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 25, 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. Foundational Bayesian sequential monitoring Cornfield (1966*a*) Cornfield (1966*b*) Neyman & Greenhouse (1967).
    2. Further development Freedman & Spiegelhalter (1989) Freedman & Spiegelhalter (1992) Spiegelhalter et al. (1993) Spiegelhalter et al. (1994) Fayers et al. (1997).
    3. Need to go further, and mention work on 3-arm Bayesian trials.
- 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 based on posterior probabilities

The Bayesian formulation allows direct inference on a parameter of interest through specification a model for the data generating mechanism and for prior distributions of the unknown quantities. Let $\mathbf{D} \in \mathbb{R}^n$ be a random variable representing the data collected in the trial with density $p(\mathbf{D}|\theta, \psi)$, where $\theta \in \Theta \subseteq \mathbb{R}^{m_1}$, $\psi \in \Psi \subseteq \mathbb{R}^{m_2}$, and $n, m_1, m_2 \in \mathbb{N}$. Suppose it is desired to test 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 $\Theta_{H_0} \cap \Theta_{H_0} = \emptyset$. Consider the posterior destribution of the parameter of interest θ given the data $\mathbf{D}$, $P(\theta \in \Theta_{H_i}|\mathbf{D}) = \int_{\Theta_{H_i}} p(\theta|\mathbf{D})d\theta$ for $i \in \{0, 1\}$, where $p(\theta|\mathbf{D}) = \int_{\Psi} p(\theta, \psi|\mathbf{D})d\psi$. The posterior distribution of the parameter of interest θ is unconditional on values of the nuisance paramter ψ. The posterior distribution of θ depends on the choice of prior distribution $\pi(\theta, \psi)$, since $p(\theta, \psi|\mathbf{D}) = p(\mathbf{D}|\theta, \psi)\pi(\theta, \psi)/p(\mathbf{D})$ by Bayes rule. Possible simplifying functional relationships between θ and ψ are conditional independence $\pi(\theta, \psi) = \pi(\theta|\psi)\pi(\psi)$ and independence $\pi(\theta, \psi) = \pi(\theta)\pi(\psi)$.

Inferences on the posterior distribution of θ will be central to all aspects of the trial procedure. *Define inference priors. Define monitoring priors.*

Summarizing Spiegelhalter et al. (1993) and others: A Bayesian may monitor data continually and stop collection when any of the following criteria have been met: a sufficiently *skeptical person* is convinced that H_1 is true, a sufficiently *enthuastic person* is convinced that H_1 is false or that the benefit of treatment is not likely to be what was expected, the probability of *eventually* proving that H_1 is true is sufficiently low, or the resources have been exhausted. *Define all but convinced.* The use of monitoring based on changing the opinion of skeptical and enthuastic priors has been described as overcoming a handicap (Freedman & Spiegelhalter (1989)), constructing "an adversary who will need to be disilusioned by the data to stop further experimentation" (Spiegelhalter et al. (1994)) and providing a brake (Fayers et al. (1997)) on the premature termination of trials.

We will illustrate pragmatic aproaches for elicitation of skeptical and enthuastic priors by linking elicitation to the set of hypotheses to be tested. 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 $\pi_S(\theta, \psi) \equiv \pi_S$ and $\pi_E(\theta, \psi) \equiv \pi_E$ such that the quantities $P(\theta \in \Theta_0 | \pi_S)$ and $P(\theta \in \Theta_1 | \pi_E)$ reflect a high amount of certainty. Based on the study design there will be many options for specification of these priors (see Section 2.1.3). 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 $P(\theta \in \Theta_1 | \mathbf{D}, \pi_S)$ is close to 1. Similarly, it is reasponable to stop the trialy for futility if, $P(\theta \in \Theta_0 | \mathbf{D}, \pi_E)$ is close to 1.

2.1.2 Inference priors

Once enrollment 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 potentially inadvisable to use either the skeptical or enthuastic prior for inference since they represent quite extreme opinions in about the hypotheses to be tested. For inference purpose, we propose use of a mixture prior constructed from the monitoring process. Define $p(\mathbf{D}|\pi) = \int L(\theta|\mathbf{D})\pi(\theta)d\theta$ 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,

$$E(\theta|\mathbf{D}, \pi_{Inference}) = \omega \times E(\theta|\mathbf{D}, \pi_S) + (1-\omega) \times E(\theta|\mathbf{D}, \pi_E)$$

for $\omega \in [0, 1]$. Choosing $\omega = 1/2$ corresponds to an agnostic inference prior. Choosing ω based on posterior model probabilities of the null and alternative hypothesis yields $\omega = p(\mathbf{D}|\pi_S/(p(\mathbf{D}|\pi_S) + p(\mathbf{D}|\pi_E))$.

2.1.3 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.

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 $\pi(\theta|D) \equiv \pi(\theta|D, \pi_S)$ denote the posterior distribution for θ based on the skeptical prior π_S and the current data D. Let $\psi \equiv \psi(D, \pi_S)$ denote the POS which is given as follows:

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

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

Cornfield, J. (1966*a*), 'A Bayesian Test of Some Classical Hypotheses, with Applications to Sequential Clinical Trials', *Journal of the American Statistical Association* **61**(315), 577.
URL: *https://www.jstor.org/stable/2282772?origin=crossref*

Cornfield, J. (1966*b*), 'Sequential Trials, Sequential Analysis and the Likelihood Principle', *The American Statistician* **20**(2), 18.
URL: *https://www.jstor.org/stable/2682711?origin=crossref*

Fayers, P. M., Ashby, D. & Parmar, M. K. B. (1997), 'Tutorial in Biostatistics: Bayesian Data Monitoring in Clinical Trials', *Statistics in Medicine* **16**(12), 1413–1430.
URL: *http://doi.wiley.com/10.1002/%28SICI%291097-0258%2819970630%2916%3A12%3C1413%3A%*
SIM578%3E3.0.CO%3B2-U

Freedman, L. S. & Spiegelhalter, D. J. (1989), 'Comparison of Bayesian with group sequential methods for monitoring clinical trials', *Controlled Clinical Trials* **10**(4), 357–367.
URL: *https://www.sciencedirect.com/science/article/pii/0197245689900019?via%3Dihub*

Freedman, L. S. & Spiegelhalter, D. J. (1992), 'Application of bayesian statistics to decision making during a clinical trial', *Statistics in Medicine* **11**(1), 23–35.
URL: *http://doi.wiley.com/10.1002/sim.4780110105*

Neyman, J. & Greenhouse, S. W. (1967), *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability.*, University of California Press.
URL: *https://projecteuclid.org/euclid.bsmsp/1200513830*

Spiegelhalter, D. J., Freedman, L. S. & Parmar, M. K. B. (1993), 'Applying Bayesian ideas in drug development and clinical trials', *Statistics in Medicine* **12**(15-16), 1501–1511.
URL: *http://doi.wiley.com/10.1002/sim.4780121516*

Spiegelhalter, D. J., Freedman, L. S. & Parmar, M. K. B. (1994), 'Bayesian Approaches to Randomized Trials', *Journal of the Royal Statistical Society. Series A (Statistics in Society)* **157**(3), 357.
URL: *https://www.jstor.org/stable/10.2307/2983527?origin=crossref*