.3	<2.0	<2.0	6.2

Animal number	Days post dose	Liver	Fat	Kidney	Muscle	Inj. site	Daily intake*
40	28	9.0	6.8	2.3	<2.0	<2.0	5.0
41	28	<2.0	<2.0	<2.0	<2.0	<2.0	1.2
42	28	4.5	4.5	2.3	<2.0	<2.0	3.1
43	28	<2.0	<2.0	<2.0	<2.0	<2.0	1.2
44	28	<2.0	<2.0	<2.0	<2.0	<2.0	1.2
45	28	2.3	4.5	<2.0	<2.0	<2.0	2.2
46	28	6.8	9.0	2.3	<2.0	<2.0	4.7
47	28	13.5	13.5	4.5	2.0	49.5	32.3
48	28	<2.0	<2.0	<2.0	<2.0	<2.0	1.2
49	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
50	35	n.a.	4.5	n.a.	n.a.	<2.0	-
51	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
52	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
53	35	n.a.	4.5	n.a.	n.a.	4.5	-
54	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
55	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
56	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
57	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
58	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
59	35	n.a.	<2.0	n.a.	n.a.	<2.0	-
60	35	n.a.	<2.0	n.a.	n.a.	<2.0	-

632 * Amount of total residue calculated by using the ratios marker/total 0.3 for liver, fat, kidney and 0.6 for injection
633 site. The food consumption figures used were 100 g liver, 50 g fat, 50 g kidney and 300 g injection site. Values
634 below the LOQ were set to one-half of the LOQ.

635 n.a.: not assayed

636 LOQ⁴: 2 µg/kg

637 Results corrected for recoveries.

638

⁴ Although this was reported as the LOD in previous iterations of this guideline, it is more likely to be the Limit of Determination (Quantification) (LOQ), rather than the Limit of Detection (LOD).

639 **Annex B**

640 Comparison to the FDA approach:

641 In order to compare the results of the equations according to Stange (5) and Graf et al. (6) to the
642 results of the FDA procedure, three data sets out of the data set for liver from Table 1 (Annex A) were
643 tested:

- 644 1. The full data set for liver (n=48).
645 2. The last 5 data of each time point for liver (n=20).
646 3. The last 3 data of each time point for liver (n=12).

647 For all three data sets the regression assumptions were met. This can be seen from Table 15.

648 **Table 15:** Test results

Data set:	1	2	3
	(n=48)	(n=20)	(n=12)
Bartlett's test	p>0.05	p>0.05	p>0.05
Cochran's test	p>0.05	p>0.05	p>0.05
Lack of fit test	P>0.05	p>0.05	p>0.05
Shapiro-Wilk test	P>0.10	p>0.10	p>0.10

649 Remark: for all calculation procedures used here, values below the LOQ were set to one-half of the
650 LOQ

651 Calculation of the tolerance limits:

652 The tolerance limits according to Stange (5) and Graf et al. (6) were calculated as described earlier
653 (section 2).

654 The calculation using the non-central t-distribution was performed as recommended by the FDA (1,2):

- 655 • calculation of the non-centrality parameter d,
656 • calculation of the 95th percentile (designated k or t_0 of the non-central t-distribution by using the
657 inverse of the non-central t-distribution function),
658 • calculation of the tolerance limit according to the equation given in the FDA guideline.

659 Since the tolerance limits for the calculation of withdrawal periods require only 95% confidence, the
660 tables provided by Owen (8) can also be used. The 95th percentile of the non-central t-distribution for
661 the given non-centrality parameter d and the given degrees of freedom (df = n-2) can be calculated
662 by using the table on page 111 in conjunction with the interpolation procedure described on page 109
663 of the Owen handbook (8). Because of the very tight tabulation of values, the interpolated figures are
664 sufficiently exact. An additional advantage is that the table, as well as the interpolation procedure, can
665 easily be integrated in any calculation program.

666 **Results:**

667 1. Data set of 48 animals, 12 per slaughter day, MRL = 30 µg/kg

668 **Table 16:** Upper 95% tolerance limits with 95% confidence

Days post dose	Non-central t-distribution (µg/kg)	Stange (µg/kg)	Graf et al. (µg/kg)
25	41.60	41.26	41.82
26	36.00	35.70	36.18
27	31.20	30.93	31.35
28	27.07	26.83	27.20
29	23.51	23.30	23.62
30	20.45	20.25	20.53

669 **Table 17:** Upper 99% tolerance limits with 95% confidence

Days post dose	Non-central t-distribution (µg/kg)	Stange (µg/kg)	Graf et al. (µg/kg)
25	91.20	90.33	92.03
26	78.72	77.94	79.41
27	68.04	67.35	68.62
28	58.88	58.26	59.36
29	51.01	50.46	51.41
30	44.24	43.74	44.57
31	38.40	37.96	38.68
32	33.36	32.96	33.60
33	29.00	28.65	29.20

670 2. Data set of 20 animals, 5 per slaughter day, MRL = 30 µg/kg

671 **Table 18:** Upper 95% tolerance limits with 95% confidence

Days post dose	Non-central t-distribution (µg/kg)	Stange (µg/kg)	Graf et al. (µg/kg)
25	37.21	36.47	38.00
26	31.98	31.32	32.63
27	27.53	26.95	28.08
28	23.75	23.23	24.21
29	20.52	20.05	20.91
30	17.76	17.33	18.08

672 **Table 19:** Upper 99% tolerance limits with 95% confidence

Days post dose	Non-central t-distribution (µg/kg)	Stange (µg/kg)	Graf et al. (µg/kg)
25	82.57	80.70	85.42
26	70.69	69.02	73.07
27	60.63	59.15	62.63
28	52.10	50.78	53.77
29	44.83	43.66	46.24

Days post dose	Non-central t-distribution (µg/kg)	Stange (µg/kg)	Graf et al. (µg/kg)
30	38.64	37.59	39.83
31	33.35	32.41	34.35
32	28.82	27.98	29.66

673 3. Data set of 12 animals, 3 per slaughter day, MRL = 30 µg/kg

674 **Table 20:** Upper 95% tolerance limits with 95% confidence

Days post dose	Non-central t-distribution (µg/kg)	Stange (µg/kg)	Graf et,al. (µg/kg)
25	88.53	85.10	94.94
26	77.93	74.76	83.45
27	68.79	65.87	73.57
28	60.89	58.19	65.03
29	54.03	51.52	57.63
30	48.04	45.72	51.17
31	42.79	40.64	45.53
32	38.18	36.19	40.58
33	34.12	32.27	36.23
34	30.53	28.82	32.39
35	27.35	25.76	28.99

675 **Table 21:** Upper 99% tolerance limits with 95% confidence

Days post dose	Non-central t-distribution (µg/kg)	Stange (µg/kg)	Graf et al. (µg/kg)
25	240.37	230.00	267.87
26	210.33	200.88	234.02
27	184.56	175.92	205.01
28	162.38	154.44	180.06
29	143.20	135.91	158.52
30	126.57	119.86	139.87
31	112.09	105.91	123.67
32	99.45	93.75	109.54
33	88.39	83.13	97.20
34	78.67	73.83	86.39
35	70.13	65.66	76.89

676 **Table 22:** Withdrawal periods obtained

Data set:	n=48		n=20		n=12	
Tolerance limits*:	95%	99% (days)	95%	99% (days)	95%	99% (days)
Non-central t-distribution	28	33**	27	32**	35**	-***
Stange	28	33**	27	32**	34**	-***
Graf et al.	28	33**	27	32**	35**	-***

677 * with 95% confidence

678 ** more or less severe extrapolation

679 *** unacceptable extrapolation

680 **Discussion:**

681 Tables 16-21 show that all three methods of calculation gave similar results. When comparing the
 682 results of the procedure using the non-central t-distribution to the others, the tolerance limits
 683 calculated according to Graf et al (6) were somewhat higher, while those calculated according to
 684 Stange (5) were somewhat lower. The time points when the tolerance limits dropped below the MRL of
 685 30 µg/kg are listed in Table 22. As it can be seen in that case, only in one data set (n=12 data set) did
 686 a difference of one day appear. The results from Table 22 also show that the evaluation of small data
 687 sets (e.g., n = 12) could result in relatively long withdrawal periods.

688 To set withdrawal periods, all three methods of calculation can be considered to be appropriate and of
 689 equal value.

690 With a view to more practical considerations, we propose the procedure according to Stange (5). This
 691 approach is not confined to $n \approx 20$, as is the procedure according to Graf et al. (6) and is much easier
 692 to perform than the FDA procedure (1,2).

693

694 **Annex C**

695 Test of the Significance of the Quadratic Time Effect:

696 In order to test linearity, checking the significance of the quadratic time effect according to Mandel (9)
 697 can be done in advance as an appropriate supplementation to the lack of fit test. The question is,
 698 whether a quadratic fit is better than the linear fit.

699 The linear model is represented by the relation $y = a + bx$, the quadratic model by $y = a + bx + cx^2$.

700 Both equations have to be fitted by the method of least squares and the residual errors ($s_{y,x}$) have to
 701 be calculated (using the \log_e -transformed residue concentrations).

702 The question is then to determine whether the residual variance of the quadratic fit is significantly
 703 smaller than the residual variance of the linear fit. It should be noted, however, that this test only
 704 shows if one model is or is not significantly better than the other one, whereas both may be
 705 inadequate.

706 If there is a significant quadratic time effect which is due to the first time point, the next step is to
 707 remove the first time point and re-run the analysis.

708 Remark: A coefficient of the quadratic term equivalent to zero (in the statistical sense) is in accordance
 709 with the statement that the linear model is the better one. A statistically significant positive coefficient
 710 has to be seen as the most likely alternative model (biphasic elimination kinetic). A statistically
 711 significant negative coefficient of the quadratic term indicates that the maximum concentration in
 712 tissues has not been reached at early time points.

713 The test of significance gave the following results for the data for liver and fat from Table 1 (Annex A):

714 1. Liver

715 Coefficient c: 0.0017 ± 0.0029 (not significantly different from zero at $P = 0.05$)

716 Residual error (linear fit): 0.9930

717 Residual error (quadratic fit): 1.0004

718 **Table 26:** Analysis of variance for liver

	Number of parameters in model	Remaining degrees of freedom (df)	Sum of squares of residuals	Mean square (SS/df)
Linear fit:	2	48-2=46	SS _L =45.3569	MS _L = 0.9860
Quadratic fit:	3	48-3=45	SS _Q =45.0339	MS _Q =1.0008
Difference		1	SS _D =0.3230	MS _D = 0.3230

719 $MS_D = 0.3230$
 720 $\hat{F} = \frac{MS_D}{MS_Q} = \frac{0.3230}{1.0008} = 0.323$
 721 $MS_Q = 1.0008$

722 $F(P = 0.05; df1 = 1, df2 = 45) = 4.06$

723 Result: The quadratic model is not significantly better than the linear model at the 5% level.

724

725 2. Fat:

726 Coefficient c: 0.0065 ± 0.0029 (not significantly different from zero at $P = 0.025$)

727 Residual error (linear fit): 1.0258

728 Residual error (quadratic fit): 0.9839

729 **Table 27:** Analysis of variance for fat

	Number of parameters in model	Remaining degrees of freedom	Sum of squares of residuals	Mean square (SS/df)
Linear fit:	2	48-2=46	SSL=48.4049	MSL=1.0523
Quadratic fit:	3	48-3=45	SSQ=43.5584	MSQ=0.9680
Difference		1	SSD= 4.8465	MSD=4.8465

730 $MS_D = 4.8465$

731 $\hat{F} = \frac{MS_D}{MS_Q} = \frac{4.8465}{0.9680} = 5.01$

732 $MS_Q = 0.9680$

733

734 $F(P = 0.05; df_1 = 1, df_2 = 45) = 4.06$

735 $F(P = 0.025; df_1 = 1, df_2 = 45) = 5.38$

736

737 Result: The quadratic model is significantly better than the linear model at the 5% level but not at the
738 2.5% level. In other words, deviation from linearity emerges.

739 Conclusion: The quadratic time significance test showed the same results as the lack of fit test (see
740 Step 5). The liver data can be considered linear. For fat, deviation from linearity emerged
741 ($0.05 > P > 0.025$). As already stated in the main part of the document, a re-calculation of the data
742 for fat excluding day 7 from calculation was not taken into account because, in our view, the linearity
743 assumption was not seriously violated.

744

745 **Annex D: Comparisons of different approaches for dealing** 746 **with values below the LOQ**

747 In January 2014 the CVMP published a concept paper (EMA/CVMP/SWP/285070/2013) proposing a
748 revision of the Note for guidance: Approach Towards Harmonisation of Withdrawal Periods, in order to
749 look again at the approach used for considering residues present at levels below the limit of
750 quantification (LOQ). The concept paper noted that the original Note for guidance recommends that a
751 value of half of the limit of quantification should be applied to data points below the limit of
752 quantification, but that since publication of the Note for guidance, more sophisticated methods for
753 dealing with levels below the limit of quantification have become available, such as the maximum
754 likelihood approach (i.e., determining the depletion curve that would maximise the likelihood of the
755 observed data).

756 Following the receipt of comments on the concept paper, the SWP undertook work comparing the
757 withdrawal periods calculated using different approaches for dealing with values below the LOQ. This
758 work indicated that the current method (assigning values below the LOQ to half the LOQ) provides
759 results that are comparable to those obtained using the maximum likelihood approach and also to
760 using data 'as measured'. This supports the view that the current approach remains appropriate and
761 that there is little to be gained by moving to an alternative. The CVMP therefore concluded that the
762 existing approach for the treatment of values below the LOQ should remain in place. However, it
763 should be noted that VICH GL49 recommends methods for determining the LOQ that are likely to make
764 this issue less of a problem (as LOQs are likely to be $< \frac{1}{2}$ MRL).

765 The work undertaken by the SWP in order to arrive at this conclusion is briefly described in this annex.

766 In a first step the SWP compared the following approaches:

- 767 (i) Omitting values below the LOQ;
- 768 (ii) Assigning a value of half the LOQ to values recorded as below the LOQ;
- 769 (iii) Using the maximum likelihood approach (i.e. the regression parameters were determined in
770 such a way that the likelihood of observing the given values above the LOQ and the given
771 frequency of values below the LOQ is maximised).

772 The results provided for liver in Annex A were used as the starting point from which to generate
773 simulated data sets (derived based on the intercept, slope and standard deviation of the original data).
774 Withdrawal periods were then derived from the (log transformed) simulated data sets either (i)
775 omitting values below the LOQ, (ii) using values of half the LOQ when recorded values were below the
776 LOQ, or (iii) using regression parameters based on the maximum likelihood approach. The original data
777 set was considered to represent reality and to yield the 'true' withdrawal period, i.e., to yield a
778 withdrawal period at the end of which 95% of all residue concentrations were, at most, as high as the
779 MRL.

780 In principle, if a sufficient number of simulated data sets is sampled and withdrawal periods derived,
781 then the frequency of withdrawal periods that are shorter than the 'true' withdrawal period should be
782 5% as, in line with the guideline, withdrawal periods should be derived in such a ways as to provide
783 95% confidence that they are not too short.

784 When withdrawal periods were derived treating values below the LOQ as described above, the
785 following results were obtained:

- 786 (i) when values below the LOQ were omitted 1.3% of estimated withdrawal periods were at most
787 as long as the 'true' withdrawal period (i.e. 98.7% were longer);

788 (ii) when values below the LOQ were replaced by a value of half the LOQ 5.6% of estimated
 789 withdrawal periods were at most as long as the 'true' withdrawal period (i.e. 94.4% were
 790 longer);

791 (iii) when the maximum likelihood approach was used to replace values below the LOQ 6.8% of
 792 estimated withdrawal periods were at most as long as the 'true' withdrawal period (i.e. 93.2%
 793 were longer).

794 In this example, the method currently used in the EU came closest to the 5% value, with the
 795 maximum likelihood approach being almost as good.

796 The above exercise was then repeated using a further four real data sets and the withdrawal periods of
 797 the simulated data sets derived treating values below the LOQ, as described above. In addition, a
 798 fourth approach was used in which withdrawal periods were derived by using the values recorded for
 799 values below the LOQ ('as measured' values).

800 For each of the four approaches withdrawal periods for the simulated data sets were derived using
 801 three different assigned LOQs (LOQ assigned so that the expected percentage of values below the LOQ
 802 was 5%, 10% or 20%) and using MRLs set to either twice the LOQ or 5 times the LOQ, resulting in six
 803 different combinations of assigned LOQ and MRL for each data set.

804 The results are summarised in the table below.

Approach for dealing with values below LOQ (BLOQ, <LOQ)						
Data set	%BLOQ	MRL	Omit	LOQ/2	As measured	Max Likelihood
A	5%	5 x LOQ	2.8	5.1	5.3	5.6
		10 x LOQ	1.9	6.4	5.4	5.5
	10%	5 x LOQ	2.2	4.5	4.7	4.8
		10 x LOQ	1.2	4.9	3.8	3.8
	20%	5 x LOQ	1.8	3.3	3.8	3.7
		10 x LOQ	1.0	4.7	3.6	3.7
B	5%	5 x LOQ	3.1	3.8	5.6	5.6
		10 x LOQ	2.1	5.9	5.1	5.4
	10%	5 x LOQ	2.4	2.8	4.3	4.2
		10 x LOQ	1.5	4.6	3.6	3.9
	20%	5 x LOQ	3.0	2.6	4.7	4.6
		10 x LOQ	1.6	4.6	3.8	4.0
C	5%	5 x LOQ	7.8	2.1	6.8	6.8
		10 x LOQ	2.6	3.0	5.6	5.8
	10%	5 x LOQ	7.6	1.2	5.7	5.6
		10 x LOQ	1.6	2.3	4.3	4.2
	20%	5 x LOQ	11.7	1.2	6.7	6.6
		10 x LOQ	1.6	1.6	3.9	3.9
D	5%	5 x LOQ	2.7	4.8	5.2	5.4
		10 x LOQ	1.6	5.8	4.5	4.8
	10%	5 x LOQ	2.1	3.9	4.2	4.2
		10 x LOQ	1.4	5.1	3.9	4.1
	20%	5 x LOQ	1.9	2.8	3.7	3.7
		10 x LOQ	1.2	4.3	3.3	3.3

805 The following observations can be made from the above table.

806 Omitting levels below the LOQ never came closest to yielding the desired frequency of 5% of
 807 withdrawal periods shorter than the 'true' withdrawal period. In most cases it was the most
 808 conservative method. This may be because omitting very low recorded residue levels will tend to make
 809 the regression line less steep.

810 Using 'as measured' values for values below the LOQ yielded good results. However, it should be noted
811 that in the simulation constant variability of (log-transformed) data was assumed. With real data sets
812 higher variability is often seen at low residue levels (as described by the Horwitz equation). Therefore,
813 the apparent appropriateness of this method could be an artifact of the simulation's simplicity. Another
814 potential difficulty with this approach is that measurements below the limit of quantification are often
815 not reported.

816 Assigning values below the LOQ as half the LOQ and using the maximum likelihood approach yielded
817 similarly appropriate results in most cases – withdrawal periods were generally similarly distributed,
818 and the fraction of withdrawal periods at most as long as the 'true' withdrawal period were similar.
819 However, for one data set (data set C) the maximum likelihood approach does appear to have yielded
820 better results.

821 Overall, the 'as observed' approach, the half LOQ approach and the maximum likelihood approach can
822 be considered to have yielded similar results, with the percentage of withdrawal periods that are too
823 short ranging from approximately 3% to less than 7%, corresponding to a confidence more than 93%
824 to approximately 97%.

825 It is acknowledged that the above investigation is limited and that further work could be undertaken to
826 further explore different approaches for dealing with values below the LOQ and for investigating
827 whether all assumptions used in derivation of withdrawal periods are supported. In reality it is likely
828 that there is not one single method that will be optimal for dealing with all data sets. Ideally, software
829 would be developed that would automatically identify and apply the most appropriate approach.
830 However, the development of such software would be a very substantial undertaking. VICH GL 49
831 (adopted by CVMP, March 2011) recommends that the LOQ for an analytical method should be
832 estimated as the mean of 20 control samples plus 6-10 times the standard deviation (SD), and then
833 confirmed, or be based on the ability of the method and the instrumentation used to detect and
834 quantify a specific analyte in a specific matrix (see Annexes 1 & 2 of GL49). Before GL49 was adopted,
835 the LOQ was routinely determined as 0.5 x MRL, leading to many results being reported as 'below LOQ'
836 (<LOQ or BLOQ). With the guideline-recommended method of determining the LOQ, it is foreseen that
837 there will be fewer data <LOQ, as the difference between LOQ and MRL would usually be greater than
838 that between 0.5 x MRL and MRL. This should lead to fewer issues around which values to use, as the
839 depletion curve would be better described.