Towards Structured Use of Bayesian Sequential Monitoring in Clinical Trials

Evan Kwiatkowski[†], Eugenio Andraca-Carrera[‡],
Mat Soukup[‡], Matthew A. Psioda[†*]

† Department of Biostatistics, University of North Carolina,
McGavran-Greenberg Hall, CB#7420,
Chapel Hill, North Carolina, U.S.A.

‡ Division of Biometrics VII, Office of Biostatistics
Center for Drug Evaluation and Research,
US Food and Drug Administration,
Silver Spring, Maryland, USA

June 20, 2019

Abstract

The text of your abstract. 200 or fewer words.

Keywords: 3 to 6 keywords, that do not appear in the title

[*]The authors gratefully acknowledge *please remember to list all relevant funding sources in the unblinded version*

1 Introduction

Things to discuss:

- 21st Century Cures Act (MATT)
- PDUFA VI reauthorization (MATT)
- Expansive work already done on sequential monitoring (EVAN – draft on 6/21)

    1. (1) Berry Montoring Accumulating Data, (2) Cornfield/Greenhouse On certain aspects, (3) Cornfield Sequential Trials, (4) A Bayesian Test of Some Classical Hypotheses, with Applications to Sequential Clinical Trials Jerome Cornfield
    2. Bayes & monitoring, based on posterior distributions
    3. First papers in Bayes sequential monitoring. Bayesian inferences not affected by frequent or continual monitoring by the likelihood principle.
    4. Papers which compare to frequentist stopping rules & increased interpretation on role of priors.
    5. Freedman/Spiegelhalter Comparison of Bayesian with group sequential: The need to overcome this '**handicap**' prevents unduly early termination.
    6. Spiegelhalter/Freedman/Parmar Bayesian approaches to randomized trials -
    7. Spiegelhalter/Freedman/Parmar Applying Bayesian ideas - predictive distributions as basis for monitoring
    8. Freedman/Spiegelhalter - choice of prior explained by showing its impact on percentiles of posterior distribution

9. Fayers/Ashby/Parmar Tutorial in biostatistics... Choosing these two priors (skeptical, enthuastic) provides a useful **brake** against the premature termination of trials.
10. Bayesian Adaptive Methods for Clinical Trials Berry, Carlin, Etc.

- Our majors contribution (EVAN – as early as possible in introduction without having the flow appear weird – draft on 6/21)
- Outline for the remaining section of the paper (EVAN – draft on 6/21)

2 Methods

As you introduce ideas that come from or extend other ideas in the literature, cite the relevant literature.

2.1 Monitoring versus Estimation Priors (EVAN – draft on 6/21)

- Define generally in terms of $\boldsymbol{\theta} = (\gamma, \boldsymbol{\psi})$ where γ is a parameter of interest and $\boldsymbol{\psi}$ is a nuisance parameter (possible vector valued).
- Define *Monitoring* Priors and *Inference* Priors.
- Make connection between Inference priors and two-part mixture prior and BMA.
- Define *Skeptical* and *Enthusiastic* monitoring priors and how each would be used.
- I would have a generic graphic to illustrate the types of priors and the mixture.

2.1.1 Bayesian hypothesis testing

Suppose the parameter space is Θ and consider testing the hypothesis $H_0 : \theta \in \Theta_{H_0}$ versus $H_1 : \theta \in \Theta_{H_1}$, where $\Theta = \Theta_{H_0} \cup \Theta_{H_0}$ and $\Theta_{H_0} \cap \Theta_{H_0} = \emptyset$. Let D denote the data collected in the experiment and let $\pi \equiv \pi(\theta)$ denote a prior distribution for θ.

2.1.2 Skeptical and enthuastic monitoring priors

Consider a-priori beliefs about the true value of θ before data is collected. Define the skeptical viewpoint as being "all but convinced" that $\theta \in \Theta_{H_0}$ and define the enthuastic viewpoint as being "all but convinced" that $\theta \in \Theta_{H_1}$. Associate these viewpoints with priors $\pi_{Skeptical}$ and $\pi_{Enthuastic}$ such that the quantities $P(\theta \in \Theta_0 | \pi_{Skeptical})$ and $P(\theta \in \Theta_1 | \pi_{Enthuastic})$ reflect a high amount of certainty. Based on the study design there will be many options for specification of these priors (see Section 2.1.4).

These viewpoints can be used in monitoring the trial once data is collected, using the posterior probability of θ under different prior specifications. It is reasonable to stop the trial for efficacy if, once presented with the data, the skeptical viewpoint as updated via Bayes rule becomes convinced of the alternative hypothesis, that is, if $P(\theta \in \Theta_1 | D, \pi_{Skeptical})$ is close to 1. Similarly, it is reasponable to stop the trialy for futility if, $P(\theta \in \Theta_0 | D, \pi_{Enthuastic})$ is close to 1. The concept of using the the skeptical and enthuastic priors for trial monitoring has been presnted in XXXX and YYYY. (Include interpretations).

2.1.3 Inference priors

Once the trial is stopped based on the monitoring priors, or at the planned end of the trial, it is necessary to use a prior to make final inference on θ. It is inadvisable to use either

the skeptical or enthuastic prior for inference since they are admittedly biased opinions in the direction of the null or the alternative for monitoring purposes. For inference purposes it is better to use an impartial prior $\pi_{Inference}$. Define $p(D|\pi) = \int L(\theta|D)\pi(\theta)d\theta$ to be the marginal likelihood for the data given the prior π. One can average the posterior means from the analyses using the skeptical and enthuastic priors, such that,

$$\pi_{Inference} = \frac{p(D|\pi_{Skeptical})\pi_{Skeptical} + p(D|\pi_{Enthuastic})\pi_{Enthuastic}}{p(D|\pi_{Skeptical}) + p(D|\pi_{Enthuastic})}$$

Let $\omega = p(D|\pi_{Skeptical}/(p(D|\pi_{Skeptical}) + p(D|\pi_{Enthuastic}))$. Then

$$E(\theta|D, \pi_{Inference}) = \omega \times E(\theta|D, \pi_{Skeptical}) + (1 - \omega) \times E(\theta|D, \pi_{Enthuastic})$$

Need to describe relation to Type I and Type II error.

2.1.4 Default parameterization of monitoring priors for common designs

Define prior distribution as $\pi(\theta|\lambda)$ where λ is a vector of hyperparameters.

Skeptical prior might be centered around the clinical inferiority boundary. Enthuastic prior might be centered around clinical superiority boundry. Variance based on tail area considerations, or by matching the skeptical prior variance.

Reference prior attempts to express no particular opinion about the treatment's merit.

Spiegelhalter et al. 1994 "an adversary who will need to be disilusioned by the data to stop further experimentation."

2.2 Futility Monitoring Using Probability of Success (EVAN – draft on 6/21)

- Futility monitoring using POS is about stopping early when their is a high likelihood of a study being inconclusive at the end of the study.

- Since the final analysis uses the *Inference* prior, POS should be based on the inference prior.

- Develop the framework for POS and show how it is a weighted average POS based on the skeptical and enthusiastic priors.

As an alternative strategy to futility analysis, one can monitor the probability of success (POS) for the trial. Let $\pi(\theta|D) \equiv \pi(\theta|D, \pi_{Skeptical})$ denote the posterior distribution for θ based on the skeptical prior $\pi_{Skeptical}$ and the current data D. Let $\psi \equiv \psi(D, \pi_{Skeptical})$ denote the POS which is given as follows:

$$\psi = E[1\{P(\theta > 0.20|D_1, D, \pi_{Skeptical}) \geq 0.95\}]$$

where the expectation is taken with respect to the posterior predictive distribution $p(D_1)$ for future data D_1 (ongoing + subjects yet to enroll).

$$p(D_1) = \int p(D_1|\theta) \times \pi(\theta|D) d\theta$$

One may stop the enrollment if ψ is sufficiently small (i.e., $\psi < 0.05$).

Stochastic curtailment in frequentist setting.

3 Examples – (EVAN)

3.1 Single-Arm Oncology Proof-of-Activity Trial w/ Binary Endpoint

For example, consider testing the hypothesis $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ where θ is a treatment effect of interest. Suppose an effect $\theta_1 > \theta_0$ is thought to be highly clinically

relevant and plausible given what limited data is available. Consider a clinical trial with a single analysis and fixed sample size chosen that there is a high probability of proving H_1 when $\theta = \theta_1$. A standard Bayesian decision rule would reject H_0 when $P(\theta > \theta_0|D) \geq 0.95$ which will result in a type one error rate of 0.05 (approximately) if $\theta = \theta_0$ when the analysis prior is non-informative (a so-called reference or flat prior).

3.2 Parallel Two-Group Superiority Trial /w Continuous Binary Endpoint

3.3 Three-Arm, Placebo Controlled Non-Inferiority Trial w/ Continuous Endpoint

4 Discussion – (MATT/EVAN)

SUPPLEMENTARY MATERIAL

5 BibTeX

References