Sample size for a parallel design CRT of a vaccine in Bioko

1. Assumptions

The trial will have two arms with a 1:1 ratio of individuals included per arm. The primary outcome is a binary classification of whether an individual is infected during the follow-up. The attack rate (i.e. proportion infected) in the absence of vaccine is 10%. The effect estimated is the overall effectiveness (i.e. including both direct and indirect effects). The key measure of the size of cluster is the number of trial participants in the cluster. If these are a large proportion of the population of the cluster then we expect an indirect effect of vaccination (herd immunity) which will lead to a higher effectiveness than the direct effect that is measured in an individually randomized trial. Contamination between arms (owing to mosquito movement), may somewhat reduce the herd immunity/indirect effect, but we anticipate that this would be rather small unless the clusters are extremely small in geographical area.

2. Sample size calculations

The standard formula for sample size for a cluster randomised trial is given in Donner, Birkett, Buck (1981). This was implemented in R using the ICC from Solarmal, a range of numbers of samples per cluster, and a range of effect sizes. This gives the number of clusters required for each arm in a two-armed trial. The key measure of the size of cluster is the number of trial participants in the cluster. If these are a large proportion of the population of the cluster then we expect an indirect effect of vaccination (herd immunity) which will lead to a higher effectiveness than the direct effect that is measured in an individually randomized trial.

3. Conclusions

The power depends mainly on the number of clusters and rather little on the number of samples per cluster. The figures Table 1a represent a reasonable **minimal** sample size, considering that the incidence may be lower. There should be some inflation of the number of clusters to allow for drop-out clusters, and a possibly lower control incidence. Table 1b is provided as an indication of what would be the corresponding values I the attack rate in the controls is lower than 10%.

Clearly a trial with small clusters is desirable, providing contamination is not excessive, but in urban Malabo a small cluster would presumably be an extremely small geographical area, making very large spill-over effects very likely.

Table 1a. Minimum Sample size required to achieve 80% power, 5% significance level, for vaccine (assuming attack rate in the controls of 10%)

		Effectiveness			
		0.75	0.5	0.333	0.167
		Total clusters required across all arms			
Number of participants per cluster	50	36	72	182	790
	100	32	64	162	702
	200	30	60	152	658
	500	30	58	146	632
	1000	28	58	144	624

Table 1a. Minimum Sample size required to achieve 80% power, 5% significance level, for vaccine (assuming attack rate in the controls of 5%)

		Effectiveness			
		0.75	0.5	0.333	0.167
		Total clusters required across both arms			
Number of participants per cluster	50	76	152	380	1662
	100	68	134	338	1476
	200	64	126	316	1384
	500	60	122	304	1330
	1000	60	120	300	1310