Power Considerations for Village RCT Odisha - Phase 1 - Diarrhea outcome

Alex Lehner

July2023

This document lays out power calculations for a binary outcome, measured at the individual level, for a cluster randomized trial (C-RCT) in Odisha, India. The outcome of interest is diarrhea (approximate incidence rate of 5%) and the treatment (inline chlorination devices for clean water) is assigned at the household level for a complete cluster (a village).

The crucial assumptions for these calculations are:

- Intra-cluster correlation (ICC) [`0.01` and `0.02`]
- (Average) size of villages (U5 children) [`30` and `50`]
- Effect size [`10%` and `20%`]

Autocorrelation of diarrhea incidence within individuals over time is considered here as well (in the form of an AR(1) process, roughly matching the values from Pickering et. al. (2019)), but it is less of a concern for power.

One of the main goals of the document is to illustrate the **trade-off between including more villages and increasing the number of survey rounds** to achieve power.

The OLS specification that is assumed regresses a dummy variable of diarrhea incidence on a treatment dummy at the village level and one fixed effect for every round - exploiting only within-round variation and thus controlling for unobserved variables that vary over time:

$$Y_{ic} = \beta_1 \; T_c + \sum_{r=1}^{3} \gamma_r \; ROUND_{ric} + \varepsilon_{ic},$$

with standard errors clustered at the village level (or the village by individual level to account for autocorrelation if the number of rounds, r, is greater than 1).

A quick comment on clustersizes. Please note that none of the following calculations assume variation in clustersize. In reality these are going to vary and thus harm statistical power. For details and some descriptices (i.e. the CV of sizes per district), please have a look at the dashboard that contains information on all census villages of Odisha. In practice, for the final randomization it could make sense to exclude very small and very big villages from the sample ex-ante in order to increase power.

To inform our calculations, we first analyze data from a clustered chlorination intervention in Bangladesh from Pickering et al (2019, The Lancet).

The document is structured as follows:

1) Relevant descriptive statistics from Pickering et. al. (2019, The Lancet)
2) Simple benchmark calculations for only one survey round (plug-in formulas)
3) Consideration of multiple rounds of data collection (simulations, DGP targeted to match data in Pickering et. al. (2019))

1) Descriptives from Pickering et. al. (2019, The Lancet)

Pickering et. al. (2019): 100 shared water points (clusters) in two low-income urban communities in Bangladesh were randomly assigned (1:1) to have their drinking water automatically chlorinated at the point of collection by a solid tablet chlorine doser (intervention group) or to be treated by a visually identical doser that supplied vitamin C (active control group). The trial followed an open cohort design; all children younger than 5 years residing in households accessing enrolled water points were measured every 2–3 months during a 14-month follow-up period (children could migrate into or out of the cluster). The primary outcome was caregiver-reported child diarrhoea (more than 2 loose or watery stools in a 24-h period [WHO criteria]) with a 1-week recall, including all available childhood observations in the analyses. Children in the treatment group had less WHO-defined diarrhoea than did children in the control group (control 216 [10.0%] of 2154; treatment 156 [7.5%] of 2073; prevalence ratio 0.77, 95% CI 0.65–0.91).

They collected 7 rounds in total and found an effect of roughly 25% (diarrhea incidence reduction). The study was block-randomized by matching pairs, i.e. there are 50 randomization blocks with one cluster each assigned to treatment and control.

Table 1: Diarrhea incidence for every round of collection, broken down by treatment and control group.

round	vitamins (control)	chlorine
1	0.0686275	0.0623853
2	0.0990826	0.0731225
3	0.1111111	0.0770791
4	0.0862745	0.0518672
5	0.0517598	0.0617021
6	0.0478360	0.0436893
7	0.0526316	0.0193548

Intra-cluster correlation (ICC) We calculate the raw ICC to be 0.0224 and 0.0167 after taking into account the blocking structure of the experiment. The sample size is not very large and we thus have reasons to believe that these estimates are not very precise. From earlier data work on child mortality with substantially larger sample sizes we obtained estimates of around 0.01. We therefore want to use 0.01 and 0.02 as benchmark values in our power calculations.

Checking counts per cluster

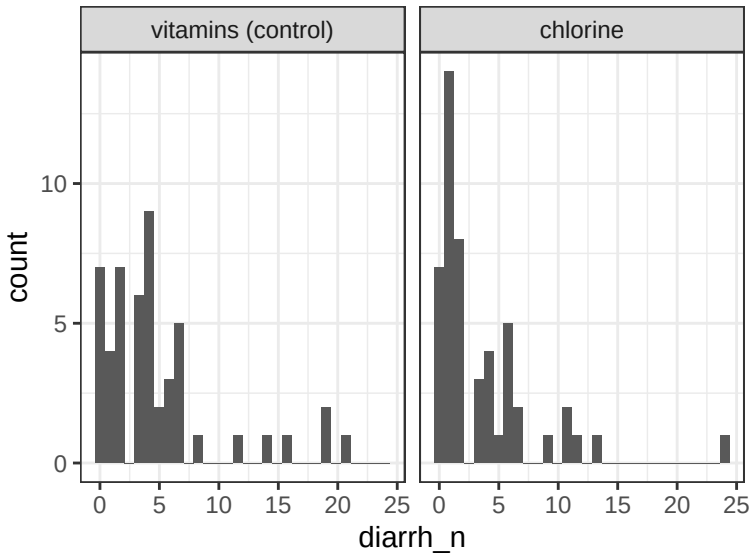


Figure 1: Diarrhea counts per cluster (pump), summed over all 7 rounds - approximately Poisson.

2) Benchmark calculations - one round only (plug-in formulas)

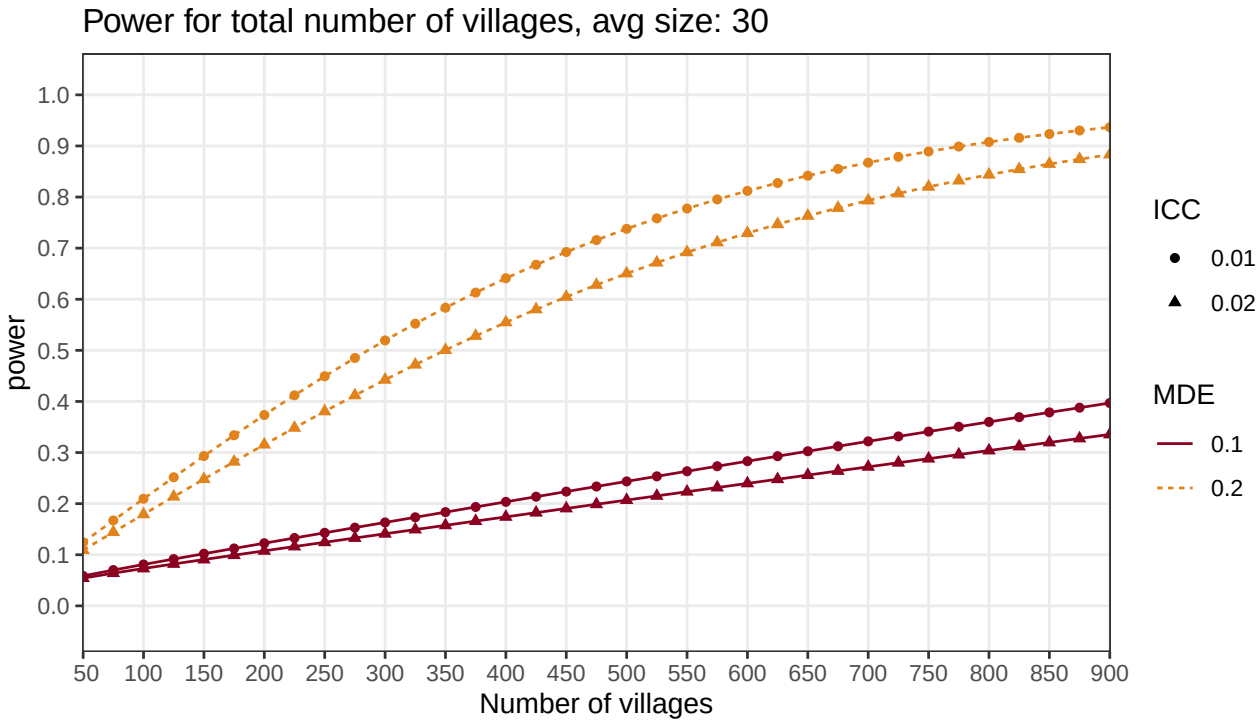
The following two plots show power as a function of the number of villages for two different average villages sizes: 30 and 50. The plots are grouped by ICC (`0.01` and `0.02`) and MDE (`10%` and `20%`). As noted in the introduction, the assumed diarrhea incidence is 5% across all specifications.

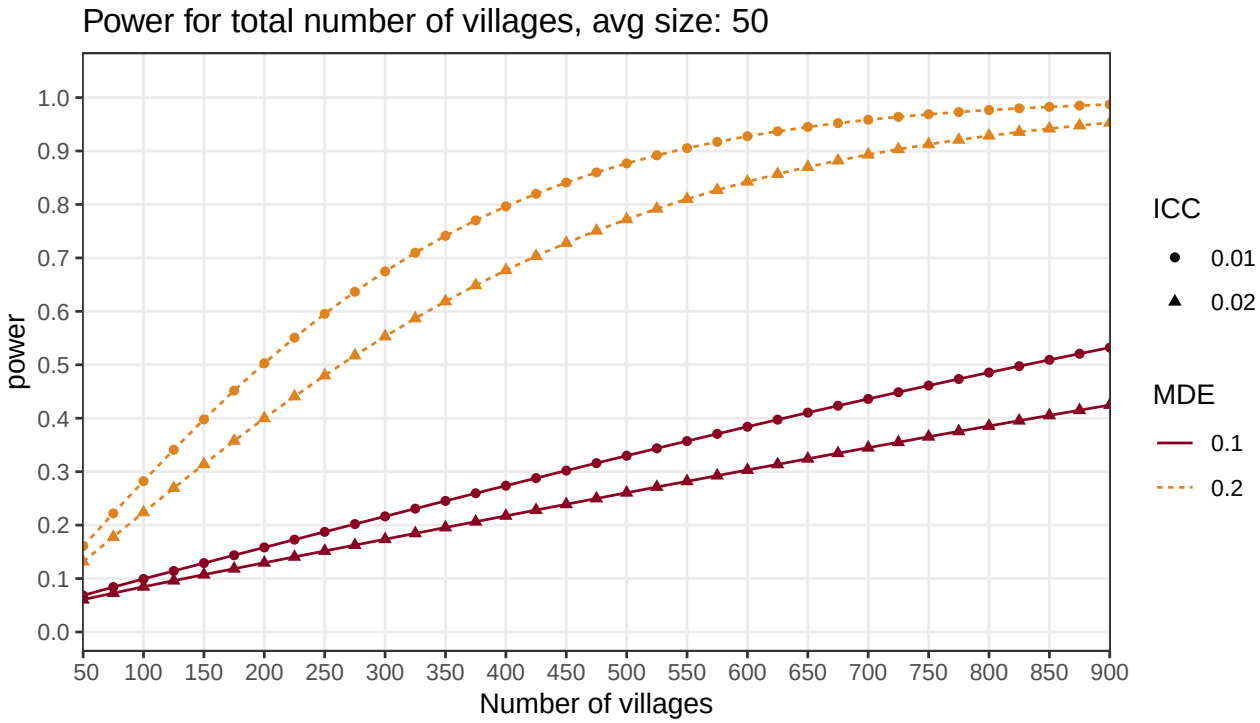
On the backend, the code uses the function `clusterPower::cpa.binary()` many times to compute each point in the plots.

The bottom line of these calculations is that for an MDE of 10% we are not going to be sufficiently powered. For the MDE of 20%, the following table summarizes the different scenarios where power is approximately 80% (extracted from the data that can be seen in the plots).

Table 2: Number of villages needed for different scenarios.

avg_clustersize	diarrhea_pc	MDE	ICC	villages_needed	power
30	0.05	0.2	0.01	575	0.8
30	0.05	0.2	0.02	700	0.8
50	0.05	0.2	0.01	400	0.8
50	0.05	0.2	0.02	525	0.8


Power for total number of villages, avg size: 30
power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900
Number of villages
ICC
0.01
0.02
MDE
0.1
0.2


Power for total number of villages, avg size: 50
power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900
Number of villages
ICC
0.01
0.02
MDE
0.1
0.2

3) Simulation: multiple rounds of data collection

The simulation code that is used here models the data generating process (DGP) as draws from a Bernoulli distribution with a cluster level random effect that induces ICC. The events within individuals across rounds are modeled with a correlated error term that induces an AR(1) process to approximately match the data in Pickering et. al. (2019). The following simulations are carried out 1,000 times for each dot that is visualized.

To showcase that the code delivers meaningful power estimates, the following plots showcase the scenario where the number of rounds in the simulations is 1. It can be seen that the resulting power curves are very similar to the results from above, verifying that they can approximate the results from the plug-in formula. Note that the curve for MDEs of 10% is a bit wobbly due to the small effect size, a larger amount of iterations would straighten them out further.

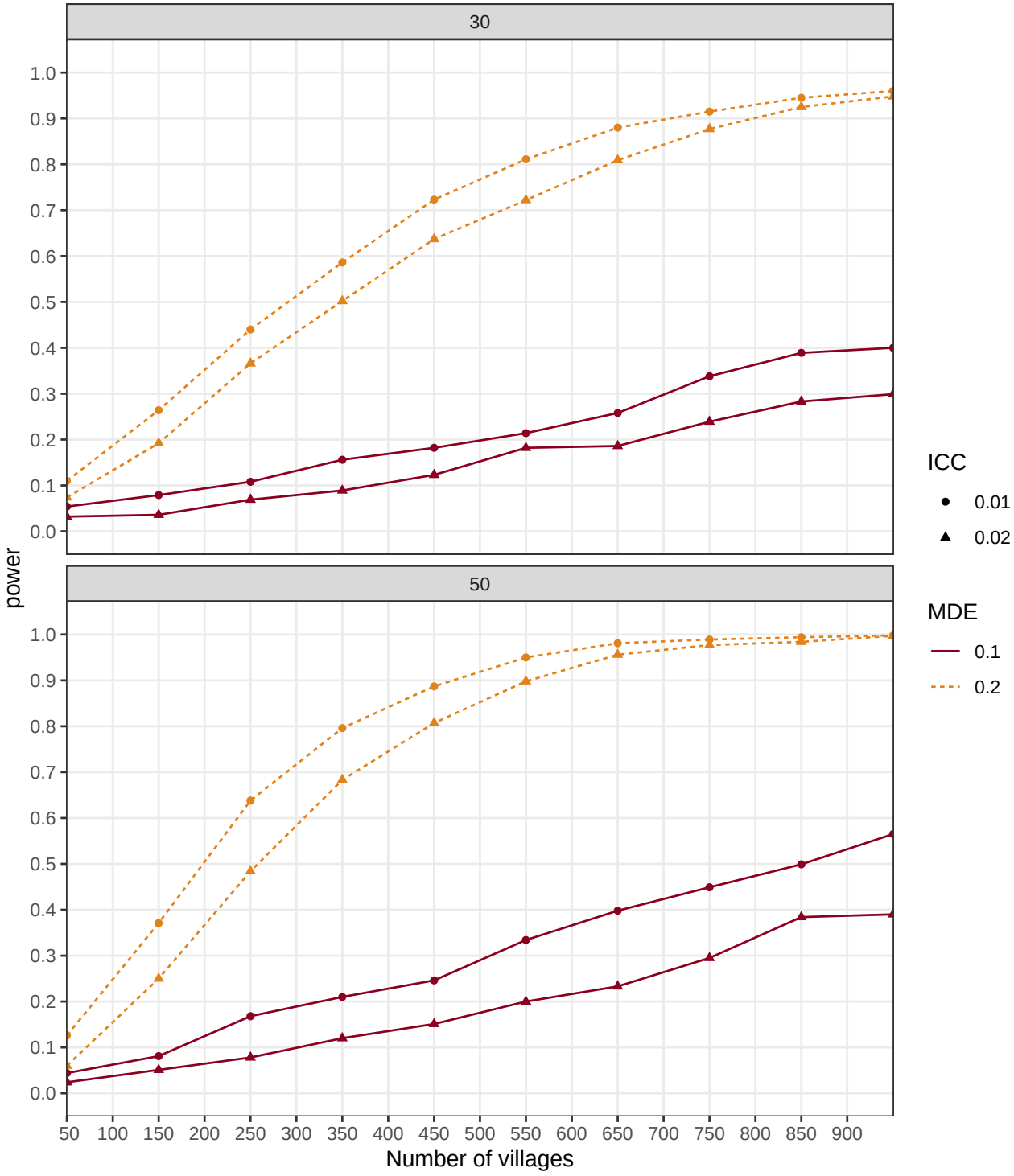


Figure 2: Showcase the working of the simulation code by setting the number of rounds to 1 and matching the results from the plug-in formulas above.