REPORT ON ACTIVITIES OF THE OECD EXPERT GROUP ON AVIAN AVOIDANCE TESTS BETWEEN 1994 AND 2014

Introduction

1. This document presents the state-of-play on OECD activities on avian avoidance testing, as of 2014. Originally intended as a harmonised OECD guidance document, however no consensus has been reached yet among countries on the level of guidance to provide, on the test design and sequence of testing, and on several other issues. Each sub-section contains a brief record of the arguments considered by the Expert Group and the conclusions reached.

Initial Considerations

Possible Role of Avoidance Testing in Avian Risk Assessment

2. In the wild, a range of different foods and habitats are available to birds. They may avoid those which are contaminated with toxicants. There are many examples of wild birds avoiding foods which contain natural chemicals whose repellency has been confirmed in tests with captive birds (e.g. Brower and Fink 1985, Buchsbaum et al. 1984, Crocker and Perry 1990, Jakubas and Gullion 1990, Kvitek 1991, Mason 1990, Mason et al. 1989b, Mason et al. 1991, Mason and Turpin 1990, Rowell-Rahier et al. 1995). Many pesticides are strongly avoided by captive birds in the 5-day dietary toxicity test (see consumption data in Hill and Camardese 1986) and in other types of test (e.g. Avery 1989, Avery and Decker 1991, Avery et al. 1993, Avery et al. 1994a, Avery et al. 1994b, Babu 1988, Bennett 1989a, b, Bennett and Prince 1981, Fryday et al. (1998), Grau et al. 1992, Grue 1982, Hart et al. (1999), Hill 1972, Kononen et al. 1986, 1987, Mason and Reidinger 1982, Robel and Morrow 1987, Rogers 1974, Schafer et al. 1983, Schafer and Brunton 1971). Furthermore, adding repellents to pesticide formulations may cause them to be avoided by birds (Mastrota and Mench 1995). It therefore seems reasonable to suppose that some pesticide products may be avoided to some extent by birds in the wild, reducing the risk of poisoning below what would otherwise be expected. Thus tests of avoidance may have a role in the later stages of risk assessment, as one of the factors to be considered in refining the preliminary assessment of risk.

Terminology

3. Several terms have been used in relation to the phenomenon of avoidance, including repellency, palatability, acceptance, anorexia and aversion. Most of these terms imply particular mechanisms, or are not applicable to the avoidance of materials other than foods. Throughout the document ‘avoidance’ is used as a general term.

Existing Methods for Testing Avoidance

4. Testing for avoidance has been requested in the past for some pesticides by regulatory authorities in France, Germany and the UK. Guidelines have been published in France and Germany (INRA 1990, BBA 1993). In addition, a number of approaches have been proposed in the scientific literature for routine use (e.g. Schafer et al. 1983, Bennett and Schafer 1988, Grau et al. 1992, Kononen et al. 1986, Kononen 1988). More recently, a programme of research in the UK has developed and evaluated methods for testing avoidance of treated seeds (Fryday et al. 1999, 2001), pellets and granules (Hart 2002). Luttik (1993)
suggested that the standard 5-day dietary toxicity test (LC$_{50}$, e.g. OECD 1984) could be modified to provide a measure of avoidance as well as toxicity. However, it has also been argued that various shortcomings of this test make it a poor guide to exposure and toxicity in the wild (Mineau et al. 1994).

5. A variety of methods have been used for testing the efficacy of chemicals for use as repellents to protect crops and other areas from birds. These methods include molecular modelling, cage tests, large pen trials and field studies (e.g. Avery 1989, Avery and Decker 1991, Avery et al. 1993, Clark and Shah 1991, Cummings et al. 1991, Mason et al. 1989a, Schafer and Brunton 1971, Starr et al. 1964). Many of the same principles apply to testing the efficacy of repellents and the ecological safety of pesticides, although there are some significant differences. Perhaps most importantly, a repellent need only be effective against one or a few target species whereas, for a pesticide, the potential for avoidance may need to be considered for a wide range of non-target species. Generally it will not be practicable to test more than a few species.

Are Tests with Captive Birds Reliable Predictors of Avoidance in the Wild?

6. Whether avoidance is effective in protecting birds from pesticides in the wild, and whether it can be reliably predicted from any type of test with captive birds, are matters of controversy. These questions need to be considered before deciding what type of avoidance tests, if any, should be used for risk assessment.

7. In a critique of the 5-day dietary test, Mineau et al. (1994) argued that, for a few relatively well studied cholinesterase-inhibiting pesticides at least, avoidance in laboratory tests is not borne out in the wild. Examples of such chemicals avoided in the laboratory include methiocarb (Dolbeer et al. 1994), bensultap (Edwards et al. 1993), imidacloprid (Grau et al. 1992, Avery et al. 1994a) and fonofos (Fryday et al. 1998, Hart et al. 1999, Hart and Cloock 1994, McKay et al. 1999). In a study of fonofos, woodpigeons (Columba palumba) feeding on arable fields in Cambridgeshire, UK, were counted once or twice per week over four winters from 1992 to 1996. Fewer fields drilled with fonofos-treated cereal seed were used by wood-pigeons than untreated (without fonofos) fields for the first week after drilling only. The extent of the avoidance reaction was related to the concentration of fonofos on the seed left exposed on field surfaces. The density of seed left exposed after drilling, and the residues of fonofos they contained, both varied widely between fields. The study concluded that fonofos-treated cereal seed may be partially avoided in the wild, but that avoidance fails to prevent poisoning under some (rare) conditions, possibly when seed density, concentrations of fonofos and levels of hunger are all above average. Only for methiocarb, used as a bird repellent on fruit, is there strong evidence of avoidance being effective in reducing risk to birds in the wild. A large number of field studies showed that this use of methiocarb caused very few bird mortalities, and that birds took significantly less of treated than of untreated crops. Even in this case the effect of avoidance in reducing risk might be relatively small, since the predicted risk would not be very high even if there were no avoidance (birds feeding exclusively on freshly-sprayed fruit would ingest approximately one lethal dose per day based on acute toxicity).

8. In some cases, strong avoidance shown by captive birds failed to protect wild birds. Diazinon has a high acute toxicity to birds, but is strongly avoided in dietary toxicity tests. In small pen tests where Canada geese were held on treated turf there were no mortalities, suggesting that avoidance was effective in reducing risk. However, a number of poisoning incidents involving waterfowl have been reported following the use of diazinon on turf and in field studies. A possible explanation is that the birds involved in these incidents fed more rapidly than captive birds and consequently ingested lethal doses before developing an avoidance response (Mineau et al. 1994). There was also no mortality in a simple semi-field study with pigeons and fonofos-treated wheat (Pascual and Hart, 1997) although in practical use a low frequency of poisoning incidents has occurred. This discrepancy is most likely explained by (a) higher level of motivation and feeding rate by free-ranging pigeons, and (b) localised areas of higher seed density in practical use, including spills. This is supported by the results of no-choice avoidance studies in which
conditions were manipulated to provide a realistic worst case feeding rate and seed density. Less severe no-choice and choice tests failed to reproduce the mortality seen in the field (Hart et al. 1999, Fryday et al. 1999). Ideally, the test design would take such effects into account.

9. CSL research shows that it IS possible to design tests to demonstrate the effect of feeding rate on avoidance and mortality in the lab. These tests are fairly simple but the conditions need to be adjusted for each new species and feeding rate. Data on feeding rates in the wild are rare: if they cannot be obtained then a (precautionary) option is to design lab studies to achieve a near-maximum feeding rate. The most difficult issue is extrapolating between species as there is currently almost no data to assess consistency of avoidance responses between species under realistic worst case conditions.

10. None of the other case studies provide enough evidence to decide whether risk was substantially reduced by avoidance in the field. In several cases, where mortalities had occurred in the field, a number of factors contributed to uncertainty. Often the unrealistic conditions of laboratory studies (and pen trials) may have made avoidance more likely than it would have been in the field (e.g. using well fed birds, or a highly nutritional food base). There was also doubt as to whether the numbers of poisoning incidents in the wild were as high as would be expected if avoidance were having no effect at all in reducing risk (e.g. dieldrin and heptachlor - see Murton and Visozo 1963). On the other hand, where mortalities had not been recorded, factors other than avoidance (e.g. pesticide residues lower than expected) might have been responsible for the reduced risk. Moreover, mortality may have been underestimated due to poor search efficiency in field studies or under-reporting of poisoning incidents. In very few cases had an attempt been made to measure avoidance behaviour directly in the wild.

11. A comparison was presented between results from avoidance tests using the BBA protocol and the occurrence of poisoning incidents in the wild for the same pesticide formulations (a range of anonymous granular pesticides, treated seeds and baits; see Appendix to this section). Poisoning incidents had been reported for some but not all the formulations which had caused mortality in BBA tests. No incidents had been reported for formulations classified as low risk based on the BBA test. These data suggest that the BBA test provides a conservative measure of risk, although it should be noted that some formulations have been little used (so there has been little opportunity for risk to be realised) and that only a small fraction of incidents are reported. The likelihood of small bird poisonings being picked up by incident schemes is particularly low, so lack of incidents is not necessarily grounds for assuming safety.

12. Considerable uncertainty remains about the extent to which avoidance is effective in reducing risk in the wild. A more detailed review of existing data is required to further assess the relationship between avoidance in laboratory and field. This might best be carried out by an OECD/SETAC expert group, as some of the data are commercially sensitive.

**Factors Which Can Affect the Extent of Avoidance**

13. A wide range of factors have been shown to influence the extent to which treated food is avoided by captive birds, including species (Esapaillat and Mason 1990, Kononen et al. 1986, 1987, Mason and Bonwell 1993, Mason et al. 1993, Schafer and Brunton 1971, Schafer et al. 1983), sex (Esapaillat and Mason 1990), age (Williams, Fairbrother and Sullivan, unpublished data), group size (Hart et al. 1999, Kononen et al. 1986), social interactions (Mason and Reindinger 1982), previous experience (Fryday et al. unpublished, Greig-Smith 1987, Starr et al. 1964), type of treated food and prior food deprivation (Fryday et al. 1999, 2001, Hart et al 1999, Pascual et al. 1999a, Thompson et al. 1981), limitations on feeding time during the exposure period (Hart et al. 1999, Fryday et al. 1999, 2001, colour of the treated food (Greig-Smith and Rowney 1987, Mason and Reindinger 1983), the type of untreated alternative food available (Avery et al. 1995, Rogers 1974), the number of choices available (Bennett 1989a, b), ambient temperature and, possibly, the perceived risk of predation (Avery et al. 1994a), method of presentation (bowl vs. tray,
possibly due to increased vapour pressure over bowl – Fryday et al. 1998), regurgitation (Pascual et al. 1999b). Some species feeding on treated seeds may greatly reduce their intake of chemical by removing the husks (Avery et al. 1994a) but it has been shown in field studies that species which dehusk do not dehusk all types of seeds, and do not dehusk every seed of types they dehusk (Prosser 1999).

14. The relative importance of the different factors differs between acute and longer-term exposures. Acute exposures occur when a lethal dose could potentially be contained in less food than is normally consumed in a single day. In this situation, factors influencing the initial rate of consumption are probably dominant (Mineau et al 1994, Hart et al 1999, Fryday et al. 1999, 2001). These can be broadly divided into factors that influence motivation to feed (e.g. by increasing hunger, or decreasing the time available for feeding) and those that create physical constraints on feeding rate (e.g. long intervals between food items due to low density, concealment, handling time etc.). All of these factors, combined with the concentration of pesticide in the food and any regurgitation, determine the likelihood that a lethal dose is accumulated in the bird before an avoidance response is elicited. In longer-term exposures it takes more than a day to consume a lethal dose. In this situation, acute factors influencing the initial bout of feeding may still be significant, but factors that influence successive exposures (especially learning or conditioned aversion, and the ease of access to alternative, untreated foods), together with the rate of detoxification and excretion, become more important in determining whether a lethal internal dose is reached.

15. The complexity of these factors and their strong context-dependency and species-dependency imply considerable uncertainty in extrapolating from the avoidance responses of captive birds to the behaviour of wild birds. It was concluded that the reliability of the extrapolation should be improved by using test conditions close to those in the wild, rather than using a standardised test design. It has been argued that existing standard test guidelines and surrogate species had proven sufficiently reliable for regulatory purposes in the past. However, it is unlikely that “standard test methods” would have predicted a risk to pigeons from fonofos due to the strong avoidance response that is only broken down by significant manipulation of test conditions to increase the motivation to feed. If the risk in the wild is actually only from spills then it could be argued that any avoidance tests at realistic mainfield densities might not be expected to cause mortalities. However, in the fonofos study the birds had to be strongly motivated to feed before mortalities occurred with spills so this should be considered to be the minimum preparation required for any study if the condition of the birds is to be considered at all realistic. Where a range of conditions may apply in the wild it is considered preferable to select those which are likely to reduce the extent of the avoidance response, to ensure a margin of safety in the subsequent risk assessment. That is, the test conditions should be realistic but tending towards the worst case. For example, choices between treated and untreated food could be made difficult rather than easy, and birds could be placed under at least moderate hunger stress. Experience with pigeons and fonofos has shown that birds must be motivated to feed far more than standard conditions provide if they are to be put at the same risk as wild birds. It would also be necessary to present the test material at a realistic density (worst case) so that the combination of motivation to feed and the physical constraints on ingestion rate (e.g. distance between seeds) combine to provide a realistic worst case feeding rate. Balancing the need for realistic test conditions against the need to minimise suffering by test animals is a major factor in test design (see later).

**Should Avoidance Be Considered in Avian Risk Assessment?**

16. Despite the uncertainty which still remains concerning the extent to which avoidance is effective in the wild, and the difficulties of extrapolation from tests with captive birds, avoidance should be considered in avian risk assessments for the following reasons:

- There are many examples of birds avoiding foods which contain natural chemicals whose repellency has been confirmed in tests with captive birds;
Many pesticides are strongly avoided by captive birds at least in favourable conditions;

If the effects of avoidance were disregarded, potentially valuable products might be unnecessarily refused registration due to over-estimating risks to birds;

Tests of avoidance can be made realistic and conservative to reduce the chance of over-estimating its effect on risk;

Both test design and interpretation of results can be improved by making better use of information on the ecology and behaviour of the species at risk;

Although avoidance may be partial, variable and context-dependent, these factors may be taken into account in assessing its effect on risk.

However, any assessment of avoidance must give proper consideration to the context-dependency and uncertainties mentioned above.

**Chemicals other than Pesticides**

The principles of avoidance testing for pesticides could equally be applied to other chemicals, taking into account the manner in which birds encounter them in the wild (e.g. in contaminated plant material, in other food, or in non-food particles; see later).

**Proposals for Avoidance Tests**

**Principles of the Tests**

**Sequence of testing**

Tests of avoidance would normally only be considered for pesticide uses for which a medium or high risk has been identified in a preliminary assessment. Usually, the concern would relate to one or more scenarios (crop, pest, manner and rate of application, species at risk) which have been fairly closely defined. The sequence by which avoidance testing might proceed is illustrated in Figure 1.

For reasons discussed above, the final assessment of how avoidance may reduce risk should be based on a realistic and severe test. This should be using suitably motivated birds offered treated material at realistic worst case densities for sufficient time that will ensure that avoidance is robust without the risk of mortality due to starvation. This could be very different for each combination of formulation type and test species. Fryday *et al.* 2001 have shown how pre-testing (without pesticide) can be used to identify appropriate test conditions for different species. Because each test will be relevant to a relatively limited range of conditions, two or more avoidance tests may be required to cover the range of scenarios which have led to the preliminary assessment of high risk.

The registrant might proceed to an avoidance test if there is already evidence of avoidance in dietary tests, if the chemical is similar in structure to known repellents, or if the physical properties of the formulation are thought likely to cause avoidance. Equally, if a dietary toxicity study has already been conducted and shows no evidence of avoidance, it is unlikely to be worth proceeding to an avoidance test unless the properties of the formulation are expected to be different.
Figure 1: Flow diagram illustrating proposed sequence for conducting and interpreting tests of avoidance

CHEMICAL IDENTIFIED AS HIGH RISK IN PRELIMINARY RISK ASSESSMENT

No evidence of avoidance

Evidence of avoidance e.g. from dietary test or other feeding studies, or from the literature

CRITICAL TEST(S)

Mortalities

Adverse sublethal effects

Avoidance not a factor in final risk assessment

Avoidance may partially reduce risk

Avoidance probably reduces risk

No adverse effects

Confirm using field studies or Risk mitigation and re-test

Ensure through post-registration monitoring
22. Even if some information on avoidance is already available from a standard dietary test, other simpler tests (e.g. USDA on-choice test for repellents) might be useful for investigating the speed of onset of avoidance and whether avoidance is due to primary repellency or other mechanisms, such as toxicosis or conditioned aversion. This information can assist in refining the risk assessment (comparing time to avoidance with time to ingestion of lethal dose), interpreting the results of the standard dietary study, or interpreting results of the subsequent avoidance test (due to the lack of data on ingestion, or if the avoidance test has been performed with fewer treatment levels; see later).

**Flexibility versus standardisation**

23. The conclusion that the avoidance test should be realistic rather than standardised implies that if the species and situations differ between countries, tests conducted for one regulatory authority may not be acceptable to another. It would also be inappropriate to draft prescriptive guidelines for the conduct of avoidance tests. The difficulty in developing standardised tests limits the progress which can be made towards OECD countries’ objectives of harmonization and mutual acceptance of data, so far as avoidance testing is concerned. Further research is therefore urgently required to improve our understanding of the factors affecting avoidance, and thus our ability to extrapolate between different conditions. Improved understanding should make it possible to develop and validate more standardised methods, at least for particular types of formulation. However, if studies are conducted at “near maximum” motivation and worst-case densities, then they may be sufficiently precautionary to be accepted by different authorities. Alternatively, studies could be conducted with several treatment levels (different levels of motivation and density) so that regulators can decide which result is relevant to them (but this approach implies using more animals). We can probably already describe a minimum set of conditions for pigeons and similar species and treated seed, and provide some preliminary guidance for sparrows, based on Fryday et al. (1999, 2001). One issue that may need consideration is the relevance of densities representing spills of treated seeds.

24. Nevertheless, greater harmonization of the general approach to avoidance testing should be possible. It is hoped that, despite the need for further research on some issues, the contents of this document are of immediate practical use both to regulators and registrants. In particular, it was hoped that those designing avoidance tests might find it useful to consider the proposals in each section below and incorporate relevant aspects into their own study protocols.

**Need for prior information, expertise and consultation**

25. The many factors which have been shown to influence avoidance by captive birds (see earlier) make the task of designing and interpreting appropriately realistic and severe tests a difficult one. It is therefore recommended that, whenever such tests are contemplated, advice should be sought from scientists with relevant specialist knowledge and experience. It is recognised that such expertise will not always be available from the existing staff of either registrant or regulator.

26. For the same reasons, a detailed dialogue between regulators and registrants is essential, especially in the planning stage. Concern was expressed that such dialogue was often not practicable. Nevertheless, the majority view was that regulators and registrants should seek to agree, before the test is conducted, that the test design is appropriate and that, providing no unexpected complications arise, both will accept the results for use in risk assessment. This discussion is also necessary on ethical grounds, as tests should not be conducted unless it is certain that the results will be accepted by regulators for use in risk assessment (to avoid conducting unnecessary tests, and to avoid having to repeat tests due to disagreements about study design).
27. The following range of detailed information would be desirable to assist the design of the test: see Table 1 below.

Table 1: Types of information which may assist the designing of appropriate tests of avoidance

<table>
<thead>
<tr>
<th>Type of information</th>
<th>Possible sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute avian toxicity of pesticide</td>
<td>Basic data set</td>
</tr>
<tr>
<td>Pesticide formulation</td>
<td>Basic data set</td>
</tr>
<tr>
<td>Crop, habitat, time of year, weather</td>
<td>Preliminary risk assessment</td>
</tr>
<tr>
<td>Species at risk</td>
<td>Preliminary risk assessment</td>
</tr>
<tr>
<td>Manner of exposure (food type/granules, etc.)</td>
<td>Preliminary risk assessment</td>
</tr>
<tr>
<td>Evidence of avoidance</td>
<td>Avoidance observation and monitoring during dietary tests</td>
</tr>
<tr>
<td>Pesticide residues in food or other material</td>
<td>Preliminary assessment (e.g. 'Kenaga nomogram'), preferably augmented by data on actual residues</td>
</tr>
<tr>
<td>Pesticide half-life in appropriate media</td>
<td>Other parts of registration data package e.g. safety to consumers</td>
</tr>
<tr>
<td>Ecology and behaviour of species at risk, e.g. typical diet, foraging range, special factors (e.g. migrants)</td>
<td>Wildlife experts and scientific literature</td>
</tr>
</tbody>
</table>

28. Before conducting an avoidance test, an attempt should be made to estimate whether the birds at risk are likely to ingest a lethal dose of the pesticide before the onset of severe symptoms, such as would preclude further feeding. If this appears likely, then it may be best to conclude that avoidance will not be effective in protecting the birds and to forego the avoidance test, unless strong primary repellency (e.g. taste aversion) is demonstrated in other tests. The calculation requires an appropriate LD$_{50}$, information on the time it takes for severe symptoms to develop, an estimate of the maximum levels of residues to be expected, and estimates of the maximum rate of feeding. A calculation of this sort was successful in predicting the occurrence of poisoning incidents involving waterfowl feeding on diazinon-treated turf, whereas laboratory and semi-field experiments indicated avoidance (Mineau et al. 1994).

29. Tests of avoidance with birds are usually conducted under relatively stress-free conditions (e.g. single birds, food freely available during daylight hours) and may not represent the behaviour of hungry wild birds which may be less easily deterred from feeding on treated seed. CSL research has shown that test conditions designed to increase the motivation to feed in pigeons (e.g. reduced available feeding time, food deprivation and increased group size) were necessary to produce a level of risk that is known to occur in the field (i.e. mortality among at least some exposed individuals). This showed that tests conducted under ‘normal’ conditions would under estimate risk by indicating that birds would strongly avoid fonofos treated seed. It could be argued that as these tests were conducted under spill conditions
(seeds on a tray at high density) they do not represent the risk provided by good practice. However, this is to miss the point that these were the minimum conditions under which birds were put at the same risk as wild birds (i.e. that at least some would eat sufficient seed to be poisoned). We do not know if wild birds are only poisoned at spills or whether normal field surface densities could present a risk – if the latter is true, even these manipulations might not be sufficient to represent wild birds. It is not being recommended that tests should be conducted at spill densities (unless the risk from spills is specifically being addressed), but that even though tests should be conducted at the realistic densities expected in the field, the birds should be prepared for the test as described. Only then can we have confidence that any avoidance response is not merely due to low motivation to feed compared to wild birds but due to the physical constraints placed on feeding rate by the low seed density.

30. Even under these test conditions there are several reasons why even these manipulations may not fully represent the motivation of wild birds:

   a) Wild bird mortalities were not only caused by spills (not known).

   b) Residues on seeds were lower in the wild (some field evidence for this).

   c) Proportion of wild birds exposed that were poisoned was higher (not known).

**Description of the Avoidance Test Method**

**Objectives**

31. To test whether mortality or adverse sub-lethal effects occur when birds are exposed to the pesticide in realistic conditions, with the purpose of deciding whether risk will be reduced to an acceptable level by avoidance in the wild.

**Test design**

32. Conditions should be realistic but also sufficiently worst-case to provide the degree of protection that regulators consider appropriate (e.g. if the conditions are "typical" or "average" then the result will underestimate risk in approximately 50% of real exposures). The requirement for realism implies that the test design must be very flexible, and that the results will be specific to a fairly narrow range of conditions. Expert knowledge and close dialogue between registrant and regulator are essential (see earlier).

33. In general, the study should be designed to offer birds the test material spread on the floor of a sufficiently large aviary, in a manner which reflects the exposure scenario of concern, for the duration of natural daylight over 4 days. For acutely toxic material a one day test is sufficient. For less toxic materials an alternative design (e.g. choice in time) should be considered, giving more weight to factors such as learning and detoxification. A range of treatment concentrations may be tested, plus a control treatment. Formulations should be used, not technical active ingredients. Positive and negative control treatments using other pesticides could be included to confirm that the test conditions are capable of distinguishing hazardous compounds which are avoided from those which are not. However, for ethical reasons positive and negative controls should not be used in every study (see below). The primary endpoints should be intake, mortality and sub-lethal effects so every effort should be made to ensure that some measure of consumption can be made. A pre-treatment phase would be needed to identify the level of diet to be provided in the treatment phase, to acclimate the birds to the diet and, if consumption is to be measured in the treatment phase, to provide a pre-treatment measure of consumption.

34. There are three broad types of exposure scenario for which radically different variations of this general design would be required:
- non-food items including most granular formulations*: this may also include coated or pillorized seeds if these are reliably treated as non-food items by birds;
- food items that are difficult to handle in a laboratory diet, such as growing plants (foliage);
- other food including treated seeds and baits and natural foods other than growing plants.

* note that some granules may be mistaken for small weed seeds (Hart 2002, SCP fosthiazate opinion).

Non-food items

35. These are materials with negligible nutritional value that are likely to be ingested incidentally or as grit. Note that granules which have some nutritional value (e.g. those with a corn cob base) should not be included in this category. During the treatment phase, birds should be presented with the test material spread on a soil surface. An adequate ration of suitable food would be spread on the same surface.

Foliage

36. Growing plants could become contaminated by spray applications, and may be ingested intentionally by birds. The possibility of bringing sufficient leafy material into the laboratory depends on the size of the avaries, the toxicity of the compound, the duration of the test and the size and numbers of birds involved. Where this is not possible it may be necessary to consider pen studies. A pen would be brought to the treated field and placed over a test site, and birds would be introduced to the pen for feeding (e.g. Fink 1979, Robel and Morrow 1987). It must be noted that there are doubts about these tests’ ability to represent birds' responses in the wild (e.g. for diazinon, Mineau et al. 1994). There is a need for research to refine and validate this method. For example it would be necessary to condition birds to appropriate level of motivation, as in laboratory studies. Omitting this probably explains the failure of the diazinon pen studies to simulate field mortalities.

37. Some degree of exposure by non-dietary routes (inhalation, dermal absorption) is possible in the avoidance test. If it is thought that exposure by these routes might contribute significantly to the overall risk, then the test design should ideally be modified to ensure that those contributions are realistic, unless they can be taken into account in other ways.

38. Comments in the following sections are considered relevant to all three categories (non-food, foliage and other food) except where otherwise indicated.

Test subjects

39. Ideally, the species most likely to be at risk would be used in the test. However, there will often be a range of species at risk, most of which it would be impractical or unacceptable to test in the laboratory. It would therefore usually be preferable to use a representative surrogate species. Careful consideration should be given to the use of domestic stock of some wild species (e.g. waterfowl) as they may have substantially greater body reserves and therefore be much more able to undergo prolonged fasting in order to avoid ingesting unfamiliar food items.

40. The species should be ecologically relevant. For example, common species that can acclimate to captivity may be selected to represent corn field birds (quail, ring-necked pheasants, red-winged blackbirds, sparrows).

41. The species should be metabolically relevant, i.e. of similar body size and diet to those at risk. Small-bodied species should be preferred for testing granules since, for some chemicals, they may require
only a very small number of granules to receive a lethal dose, which restricts their opportunity to develop learned avoidance. Care should be taken to select species that select grit particles of sizes which overlap the size of the pesticide granules. These will typically be the smaller bird species (Best 1992).

42. Body condition should be equivalent to that at the time of year when exposure could be expected. However, achieving a target level of body condition may be problematic. When attempting to reduce body condition by food restriction this may not have the desired effect as birds may limit weight loss (e.g. Pascual et al. 1999ab). Particular care will be necessary in simulating the body condition of birds during migration, where this is relevant (e.g. in the case of waterfowl on turf).

43. If the principal endpoint is toxic effects or mortality then it is essential that the species is pre-tested to check its LD$_{50}$ (approximate LD$_{50}$ is sufficient). This should be used to check whether the LD$_{50}$ is low enough for the animal to get a lethal dose in one day’s food. If this is not the case, then the lack of effects in the study would be interpreted as avoidance when in fact it is simply low sensitivity. This is very important because it would not tell you whether more sensitive species would show avoidance.

Accommodation (housing)

44. The number of animals per pen should take account of the group size which is ecologically relevant for the test species, statistical considerations, the need to measure individual consumption (requiring individual housing), and practical limitations on pen size. Group size should not be chosen simply on size of bird (e.g. large birds housed individually and small birds in groups).

45. Pen size would be dependent on species, number of animals per cage, the area required for sufficient spreading out of the test material (see later), and animal welfare considerations. In testing the repellency of treated seeds by comparing bird use of treated and untreated plots within a large flight pen, Daneke and Avery (1989) found that larger plots (78 or 108 m$^2$) were more readily discriminated and gave more consistent results.

46. In the case of foliage, suitable plants could be grown and sprayed in permanent aviaries. The size of aviary required is obviously dependent on the toxicity of the material, the type of plant, the duration of the test and the size/numbers of birds involved. In practice, the nutritional value of such foods may be too low for adequate supplies to be grown in the space available. In this case temporary pens may be constructed on field plots containing growing plants which have been sprayed with the test material. However this type of pen study has its own set of practicality issues.

47. It will usually be more practical to use outdoor pens, due to the space required. Ambient conditions will influence results, so concurrent control birds will be needed (see later).

Time of year

48. Seasonal factors could affect bird condition and behaviour, particularly feeding rates. Consequently, it may often be preferable to conduct the tests at the time of year expected in real exposures. However, simply conducting studies at the correct season will have little effect if birds are unstressed and fed ad lib. It may be sufficient to manipulate other conditions (e.g. deprivation) to ensure motivation and feeding rate simulate exposure at the relevant time of year.

Preparation of animals

49. A five- to seven-day period for acclimation to test conditions, or longer if required to stabilise daily food consumption, was recommended. In practice longer periods of up to several weeks will often be
required to ensure that birds are in suitable condition for testing (Fryday et al. 2001). This is even more important for birds obtained from the wild.

50. Food presented during acclimation may be standard aviary diet. In tests for foliage and other foods, it may be desirable to acclimatise birds to the alternative food to be presented in the treatment phase (see below) if this food is a normal constituent of the diet in the wild. It is essential to pre-acclimatise birds to uncontaminated versions of the test food if they would (in a realistic worst case) encounter it in the field. Otherwise risk may be under-estimated due to a simple neophobic response that might not occur in the field. This applies to treated seeds, contaminated natural seeds, contaminated plant material and contaminated invertebrates. In addition, if wild birds could encounter seeds and/or grit of similar sizes and appearance to a granule to be tested, then they should be acclimatised to them before the test period (see Hart 2002). This might also apply to some types of pellet or coated seed. In the case of treated seed those designing tests should consider whether seed presented in the acclimation phase should be dyed (but untreated) if most seed available in the wild was dyed. This might be necessary in some cases to prevent responses to seed colour having more influence in the test than they would in the wild. This is an issue which requires further research. Wild caught animals also need acclimatisation because their experience in the wild is uncertain. These issues require case-by-case consideration. Also, any pre-acclimatisation needs to be clearly documented in the study report so that others can consider it when using the results.

51. Birds should be fasted for an appropriate period every night to simulate normal diurnal restrictions on feeding activity. Wild birds usually have higher energy requirements than captive birds and may show higher consumption rates, which would increase the risk of ingesting a lethal dose before developing an avoidance response (Mineau et al. 1994). Artificially shortening the feeding period (e.g. Rogers 1974, Hart et al. 1999, Fryday et al. 1999, 2001) may provide a means of increasing consumption rates to simulate such effects. The feeding period may therefore need to be further restricted to simulate limitations on feeding which wild birds experience during daylight hours, for example as a result of predation risk, human disturbance or environmental factors (e.g. bad weather or snow cover). CSL research (Hart et al. 1999, Fryday et al. 1999, 2001) shows that to achieve realistic worst case feeding rates, preparation of birds should include calibrated amounts of fasting and manipulation of feeding periods prior to the exposure (test) period.

Test material

52. As the emphasis is on realism, the test material in the avoidance test would usually be the formulation. This implies that separate tests may be required for different formulations, unless there is good reason to believe the results would be equivalent.

Avoidance Test Procedure

Treatment levels

53. Consultation between regulator and registrant is particularly important in deciding the number and setting of treatment levels because of their effect on the cost of the test, the number of animals required, and on the potential to extrapolate results to other concentrations relevant to the risk assessment.

54. It is expected that a bird would encounter a range of doses in the wild. This can occur because of patchy spray applications, variability in the formulation process for seeds (e.g. Jahn 1991), granules and baits, and decay of residues over time. Even if the maximum residue concentration is repellent, lower concentrations may be sufficiently palatable for the animal to ingest a lethal dose. A range of concentrations should therefore be selected to provide a reasonable chance of detecting such a hazard,
taking account of the range of residue levels expected in the field. It may be desirable to collect residue data from the field specifically for this purpose if existing data and estimates are inadequate.

55. The range of concentrations should include the maximum expected in the field. This may somewhat exceed those predicted for the recommended application rate, to allow for variability in application techniques (this may apply to seed treatments, granules and baits as well as sprays).

56. Up to four treatment concentrations might often be appropriate.

57. There may be some scenarios that present the possibility of sequential exposure to the same product, possibly at differing concentrations, for example when it is used in several fields in one locality. Research has suggested that, for some compounds, sub-lethal exposure may reduce ability to avoid subsequent exposures (Bussiere et al. 1989). It was suggested that if there were cases where this was a significant concern, treatments might be offered in a 'flip-flop' design, with some birds receiving a high concentration followed by a low concentration and other birds receiving the treatments in the reverse order. However, it would be difficult to determine what combinations would actually be encountered in the wild.

Controls

58. There should generally be one untreated control treatment, in which birds are kept under identical conditions to those in other treatments (including diet) except that the test material is omitted. This is required in order to assist in the interpretation of mortality and sub-lethal effects seen in the other treatments, and to determine whether the test conditions are biasing the test by deterring consumption even in the absence of the test material. In the case of non-food items, if the items are the formulation (e.g. granules) they should be completely omitted rather than being replaced with untreated items. The same number of replicates should be performed for the control as for other treatments, and the same observations should be made.

59. There is also possibility of using positive and negative controls to confirm that the test was capable of correctly distinguishing between chemicals which would be avoided in the wild (e.g. methiocarb) and those which would not (e.g. dieldrin, anticoagulant rodenticides). However, provided the recommended test design were successfully validated (see later), it would be preferable not to include such controls in every test for reasons of welfare and economy. However, in some cases the test design may have to be modified so much (to represent particular risk scenarios) that it requires additional validation. If negative or positive controls are to be included, their nature should be the subject of detailed discussion between regulator and registrant as they would be crucial to the interpretation of the test results.

Fasting

60. Birds should continue to be fasted for an appropriate period every night to simulate normal diurnal restrictions on feeding activity. Depending on the design of the study any restriction of the feeding period may also need to be continued. The degree of fasting appropriate will be dependent on the size of the bird tested and needs careful calibration along with other factors to achieve relevant feeding motivation/rates and prevent risk of starvation (see Fryday et al. 1999, 2001 and Preparation section above).

Duration of treatment and post-treatment periods

61. The treatment period should be at least long enough to ensure that, if the test material is not avoided, a lethal quantity will be ingested. This may require a short term acute test of only one day (perhaps only a few hours for small species) or more long term chronic studies lasting several days using a
choice in time approach. Longer test periods may also allow detection of delayed effects (e.g. due to mode of action or partial avoidance), delayed learning or, in the case of an early response, habituation. Even if treated food in the wild is typically available for only a short period on any one field, it is likely to be available over longer periods on neighbouring fields within the foraging range of individual birds. The relevance of each for a particular pesticide would be based on a comparison of available LD\textsubscript{50} data with predicted exposures. Some pesticides might require both acute and chronic tests.

62. Test diets should be presented for the whole of the feeding period each day. Non-food items could be made available continuously.

63. Post-treatment period: Birds should be maintained on untreated food for a post-treatment period of 4 days to allow for observation of any delayed symptoms of intoxication, or longer if the test material is thought to be capable of causing more delayed effects.

Early termination of test on welfare grounds

64. In the interests of animal welfare, the test should be terminated early if at any point sufficient mortality or other adverse effects have occurred to enable the regulatory decision to be made (i.e. to demonstrate that avoidance should not be a significant factor in risk assessment; see later).

65. In addition, it would be desirable to remove individual subjects if they become moribund or suffer convulsions.

Presentation of test material

66. For the purpose of realism, the test material should be presented in a way which is representative of what the animal is expected to encounter in the field.

67. In general the test material should be spread out rather than presented in pots, hoppers or small trays. For treated seeds, pellets and granules, density is probably a key factor limiting the frequency of acute mortality in the field. Choice of test densities therefore requires careful consideration of (a) what occurs in field, (b) what level of protection is required (how worst case), (c) relevance of spills, (d) practical limitations due to size of aviary. Where limitations on space are constrain the density of the test material to much higher levels than is usual in the field, it must be low enough to avoid vapour pressures of the test compound significantly higher than those in the field. Avoidance of higher vapour pressure would also be assisted by conducting the test outdoors, or otherwise ensuring high levels of air exchange. A re-analysis of existing data on consumption in LC\textsubscript{50} tests, where test diet is presented in pots, shows a significant negative correlation with vapour pressure, suggesting an element of avoidance which is unlikely to occur in the field (Fryday \textit{et al.} 1998). Note that, especially if consumption is not measured, area x density needs to provide sufficient treated material to provide toxic doses for all subjects in the group. Otherwise it will be uncertain whether lack of mortality is due to avoidance or simply to insufficient test material.

68. Foliage should be presented as growing plants, as discussed earlier.

69. For realism other food types and non-food items may be presented against a background representative of those expected in the field such as soil of an appropriate type, or turf. However, this may not always be appropriate for the aims of the study. The key factors to consider are (a) what influence the background has on feeding rate, and (b) the high desirability of enabling recovery and re-weighing of unconsumed material. If an appropriate realistic worst case feeding rate can be achieved on a simple artificial background (e.g. metal tray or concrete floor) which facilitates re-weighing, then this would be preferable. This has been found to be satisfactory for treated seeds (Fryday 1999, 2001) and pellets (Hart...
For granules, detection against a naturalistic background is more important. One approach is the use of controlled mixtures of fine sand and grit that enabled recovery and reweighing of uningested granules (Hart 2002). However, this method is laborious, and the colour of the background material would need to be varied to provide a realistic worst case in terms of contrast with particular types of granules. Consideration should also be given to grit content, moisture content etc. when selecting the background for a particular study. These also influence the extent to which the scenario is worst case – e.g. moisture may affect residue losses, and low grit availability increases the chance of granule ingestion. In addition it is vital that full details of the background substrate are reported in the study report, so that others can evaluate the severity of the test conditions.

70. In the case of non-food items, an appropriate untreated diet should be scattered onto the same area as the test material. The surface may be clay silt or peaty soil, depending on what is relevant for the crops on which the pesticide is used. Food and treated material should be spread separately to avoid contamination of the food (BBA 1993). The food should be representative of the birds’ diet in the wild. In particular, a seed diet should be provided for granivores as it may increase the use of grit and the release of active substance from pesticide granules in the gizzard.

71. In tests of granular pesticides, treated granules should be spread on a soil floor. The density of granules and their incorporation into the soil is a matter requiring careful consideration and agreement between registrant and regulator. For example, as a reasonable worst case the density might be set equivalent to the maximum application rate for field use, assuming no incorporation into the soil. Although this would be higher than the surface densities achieved by many methods of application, it would be lower than occurs in spillages and when turning at row ends. In addition, it may be desirable to use artificially high densities to increase the chance of detecting accidental ingestion of granules (mortality of a few per cent might be considered significant in the wild, but would be unlikely to be detected in a relatively small test group).

72. The soil used for the background in tests with granular pesticides should also be considered carefully. If it contains large proportions of grit (e.g. sandy soils) this may reduce intentional ingestion of the test granules. Grit content should represent the lower end of that likely to be encountered when the granules are used in the field, such as a clay silt. Bird preferences for different grit types have been the subject of a number of research studies (Best 1992, Best and Gionfriddo 1991, 1994a, 1994b, Gionfriddo and Best 1995).

73. Ideally, test materials should be removed and replaced daily to ensure consistency in test conditions between days. Alternatively, birds could be moved to freshly prepared cages or given access to different areas of the same cage on successive days. A possible exception is granules, if they are provided in great excess and do not suffer significant decay of the toxic ingredients. To facilitate removal of test materials other than granules, consideration could be given to offering them on large trays or other hard surfaces, provided it is agreed that the decrease in realism is unlikely to affect the result.

**Provision of grit**

74. For species that feed primarily on seed, the subjects may require grit to break down their diet. If none is available, the birds may be unable to maintain normal weights and may not absorb as much of the active ingredient as they would in the wild. In general, therefore, an adequate supply of grit should be provided either in soil or separately (e.g. in pots).

75. In tests of grit-like non-food materials the only alternative grit should be that available in the soil (see above).
76. Birds deprived of a grit source will tend to conserve the grit remaining in their gizzards. If their grit load becomes depleted, however, they will increase their rate of grit consumption when a new source is presented. Consideration needs to be given to whether grit should be withheld for a period prior to tests with granules to make these tests more severe. At least it seems prudent to avoid offering an unrealistic excess of grit (e.g. presented in pots) during the pre-treatment period. See Hart 2002 for experimental studies of the effect of grit deprivation on granule ingestion by partridges and sparrows. This also includes a recommendation that to achieve a realistic worst case, if wild birds could encounter seeds and/or grit of similar sizes and appearance to the granule to be tested, then they should be acclimatised to them before the test period.

Availability of alternative food during tests

77. This is a crucial aspect of test design. The availability of alternative foods varies widely in the wild. There will nearly always be some alternative to contaminated food in the wild, but it may be unpalatable, far away, difficult to find, etc.

78. Standard dietary toxicity tests provide no choice, which some consider unrealistically severe. However, the severity of the existing dietary study is due to the duration of the no-choice period (five days). For acute exposures, no choice is a realistic worst case (especially for the critical first bout of feeding). For longer-term exposures, some degree of choice is likely but in the wild the alternative food will often not be in the same place as the treated material and will require the birds to move to another patch. Simultaneous provision of treated and untreated food (mixed or in adjacent areas) under-estimates risk in this scenario. Providing realistic distance between treated and untreated foods is not practical in test conditions, so we consider the "choice in time" approach as the most reasonable compromise for chronic avoidance tests. The timing of treated and untreated feeding periods for a choice-in-time approach would vary with size and resilience of the species and would require careful calibration to achieve realistic severity without significant starvation risk.

80. Another concern about simple two-choice tests was that, if the alternatives were presented in two large, separate patches, the birds may make very few active choices at the start of each feeding bout. It is preferable for alternatives to be presented in mixtures or chequered arrangements, which would require the animal to make more choices, increasing the severity of the test. However, the realism of this is debatable; choices between treated and untreated fields may be more like the simple two-choice test, but variation in application rates within fields may be high enough to resemble the mixture presentation.

81. Presentation in mixtures may has been used in the past on the basis that it increases the severity of the test by making it difficult for birds to selectively avoid test materials unless they have significant primary repellency, such as an aversive taste, or at least a distinctive taste which facilitates secondary repellency (conditioned aversion). However there seems to be no hard evidence for this and it is possible that the use of mixtures will in fact decrease severity because the exposure is diluted by the untreated portion – this also means that a larger amount of food has to be presented to ensure a lethal dose is available, and raises the possibility that mortality may be prevented simply by the dilution of the exposure rather than by avoidance. If the aim of the test is to determine whether there exists an avoidance response strong enough on its own to deter feeding then reducing initial exposure will make it difficult to measure this. In work with pigeons and fonofos treated seed (Hart et al. 1999) it was found that several factors affect the risk of poisoning of birds including availability of freshly drilled fields, low preference for these fields in relation to other feeding sites and lower residue levels along with avoidance of seed when
encountered. Perhaps this test should just address the strength of the avoidance response with other factors addressed by looking at the ecology of the species separately to determine how often they are likely to encounter the treated material. Providing multiple choices in an attempt to be realistic will likely hamper attempts to measure the avoidance response alone and in any case is unlikely to be realistic given the small scale on which tests will be conducted in terms of the distribution of alternative foods in the real world.

82. In most cases the presentation of mixtures would also make it impractical to measure consumption of treated and untreated material separately (see below).

83. It is therefore recommended that no-choice designs be used with the duration of the no-choice period adapted as necessary for the type of test (acute or chronic) and the metabolic requirements of the test species (e.g. small birds).

Non-food items

84. There is no need to provide untreated granules. Alternative grit would be available in the soil substrate (see above), interspersed with the test granules.

Foliage

85. If an untreated alternative is provided, then it and the treated plants would be presented in two or more separate plots within each aviary. A chequered arrangement would increase the frequency of choices made by the birds, but is difficult to arrange. It may be preferable to ensure severity of the test by providing no alternative to the treated foliage. Provision of standard aviary diet in hoppers, etc. as an alternative would probably reduce severity too far, unless the amount offered were substantially less than the normal daily intake. Research is necessary to investigate these alternatives.

Measurements

86. Measurements which should be made in the test include body weight, feed consumption, mortality, post-mortem investigations, sub-lethal effects, environmental conditions, and residue in feed (see Table 2 hereafter).
Table 2: Measurements which should be made in the avoidance test

<table>
<thead>
<tr>
<th>Measurement</th>
<th>When measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body weight</td>
<td>End of acclimation/beginning of pre-treatment; end of pre-treatment/beginning of treatment period; end of treatment period (all at the same time each day)</td>
</tr>
<tr>
<td>Consumption of untreated and treated food, corrected for spillage</td>
<td>Daily during acclimation period to assess stabilisation of intake; once or twice daily in treatment period; during post-treatment period if required</td>
</tr>
<tr>
<td>Mortality</td>
<td>As it occurs; especially note exact day and time and cause of death, as revealed by appropriate post-mortem investigations</td>
</tr>
<tr>
<td>Post-mortem investigations</td>
<td>Information on the pathological macroscopic findings from post-mortem investigations (e.g. stomach and crop content)</td>
</tr>
<tr>
<td>Sub-lethal effects</td>
<td>Signs of overt toxicity and behavioral changes such as ataxia, wing droop or general debilitation as they occur, plus more detailed measurements if required to assess mechanism of avoidance</td>
</tr>
<tr>
<td>Residue in feed</td>
<td>At intervals sufficient to detect chemical deterioration</td>
</tr>
<tr>
<td>Environmental conditions (including temperature and daylength)</td>
<td>Daily (minimum and maximum temperatures)</td>
</tr>
</tbody>
</table>

87. Observations for mortality and sub-lethal effects must be made at the end of each day. Additional observations should be made during the day to enable prompt removal of subjects on animal welfare grounds should this prove necessary. This requires that means are available for making observations without disturbing the birds, e.g. from a hide or blind. Methods for observations for sub-lethal effects could be guided by data on effects recorded in acute, chronic and reproductive studies if these are available.

88. If the test is outdoors, then you should report weather conditions. Considerations are needed in the study plan as to what weather conditions would be unacceptable for the test period (e.g. heavy rain).

89. Total daily food consumption should generally be measured during acclimation and pre-treatment periods to assist in deciding quantities to offer during the treatment period (see above).

90. Total daily food consumption should be measured in the test period to confirm that subjects are feeding normally under the test conditions and to provide a means of detecting generalised reductions in intake (anorexia). If direct measurement of intake is impracticable, feeding behaviour may be monitored by remote observation or using video-recordings, although the correlation between feeding movements and ingestion may be weak and could differ between treated and untreated groups. If this is the case then weighing birds could be considered as a means of detecting any significant reduction in feeding.

91. Daily ingestion of treated and untreated materials in the treatment period should be measured separately if at all possible to quantify exposure and confirm that survival/lack of symptoms is indeed due to avoidance and not just low sensitivity. For mixtures this may require sorting of mixed test materials by hand or other means. Use of no-choice and choice-in-time designs should make measurement feasible for most materials. This could also be considered for most other types of material (including pellets and
pillorized seeds). While recovery of granules would be more difficult methods have been developed that enables recovery of granules from an artificial soil background (Hart 2002). It may be sufficient to sort a sub-sample from each cage, and to do so only on selected days. If such subsamples are taken, it should be done in a way that ensures they are representative of the whole presentation area (e.g. by thorough mixing of the whole quantity before subsampling), in case birds foraged unevenly.

92. Other methods of assessing relative intake could be considered. Chemical markers (Savarie et al. 1992) might be used to assess the proportion of test material remaining, provided the markers were known not to affect palatability or toxicity. Alternatively, physiological biomarkers (e.g. cholinesterase inhibition for organophosphates) might provide a measure of intake of the test material, if they could be adequately calibrated. However, such markers are unlikely to detect subtle differences in intake and large differences can probably be detected in other ways (e.g. behaviour).

Endpoints

93. In this test the primary endpoints are intake, mortality and sub-lethal effects, which are used to assess whether avoidance may contribute to reduction of mortality in the wild. Interpretation of the results will require expert knowledge (see later). Without measurement of consumption in some way it is not possible to confirm the role of avoidance. No adverse effects could be due to low sensitivity or low concentration.

94. If no mortality or sub-lethal effects occur, data on body weights and ingestion may be helpful in assessing whether the reduced risk was due to avoidance rather than some other cause (e.g. if the pesticide is significantly less toxic to the species used in the avoidance test than to species used in standard toxicity tests).

Statistical aspects

95. Those conducting tests should obtain expert statistical advice on both experimental design and data analysis. In addition, it would be desirable for regulators and registrants to agree in advance on the role of statistical significance in interpreting the results (see later) and the degree of statistical power required.

96. For continuous data (e.g. body weights, consumption) 5-8 replicates (i.e. cages per treatment) would usually be adequate. This number is unlikely to be adequate for testing the statistical significance of differences in frequency data (e.g. mortality, qualitative symptoms of toxicosis) unless the number of animals per cage is substantial (5-10) or the differences are very large.

97. Stratification or blocking of subjects should be used if this helps to improve precision and reduce the numbers of animals required.

98. Comparisons between formulations (e.g. for development purposes or in comparative risk assessment procedures) would pose different statistical requirements than examining only one formulation.

Interpretation

99. Key outputs to necessary to aid in interpretation of avoidance tests include the following:

a) an assessment of extent of avoidance and influence on effects within the study.

b) characterisation of the conditions under which sublethal and/or lethal effects occurred in the study.
c) extrapolation of a and b to other conditions and other species.

d) discussion of uncertainties affecting the above.

100. It is essential to have some way of reliably distinguishing toxicant-induced mortalities from any background mortality which might occur in the test conditions (though this should be negligible if healthy birds are used). Where possible, the cause of death should be determined by appropriate post-mortem investigations and residue analysis as would be used in post-registration monitoring to confirm exposure. If background mortality cannot be eliminated and cause of death cannot be reliably determined, then the decision will have to be based on a statistical comparison of mortality in treated and control groups. This should be avoided, as it is likely to greatly increase the number of animals required.

101. Avoidance may remain a factor in the risk assessment despite mortalities in the avoidance test if the frequency of mortality is low, and if there is direct evidence (e.g. from consumption data) that it has been reduced by partial avoidance. However, the level of mortality in the test (a) is relevant only to corresponding conditions in the field, and could be greater or lower in other conditions, and (b) is subject to substantial sampling uncertainty due to individual variation, unless the number of birds tested is large. Therefore deductions about the level of mortality in the field should be made only with great care.

102. Consideration should be given to whether sub-lethal effects are adverse (toxicosis) or the consequences of potentially adaptive responses (e.g. reduced food intake and associated loss of body weight). Results on consumption in other tests (e.g. dietary test) may assist in making this judgement. The severity of adverse effects could be taken into account in two ways: would they be unacceptable in themselves? Note: for interpretation of acceptability of sublethal effects they should also be interpreted considering ecological consequences of effects that lead to a higher predation risk for prey animals such as general debilitation or impaired motility (secondary or indirect mortality) and what do they imply regarding the extent to which the risk of mortality is reduced by avoidance? Adaptive effects and adverse effects judged to be minor would be treated as 'no effects' in the table above.

103. Evidence on the mechanism of avoidance (e.g. from the dietary test) may be taken into account, but requires particular care and expert interpretation. Avoidance may be caused by primary repellency (reaction to sensory stimuli such as bad taste or smell) or a learned aversion which develops after post-ingestional illness. The first may be overcome by habituation or hunger, whereas the latter may not have time to develop if the speed of ingestion is high (e.g. due to hunger) or the source of the toxicant cannot be distinguished. More weight might be placed on avoidance for chemicals shown to operate through both mechanisms (at realistic exposure levels). It is essential that consideration is given to whether it might be possible in some circumstances for rapidly-feeding birds to ingest a lethal dose before developing an avoidance response. Generally this should be considered in advance and should be addressed by ensuring that the study conditions are sufficiently worst-case to provide the level of protection required by regulators. This includes having a separate, no-choice design for acute exposures, in which conditions are calibrated to generate near-maximum (but realistic worst case) rates of feeding; and ensuring that hunger and sensory cues for the contaminated material represent a realistic worst-case in the chronic (choice in time) study. If the possibility of such an effect is recognised after avoidance testing, despite a 'no effects' or 'sub-lethal effects' test outcome, the regulator would be justified in asking for further tests with the relevant species and conditions. Further research is required to identify the circumstances under which such effects may occur.

104. Those interpreting the results of the avoidance test should be alert to the possibility that exposure may be due in part to non-dietary routes (inhalation, dermal absorption), as noted earlier. This possibility would seem most likely for birds feeding on sprayed plants (or on invertebrates on or amongst sprayed plants) or walking on soil with high granule densities. Preening of contaminated feathers may lead to
further oral exposure (Driver et al. 1991). This possibility would seem most likely for birds feeding on sprayed plants (or on invertebrates on or amongst sprayed plants) or walking on soil with high granule densities. These effects may be more prolonged due to slower uptake via the dermal route (Driver et al. 1991, Henderson et al. 1994). Mineau (2002) found that for organophosphate and carbamate pesticide sprays, estimates of field mortalities were significantly improved if dermal and inhalation toxicities were taken into account. Research is needed to explore how best to adjust study design to include realistic dermal exposure.

105. The avoidance test should be designed to represent a specific use scenario, usually one which led the regulator to consider that a high risk might occur. It might be desirable to extrapolate the result to a range of other conditions, for example:

1) other avian species;
2) other concentrations of active ingredient in test substrate;
3) other substrates (including new formulations, where substrate is a formulation);
4) other crops;
5) other environmental conditions.

106. Type (1) above is nearly always necessary to some extent but our knowledge base for doing it is very sparse indeed. There are few if any examples of multiple species being tested for avoidance with the same pesticide under comparable, realistic conditions. It may be that the ratio between the threshold doses for avoidance and lethality is similar across species, but this has not been tested. Furthermore, the magnitude of this ratio that is required to prevent mortality depends on many ecological and behavioural factors (e.g. feeding rates) that also vary widely between species. To derive reliable generic factors for extrapolating avoidance would require avoidance data for many species and pesticides (comparable to that used by Luttik and Aldenberg 1997 to derive extrapolation factors for the LD₅₀). This will not be available for a long time (if ever). In the meantime however we still need some basis for confidence that, when avoidance is a significant factor, our assessment of it provides adequate protection for the range of species exposed in the field. In the absence of reliable means of extrapolation the only alternative is to test enough relevant species to satisfy the risk manager that other species are adequately protected. We should consider applying statistical methods to quantify this (in effect, an SSD for the effectiveness of avoidance). Protection against unexpectedly large species differences may be provided to some extent by requirements for field tests or post-registration monitoring, as discussed above but passive post registration monitoring is probably not efficient enough for this, as it would usually take some time for a body of evidence to build up. Type (2) is also often necessary. It may be achieved by interpolation between the treatment levels in the avoidance (provided they are not too widely separated), or by considering the results of the avoidance test in conjunction with results of other tests (if these were done, and included a wider range of concentrations). Types (3) and (4) might be desirable to assess a change in formulation or use proposed by the registrant subsequent to the original test being performed. In such cases the regulator would be justified in requesting a new test if the new use or formulation were sufficiently different to make extrapolation unsound. Type (5) should be considered at the outset. Extrapolation between environmental conditions may be unnecessary if test conditions are calibrated to generate near-maximum feeding rates for each species. If the range of conditions is too wide to be represented by a single test, then more should be done.

107. In deciding on the number of tests required, both regulator and registrant should be conscious of the need to avoid unnecessary testing for both economic and animal welfare reasons. It will be preferable
for the registrant to carry out tests one at a time and consult the regulator after each one, as the results may indicate that the subsequent tests are no longer necessary.

**Protection against False Positives**

108. A false positive would occur if the test outcome indicated that avoidance would reduce risk when, in the field, it did not. Aspects of the guidance which contribute to protecting against a false positive are:

- realism of test species and conditions;
- some test conditions tending to worst case, e.g. highest treatment level, density of exposed granules;
- subjects in nutritional condition representative of field scenario;
- pre-fasting overnight;
- ensuring that birds are motivated to feed sufficiently before testing
- pre-conditioning to test substrate without active ingredient, where this is realistic;
- special consideration of the possibility of unusually rapid feeding;
- use of the untreated control treatment to check for spurious reductions in consumption;
- validation of the test method with a range of repellent and non-repellent pesticides (see below).

109. Despite the precautions listed above, it would not be prudent to rely solely on any test of avoidance using captive birds, no matter how conservative. It is very important that post-registration monitoring is used to detect false positives and provide the opportunity for corrective action. However, passive post registration monitoring is probably not efficient enough for this, as it would usually take some time for a body of evidence to build up. When false positives are detected, their causes should be carefully investigated to decide whether changes in test design are necessary.

**Animal Welfare**

110. The welfare of test subjects must be considered at several points during study design, including the choice of species, group size and pen size, degree of replication, and early termination of the test as a whole or for individual subjects (see above).

111. For primary (but not secondary) repellents an assessment of avoidance could be based on reductions in consumption alone, without allowing substantial intoxication, but considered that this would be unconservative and could result in the regulatory process failing to prevent mortalities occurring in the wild.

112. Those designing individual tests should seek to improve the welfare of subjects as far as possible without prejudicing the essential function of the test.

113. Testing should be avoided wherever extrapolation provides a reliable alternative.
References


BBA. 1993. Guidelines for testing plant protection products in the authorization procedure. Part IV, 25-1, Testing of baits, granules and treated seeds for hazards to birds - acceptance tests (2nd edition). Published in German by the Department of Plant Protection Products and Application Techniques of the Federal Biological Research Centre for Agriculture and Forestry, Germany.


Pascual, J.A., Fryday, S.L. and Hart, A.D.M. 1999a “Effects of food restriction on food avoidance and risk of acute poisoning of captive feral pigeons from fonofos-treated seeds.” Archives of Environmental Contamination and Toxicology, Vol. 37, pp. 115-124


APPENDIX

EXPERIENCE WITH ACCEPTANCE TESTS WITH BIRDS ACCORDING TO BBA-GUIDELINE VI 25-1

G. Joermann

Biologische Bundesantalt für Land- und Forstwirtschaft, BBA (Federal Biological Research Centre for Agriculture and Forestry), Braunschweig, Germany

The repellency test according to BBA-guideline VI 25-1 was introduced in the German authorisation procedure in the early 1980s. Since then, 42 studies have been submitted. The number is relatively low, because the test is conducted only in the case of toxic granules, seeds and baits. Unfortunately the results of the studies are confidential according to present legal regulations in Germany, and therefore can only be presented in an anonymous form.

The results are presented in tabular form on the next page. Often more than one study was submitted for a certain product (different species or different types of seeds); Such studies are combined in the table. Some of the products did not gain authorisation in Germany (due to different reasons).

ai: The active ingredients (or combinations) are numbers; a certain ai may appear in two categories.

Mortality: In test A the test substance and standard diet are offered in a ratio of 3:1; in test B the ratio is 1:9. If mortalities occur in both tests, the risk is considered as high; if mortalities only occur in test A, the risk is considered as medium. However, this is only a crude interpretation because the actual risk in the field depends on dosage (items per area) and technique (incorporation or surface application).

Incidents reported in Germany (Joermann and Gemmeke 1994): The entries in the table are related to the product in question. However, the reporting system is not well developed in Germany, so the absence of incidents is not proof of safety. Some products are new on the market or used on a small scale, so that incidents would not be expected.

Incidents reported in other countries: Information is derived, for instance, from the UK (e.g. Fletcher et al. 1994) and the USA (e.g. Smith 1987). The cases refer to the same type of formulation, but not necessarily to the same product.
Table 5A1
Comparison of the results of BBA acceptance tests and data from post-registration surveillance

<table>
<thead>
<tr>
<th>Type of formulation</th>
<th>Mortality:</th>
<th>Authorisation in Germany</th>
<th>Incidents:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test A</td>
<td>Test B</td>
<td>Germany</td>
</tr>
<tr>
<td>Granules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ai</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Seeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Baits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

¹ incidents hardly to be expected (new products, small scale of use)