Some New Aspects of Dose-Response Models with Applications to Multistage Models Having Parameters on the Boundary

Bimal Sinha
Department of Mathematics & Statistics
University of Maryland, Baltimore County, USA

Leonid Kopylev & John Fox
National Center for Environmental Assessment
US Environmental Protection Agency

Outline of this presentation

Introduction to NCEA risk assessment quantitative methods and why boundaries of parameter space matter

Discuss history and published results, in particular Self and Liang (1987) in JASA

More detailed outline

NCEA Risk Assessment Methodology

- Quite often, risk assessments for various toxic substances can only use data from animal bioassays. No human data on that toxic substance is available or the human data lacks exposure measurement suitable for quantifying risks.
- Typical bioassay data comes from a 2-year rodent experiment consisting of several dose groups and a control group (usually 3-6 groups total). Normally, there are 10-50 animals in each group.
- In most cases, only summary data from the bioassay is available, not data on individual animals.
- A number of dose-response models are used, many with natural constraints on parameters due to biological understanding of dose-response.
- The goal is to estimate risk at very low doses that are applicable to human risks, accounting for and, as much as possible, quantifying associated uncertainty.

Models Used in Risk Assessments for dichotomous data from bioassays

Commonly used model for quantal responses:

Multistage Cancer Model

$$P(d, \boldsymbol{\theta}) = 1 - exp(-\sum_{j=0}^{k} \theta_j d^j), \qquad \boldsymbol{\theta} \geq 0$$

Loglogistic Model

$$P(d, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \gamma + \frac{1-\gamma}{1 + e^{-(\alpha+\beta * ln(d))}} \quad \boldsymbol{\beta} \geq 1, \boldsymbol{\gamma} \in [0;1]$$

Logprobit Model

$$P(d, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \gamma + (1-\gamma) * \Phi(\alpha + \beta * ln(d)) \quad \boldsymbol{\beta} \geq 1, \ \boldsymbol{\gamma} \in [0;1]$$

Weibull Model

$$P(d, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \gamma + (1-\gamma) * (1 - \beta * e^{-\beta * d^{\alpha}}) \quad \boldsymbol{\alpha} \geq 1, \ \boldsymbol{\gamma} \in [0;1]$$

Benchmark Dose method

The benchmark dose (BMD) method (Crump, 1984) consists of estimating a lower confidence limit (BMDL) for the dose associated with a specified increase γ in adverse response (i.e., increased risk) above the background level.

$$Extra\ Risk: \qquad ER = \frac{P(d,\boldsymbol{\theta}) - P(0,\boldsymbol{\theta})}{1 - P(0,\boldsymbol{\theta})}$$

$$Benchmark\ Dose\ (BMD): \quad d*: \gamma = \frac{P(d*,\boldsymbol{\theta}) - P(0,\boldsymbol{\theta})}{1 - P(0,\boldsymbol{\theta})}$$

Profile likelihood is commonly used for estimating BMDL; Wald method had been shown in the literature not to work well, in particular when parameters are on the boundary

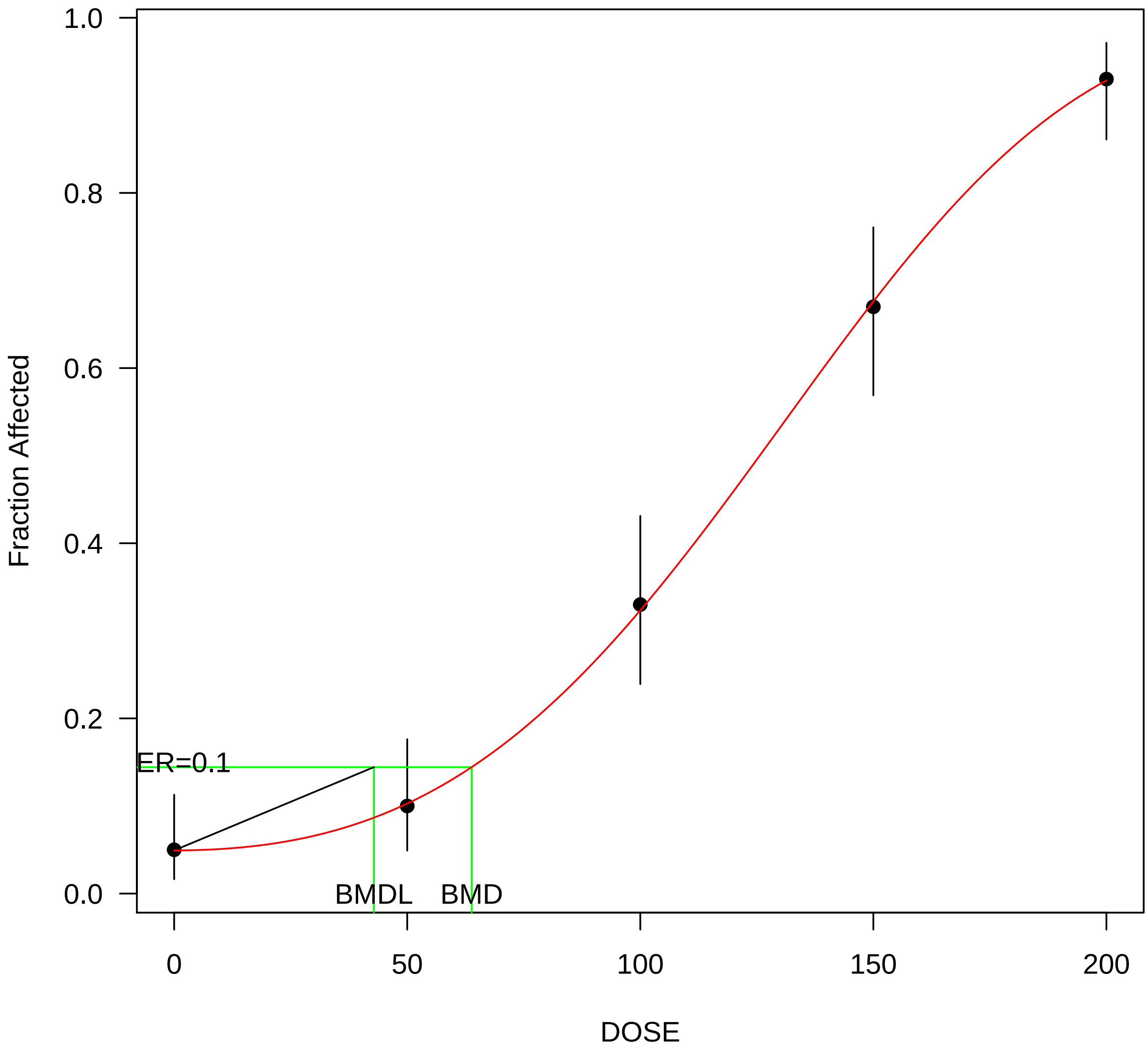

Figure 1: Benchmark Dose (BMD) Methodology

The Literature

1. Chernoff, H. (1954). ‘On the distribution of the likelihood ratio’. **Annals of Math Stat**, **25**, 573-578.
2. Feder, P.I. (1968). ‘On the distribution of the log likelihood ratio test statistic when the true parameter is *near* the boundaries of the hypothesis regions’. **Annals of Math Stat**, **39**, 2044-2055.
3. Self, S.G. and K.-Y. Liang. (1987). ‘Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions’. **JASA**, **82**, 605-610.
4. Sen, P.K. and M.J. Silvapulle. (2003).‘An appraisal of some aspects of statistical inference under inequality constraints’. **J. of Stat Planning and Inference**, **107**, 3-43.
5. Sinha B., Kopylev L., Fox J. (2008) ‘Some new aspects of dose-response multistage models with applications’. *Environmental and Ecological Statistics, Platinum Jubilee Conference of ISI,* **World Scientific, Singapore** (in press) Earlier version `http://www.math.umbc.edu/~kogan/technical_papers/2007/Sinha_Kopylev_Fox.pdf`

Self and Liang (1987)

- Described asymptotic distribution of MLE when some of the parameters have true values on the boundary

  NB Some of the components of asymptotic distribution are not necessarily Gaussian

- For LRT established that the asymptotic distribution depends solely on the four-tuple $(Ib,\ I\bar{b},\ Nb,\ N\bar{b})$, where

  Ib parameters of interest: true values on the b-dary

  $I\bar{b}$ parameters of interest: true values not on the b-dary

  Nb nuisance parameters: true values on the b-dary

  $N\bar{b}$ nuisance parameters: true values not on the b-dary

Self and Liang (1987)

For example:

(0, s, 0, p-s) χ^2_s

(1, 0, 0, p-1) 50:50 mixture of χ^2_0 & χ^2_1

(1, 1, 0, p-2) 50:50 mixture of χ^2_1 & χ^2_2

(1, 0, 1, 0) mixture of several χ^2 and not χ^2 distributions

True value of a single nuisance parameter being on the boundary or not may tremendously complicate asymptotic distribution.

Outline of the presentation

Derive asymptotic distribution of LRT in case of (0, 1, m, p-m-1) - parameter of interest is not on the boundary, but m nuisance parameters are on the boundary.

Apply derived distribution of LRT and results on asymptotic distribution of MLE (Self and Liang, 1987) to multistage cancer model

Example and Simulations Results

Concluding Remarks

Other Statistical Problems in NCEA

New result on Boundary value problems

Proposition. Consider the general problem of testing the null hypothesis for a boundary point θ_{10}

$$H_0 : \theta_1 = \theta_{10} \quad versus \quad H_1 : \theta_1 > \theta_{10},$$

and suppose we compute the LRT by maximizing $L(\boldsymbol{\theta}|\mathbf{X})$ wrt $\boldsymbol{\theta} \in \Omega$ (Ω is quite general) and also under H_0. Write

$$\Omega_0 = \{\boldsymbol{\theta} : \theta_1 = \theta_{10}, \theta_2, \cdots, \theta_p \; unspecified.\}$$

Then the null distribution of the profile log likelihood based LRT using $\mathbf{X}$, namely, the null distribution of

$$-2ln \left[\frac{max_{\boldsymbol{\theta} \in \Omega_0} L(\boldsymbol{\theta}|\mathbf{X})}{max_{\boldsymbol{\theta} \in \Omega} L(\boldsymbol{\theta}|\mathbf{X})} \right] \tag{1}$$

is asymptotically equivalent to the distribution of the profile log likelihood based LRT using $\mathbf{Z}$ under $N[\boldsymbol{\theta}, \Sigma]$, namely,

New result on Boundary value problems

the distribution of

$$min_{\boldsymbol{\theta}\in\Omega_0^*}Q(\boldsymbol{\theta}|\mathbf{Z}) - min_{\boldsymbol{\theta}\in\Omega^*}Q(\boldsymbol{\theta}|\mathbf{Z}) \tag{2}$$

when $\boldsymbol{\theta} = \mathbf{0}$. Here

$$Q(\boldsymbol{\theta}|\mathbf{Z}) = (\mathbf{Z} - \boldsymbol{\theta})'[\Sigma]^{-1}(\mathbf{Z} - \boldsymbol{\theta}) \tag{3}$$

is the exponent of the normal likelihood of $\mathbf{Z}$ with mean $\boldsymbol{\theta}$ and dispersion Σ. Minimum under $\boldsymbol{\theta} \in \Omega_0^*$ is computed when $\theta_1 = 0$ and $\theta_2, \cdots, \theta_p$ are unspecified in $(-\infty, \infty)$, and minimum under $\boldsymbol{\theta} \in \Omega^*$ is computed when all the parameters $\boldsymbol{\theta}$ are unspecified in $(-\infty, \infty)$. In other words, quite generally, Ω^* is a translation of Ω by Ω_0 so that Ω^* includes the point $\mathbf{0}$ in the parameter space of $\mathbf{Z}$.

New result on Boundary value problems

Remark. The null hypothesis above does not mention anything about nuisance parameters. It is quite possible some of the nuisance parameters may lie in a parameter spaces containing boundary points. When this happens, it is implied that the minimization wrt $\boldsymbol{\theta}$ in the quadratic form $Q(\boldsymbol{\theta}|\mathbf{Z})$ is done under restriction that corresponding nuisance parameter $\theta_i \geq 0$.

Remark. This reduction of the original inference problem based on $\mathbf{X}$ with an arbitrary distribution to a canonical form using $\mathbf{Z}$ which has a normal distribution, though only asymptotically valid, is a key feature of the asymptotic theory and the spirit of all the earlier works of all the authors.

It turns out that, as one can expect, the form of the LRT becomes quite complex with the increase in the number of nuisance parameters which lie on the boundary, so we'll show here only case of 1 boundary point

One boundary point

θ_1 is an interior point and θ_2 is a boundary point. To derive the LRT for $H_0 : \theta_1 = \delta + \theta_{10}$ versus $H_1 : \theta_1 \neq \delta + \theta_{10}$ for some $\delta > 0$ on the basis of $\mathbf{X}$, note that, asymptotically, this is equivalent to testing $H_0 : \theta_1 = \delta$ versus $H_1 : \theta_1 \neq \delta$ for $\delta > 0$ based on $\mathbf{Z}$.It is clear that what matters is the quadratic form $Q(\boldsymbol{\theta}|\mathbf{Z})$ given by

$$Q(\boldsymbol{\theta}|\mathbf{Z}) = (\mathbf{Z} - \boldsymbol{\theta})'\Sigma^{-1}(\mathbf{Z} - \boldsymbol{\theta}). \tag{4}$$

Writing

$$Q(\boldsymbol{\theta}|\mathbf{Z}) = Q(\theta_1, \theta_2|Z_1, Z_2) + Q(Z_3, \cdots, Z_p|Z_1, Z_2; \boldsymbol{\theta})$$

where the first part is the marginal bivariate quadratic of (Z_1, Z_2) and the second part is the $(p-2)$-dimensional conditional quadratic of $(Z_3, \cdots, Z_p)$, given (Z_1, Z_2), it follows from Self and Liang (1987) that due to the interior nature of the parameters $\theta_3, \cdots, \theta_p$, the only part we need to study is the first part, and maximization of the likelihood corresponds to finding the two minimums of the first part, one under the union of null and alternative hypotheses, and the other under the null hypotheses

One boundary point

After minimization in 4 separate region and combining, we get $-2ln(LRT) = W$ ($Z_{2.1}$ are usual residuals) as

$$
\begin{aligned}
W &= (Z_1 - \delta)^2/\sigma_{11}, \quad if Z_2 > 0, Z_{2.1} > 0 \\
&= (Z_1 - \delta)^2/\sigma_{11} - Z_2^2/\sigma_{22}, \quad if Z_2 < 0, Z_{2.1} > 0 \\
&= Z_{1.2}^2/\sigma_{11}(1-\rho^2), \quad if Z_2 < 0, Z_{2.1} < 0 \\
&= \frac{(Z_1 - \delta)^2}{\sigma_{11}} + \frac{Z_{2.1}^2}{\sigma_{22}(1-\rho^2)}, \quad if Z_2 > 0, Z_{2.1} < 0.
\end{aligned}
$$

It is easy to verify that when $\rho = 0$, W reduces to $(Z_1 - \delta)^2/\sigma_{11}$ which has a chi-square distribution with 1 d.f. under H_0, a familiar result.

One boundary point

Theorem. The cdf $G(w)$ of W, for $0 < w < \infty$, is given by the sum of four parts:

$$First\ part = \int_0^{\rho\sqrt{w}} \left[\int_{-\sqrt{\{(w-v^2)(1-\rho^2)\}}}^{v\rho+\sqrt{w}} N(0,1)dx \right] N(0,1)dv$$

$$Second\ part = \int_{\rho\sqrt{w}}^{\infty} \left[\int_{v\rho-\sqrt{w}}^{v\rho+\sqrt{w}} N(0,1)dx \right] N(0,1)dv$$

$$Third\ part = \int_{-\rho\sqrt{w/(1-\rho^2)}}^{0} \left[\int_{-\sqrt{w(1-\rho^2)}}^{v\rho+\sqrt{w+v^2}} N(0,1)dx \right] N(0,1)dv$$

$$Fourth\ part = \left[\int_{-\sqrt{w(1-\rho^2)}}^{\sqrt{w(1-\rho^2)}} N(0,1)dx \right] \cdot \left[\int_{-\infty}^{-\rho\sqrt{w/(1-\rho^2)}} N(0,1)dx \right]$$

The above distribution can be used to get a cut-off point of the statistic $-2ln(LRT)$ which can then be used to carry out the LRT based on $\mathbf{X}$. *Quite surprisingly*, our simulations indicate that the above distribution does *not* depend on ρ!

Linear Multistage Model

$\pi_i(\boldsymbol{\theta}|d_i) = 1 - e^{-[\theta_0 + d_i\theta_1]}$, which readily gives

$$AR(d*) = 1 - e^{-[\theta_0 + d*\theta_1]}$$
$$ER(d*) = 1 - e^{-d*\theta_1}$$
$$BMD = (1/\theta_1)ln(1/(1-\gamma)).$$

Fisher information matrix $I(\theta_0, \theta_1)$: under the assumption of independent binomial distributions of the X_i's:

$$I(\theta_0, \theta_1) = \begin{pmatrix} \sum_{i=0}^{m} \frac{1-\pi_i(\boldsymbol{\theta})}{\pi_i(\boldsymbol{\theta})} & \sum_{i=0}^{m} d_i \frac{1-\pi_i(\boldsymbol{\theta})}{\pi_i(\boldsymbol{\theta})} \\ \sum_{i=0}^{m} d_i \frac{1-\pi_i(\boldsymbol{\theta})}{\pi_i(\boldsymbol{\theta})} & \sum_{i=0}^{m} d_i^2 \frac{1-\pi_i(\boldsymbol{\theta})}{\pi_i(\boldsymbol{\theta})} \end{pmatrix}.$$

When $\theta_0 > 0, \ \theta_1 > 0$, we clearly have a *regular* parametric scenario. By applying the standard asymptotic theory:

$$\sqrt{N}(\hat{\theta}_{0N} - \theta_0, \hat{\theta}_{1N} - \theta_1) \rightarrow N_2[(0,0), \Sigma(\theta_0, \theta_1)]$$

$$\sqrt{N}[(\hat{\theta}_{0N} + d*\hat{\theta}_{1N}) - (\theta_0 + d*\theta_1)] \rightarrow N[0, (1, d*)\Sigma(\theta_0, \theta_1)(1, d*)'].$$

Inference about AR, ER and BMD is straightforward.

Linear Multistage Model

When $\theta_0 = 0$, have the case of one parameter point being on the boundary, and hence the regular asymptotic distributional and inferential results mentioned above are *not* true. Using Self and Liang (1987), we get

$$\sqrt{N}(\hat{\theta}_{0N}, \hat{\theta}_{1N} - \theta_1) \rightarrow (\hat{\theta}_0(Z_0, Z_1), \hat{\theta}_1(Z_0, Z_1))$$

where $\mathbf{Z} = (Z_0, Z_1) \sim N[\mathbf{0}, \hat{\Sigma}]$ and

$$\begin{aligned} \hat{\theta}_0(Z_0, Z_1) &= 0 \quad if \quad Z_0 \leq 0 \\ &= Z_0 \quad if \quad Z_0 > 0 \end{aligned}$$

and

$$\begin{aligned} \hat{\theta}_1(Z_0, Z_1) &= Z_{1.0} \quad if \quad Z_0 \leq 0 \\ &= Z_1 \quad if \quad Z_0 > 0. \end{aligned}$$

$Z_{1.0} = Z_1 - \rho\sigma_1 Z_0/\sigma_0$: standard residual of Z_1 on Z_0.

Linear Multistage Model

Returning to the inference problems, although we still have the same point estimates as before, namely, $\hat{AR}(d*) = 1 - e^{-[\hat{\theta}_{0N}+d*\hat{\theta}_{1N}]}$, $\hat{ER}(d*) = 1 - e^{-d*\hat{\theta}_{1N}}$ and $B\hat{M}D = (1/\hat{\theta}_{1N})ln(1/(1-\gamma))$, the asymptotic distributions of these estimates are not normal any more, and in fact are obtained by replacing $\hat{\theta}_{0N}$ and $\hat{\theta}_{1N}$ by $\hat{\theta}_0(Z_1, Z_2)$ and $\hat{\theta}_1(Z_1, Z_2)$, respectively, and deriving the resultant distributions by simulation.

What we have described so far is the Wald approach. For the derivation of the LRT for hypotheses about $ER(d*)$ and BMD, since these quantities involve only θ_1, it follows from the general theory discussed above that irrespective of the nature of θ_0, whether a boundary point or not, the asymptotic distribution of $-2ln(LRT)$ remains as χ^2 with 1 d.f. We now will compare the Wald test with the LRT test in a few cases.

Monte-Carlo Simulations

Imitating standard animal bioassay: control group and 3 dose groups of 50 animals; desired coverage 90%

"Self-Liang" column shows coverage of the procedure such that for each simulation, depending on the estimated value of the corresponding parameter being on the boundary or not, interval computed according to the asymptotic distribution derived in Self and Liang or Wald interval was used.

Results were similar to described below for Extra Risk (ER) and non-background parameters of multistage model).

Monte-Carlo Simulations

Cubic multistage model. Several scenarios: some of the parameters or their estimates are on the boundary.

Coverage and Length of Confidence Intervals

Parameter True Values	Wald		Self-Liang		LRT	
	CI %	Length	CI %	Length	CI %	Length
$\theta_0 = 0.1; \theta_1 = 0$ $\theta_2 = 0; \theta_3 = 2$	98.8	0.569	92.7	0.240	90.3	0.184
$\theta_0 = 0; \theta_1 = 1$ $\theta_2 = 0; \theta_3 = 3$	94.9	0.920	89.6	0.122	89.9	0.120
$\theta_0 = 0; \theta_1 = 2$ $\theta_2 = 0; \theta_3 = 0.5$	100	0.730	86.6	0.055	91.9	0.074
$\theta_0 = 0.2; \theta_1 = 0.01$ $\theta_2 = 0.01; \theta_3 = 1$	99.9	1.8368	93.9	0.387	92.1	0.303

Example

National Toxicological Program (NTP, 1993): B6C3F1 mice exposed to 1,3-Butadiene. The outcome was heart hemangiosarcomas. This data was analyzed by Bailer and Smith (1994, Risk Analysis).

Dose (ppm)	Number at risk	Number of tumors
0	50	0
6.25	50	0
20	50	0
62.5	49	1
200	50	21

A 3-stage multistage model fits well to this data and two parameters θ_0 and θ_1 are estimated to be on the boundary.

Example

Upper bound of both Wald and Self-Liang confidence intervals agree well to each other and to both nonparametric and parametric bootstrap confidence bounds (Bailer and Smith, not shown). However, for all 3 doses, the Wald 90% confidence interval contains 0, but confidence interval calculated according to Self and Liang (1987) does not. Also, Self-Liang interval is shorter than Wald's.

Wald and Self-Liang Confidence Intervals for Extra Risk at three doses

Dose (ppm)	Wald Confidence Interval	Self-Liang Confidence Interval
2.00	(0; 4.476E-5)	(2.143E-6; 4.471E-5)
0.20	(0; 5.027E-7)	(1.804E-8; 4.982E-7)
0.02	(0; 1.060E-8)	(1.770E-10; 1.063E-8)

Multistage models, 4 dose groups

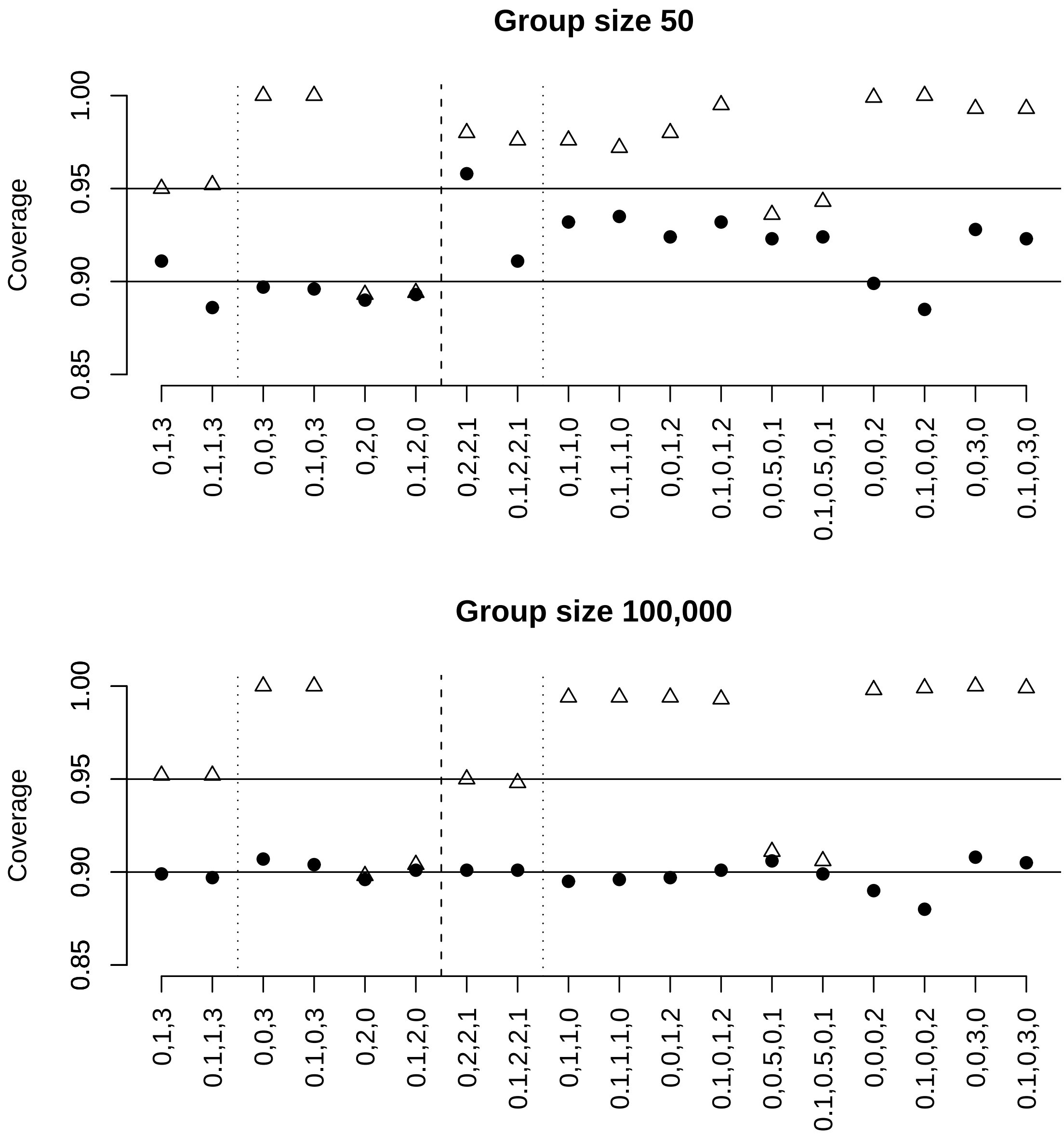

Figure 2: One and two-sided coverage for the multistage model

Multistage models, 6 dose groups

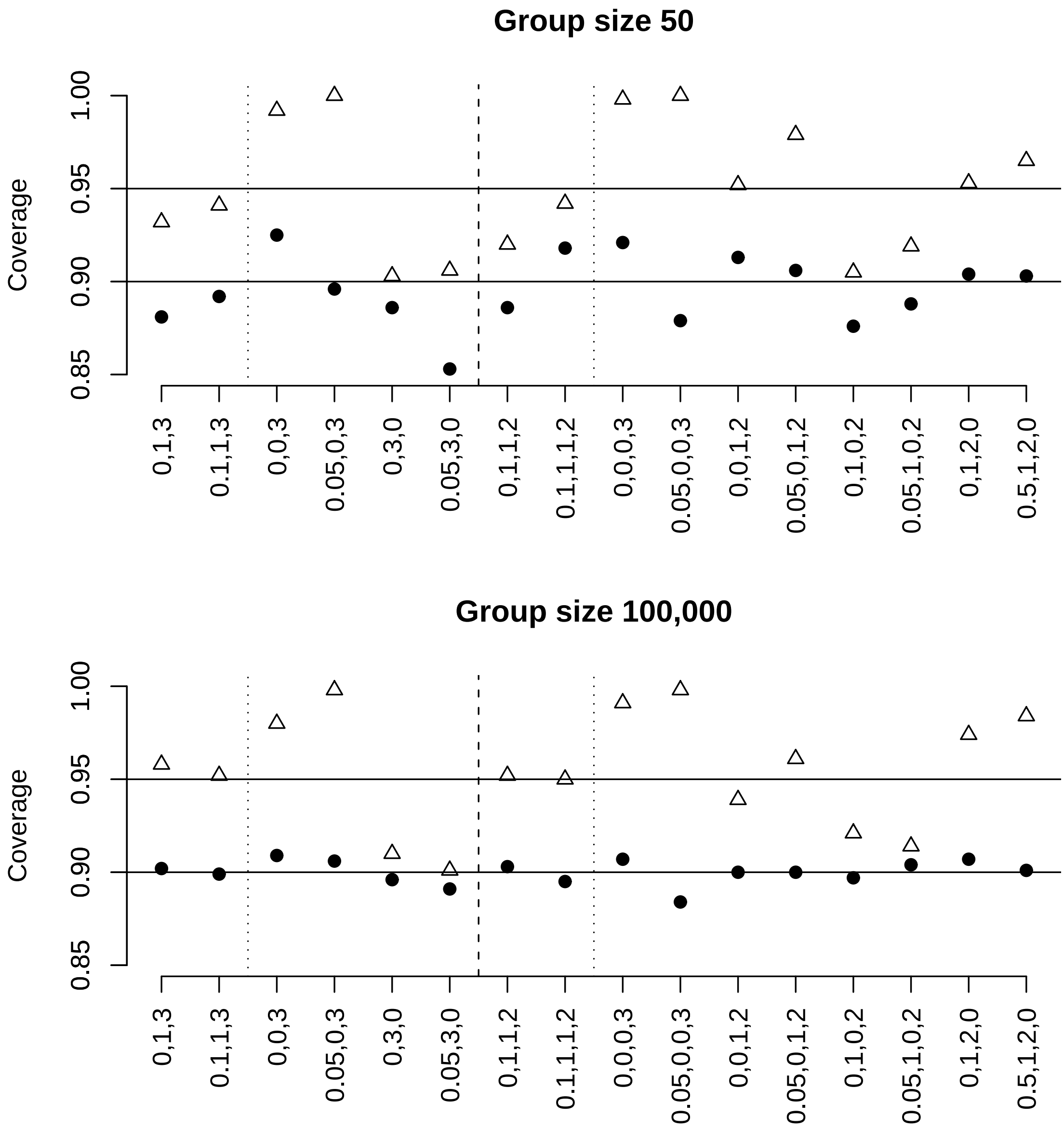

Figure 3: One and two-sided coverage for the multistage model

Log-logistic model, 6 dose groups

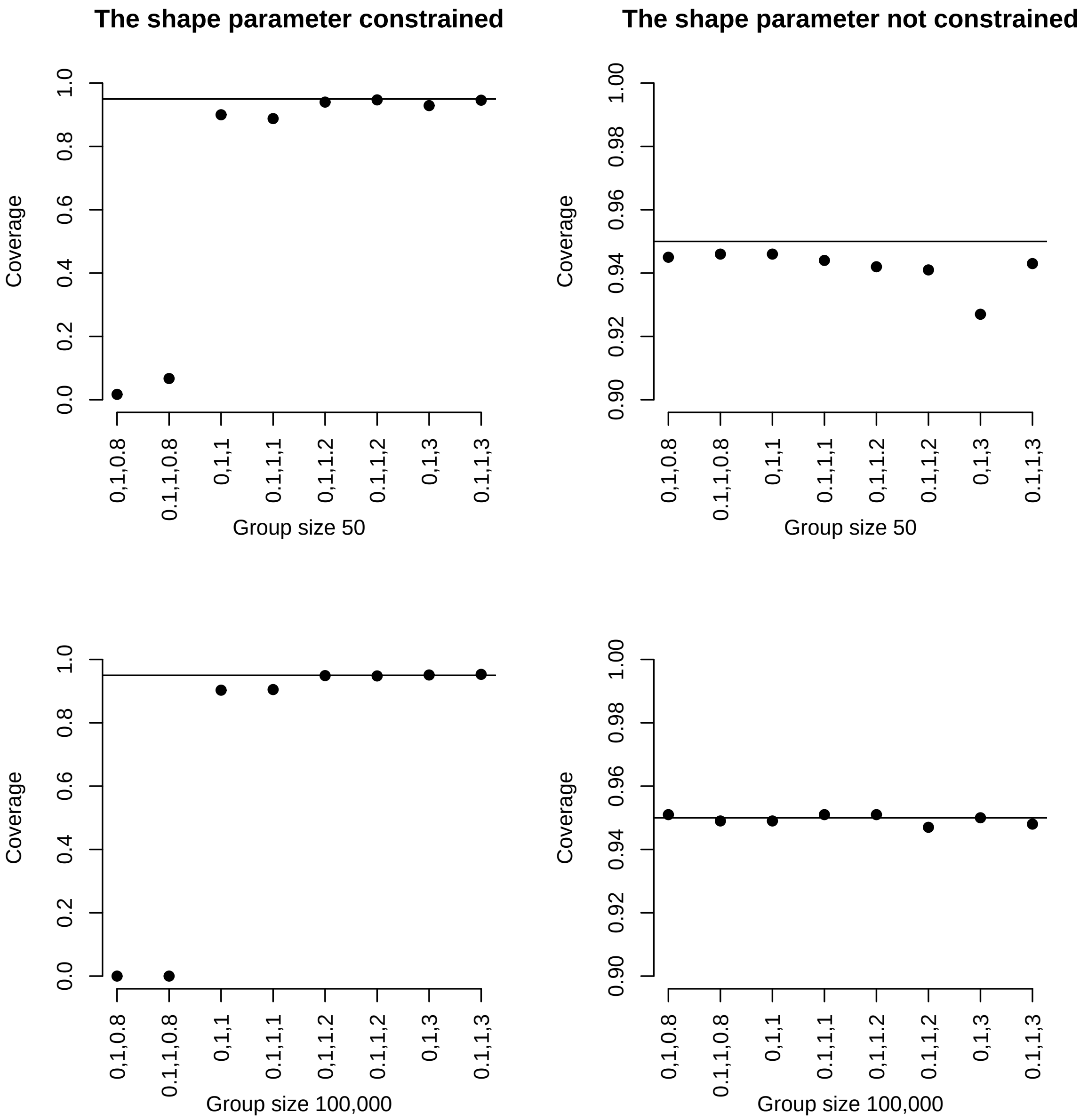

Figure 4: One-sided coverage for the log-logistic model

Concluding Remarks

- Long-standing questions about the accuracy and validity of various asymptotic CI methods for risk estimates from dose response models.
- The Self-Liang method improves upon Wald's for all the parameters and performs comparably with the LRT intervals, the latter being slightly shorter.
- Programming the profile likelihood method requires complicated routines for nonlinear optimization with inequality constraints. In contrast, Self and Liang's improvement on the Wald method can be programmed very easily, as it involves only inversion of the information matrix and simulation of a multivariate normal distribution.
- We investigated one-sided coverage of PL only. A next step would be to examine one-sided intervals by other approaches (Self-Liang, bootstrap, MCMC) when some parameters are on the boundary. MCMC methods (Gelfand, Smith, Lee JASA 1992) seems an attractive alternative,as they naturally incorporate constraints.