Generalizing boundaries for triangular designs, and efficacy estimation at extended follow-ups

Dominic Magirr, AstraZeneca (previously Lancaster University)

Acknowledgements:

Annabel Allison

Neal Alexander

Tansy Edwards

Raymond Omollo

Fabiana Alves

Plan

Background

- Visceral leishmaniasis.
- NCT01067443.

Design

- The origins of the triangular test.
- Fixing sample size at the first interim.

Estimation

- Sources of bias.
- Shrinkage estimation.

Summary

Visceral leishmaniasis

- Parasitic disease transmitted by sandflies and is fatal if left untreated.
- Symptoms are fever, weight loss, fatigue, anemia, swelling of the liver and spleen.
- Worldwide distribution: Asia, East Africa, South America, Mediterranean.
- Estimated 200,000 – 400,000 incident cases per year.
- Several treatments exist, but variety of problems (safety, resistance, cost).

NCT01067443

- Run by the Drugs for Neglected Diseases initiative.
- Three experimental treatment arms.
  - Miltefosine
  - AmBisome + Miltefosine
  - AmBisome + sodium stibogluconate
- Primary endpoint: cure at day 28 (yes / no).
- Secondary endpoint: cure at day 210 (yes / no).

Omollo *et al. Trials* 2011, **12**:166
http://www.trialsjournal.com/content/12/1/166

TRIALS

STUDY PROTOCOL Open Access

Safety and Efficacy of miltefosine alone and in combination with sodium stibogluconate and liposomal amphotericin B for the treatment of primary visceral leishmaniasis in East Africa: study protocol for a randomized controlled trial

Raymond Omollo[1], Neal Alexander[2], Tansy Edwards[2], Eltahir AG Khalil[3], Brima M Younis[3], Abuzaid A Abuzaid[3], Monique Wasunna[1,4], Njenga Njoroge[4], Dedan Kinoti[4], George Kirigi[4], Thomas PC Dorlo[5,6], Sally Ellis[7], Manica Balasegaram[7] and Ahmed M Musa[3*]

Challenge for us

1. Could we suggest improvements to the trial design?

   A sequential design (triangular test) was used, but complicated by the need to collect pharmacokinetic data from a minimum of 30 patients.

2. Could we suggest a better estimator for cure rate at day 210?

   An ad-hoc “probability tree estimator” had been used, but properties unclear.

What was the design?

- Consider each new treatment separately.
- Standard of care assumption: 75% probability of cure at day 28.
- Let p denote the probability of cure at day 28.
- Score:

$$S = n \times (\hat{p} - 0.75)$$

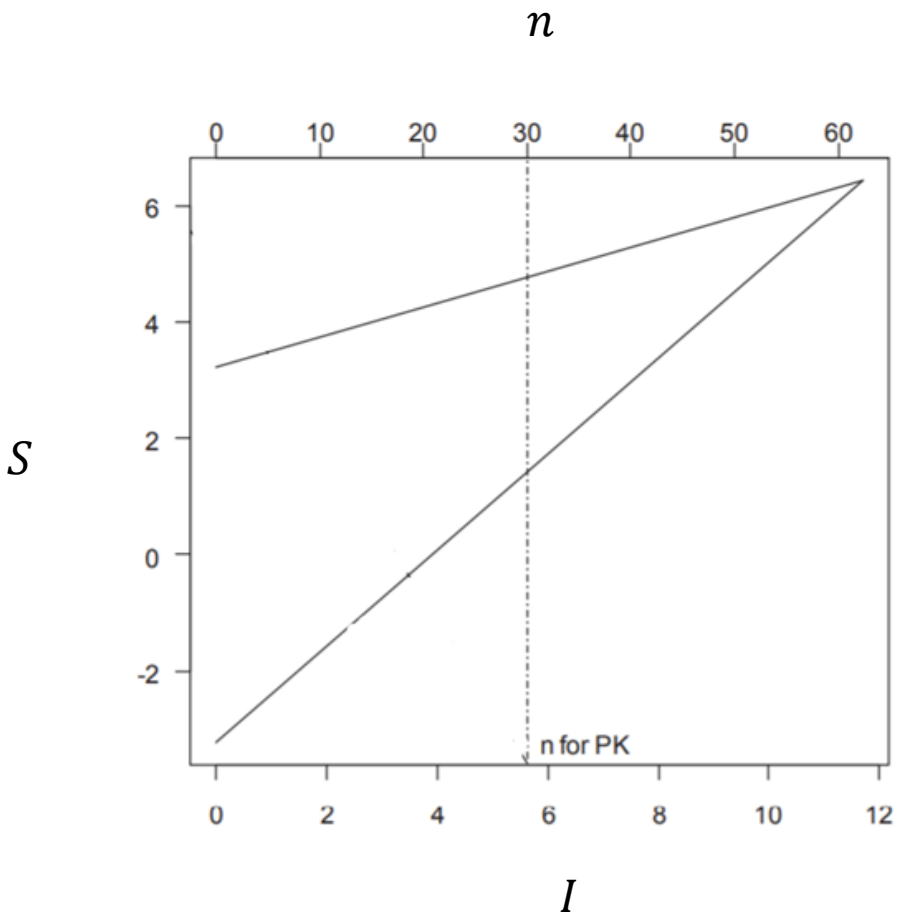
- Information:

$$I = n \times 0.75 \times (1 - 0.75)$$

- Approximately

$$S \sim N(\theta I, I),$$

$$\theta = \log\left(\frac{p}{1-p}\right) - \log\left(\frac{0.75}{1-0.75}\right).$$



Where does this come from?

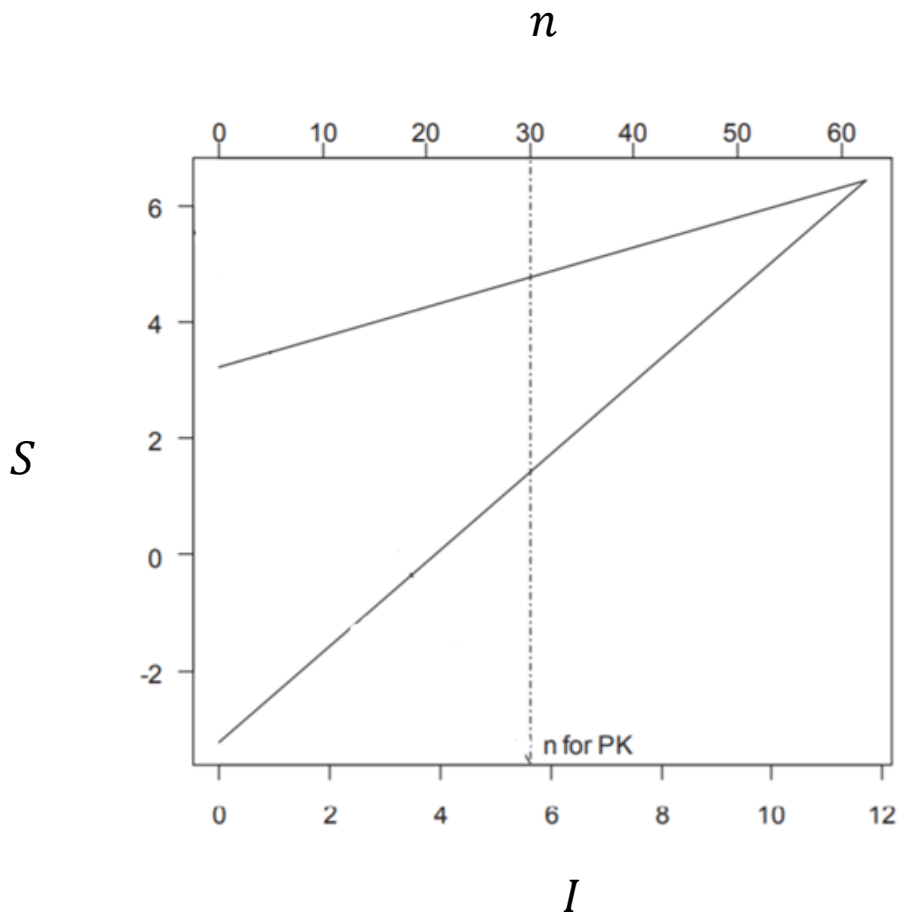
Sequential Probability Ratio Test (Wald, 1945)

↓

Sequential Probability Ratio Test 2
(Anderson, 1960; Lorden, 1976)

↓

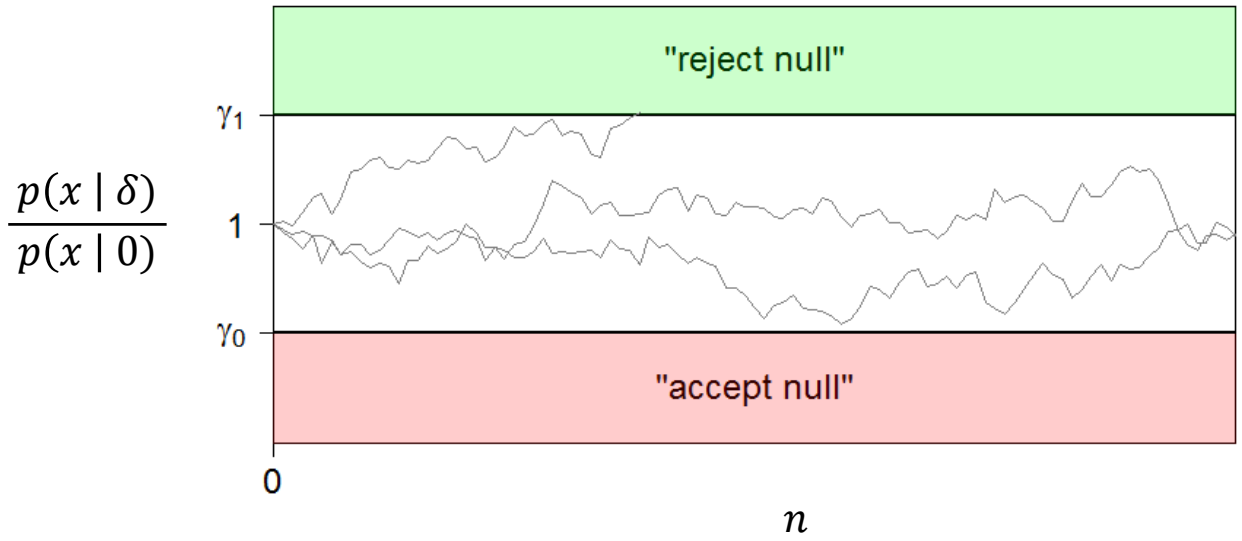
Group-sequential triangular test
(Whitehead & Stratton, 1983)



See Jennison & Turnbull (2000), p105.

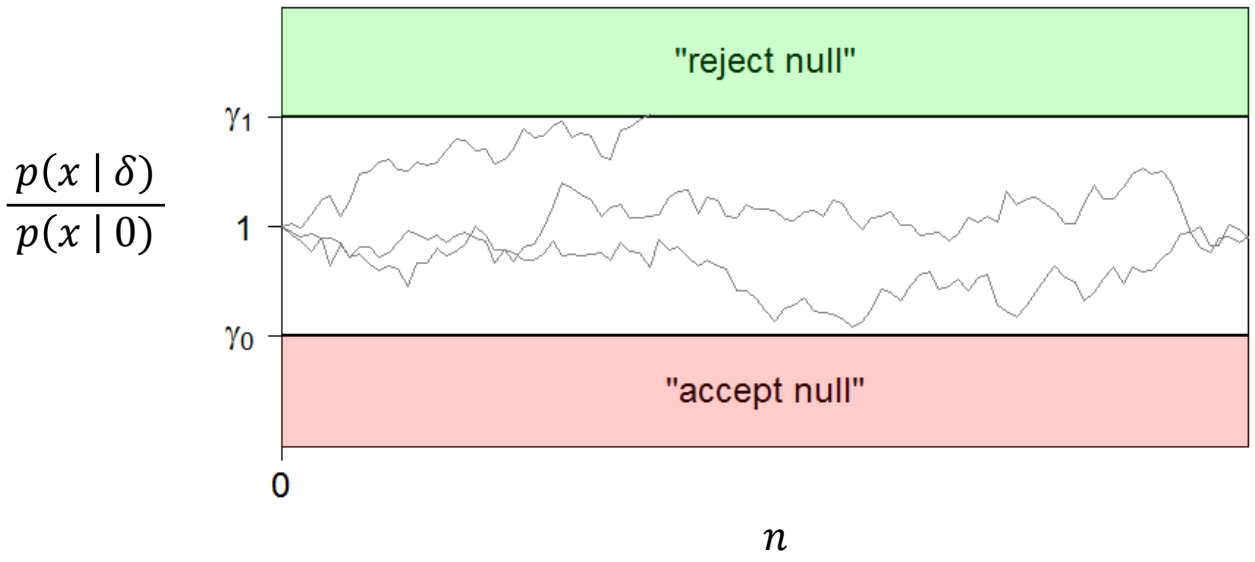
Sequential Probability Ratio Test (Wald, 1945)

- Data: x
- Model: p(x | θ)
- Null hypothesis: θ = 0
- Alternative: θ = δ



How to choose upper boundary

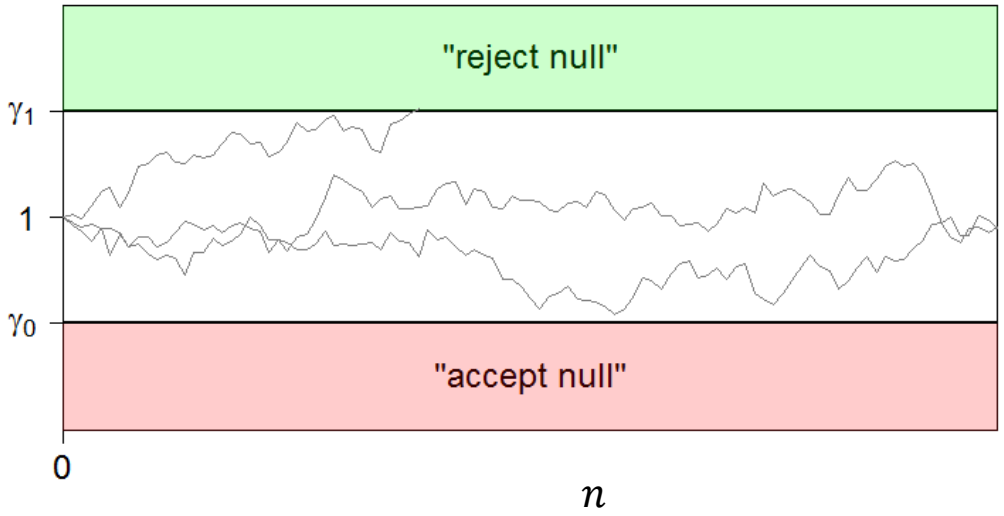
$$1-\beta = \int_{\text{x hits "reject"}} p\,(x \mid \delta)\,dx = \int_{\text{x hits "reject"}} \frac{p(x \mid \delta)}{p(x \mid 0)} p(x \mid 0)\,dx = \gamma_1 \int_{\text{x hits "reject"}} p\,(x \mid 0)\,dx = \gamma_1 \alpha$$



$$\gamma_1 = \frac{1-\beta}{\alpha}$$

How to choose lower boundary

$\frac{p(x \mid \delta)}{p(x \mid 0)}$



$$\gamma_0 = \frac{\beta}{1-\alpha}$$

$$1-\alpha = \int_{\text{x hits "accept"}} p\,(x \mid 0)\,dx = \int_{\text{x hits "accept"}} \frac{p(x \mid 0)}{p(x \mid \delta)} p(x \mid \delta)\,dx = \gamma_0^{-1} \int_{\text{x hits "accept"}} p\,(x \mid \delta)\,dx = \gamma_0^{-1}\beta$$

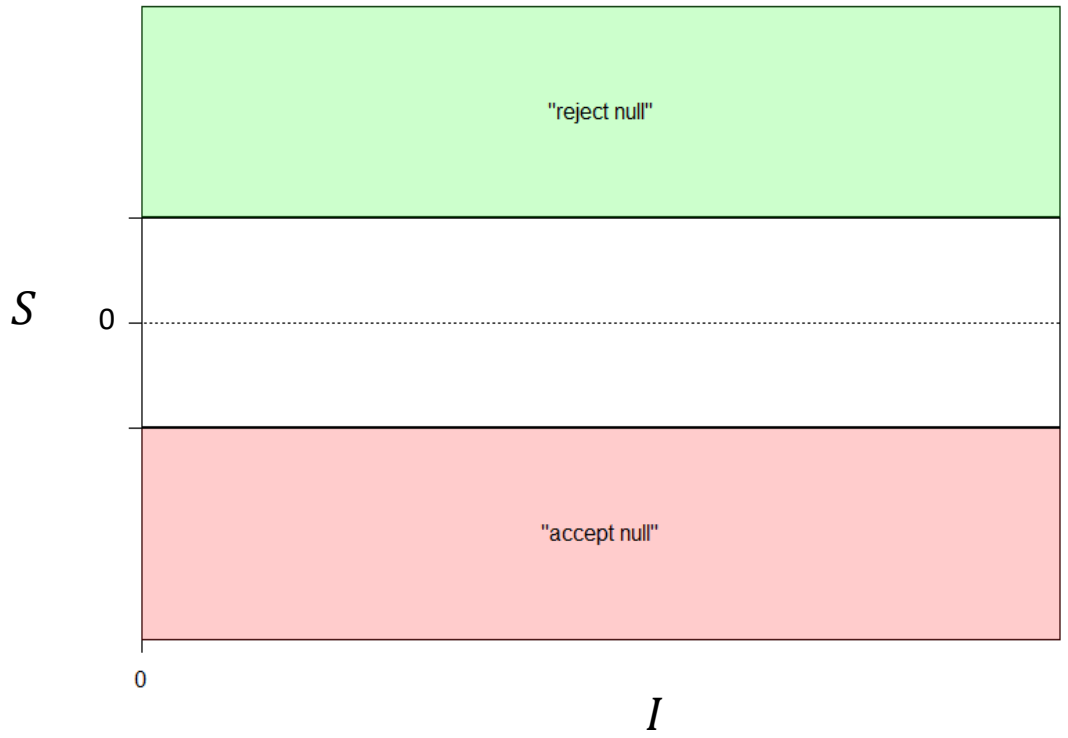
Symmetry & Gaussian likelihood

- Transform $\varphi = \theta - \frac{\delta}{2}$.
- Test H_0: $\varphi = -\frac{\delta}{2}$ vs H_1: $\varphi = \frac{\delta}{2}$.
- $\beta = \alpha$.
- Data x is summarized by a score statistic, S,

$$S \sim N(\varphi I, I).$$

- Log-likelihood ratio

$$\frac{p(S \mid \delta/2)}{p(S \mid -\delta/2)} = \delta \times S.$$



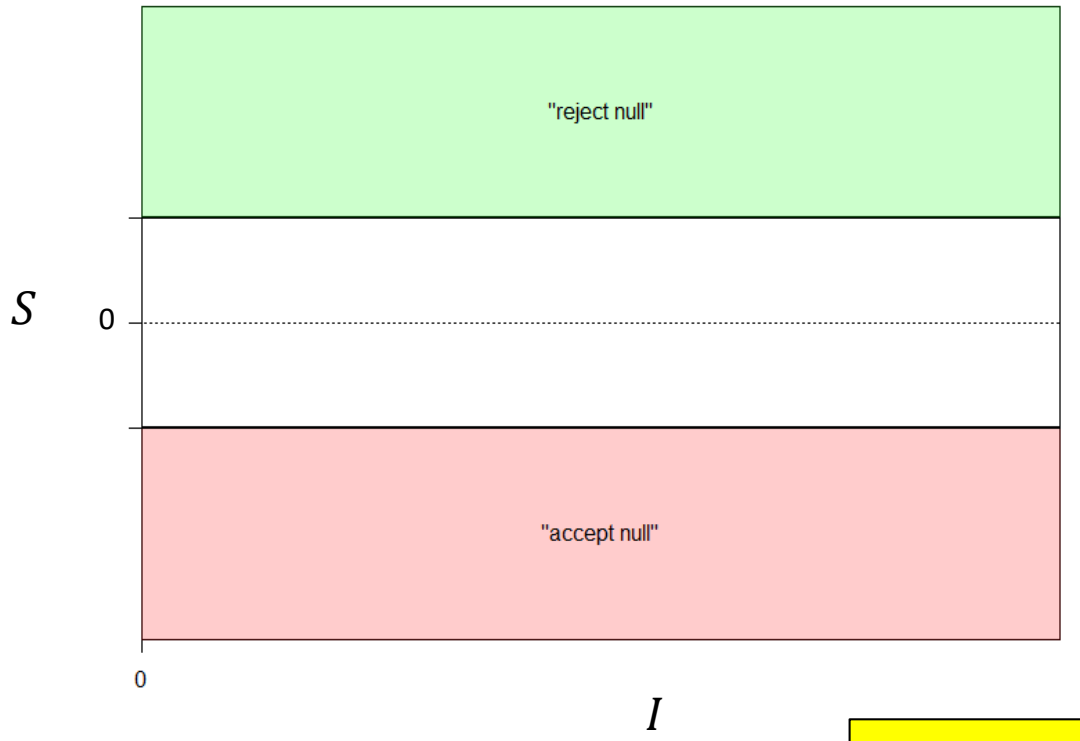
Symmetry & Gaussian likelihood

- Transform $\varphi = \theta - \frac{\delta}{2}$.
- Test $H_0: \varphi = -\frac{\delta}{2}$ vs $H_1: \varphi = \frac{\delta}{2}$.
- $\beta = \alpha$.
- Data x is summarized by a score statistic, S,

$$S \sim N(\varphi I, I).$$

- Log-likelihood ratio

$$\frac{p(S \mid \delta/2)}{p(S \mid -\delta/2)} = \delta \times S.$$

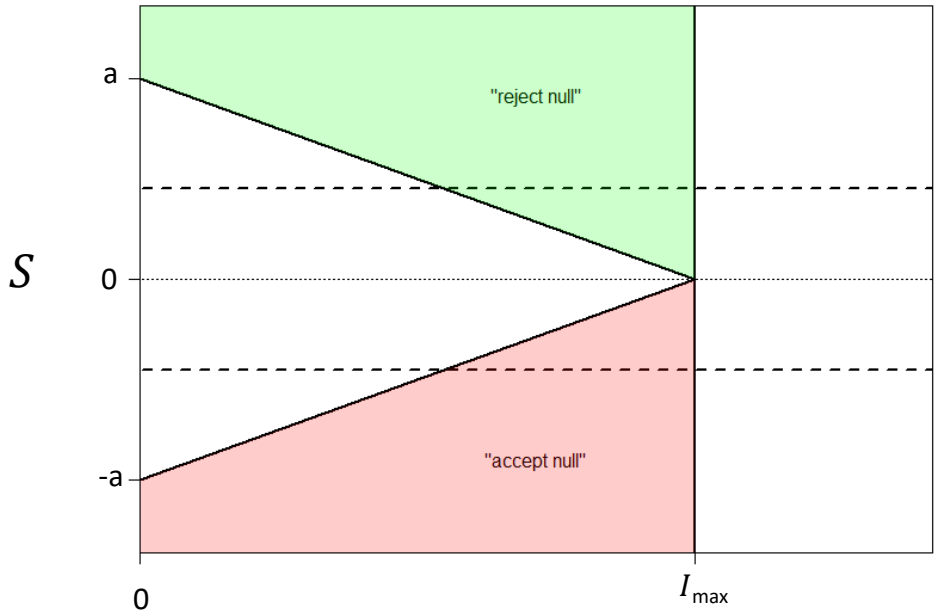


Large expected sample size when ϕ = 0

SPRT-2 (Anderson, 1960; Lorden, 1976)

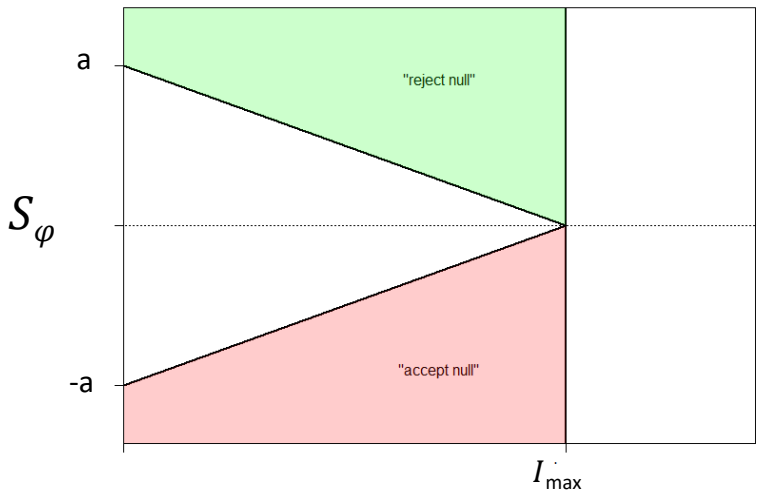
- Simplest possible modification to produce a design with a finite sample size.
- It turns out to have an exact solution:
  - $I_{max} = \frac{8}{\delta^2}\log\left(\frac{1}{2\alpha}\right)$
  - $a = \frac{2}{\delta}\log\left(\frac{1}{2\alpha}\right)$
- It turns out that it minimizes:

$\max_{\phi} E(\text{ sample size } | \phi)$.

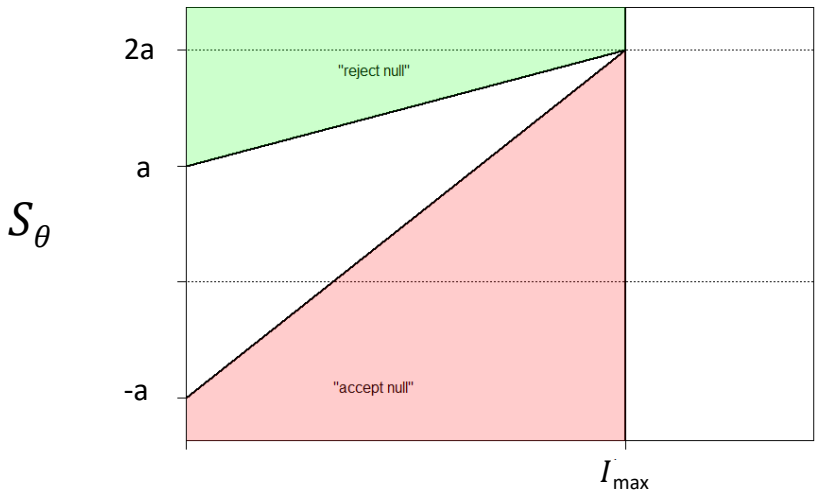


Transform back to θ parameterization

$$H_0:\ \varphi = -\frac{\delta}{2} \quad vs \quad H_1:\ \varphi = \frac{\delta}{2}$$

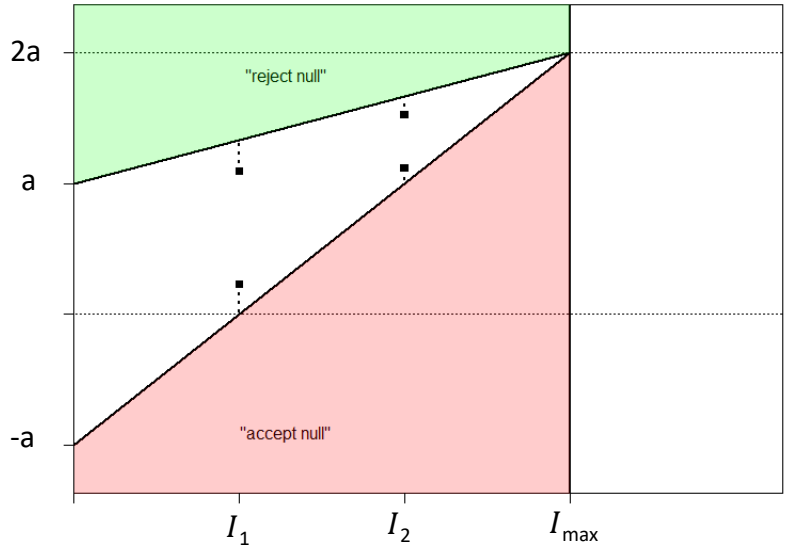


$$H_0:\ \theta = 0 \quad vs \quad H_1:\ \theta = \delta$$



Group-sequential triangular test (Whitehead & Stratton, 1983)

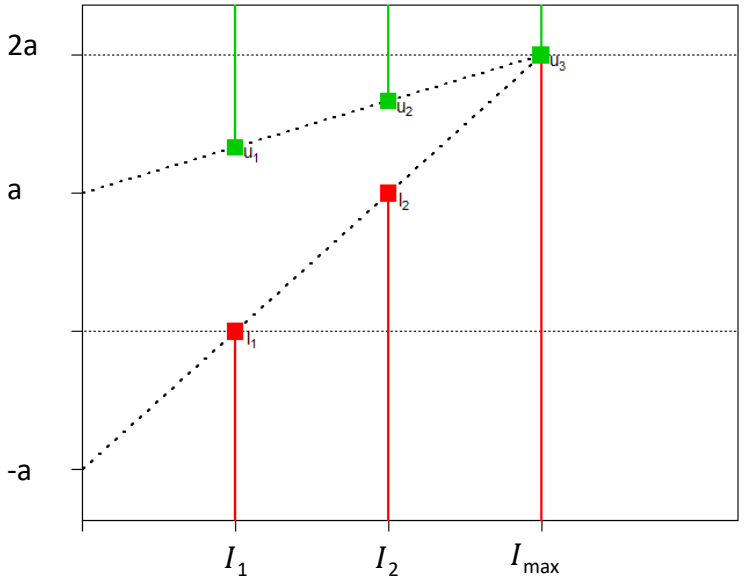
- SPRT-2 assumes continuous monitoring.
- If monitoring is discrete, we won't achieve the nominal type 1 error rate and power.
- Whitehead & Stratton (1983) proposed an adjustment.
- Still based on quite simple formulae. Doesn't require specialist software.



“Modern” group-sequential triangular test

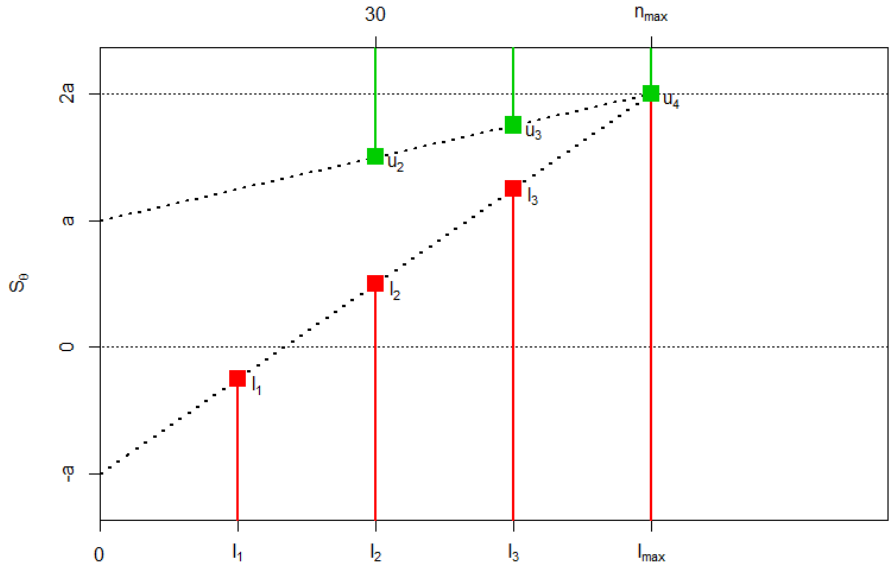
- Shape of stopping boundary is kept fixed.
- a and I_{max} are chosen such that type I error is α and power is 1 – β.
- Numerical integration using the joint distribution of the test statistics:

$$\begin{pmatrix} S_1 \\ S_2 - S_1 \\ \vdots \\ S_k - S_{k-1} \end{pmatrix} \sim N\left(\begin{pmatrix} \theta I_1 \\ \theta(I_2 - I_1) \\ \vdots \\ \theta(I_k - I_{k-1}) \end{pmatrix}, \begin{pmatrix} I_1 & 0 & \cdots & 0 \\ 0 & I_2 - I_1 & & \\ \vdots & & \ddots & \\ 0 & & & I_k - I_{k-1} \end{pmatrix} \right)$$



Fixing the first interim at n = 30

- Keep the shape fixed.
- Fix the first efficacy boundary at n = 30.
- Use numerical integration and root finding to solve for a and I_{max}.
- R code published online with our paper (Allison et al., 2015, *Trials* 16.1 522).



	(3 analyses)	(4 analyses)	(7 analyses)
Maximum sample size	62	62	66
Expected sample size under H_0	36	31	30
Expected sample size under H_1	40	40	39
Type I error	0.0479	0.0482	0.0484
Power	0.893	0.889	0.894

An equivalent fixed sample size design would require approximately n = 50.

Challenge for us

1. Could we suggest improvements to the trial design?

   A sequential design (triangular test) was used, but complicated by the need to collect pharmacokinetic data from a minimum of 30 patients.

2. **Could we suggest a better estimator for cure rate at day 210?**

   An ad-hoc “probability tree estimator” had been used, but properties unclear.

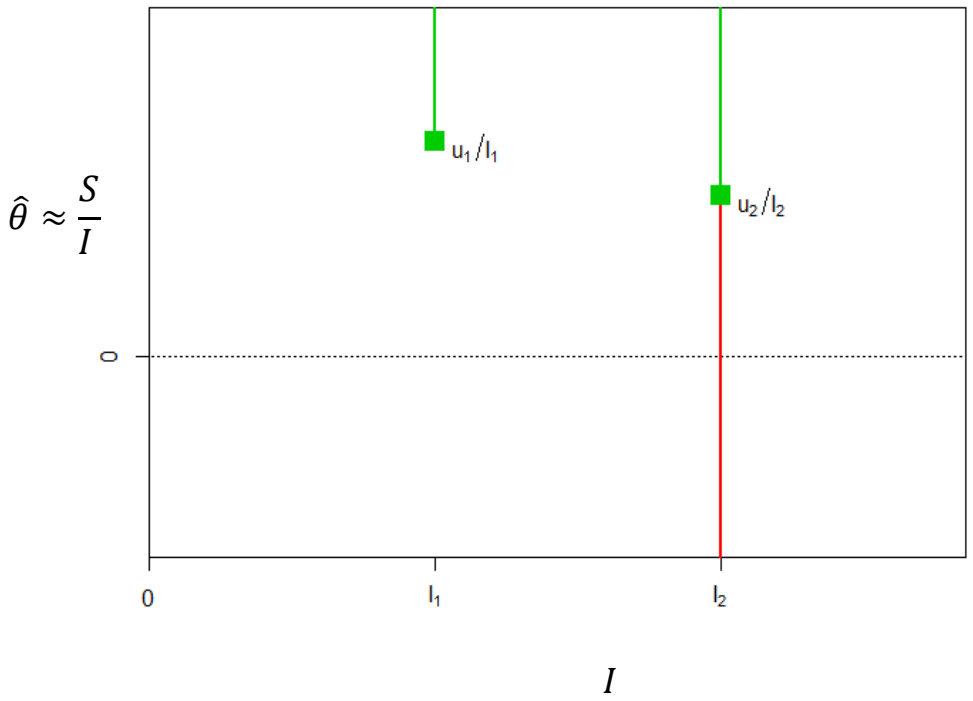
Bias of standard estimates under sequential analysis

- Let $\hat{\theta}^*$ denote the estimate of θ at the (random) stopping time.

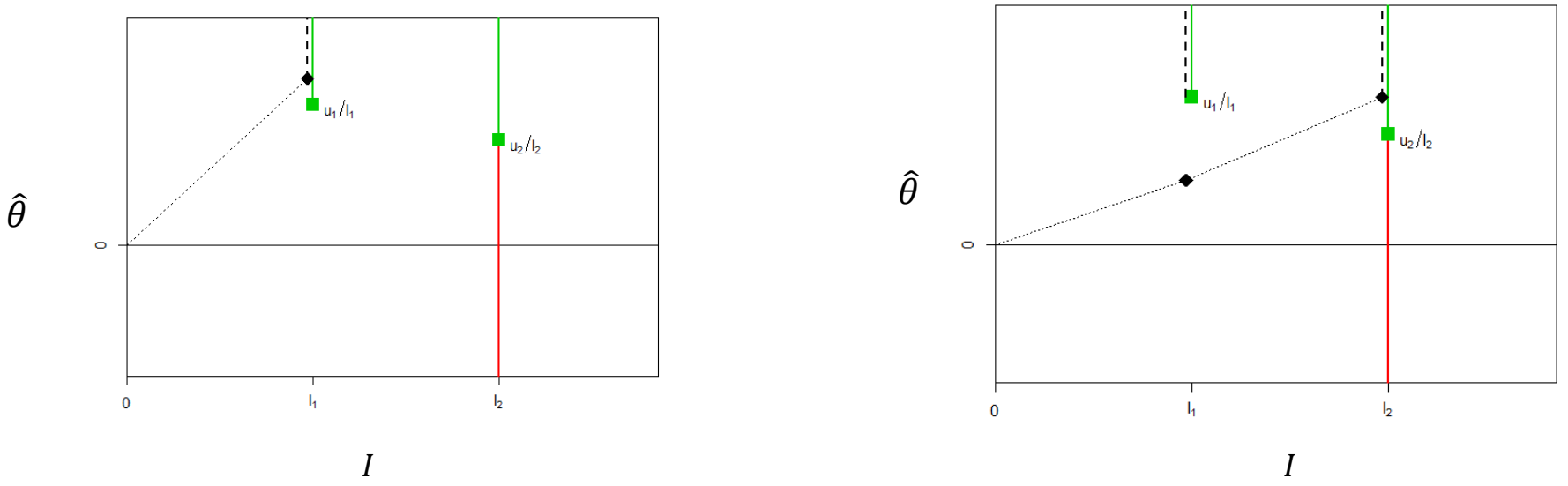
$$E(\hat{\Theta}^* \mid \hat{\Theta}_1 = \hat{\theta}_1) = \begin{cases} \hat{\theta}_1, & \hat{\theta}_1 > \frac{u_1}{I_1} \\ \hat{\theta}_1 \frac{I_1}{I_2} + \theta\left(1 - \frac{I_1}{I_2}\right), & \hat{\theta}_1 < \frac{u_1}{I_1} \end{cases}$$

- Random highs, stay high.
- Random lows, regression to the mean.
- Overall,

$$E(\hat{\Theta}*) = \int E(\hat{\Theta}* \mid \hat{\Theta}_1 = \hat{\theta}_1) f(\hat{\theta}_1)\, d\hat{\theta}_1 > \theta.$$



Median unbiased estimation



- P-value function: $pr(\text{x}, \ \theta) = pr(\text{“more extreme data than x”} \mid \theta)$
- Median-unbiased estimate: $\tilde{\theta}$ such that $pr(\text{x}, \ \tilde{\theta}) = 0.5$

Probability tree estimator (Omollo et al, 2011)

- The MLE $\hat{p}_{28}$ is a biased estimate of $p_{28.}$
- The MLE $\hat{p}_{210}$ will be highly correlated with $\hat{p}_{28}$, therefore also biased.
- Attempt to correct for bias using

$$\tilde{p}_{210} = r \times \tilde{p}_{28} + s \times (1 - \tilde{p}_{28})$$

where:

- $\tilde{p}_{28}$ is a median-unbiased estimate of p_{28}
- r is the proportion of patients with cure at day 210, out of those cured at day 28.
- s is the proportion of patients with cure at day 210, out of those not cured at day 28.

Treatment selection bias

We're likely to focus on the best-performing treatment.

- Policy recommendations.
- Planning future studies.

Without adjustment, this will systematically overestimate treatment effect:

$$E\{\max(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)\} > \mathrm{E}(\hat{\theta}_j) \text{ for j} = 1,2,3.$$

Bayesian shrinkage estimation

Idea

$\hat{\theta}_j^s = w \times \hat{\theta}_j + (1 - w) \times \bar{\theta}$

where:

- $\bar{\theta}$ is an overall mean.
- w is a data-dependent weight.

Full details

$$
\begin{array}{lll}
P(Y_{i,j} = 1) & = p_{210,j}, & i = 1, \ldots, n_j; j = 1, 2, 3, \\
p_{210,j} & = \Phi(\theta_j), & j = 1, 2, 3, \\
\theta_j | \mu, \tau^2 & \sim \mathcal{N}(\mu, \tau^2), & j = 1, 2, 3, \\
\mu & \propto 1, & \\
\tau^2 & \sim \mathcal{IG}(\alpha, \beta), & \\
\alpha & = 2, & \\
\beta & = 0.3, &
\end{array}
$$

Bias and RMSE

Scenario 1

$p_{28,1} = p_{28,2} = p_{28,3} = 0.75$

Scenario 3

$p_{28,1} = p_{28,2} = 0.75$
$p_{28,3} = 0.9$

Cases:

1. No relapses, no slow responders.
2. No relapses, 33 % slow responders.
3. 25 % relapses, no slow responders.
4. 25 % relapses, 33 % slow responders.

Summary

- The triangular design has good properties. It will (almost) minimize the maximum expected sample size across all parameter values.
- We can adapt it to match practical requirements of studies.
- Shrinkage estimation is an attractive approach to reduce bias and mean-squared-error, arising from sequential analysis and treatment selection.
- R package “gentri” is available online with the paper. https://doi.org/10.1186/s13063-015-1018-1

References

Allison, Annabel, et al. "Generalizing boundaries for triangular designs, and efficacy estimation at extended follow-ups." *Trials* 16.1 (2015): 522.

Omollo, Raymond, et al. "Safety and efficacy of miltefosine alone and in combination with sodium stibogluconate and liposomal amphotericin B for the treatment of primary visceral leishmaniasis in East Africa: study protocol for a randomized controlled trial." *Trials* 12.1 (2011): 166.

Wald, Abraham. "Sequential tests of statistical hypotheses." *The annals of mathematical statistics* 16.2 (1945): 117-186.

Anderson, Theodore W. "A modification of the sequential probability ratio test to reduce the sample size." *The Annals of Mathematical Statistics* (1960): 165-197.

Lorden, Gary. "2-SPRT's and the modified Kiefer-Weiss problem of minimizing an expected sample size." *The Annals of Statistics* (1976): 281-291.

Whitehead, John, and Irene Stratton. "Group sequential clinical trials with triangular continuation regions." *Biometrics*(1983): 227-236.

Jennison, Christopher, and Bruce W. Turnbull. *Group sequential methods with applications to clinical trials*. Chapman and Hall/CRC, 1999