	(max. 41.4%). For 28% of the pesticides 99.9% of the species will have the assumed level of protection. For birds, the median estimate of the fraction of species exposed above their LD50 for the first tier scenario (AF = 10) is 3.0% on average when the AF is applied to the lower of the toxicity values for the two standard test species. For 11% of the pesticides the median estimate is ≥ 10% (max. 15.7%). When the AF is applied instead to the geometric mean of the toxicity values for the two standard species, the median estimate of the fraction of species not covered by the AF is increased to 7.4% on average; and for 31% of the pesticides this fraction is ≥ 10% (max. 33.4%). This variation in the level of protection should be considered when defining the assumptions, assessment factors and decision criteria in regulatory risk assessment.

Title:

Variation in the level of protection afforded to birds and crustaceans exposed to different pesticides under standard risk assessment procedures.

Running title:

Variation in levels of protection in risk assessment

Robert Luttik[*, †], Andy Hart[‡], Willem Roelofs[‡], Peter Craig[§] and Pierre Mineau[⁄⁄]

[†] National Institute of Public Health and the Environment, Bilthoven, The Netherlands
[‡] The Food and Environment Research Agency, Sand Hutton, UK
[§] Department of Mathematical Sciences, Durham University, Durham, UK
[⁄⁄] Science and Technology Branch, Environment Canada, Ottawa, Canada

Robert.Luttik@RIVM.NL
andy.hart@fera.gsi.gov.uk
willem.roelofs@fera.gsi.gov.uk
P.S.Craig@durham.ac.uk
Pierre.Mineau@ec.gc.ca

* To whom correspondence may be addressed: Robert.Luttik@RIVM.NL

Abstract

First tier risk assessment for pesticides is often based on the quotient of the toxicity divided by the predicted environmental concentration or dose. This ratio is compared to a fixed assessment factor (AF) to decide whether the pesticide is to be allowed on the market or whether further research is needed. Often, a high value (e.g. the 90th percentile) is assumed for the predicted environmental concentration, and the lowest available value is chosen to represent toxicity; yet, the real level of protection is not known. Therefore, it is also not known whether the first tier is conservative enough or too conservative. By using two large toxicity databases and by assuming a log-logistic species sensitivity distribution for each pesticide, the % of species not covered by the AF is estimated in the scenario where exposure is at the maximum level allowable in the first tier. In the case of crustaceans, the median estimate of the fraction of species not covered by the AF of 100 in the first tier scenario is 3.4% on average for 72 pesticides. In other words, on average, 3.4% of the crustacean species will be exposed above their LC50 value in 10% of receiving surface waters receiving the maximum allowable exposure to an individual pesticide. The estimated level of protection varies widely between pesticides. For 10% of the pesticides, the estimated fraction of species not covered is ≥ 10% (max. 41.4%). For 28% of the pesticides 99.9% of the species will have the assumed level of protection. For birds, the median estimate of the fraction of species exposed above their LD50 for the first tier scenario (AF = 10) is 3.0% on average when the AF is applied to the lower of the toxicity values for the two standard test species. For 11% of the pesticides the median estimate is ≥ 10% (max. 15.7%). When the AF is applied instead to the geometric mean of the toxicity values for the two standard species, the median estimate of the fraction of species not covered by the AF is increased to 7.4% on average; and for 31% of the pesticides this

fraction is ≥ 10% (max. 33.4%). This variation in the level of protection should be considered when defining the assumptions, assessment factors and decision criteria in regulatory risk assessment.

Keywords: Level of protection, pesticides, birds, crustaceans

Introduction

First tier risk assessment for pesticides (e.g. pesticides and biocides) is often based on the quotient of the toxicity divided by the predicted environmental concentration or dose (PEC or PED). This ratio is compared to a fixed assessment factor (AF) to decide whether the pesticide is to be allowed on the market or whether further research or consideration is needed. Often, a high value (e.g. the 90th percentile) is assumed for the predicted environmental concentration, and the lowest available value is chosen to represent toxicity; yet, the real level of protection is not known. Therefore, it is also not known whether the first tier is conservative enough or too conservative.

From preliminary research it is known that the use of fixed AFs will result in different levels of protection (EFSA, 2005). That analysis suggested that the current first tier procedure is markedly more conservative for fish than for crustaceans and insects (see Table 3 in EFSA, 2005). Although not assessed in EFSA (op.cit.) there are also indications that the level of protection can differ substantially among pesticides within a group of organisms.

The first aim of this article is to provide information on the level of protection that is achieved by the first tier risk assessment when decisions are made for the authorisation of sprayed pesticides, e.g. for birds and for aquatic invertebrates. The overall approach is to use species sensitivity distributions (SSDs) to estimate the fraction of species not covered by the AF for each pesticide in a database, taking the exposure to be the maximum permissible under the first tier regulation, and to examine the variation in this fraction between pesticides for both birds and crustaceans. Uncertainty arising from small numbers of tested species for pesticides is calculated and shown in the figures.

Unfortunately, the guiding European Directive 91/414/EEC does not contain an explicit definition of the level of protection required when assessing risks to birds and to aquatic invertebrates (nor for other environmental risks). However, the directive describes to a certain extent how first tier risk assessment should be carried out, e.g. by defining the number of species to be tested and what assessment factor should be used. Acute risk assessment for birds in Europe requires only one bird species to be tested (LD50), either a quail species or the mallard duck. In most cases, however, two toxicity tests are available for birds, often for the bobwhite quail and the mallard duck (US EPA requirements). In that case in Europe the lower of the two toxicity values is used in the risk assessment. For aquatic risk assessments in the EU often only one invertebrate species is tested, e.g. the daphnid *Daphnia magna* (LC50). The European Directive does not provide guidance for the level of conservatism in the exposure assessment either. However, in the surface water document of the Forum for the Co-ordination of Pesticide Fate Models and their Use (FOCUS) the drift values for one or more applications are chosen in such a way that the models aim to obtain an overall 90th percentile drift loading for the entire season in the receiving surface water (i.e ditch, pond or river as defined in FOCUS (2001)). The first tier approach for birds, which assumes that all exposure is from the dietary route, is similarly based on the 90th percentile of the residue data, i.e. for about 90% of pesticide applications, residues on food items after spraying will be lower (EC, 2002)). Note that birds are assumed to eat 100% of their diet from the treated field and that no degradation of the pesticide is taken into account.

The acute Toxicity/Exposure ratio (TER) is compared with values specified in Annex VI of Directive 91/414/EEC, e.g. 100 for aquatic organisms or 10 for birds. These values can be regarded as assessment factors (AF) that allow for various uncertainties affecting the TER. There is no explicit documentation or justification for those choices. They are generally interpreted as relating only to uncertainties affecting the estimation of toxicity and are not intended to account for uncertainties in the estimation of exposure. It is often assumed that part of the AFs are accounting for between species extrapolation (i.e. from the test species to more sensitive species in the field) but it is not clear what proportion of the AF is assigned to this nor whether this is even sufficient. If the TER is lower than the relevant assessment factor, then authorisation may not be granted unless an appropriate (higher tier) risk assessment demonstrates that the risk is acceptable.

EFSA (2005) investigated whether the size of the AF could be reduced in order to maintain the level of protection when more data become available. As only the lowest toxicity value is used in the risk assessment, additional toxicity tests would lead to a more conservative risk assessment if the extra species were more sensitive than the standard test species. However, more data should allow for a better estimate of the risk and could reduce the uncertainty associated with the risk assessment rather than leading to more conservative risk estimates. Logically, there is no benefit to notifiers for providing additional data if the risk assessment is not altered and the most sensitive species continues to be used. Instead of reducing the AF, EFSA (2005) proposed that, as more data than the required minimum single species test become available, the geometric mean of the available toxicity data should be used. This, with the same exposure estimate and current assessment factor, would provide for the same level of protection albeit with a smaller level of uncertainty.

The second aim of this article is to demonstrate the implication for first tier risk assessment of implementing the advice to apply the standard AF to the geometric mean instead of the lower of the toxicity values for the two standard test species.

Methods

One way to assess the level of protection is to estimate for each pesticide the fraction of species not covered by the AF at the highest exposure (concentration/dose) that is considered to be safe in the current regulatory scheme, i.e. when the TER is equal to the AF. As the TER is based on the lowest toxicity value, another way to interpret this is to consider a regulatory 'safe' exposure which equals the ratio of the lowest toxicity value of the standard test species and the AF. The 'safe' exposure can be compared to a species sensitivity distribution (SSD) for a pesticide, based on all toxicity data available, in order to estimate what fraction of species is not covered by the AF; by 'not covered', we mean that the toxicity value for a species lies below the 'safe' exposure.

For invertebrates the 'safe' exposure is equal to PEC = L(E)C50/100 and for birds it equals PED = LD50/10. Under the directive 91/414/EEC Daphnia magna is used as a representative invertebrate. Sometimes additional invertebrate species may be a core data requirement, but for the following calculations, it is assumed that Daphnia magna is the only tested invertebrate species in the dossier.

The aquatic database used for the calculations in this paper is a research database of the National Institute of Public Health and the Environment (RIVM) in the Netherlands and is described in De Zwart (2002). From this database only the acute data for crustaceans and pesticides were used. The avian database on acute toxicity data for pesticides was made available by Environment Canada. Methods used to assemble this database were outlined in Mineau et al. 2001. The current database was updated in January 2007. Only pesticides with data on 4 or more species were used for this research. Where more than one toxicity value was available for a species for the same pesticide, the values were handled in the following way:

- censored values (where it is known only that the real toxicity value exceeded or fell below some threshold) were omitted unless the value was right-censored and was the maximum recorded or the value was left-censored and was the minimum;
- where only a single censored value remained that measurement was treated as censored in SSD calculations but was omitted from goodness-of-fit tests;
- otherwise the geometric mean was taken of the values.

Assuming a log-normal species sensitivity distribution (SSD), the % of species not covered by the AF can be estimated using ETX 2.0 (van Vlaardingen et al, 2004). However, ETX does not allow for censored data nor for other distribution families. Using our own programs written in R (R Core Development Team, 2010), we have carried out the same Bayesian calculation as ETX for log-normal SSDs but have extended the calculation to log-logistic and Weibull SSDs and to take censored data fully into account; for pesticides where there is no censored data, we have verified that the results obtained are the same for log-normal SSDs as those obtained using ETX. For each pesticide, the method provides a credible interval for the fraction of species not covered in addition to the median estimate of the fraction.

Goodness-of-fit was tested with Anderson-Darling tests for log-normality, log-logistic and Weibull distributions; precise P-values (not provided by ETX) were obtained by large-scale Monte Carlo simulation. For some chemicals, we also tested fit of log-skew-normal, inverse Burr and mixtures of log-normal distributions. Goodness-of-fit testing for many samples needs careful interpretation as one would expect some percentage of significant P-values even were the null hypothesis to be correct. We applied Fisher's method (Fisher, 1925), the combined probability test, to obtain a single overall P-value for the null hypothesis for multiple samples.

Results

In the aquatic and avian databases respectively, 78 and 62 pesticides were tested with 4 or more species. The results for the Anderson-Darling test are presented in Table 1. Because fewer than eight toxicity data are available for 43 out of 78 pesticides for the crustacean database, and for 31 out of 62 pesticides for the avian database, one might argue that the information to assess goodness-of-fit is too limited to reach firm conclusions. Even at much larger sample sizes, it is easy to show by Monte Carlo analysis that selecting the best-fitting distribution for an individual sample gives the wrong answer for a high proportion of samples and so we have not attempted to identify a particular distribution for each pesticide. Overall, by consideration of the distribution of per-pesticide P-values, we found that both databases slightly favoured the log-logistic distribution over the log-normal and that those both fitted better than the Weibull. For both databases, Fisher's method yields a highly significant overall P-value. The avian log-logistic overall P-value is 0.006 which becomes 0.12 on omission of the two pesticides having the most extreme individual P-values; detailed examination of those two pesticides suggests that a mixture of log-normal distributions gives the only satisfactory fit amongst the distributions we considered. The overall P-value for the crustaceans is 0.000027 and one would need to omit 7 pesticides to raise it above 0.05. Interestingly, several of those pesticides have sample sizes below 8. In what follows, we present the results of SSD calculations using the log-logistic distribution; in all figures, we highlight all data-sets having individual P-values below 0.05. Using log-normal and Weibull SSDs instead of log-logistic, we have produced and examined the same tables and figures as presented here; they are omitted in the interest of brevity and because we judge that the overall qualitative conclusions of the article would be the same.

For crustaceans, in the scenario in which the PEC is equal to the L(E)C50/100, the median estimate of the fraction of species not covered by the AF is on average 3.4% with a standard deviation of 7.0 (see Table 2). In other words, on average 3.4% of the crustacean species will be exposed above their L(E)C50 value in 10% surface waters receiving the maximum allowable exposure to an individual pesticide .

However, the achieved level of protection varies widely among pesticides. For 10% of the pesticides the fraction of species not covered by the AF is estimated to be equal or greater than 10% (see Table 3 and Figure 1). And at the other end of the distribution, for 15% of the pesticides, the estimated fraction of species not covered is less than 0.01%. The maximum estimated fraction not covered was 41%.

For birds, when the assessment is based on the lower toxicity value of the two standard tested bird species (bobwhite and mallard), the median estimate of the fraction of species exposed above their LD50 for the scenario where PED is equal to the LD50/10 is on average 3.0% with a standard deviation of 4.3 (see Table 2). In other words, on average, 3.0% of the bird species will be exposed above their LD50 value in 10% of pesticide applications at the maximum allowable exposure level.

For 11% of the pesticides the fraction of species not covered is estimated to be equal or greater than 10% (see Table 3 and Figure 2). And on the other hand the estimate is less than 0.01% for only 10% of the pesticides, The avian results are markedly different when the assessment is based on the geometric mean of the two available standard bird toxicity values rather than the lower value. Now the estimated fraction of species not covered following the usual first tier scenario is on average 7.4% with a standard deviation of 8.9 (see Table 2). For 31% of the pesticides the fraction of species not covered is estimated to be equal to or greater than 10% (see Table 3 and Figure 3) and for only 6% of the pesticides are fewer than 0.01% of the species exposed above their LD50.

In addition, in the figures, estimated lower and upper bounds are presented for the fraction of species not covered by the AF. For each pesticide, the uncertainty surrounding the fraction not covered is considerable due to the limited number of toxicity data.

We investigated the possibility of an association between sample size and estimated fraction of species not covered, using Spearman rank correlation and also by classifying each variable into three categories and applying the standard test of association in a contingency table. We found no evidence of association for the avian analyses. For crustaceans, there is evidence that higher sample sizes do not lead to very low estimates of fraction not covered.

Discussion

The results suggest that if the AFs used in current risk assessments are intended to account for the extrapolation of toxicity from tested to untested species, the level of protection they provide varies widely between pesticides. For example, for crustaceans the fraction of species not covered by the AF varied from less than 0.01% of species to about 40% (Table 2). This variation is caused by at least three factors:

1. Variation between pesticides in the sensitivity of the standard test species relative to other species (i.e. their position in the SSD).
2. Variation between pesticides in the standard deviation of the SSD (Levene's test for equality of variances between pesticides, $p < 0.001$ for both birds and crustaceans), and
3. Variation due to measurement uncertainty (between-lab or between-study variation).

Both the datasets considered here involve many different classes of pesticides. The avian dataset includes 62 pesticides from 18 classes of pesticides including 27 organophosphorous pesticides (OPs), 10 carbamates, and 7 organochlorines; the crustacean dataset includes 78 pesticides from 23 classes of pesticides including 26 OPs, 13 organochlorines, 8 carbamates and 5 pyrethroids. The variability shown by the analysis can be expected to apply generally, although the degree of variation may differ to some extent between classes. For example, variation in toxicity among crustacean species tends to be higher for insecticides (mean standard deviation = 0.85) than for herbicides (mean sd = 0.52, Mann-Whitney U = 221, p = 0.002**).** EFSA (2005) and Whiteside et al. (2008) also report different standard deviations for different taxa and chemical classes. This implies that the level of protection provided by the standard AFs depends to some extent on chemical class and taxonomic group, and varies also among individual pesticides within chemical class. Using the Mann-Whitney test, we investigated the possibility of association between chemical class and the estimated fraction of species not covered. Considering those pesticides classified as herbicide, fungicide or insecticide, we found no evidence of association for the bird analyses and some evidence (p=.04) for crustaceans (median estimate of fraction not covered: 0.4% for fungicides, 0.06% for herbicides and 1.4% for insecticides). No associations were found for pesticides classified as organophosphorus, organochlorine, carbamate or pyrethroid.

This analysis has examined the conservatism of one part of the risk assessment: the extent to which the AFs account for the variation of the toxicity endpoint among species. The overall level of protection is also influenced by the conservatism of other aspects of the risk assessment, including the choice of toxicity endpoint (e.g. LD50 vs. NOEL) and the exposure assessment. It has been argued that, overall, the level of protection afforded by current regulatory practices is high, due to the size of the chosen AFs, conservative assumptions in the exposure assessment and other factors (e.g. for birds, the potential for avoidance and metabolism to reduce the risk of effects). On the other hand, in the European legislative process, exposure is assumed to be only from the dietary route and this is known to be incorrect (EFSA , 2008). It is beyond the scope of this paper to quantify actual levels of protection when all these factors are considered. However, an indication of the overall level of protection for birds is provided by an analysis of field data shown in Figure 4 of EFSA (2008)), an update of the analysis first reported by Mineau (2002). That analysis indicates that there is evidence of bird mortality caused by the applied pesticides from field studies when the TER is close to 10 (i.e. allowing for an AF of 10). This suggests that, overall, the acute risk assessment for birds may not be conservative and may well not overestimate the likelihood of effects in the field.

As would be expected, the fraction of bird species not covered by the existing AF is higher when the TER is based on the geometric mean of the LD50s for the two standard test species than when the TER is based on the lower (more sensitive) of these (average percent of species not covered = 7.4% and 3.0% respectively, Table 2).

Using the minimum value introduces an increasingly conservative bias in the risk assessment when additional species are tested, whereas the geometric mean maintains the same level of protection on average (EFSA, 2005). This means that, if the first tier assessment is designed to achieve an appropriate level of protection overall (e.g. by comparison with field data as in EFSA 2008, or in relation to a specific percentile of variation in toxicity such as the HC5) when testing a single species, using the geometric mean will maintain that level of protection when additional species are tested. We found that the average percent of species not covered is 6.3% using bobwhite alone (standard deviation 9.0) and 15.7% using mallard duck alone (standard deviation 21.7) so that the geometric mean compromises between them. The difference between outcomes using mallard duck and bobwhite quail is attributable to a tendency for the latter to have a lower LD50 than the former.

Conclusions

We have quantified the extent to which AFs used in the risk assessment actually do account for the extrapolation in toxicity from tested to untested species and we have found that there is wide variation in outcome between pesticides for the databases we have used. The variation is a consequence of the small number of species tested. This variation in the level of protection could be reduced by testing more species, but such testing is avoided in routine risk assessment for both ethical and economic reasons. The consequence is that the estimated fraction of species not covered by the AF is subject to considerable uncertainty for any individual pesticide and it is unlikely that this component of uncertainty will be reduced substantially.

The overall level of protection achieved by the risk assessment as a whole depends on the conservatism of other aspects, including the assumptions used in estimating exposure. It is clear also that the overall level of protection may be lower for some classes of pesticides than others.. In any case, the uncertainties documented above should be taken into account when defining the assumptions, assessment factors and decision criteria to be used in regulatory risk assessment, to ensure that the levels of protection achieved meet policy objectives.

References

De Zwart D. 2002. Observed regularities in SSDs for aquatic species. In: Posthuma L, II Suter GW, Traas TP. Species sensitivity distributions in ecotoxicology. Boca Raton, Florida, USA, Lewis Publishers, p 133-154.

[EFSA] European Food Safety Authority. 2005. Opinion of the Scientific Panel on Plant health, Plant protection products and their Residues (PPR) on a request from EFSA related to the assessment of the acute and chronic risk to aquatic organisms with regard to the possibility of lowering the uncertainty factor if additional species were tested. 14 December 2005. *EFSA Journal* 301: 1-45. http://www.efsa.europa.eu/en/scdocs/doc/301.pdf

[EFSA] European Food Safety Authority. 2008. Risk Assessment for Birds and Mammals - Revision of Guidance Document under Council Directive 91/414/EEC (SANCO/4145/2000 – final of 25 September 2002) [1] - Scientific Opinion of the Panel on Plant protection products and their Residues (PPR) on the Science behind the Guidance Document on Risk Assessment for birds and mammals. *EFSA Journal* 734: 1-181. http://www.efsa.europa.eu/en/scdocs/doc/734.pdf

[EC] European Community. 2002. Guidance document on risk assessment for birds and mammals under Council Directive 91/414/EEC SANCO/4145/2000 – Final 25 September 2002, 74 p. http://ec.europa.eu/food/plant/protection/evaluation/guidance/wrkdoc19_en.pdf

Fisher RA. 1925. Statistical Methods for Research Workers. Edinburgh, Oliver and Boyd.

[FOCUS] Forum for the coordination of pesticide fate models and their use. 2001. "FOCUS Surface Water Scenarios in the EU Evaluation Process under 91/414/EEC". Report of the FOCUS Working Group on Surface Water Scenarios, EC Document Reference SANCO/4802/2001-rev.2. p 245. http://focus.jrc.ec.europa.eu/sw/docs/FOCUS_SWS_Final_Report.doc

Mineau P, Baril A, Collins BT, Duffe J, Joerman G, Luttik R. 2001. Reference values for comparing the acute toxicity of pesticides to birds. *Reviews of Environmental Contamination and Toxicology* 170:13-74.

Mineau P. 2002. Estimating the probability of bird mortality from pesticide sprays on the basis of the field study record. *Env. Toxicol. Chem.* 21: 1497-1506.

R Development Core Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/

Van Vlaardingen PLA, Traas TP, Wintersen AM, Aldenberg T. 2004. ETX 2.0. A program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data. Bilthoven, the Netherlands: National Institute for Public Health and the Environment (RIVM). Report no. 601501028/2004, p 68.

Table 1 Goodness of fit using the Anderson Darling test for the log-logistic distribution

Group	Range of P-values				
	<.01	.01-.025	.025-.05	.05-.1	.1-1
	% of pesticides				
Crustaceans	4	6	1	14	74
Birds	5	2	2	10	82

Table 2 Summary of estimated fractions of species not covered by the assessment factor (AF).

Group	Toxicity endpoint	Estimated fraction of species not covered (%)				Number of compounds
		mean	std	minimum	maximum	
Crustaceans	LC50	3.4	7.0	< 0.0001	41.4	78
Birds (applying AF to lower toxicity value)	LD50	3.0	4.3	< 0.0001	15.7	62
Birds (applying AF to geometric mean of toxicity values)	LD50	7.4	8.9	< 0.0001	33.4	62

Table 3 **Distribution between pesticides of the estimated fraction of species not covered by the assessment factor (AF)**

Group	Range of estimate of fraction not covered					n
	< 0.01%	0.01-0.1%	0.1 - 1%	1-10%	>10%	
	Percentage of pesticides in range					
Crustaceans	15	13	22	40	10	78
Birds (applying AF to lower toxicity value)	10	10	32	37	11	62
Birds (applying AF to geometric mean of toxicity values)	6	6	23	34	31	62

Figure 1 Estimated fraction (%) of crustacean species not covered by assessment factor, i.e. with EC50 < (daphnia magna EC50)/100. Number of tested species shown at left; * indicates compound failing log-logistic goodness-of-fit test at 5% level.

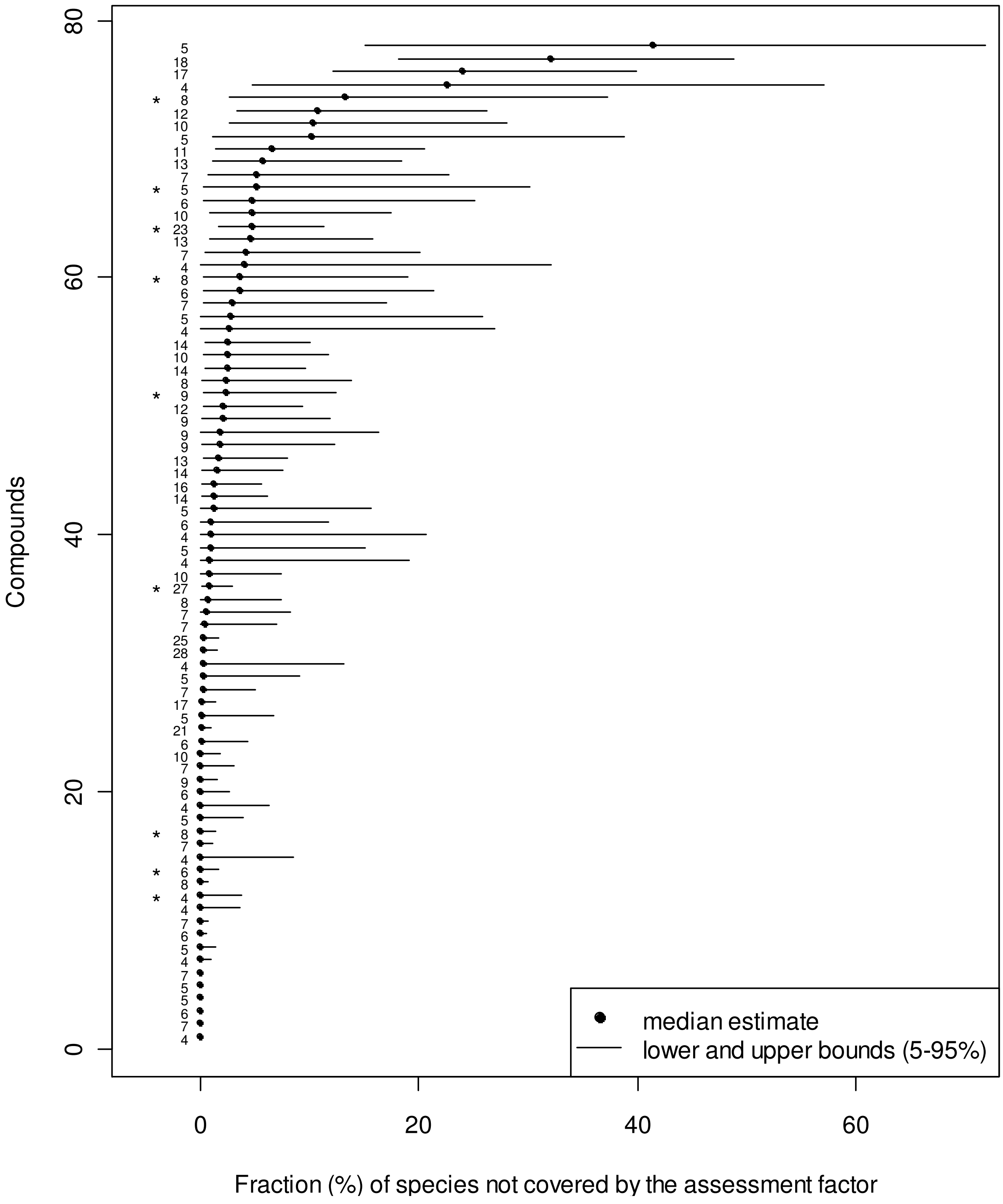

Figure 2 Estimated fraction (%) of bird species not covered by assessment factor, i.e. with LD50<(lower of bobwhite quail and mallard duck LD50s)/10. Number of tested species shown at left; * indicates compound failing log-logistic goodness-of-fit test at 5% level.

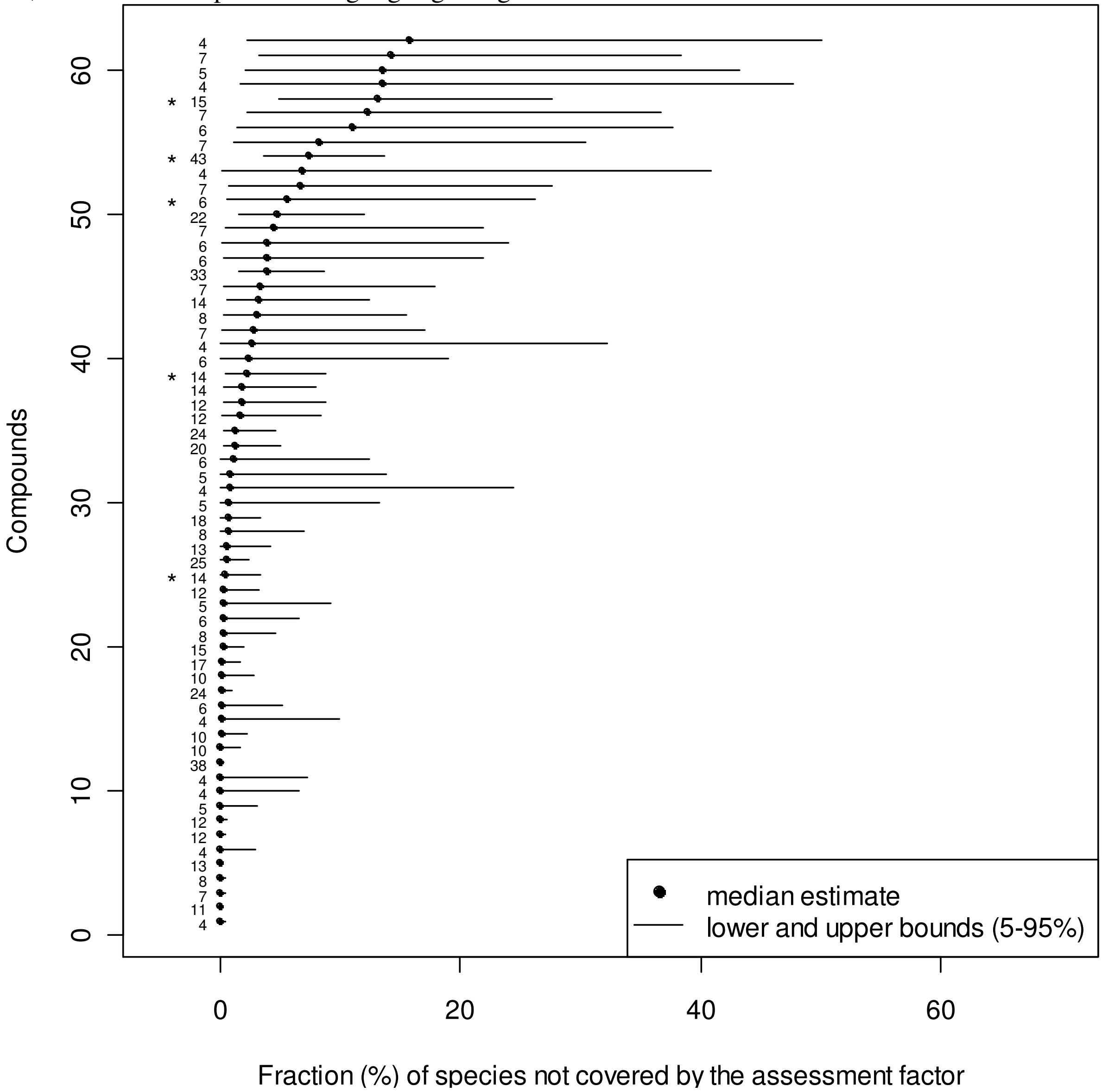

Figure 3 Estimated fraction (%) of bird species not covered by assessment factor, i.e. with LD50<(geometric mean of bobwhite quail and mallard duck LD50s)/10. Number of tested species shown at left; * indicates compound failing log-logistic goodness-of-fit test at 5% level.