Duncan T. Wilson



Clinical Trials Research Unit Leeds Institute of Clinical Trials Research University of Leeds

d.t.wilson@leeds.ac.uk

18th July 2017

4日 → 4団 → 4 三 → 4 三 → 9 Q ○

Duncan T. Wilson

MDG: Bayesian trial design

Background
Bayesian methods for trial design

Bayesian designs for CI and pilot trials References

Previously...



- Last time, we talked about using the limited data provided by a pilot trial, in conjunction with prior knowledge, to learn about the ICC and use this knowledge to better design the main trial.
- How do we design trials to test hypotheses, using full distributions describing unknown parameters as opposed to point estimates or hypothesised values?
- Only after this has been described can we think about optimal pilot trial design.

- Background
- 2 Bayesian methods for trial design
- Bayesian designs for CI and pilot trials

Bayesian methods for trial design

- 4 ロ ト 4 間 ト 4 重 ト 4 重 ト 9 Q (P

Duncan T. Wilson

MDG: Bayesian trial design

Background

Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Motivation



Power for a trial with clustering:

$$1 - \beta = \Phi\left(\sqrt{\frac{n\delta^2}{2\sigma^2[1 + (m-1)\rho]}} + z_{\alpha/2}\right),\,$$

n = number of patients per arm

m = number of patients per cluster

 $\delta = \text{treatment effect}$

$$\sigma_B^2 + \sigma_W^2 = \sigma^2 = \text{total variance}$$

$$\frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} = \rho = \text{intracluster correlation coefficient (ICC)}$$

Motivation



Power is defined with respect to some hypothesis H_1 ,

Power =
$$Pr[\text{reject } H_0 \mid H_1],$$

where H_1 specifies the values of all the parameters in our model:

$$H_1: \delta = \delta^*, \sigma = \sigma^*, \rho = \rho^*.$$

We never know the true value of ρ . How do we choose ρ^* , and what are the implications of our choice?



Duncan T. Wilson

MDG: Bayesian trial design

Background

Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Motivation



For a fixed target power, a higher ICC will lead to a higher sample size requirement:

$$n = n_i[1 + (m-1)\rho^*]$$

 n_i = sample size required when clustering is ignored

So, required sample size increases linearly with our chosen ρ^* .

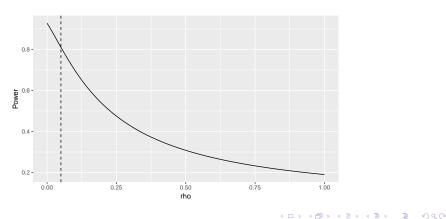
Background

Bayesian methods for trial design Bayesian designs for CI and pilot trials

Motivation



For a fixed sample size n, a higher than expected ICC will lead to a lower than expected power:



Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials

Practical solution



Example (SHIFT)

"We anticipate that the level of clustering will be low for this particular trial - possibly around 0.01 but no higher than 0.05 \dots assuming an ICC of 0.05 effectively reduces the power from 90% to around 75%. If the ICC were as low as 0.01, then this reduces the power to around 85%."

Background

Bayesian methods for trial design Bayesian designs for CI and pilot trials

Practical solution



Trial design now considers another hypothesis:

Minimise sample size n

Minimise type I error rate $\mid H_0$

Maximise power $\mid H_1 : \rho = 0.01$

Maximise power $\mid H_2 : \rho = 0.05$

We down-weight the importance of H_2 compared to H_1 because we believe it to be less likely, but we do this informally.



Duncan T. Wilson

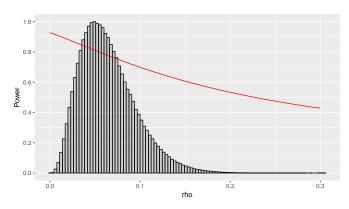
MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Bayesian alternative



An alternative: express our belief about the true value of ρ using a probability distribution, e.g. $\rho \sim Beta(4,58)$:

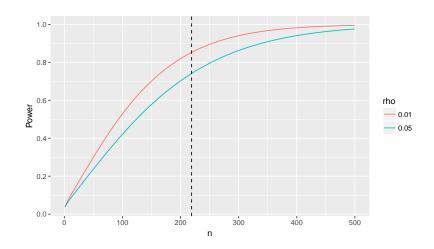


MDG: Bayesian trial design

Background
Bayesian methods for trial design
Bayesian designs for CI and pilot trials
References

Practical solution





Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials

Bayesian alternative



4□ > 4個 > 4 = > 4 = > = 900

Should we use probability to express our uncertainty? Axioms¹:

- The relation \succeq , where $A \succeq B$ means A is at least as likely as B, is a weak ordering (i.e. complete and transitive) on S.
- ② If A, B and D are such that $AD = BD = \emptyset$, then $A \succeq B \iff A \cup D \succeq B \cup D$.
- **3** $A \succeq \emptyset$ for any A. Furthermore, $S \succ \emptyset$.
- **4** If $A_1 \supset A_2 \supset ...$ is a decreasing sequence of events and B is such that $A_i \succeq B$ for all i, then $\bigcap_{i=1}^{\infty} A_i \succeq B$.
- **1** There exists a random variable $X \sim Unif(0,1)$.

Together, these imply that there is a unique probability distribution P that agrees with the relation \succeq .

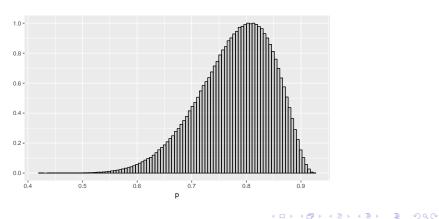


¹DeGroot 1970.

Bayesian alternative



Uncertainty about $\rho \sim Beta(4,58)$ propagates to uncertainty about power:



Duncan T. Wilson

MDG: Bayesian trial design

Bayesian methods for trial design

Assurance



Key paper: O'Hagan, Stevens and Campbell (2005), "Assurance in clinical trial design", Pharmaceutical statistics.

- If event A denotes a 'successful' trial, and this is influenced by the true value of parameter(s) θ , what is Pr[A]?
- Unconditional, so we need to integrate out any parameters.

assurance =
$$Pr[A] = \int Pr[A \mid \theta]p(\theta)d\theta$$

 $p(\theta) = \text{probability density function of } \theta$

Bayesian methods for trial design Bayesian designs for CI and pilot trials

- Background
- 2 Bayesian methods for trial design
- Bayesian designs for CI and pilot trials



Duncan T. Wilson

MDG: Bayesian trial design

Bayesian methods for trial design

Assurance



Example (Unknown treatment effect)

For a two-sample comparison of normal means, suppose we are uncertain what the true treatment difference δ is. We think $\delta \sim N(0.2, 0.25^2)$, and know that the common variance is $\sigma^2 = 1$. Our trial has n = 500 in each arm, and will test at a level of $\alpha = 0.025$ (one-sided). What is $Pr[\text{reject } H_0]$?

- **1** Sample $\delta^i \sim N(0.2, 0.25^2)$ for i = 1, ..., N
- **2** Calculate powers $\beta(\delta^i)$
- Ompute the Monte Carlo estimate:

$$Pr[\text{reject } H_0] \approx \frac{1}{N} \sum_{i=1}^{N} \beta(\delta^i) = 0.62.$$
 (1)

Assurance



Example (Unknown treatment effect)

For a two-sample comparison of normal means, suppose we are uncertain what the true treatment difference δ is. We think $\delta \sim N(0.2, 0.25^2)$, and know that the common variance is $\sigma^2 = 1$. Our trial has n = 500 in each arm, and will test at a level of $\alpha = 0.025$ (one-sided). What is $Pr[\text{reject } H_0]$?

- Compare the unconditional power of 0.62 with the conditional power at $\delta^* = 0.2$ of 0.88.
- Is this good enough? Should we adjust the sample size?

◆ロト→御ト→きト→き ● り9℃

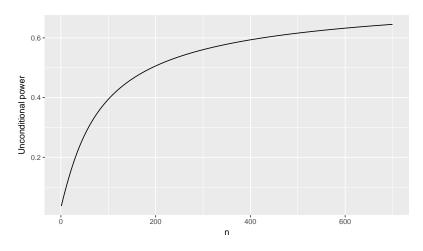
Duncan T. Wilson

MDG: Bayesian trial design

Background **Bayesian methods for trial design** Bayesian designs for Cl and pilot trials References

Optimality





Duncan T. Wilson



MDG: Bavesian trial design

Background
Bayesian methods for trial design
Bayesian designs for CI and pilot trials
References

Optimality



Given some measure(s) of the quality of a proposed trial design, how do we choose the best?

- Try to set some kind of standard threshold analogous to a power of 80%.
- Avoid thresholds and consider trade-offs²;
- Maximise performance per-patient³;
- Maximise expected utility⁴, possibly through an economic model⁵;

<□ > <**□** > < **□** > < **□** > < **□** > < **□** > < **□** > < **□** < **○** < ○

Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Assurances



The unconditional probability of rejecting the null hypothesis was the first hybrid Bayesian metric to be considered, but others have been proposed:

- Prob. rejecting the null and the true effect being meaningful⁶;
- Prob. rejecting the null in *two* separate phase III trials⁷;
- Prob. of a positive result in an updated meta-analysis⁸.

'Fully Bayesian' criteria have also been proposed:

- Prob. posterior probability the treatment effect is meaningful⁹;
- Conditional distribution of the treatment effect given a significant result¹⁰.



²Bacchetti, McCulloch, and Segal 2008.

³Stallard 2012.

⁴Lindley 1997

⁵Patel et al. 2013.

⁶O'Hagan, Stevens, and Campbell 2005

⁷Zhang and Zhang 2013.

⁸Sutton et al. 2007

⁹Ibrahim et al. 2014.

¹⁰Walley et al. 2015.

Uncertainty



Various papers have addressed uncertainty in ...

- Treatment effects only 11;
- Nuisance parameters only, including ICCs¹² and parameters in survival models¹³:
- Samples size¹⁴;
- Adherance rates¹⁵.

For trials of complex interventions following a pilot we might also expect uncertainty in recruitment and data collection rates 16 .

Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Computation



- Some papers focus on conjugacy²²;
- The more flexible approach is through simulation²³, sampling test statistics if possible or failing that, individual patient data;
- Bayesian (unconditional) simulation is no trickier than frequentist (conditional) simulation;

Duncan T. Wilson

• But simulation is time-consuming - if the design problem is complex, we need efficient optimisation methods.

MDG: Bavesian trial design



Bayesian methods for trial design Bayesian designs for CI and pilot trials

Priors



How do we define the priors to be used in these calculations?

- Using early phase trial data (empirical Bayes)¹⁷;
- Using other historical data, perhaps from a meta-analysis 18;
- Using expert judgement¹⁹;
- Using defaults, e.g. pessimistic and optimistic priors²⁰;
- If using a fully Bayesian approach, consider differentiating between 'design' and 'analysis' priors²¹;

Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design ayesian designs for CI and pilot trials References

Interpretation



- Can we have distributions for nuisance parameters but condition on treatment effects?²⁴
- Can we have multiple priors?²⁵
- Can we interpret a distribution of treatment effects from a frequentist perspective?²⁶



¹¹O'Hagan, Stevens, and Campbell 2005.

¹²Turner, Toby Prevost, and Thompson 2004.

¹³Ren and Oakley 2013.

¹⁴Ambrosius and Mahnken 2010.

¹⁵Fav, Halloran, and Follmann 2006

 $^{^{16}\}mbox{Avery et al. }2017.$

²²Ibrahim et al. 2014.

²³Wang et al. 2013.

¹⁷ Jiang 2011.

¹⁸Sutton et al. 2007

¹⁹Ren and Oakley 2013.

²⁰Spiegelhalter, Abrams, and Myles 2004.

²¹Wang and Gelfand 2002.

²⁴Turner, Toby Prevost, and Thompson 2004.

²⁵Chen et al. 2011; Kirby et al. 2012.

²⁶Gtte et al. 2015; Julious and Owen 2006

Challenges



- Background
- 2 Bayesian methods for trial design
- 3 Bayesian designs for CI and pilot trials

4日 > 4個 > 4 国 > 4 国 > ■ 9 9 0

Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Challenges



- Complex models, e.g. recruitment, **missing data**, adherance, multilevel outcomes, multivariate outcomes³⁰
- Complex design problem, e.g. multilevel sample sizes, multivariate acceptance regions, choice of primary endpoint
- Qualitative gap between pilot and main trial settings, including changes to the intervention itself³¹
- Computational burden of searching for optimal pilot designs³²

MDG: Bavesian trial design

- Complex models, e.g. recruitment, missing data, adherance, multilevel outcomes, multivariate outcomes²⁷
- Complex design problem, e.g. multilevel sample sizes, multivariate acceptance regions, choice of primary endpoint
- Qualitative gap between pilot and main trial settings, including changes to the intervention itself²⁸
- Computational burden of searching for optimal pilot designs²⁹

Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Example



Example (REACH(ish))

One of the goals of out pilot study was to assess two potential primary outcomes for the main trial. We weren't sure about what treatment effects we might see, the variability in the outcomes, or how complete the data are likely to be. Following the pilot, a Bayesian analysis gives us posterior distributions on the relevant parameters.

Endpoint	$\hat{\delta}$	$\hat{\sigma}^2$	\hat{p}_{miss}
Α	0.145	1.16	0.75
В	0.18	1.08	8.0



³⁰ Landau and Stahl 2013.

³¹ Hampson et al. 2017.

³²Strong2015.

²⁷Landau and Stahl 2013.

²⁸Hampson et al. 2017.

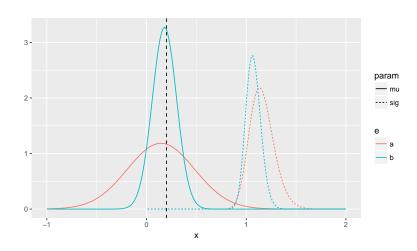
²⁹Strong2015.

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Example

0.00 -

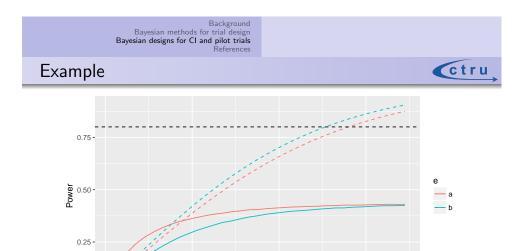




Duncan T. Wilson

MDG: Bayesian trial design

◆ロト→御ト→きト→き ● り9℃



 $510 \rightarrow 370$ patients per arm.

600

Duncan T. Wilson MDG: Bayesian trial design

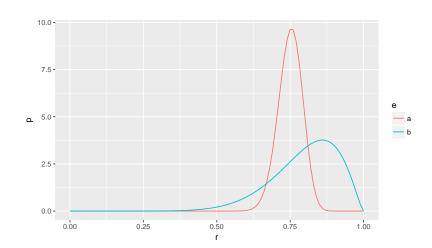
400

n

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References

Example





Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References



Ambrosius, Walter T, and Jonathan D. Mahnken (2010). "Power for studies with random group sizes". In:

Statistics in Medicine, n/a-n/a. DOI: 10.1002/sim.3873. URL: https://doi.org/10.1002%Fsim.3873



Avery, Kerry N L et al. (2017). "Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies". In: BMJ Open 7.2, e013537. DOI: 10.1136/bmjopen-2016-013537. URL: https://doi.org/10.1136%2Fbmjopen-2016-013537.



Bacchetti, Peter, Charles E. McCulloch, and Mark R. Segal (2008). "Simple, Defensible Sample Sizes Based on Cost Efficiency". In: Biometrics 64.2, pp. 577–585. DOI: 10.1111/j.1541-0420.2008.01004_1.x. URL: http://dx.doi.org/10.1111/j.1541-0420.2008.01004_1.x.



Chen, Ming-Hui et al. (2011). "Bayesian Design of Noninferiority Trials for Medical Devices Using Historical Data". In: Biometrics 67.3, pp. 1163–1170. DOI: 10.1111/j.1541-0420.2011.01561.x. URL: https://doi.org/10.1111%2Fj.1541-0420.2011.01561.x.



DeGroot, Morris H. (1970). Optimal Statistical Decisions.



Fay, Michael P., M. Elizabeth Halloran, and Dean A. Follmann (2006). "Accounting for Variability in Sample Size Estimation with Applications to Nonadherence and Estimation of Variance and Effect Size". In: Biometrics 63.2, pp. 465–474. DOI: 10.1111/j.1541-0420.2006.00703.x. URL: https://doi.org/10.1111%2Fj.1541-0420.2006.00703.x.



Gtte, Heiko et al. (2015). "Sample size planning for phase II trials based on success probabilities for phase III". In:
Pharmaceutical Statistics 14.6, pp. 515–524. DOI: 10.1002/pst.1717. URL:
https://doi.org/10.1002%2Fpst.1717.

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References



Hampson, Lisa V et al. (2017). "A framework for prospectively defining progression rules for internal pilot studies monitoring recruitment". In: Statistical Methods in Medical Research 0.0. PMID: 28589752, p. 0962280217708906. DOI: 10.1177/0962280217708906. eprint: http://dx.doi.org/10.1177/0962280217708906. URL: http://dx.doi.org/10.1177/0962280217708906.



Ibrahim, Joseph G. et al. (2014). "Bayesian probability of success for clinical trials using historical data". In: Statistics in Medicine 34.2, pp. 249–264. DOI: 10.1002/sim.6339. URL: https://doi.org/10.1002%2Fsim.6339.



Jiang, Kaihong (2011). "Optimal Sample Sizes and Go/No-Go Decisions for Phase II/III Development Programs Based on Probability of Success". In: Statistics in Biopharmaceutical Research 3.3, pp. 463–475. DOI: 10.1198/sbr.2011.10068. URL: https://doi.org/10.1198/2Fsbr.2011.10068.



Julious, Steven A. and Roger J. Owen (2006). "Sample size calculations for clinical studies allowing for uncertainty about the variance". In: *Pharmaceutical Statistics* 5.1, pp. 29–37. DOI: 10.1002/pst.197. URL: http://dx.doi.org/10.1002/pst.197.



Kirby, S et al. (2012). "Discounting phase 2 results when planning phase 3 clinical trials". In: Pharmaceutical statistics 11.5, pp. 373–385.



Landau, Sabine and Daniel Stahl (2013). "Sample size and power calculations for medical studies by simulation when closed form expressions are not available". In: Statistical Methods in Medical Research 22.3, pp. 324–345. DOI: 10.1177/0962280212439578. eprint: http://smm.sagepub.com/content/22/3/324.full.pdf+html. URL: http://smm.sagepub.com/content/22/3/324.abstract.



Lindley, Dennis V. (1997). "The choice of sample size". In: Journal of the Royal Statistical Society: Series D (The Statistician) 46.2, pp. 129–138. ISSN: 1467-9884. DOI: 10.1111/1467-9884.00068. URL: http://dx.doi.org/10.1111/1467-9884.00068



Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References



Walley, Rosalind J. et al. (2015). "Advantages of a wholly Bayesian approach to assessing efficacy in early drug development: a case study". In: Pharmaceutical Statistics 14.3, pp. 205–215. DOI: 10.1002/pst.1675. URL: http://dx.doi.org/10.1002/pst.1675.



Wang, Fei and Alan E. Gelfand (2002). "A Simulation-Based Approach to Bayesian Sample Size Determination for Performance under a Given Model and for Separating Models". English. In: Statistical Science 17.2, pp. 193–208. ISSN: 08834237. URL: http://www.jstor.org/stable/3182824.



Wang, Yanping et al. (2013). "Evaluating and utilizing probability of study success in clinical development". In: Clinical Trials 10.3, pp. 407–413. DOI: 10.1177/1740774513478229. URL: https://doi.org/10.1177%2F1740774513478229.



Zhang, Jianliang and Jenny J. Zhang (2013). "Joint probability of statistical success of multiple phase III trials". In: Pharmaceutical Statistics 12.6, pp. 358–365. DOI: 10.1002/pst.1597. URL: https://doi.org/10.1002%2Fpst.1597.



Duncan T. Wilson

MDG: Bayesian trial design

Background Bayesian methods for trial design Bayesian designs for CI and pilot trials References



O'Hagan, Anthony, John W. Stevens, and Michael J. Campbell (2005). "Assurance in clinical trial design". In: Pharmaceutical Statistics 4.3, pp. 187–201. ISSN: 1539-1612. DOI: 10.1002/pst.175. URL: http://dx.doi.org/10.1002/pst.175.



Patel, Nitin R et al. (2013). "A mathematical model for maximizing the value of phase 3 drug development portfolios incorporating budget constraints and risk". In: Statistics in medicine 32.10, pp. 1763–1777. DOI: 10.1002/sim.5731



Ren, Shijie and Jeremy E. Oakley (2013). "Assurance calculations for planning clinical trials with time-to-event outcomes". In: Statistics in Medicine 33.1, pp. 31–45. DOI: 10.1002/sim.5916. URL: https://doi.org/10.1002%2Fsim.5916.



Spiegelhalter, David J., Keith R. Abrams, and Jonathan P. Myles (2004). Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Wiley.



Stallard, Nigel (2012). "Optimal sample sizes for phase II clinical trials and pilot studies". In: Statistics in Medicine 31.11-12, pp. 1031-1042. ISSN: 1097-0258. DOI: 10.1002/sim.4357. URL: http://dx.doi.org/10.1002/sim.4357.



Sutton, Alexander J. et al. (2007). "Evidence-based sample size calculations based upon updated meta-analysis".

In: Statistics in Medicine 26.12, pp. 2479–2500. DOI: 10.1002/sim.2704. URL: https://doi.org/10.1002%2Fsim.2704.



Turner, Rebecca M., A. Toby Prevost, and Simon G. Thompson (2004). "Allowing for imprecision of the intracluster correlation coefficient in the design of cluster randomized trials". In: Statistics in Medicine 23.8, pp. 1195–1214. ISSN: 1097-0258. DOI: 10.1002/sim.1721. URL: http://dx.doi.org/10.1002/sim.1721.



Duncan T. Wilson

MDG: Bayesian trial design