A Framework for Adaptively Incorporating External Evidence in Sequentially Monitored Trials

Matthew A. Psioda

Department of Biostatistics, University of North Carolina at Chapel Hill

September 23rd, 2021

Joint work with Evan Kwiatkowski, Eugenio Andraca-Carrera, and Mat Soukup
(In partial fulfillment of doctoral thesis requirements for Evan Kwiatkowski)

Outline

Motivation – Pediatric Therapeutic Development

Background

- The Pediatric Research Equity Act (PREA) and prior legislation has spurred pediatric research to inform new drug labeling by allowing the FDA to request/require trials in pediatric disease populations [Avant *et al.*, 2018].

- At this time, both the EMA/FDA require pediatric research plans at certain stages of the application for authorization of new medicines.

- Significant challenges often occur in pediatric trials necessitating timeline extensions and/or modifications to original designs developed to meet FDA post marking requirements (PMRs).

- Pediatric trials face numerous barriers (e.g., small participant pools and ethical/parental concerns related to participation [Greenberg *et al.*, 2018]) and are often unable to produce substantial evidence of efficacy.

Motivating Example – Belimumab Development Program I

- Based on data generated by two well-controlled pivotal trials in adult systemic lupus erythematosus (SLE), the BLISS-52 and BLISS-76 trials, the FDA approved belimumab for the treatment of adults with active, seropositive SLE (who are already on standard therapy).

		Study 1056			Study 1057	
	Placebo N=275	Belimumab 1 mg/kg N=271	Belimumab 10 mg/kg N=273	Placebo N=287	Belimumab 1 mg/kg N=288	Belimumab 10 mg/kg N=290
Response, n (%)	93 (34)	110 (41)	118 (43)	125 (44)	148 (51)	167 (58)
Observed difference	-	7%	9%	-	8%	14%
Odds ratio (95% CI)	-	1.3 (0.9, 1.9)	1.5 (1.1, 2.1)	-	1.6 (1.1, 2.2)	1.8 (1.3, 2.6)

Source: Review of BLA 125370 Belimumab IV dated February 18, 2011.

Figure: BLISS Trials Primary Endpoint Data: SRI Response Rate

Belimumab Development Program II

- As part of the FDA's 2011 approval of belimumab for treatment of adult SLE, a pediatric PMR was issued under PREA.
- This PMR was to conduct a phase II, multicenter study to evaluate the safety, efficacy, and pharmacokinetics of belimumab in a pediatric SLE population.
- The 2011 PMR required a trial with $N = 100$ patients but was reissued in 2016 to require $N = 70$ patients due to enrollment challenges.

	Placebo	Belimumab 10 mg/kg
Unadjusted Response n/N (%)	17/39 (43.6)	28/53 (52.8)
Observed difference	-	9.2
Odds ratio (95% CI)	-	1.5 (0.6, 3.3)

Figure: PLUTO Trial Primary Endpoint Data: SRI Response Rate

Belimumab Development Program III

- For the PLUTO trial, there could be no preconception that the trial was adequate to produce substantial evidence of efficacy on its own.
- Required sample size for 90% power $\rightarrow N = 760$.
- Nonetheless, the FDA concluded that the safety/efficacy data provided by the PLUTO trial, *taken together with the efficacy/safety data from the BLISS trials*, supported approval of belimumab for pediatric use [U.S. Food and Drug Administration, 2018].
- As a part of the multidisciplinary review, a *post-hoc* Bayesian analysis was requested that borrowed information from the BLISS trials.

Bayesian Sequential Monitoring

Philosophy

- Summarizing Spiegelhalter *et al.* [1994] and others, a Bayesian may monitor data continually and stop collection when any of the following criteria have been met:
  - A sufficiently skeptical observer becomes convinced H_1 is true.
  - A sufficiently enthusiastic observer becomes convinced H_1 is unlikely to be true *or* that the benefit of treatment is not what was hoped.
  - Resources allocated for the study have been exhausted.
- One can analyze accumulating data as frequently as is logistically feasible until one of these criteria is met.

Evidence from the Bayesian Perspective I

- Consider testing the hypotheses

$$H_0 : \theta \leq \theta_0 \text{ versus } H_1 : \theta > \theta_0$$

  where

  - θ is a parameter of interest (e.g., difference in response rates), and
  - θ_0 is a chosen constant (e.g., most commonly $\theta_0 = 0$).

- Suppose there exists a constant $\theta_1 > \theta_0$ that is thought to be a clinically meaningful and also a plausible.
  - $\theta_1 = 0.118$ based on the BLISS trial data

Evidence from the Bayesian Perspective II

- A standard Bayesian decision rule is given by

$$P_\pi\left(\theta > \theta_0 \middle| \mathbf{D}\right) > 1 - \epsilon,$$

  based on some prior distribution $\pi(\theta)$ and data $\mathbf{D}$.

- One may formally define the concept of substantial total evidence in favor of a claim (e.g., $\theta > \theta_0$) as the posterior probability of the claim exceeding $1 - \epsilon$.

- Total information = information from likelihood + prior information

- Sequential monitoring rationale:

  - Skeptic: substantial total evidence that $\theta > \theta_0 \implies$ enrollment stop.
  - Enthusiast: substantial total evidence that $\theta < \theta_1 \implies$ study stop.

Skeptical and Enthusiastic Priors I

- Define a skeptical observer as someone who ...

  (i) believes that θ_0 is the most likely value of θ.

  (ii) is all but convinced that $\theta < \theta_1$.

  This is operationalized through a skeptical prior $\pi_S(\theta)$ satisfying $\text{argmax}_\theta \ \pi_S(\theta) = \theta_0$ and $P_S(\theta < \theta_1) = 1 - \epsilon$.

- Define an enthusiastic observer as someone who ...

  (i) believes that θ_1 is the most likely value of θ.

  (ii) is all but convinced that $\theta > \theta_0$.

  This is operationalized through an enthusiastic prior $\pi_E(\theta)$ satisfying $\text{argmax}_\theta \ \pi_E(\theta) = \theta_1$ and $P_E(\theta > \theta_0) = 1 - \epsilon$.

Skeptical and Enthusiastic Priors II

- For prior construction, we propose the flexible family of generalized normal distributions – $\mathcal{GN}(\mu, \alpha, \beta)$

$$f(\theta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left\{-\left(\frac{|\theta - \mu|}{\alpha}\right)^{\beta}\right\}$$

  where

  - μ is a location parameter,
  - $\alpha > 0$ is a scale parameter, and
  - $\beta > 0$ is a shape parameter.

- This family of distributions accommodates a variety of shapes – both very peaked and very flat – and includes the normal distribution as a special case.

Skeptical and Enthusiastic Priors III

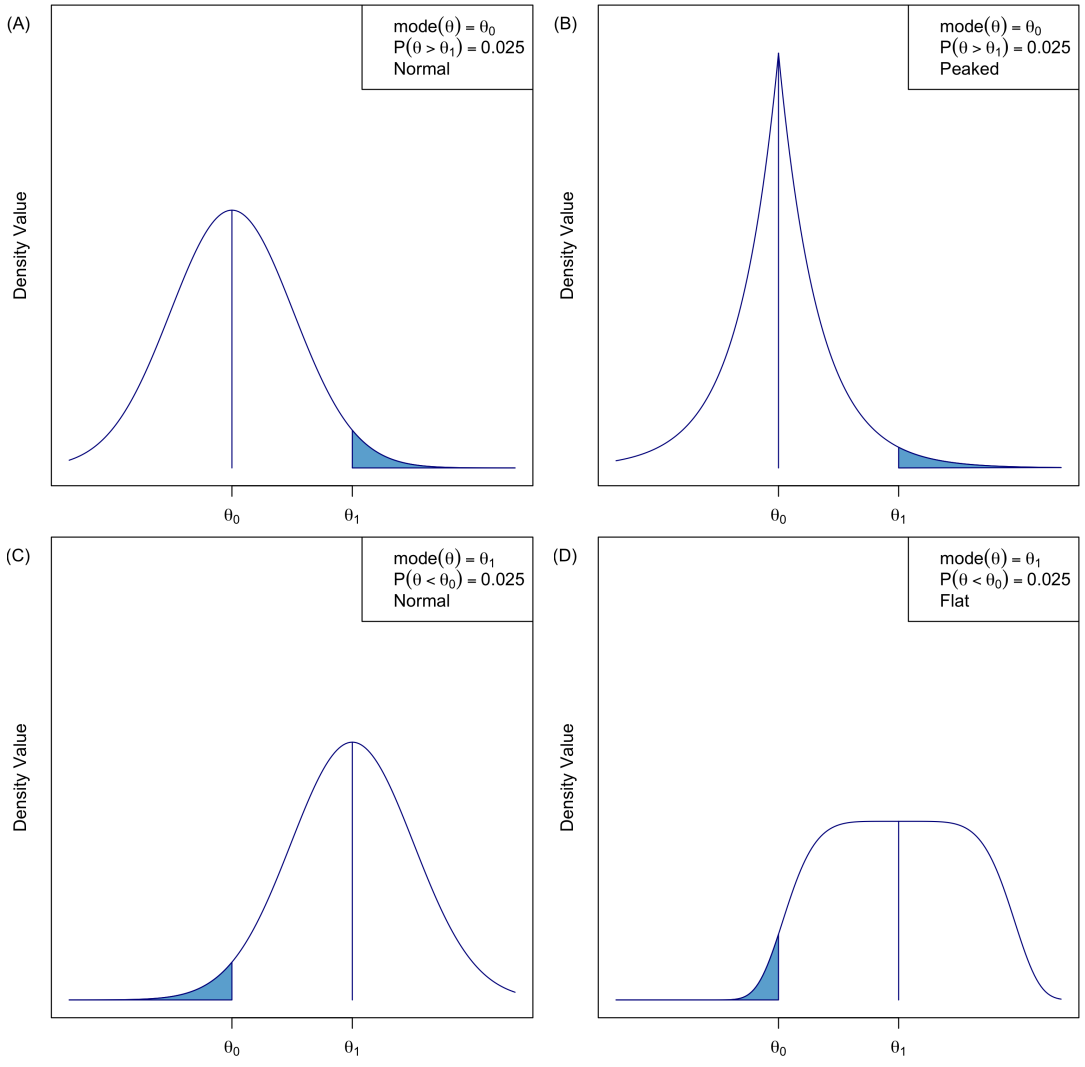

Figure: A, Default skeptical prior. B, Peaked skeptical prior. C, Default enthusiastic prior. D, Flattened enthusiastic prior.

Adaptive Monitoring Prior I

- One may define an *adaptive monitoring prior* for efficacy monitoring as the mixture distribution

$$\pi_{AMP}(\theta) = \omega \cdot \pi_E(\theta) + (1 - \omega) \cdot \pi_S(\theta)$$

- We take $\omega = (1 - \delta) \times \psi^{(E)}(\mathbf{D}_{\text{obs}})$, where
  - $\psi^{(E)}(\mathbf{D}_{\text{obs}}) \in [0, 1]$ is a measure of congruency between the observed pediatric data $\mathbf{D}_{\text{obs}}$ and the enthusiastic prior, and
  - $\delta \in [0, 1]$ represents the fraction of skepticism to be overcome by the pediatric data alone.

- $\delta = 1 \implies \pi_{AMP}(\theta) = \pi_S(\theta)$ regardless of $\psi^{(E)}(\mathbf{D}_{\text{obs}})$

Adaptive Monitoring Prior II

- Similar to Psioda and Xue [2020], we weight the enthusiastic prior based on congruency of the pediatric data with the enthusiastic prior predictive distribution using a Bayesian p-value (Box [1980]).
- The prior-predictive distribution for reflects the probability of observing data $\mathbf{D}$ given the enthusiastic prior.

$$p(\mathbf{D}) = \int p(\mathbf{D}|\theta)\pi_E(\theta)d\theta$$

- Let $\mathbf{D}_{\text{obs}}$ be the observed data at some point in time in an ongoing trial. Box's p-value is defined as the following:

$$\psi(\mathbf{D}_{\text{obs}}) = \int p(\mathbf{D})1[p(\mathbf{D}) \leq p(\mathbf{D}_{\text{obs}})]d\mathbf{D},$$

where $1[A]$ is an indicator that the event A is true.

An Illustration Based on the PLUTO Trial

Returning to the PLUTO Trial

- From the BLISS trial data, we elicited an enthusiastic prior that reflected a difference in response probabilities equal to $\theta_1 = 0.118$.
- A weakly informative prior was elicited for the placebo group response probability (denoted by η) centered at approximately $\eta_0 = 0.38$ (consistent with the BLISS data).

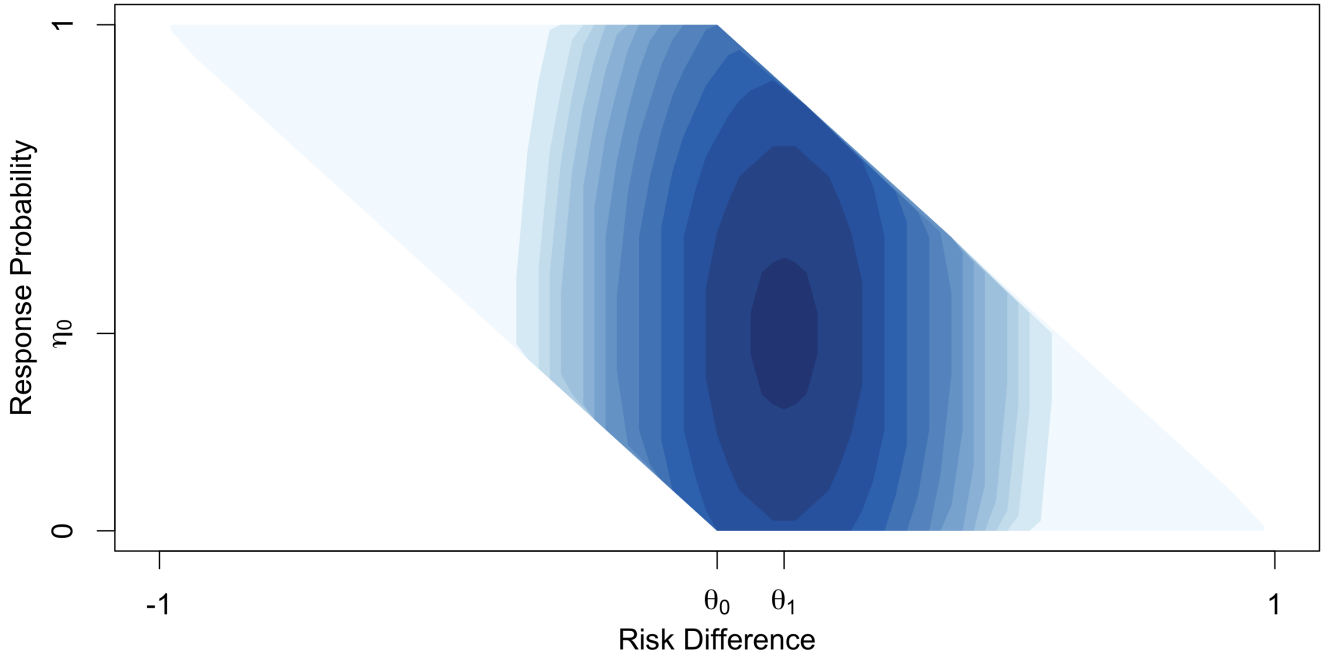

Figure: $\pi(\theta, \eta) = \pi(\theta) \times \pi(\eta|\theta)$ over $-1 < \theta < 1$ and $0 < \theta + \eta < 1$.

Final PLUTO Trial Data

- Ultimately, 93 patients were enrolled over approximately 52.5 months (approximately 1 patient enrolled per 17 days).
- Clinical response was observed in 28 of 53 (52.8%) of patients randomized to belimumab and in 17 of 40 (43.6%) of patients randomized to placebo.
- Posterior characteristics for response rate difference (using non-informative priors)
  - Posterior Mean $\rightarrow$ 0.09
  - 95% Credible Interval $\rightarrow$ [-0.11, 0.30]

Re-analysis of Accumulating PLUTO Trial Data

Table: Summary characteristics of re-analysis of PLUTO trial.

δ	SS (I/F)	$\psi^{(E)}(\mathbf{D}_{\text{obs}})$ (I/F)	ω (I/F)	$P(\theta > \theta_0 = 0)$ (I/F)
0.00	62 / 90	0.914 / 0.965	0.914 / 0.965	0.980 / 0.979
0.05	64 / 92	0.876 / 0.934	0.833 / 0.887	0.976 / 0.962
0.10	76 / 92	0.941 / 0.934	0.847 / 0.841	0.975 / 0.951
0.15	92 / 92	0.934 / 0.934	0.794 / 0.794	0.940 / 0.940
0.20	92 / 92	0.934 / 0.934	0.747 / 0.747	0.928 / 0.928
0.25	92 / 92	0.934 / 0.934	0.701 / 0.701	0.917 / 0.917

SS = Sample Size, I/F = Interim/Final.

Box's P-value Using Enthusiastic Prior $\psi^{(E)}(\mathbf{D}_{\text{obs}})$

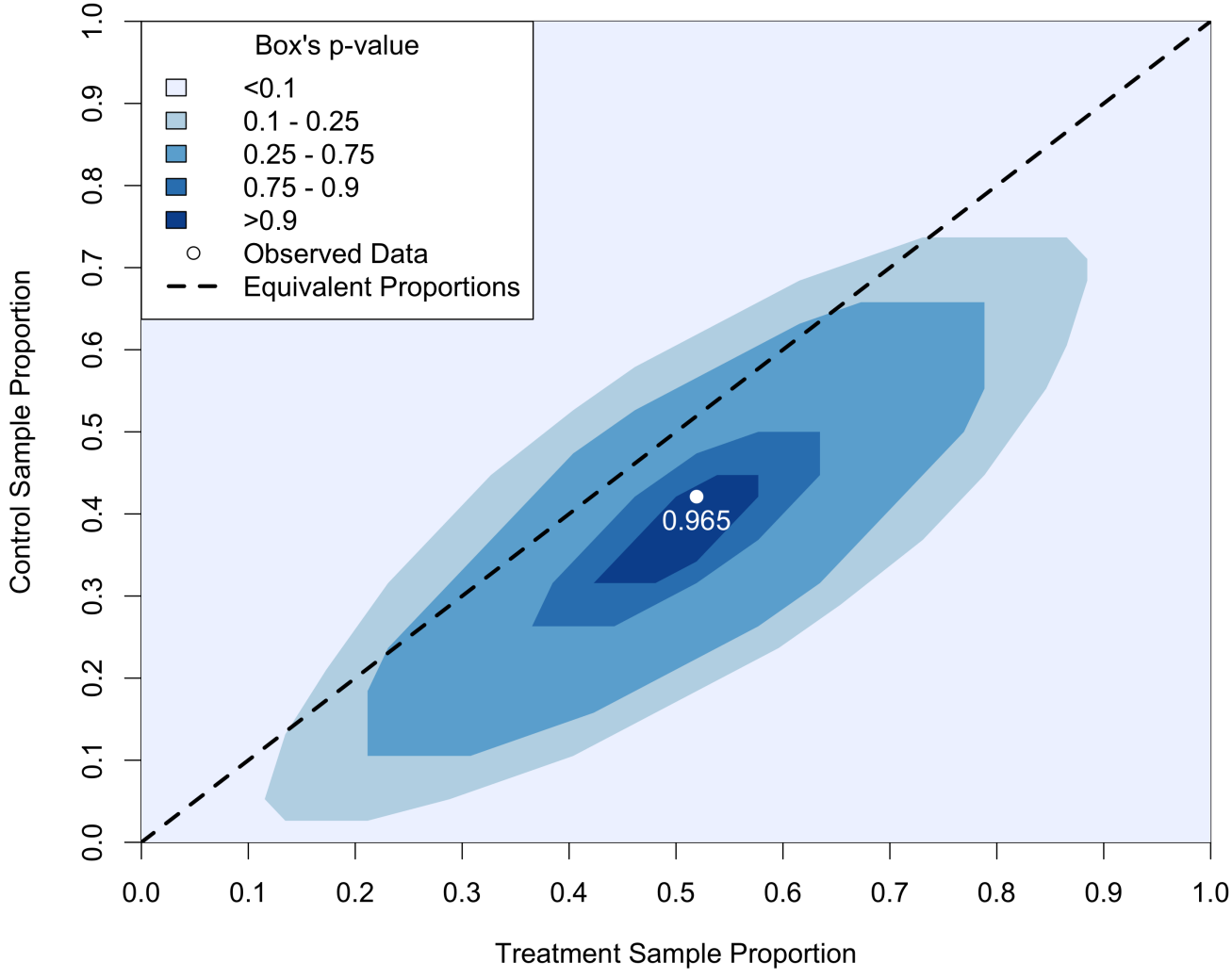

Figure: A, Values of $\psi^{(E)}(\mathbf{D}_{\text{obs}})$ by control and treatment sample proportions at the final analysis with 90 subjects for the $\delta = 0$ case.

Posterior Probability of Treatment Efficacy

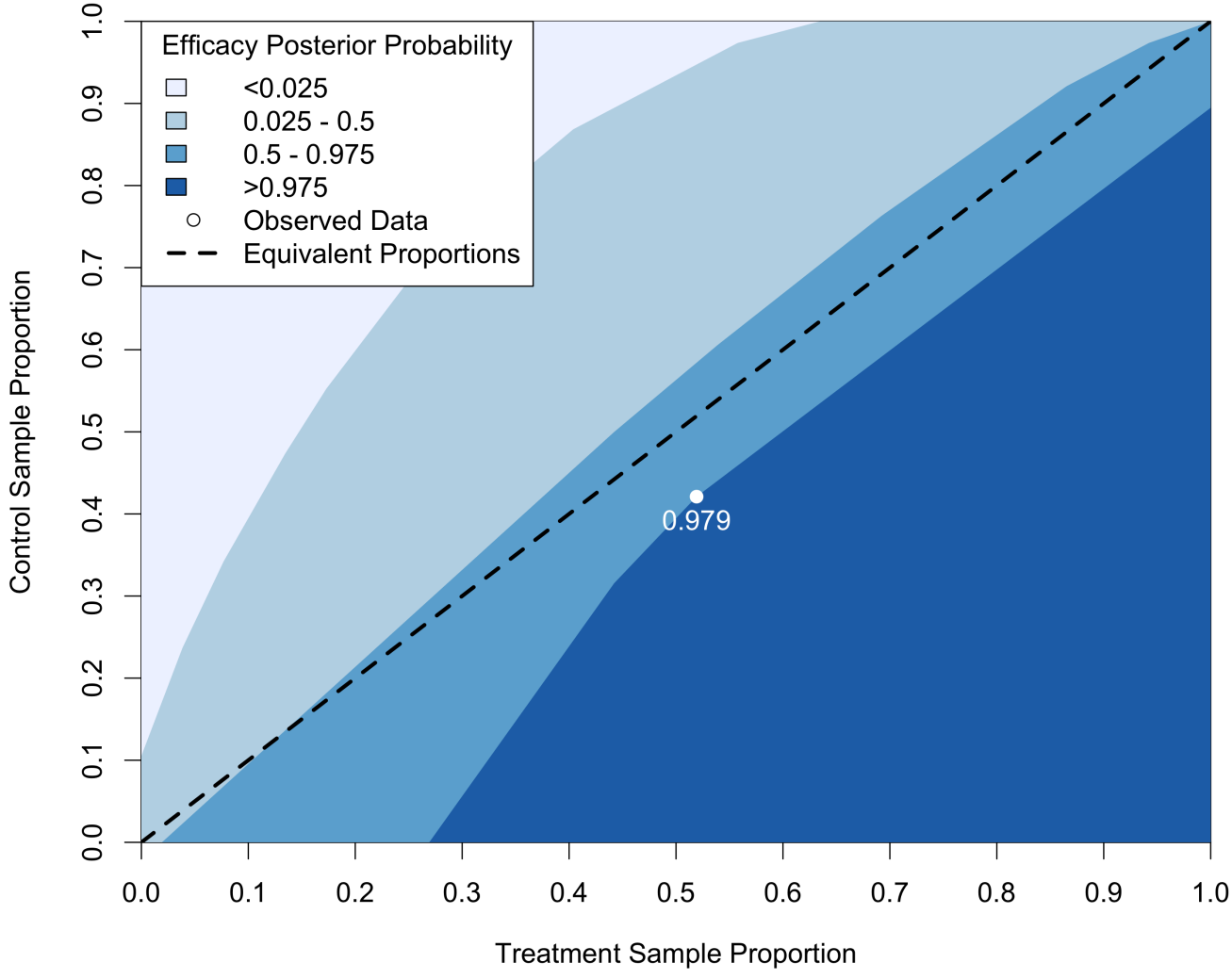

Figure: B, $P(\theta > \theta_0 = 0)$ by control and treatment sample proportions at the final analysis with 90 subjects for the $\delta = 0$ case.

Closing Remarks

Closing Remarks

- The formulation of the enthusiastic prior enforces that there be residual uncertainty that the null hypothesis is true; it demonstrates strong belief about effectiveness of the treatment yet is still consistent with a modicum of equipoise.

- Even in the most liberal case where $\delta = 0$, the residual uncertainty that the null hypothesis is true reflected in the adaptive monitoring prior cannot be less than that reflected in the enthusiastic prior.

- This is a critical feature of the design as it enforces the requirement that observed data must demonstrate some degree of efficacy on their own to justify stopping enrollment early.

References I

Avant, D., Wharton, G. T., and Murphy, D. (2018). Characteristics and Changes of Pediatric Therapeutic Trials under the Best Pharmaceuticals for Children Act. *Journal of Pediatrics*, **192**, 8–12.

Box, G. E. P. (1980). Sampling and Bayes' Inference in Scientific Modelling and Robustness. *Journal of the Royal Statistical Society. Series A (General)*, **143**(4), 383–430.

Greenberg, R. G., Gamel, B., Bloom, D., Bradley, J., Jafri, H. S., Hinton, D., Nambiar, S., Wheeler, C., Tiernan, R., Smith, P. B., Roberts, J., and Benjamin, D. K. (2018). Parents' perceived obstacles to pediatric clinical trial participation: Findings from the clinical trials transformation initiative. *Contemporary Clinical Trials Communications*, **9**, 33–39.

Psioda, M. A. and Xue, X. (2020). A Bayesian Adaptive Two-Stage Design for Pediatric Clinical Trials. *Journal of Biopharmaceutical Statistics*, pages 1–18.

References II

Spiegelhalter, D. J., Freedman, L. S., and Parmar, M. K. B. (1994). Bayesian Approaches to Randomized Trials. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **157**(3), 357.

U.S. Food and Drug Administration (2018). Benlysta® multi-disciplinary review and evaluation. *[accessed 4/1/2020]*.