SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 22, No. 9, pp. 2020–2029, 2003

Printed in the USA
0730-7268/03 $12.00 + .00

ANALYSIS OF FIELD AND LABORATORY DATA TO DERIVE SELENIUM TOXICITY THRESHOLDS FOR BIRDS

WILLIAM J. ADAMS,*† KEVIN V. BRIX,‡ MELANIE EDWARDS,§ LUCINDA M. TEAR,§ DAVID K. DEFOREST,§ and ANNE FAIRBROTHER§

†Kennecott Utah Copper Corporation, 8315 West 3595 South, P.O. Box 6001, Magna, Utah 84044, USA
‡EcoTox, 721 Navarre Avenue, Coral Gables, Florida 33134, USA
§Parametrix, 5808 Lake Washington Boulevard Northeast, Suite 200, Kirkland, Washington 98033, USA

(*Received* 11 *March* 2002; *Accepted* 19 *December* 2002)

Abstract—In this paper, we critically evaluate the statistical approaches and datasets previously used to derive chronic egg selenium thresholds for mallard ducks (laboratory data) and black-necked stilts (field data). These effect concentration thresholds of 3%, 10% (EC10), or 20% have been used by regulatory agencies to set avian protection criteria and site remediation goals, thus the need for careful assessment of the data. The present review indicates that the stilt field dataset used to establish a frequently cited chronic avian egg selenium threshold of 6 mg/kg dry weight lacks statistical robustness (r^2 = 0.19–0.28 based on generalized linear models), suggesting that stilt embryo sensitivity to selenium is highly variable or that factors other than selenium are principally responsible for the increase in effects observed at the lower range of this dataset. Hockey stick regressions used with the stilt field dataset improve the statistical relationship (r^2 = 0.90–0.97) but result in considerably higher egg selenium thresholds (EC10 = 21–31 mg/kg dry wt). Laboratory-derived (for mallards) and field-derived (for stilts) teratogenicity EC10 values are quite similar (16–24 mg/kg dry wt). Laboratory data regarding mallard egg inviability and duckling mortality data provide the most sensitive and statistically robust chronic threshold (EC10) with logit, probit, and hockey stick regressions fitted to laboratory data, resulting in mean egg selenium EC10 values of 12 to 15 mg/kg dry weight (r^2 = 0.75–0.90).

Keywords—Selenium toxicity thresholds Avian Egg

INTRODUCTION

The toxic effects of selenium were first described in South Dakota (USA), where chickens (*Gallus domesticus*) were observed to have poor reproductive success as a result of chick teratogenesis and mortality when fed grains from specific sources [1]. Subsequent studies identified elevated selenium content in the grains as the cause of the observed effects [2]. During the 1980s, a severe reduction in reproductive success of aquatic birds at Kesterson Reservoir (CA, USA) prompted a series of studies that identified seleniferous water from subsurface agricultural drainage as the causative factor. Again, teratogenesis and chick mortality were the primary toxicological effects, although selenium concentrations were sufficiently elevated to also cause effects in adult birds [3].

Like other metals and metalloids, the primary avian exposure pathway for selenium is the diet [4,5]. Aquatic invertebrates and plants bioaccumulate selenium via the water and their diet, and aquatic birds that feed on these invertebrates and plants during the breeding season transfer a significant fraction of their dietary intake to the eggs [6,7]. At sufficiently high egg selenium concentrations, teratogenic effects on developing embryos can result [8,9].

Selenium interferes with embryo development and, at sufficiently high concentrations, can cause gross abnormalities, such as anophthalmia, incomplete beak development, brain defects (hydrocephaly and exancephaly), foot defects, and other terata [8,9]. More subtle teratogenic effects, such as enlarged hearts, edema, liver hypoplasia, and gastroschisis, also occur but rarely are recorded in field studies. Any of these effects can lead to reduced embryo and chick survival.

Several toxicological endpoints are relevant for selenium, including teratogenesis, egg inviability, clutch inviability (i.e., one or more inviable eggs in a clutch), and chick mortality. These endpoints have been characterized as a function of individual egg selenium concentrations, mean egg selenium concentrations for a clutch, and clutch or hen responses. Differences in the way the dose–response relationship is characterized are important in interpreting the thresholds proposed by different researchers. The effects of these differences will be discussed later.

Since the events at Kesterson, several laboratory and field studies have been conducted to estimate egg selenium toxicity thresholds for birds. The importance of the threshold is that it often is used as a concentration of concern for regulatory decisions. The U.S. Fish and Wildlife Service has reviewed the extant data several times during the past few years and consistently recommended an egg selenium threshold of 6 to 7 mg/kg dry weight based on field data for black-necked stilt (*Himantopus mexicanus*) clutch inviability [10,11].

More recently, Fairbrother et al. [12,13] developed mean egg selenium thresholds for teratogenesis and duckling mortality by pooling toxicity data from several laboratory studies with mallards (*Anas platyrhynchos*). Fairbrother et al. also developed dose–response relationships for these pooled data. They concluded that the mean egg selenium threshold for duckling mortality was 16 mg/kg dry weight. They further questioned whether the field-based threshold (6 mg/kg dry wt) developed by the U.S. Fish and Wildlife Service was reliable, noting that field- and laboratory-based thresholds (10% effect

* To whom correspondence may be addressed (william.adams@riotinto.com).

concentration [EC10]) for the less-sensitive threshold of teratogenesis were similar but that the chick/duckling mortality endpoints were dissimilar. They hypothesized that the data regarding teratogenesis were in agreement because it was a selenium-specific endpoint that could be measured accurately in the field. In contrast, the chick/duckling mortality endpoint was not selenium-specific and was subject to confounding factors, such as weather, starvation, and other contaminants, when measured in the field. They further hypothesized that the lower threshold predicted from field data was confounded by one or more of these factors and concluded that until these data could be thoroughly scrutinized, the laboratory-derived threshold of 16 mg/kg dry weight should be used.

Since the publication of Fairbrother et al. [13], the raw data for the field-based thresholds have been released to the public for review [10]. Additionally, Ohlendorf [14] recently reexamined the available laboratory toxicity data using logistic regression and concluded that the mean egg selenium threshold (EC10) is 12.8 mg/kg dry weight. Ohlendorf pointed out that Fairbrother et al. [13] failed to include three available data points from laboratory mallard studies that could affect the results of their threshold analysis. The purpose of our study was to critically review the field data, update the analysis of Fairbrother et al. to include the missing data, and compare and contrast the field and laboratory data.

MATERIALS AND METHODS

Toxicological endpoints

Before comparing field and laboratory data, it is necessary to define the different toxicological endpoints that have been used to characterize the effects of selenium on aquatic birds. Chick mortality is thought to be the most sensitive chronic toxicological endpoint [10,13]. This endpoint evaluates net reproductive success, taking into account chicks affected by teratogenesis, egg inviability, and mortality up to 7 d posthatch. However, because of practical limitations in the field, evaluation of chick mortality typically is not feasible because it is difficult to accurately monitor chicks posthatch. Hence, available field data do not consider this endpoint.

Egg inviability is slightly less sensitive than chick mortality [15], whereas teratogenesis tends to be less sensitive than chick mortality by a factor of 1.5- to 3-fold [10,12]. A fourth endpoint, clutch inviability, is toxicologically equivalent to egg inviability but is evaluated on a henwise (i.e., clutchwise) basis rather than on an individual-egg basis. This endpoint has been used only for field data. It is based on the premise that individual sibling chicks within a clutch are not independent samples, so the hen (i.e., clutch) represents the lowest statistical unit. In developing clutch inviability data, a clutch is considered to be inviable (i.e., affected) if at least one egg did not hatch [10].

Teratogenicity field data are available for ducks, black-necked stilts, and American avocets (*Recurvirostra americana*). The duck data are a composite of information on *Anas* spp. (gadwalls, mallards, pintails, and shovelers), *Aythya* spp. (redheads and canvasbacks), and the ruddy duck (*Oxyura jamaicensis*). In contrast, field clutch inviability data are available only for black-necked stilts. A limited number of field studies have also monitored clutch inviability and egg inviability in concert. Skorupa [10] evaluated individual egg data from eight nesting neighborhoods, each of which contained from 18 to 42 nests. Using these data, Skorupa developed a relationship between the probability of a given egg selenium concentration resulting in an inviable clutch or an inviable egg. Using this relationship, Skorupa developed the following equation to relate the probability of an inviable egg to the probability of an inviable clutch:

$$AE{:}AH \text{ ratio} = 0.047 + 0.423 \times \log(AH) \quad (1)$$

where

$$AE = \% \text{ affected eggs}$$

$$AH = \% \text{ affected hens (or clutch)}$$

We used the above equation in our analyses of field data on stilt clutch inviability to calculate the *AE*:*AH* ratio. The *AE*:*AH* ratio for a given egg selenium concentration was then multiplied by the probability of an inviable clutch to estimate the associated probability of an inviable egg. This effectively normalized the field and laboratory data to the same endpoint. This technique was used only in the hockey stick regression analysis for reasons that will be discussed later.

Table 1. Study areas for field data located in the United States

	Teratogenesis			Inviability
Study area[a]	Avocet	Stilt	Duck	Stilt
Bowdoin NWR, MT	X			
Grasslands Water District, CA	X	X	X	
Gunnison River Area, CO			X	
Honey Lake Valley, CA	X	X		
Kendrick Irrigation Project, WY	X			
Kesterson Reservoir, CA	X	X	X	X
Ouray NWR, UT			X	
Red Rock Ranch, CA		X		
Salton Sea, CA				X
San Juan River Area, NM	X		X	
Sun River Area, MT	X			
Tulare Basin, CA	X	X	X	X
Volta State Wildlife Area, CA	X	X	X	X

[a] NWR = National Wildlife Refuge.

The source of the field datasets was the raw data published in the appendices of Skorupa. These binomial data were collected by various researchers from 1983 through 1996 from study areas listed in Table 1. Some of the locations were systematically sampled; others were not. All data were pooled regardless of sampling approach, as in Skorupa.

Statistical analysis of laboratory and field data

The analysis of Fairbrother et al. [13] was updated by including the three data points from Heinz and Hoffman [16,17] that were previously excluded. Other than inclusion of these additional data, the analysis was performed here as previously described by Fairbrother et al. [12,13], except that the dataset was also analyzed using hockey stick regression, as described later in this paper. Table 2 provides the raw laboratory data used in this reanalysis.

Initially, we evaluated the field data using the same methods and datasets as Skorupa [10] for comparative purposes. Skorupa used generalized linear models (GLiMs) and a three-point moving average in separate analyses to evaluate goodness-of-fit and to estimate toxicity thresholds (e.g., EC10). We used a similar approach and then considered an alternative method that involved binning the field data and using hockey stick regressions to estimate the toxicity threshold. For comparison,

Table 2. Laboratory egg selenium toxicity data used in reanalysis of Fairbrother et al. [12,13]

Mean egg selenium (mg/kg dry wt)[a]	% Teratogenesis in eggs that did not hatch	% Fertility[b]	% Fertile eggs that hatched[c]	% All eggs that hatched	% Duckling survival	% Duckling mortality for all eggs laid[d]	% Duckling mortality for all eggs laid (Abbott's-corrected)	Reference
0.6	0.6	98.6	59.6	58.8	99.4	41.6	0	[15]
0.93	NA[e]	88	62	55	74.2	59.5	0	[36]
1.16	6.1	96.0	44.2	42.4	96.2	59.2	0	[17]
1.35	1.3	96.7	41.3	39.9	98.4	60.7	0	[16]
1.4	1	100	91.4	91.4	82.5	24.6	0	[36]
2.8	0.9	96.3	70.7	68.1	97.7	33.5	0	[15]
5.3	0.5	97.0	60.0	58.2	98.9	42.4	1.5	[15]
11.3	1.4	97.9	53.4	52.3	99.7	47.9	10.8	[15]
12.1	NA	90	61	55	78.1	57.1	0	[37]
24.5	NA	89	41	36	60.9	77.8	45.0	[37]
25.1	36.2	99.1	24.0	23.8	66.0	84.3	62	[17]
29.4	28.2	99.1	6.4	6.3	20.0	98.7	97	[16]
30.4	24.6	95.6	7.6	7.3	47.2	96.6	91	[16]
36.7	6.8	96.6	36.9	35.6	80.6	71.3	50.8	[15]
42	57.5	100	8.5	8.5	10.0	99.2	98.9	[36]
60	67.9	99.6	2.2	2.2	0	100	100	[15]

[a] If mean egg selenium (MES) was reported as wet weight, then MES as dry weight was calculated as $MES_{wet\ weight} \div$ fraction solids (using % moisture from study).
[b] Percentage fertility not reported in Stanley et al. [36], who only reported that selenium had no effects on fertility. Assumed fertility was 100%.
[c] In Stanley et al. [36], it is unclear if hatching success is reported for all eggs or for fertile eggs. Currently, it is assumed it is for fertile eggs.
[d] Duckling mortality assessed after 6 d by Heinz et al. [15] and by Heinz and Hoffman [16], after 7 d by Heinz and Hoffman [17], and after 14 d by Stanley et al. [36,37]. Ducklings in the Heinz et al. [16] study were not fed selenium diets; ducklings in the Stanley et al. [36,37] studies were fed the same selenium diets as their parents.
[e] Percentage of teratogenic embryos not reported; it was only stated that teratogenesis was not significantly different than in controls.

hockey stick regressions were also fitted to the laboratory data. The methods for each of these models are described below.

Generalized linear models. For stilts, avocets, and ducks, each measured egg selenium concentration (mg/kg dry wt) was paired with a 0 or a 1 to indicate the absence (0) or the presence (1) of teratogenesis in the embryo. For stilt clutch inviability, egg selenium was measured in one egg, and 0 or 1 indicated whether one or more sibling eggs did (0) or did not (1) hatch. Binary logistic regression models (logistic link, binomial error) were used to fit the relationship between egg selenium concentrations and the binary responses, teratogenesis, and clutch inviability. We did not apply the *AE*:*AH* ratio equation to the stilt clutch inviability data, because as discussed below, the r^2 for this model was very low and further extrapolation of the data was not considered to be appropriate. When data indicated teratogenesis or clutch inviability at background selenium levels, the methods of Bailer and Oris [18] were used to estimate the percentage decrease in clutch viability relative to the background level of effects. Models were fitted using both SPlus [19] and SPSS [20] as a quality control, and results were compared to those reported by Skorupa [10].

In binary logistic regression, maximum likelihood methods are used to compare the fit of a model with an independent variable of interest to the fit of a model without that variable [21]. In maximum likelihood methods, the significance of the proposed independent, explanatory variable is assessed by comparing the likelihood of the model with the term (expressed as deviance) to the likelihood of the model without the term. A difference in deviances is assumed to be chi-squared distributed, with degrees of freedom being equal to the difference between the number of parameters used in the two models.

In this case, the significance of selenium in explaining teratogenicity and clutch inviability was assessed by comparing the difference between the model deviance with an intercept only and the model deviance with an intercept and a slope to a chi-squared statistic with one degree of freedom. If the model choice is correct, then maximum likelihood procedures simultaneously assess both the significance of the independent variable and the goodness-of-fit of the model, in that the addition of an independent variable to the model will not be significant if it does not represent a significant improvement of fit compared to the less-complicated model. In the more familiar cases of analysis of variance and general linear regression, maximum likelihood and least-squares methods are identical, and goodness of fit is expressed as the familiar r^2 statistic that describes the percentage of variance that is explained by the model. For GLiMs, a variety of summary statistics have been developed that can be considered as analogous to the r^2 statistic (referred to as pseudo r^2 by Hagle and Mitchell [22]). Because none of these statistics is exactly equivalent to the percentage of variance explained and each reveals a slightly different perspective on the data, we report two measures: First the Nagelkerke r^2 (r^2_M in Menard [23]), and then the Dhrymes measure (Hagle and Mitchell [22] from Dhrymes [24]; also called r^2_L in Menard [23]). These two measures can be expressed as

$$\text{Nagelkerke} \qquad r^2 = \frac{1 - (L_0/L_1)^{2/N}}{1 - (L_0)^{2/N}} \tag{2}$$

$$\text{Dhrymes} \qquad r^2 = 1 - \frac{\ln L_1}{\ln L_0} \tag{3}$$

where

L_0 = likelihood of model with intercept only

L_1 = likelihood of full model (intercept and slope)

We did not utilize the frequently used Hosmer and Le-

meshow [25] goodness-of-fit test, which is based on the chi-squared test, because small numbers of teratogenic eggs for all species meant that many of the cells of the contingency table had counts of less than five, and this violated the assumptions of the test. Discussion of these and other methods of assessing model fit can be found in any number of texts, articles, and software manuals about GLiMs (e.g., [19–23,25–28]).

Three-point moving average. In addition to the GLiM analysis, we also analyzed the black-necked stilt field clutch inviability data using a three-point moving average, as in Skorupa [29]. Egg selenium concentrations were rounded to whole numbers, and a response rate was calculated for each. The response rate was calculated as the number of inviable clutches divided by the number of clutches assessed at a given (rounded) selenium concentration. A three-point moving average response rate was calculated as the average of the calculated response rates for egg selenium concentrations at 1 mg/kg (dry weight) lower than the given level, at the given level, and at 1 mg/kg higher than the given level. Skorupa used this method to further support a threshold of 6 mg/kg but only presented results for selenium concentrations from 4 to 9 mg/kg. We applied similar and additional analyses to the full range of consecutive whole-mg/kg selenium egg concentrations (2–160 mg/kg dry wt) reported in the Appendix of Skorupa [10].

Hockey stick regressions. As another approach for evaluating and interpreting data regarding the effects of selenium, we fit hockey stick regression models to the field (stilt) and laboratory (mallard) dose–response relationships. This type of model has been used to define a threshold when an underlying background level of response is unrelated to the dose, as is the case for many toxicological datasets [30–32].

The response endpoints in the field and the laboratory studies were equated for ease of comparison. The field data for clutch inviability were based on a clutchwise (i.e., henwise) response, whereas the laboratory data were based on individual responses. For the clutchwise response, a positive response was recorded when one or more eggs were observed not to hatch, which is consistent with Skorupa [10]. We used the *AE*: *AH* ratio equation described earlier (Eqn. 1) to normalize the data to individual egg responses [10].

The laboratory studies combined by Fairbrother et al. [12,13] measured duckling survival up to 7 d posthatch, whereas the field data reported by Skorupa [10,29] did not consider chick survival posthatch. In the present study, we used the egg inviability endpoint for the laboratory data without considering posthatch survival so that we could make direct comparisons between the field and laboratory data. However, our final analysis of the laboratory data (independent of the comparison to the field data) was based on posthatch duckling survival to provide an appropriately conservative threshold.

A background or control response rate was observed in both field and laboratory datasets. This is particularly important for the egg inviability endpoint, because the response rate was significantly greater than zero. To account for control responses, the datasets were normalized using Abbott's formula [33]. For the field data, the response observed in eggs with selenium levels less than 3 mg/kg dry weight was considered as background for the normalization procedure. Use of Abbott's formula is necessary to separate the selenium-related response from the background response, but it also reduces response variance at background and low-exposure concentrations (any response less than the mean background response is set to zero). Consequently, to some extent, Abbott's formula artificially enhances the ability of the hockey stick regression to define a threshold as compared to a noncorrected dataset.

A final difference between the field and laboratory hatchability data is that the field data were based on measured responses in stilts, whereas the laboratory studies measured mallard responses. Based on field data regarding teratogenesis collected on both species, mallards are approximately 50% more sensitive than stilts [10]. Whether the same relative sensitivities apply to other endpoints (e.g., clutch inviability) is unknown.

In the field, birds are exposed to an uncontrolled range of dietary selenium concentrations, as opposed to the laboratory studies, which exposed multiple hens to specific concentrations. To model the field data similarly to the laboratory data, the field selenium concentrations were binned into treatment categories, and a percentage response was calculated for each bin. Five different binning schemes (A through E) were used to evaluate the sensitivity of the dose–response relationship to a specific binning scheme. The rate of teratogenicity in ducks, stilts, and avocets was calculated as the number of teratogenic eggs divided by the number of eggs assessed within each binned selenium range. In all binning schemes, the response rate was paired with the median egg selenium concentration within the respective bin. Data evaluations using binning schemes were performed on the field and laboratory datasets for both teratogenicity and clutch inviability.

Binning scheme A grouped egg selenium concentrations according to a geometric sequence: 0 to <2, 2 to <4, 4 to <8, 8 to <16 mg/kg dry weight, etc. Binning scheme B grouped responses by every 5 mg/kg dry weight in measured egg selenium, resulting in bins of 0 to <5, 5 to <10, 10 to <15, and 15 to <20 mg/kg dry weight, etc. Binning scheme C grouped a similar number of responses per bin, approximately 11 eggs (12 bins) for teratogenesis and 21 eggs (19 bins) for inviability. Scott's normal approximation was used to calculate the most informative number of bins from each dataset [34]. The resulting binning scheme for stilt teratogenesis used median selenium concentrations of 1.6, 2.25, 3.9, 5.1, 7.3, 9.65, 11, 16, 20, 23, 27, and 35 mg/kg dry weight. The binned median selenium concentrations for stilt egg inviability were 2.1, 2.85, 3.4, 4.2, 5.1, 6.5, 7.5, 9.65, 16, 18, 22, 28, 33, 37, 41, 48.5, 54, 59, and 72.5 mg/kg dry weight. Binning scheme D was similar to binning scheme A but incremented initially by 3 mg/kg dry weight (i.e., 0 to <3, 3 to <6, 6 to <12, 12 to <24 mg/kg dry wt, etc.). This scheme results in a bin break at the threshold of 6 mg/kg dry weight proposed by Skorupa [10,29]. Binning scheme E used the breakpoints identified by the U.S. Department of the Interior [11], namely median selenium concentrations of 2.7, 9.2, 21.5, 31, and 76.5 mg/kg dry weight for teratogenesis and of 3.2, 7.65, 22, 42.5, 65, 74, and 97.5 mg/kg dry weight for egg inviability. Each binning scheme has benefits and drawbacks, so they should be interpreted collectively.

Using all binning schemes, considerable variance was seen in the relationship between egg selenium concentration and the measured response. A true dose–response relationship did not appear to begin until higher concentrations were reached, although the exact point at which the relationship began depended on the binning scheme. Hockey stick regression models were fit to describe the variance in the response at low selenium concentrations, to estimate the inflection concentration at

which the dose–response relationship began, and to describe the dose–response relationship above the inflection point.

Hockey stick is a form of nonlinear model in which the independent variable is broken into two nonoverlapping ranges and separate linear regression models are fit to the response variable in these two subsets [30,32]. The two models are connected at an inflection point (quantification of the breakpoint [τ]) by constraining the second intercept to be a function of the first model and τ. This allows separate relationships to be fit to different subsets of the independent variable, but all data are used to estimate the slope and intercept of each line as well as the inflection point at which they connect. The general form for this model can be expressed as

$$y = \begin{cases} a + bx & x \le \tau \\ c + dx & x \ge \tau \end{cases} \quad (4)$$

When $c = a + \tau(b - d)$, y will be continuous at all values of x, and the final model will be of the form

$$y = \begin{cases} a + bx & x \le \tau \\ a + b\tau + d(x - \tau) & x \ge \tau \end{cases} \quad (5)$$

Hockey stick models are frequently used to describe data with a threshold response [31,32]. As part of model fitting, maximum likelihood methods are used to determine whether a significant linear relationship exists at values of the independent variable that are less than the inflection point. If no relationship exists between the independent and dependent variable when the independent variable is less than τ (i.e., $b = 0$), then the model can be simplified to

$$y = \begin{cases} a & x \le \tau \\ a + d(x - \tau) & x \ge \tau \end{cases} \quad (6)$$

In this application, the intercept estimates the mean response associated with selenium concentrations less than τ or, in other words, the background level of response (teratogenic or inviable eggs) that is not related to the presence of selenium. The value of τ is the selenium concentration at which a significant relationship begins between selenium and response. As a result, the minimum detectable effects concentration can be estimated as the ECτ.

In all cases, the selenium concentrations were analyzed on a log scale to distribute the data more uniformly over the range of concentrations assessed, allowing each data point to contribute more equally to the fit of the curve [32]. The datasets were truncated at the lowest concentration at which the maximum observed response occurred to represent the true slope of the dose–response relationships. Models were fit in SPSS [20]. An r^2 value that assumed no intercept was used to assess the fit of the models:

$$r^2 = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} y_i^2} \quad (r^2 \text{ for a no-intercept model}) \quad (7)$$

RESULTS

Updated analysis of laboratory data for mallards

Inclusion of three data points from Heinz and Hoffman [16,17] in the mallard laboratory dataset previously analyzed by Fairbrother et al. [13] did make a slight difference in estimates of the EC10 for teratogenesis and duckling survival 7 d posthatch. Previously, Fairbrother et al. [13] reported an estimated EC10 of 26 mg/kg dry weight for teratogenesis and of 16 mg/kg dry weight using both probit and logit models for duckling survival. Analysis of the revised dataset resulted in an EC10 of 23 mg/kg dry weight for teratogenesis and of 14 to 15 mg/kg dry weight (logit, probit models) for mallard duckling survival (Figs. 1 and 2). One data point reported by Heinz et al. [35] was not included in the analysis (as discussed by Fairbrother et al. [12]), because the mean egg selenium concentration (15.9 mg/kg dry wt) associated with duckling mortality (i.e., 76%, corrected for control mortality) was higher, by a factor of seven- or eightfold, than would be expected as compared to all other data in the dataset. When the data point is included in the dataset, the calculated mallard duckling mortality EC10s for mean egg selenium are slightly lower and range from 13 to 14 mg/kg dry weight using probit and logit models, respectively. In addition to the probit and logit models, we also fitted the data to a hockey stick regression (Fig. 3). Analysis of the dataset via this method resulted in a slightly lower mean egg selenium EC10 of 12 mg/kg dry weight ($r^2 = 0.90$).

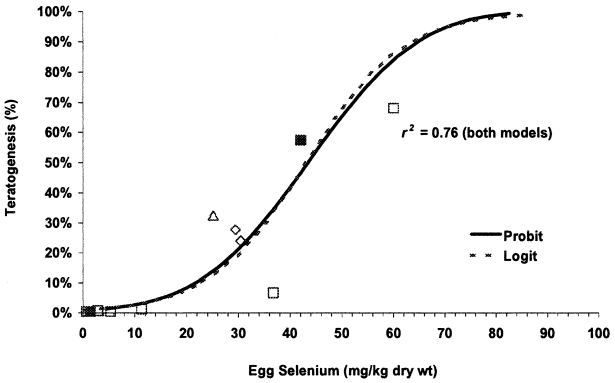

Fig. 1. Logit and probit models fit to mallard duck laboratory teratogenicity data obtained from several studies (□, Heinz et al. [15]; ◊, Heinz and Hoffman [16]; △, Heinz and Hoffman [17]; ■, Stanley et al. [36]).

GLiMs applied to field data for black-necked stilts

Analysis of the field data showed teratogenesis to be a less-sensitive endpoint than stilt egg inviability, as expected from previous studies. The proportions of teratogenic eggs in the field samples were much lower than the proportion of inviable eggs (1–11% for teratogenesis and 29% for egg inviability). It is interesting to note that the range of concentrations over which effects were found for teratogenicity was much narrower

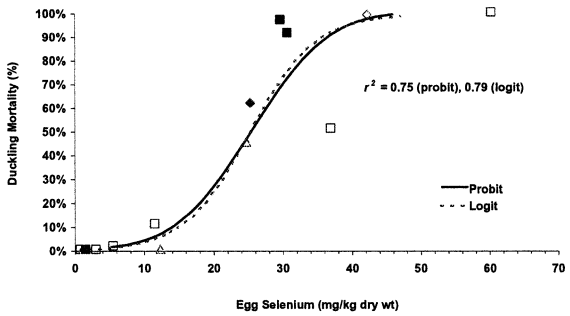

Fig. 2. Logit and probit models fit to laboratory mallard duckling survival at 7 d posthatch (□, Heinz et al. [15]; ◊, Stanley et al. [36]; △, Stanley et al. [37]; ■, Heinz and Hoffman [16]; ◆, Heinz and Hoffman [17]).

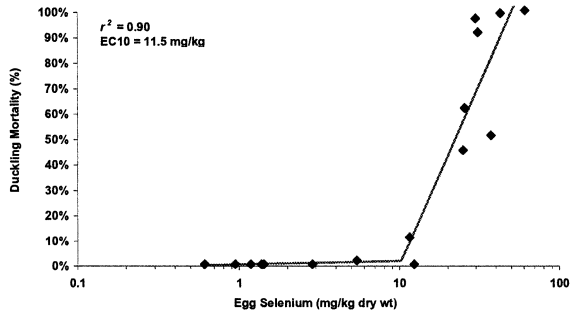

Fig. 3. Hockey stick regression of laboratory mallard duckling survival at 7 d posthatch. EC10 = 10% effect concentration.

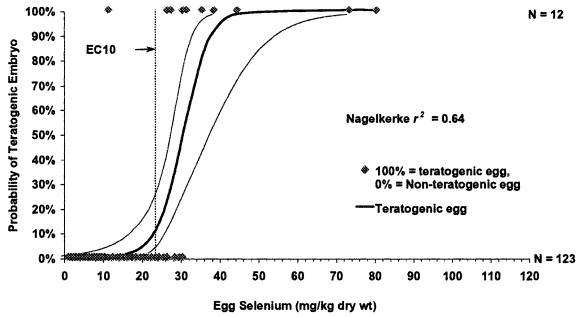

Fig. 4. Logistic model fit to duck teratogenicity field data. EC10 = 10% effect concentration. Dataset from Skorupa [29].

than the range for the egg inviability data. In fact, egg inviability occurred over almost the entire range of egg selenium concentrations, including very low concentrations. This caused the slope of the logistic model for inviability to be more shallow, and the goodness-of-fit measures to be lower, than the respective statistics for any of the species teratogenicity models (Table 3).

In the present analysis, the data as well as the model curves are shown, and the goodness-of-fit measures, Nagelkerke r^2 and Dhrymes r^2, are presented for all three models (Figs. 4–6). This yielded interesting results, in that the goodness-of-fit estimates (r^2) were much lower for the inviability logistic model than for the teratogenesis logistic models. Nagelkerke r^2 ranged from 0.514 (avocet) to 0.639 (duck) for the teratogenesis models but was only 0.284 for the stilt clutch inviability data (Table 3). Dhrymes r^2 ranged from 0.468 (stilt) to 0.567 (duck) for the teratogenesis models but was only 0.184 for the stilt clutch inviability model (Table 3). The low r^2 values for the stilt chick inviability model goodness-of-fit measures indicate either that the fitted logistic curve does not characterize the data well or that considerable scatter around the model exists.

Three-point moving average applied to field data

Three-point moving average responses were calculated for each whole-number egg selenium concentration (2, 3, 4, . . . mg/kg dry wt). Initially, a plot of the dose–response relationship for concentrations up to 9 mg/kg dry weight for stilt clutch inviability (Fig. 7) [29] showed an increase in the response from 6 to 17% between 6 and 7 mg/kg dry weight, providing support for the proposed egg selenium threshold of 6 mg/kg dry weight. However, extending the analysis beyond 9 mg/kg dry weight puts this increase in perspective and demonstrates that it is within the background variability of the dose–response relationship for concentrations less than 30 mg/kg dry weight (Fig. 7). For example, the increase in response between 6 and 7 mg/kg dry weight is identical to that between 27 and 28 mg/kg dry weight (i.e., from 6–17%). In general, the response is variable for low selenium concentrations, with no consistent relationship beginning until concentrations near 30 mg/kg dry weight (Fig. 7). Presentation of the entire dataset allows for an interpretation of the data that differs from that published by Skorupa [29].

Hockey stick regressions applied to field and laboratory data (egg inviability data)

Constrained hockey stick models, with zero slope at concentrations less than τ, fit all the datasets and produced r^2 values generally greater than 0.90. The r^2 values adjusted for the number of model parameters supported use of the hockey stick model with only an intercept for selenium concentrations less than τ, as opposed to a separate, additional slope estimate. Tables 4 and 5 present the ECτ, the egg selenium concentration (mg/kg dry wt) at τ, the model EC10, and the r^2 value for the teratogenic and inviability data, respectively. The ECτ represents the minimum effect concentration the model could distinguish from background for that binning scheme.

For the teratogenicity endpoint, which is selenium-specific, the ECτ occurred at approximately the 1% effect level (EC1) for all binning schemes except for the laboratory data, for

Table 3. Binary logistic model summary statistics for teratogenesis (duck, stilt, and avocet) and clutch inviability (stilt) using field data

	Teratogenesis			Inviability
	Duck egg	Stilt egg	Avocet egg	Stilt clutch
No. of cases	135	608	572	409
Not affected (0)	123	541	564	290
Affected (1)	12	67	8	119
Model statistics				
Intercept (standard error)	−8.97 (2.34)	−6.13 (0.58)	−7.48 (1.18)	−2.32 (0.22)
Slope (standard error)	0.30 (0.09)	0.11 (0.01)	0.07 (0.01)	0.05 (0.01)
−2 Ln(likelihood)				
Intercept only in model	80.99	421.86	84.20	493.26
Intercept and slope in model	35.11	224.54	42.46	402.54
Reduction	45.88	197.32	41.74	90.72
Goodness of fit				
Nagelkerke r^2	0.639	0.554	0.514	0.284
Dhrymes r^2	0.567	0.468	0.496	0.184

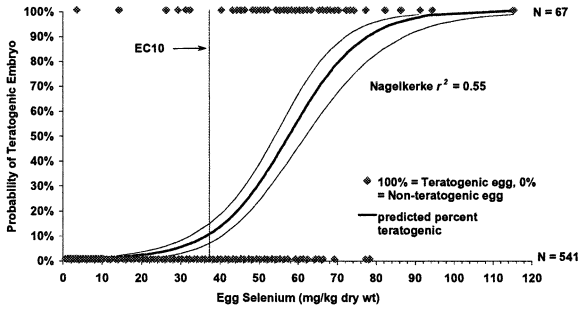

Fig. 5. Logistic model fit to stilt teratogenicity field data. Dataset obtained from Skorupa [29]; EC10 = 10% effect concentration.

which it was approximately 0.1%. Egg selenium concentrations corresponding to ECτ ranged from 13.5 to 23.4 mg/kg dry weight for the different binning schemes applied to the field data and was 17.4 mg/kg dry weight for the laboratory data. The calculated EC10s for the field data ranged from 15.5 to 24.3 mg/kg dry weight, and the laboratory EC10 was 21 mg/kg dry weight (Table 4). In each case, the EC10s were greater than ECτ, indicating that the EC10s do not fall within the range of background responses. A representative example of the hockey stick model (binning scheme A) is shown in Figure 8. Several inferences can be made from these data. First, the laboratory teratogenicity data appear to be slightly less sensitive than the field data. Second, once egg selenium concentrations are sufficiently high to elicit a teratogenic response, the response is very steep, because the ECτ, which approximates the EC1 for these data, is nearly the same as the estimated EC10.

The stilt clutch inviability data were adjusted to account for effects on individual eggs using the *AE*:*AH* ratio equation (Eqn. 1). Figure 9 provides a representative example of the dose–response relationship observed for the egg inviability endpoint using hockey stick regression. Selenium egg concentrations at τ ranged from 19.5 to 44.7 mg/kg dry weight for the different binning schemes, and EC10 values for stilt egg inviability field data ranged from 20.9 to 31.0 mg/kg dry weight (Table 5). The ECτ was in the range of 7.3 to 15.6%, with binning schemes B and E having ECτ values greater than 10%, which precluded estimation of an EC10. In comparison to the field data, the egg selenium concentration at τ was 10.2 mg/kg dry weight for the laboratory data, and the EC10 was 12.3 mg/kg dry weight. Inferences from these analyses are that the laboratory egg inviability data are more sensitive than the field data, the slope of the dose–response line for the laboratory data is much steeper than the field data, and the egg selenium concentration associated with field egg inviability is fairly high

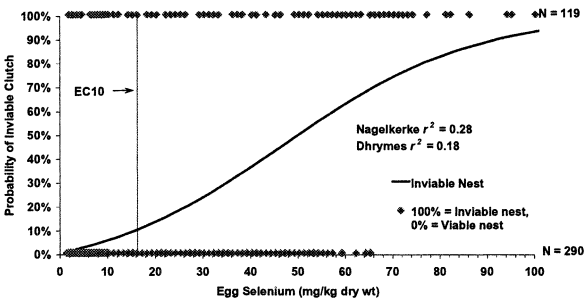

Fig. 6. Logistic model fit to stilt clutch inviability field data. Dataset obtained from Skorupa [29]. EC10 = 10% effect concentration.

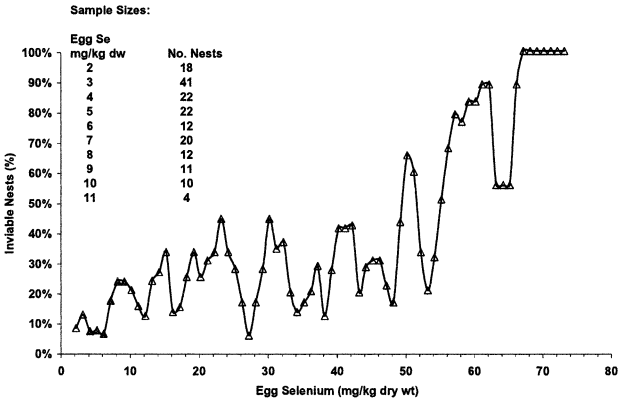

Fig. 7. Comparison of clutch inviability with egg selenium concentrations using the complete dataset from Skorupa [29]. EC10 = 10% effect concentration.

because of the variability in the field dataset (EC10s of 21–31 mg/kg dry wt, depending on the binning scheme).

DISCUSSION

The use of mean egg selenium thresholds for regulating point and nonpoint sources of selenium in water has become increasingly common in the western United States. The state-of-the-science and regulatory approaches for protecting wildlife and aquatic life resources have been evolving in the direction of incorporating tissue residue–based thresholds or criteria. Toxicologically, a tissue residue–based endpoint appears to be the most appropriate, because it integrates the effects of multiple factors (e.g., selenium speciation, biotransformation, species feeding habits) that ultimately influence exposure to sensitive receptors (e.g., birds and fish). It is critical, therefore, that the mean egg selenium threshold used for regulatory purposes be scientifically robust, as has been the case in developing regulatory approaches for water-quality and wildlife criteria.

Table 4. Constrained hockey stick models fit to laboratory and field teratogenesis data

Binning scheme	r^2	Egg selenium at τ (mg/kg dry wt)	ECτ (%)[a]	EC10 (mg/kg dry wt)[b]
A[c]	1.00	21.9	0.67	23.0
B[d]	0.98	23.4	1.11	24.3
C[e]	0.99	20.9	0.96	23.3
D[f]	1.00	18.6	0.81	21.3
E[g]	0.83	13.5	1.11	15.5
Lab (mallard)	0.79	17.4	0.15	21.3

[a] ECτ = percentage effects predicted to occur at quantification of the breakpoint (τ).
[b] EC10 = 10% effect concentration.
[c] Binning scheme A = 0 to <2, 2 to <4, 4 to <8, 8 to <16 mg/kg dry weight, etc.
[d] Binning scheme B = 0 to <5, 5 to <10, 10 to <15, 15 to <20 mg/kg dry weight, etc.
[e] Binning scheme C used 11 eggs/bin and the median egg selenium concentration for 12 bins for teratogenicity (see *Materials and Methods*).
[f] Binning scheme D = 0 to <3, 3 to <6, 6 to <12, 12 to <24 mg/kg dry weight, etc.
[g] Binning scheme E used the break points identified by the U.S. Department of the Interior [11], namely median selenium concentrations of 2.7, 9.2, 21.5, 31, and 76.5 for stilt teratogenesis.

Table 5. Constrained hockey stick models fit to black-necked stilt field egg inviability data

Binning scheme	r^2	Egg selenium at τ (mg/kg dry wt)	ECτ (%)[a]	EC10 (mg/kg dry wt)[b]
A	0.94	19.5	7.3	20.9
B	0.93	44.7	15.6	NE[c]
C	0.90	30.9	9.8	31.0
D	0.97	30.2	9.5	30.7
E	0.95	43.7	14.7	NE
Lab (mallard)	0.88	10.2	0.0	12.3

[a] ECτ = percentage effects predicted to occur at quantification of the breakpoint (τ).
[b] EC10 = 10% effect concentration.
[c] NE = not evaluated.

In this paper, we have critically reviewed all available data that could be used to derive an avian mean egg selenium chronic toxicity threshold. A threshold of 6 mg/kg dry weight based on a clutch inviability endpoint for black-necked stilts using field data primarily collected from central California has been proposed [10]. This threshold was defined using a GLiM approach to calculate the EC3 and also employed a three-point moving average analysis as supportive evidence. We have compared and contrasted the statistical reliability of this threshold with results determined using a variety of models (GLiM, three-point moving average, and hockey stick regressions) for both field and laboratory data.

Examination of the clutch inviability endpoint using GLiM analysis indicates that using a binary logistic model, a relatively limited amount of the variability observed in field responses (19–28%) is explained by selenium concentrations in eggs. The low r^2 is not unexpected for an endpoint generated from field data in which the effect (reduced hatching success) is not selenium-specific and can be confounded by other factors, such as weather, disease, and other contaminants. In contrast, for the selenium-specific teratogenicity endpoint, r^2 values in the range of 60% were observed for both species of birds using field data and GLiMs. The r^2 values were somewhat higher using hockey stick regression analysis. These values are much more consistent with the level of precision expected from field data when a relationship is known to exist. Given the low r^2 observed for the field clutch inviability data, we believe that these data lack the necessary robustness on which to derive a chronic threshold.

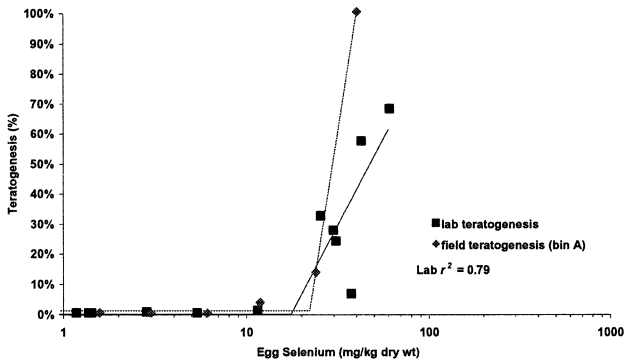

Fig. 8. Hockey stick model fit to laboratory mallard and field stilt teratogenicity data (binning scheme A). EC10 = 10% effect concentration.

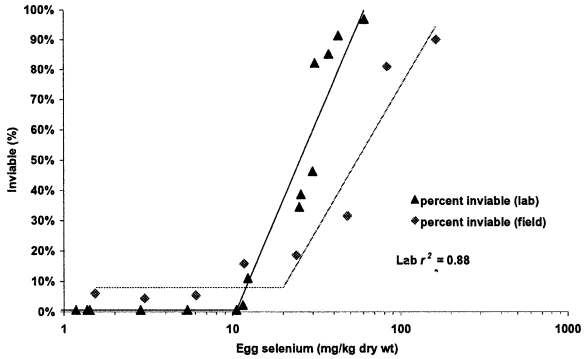

Fig. 9. Hockey stick model fit to laboratory mallard and field stilt egg inviability data (binning scheme A).

A second analysis of the clutch inviability data used by Skorupa [29] to support a threshold of 6 mg/kg dry weight involved plotting inviability as a function of a three-point moving average of egg selenium concentrations. The present analysis shows a significant increase in clutch inviability between 6 and 7 mg/kg dry weight, as reported by Skorupa [29]. However, plotting the entire dataset shows an equally significant decrease in clutch inviability at 12 mg/kg dry weight, with continued oscillations in the response variable until approximately 30 mg/kg dry weight, when a consistent increase in clutch inviability is observed. Analysis of the dataset across all egg selenium concentrations does not support 6 mg/kg dry weight as a threshold.

As a final analysis, we binned the egg selenium data in multiple ways, converted the data from a clutch-based to an egg-based response, and used hockey stick regression to determine the egg selenium concentration at which egg inviability increased above background levels. Several important inferences can be made from this analysis. Most important is quantification of the breakpoint (τ) in the regression. This is the point at which the dose–response relationship becomes distinguishable from background and, as a result, defines the effect concentration (i.e., EC*x* defined here as ECτ) that can be estimated reliably for the dataset. This is somewhat analogous to determining the minimum significant difference in hypothesis testing.

Results from this analysis indicate τ to be in the range of the EC7 to EC16 for the field egg inviability data. This is important, because the proposed threshold from Skorupa [10] is 6 mg/kg dry weight based on an EC3. Our analysis indicates that the EC3 is less than τ and cannot be reliably distinguished from background for any of the response scenarios. The EC10, which is a more commonly used toxicity threshold, can be estimated reliably for three of the five binning schemes. We believe this analysis adds to the weight of evidence that an EC3 of 6 mg/kg dry weight cannot be reliably distinguished from naturally occurring effects based on the available data.

A comparison of the EC10 values for egg inviability to the same effect level for the teratogenicity data also is of interest. As has been well documented in the laboratory, egg inviability is a significantly more sensitive endpoint than teratogenesis [13,15]. However, evaluation of the hockey stick regressions for the field data shows a different trend, with two of the three binning schemes for which an egg inviability EC10 could be estimated indicating teratogenesis to be more sensitive than egg inviability. We do not interpret this to mean that teratogenesis is actually the more sensitive endpoint, because lab-

Table 6. Summary of laboratory and field-derived selenium thresholds for avian species

Endpoints	Lab/field data	EC10 (mg/kg dry wt)[a]	Type of analysis	r^2	Reference
Teratogenicity					
Stilt teratogenicity	Field	37	Logit[b]	0.55	[10]
Stilt teratogenicity	Field	16–24	Hockey stick[c]	0.83–1.0	Present study
Duck teratogenicity	Field	23	Logit[b]	0.64	[10]
Mallard teratogenicity	Lab	26	Weibull[c]	0.84	[13]
Mallard teratogenicity	Lab	23	Logit[d] & Probit[b]	0.76	Present study
Mallard teratogenicity	Lab	21	Hockey stick[c]	0.79	Present study
Inviability/mortality					
Stilt clutch inviability	Field	16	Logit[b]	0.28	[10]
Stilt egg inviability	Field	21–31	Hockey stick[c]	0.90–0.97	Present study
Stilt egg inviability	Field	24	Logit[b]	N/A	Present study
Mallard egg inviability	Lab	12	Hockey stick[c]	0.88	Present study
Mallard duckling mortality	Lab	16	Probit[e]	0.86	[13]
Mallard duckling mortality	Lab	14	Probit[b]	0.75	Present study
Mallard duckling mortality	Lab	15	Logit[b]	0.79	Present study
Mallard duckling mortality	Lab	12	Hockey stick[c]	0.90	Present study

[a] EC10 = 10% effect concentration.
[b] Generalized linear model (GLiM) with logit or probit link and binomial error.
[c] Nonlinear model.
[d] EC10.
[e] GLiM with probit transformation.

oratory data and sound underlying toxicological principles support why egg inviability would be more sensitive. Rather, this highlights the considerable variability in the field egg inviability data and the subsequent inability to distinguish selenium-related effects from background until excessively high egg selenium concentrations are observed.

Finally, a comparison of EC10 values for the mallard laboratory data evaluated using hockey stick regression, probit, and logit analyses showed similar results for egg inviability and duckling mortality (12–16 mg/kg dry wt) (Table 6). These analyses indicate that the laboratory egg inviability and duckling mortality endpoints are more sensitive than teratogenesis, which is consistent with previously drawn conclusions.

In summary, a review of the key teratogenicity endpoints for stilts and ducks using laboratory and field data indicates fairly good agreement between the EC10 for all groups. The EC10 ranges from 16 to 24 mg/kg dry weight (present study) across species groups and laboratory and field data (Table 6). The hockey stick regression analyses indicate that the threshold egg concentration of selenium for teratogenic effects in the field for stilts might be as low as 15.5 mg/kg dry weight. In contrast, the hockey stick regression analyses for stilt egg inviability using field data indicate that the mean egg selenium threshold (EC10) lies in the range of 21 to 31 mg/kg dry weight, which is considerably higher than has been previously reported. These elevated EC10 values reflect the variability that exists in the field dataset for egg inviability; they do not represent the true potential for effects on egg inviability because of selenium exposure. Evaluation of the egg inviability and duckling mortality endpoints for mallard laboratory data indicates that the mean egg selenium EC10 ranges from 12 mg/kg dry weight based on hockey stick regression analysis to 14 to 15 mg/kg dry weight based on logit and probit and analyses.

Given that a mean egg selenium threshold in the range of 12 to 14 mg/kg dry weight appears to be the most appropriate, the question of how such a threshold should be used requires comment. First, it should be noted that this threshold is based on data for mallards, the most sensitive species of the few that have been tested to date. Application of this threshold to stilts or avocets may not be appropriate. This is particularly true for avocets, which appear to be substantially less sensitive than mallards.

Second, the question is raised regarding what exposure data should be used for comparison to the threshold. Given that the threshold is a mean egg selenium value, it will be most appropriate to compare the threshold against a mean egg selenium value for a given study area. However, caution must be used in defining a study area, because significant heterogeneity in terms of dietary exposure and subsequent egg selenium concentrations could lead to inaccurate assessments of potential selenium impacts. An evaluation of the Heinz et al. [15] dataset, which provides a significant portion of the data used to derive the threshold, indicates a mean coefficient of variation in egg selenium concentrations within treatment levels of 25% (range, 20–29% for all treatments). We believe it is appropriate for field data to exhibit the same level of homogeneity (i.e., coefficient of variation $\leq$ 30%) for comparison to the mean egg selenium threshold. For study locations with variability in egg selenium concentrations greater than 30%, it would be more appropriate to evaluate subgroups, as the high variability indicates subgroups may be at greater risk than indicated by the mean egg selenium concentration for the group as a whole.

In conclusion, considering the weight of evidence described above, we believe that the field data used to derive an egg selenium threshold of 6 mg/kg dry weight are seriously confounded by factors other than selenium. This leads to excessive variability; consequently, the field clutch/egg inviability data are not sufficiently reliable for establishing a chronic egg selenium threshold for effects. In contrast, evaluation of the lab-

oratory chronic data for mallards using probit, logit, and hockey stick regressions provide an EC10 mean egg selenium threshold that can be estimated reliably. Given the controlled conditions under which the laboratory data were derived and the similarity of EC10 values obtained using multiple statistical analyses, a mean egg selenium threshold for effects of 12 to 14 mg/kg dry weight based on laboratory data appears to be both appropriate and conservative.

Acknowledgement—The authors acknowledge Kennecott Utah Copper for funding this study, an anonymous reviewer for providing helpful comments, and Debbie Fetherston for assistance in preparing this manuscript.

REFERENCES

1. Franke KW, Tully WC. 1935. A new toxicant occurring naturally in certain samples of plant foodstuffs. V. Low hatchability due to deformities in chicks. *Poult Sci* 14:273–279.
2. Poley WE, Moxon AL, Franke KW. 1937. Further studies on the effects of selenium poisoning on hatchability. *Poult Sci* 16:219–225.
3. Ohlendorf HM, Kilness AW, Simmons JL, Stroud RK, Hoffman DJ, Moore JF. 1988. Selenium toxicosis in wild aquatic birds. *J Toxicol Environ Health* 24:67–92.
4. Adams WJ, Brix KV, Cothern KA, Tear LM, Cardwell RD, Fairbrother A, Toll JE. 1998. Assessment of selenium food chain transfer and critical exposure factors for avian wildlife species: Need for site-specific data. In Little EE, Delonay AJ, Greenberg BM, eds, *Environmental Toxicology and Risk Assessment,* Vol 7. American Society for Testing and Materials, Philadelphia, PA, pp 312–342.
5. Lemly AD, Smith GJ. 1987. Aquatic cycling of selenium: Implications for fish and wildlife. U.S. Fish and Wildlife Service, Washington, DC.
6. Hothem RL, Ohlendorf HM. 1989. Contaminants in foods of aquatic birds at Kesterson Reservoir, California, 1985. *Arch Environ Contam Toxicol* 18:773–786.
7. Ohlendorf HM, Kilness AW, Simmons JL, Stroud RK, Hoffman DJ, Moore JF. 1993. Food-chain transfer of trace elements in wildlife. In Allen RG, Neale CMU, eds, *Management of Irrigation and Drainage Systems: Integrated Perspectives.* American Society of Civil Engineers, Park City, UT, pp 596–803.
8. Hoffman DJ, Ohlendorf HM, Aldrich TW. 1988. Selenium teratogenesis in natural populations of aquatic birds in Central California. *Arch Environ Contam Toxicol* 17:519–525.
9. Hoffman DJ, Heinz GH. 1988. Embryotoxic and teratogenic effects of selenium in the diet of mallards. *J Toxicol Environ Health* 24:477–490.
10. Skorupa JP. 1998. Risk assessment for the biota database of the National Irrigation Water Quality Program. U.S. Fish and Wildlife Service, Sacramento, CA.
11. U.S. Department of the Interior. 1998. Guidelines for interpretation of the biological effects of selected constituents in biota, water, and sediment Selenium. National Irrigation Water Quality Program Information Report 3. Denver, CO, pp 139–184.
12. Fairbrother A, Brix KV, Toll JE, McKay S, Adams WJ. 1999. Egg selenium concentrations as predictors of avian toxicity. *Hum Ecol Risk Assess* 5:1229–1253.
13. Fairbrother A, Brix KV, De Forest DK, Adams WJ. 2000. Egg selenium thresholds for birds: A response to J. Skorupa's critique of Fairbrother et al. 1999. *Hum Ecol Risk Assess* 6:203–212.
14. Ohlendorf HM. 2002. Ecotoxicology of selenium. In Hoffman DJ, ed, *Handbook of Ecotoxicology.* Lewis, Boca Raton, FL, USA, pp 465–500.
15. Heinz GH, Hoffman DJ, Gold LG. 1989. Impaired reproduction of mallards fed an organic form of selenium. *J Wildl Manag* 53: 418–428.
16. Heinz GH, Hoffman DJ. 1996. Comparison of the effects of seleno-L-methionine, seleno-DL-methionine, and selenized yeast on reproduction of mallards. *Environ Pollut* 91:169–175.
17. Heinz GH, Hoffman DJ. 1998. Methylmercury chloride and selenomethionine interactions on health and reproduction in mallards. *Environ Toxicol Chem* 17:139–145.
18. Bailer AJ, Oris JT. 1997. Estimating inhibition concentrations for different response scales using generalized linear models. *Environ Toxicol Chem* 16:1554–1559.
19. Mathsoft. 1997. *SPlus Software.* Seattle, WA, USA.
20. SPSS. 1999. *SPSS Software.* Chicago, IL, USA.
21. McCullagh P, Nelder JA. 1989. *Generalized Linear Models,* 2nd ed. Chapman & Hall, New York, NY, USA.
22. Hagle TM, Mitchell GE. 1992. Goodness-of-fit measures for probit and logit. *Am J Pol Sci* 36:762–784.
23. Menard S. 2000. Coefficients of determination for multiple logistic regression analysis. *American Statistics* 54:17–24.
24. Dhrymes PJ. 1986. Limited dependent variables. In Griliches Z, Intrilagator MD, eds, *Handbook of Econometrics.* North-Holland, Amsterdam, The Netherlands, pp 1567–1631.
25. Hosmer DW, Lemeshow S. 1989. *Applied Logistic Regression.* John Wiley, New York, NY, USA.
26. Aitkin M, Anderson D, Francis B, Hinde J. 1989. *Statistical Modeling in GLIM*. Oxford University Press, New York, NY, USA.
27. Christensen R. 1990. *Log-Linear Models.* Springer-Verlag, New York, NY, USA.
28. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. 1997. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 16:965–980.
29. Skorupa JP. 1998. Selenium poisoning in fish and wildlife in nature: Lessons from 12 real-world examples. In Frankenberger WT, Engberg RA, eds, *Environmental Chemistry of Selenium.* Marcel Dekker, New York, NY, USA, pp 315–354.
30. Cox C. 1987. Threshold dose–response models in toxicology. *Biometrics* 43:511–523.
31. Horness BH, Lomax DP, Johnson LL, Meyers MS, Pierce SM, Collier TK. 1998. Sediment quality thresholds: Estimates from hockey stick regression of liver lesion prevalence in English sole (*Pleuronectes vetulus*). *Environ Toxicol Chem* 17:872–882.
32. Piegorsch WW, Bailer AJ. 1997. *Statistics for Environmental Biology and Toxicology.* Chapman & Hall, London, UK.
33. Finney DJ. 1971. *Probit Analysis,* 3rd ed. Cambridge University Press, London, UK.
34. Palisade. 1997. *Best Fit.* Newfield, NY, USA.
35. Heinz GH, Hoffman DJ, Krynitsky AJ, Weller DMG. 1987. Reproduction in mallards fed selenium. *Environ Toxicol Chem* 6: 423–433.
36. Stanley TR, Spann JW, Smith GJ, Roscoe R. 1994. Main and interactive effects of arsenic and selenium on mallard reproduction and duckling growth and survival. *Arch Environ Contam Toxicol* 26:444–451.
37. Stanley TR, Smith GJ, Hoffman DJ, Heinz GH, Rosscoe R. 1996. Effects of boron and selenium on mallard reproduction and duckling growth and survival. *Environ Toxicol Chem* 15:1124–1132.