Ultrasound in Med. & Biol., Vol. 30, No. 10, pp. 1289-1295, 2004

0301-5629/04/$–see front matter

ELSEVIER

doi:10.1016/j.ultrasmedbio.2004.07.007

● *Original Contribution*

EXCESS RISK THRESHOLDS IN ULTRASOUND SAFETY STUDIES: STATISTICAL METHODS FOR DATA ON OCCURRENCE AND SIZE OF LESIONS

DOUGLAS G. SIMPSON,* MOON-HO HO,† YAN YANG,* JIANHUI ZHOU,* JAMES F. ZACHARY‡ and WILLIAM D. O'BRIEN, JR.§

*Department of Statistics, University of Illinois, Champaign, IL, USA; †Department of Psychology, McGill University, Montreal, QUE, Canada; ‡Department of Veterinary Pathobiology; and §Bioacoustics Research Laboratory, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL, USA

(*Received* 7 *July* 2003; *revised* 18 *June* 2004; *accepted* 8 *July* 2004)

Abstract—Concerns about the safe use of clinical ultrasound (US) at diagnostic pressure levels (below a mechanical index, or MI, = 1.9) have stimulated considerable research in US risk assessment. The objective of the present study was to develop probability-based risk thresholds for US safety studies, to present statistical methods for estimating the thresholds and their standard errors and to compare these methods with the analysis based on a piecewise linear ("hockey stick") model. The excess risk at exposure level $x > 0$ was defined as the relative increase in the probability of a lesion at that level compared with the background probability of a lesion at exposure $x = 0$. The risk threshold was then defined as the exposure level at which the excess risk exceeded a specified level (e.g. 5% or 50%). Thus, given pressure-dependent estimates of the excess risk, the thresholds were estimated by solving the risk equation to obtain the pressure at which the target level of excess risk occurs. Threshold estimates of this type have been developed extensively in the literature for incidence (presence or absence) data. Only recently, however, have excess risk threshold estimates been derived for data in which lesion size (depth, surface area) is measured if present and a zero is recorded if the lesion is absent. Tobit regression was used to estimate pressure-dependent percentiles of the size distribution, and the excess risks were estimated from the tobit probability of a positive-valued response. The tobit model provides a well-established approach to modeling data constrained to be nonnegative. Solving the risk equation for the tobit model leads to risk threshold estimates that incorporate the information on size of observed lesions. Results using these probability-based risk estimates were compared with results for a piecewise linear ("hockey stick") model, which has also been used in the US safety literature, although it does not explicitly address the nonnegativity constraint in the sampling model. The comparisons were carried out for data from two previously published studies, from different laboratories, on US-induced lung hemorrhage. The thresholds derived from logistic regression of lesion occurrence and tobit regression of lesion size were quite consistent with each other and within sampling error. The hockey stick thresholds, defined as the exposure level at which the piecewise linear model for the probability of the expected size of a lesion bends upward, corresponded to quite different excess risk values for incidence (lesion occurrence) compared with size (lesion surface area or depth), although these methods have been developed previously for both types of data. The use of probability-based excess risk thresholds is recommended to obtain consistent incidence vs. size thresholds and to ensure that the thresholds are well-defined and interpretable independent of the details of the statistical model. (E-mail: dgs@uiuc.edu) **© 2004 World Federation for Ultrasound in Medicine & Biology.**

*Key Words***Excess risk, Statistical comparison, Thresholds, Ultrasound risk, Lung hemorrhage, Logistic regression, Tobit regression.**

INTRODUCTION

Concerns about the safe use of clinical ultrasound (US) at diagnostic pressure levels (i.e., below a mechanical index, or MI, = 1.9) have stimulated considerable research in threshold estimation and US risk assessment. A number of studies in different animal species have been performed to estimate occurrence and size thresholds for US-induced lung hemorrhage (Zachary and O'Brien 1995; Baggs et al. 1996; Raeman et al. 1996; Dalecki et al. 1997a, 1997b; O'Brien and Zachary 1997; O'Brien et al. 2001a; Zachary et al. 2001). In these experiments, lung was exposed to a focused US beam. Groups of animals were exposed to US at different pressure levels according to standard experimental designs. The range of US pressures within an ex-

Address correspondence to: Professor Douglas G. Simpson, Department of Statistics, University of Illinois, 725 S. Wright Street, Champaign, IL 61820 USA. E-mail: dgs@uiuc.edu

periment allowed for the estimation of a threshold for the response. In addition, by introducing variations in other conditions, such as the frequency, beam width and species and age of animal, it has been possible to investigate the association between these other factors and the susceptibility to US-induced lung hemorrhage.

The purpose of the present paper was to propose probability-based excess risk thresholds for US research, to present statistical methods for estimating the risks and their uncertainties and to compare the proposed probability-based methods with a commonly used existing method. Published data from two laboratories provide an empirical basis for comparison and illustration of the methods.

The earliest risk model in the US safety literature, sometimes referred to as the "hockey stick" model, entails fitting a piecewise linear model to lesion size or incidence as a function of US exposure (Baggs et al. 1996; Raeman et al. 1996; Dalecki et al. 1997a, 1997b). The mean response is assumed to be flat up to a threshold value and to increase linearly beyond the threshold. The model is characterized by three parameters: they are the constant background level (below the threshold), the threshold value and the slope of the expected response above the threshold. The model has a physical constraint that the expected lesion size or probability of occurrence of a lesion is nonnegative. The probability model has an additional constraint that the probability of a lesion cannot exceed 1. The point at which the hockey stick model bends upward as a function of US exposure (usually a peak acoustic pressure) has been taken as the threshold for the response and has been suggested as a target value for setting safe levels of US exposure (Baggs et al. 1996).

An alternative model has been employed in several recent studies (Zachary et al. 2001; O'Brien et al. 2001, 2003). In these studies, a logistic regression model (Agresti 2002) is used to model the probability of occurrence of a lesion as a function of US exposure and a tobit regression model (Amemiya 1984) is used to model the size of the lesion. The logistic regression model implies that the log-odds of a lesion depend linearly on the exposure level, whereas the tobit model implies that the median lesion size (depth, area) follows a piecewise linear model as a function of exposure (O'Brien et al. 2003). The error distribution for the tobit model is a Gaussian distribution truncated below at 0. The tobit model for lesion size provides both an exposure-dependent probability of a lesion and an expected size of the lesion conditional on its occurrence. In statistical terminology, the tobit model has the structure of a censored regression model, so techniques for censored regression enable the estimation of the model (Galfalvy and Simpson 1999). Using logistic regression or tobit regression, the risk threshold is defined as the exposure level giving a specified probability of a response. For safety studies, the risk threshold is set at a relatively low probability such as 5%, whereas, for effect studies, a higher threshold such as 50% may be more relevant. Although the estimation of the threshold depends on the model, its definition is independent of the model. In the toxicology literature, the dose leading to an effect probability of 100p% is often called the ED100p (effective dose) (Simpson et al. 1996), where p is a specified probability of an adverse event (e.g., 0.05 or 0.50).

Variation in threshold estimates between the studies cited above raises the question of whether the difference is due to differences in statistical methods, differences in experimental protocols, or a combination of these differences. To address this issue, the present study compares statistical results for the two approaches applied to the same data. This comparison sheds light on the impact of the statistical models and provides information on the interpretation of results between methods and studies.

MATERIALS AND METHODS

The probability-based threshold modeling approach is defined by selecting a specific excess risk of lesions compared with the probability of lesions in the absence of exposure, that is, the background rate (Chen and Kodell 1989; Ryan 1992; Xie and Simpson 1999). Let $P(x)$ denote the probability of lesion occurrence for exposure level x. Then the background probability is $P(0)$, the excess risk at exposure level x is $P(x) - P(0)$, and the excess risk due to exposure is given by:

$$r(x) = \frac{P(x) - P(0)}{1 - P(0)}. \quad (1)$$

The ED100p is the "effective dose" that gives a 100p% excess risk due to exposure (i.e., it is the solution of the equation):

$$r(\mathrm{ED}100p) = p, \quad (2)$$

or, equivalently, the solution of the equation $P(\mathrm{ED}100p) = P(0) + p\{1 - P(0)\}$. If the background rate, $P(0)$, is zero, then $r(x) = P(x)$ and ED100p is the exposure level such that $P(\mathrm{ED}100p) = p$. Two commonly used thresholds are the ED50, corresponding to a 50% excess risk due to exposure, and the ED05, the exposure level giving a 5% excess risk probability.

If the measured response is the presence or absence of a lesion, then ED100p can be estimated by first fitting a logistic regression model (Agresti 2002) to estimate the functional form for $r(x)$ in eqn (1) and then solving eqn (2) for ED100p. The logistic regression model assumes the log-odds of $P(x)$ (the "logit") is a linear function of the exposure x in megapascals (MPa). The corresponding probability expression is:

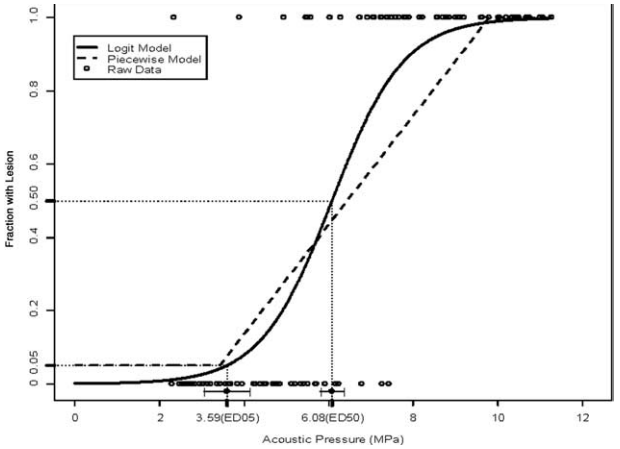

Fig. 1. Occurrence of lesions in neonate pigs (data from O'Brien et al. 2003) with estimated probability curves using logistic regression (including ED05 and ED50 with 1 SD bar for each) and piecewise linear regression [lesion incidence data are coded as 0 = no lesion, 1 = lesion].

$$P_L(x) = \frac{\exp(\alpha_0 + \alpha_1 x)}{1 + \exp(\alpha_0 + \alpha_1 x)}. \quad (3)$$

The model parameters α_0 and α_1 were estimated by maximum likelihood using the function glm in the R statistical programming language (Ihaka and Gentleman 1996). ED100p was computed by solving eqn (3) to obtain:

$$\text{ED100}p = \frac{\log\left(\frac{q}{1-q}\right) - \alpha_0}{\alpha_1}, \quad (4)$$

where $q = p + (1 - p)(1 - (1 + e^{\alpha_0})^{-1})$, and inserting the parameter estimates from the fitted model in eqn (4). Standard errors of the parameter estimates were computed by a standard Taylor series approximation. Figure 1 illustrates how the logistic regression model is used to estimate risk thresholds.

If the measured response is continuous, then the threshold can be estimated by fitting a tobit regression model (Amemiya 1984; O'Brien et al. 2001b, 2003) and then using the probability of lesion occurrence implied by the tobit model to estimate $r(x)$ and solve for the ED100p threshold (O'Brien et al. 2001b). Let Y denote the lesion depth or surface area. The tobit model entails that Y has the distribution of zero-censored response variable from a Gaussian linear model. In particular,

$$\begin{cases} Y = \max(0, Z) \\ Z \sim N(\beta_0 + \beta_1 x, \sigma^2), \end{cases} \quad (5)$$

where $N(\mu,\sigma^2)$ denotes the normal distribution with mean = μ and variance = σ^2. The tobit model in eqn (5) implies that, conditional on x, the 100pth percentile of the distribution of Y is

$$Q_{100p}(x) = \max(0, \beta_0 + \beta_1 x + z_p \sigma), \quad (6)$$

where z_p denotes the 100pth percentile of the standard normal distribution, $N(0,1)$. The tobit model also implies that the probability of a lesion, $P(x)$, is given by:

$$P_T(x) = \Pr(Y > 0) = \Pr(Z > 0) = \Phi\left(\frac{\beta_0 + \beta_1 x}{\sigma}\right), \quad (7)$$

where Φ is the standard normal distribution function given by $\Phi(t) = \int_{-\infty}^{1}(1/\sqrt{2\pi})e^{-u2/2}$ ***du*** for which numerical computing algorithms are widely available. Hence, the estimated tobit model can be used to estimate the excess risk and to solve for the risk threshold (O'Brien et al. 2003).

The conditional percentile curve in eqn (6) has a piecewise linear form, and the bend in the curve occurs precisely at the point where the probability of a lesion equals $1 - p$. The conditional median of Y, for example, corresponds to $z_{0.50} = 0$ and $Q_{50}(x) = \max(0, \beta_0 + \beta_1 x)$. As illustrated in Fig. 2, the bend in the 50th percentile curve occurs at the exposure where the probability of a lesion equals 50%.

Using these equations, the ED100p can be estimated from the tobit percentile curve. The tobit-based background ($x = 0$) probability of a lesion is $\Phi(\beta_0/\sigma)$. It therefore follows, from eqn (1) that an excess risk of $r(x) = p$ corresponds to a lesion probability of $p(x) = q = p + (1$

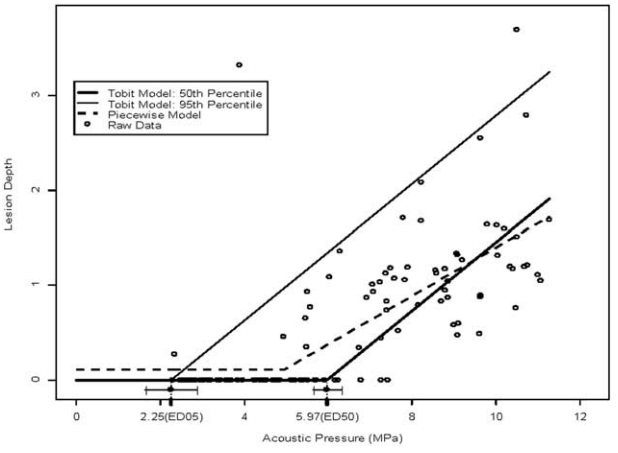

Fig. 2. Lesion depth in neonate pigs (data from O'Brien et al. 2003) with regression models superimposed: the tobit model gives an estimate of median lesion depth with ED05, ED50 and their SD bars, whereas the piecewise linear model gives an estimate of mean lesion depth.

$- p)\Phi(\beta_0/\sigma)$. Inserting the expression for $P_T(x)$ from eqn (7) into eqn (2) and solving for ED100p yields the formula,

$$ED100p = \frac{z_q\sigma - \beta_0}{\beta_1}, \quad (8)$$

where $q = p + (1 - p)\Phi(\beta_0/\sigma)$ and z_q is the 100qth percentile of the standard normal distribution. The tobit regression model was estimated using the function survreg in R (Ihaka and Gentleman 1996) and the threshold estimates derived from it were computed using the matrix operations in R. Figure 2 illustrates how the tobit regression model gives threshold estimates using the extra information on lesion size.

The methods proposed above were compared with the piecewise ("hockey stick") linear model developed by the acoustics group at Rochester, NY, USA (Baggs et al. 1996; Dalecki et al. 1997a, 1997b). For continuous outcomes, such as lesion surface area, the model assumes a constant expected lesion size up to a threshold exposure value beyond which the expected size is assumed to increase linearly with the acoustic pressure. For 0 to 1 outcomes, such as incidence of lesions, the probability of the outcome (occurrence of a lesion) is assumed to 1. follow the piecewise linear form, with a constant background probability of lesions, 2. have a linear increase in the probability above a threshold value, and 3. have a plateau at the exposure value where the probability reaches 100%. For continuous outcomes, the model was estimated by nonlinear least squares using the SAS NLIN procedure. For 0 to 1 outcomes, maximum likelihood estimates were computed using the iterative weighted-least-squares algorithm capability of the SAS NLIN procedure. The iterative weighted-least-squares algorithm for maximum likelihood estimation for this type of model is described in Cox (1984). Previous studies using this approach define the estimated threshold as the exposure level where the response function bends upward (Baggs et al. 1996; Dalecki et al. 1997a, 1997b).

In this model, the response Y (lesion occurrence or size) was assumed to follow a piecewise linear model for the expected value:

$$E(Y) = \gamma_0 + \gamma_1 \max(0, x - \tau), \quad (9)$$

where x is the acoustic pressure, γ_0 is the background value, γ_1 is the superthreshold slope, and τ is the location of the threshold. The nonlinearity of the model results from the product of γ_1 and τ in eqn (9). The interpretation of the parameters depends on whether Y denotes the size (depth, surface area) of a lesion or a 0/1 valued indicator for the presence of a lesion (1 for presence of a lesion and 0 for absence). If Y denotes the size, then eqn (9) is a model for the expected size. If Y is a 0 to 1 indicator, then $E(Y) = P(Y = 1)$. The threshold value, τ, has been suggested as a risk target (Baggs et al. 1996; Dalecki et al. 1997a, 1997b). This parameter corresponds to the bend in the hockey stick model. Figures 1 and 2 include the fitted hockey stick models for comparison to the logistic regression and tobit regression models.

The excess risks for the "hockey stick" thresholds were estimated by inserting their values for x in eqn (1) with $P(x)$ estimated by logistic regression for incidence data and by tobit regression for size data. The resulting probabilities provide a basis for comparing hockey stick thresholds with the thresholds based on the logistic and tobit regressions.

Data from two published studies were analyzed to compare statistical results and differences in effects due to different protocols. The first analysis was for data from a study of US-induced lung hemorrhage in mice (Dalecki et al. 1997b). Mice in three age groups (neonatal, juvenile, adult) were exposed to pulsed US with a range of peak positive acoustic pressures from 0 (sham) to 3.0 MPa. For each age category, there were 40 neonates, 20 juveniles and 10 adults at each exposure level (sham, 0.5 MPa, 1.0 MPa, 1.5 MPa, 2.0 MPa and 3.0 MPa). Each mouse was exposed for 3 min. Neonates were exposed at 10°C (cold) or 37°C (warm). Because the cold neonate data were reported to be unreliable, the analysis herein includes only data on 120 warm neonates, 120 juveniles and 60 adults. Acoustic pressures were corrected for chest wall attenuations as reported in the original study. The attenuation coefficients were 0 dB for neonates, 0.4 dB for juveniles and 1.5 dB for adults. Details of the protocol are given in Dalecki et al. (1997b).

The second data source was a study of US-induced lung hemorrhage in crossbred pigs (O'Brien et al. 2003). Pigs in three age groups (neonate, juvenile, young adult) were exposed to pulsed US that was focused on the lung. In this study, the exposure parameter was the in situ peak rarefactional pressure calculated from the in vitro peak rarefactional pressure by correcting for chest wall attenuation as described in O'Brien et al. (2003). Evaluation of the lungs for lesion occurrence and size was "blind" to the exposure conditions. Both lungs were exposed for each animal, but, in the present study, only the first lung exposure was analyzed to focus attention on the threshold estimation rather than on the peripheral issues of the order of exposure and the correlation between lungs from the same animal. A total of 301 pigs were exposed, including 108 neonates, 97 juveniles and 96 young adults. O'Brien et al. (2003) detailed the computations of attenuation due to chest wall thickness, along with descriptions of the age groupings and the US exposure protocol.

Table 1. Analysis of lesion incidence for Dalecki et al. (1997b) data: ED05, ED50, piecewise linear threshold and its excess risk (computed from logistic regression).

Data source	ED05 [SE] (MPa)	ED50 [SE] (MPa)	Piecewise [SE]* (MPa)	Piecewise excess risk
Warm neonate mice	0.29 [0.24]	1.56 [0.15]	0.63 [0.28]	0.14
Juvenile mice	0.56 [0.18]	1.47 [0.09]	0.92 [0.08]	0.15
Adult mice	0.68 [0.12]	1.00 [0.07]	0.72 [0.10]	0.07

* Piecewise threshold values and standard errors reported by Dalecki et al. (1997b).

RESULTS

Threshold estimates for lung hemorrhage in the age-dependent mouse study (Dalecki et al. 1997b) are summarized in Tables 1 and 2.

Table 1 compares the age-dependent ED05, ED50 and piecewise linear (hockey stick) thresholds for incidence of lesions. Also shown are the logistic regression-based estimates of excess risk for the piecewise linear thresholds. The piecewise linear thresholds for incidence are close but somewhat higher than the ED05 and the estimated excess risks at the incidence threshold values range from 7% to 15%.

Table 2 compares age-dependent ED05, ED50 and piecewise linear thresholds derived from the surface areas of the lesions. Note that ED05 and ED50 values for lesion area in the second and third columns of Table 2 are within sampling errors of the corresponding columns for incidence in Table 1. This result is an indication that the logistic and tobit models provide consistent fits to the data (O'Brien et al. 2003, page 158). In addition, the agreement between the estimated thresholds (ED05 and ED50) from logistic models and those from tobit models illustrates the model-independent characteristic of the probability-based risk thresholds because similar threshold estimates are provided, no matter what model is adopted. In contrast to the piecewise linear thresholds for lesion incidence, which are close to ED05, those for lesion area are closer to the ED50, with estimated excess risks ranging from 47% to 74% at the threshold values.

Table 2. Analysis of lesion area (mm^2) for Dalecki et al. (1997b) data: ED05, ED50, piecewise linear threshold and its excess risk (computed from tobit regression).

Data source	ED05 [SE] (MPa)	ED50 [SE] (MPa)	Piecewise [SE]* (MPa)	Piecewise excess risk
Warm neonate mice	0.27 [0.14]	1.25 [0.10]	1.35 [0.05]	0.56
Juvenile mice	0.58 [0.10]	1.27 [0.07]	1.24 [0.10]	0.47
Adult mice	0.88 [0.03]	1.00 [0.02]	1.05 [0.04]	0.74

* Piecewise threshold values and standard errors reported by Dalecki et al. (1997b).

Table 3. Analysis of lesion incidence for O'Brien et al. (2003) data: ED05, ED50, piecewise linear threshold and its excess risk (computed from logistic regression).

Data source	ED05 [SE] (MPa)	ED50 [SE] (MPa)	Piecewise [SE] (MPa)	Piecewise excess risk
Neonate pigs	3.59 [0.53]	6.08 [0.27]	3.44 [0.55]	0.042
Juvenile pigs	5.87 [0.60]	8.78 [0.28]	5.91 [0.25]	0.052
Young adult pigs	2.84 [0.39]	4.70 [0.18]	3.12 [0.30]	0.076

The thresholds *in situ* peak rarefactional pressure values.

Threshold estimates for lung hemorrhage in the age-dependent pig study (O'Brien et al. 2003) are summarized in Tables 3 to 5. The study included 2.1 ± 0.3 kg (4.9 ± 1.6-day-old) neonate pigs, 10 ± 1.1 kg (39 ± 5-day-old) juvenile pigs and 20 ± 1.2 kg (58 ± 5-day-old) young adult pigs.

Table 3 compares the age-dependent ED05, ED50 and piecewise linear thresholds for lesion incidence. For these data, the piecewise linear threshold for incidence is virtually the same as the ED05 in each age group, with estimates of excess risk ranging from 4% to 8%.

Table 4 compares age-dependent ED05, ED50 and piecewise linear thresholds for lesion depth. Similar to the results for the age-dependent mouse study (Dalecki et al. 1997b), here the ED05 and ED50 thresholds based on lesion depth are close to the values based on lesion incidence, indicating again the consistency of the logistic and tobit methods. The piecewise linear thresholds have estimated excess risks ranging from 26% to 33%. These values are considerably lower than the values in Table 2 for the mouse study, illustrating the data-dependence of the excess risk associated with the piecewise linear

Table 4. Analysis of lesion depth for O'Brien et al. (2003) data: ED05, ED50, piecewise linear threshold and its excess risk (computed from tobit regression).

Data source	ED05 [SE] (MPa)	ED50 [SE] (MPa)	Piecewise [SE] (MPa)	Piecewise excess risk
Neonate pigs	2.25 [0.61]	5.97 [0.33]	4.96 [0.62]	0.328
Juvenile pigs	5.88 [0.49]	8.69 [0.27]	7.61 [0.44]	0.264
Young adult pigs	2.60 [0.38]	4.75 [0.20]	3.93 [0.32]	0.264

The thresholds *in situ* peak rarefactional pressure values.

Table 5. Analysis of lesion surface area for O'Brien et al. (2003) data: ED05, ED50, piecewise linear threshold and its excess risk (computed from tobit regression).

Data source	ED05 [SE] (MPa)	ED50 [SE] (MPa)	Piecewise [SE] (MPa)	Piecewise excess risk
Neonate pigs	2.10 [0.58]	5.62 [0.34]	4.62 [0.55]	0.321
Juvenile pigs	5.46 [0.52]	8.39 [0.29]	6.33 [0.56]	0.124
Young adult pigs	2.58 [0.37]	4.66 [0.20]	3.87 [0.30]	0.268

The thresholds *in situ* peak rarefactional pressure values.

threshold. Table 5 provides a similar comparison for the thresholds derived from lesion square root surface area. Here, the tobit regression model has been fit to the square-root of surface area instead of the surface area itself to improve the fit of the model. In Table 5, the estimated excess risks associated with the piecewise linear thresholds are also between 5% and 50%.

Figure 1 demonstrates graphically how the logit model (logistic regression) gave estimates of ED05 and ED50 and how these results differed from the hockey stick results. For these data, the background rate $P(0) \approx 0$ and the excess risk for a given pressure value equals the probability of a lesion at that pressure value. Solving eqn (2) for ED50 (with $p = 0.50$) with $P(0) = 0$ was therefore numerically equivalent to finding the acoustic pressure where the logit-based probability curve crosses 0.50. This is shown in the graph to occur at 6.08 MPa. Solving eqn (2) for ED05 (with $p = 0.05$) amounts to finding the acoustic pressure (3.59 MPa) where the excess probability reaches 0.05. The threshold calculations were done numerically using eqn (4). The error bars are $\pm$ 1 SE.

Figure 2 demonstrates how the tobit model gave estimates of ED05 and ED50 using the extra information on lesion size. The conditional percentile curves, $Q_{50}(x)$ and $Q_{05}(x)$, are shown as functions of x = acoustic pressure in MPa. The points where these curves bend are, respectively, ED50 = 5.97 MPa and ED05 = 2.25 MPa. The ED50 value is quite close to the value estimated in Fig. 1 from the logistic regression. The ED05 value is somewhat lower than the logit-based estimate, although it appears to be within sampling error. The figure includes the measured lesion depths in mm with 0 mm shown for animals without lesions. The tobit-based background probability of a lesion is essentially zero and the excess risk, $r(x)$, is equal to the unadjusted risk, $P(x)$, in this example. The thresholds were computed numerically using eqn (8). The fitted line for the hockey stick model is also included for comparison. The point where it rises above zero is the "hockey stick" threshold.

DISCUSSION

The use of well-established probability-based threshold estimates provides the means for defining risk independently of the model being fit. Hence, the ED05 value can be estimated from logistic regression of event incidence, or from tobit regression of event magnitude, or by other means, but the definition of the threshold is the same regardless of the model and the event endpoint. What differs between models is the method of estimating the probabilities $P(x)$ and $P(0)$ in eqns (1) and (2). Provided that the different models are adequate for the relevant data (e.g., the logistic model is adequate for the percentage of lesions and the tobit model is adequate for size), similar threshold estimates are to be expected from them. This outcome relates to the close connection between tobit regression and probit analysis of binary outcomes, as well as the similarity between ED100p computed via probit regression and logistic regression (Simpson et al. 1996).

In contrast to the probability risk-based approach, the piecewise ("hockey stick") linear-threshold approach results in the threshold definition being model-dependent and data-type-dependent. An additional technical issue is that, unlike the tobit model for size, the hockey stick model takes no account of the lower bound of zero on the measurements. This bounding implies that the errors from the hockey stick model are asymmetric and have a distribution depending on the parameters, so the optimality assumptions of least-squares estimation are violated. The tobit model explicitly addresses the bounding issue by including the bounding below (at 0 mm) in the sampling model.

The results for incidence and size of lesions demonstrate that quite different excess risks are associated with the piecewise linear thresholds for percentage of lesions vs. size of lesions using the hockey stick model. Based on the results for the two data sets considered above, the piecewise linear thresholds for percentage of lesions appear to behave like the ED05, whereas those for lesion size appear to behave more like an ED25-ED50.

Dalecki et al. (1997a, 1997b) estimated threshold using the piecewise linear approach, obtaining values considerably lower than the ED05 thresholds reported by O'Brien et al. (2003). The results reported here provide comparisons between two different threshold-estimation methods for the same data and between similar threshold estimates applied to different experimental data. By comparing the same statistical methods across both sets of experimental data, these results demonstrate the importance of the differing experimental protocols for Dalecki et al. (1997a, 1997b) and O'Brien et al. (2003) as a source of nonsampling variation. One might also conjecture a difference associated with small animals vs. large

animals, but Dalecki et al. (1997a, 1997b) reported similar thresholds for mice and pigs, using the same analysis for both.

In the comparison between logit-based thresholds derived from incidence data and tobit-based thresholds derived from size data, it can be expected that the tobit-based estimates are more precise, if the censored Gaussian model holds, but the logit-based estimates are more robust. In Tables 1 and 2, for example, the tobit thresholds yield smaller standard errors than the logit thresholds. In Tables 4 and 5, however, the precision of the tobit estimation of ED05 for neonate pigs is reduced by an unusually large lesion at an exposure below 4 MPa, as can be seen in Fig. 2. This outlier appears to have driven down the tobit-based ED05 estimate compared with the logit-based estimate. It should be noted, however, that the 95% confidence intervals ($\pm$ 2 SE) for the logit and tobit-based estimates are overlapping.

In general, the use of probability-based excess risks and thresholds is recommended to provide consistent risk targets across experiments and data types, and the logit and tobit models are recommended to properly model 0 to 1 and zero-bounded response data.

Acknowledgements—This work was supported by an NSF grant (DMS-0073044) awarded to D.G. Simpson and by NIH (grant EB02641, formerly HL58218) awarded to W.D. O'Brien, Jr. and J.F. Zachary.

REFERENCES

Agresti A. Categorical Data Analysis. John Wiley & Sons: New York.

Amemiya T. Tobit models: A survey. J Econometrics 1984;24:3–61.

Baggs R, Penney DP, Cox C, et al. Thresholds for ultrasonically induced lung hemorrhage in neonatal swine. Ultrasound Med Biol 1996;22:119–128.

Chen JJ, Kodell RL. Quantitative risk assessment for teratological effects. J Am Statist Assoc 1989;84:966–971.

Cox C. Generalized linear models—The missing link. Appl Statist 1984;33:18–24.

Dalecki D, Child SZ, Raeman CH, Cox C, Carstensen EL. Ultrasonically induced lung hemorrhage in young swine. Ultrasound Med Biol 1997a;23:777–781.

Dalecki D, Child SZ, Raeman CH, et al. Age dependence of ultrasonically induced lung hemorrhage in mice. Ultrasound Med Biol 1997b;23:767–776.

Galfalvy H, Simpson DG. Infrastructure degradation: An application of censored regression models. American Statistical Association, Alexandria, VA ASA Proceedings of the Section on Physical and Engineering Sciences. 1999:242–247.

Ihaka R, Gentleman R. R: A language for data analysis and graphics. J Comput Graph Stat 1996;5:299–314.

O'Brien WD Jr.. Zachary JF. Lung damage assessment from exposure to pulsed-wave ultrasound in the rabbit, mouse, and pig. IEEE Trans Ultrason Ferroelec Freq Control 1997;44:473–485.

O'Brien WD Jr., Frizzell LA, Schaeffer DJ, Zachary JF. Superthreshold behavior of ultrasound-induced lung hemorrhage in adult mice and rats: Role of pulse repetition frequency and pulse duration. Ultrasound Med Biol 2001a;27:267–277.

O'Brien WD Jr., Simpson DG, Frizzell LA, Zachary JF. Superthreshold behavior and threshold estimation of ultrasound-induced lung hemorrhage in adult rats: Role of beamwidth. IEEE Trans Ultrason Ferroelec Freq Control 2001b;48:1695–1705.

O'Brien WD Jr., Simpson DG, Ho M, et al. Superthreshold behavior and threshold estimation of ultrasound-induced lung hemorrhage in pigs: Role of age dependency. IEEE Trans Ultrason Ferroelec Freq Control 2003;50:153–169.

Raeman CH, Child SZ, Dalecki D, Cox C, Carstensen EL. Exposure-time dependence of the threshold for ultrasonically induced murine lung hemorrhage. Ultrasound Med Biol 1996;22:139–141.

Ryan L. Quantitative risk assessment for developmental toxicity. Biometrics 1992;48:163–174.

Simpson DG, Carroll RJ, Zhou H, Guth DJ. Interval censoring and marginal analysis in ordinal regression. J Agricult Biol Environ Stat 1996;1:354–376.

Xie M, Simpson DG. Regression modeling of ordinal data with non-zero baselines. Biometrics 1999;55:308–316.

Zachary JF, O'Brien WD Jr. Lung lesion induced by continuous- and pulsed-wave (diagnostic) ultrasound in mice, rabbits, and pigs. Vet Pathol 1995;32:43–54.

Zachary JF, Sempsrott JM, Frizzell LA, Simpson DG, O'Brien WD Jr. Superthreshold behavior and threshold estimation of ultrasound-induced lung hemorrhage in adult mice and rats. IEEE Trans Ultrason Ferroelec Freq Control 2001;48:581–592.