Odisha Pilot - Sampling Variation of Diarrhea Outcome

Alex Lehner

2023-08-14

The ILC Pilot in Odisha will have a fixed sample of 20 villages. We want to get a sense for how precise and informative diarrhea prevalence estimates from an endline census in all villages will be. The assumed diarrhea prevalance is **five percent** as per the latest NFHS survey at a district level.

There will be substantial sampling variation depending on the size of the villages. Therefore, I show the variation of the prevalence rate across different village sizes from 5 to 50 U5 children per village in increments of 5. Vaishnavi mentioned that a total village size of 500 people might be already too big because their water systems might get complex. Therefore, 50 U5 children could already be too high of an upper bound for what is a realistic number of U5 children that can be expected.

For each village size, the graphic displays the average total prevalence over all 20 villages. 100 simulations within each village size category are displayed (village size always constant) and 95 percent confidence intervals are displayed with red errorbars. These confidence intervals are printed in a separate table below the plot for every village size category.

The results suggest that **if diarrhea prevalance is five percent**, having somewhere between 25 and 40 U5 children per village should keep the margins of error reasonably in check. Obviously, a second census (at baseline) will give you another draw and thus more certainty - which would also allow you to assess the usefulness of multiple rounds of diarrhea censuses.

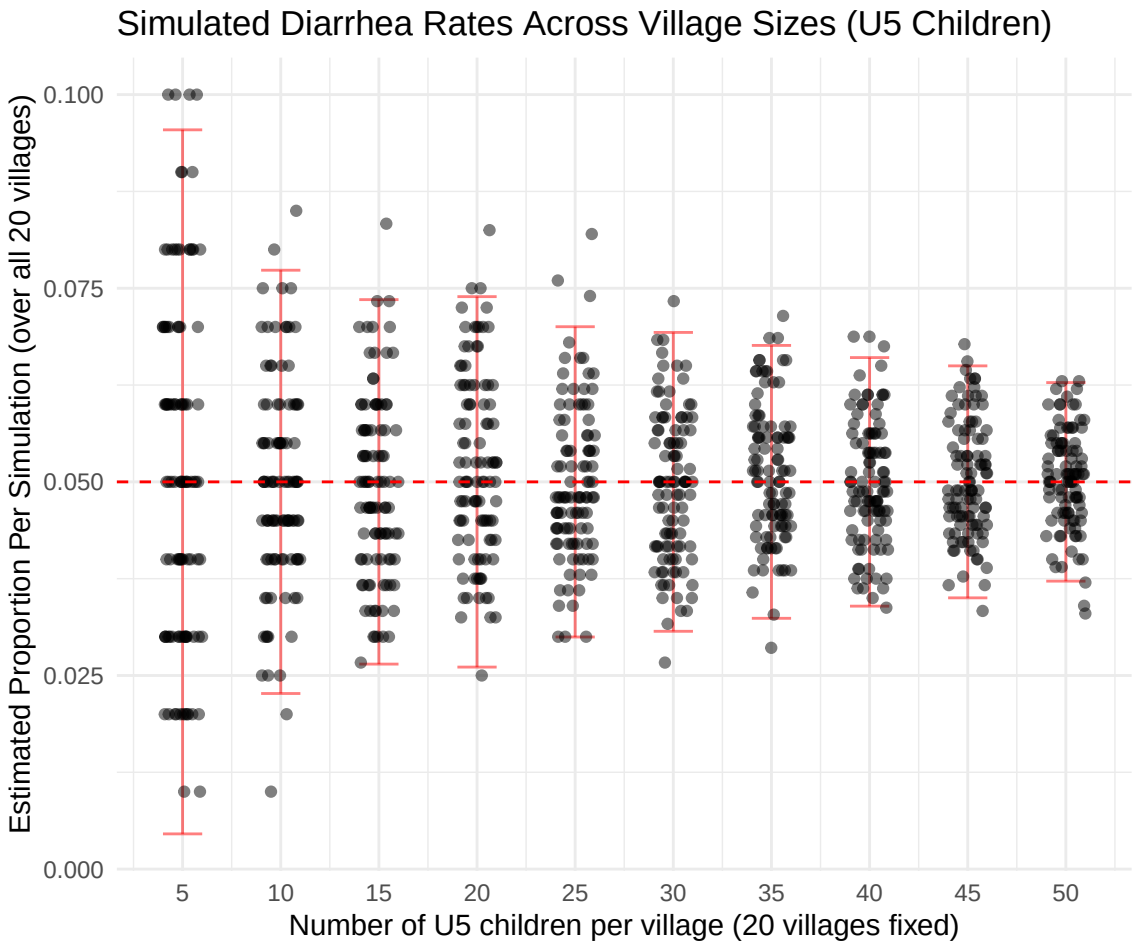

Table 1: 95percent confidence intervals for every village size (U5 children) as displayed by the errorbars in the plot above

GroupSize	CI_lower	CI_upper
5	0.005	0.095
10	0.023	0.077
15	0.026	0.074
20	0.026	0.074
25	0.030	0.070
30	0.031	0.069
35	0.032	0.068
40	0.034	0.066
45	0.035	0.065
50	0.037	0.063