MDG: Estimating ICCs in pilot trials

Duncan T. Wilson

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

d.t.wilson@leeds.ac.uk

25th October 2016

Overview

Suppose we want to design (i.e. choose sample size) of a phase III cluster randomised trial. We have the usual formula

$$n = 2\frac{\sigma^2}{\delta^2}(z_{1-\beta} - z_{\alpha.2})^2[1 + (m-1)\rho], \tag{1}$$

where ρ is the ICC, m is the cluster size, and $[1 + (m-1)\rho]$ is known as the *design effect*. Equivalently, the power of a trial is given by

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

$$1 - \beta = \Phi\left(\sqrt{\frac{n\delta^2}{2\sigma^2[1 + (m-1)\rho]}} + z_{\alpha/2}\right) \tag{3}$$

- Suppose we know σ^2, the total variance. How sensitive is power to the true value of the ICC ρ?
- Going further: given a distribution for ρ, what will the distribution of power look like?

David J. Spiegelhalter (2001)
Bayesian methods for cluster randomized trials with continuous responses
Statistics in Medicine, 20, 435-452.

- So, important to have a good estimate of ρ to design the trial.
- How to get a good estimate?
- Through an external pilot?
- E.g. TIGA-CUB, 'an investigation of the clustering effect (ICC) relating to the Child Psychotherapists will be carried out to aid sample size calculation'.

Sandra M. Eldridge et al. (2015)
How big should the pilot study for my cluster randomised trial be?
Statistical Methods in Medical Research, 25, 1039-1056.

- Conclusion: pilot studies aren't big enough to generate reliable ICC estimates.
- Instead, we should build estimates from a number of other sources, e.g. other trials, observational studies, papers publishing big lists of ICC estimates, ...
- Makes sense - but how? Meta analyses, evidence synthesis?
- However we do it, is it going to be much more work to define a prior distribution rather than a point estimate?
- If we have a prior, should we still ignore pilot data?
- If we don't want to ignore pilot data, how should we choose its sample size?

Suppose we have an (informative) prior distribution $p(\rho)$ for our ICC, e.g. $\rho \sim Beta(1.5, 10)$.

Suppose we have an (informative) prior distribution $p(\rho)$ for our ICC, e.g. $\rho \sim Beta(1.5, 10)$. If we conduct our pilot trial and observe x, we can get the posterior distribution for ρ using Bayes' rule:

$$p(\rho \mid x) = \frac{p(x \mid \rho)p(\rho)}{p(x)}. \tag{4}$$

Now, how do we use this to choose the sample size of the main trial? We can follow Spiegelhalter (and others) and ask that the trial is going to be powered (say, at 90%) with at least a certain probability (say, 0.8).

So, after the pilot trial gives us x, we choose the smallest k such that

$$\int_0^1 f(\rho, k) p(\rho \mid x) d\rho \geq 0.8, \tag{5}$$

where $f(\rho, k)$ is an indicator function which returns 1 if a trial of k clusters and an ICC of ρ is powered, and 0 otherwise.

Before the pilot trial, x is unknown. We want to derive the joint predictive distribution of the main trial's sample size and its power. We can generate samples:

1. Sample x_i, ρ_i from $p(x, \rho) = p(x \mid \rho)p(\rho)$
2. Use x_i to get the associated main trial sample size k_i as above
3. Calculate the power of the main trial, $1 - \beta_i$ using the earlier formula and the sampled ρ_i
4. Do this to generate lots of $k_i, 1 - \beta_i$ pairs

Compare with two other ways to choose the main trial k:

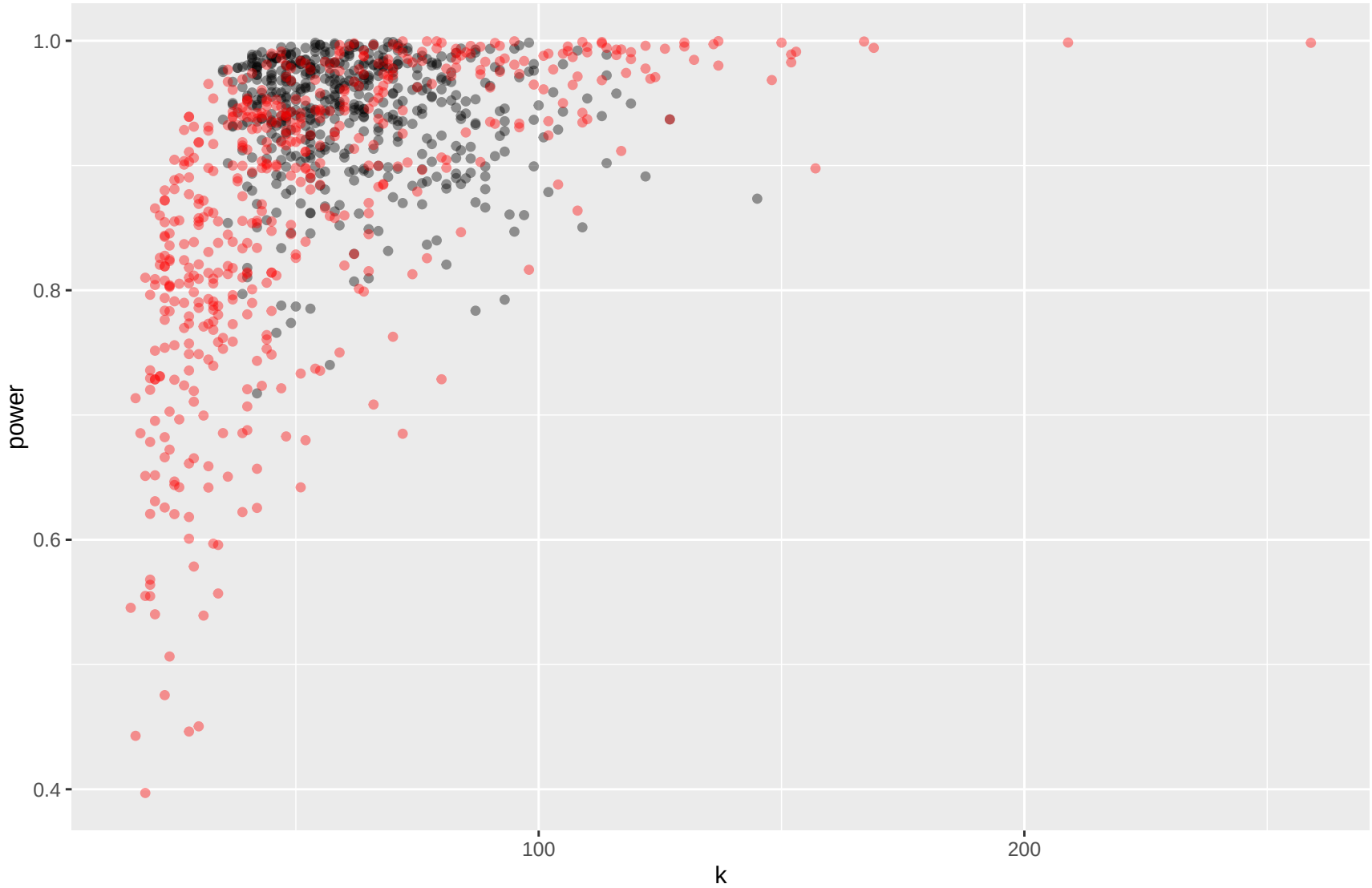
1. As in Eldridge et al., use the pilot data to compute an estimate $\hat{\rho}$ and plug this into the usual formula;
2. Ignore the pilot data and use only your prior knowledge to form a point estimate $\hat{\rho}$, e.g. choosing the mean / median / mode of the prior, then plug into the usual formula.

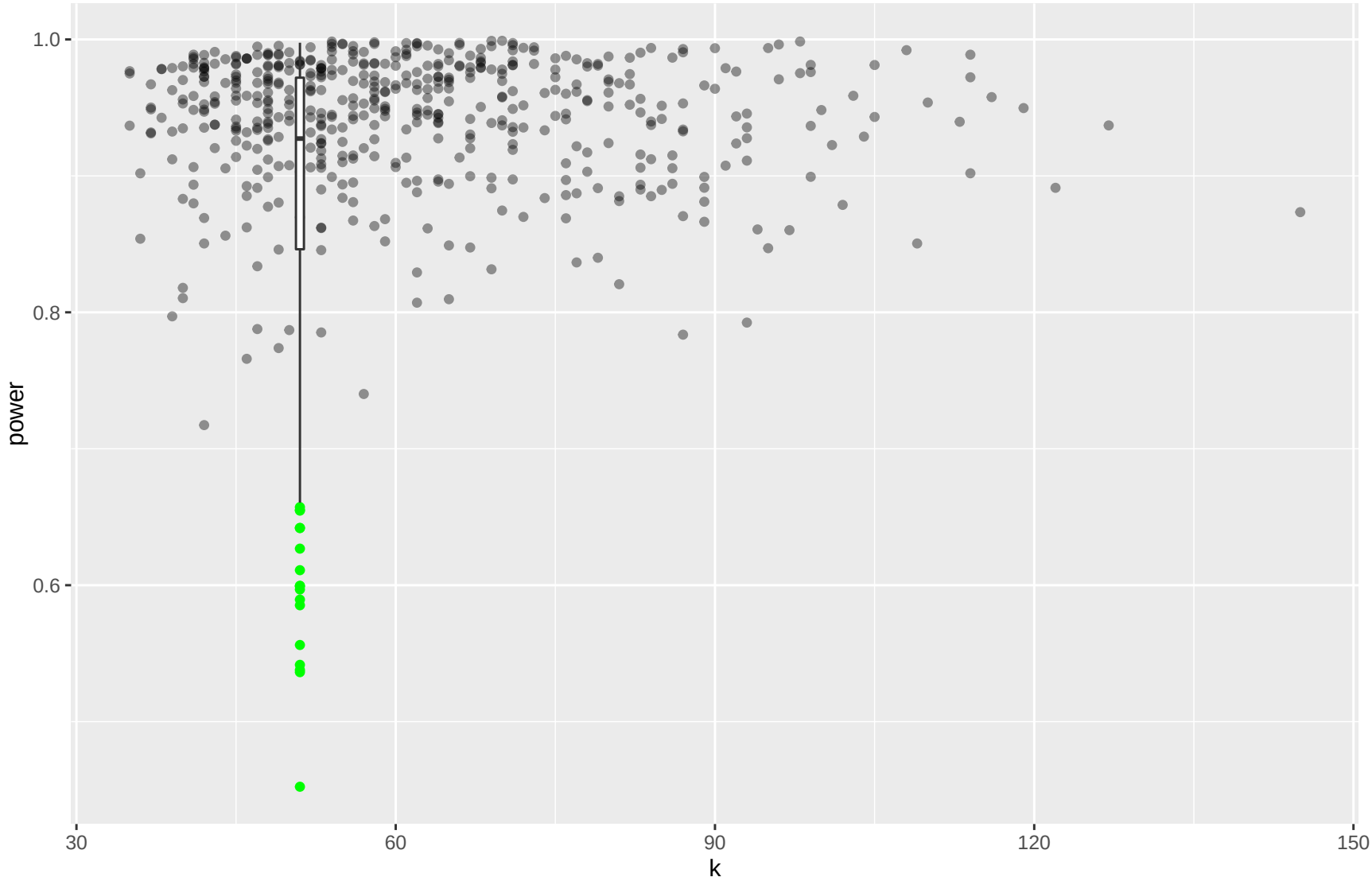
Compare with other pilot sample sizes k_p.

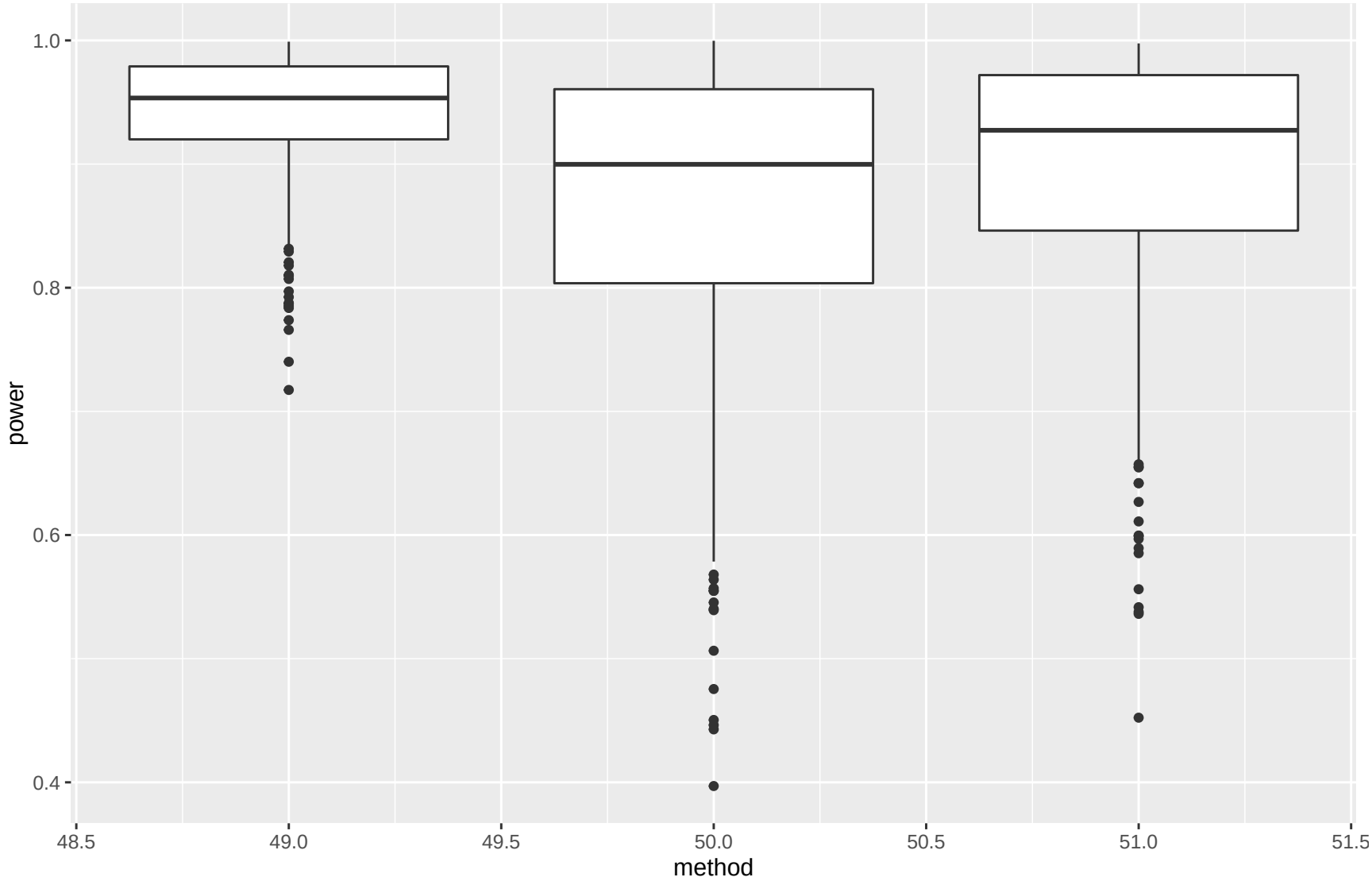
As a *very* tentative first step, a quick implementation lets us look at the joint predictive distribution of the main trial sample size k and its power.

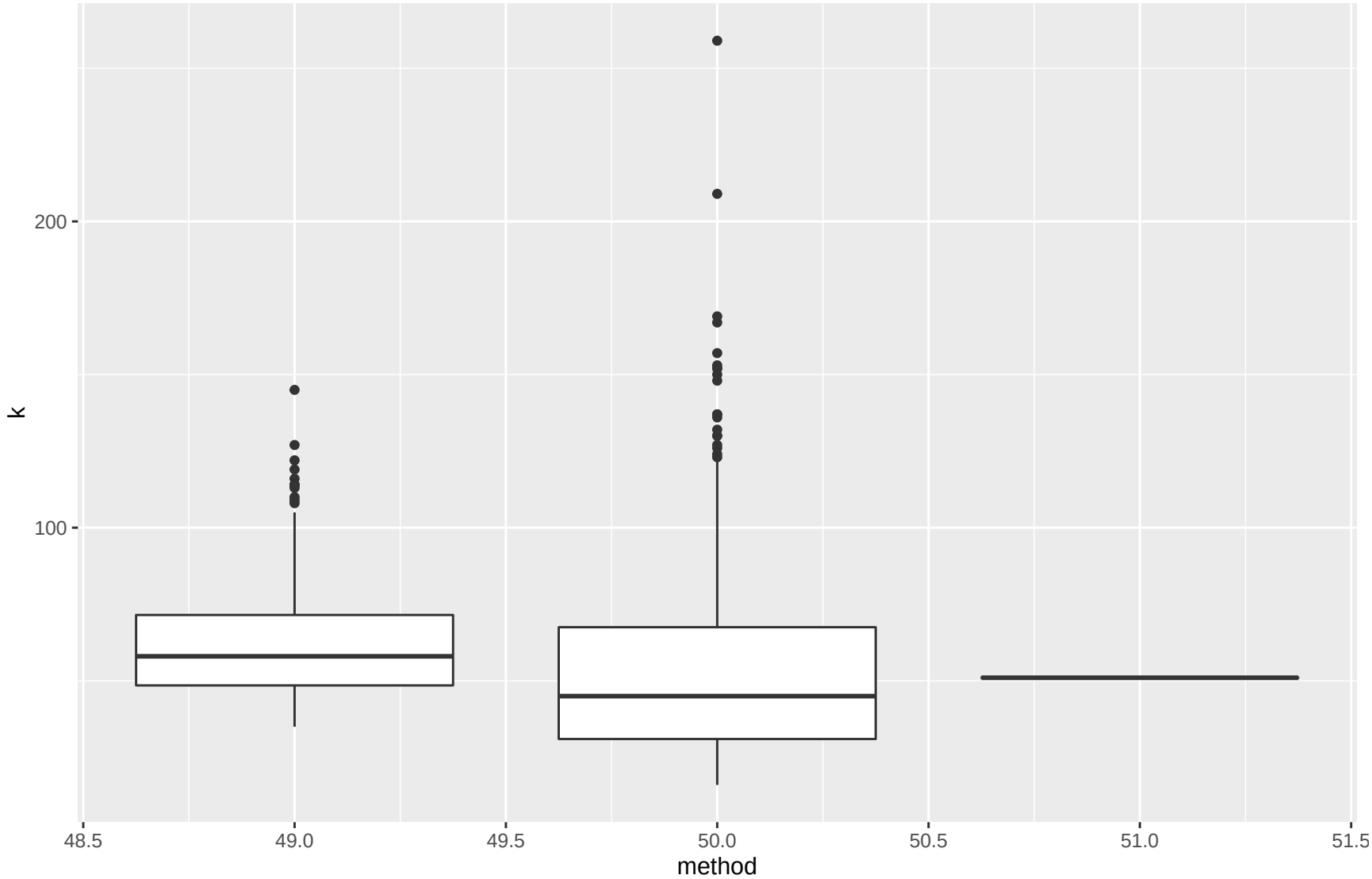
- Pilot sample size is $k_p = 10$.
- Cluster size in both trials is fixed at $m = 10$.
- In the main trial we want to detect a standardised effect size of $\delta/\sigma = 0.3$ with a power of 0.9.
- We take 500 samples of the joint distribution for $k, 1 - \beta$.

	Bayesian	Using pilot	Ignoring pilot
Prob underpowered	0.18	0.51	0.39
Mean power	0.94	0.87	0.89
Mean k	62	54	51


1.0
0.8
0.6
0.4
power
100
200
k


1.0
0.8
0.6
power
30
60
90
120
150
k


1.0
0.8
0.6
0.4
power
48.5
49.0
49.5
50.0
50.5
51.0
51.5
method


200
100
k
48.5
49.0
49.5
50.0
50.5
51.0
51.5
method

Comparisons are between joint distributions on $(k_p + k, \beta)$ - how to decide which is better?

Defining scope and questions to answer:

- Only uncertainty in the ICC, or in both variance components?
- Only cluster randomised trials, or all nested or partially nested designs?
- When will the different methods be better than each other?
- When will full elicitation be worthwhile?
- When will a pilot trial be worthwhile (purely as a means of ICC estimation)? When would it be better to try and elicit a less variable prior?

Making the computations feasible:

- Currently we optimise k for the main trial, and for each optimisation iteration we integrate over $p(\rho \mid x)$.
- We then need to do this many times using different samples of $p(x, \rho)$ for some pilot sample size k_p.
- We then want to optimise over k_p.
- Similar to decision-theoretic sample size determination - Lindley 1997.

Extending to other cases:

- Estimating several nuisance parameters? Multivariate prior elicitation?
- Incorporating uncertainty in the effect size?
- Likelihood function not available? Approximate Bayesian Computation (ABC)?