

DRAFT GUIDELINE 223

March 2007

OECD GUIDELINES FOR THE TESTING OF CHEMICALS

PROPOSAL FOR A NEW GUIDELINE 223

Avian Acute Oral Toxicity Test

INTRODUCTION

1. This Test Guideline is designed to estimate the acute oral toxicity of substances to birds. This Guideline provides a sequential testing procedure that optimises the placement of doses and matches the precision of the endpoint with the precision required. The method has been designed to minimise the numbers of birds used.

2. This Guideline started development at the SETAC/OECD Workgroup on avian toxicity testing following a workshop held in Pensacola, United States, in 1994 (1) with subsequent open SETAC and closed OECD Expert Group meetings in Europe and the United States to develop and optimise the sequential testing design. The sequential testing design has been developed with extensive statistical validation (2).

INITIAL CONSIDERATIONS

3. The information required by different hazard assessment schemes may vary considerably. To satisfy these various needs, the following three tests are described:

- **Limit dose test** – this is the preferred test when toxicity is expected to be low and lethality is unlikely at the limit dose. The limit dose must be adequate for assessment purposes, and it is usually 2000 mg/kg body-weight.
- **Full test: LD50-only test** – this is the preferred test when an estimate of the median lethal dose is required but neither the slope of the dose response curve or the confidence interval for the LD50 is required. This may be the appropriate test to estimate a percentile of a species sensitivity distribution of LD50s and to provide information for product labelling purposes.
- **Full test: Dose-response test** – this is the preferred test when the slope of the dose response curve and/or the confidence interval is required in addition to an estimate of the LD50.

4. Definitions used in this Guideline are given in Annex 1.

PRINCIPLE OF THE TEST

5. The test is divided into a number of discrete stages. At each stage a number of birds are simultaneously given a dose (mg/kg body weight) of the test substance into the crop or proventriculus. At each stage, individual birds may receive different doses or doses may be replicated. The recommended strategy for testing materials that are unlikely to present a significant hazard is to perform a test with multiple birds dosed at the limit dose. If toxicity is expected the recommended strategy is to use a sequential design with 2, 3 or 4 stages. Stages 1 and 2 require non replicated doses, whilst stages 3 and 4 require replicated doses. In the first stage, the range of doses is based on the best available estimate of the LD50 – for example, the rodent LD50. Doses for subsequent stages are determined based on the mortalities observed in all previous stages, so that the estimation of the LD50 and the slope of the dose-response curve is carried out at the same time.

6. After dosing, the birds are observed for a 14 day period in order to measure mortality. It may be necessary to extend the observation period depending on evidence of delayed effects. The staged test design is easiest to apply to chemicals that produce death within a few days. In such cases it is not necessary to wait 14 days before moving to the next stage. Mortality observed after only three days may be used to determine doses for the following stage.

7. Mortality is the primary endpoint in this study and background mortality is presumed to be negligible. Controls are required to monitor the health and husbandry of test birds to ensure that the ability of the study to provide reliable results is not compromised.

8. Five untreated control birds from the same hatch will be included in the Limit test or the final stage of the Full test.

Control birds will be sham dosed with the same carrier (or capsule), maintained under the same conditions and weighed prior to dosing and 3 and 14 days after dosing. If there are deaths among control birds that might be attributed to disease or mishandling of animals, the study will be repeated. However, incidental deaths that result from self inflicted injuries such as broken legs or abrasions will not be considered to be a cause for repeating a study.

9. During the test, animals obviously in pain or showing signs of severe distress should be humanely killed.

DESCRIPTION OF THE TEST METHODS

Selection of bird species

10. Captive bred species with a low propensity to regurgitate are preferred because background mortality levels are very low. Frequently used species which fit these requirements are bobwhite quail, *Colinus virginianus* and Japanese quail, *Coturnix coturnix japonica* (Galliform).

11. It may be necessary to test additional species in order to take account of the species sensitivity. A short list of species that may be considered, in addition to quail, includes mallard, *Anas platyrhynchos* (Anseriform), pigeon, *Columba livia* (Collumbiform), zebra finch, *Poephila guttata* (Passeriform) and budgerigar, *Melopsittacus undulatus* (Psittaciform). This list is not intended to limit the recommended species but to provide guidance on some relatively robust and laboratory bred species from different Orders.

12. Birds should be in mature plumage but not in breeding condition. Wild phenotypes are preferred, where possible. Captive bred birds should be from the same source and breeding population, and when possible, breeding history should demonstrate periodic out-breeding to maintain genetic heterogeneity.

13. Birds should be drawn at random from a group comprising a single sex. Either sex may be used. If it is not possible to separate sexes by plumage, birds can be drawn at random from the whole population. If it is necessary to check for sensitivity due to sex an additional stage may be added using 4-6 birds of the opposite sex covering the critical dose range (2-3 sequential doses either side of the estimated mean).

Housing and test conditions

14. Individual caging is preferred to identify animals regurgitating the dose and to prevent fighting. However, group caging may be used if it improves animal welfare in the case of sociable species. Housing conditions should be within optimal limits for the test species and minimum cage sizes recommended are 3000cm² for pigeon; 2000cm² for mallard; 1000cm² for quail and 500cm² for budgerigar and zebra finch or larger to comply with national guidelines for animal welfare if different. Cage floors should be constructed of mesh, large enough to allow faeces to fall through, but not restrictive on the bird's movements. Pigeons, zebra finches and budgerigars require perches.

15. The test environment may be under controlled conditions or at ambient temperature and humidity. Temperatures within the range 15-27⁰C are suitable for quail and duck but should fluctuate as little as possible during the test stages. Ventilation should be sufficient to supply at least 10 changes of air/hour. The photoperiod for quail and mallard should be 8 hours light and 16 hours dark. For other species it may be necessary to increase the light phase to 10 hours. Fresh food and water should be provided *ad libitum*. Commercial gamebird diets can be used, but they must be nutritionally appropriate for the species used. Medication should be avoided within 14 days prior to dosing, during dosing and during the observation period. Diets and water should be periodically analysed to check for impurities.

Preparation of birds

16. Birds should be uniquely identified. Acclimatisation to test conditions and diet prior to dosing should be at least 14 days for cage reared birds. Normally, wild caught birds need longer acclimatisation periods. All birds must be in healthy condition and should not be used if greater than 5% of cage-reared and greater than 10% of wild test birds die during the acclimatisation period. If wild birds do not acclimatise they should be released. Cage-reared birds should be of approximately the same age.

Preparation of doses

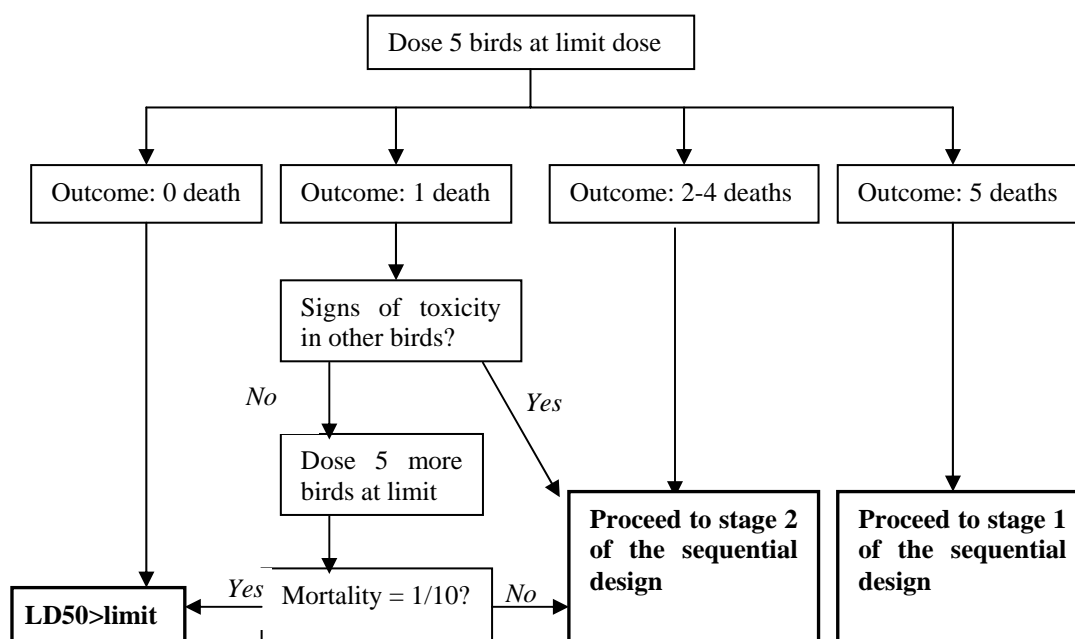
17. The test substance should be dissolved or suspended in a suitable vehicle or administered in a capsule. If the substance is dissolved or suspended it is recommended that, whenever possible, the use of an aqueous solution or suspension be considered first, followed by consideration of a solution or emulsion in oil (e.g. corn oil) and then by possible solution in other vehicles. For vehicles other than water, the toxicity of the vehicle must be known and should not cause vomiting. The dose will be determined based on body weight measured within 24 hours of dosing.

PROCEDURE

Limit test

18. The limit test should be employed unless available information suggests that the LD50 may be below the limit dose (typically 2000 mg/kg). Figure 1 describes the procedure to be followed according to the mortality observed. The limit test design consists of dosing 5 animals simultaneously at the limit dose. Birds are given a single oral dose of the test substance (mg/kg body weight) into the crop or proventriculus, and then observed for 14 days. If no mortality occurs, it can be concluded at the 95% confidence level, that the LD50 is above the limit dose.

Figure 1: Limit Test Procedure



19. If only one death is observed, and no signs of toxicity are observed in other birds, then five more birds can be dosed at the limit, to attempt to demonstrate with adequate confidence (i.e. 95%) that the LD50 is above the limit dose. Additional dosing can begin before the 14 day observation period is complete. If there is only one death in the total of 10 birds, then it can be concluded that at the 95% confidence level the LD50 is above the limit dose.

20. If the observed mortality is 1 out of 5 birds and there are signs of toxicity in other birds, or if there are 2 or more mortalities among 10 birds, or if there are between 2 and 4 mortalities among 5 birds, use the sequential design described later, but starting with stage 2. If mortality is complete (i.e. all 5 birds have died), use the sequential design but starting with stage 1. Additional dosing can begin before the 14 day period of observation is complete.

Table 1. Working estimate of LD50 for use in stage 2 of the sequential design derived from mortality in a limit test at 2000mg/kg.

Mortality (%)	10	20	30	40	50	60	70	80	90
Working Est. LD50	3609	2947	2546	2247	2000	1780	1571	1357	1108

21. To proceed from the limit test to stage 2 of the sequential design, an working estimate of the LD50 is needed to select doses at this stage. Appropriate estimates for this use are given in Table 1. Using the appropriate working estimate of the LD50, the lowest and highest doses can be calculated as described in the directions for stage 2 (described in Annex 3). The calculated high dose (*hdose*) will generally need to be adjusted (reduced) to the highest level that is consistent with practicality and animal welfare constraints. It should be noted that in some circumstances, the LD50 cannot be estimated without using doses above the limit dose. In addition, if the outcome of the limit test is complete mortality, an estimate of the lower bound of the LD50 is also required. Because there are constraints on the use of very high doses of test substance, it may not always be possible to estimate the LD50 for slightly toxic substances. If mortality is complete (i.e. all 5 birds have died), it is not possible to use the limit test data to estimate the LD50 value, so proceed to stage 1.

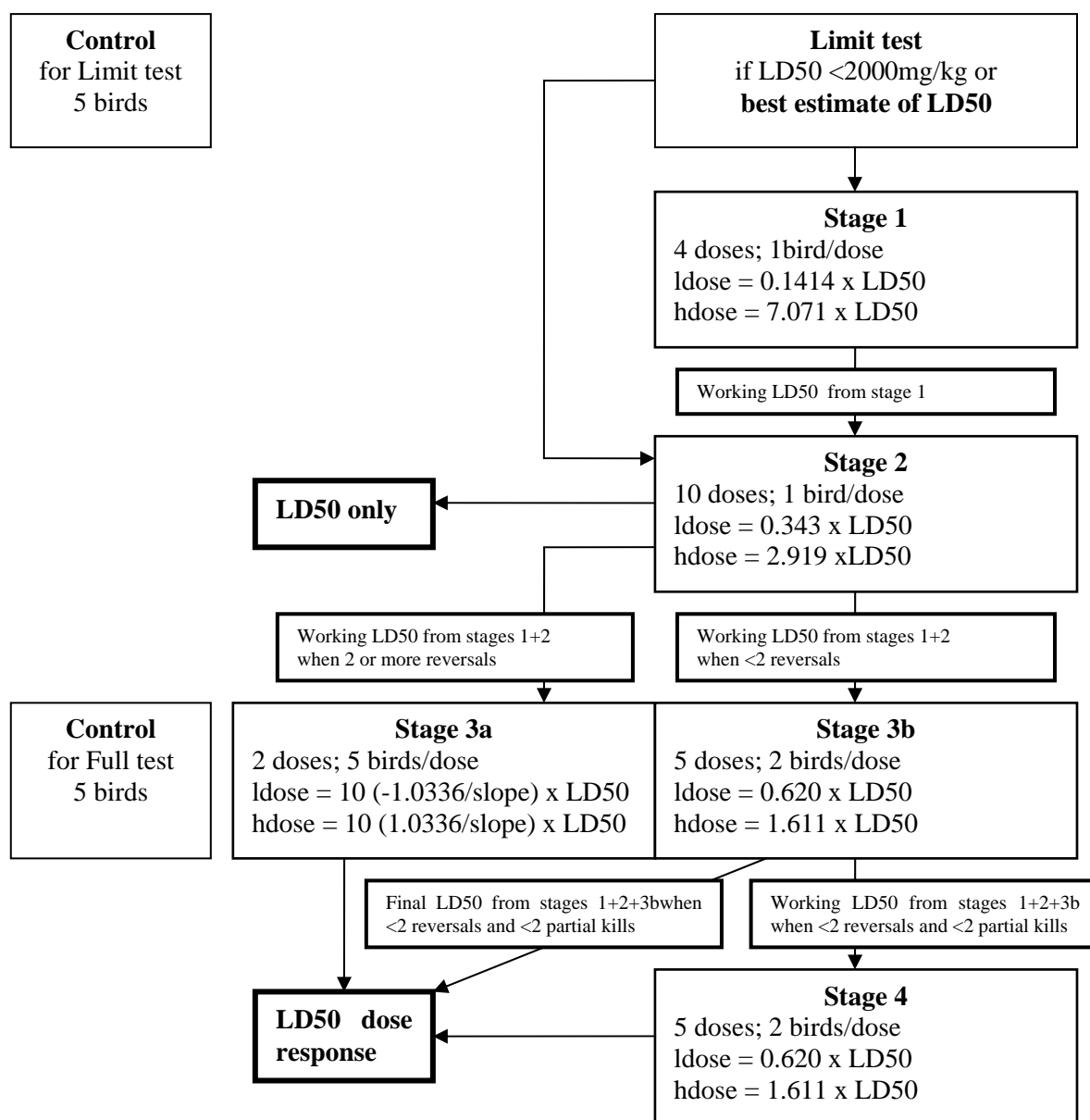
Full test

22. There are two kinds of full test: LD50-only test and dose-response test. Both tests employ sequential test designs where animals are dosed at different times (stages) to allow the use of accumulated information to position doses along the dose response curve. The LD50-only test and the dose-response test differ primarily in the total number of birds used, and in the number of dosing stages. The rationale and statistical basis of the full test can be found in Annexe 2. A detailed description can be found in Annexe 3. A flow chart is presented in Figure 2.

23. The full dose response test designs have been shown to have adequate performance properties for estimating LD50, confidence intervals, and dose response slopes. The dose-response test incorporates three or four dosing stages. However, the recommended specifications for the LD50-only test call for only two dosing stages, with the consequence that the confidence in the LD50 estimates is less, and that the LD50-only design is unsuited for estimation of the dose response slopes.

24. At each stage in both full test designs, one or more birds are given a single oral dose (mg/kg body weight) of the test substance using doses that are expected to bracket the evolving working estimate of the LD50. Birds are observed for 14 days, but selection of doses for subsequent stages is typically based on results after 3 days. This interval may be reduced if birds quickly show signs of recovery or extended if delayed mortality is expected or observed. The full dose response test can be initiated using information gained from a failed limit test (LD50 <2000mg/kg). Alternatively, for compounds of suspected high toxicity, testing may be initiated by using a 1st dosing stage where each bird is given a different dose, so that doses are selected to cover the best available estimate of the LD50 (e.g. based on the rodent or other bird species' LD50). Depending on the outcome of the 1st stage, the doses bracketing the working estimate of the LD50 are determined for the 2nd stage as described in Annexes 2 and 3, and 1 bird is dosed at each of ten doses. If there is a working estimate of the LD50 from a failed limit test, the full dose-response test starts with stage 2 also. The process continues to a 3rd stage and possibly a 4th in the dose-response test or if added precision is needed in the estimate of the LD50. Observations of deaths that are clearly not treatment related, (e.g. physical injury) should be excluded from calculations. The LD50-only test always begins with a 1st dosing stage in which each bird is given a different dose, and stops after the 2nd dosing stage.

Figure 2: Full Test Procedure



Administration of dose

25. The test substance is administered in a single dose by gavage or capsules. The dosing volume must remain constant with respect to body weight and should not exceed 10ml/kg body weight. Birds should be fasted for 12-15 hours overnight immediately prior to dosing. Shorter fasting periods may be necessary for smaller and wild caught species. Regurgitation must be recorded. The addition of a non-toxic coloured food dye which contrasts with faeces will allow regurgitation to be more easily recognised.

26. Regurgitation compromises the evaluation of toxicity and is a feature of acute oral toxicity testing in birds. It may be related to the dosing technique or characteristics of the test substance. The frequency of regurgitation may be reduced by lowering the dose volume or by changing carriers.

Observations

27. Birds are observed individually during the first 2 hours after dosing for regurgitation and the onset of clinical signs, on at least 3 evenly spread additional occasions during the first 24 hours for clinical signs and at least daily thereafter for a total of 14 days. However, the duration of the observation period should not be fixed rigidly. It should be determined by the cessation of clinical symptoms and death and may thus be extended when considered necessary.

28. Observations on each individual should include regurgitation, signs of intoxication and remission, abnormal behaviour, bodyweight, mortality and time to death.

29. Birds may be weighed before dosing, then 3, 7 and 14 days after dosing (or later depending on the duration of the study) to determine weight change. Food consumption may be measured on days 1, 3, 7 and 14 days after dosing. Gross pathology should be undertaken on all birds from each treatment group to help identify incidental mortalities and obvious symptoms of toxicity.

DATA AND REPORTING

Data

30. Individual bird data should be provided and summarised in tabular form, showing the dose, number of birds tested, signs of toxicity, death and numbers sacrificed for humane reasons, time of death of individual birds, a record of the time to onset and cessation of clinical signs, bodyweight change, food consumption and gross pathological findings.

Calculation of the LD50

31. The methods for estimating the LD50 (median lethal dose) are the same for staged tests as for other types of dose response tests. Certain features of the staged designs are based on the assumptions that the underlying form of the dose response curve approximates the probit model, so the use of a probit regression model (with the logarithm of dose as the independent variable) to estimate the LD50 is appropriate. For the probit model, the maximum likelihood estimate of the $\text{Log}(\text{LD50})$ is $-a/b$, where a is the intercept and b is the slope. The logistic model is very similar in shape to the probit model, and can be used in its place in the sequential design. Furthermore, when the mortality data from all stages of a study are pooled for analysis, other models may also be appropriate if the probit model fails to fit the observations.

32. Standard maximum likelihood methods used to fit models to the data can be used only if there are two doses that display partial mortality (i.e. neither 0 nor 100%), or there is a reversal in trend (i.e. mortality at a lower dose is greater than mortality in one of the higher doses) somewhere in the observed responses. A benefit of the sequential design is that these conditions will usually be met. On the rare occasions they are not met, various interpolation and moving average methods may be used to estimate the LD50. A description of appropriate methods can be found in (2).

33. Confidence intervals for the LD50 can be obtained using Fieller's theorem, (Fieller 1940) likelihood ratio methods (Crump and Howe, 1985, Piegorsch and Bailer 1997), or by binomial methods (Chapman et. al. 2002). Many standard statistical packages have built in routines to estimate the LD50 and

its confidence interval (SAS©, LogXact©, Toxstat©, BMD5 [available from USEPA - calculates lower confidence limit only]).

34. It should be recognized that it is often possible to calculate the slope and confidence intervals for the LD50 from the results of the LD50-only test. However, due to the small number of replicates, these estimates may be unreliable. Thus, the slope and confidence intervals for the LD50 should be reported only for the full dose-response test. Some measure of goodness of fit of the regression model (e.g. Pearson goodness of fit, likelihood ratio Chi-square) should be reported for each dose-response test. Reporting of a measure of goodness of fit is optional for the LD50-only test.

Report

35. The report should contain the following minimum information to confirm compliance with the Guideline and test results:

test substance

- identification
- batch and lot number
- purity
- stability at room temperature
- volatility

test method and system

- test type
- test species, source, strain, age, weight, health
- description of test method

conduct of test

- test groups and design (no treatments and replicates, individual or group caging)
- acclimation and assignment procedures (duration, randomisation)
- dose method (gavage/capsule, carrier/solvent, volume/bird as % body weight)
- housing conditions (type, size, pen materials, floor covering, temperature, humidity,

photoperiod, light intensity)

- food and water (availability, identification, source, composition, calorific value, results of contaminant analysis)
- frequency, duration and method of observations (health/mortality, body weight, food

consumption)

- description of statistical methods

results of test

- mortality (time to death, clinical symptoms, calculation of LD₅₀)
- onset (minutes) and cessation of clinical signs
- gross pathological examination

LITERATURE

- (1) OECD 1996. Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996). OECD/GD (96) 166.
- (2) Chapman, P.F., R. Dark, T. A. Springer. 2007. Avian Acute Oral Toxicity: A Statistical Evaluation of Sequential Experimental Designs. In preparation.
- (3) Crump, K.S. and R. Howe. 1985. A review of methods for calculating confidence intervals in low dose extrapolation. In: D. Krewski, Ed. Toxicology and Risk Assessment. CRC Press. Canada.
- (4) Efron, B. and R. Tibshirani. 1993 Introduction to the bootstrap. Chapman and Hall. New York.
- (5) Neyer, B.T. 1994. A D-Optimality Based Sensitivity Test, Technometrics, 36:61-70.
- (6) Piegorsch, W. W. and A.J. Bailer. 1997. Statistics for environmental biology and toxicology. Chapman and Hall. London.

ANNEX 1

DEFINITIONS

Acute oral toxicity refers to those adverse effects occurring following oral administration of a single dose of a substance, or multiple doses given within 24 hours.

Dose is the amount of test substance administered. Dose is expressed as weight (g, mg) or as weight of test substance per unit weight of test animal (e.g. mg/kg).

LD50 (median lethal oral dose), is a statistically derived single dose of a substance that can be expected to cause death in 50 per cent of animals when administered by the oral route. The LD50 value is expressed in terms of weight of test substance per unit weight of test animal (mg/kg).

Limit dose refers to a dose at an upper limitation on testing.

Probit is an abbreviation for the term “probability integral transformation” and a probit dose-response model permits a standard normal distribution of expected responses (i.e., one centred to its mean and scaled to its standard deviation) to doses (typically in a logarithmic scale) to be analyzed as if it were a straight line with slope the reciprocal of the standard deviation. A standard normal lethality distribution is symmetric; hence, its mean is also its true LD50 or median response.

Slope (of the dose-response curve) is a value related to the angle at which the dose response curve rises from the dose axis. In the case of probit analysis, when responses are analyzed on a probit scale against dose on a log scale this curve will be a straight line and the slope is the reciprocal of the standard deviation of the underlying test subject tolerances, which are assumed to be normally distributed.

Stage (in a sequential design) refers to a period during an experiment in which a number of birds are dosed simultaneously, and observed for a period of time. Provisional estimates of LD50 and slope from the previous stage may influence the design of the stage. Likewise the data resulting from the stage may influence the design of succeeding stages.

ldose is the lowest dose used during a particular stage (mg/kg body weight).

hdose is the highest dose used during a particular stage (mg/kg body weight).

step is the multiplication factor used in calculating individual doses.

Initial estimate of the LD50 is an estimate made prior to the start of the study and around which the 4 doses in stage 1 are arranged.

Working estimate of LD50 and slope are provisional estimates made before the end of a stage and which are used to determine doses in following stages. Typically working estimates are based on observations from all stages that have been started and made 3 days after the start of the current stage.

Final estimates of LD50 and slope are made after the study has been stopped. They are based on observations from each completed stage, normally 14 days or longer.

Partial kills can occur when a single dose is given to more than 1 bird. If some birds die and some birds survive this outcome is called a partial kill. 2 partial kills occurs when 2 different doses exhibit partial kill.

Reversals are analogous to partial kills for non-replicated single birds receiving a dose and occur when lower doses result in a kill while higher doses result in survival. This is likely to occur in the region of the LD50. So for example, if 0 represents survival and 1 represents kill and we list responses in order of increasing dose then a sequences such as 0001011111, 000110111 both exhibit a single reversal. 0001010111 and 00011010011 both exhibit 2 reversals.

ANNEX 2

GENERAL DESCRIPTION, RATIONALE AND STATISTICAL BASIS OF THE SEQUENTIAL DESIGN

Statistical Background

1. The philosophy underlying tolerance distributions is that an individual bird will die if it receives a dose above a certain value and will survive if the dose is less than this value. The specific value is called a tolerance and is assumed to be fixed for an individual bird, but to vary between birds. Thus if we have a population of birds we can speak of a distribution of tolerances, or a tolerance distribution.

2. In order to estimate the tolerance distribution from a sample of birds we fit a statistical model. If we assume that the tolerances follow a normal distribution we fit a probit model which takes the form:

$$\text{Probit}(p) = \alpha + \beta \cdot \log(d)$$

Where: p is the probability that the tolerance of an individual bird is less than dose d – i.e. the probability that a bird receiving dose d will survive.

α and β are parameters representing the intercept and slope of a straight line relationship between probit (p) and $\log(d)$.

3. The probit model is fitted to test data in order to obtain estimates of the parameters α and β which we call a and b respectively. The estimate of the mean of the tolerance distribution of the population of birds (called the LD50) can then be determined from the equation:

$$\text{Estimate } (\log(\text{LD50})) = -a/b,$$

And the variance of the tolerance distribution can be estimated by:

$$\text{Variance} = 1/b.$$

4. Once the tolerance distribution has been estimated (i.e. once estimates of α and β have been obtained) percentiles of the distribution can be computed (dose levels corresponding to specific values of p).

5. Methods for estimating LD50, slope and confidence intervals for a variety of different experimental outcomes are recommended in (2). The slope can only be estimated if there are at least 2 partial kills or at least one dose response inversion. If a single dose is given to two or more birds, a partial kill occurs when the percent death is greater than 0% and less than 100%, i.e. not all birds die and not all survive. If single doses are only given to single birds, then a dose response inversion happens when a lower dose results in death but an even higher dose results in survival.

Description of the Sequential Design

1. This design consists of several stages, in each of which a number of birds are dosed simultaneously and observed for a period of time. The advantage of such a design is that information gathered in earlier stages, notably working estimates of LD50 and slope of the dose response curve, is used to help design the next stage, in particular placing of doses. The generic design is completely flexible in respect of number of stages, number of doses per stage, and number of birds per dose. The specific design described here has been optimised by means of computer simulations and consists of 4 birds in the first and 10 birds in later stages, whilst the number of doses and birds per dose varies between stages, as will be explained. At the start of the test it is assumed that we can make some initial estimate of the value the LD50. In practice this estimate may be made using knowledge of the chemistry and toxicity of similar active ingredients, and also from the results of tests on other bird species or mammals. Alternatively the initial estimate can be derived from a limit test. The test generally has at least two stages, and may be extended to three or more. The first three stages differ considerably from each other. The first is a 'ranging' stage designed to confirm and improve an initial guess at the LD50. The second is designed to more accurately locate the LD50. The third (and any subsequent stage) is intended to provide estimates of the confidence interval and slope of the dose response curve or to further reduce uncertainty in the estimate of the LD50.

2. Because of the way that the dosing of stages is staggered, working estimates of the LD50, and slope are recalculated repeatedly, and each time the recalculations are performed the number of groups with partial mortality or the number of reversals in mortality will change. Working estimates may be obtained after 3, 6 and 9 days but in all stages observations are continued until the end of day 14. At the very end of the study final estimates of LD50 and slope (and confidence intervals, if appropriate) are determined from the combined set (i.e. from all stages) of 14-day assessments. So, for example, 3 days after the start of stage 1 working estimates are used to design stage 2. 3 days after the start of stage 2 (and 6 days after the start of stage 1) working estimates are used to design stage 3. 3 days after the start of stage 3 (which is 9 days after the start of stage 1, and 6 days after the start of stage 2) working estimates are used to design stage 4, if necessary. In the discussion below, it should be clear from the context whether 3, 6, 9 day or 14 day observations are intended. This is because final estimates are based on 14 day assessments (from all stages) and working estimates, which are used for designing following stages, are based on interim assessments.

3. If a test material that is expected to have low toxicity is to be tested, the study may be initiated with a limit test. If the limit test is performed, 5 birds are given a single dose, typically 2000 mg/kg. If no bird dies then it may be assumed (with 95% confidence) that the LD50 is greater than the limit dose. If 1 bird dies, an additional group of 5 birds may be tested. If none of the additional 5 die, then again it can be assumed that the LD50 is greater than the limit dose. If between 2 and 4 birds die when 5 are dosed, or between 2 and 6 die when 10 birds are dosed, then a working estimate of the LD50 can be obtained from the failed limit test, and the multi-dose stage 1 described below can be omitted. In this case the working estimate of the LD50 for use in stage 2 is determined from Table 1. If all five of the initially dosed birds die, then stage 1 cannot be omitted. Limit test results can be combined with results from stages 1, 2, etc in order to obtain working or final estimates of LD50 and slope.

4. In stage 1, 4 doses are equally spaced on a log scale around the initial estimate of the LD50. The ratio of the highest to the lowest dose is set to 50, the value of 50 being based on extensive computer simulations (2). Each dose of the test substance is given to a single bird. (In computer simulations it was found that 4 doses and a very large ratio of high to low dose gave an adequate working estimate of the LD50 and preserved birds for later stages where they were more useful in estimating the slope.) A working estimate of the LD50 is computed as the geometric mean of the transition doses (see Table 2 in Annexe 3).

This method of estimating the LD50 has several advantages. First, it can provide a reasonable estimate given very little information, and second, it can be adapted to provide a working estimate of the LD50 when there is mortality at the lowest dose or survival at the highest dose. Either of these cases suggests that the actual LD50 may be outside of the range of the test concentrations. It is then assumed that if there was not survival at the low dose, there would be at a transition to survival at low dose/step. The geometric mean of transitions, including these values is then taken as the estimate the LD50. Likewise if there was no mortality at the high dose, it is assumed that there would be at a transition to mortality at high dose*step. The geometric mean of transitions, including these values, is then taken as the working estimate the LD50.

5. The purpose of stage 2 is to refine the estimate of the LD50 and to obtain an initial estimate of the slope. 10 doses are equally spaced on the log scale about the working estimate of the LD50 from stage 1. Assuming a slope of 5 (and assuming that the tolerance distribution is a normal distribution with mean equal to the LD50 and standard deviation equal to 1/slope) the extreme doses are placed at those points corresponding to 1 and 99 percent kill. (As justification for using a value of 5 for slope, a review of the EPA "one-liner" database suggests that, for the historical record of pesticide tests, the modal slope for avian acute tests is 5.) Each dose is given to a single bird. For the LD50 only test, observations are continue for a total 14 days and a final estimate of the LD50 is obtained by fitting a probit model to the combined data from limit test and from stages 1 and 2 . For the dose response test a probit model is fitted to the combined data from limit test and from stages 1 and 2 in order to obtain working estimates of the LD50 and, if possible, the slope. In addition the number of reversals in the combined results from stages 1 and 2 is counted. If the number of reversals is 2 or greater and a working estimate of the slope is obtained, then stage 2 is followed by stage 3a. Otherwise stage 2 is followed by stage 3b.

6. The purpose of stage 3a is to refine and improve the working estimates of LD50 and slope obtained from stage 2, choice of doses is based on these working estimates. If the working estimate of slope from stage 2 is either greater than fifteen or less than one, then it is set to be fifteen or one respectively. These limits on slope estimates are needed because maximum likelihood estimates using small numbers of observations may give extreme (probably grossly inaccurate) slope estimates. Only two doses are used - half of the birds are given a lower dose corresponding to 15% kill, and the remaining birds are given an upper dose corresponding to 85% kill. Observations are continue for 14 days and a final estimates of the LD50 and slope are obtained by fitting a probit model to the combined data from limit test and from stages 1, 2 and 3a. The rationale for placing doses at the 15th and 85th percentiles follows from a statistical concept called D-optimality. This form of optimality gives the best simultaneous estimates of LD50 and slope (in the sense that the width of confidence intervals for LD50 and slope are simultaneously minimised).

7. In stage 3b, 5 doses are equally spaced on a log scale around the working estimate of LD50 obtained from stage 2. The slope is again assumed to be 5, and extreme doses are again placed at responses corresponding to 15% and 85% kill. A probit model is fitted to data from limit test and stages 1,2 and 3b in order to obtain working estimates of the LD50 and slope, and number of reversals and number of partial kills are counted. If the number of reversals is 2 or more and/or the number of partial kills is 2 or more, and the slope can be estimated, then the study can be stopped. If these conditions are not met then stage 4 can be run. If the study is stopped, observations are then continued until a total of 14 days is completed. Final estimates of LD50 and slope are then obtained by fitting a probit model to the combined data from limit test and stages 1, 2, and 3b.

8. Stage 4 is simply a repeat of stage 3b. Doses are equally spaced as described for stage 3b around the working estimate of the LD50 obtained from stage 3b.

ANNEX 3

DETAILED DESCRIPTION OF THE SEQUENTIAL DESIGN

The design consists of a number of stages, i.e. periods during an experiment in which a number of birds are dosed simultaneously and observed for a period of time. If an LD50 estimate only is required, a 2 stage 14 bird design, with 4 birds in the first stage and 10 in the second, is recommended. Alternatively, if both LD50 and slope estimates and/or a confidence intervals are required, then a 3 or 4 stage design, with 4 birds in the first stage and 10 birds in each remaining stage, is recommended. At stage 3 there are two alternatives, 3a and 3b, the choice being determined by the outcome of stage 2. Depending on the outcome of stage 3b there may also be a need to proceed to a stage 4. An explanation of the logic and statistical basis of the design is given in Annex 2. Up to four stages of the design are described below.

The following parameters are used in this description:

ldose the lowest dose used during a particular stage (mg/kg body weight).

hdose the highest dose used during a particular stage (mg/kg body weight).

step the multiplication factor used in calculating individual doses.

Limit Test: The test is described in the main body of the guideline, and in Annex 2. If the result of the test permits moving directly to stage 2, then a working estimate of the LD50 for use in stage 2 is taken from Table 1 (also in main body of the guideline).

Stage 1: 4 doses are equally spaced on a log scale around the initial estimate or guess of the LD50.

(1) Calculate

$$ldose = 0.1414 \times (\text{initial estimate of LD50}).$$

$$hdose = 7.071 \times (\text{initial estimate of LD50}).$$

If *hdose* is greater than 3330 then set *hdose* = 3330 (may be less if limited by physical constraints) and recalculate the lowest dose as $ldose = hdose / 50$.

(2) Calculate: $step = 50^{1/3}$.

(3) Calculate the second and third doses:

$$dose\ 2 = ldose \times step, \quad dose\ 3 = dose\ 2 \times step = ldose \times step \times step.$$

(4) Give each of the four doses to a single bird.

(5) Observe birds for a period (typically 3 days), and note whether each bird is dead or alive. Compute a working estimate of the LD50 as the geometric mean of the doses that produce a transition from survival to death (see Table 2 below).

(6) Continue to observe the birds for a further 11 days (14 days in total).

Stage 2: 10 doses are equally spaced on a log scale around the working estimate of the LD50 obtained from stage 1.

(1) Calculate: $ldose = 0.3425 \times LD50$ and $hdose = 2.919 \times LD50$. If *hdose* is greater than 3330 then set *hdose* = 3330 (may be less if limited by physical constraints) and recalculate the lowest dose as $ldose = hdose / 8.5$.

(2) Calculate: $step = (hdose / ldose)^{1/9}$

(3) Calculate the 8 intermediate doses: $dose\ i = ldose \times step^{(i-1)}$, for *i* = 2 to 9.

- (4) Give each of the 10 doses to a single bird.
- (5) If an LD50 only test is being run, observe the birds for a total of 14 days and count the number dead and alive at the end of that period. Fit a probit model to the combined stage 1 and stage 2 14 day mortality data in order to obtain a final estimate of the LD50. It may also be possible to obtain a final estimate of the slope, but this is not a requirement of the LD50 only test.
- (6) If a full dosed response test is being run, observe the birds for a period (typically 3 days), and note whether each bird is dead or alive. Count the number of reversals and fit a probit model to the combined mortality data from stages 1 and 2 in order to obtain a working estimate of the LD50 and, if possible, the slope. If two or more reversals are observed and a working estimate of the slope has been obtained proceed to stage 3a, otherwise proceed to stage 3b.
- (7) Continue to observe the birds for a further 11 days (14 days in total).

Stage 3a: Two doses are equally spaced on a log scale around the working estimate of the LD50 obtained from Stages 1 and 2.

- (1) Calculate: $ldose = 10^{(-1.036 / slope)} \times LD50$ and $hdose = 10^{(1.036 / slope)} \times LD50$.
If an estimate of the slope was not obtained from stages 1 and 2, assume it to be 5. If the estimate of slope from stage 2 is either greater than fifteen or less than one, then set it to fifteen or one respectively. If $hdose$ is greater than 3330 then set $hdose = 3330$ (may be less if limited by physical constraints) and recalculate the lowest dose as $ldose = hdose / 10^{(2.072 / slope)}$.
- (2) Give each dose to 5 birds.
- (3) Observe the birds for 14 days and note whether each bird is dead or alive. Fit a probit model to the combined mortality data from stages 1, 2 and 3a, in order to obtain final estimates of both the LD50 and the slope. Then stop the study.

Stage 3b: 5 doses are equally spaced on a log scale around the working estimate of the LD50 obtained from Stages 1 and 2.

- (1) Calculate: $ldose = 0.6205 \times LD50$ and $hdose = 1.6113 \times LD50$. If $hdose$ is greater than 3330 then set $hdose = 3330$ (may be less if limited by physical constraints) and recalculate the lowest dose as $ldose = hdose / 2.6$.
- (2) Calculate: $step = (hdose / ldose)^{1/4}$.
- (3) Calculate the 3 intermediate doses: $dose_i = ldose \times step^{(i-1)}$, for $i = 2$ to 4 .
- (4) Give each dose to 2 birds.
- (5) Observe the birds for a period (typically 3 days) and note whether each bird is dead or alive. Fit a probit model to the combined mortality data from stages 1, 2 and 3b in order to obtain a working estimate of the LD50 and slope.
- (6) If 2 or more reversals are observed, or 2 or more partial kills, and a working estimate of the slope has been obtained then the study can be stopped. Continue observing the birds until the end of the 14th day and fit a probit model to the combined data from stages 1, 2 and 3b, in order to obtain final estimates for both LD50 and slope.
- (7) If conditions in 6 above are not satisfied proceed to stage 4, but also continue observing the birds until the end of the 14th day.

Stage 4: 5 doses are equally spaced on a log scale around the working estimate of the LD50 obtained from Stages 1,2 and 3b.

- (1) Calculate: $ldose = 0.6205 \times LD50$ and $hdose = 1.6113 \times LD50$. If $hdose$ is greater than 3330 then set $hdose = 3330$ (may be less if limited by physical constraints) and recalculate the lowest

dose as $ldose = hdose / 2.6$.

- (2) Calculate: $step = (hdose / ldose)^{1/4}$.
- (3) Calculate the 3 intermediate doses: $dose_i = ldose \times step^{(i-1)}$, for $i = 2$ to 4.
- (4) Give each dose to 2 birds.
- (5) Observe the birds for 14 days and note whether each bird is dead or alive. Fit a probit model to the combined data from stages 1, 2, 3b and 4 in order to obtain final estimates of both the LD50 and the slope. Then stop the study.

Table 2. Example of how the approximate LD50 is computed from the 4 doses in stage 1 by computing the geometric mean of the doses that produced a transition from survival to death. Survival is represented by O and death by X.

dose1	dose2	dose3	dose4	Approx. LD50 Estimate
O	O	O	O	$(dose4 \times dose5)^{1/2}$
O	O	O	X	$(dose3 \times dose4)^{1/2}$
O	O	X	O	$(dose2 \times dose3 \times dose4 \times dose5)^{1/4} = (dose3 \times dose4)^{1/2}$
O	X	O	O	$(dose1 \times dose2 \times dose4 \times dose5)^{1/4} = dose3$
X	O	O	O	$(dose0 \times dose1 \times dose4 \times dose5)^{1/4} = (dose2 \times dose3)^{1/2}$
O	O	X	X	$(dose2 \times dose3)^{1/2}$
O	X	X	O	$(dose1 \times dose2 \times dose4 \times dose5)^{1/4} = dose3$
X	X	O	O	$(dose0 \times dose1 \times dose4 \times dose5)^{1/4} = (dose2 \times dose3)^{1/2}$
O	X	O	X	$(dose1 \times dose2 \times dose3 \times dose4)^{1/4} = (dose2 \times dose3)^{1/2}$
X	O	X	O	$(dose0 \times dose1 \times dose2 \times dose3 \times dose4 \times dose5)^{1/6} = (dose2 \times dose3)^{1/2}$
X	O	O	X	$(dose0 \times dose1 \times dose3 \times dose4)^{1/4} = dose2$
O	X	X	X	$(dose1 \times dose2)^{1/2}$
X	O	X	X	$(dose0 \times dose1 \times dose2 \times dose3)^{1/4} = (dose1 \times dose2)^{1/2}$
X	X	O	X	$(dose0 \times dose1 \times dose3 \times dose4)^{1/4} = dose2$
X	X	X	O	$(dose0 \times dose1 \times dose4 \times dose5)^{1/4} = (dose2 \times dose3)^{1/2}$
X	X	X	X	$(dose0 \times dose1)^{1/2}$

Note. Even though only 4 doses (dose1 through to dose 4) are used in the test, values for dose0 and dose5 are mentioned in the table. The values that should be used for these doses are one step up or down from the actual test doses. That is, $dose0 = dose1 / step$ and $dose5 = dose4 \times step$. Dose0 must be added to the computation of the approximate LD50 when mortality occurs at the lowest test dose, and dose5 is added when there is survival at the highest test dose.