Discussion of: Uncertainty Quantification for Dose-Response Models Using Probabilistic Inversion with Isotonic Regression: Bench Test Results

Thomas A. Louis PhD, Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
www.biostat.jhsph.edu/~tlouis/
tlouis@jhsph.edu

Points of Agreement

- One needs to compute and report all relevant uncertainties
  - Both *a priori* and *a posteriori*
  - Both "sampling" and "non-sampling"
  - In general, the Bayesian formalism does this best
- One needs to use models with sufficient flexibility to honor the data, but with sufficient constraints to honor biology (or physics, or sociology, or . . .)
  - If you really (really) believe in monotonicity, enforce it
- One needs to be vigourous in model criticism

Points of Disagreement

- Roger's SAU is not my SAU and is not SAU at least not for WEES (well-educated and experienced statisticians)
  - WEESs would use isotonic regression with an appropriate Bayesian or frequentist uncertainty analysis
- Roger's Bayes is not my Bayes; it is one example
- I have issues with some "tent poles"
- Probabilistic Inversion may have a role in approximating a fully Bayesian approach in problems that need it, but should serve only that role
  - need to discretize
  - how to handle $\mathbf{0}$s?

Tent Poles

1. **Observational Uncertainty:** There must be some antecedent uncertainty over observable phenomena which we try to capture via distributions on parameters of DR models (I agree)
2. **Integrated Uncertainty Analysis:** the uncertainty in DR must be captured via a joint distribution over parameters which does not depend on dose (partial agreement)
   - No model can be completely global
3. **Monotonicity:** Barring toxicological insights to the contrary, we assume that the probability of response cannot decrease as dose increases (partial agreement)
   - There may well be toxicological insights to the contrary, such as doses beyond the MTD
   - Do you really want absolutely strict enforcement of monotonicity (see below)

Model-based Uncertainty and Model Uncertainty

- Model based uncertainty assessment is "easy"
- But when there are little or no data to inform about models (when the likelihood is not informative about model choice) prior uncertainties migrate directly to the posterior distribution
- A priori uncertainty is not changed and so one can inflate uncertainty with little or no empirical control
- This uncertainty propagation operates in "what if" analyses or in Bayesian Model Averaging

A low-dose extrapolation example

- Linearity or other shape depends on the dose metameter
- Shape also depends on the fundamental relation between exposure and dose
- Many models fit the observed data well, but give radically different estimated safe doses
- Sometimes relatively *ad hoc* approaches are used

Model Uncertainty

Liver Tumors in Rats consuming 2AAF

Low dose model	Safe Dose (10^{-8} elevation)
Linear	10^{-5}
Multi-stage*	
Weibull*	10^{-2}
Gamma*	
Logit	
Probit	10^{-1}

* = model fits observed data

Enforcing Monotonicity: Binomial Example

- The standard binomial estimates aren't appropriate, if you really believe in monotonicity or put a lot of weight on it
- Use isotonic regression or a Bayesian version

$$
\begin{aligned}
\boldsymbol{P_d} &= \boldsymbol{pr}(\text{ response } \mid \text{ dose } = \boldsymbol{d}) \\
\mathbf{P} &= (\boldsymbol{P_0}, \ldots, \boldsymbol{P_K}) \sim \boldsymbol{G} \\
\text{monotone:} \quad & \boldsymbol{G}(\boldsymbol{P_0} \leq \boldsymbol{P_1} \leq, \ldots, \boldsymbol{P_K}) = 1
\end{aligned}
$$

- Dirichlet process on increments, Gamma process, ...

The Binomial Example: $K = 1$
Monotone

Let

$$P_0 = \frac{e^{\mu}}{1 + e^{\mu}}$$

$$P_1 = \frac{e^{\mu+\theta}}{1 + e^{\mu+\theta}}$$

require,

$$p_G(\theta \geq 0) = 1$$

But, what if you see convincing non-monotonicity?

- What would you do if you had very large sample sizes and still saw non-monotonicity?
- How are you going to make the decision to switch gears?
- Wouldn't it be better to build in *a priori* a (small) possibility of non-monotonicity?

For $K = 1$, let

$$\theta = = \begin{cases} -|\eta_1| & w.p.\ \epsilon \\ \quad |\eta_2| & w.p.\ (1-\epsilon) \end{cases}$$

where,

$$\eta_j = N(\mu_j, \sigma_j^2)$$

Probabilistic Inversion

- Useful in many contexts
  - Poole D, Raftery AE (2000). Inference for Deterministic Simulation Models: The Bayesian Melding Approach. JASA, 95: 12441255
  - Yin Y, Louis TA (2007). Optimal Constrained Bayesian Updating to Match a Target Output Distribution and Retain Stochastic Model Structure (submitted).
- But, difficult to get right for Roger's goals, especially in complex situations

Consider a simple situation

$$\begin{aligned} \boldsymbol{\theta} &\sim \boldsymbol{G} \text{ (unknown)} \\ \boldsymbol{Y} \mid \boldsymbol{\theta} &\sim \boldsymbol{N}(\boldsymbol{\theta}, \boldsymbol{\sigma^2}),\ \boldsymbol{\sigma^2} \text{ (known)} \\ \hat{\boldsymbol{\theta}} &= \boldsymbol{Y} \end{aligned}$$

- "Simulating" from $\hat{\boldsymbol{\theta}}$ induces a $\boldsymbol{G}$ with variance $= \mathbf{0}$
- Need a two-stage simulation
  1. Sample $\boldsymbol{\theta}$ from $\boldsymbol{N}(\boldsymbol{Y}, \boldsymbol{\sigma^2})$
  2. Generate a $\boldsymbol{Y}^*$
  3. Repeat 1 and 2

Try for Bayes

- Definitely use Bayes to identify the goal(s) even if you then have to approximate
- But, beware of statistical mischief

The Perils of Bayesian Model Averaging

An air pollution and health example

- Koop, and Tole 2004 Measuring the health effects of air pollution: to what extent can we really say that people are dying from bad air? J Environ Econ Mgmt, 47: 30-54

  show that using Bayesian Model Averaging and a broad range of candidate models, there is a great deal of uncertainty about the magnitude of the air pollution effect

- Thomas, Jerrett, Kuenzli, Louis, Dominici, Zeger, Schwartz, Burnett, Krewski, Bates (2007) Bayesian Model Averaging in Time Series Studies of Air Pollution and Mortality. Journal of Toxicology & Environmental Health, Part A, 70: 1-5

  rebut K&T by showing that the data provide very little information on model choice and that you need to be careful in selecting candidate models and priors

  - Ditto for any frequentist approach to model selection

- Statistical "wrappers" can be used to package mischief

There are limits to what we can do

- For example, model choice will always have unquantifiable aspects
- We can bring in expert opinion/ knowledge on mechanism, genomics, discount rates, ...
- But, how do we rate the "credibility" of an expert?
- There are limits to all formalisms

Summary

- Roger provokes us to do better
- Many of his themes are right on the mark
- I disagree with others and with proposed improvements
- I do agree that we need to take BMD "to the next level"