# Clinical trial monitoring with Bayesian hypothesis testing

John D. Cook
Valen E. Johnson

August 6, 2008

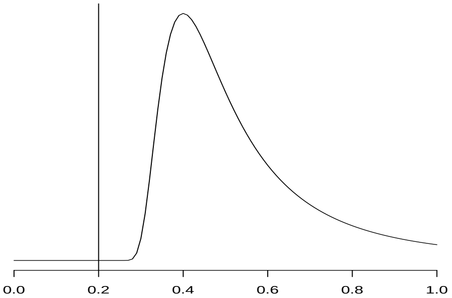
# Estimation and Testing

- Bayesians typically approach a clinical trial as an estimation problem, not a test.
- Possible explanation: poor operating characteristics ...
- Unless you choose your alternative prior well.

# Local prior operating characteristics

- Point null hypothesis versus alternative prior that assigns positive probability to the null
- When simulating from the alternative, Bayes factor in favor of alternative grows like $e^n$.
- When simulating from the null, Bayes factor in favor of null grows like $n^{1/2}$.
- Hard to ever reject the null.

# Inverse moment priors (iMOM)




$$\pi_1(\theta) \propto (\theta - \theta_0)^{-\nu-1} \exp\left(-\lambda(\theta - \theta_0)^{-2k}\right) [\theta > \theta_0]$$

# iMOM Convergence rates

When simulating from alternative

$$p \lim_{n\to\infty} n^{-1} \log BF_n(1|0) = c > 0.$$

(Well known result.)

When simulating from null,

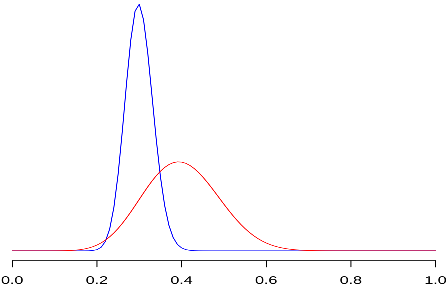
$$p \lim_{n\to\infty} n^{-k/(k+1)} \log BF_n(1|0) = c < 0.$$

(New result.)

# Thall-Simon method

- Historical standard: $\theta_S \sim \text{Beta}(a_S, b_S)$. Parameters $a_S$ and $b_S$ large.
- Experimental treatment: $\theta_E \sim \text{Beta}(a_E, b_E)$ a priori, $a_E$ and $b_E$ small.
- Stop for inferiority if $P(\theta_E < \delta + \theta_S \mid \text{data})$ is large.
- Stop for superiority if $P(\theta_E > \theta_S \mid \text{data})$ is large.
- Operating characteristics degrade without $\delta > 0$.
- Inconsistent in limit: both stopping rules could apply.

# Thall-Simon plot




Beta(60, 140) historical, Beta(12, 18) experimental

# Comparing Bayes factor with Thall-Simon

Historical response 20%, alternative 30%. Fifty patients maximum.

Bayes factor design:

- $H_0$: $\theta = 0.2$
- $H_1$: iMOM prior with mode 0.3.
- Stop for inferiority if $P(H_0 \mid \text{data}) > 0.9$.
- Stop for superiority if $P(H_1 \mid \text{data}) > 0.9$.

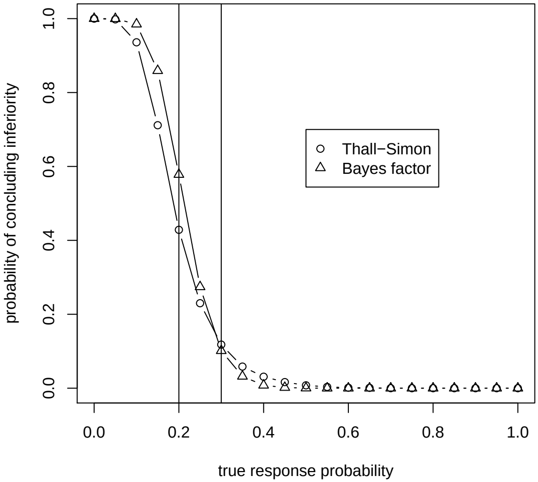
# Comparing Bayes factor with Thall-Simon, cont.

Thall-Simon design:

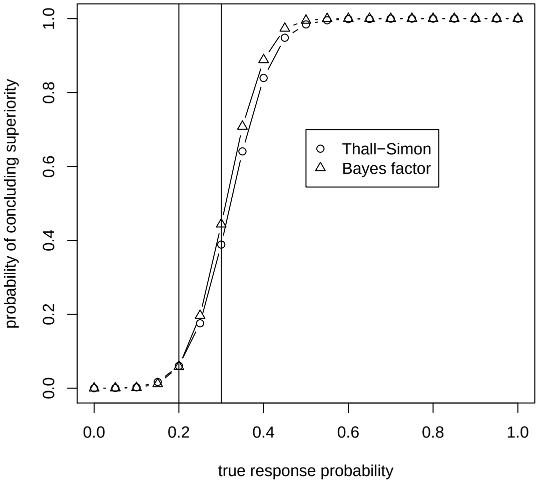
- $\theta_S \sim$ Beta(200,800)
- $\theta_E \sim$ Beta(0.6, 1.4) *a priori*
- Stop for inferiority if $P(\theta_S > 0.1 + \theta_E \mid \text{data}) > 0.976$.
- Stop for superiority if $P(\theta_E > \theta_S \mid \text{data}) > 0.99$.

Calibrated to match probability of stopping for wrong reason at null and alternative.

# Stopping for inferiority



# Stopping for superiority



# Thall-Wooten time-to-event method

- Analogous to Thall-Simon method for binary outcomes.
- $t \mid \theta \sim$ exponential with mean $\theta$, $\theta \sim$ inverse gamma
- Stop for inferiority if $P(\theta_S + 0.1 > \theta_E \mid \text{data})$ large ...
- Stop for superiority if $P(\theta_E > \theta_S \mid \text{data})$ large

# Comparing Bayes factor and Thall-Wooten method

Standard treatment 6 months PFS, alternative 8 months, maximum 50 patients

Bayes factor design:

- $H_0$: $\theta = 6$
- $H_1$: iMOM prior with mode 8.
- Stop for inferiority if $P(H_0 \mid \text{data}) > 0.9$.
- Stop for superiority if $P(H_1 \mid \text{data}) > 0.9$.

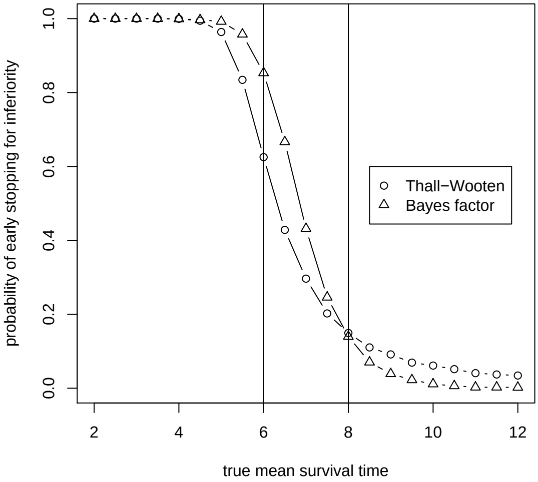
# Comparing Bayes factor and Thall-Wooten method, cont.

Thall-Wooten design:

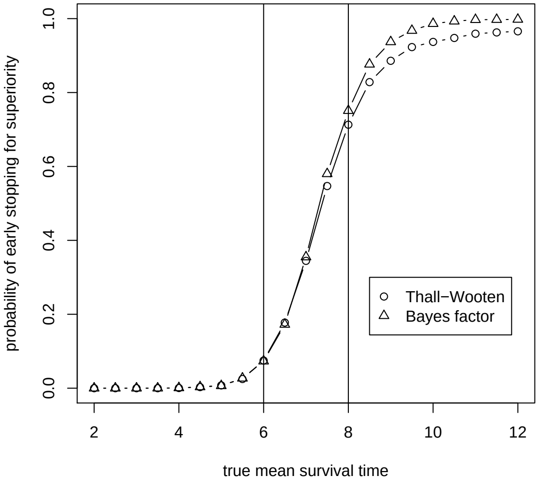
- $\theta_S \sim$ Inverse Gamma (20,1200)
- $\theta_E \sim$ Inverse Gamma(3, 12) *a priori*
- Stop for inferiority if $P(\theta_S + 2 > \theta_E \mid \text{data}) > 0.976$.
- Stop for superiority if $P(\theta_E > \theta_S \mid \text{data}) > 0.93$.

Calibrated to match probability of stopping for wrong reason at null and alternative.

# Stopping for inferiority



# Stopping for superiority



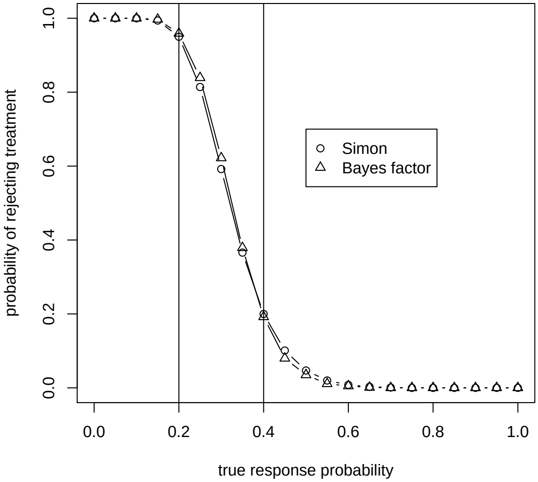
# Comparison with Simon two-stage design

Simon two-stage design to test null response rate 0.20 versus alternative rate 0.40.

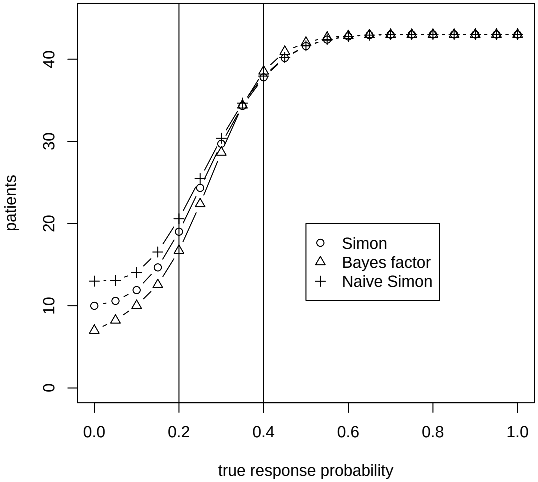
Reject 95% of the time under null, 20% under alternative.

Maximum of 43 patients: 13 in first stage, 30 in second stage.

# Comparison with Simon two-stage design: rejection probability



# Comparison with Simon two-stage design: patients used



# References

- Non-Local Prior Densities for Default Bayesian Hypothesis Tests by Valen E. Johnson and David Rossell http://www.bepress.com/mdandersonbiostat/paper42
- On the Bayesian Design of Clinical Trials Using Hypothesis Tests by Valen E. Johnson and John D. Cook http://www.bepress.com/mdandersonbiostat/paper47
- Bayes factor software: cook@mdanderson.org
- Thall-$*$ software: http://biostatistics.mdanderson.org