Competing effects of indirect protection and clustering on the power of cluster-randomized controlled vaccine trials

Authors: Matt D.T. Hitchings, Marc Lipsitch, Steven E. Bellan

Author affiliations: Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, Massachusetts (Matt D.T. Hitchings, Marc Lipsitch); Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, Massachusetts (Marc Lipsitch); Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, Georgia (Steven E. Bellan); Center for Ecology of Infectious Diseases, University of Georgia, Athens, Georgia (Steven E. Bellan).

Financial support: This work was supported by Award Number U54GM088558 from the National Institute Of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of General Medical Sciences or the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. S.E.B. was supported by a National Institute of Allergy and Infectious Diseases K01 award (K01AI125830).

Conflicts of interest: ML declares research support through his institution from Pfizer and PATH Vaccine Solutions, consulting income (donated in full to charity) from Affinivax, Pfizer and Merck, and consulting income from Antigen Discovery, Inc. The other authors declare that no competing interests exist.

Data availability: Code that can be used to generate data presented in this study is available in the Supplementary Material.

Word counts:

Abstract: 200

Main text: 3,998

Abstract: Sample-size theory for cluster-randomized trials (cRCTs) has been developed with a focus on clustering, which leads to two general rules of cRCT design: (i) a cRCT is less statistically efficient than an individually-randomized trial (iRCT) of the same size, and (ii) for a cRCT, it increases statistical power to enroll more, smaller clusters than fewer, larger clusters. However, when vaccination gives indirect protection to all individuals, trial design affects two other determinants of sample size: measured effect size and disease incidence. We test the generality of these rules, simulating trials in small communities in an emerging epidemic. We estimate vaccine effect and power, varying cluster size by changing the size of underlying communities and the proportion of each community enrolled. Across a range of parameters, we find no evidence that rule (i) is broken, though the efficiency difference can be small. When changing community enrollment proportion, increased indirect protection gained by enrolling a higher proportion counters the loss of power from enrolling larger clusters, and rule (ii) need not hold. When changing community size, increased case importation into larger communities partially or fully mitigates loss of power from enrolling larger clusters. Consideration of indirect effects is important for cRCT design.

Main Text:

Cluster-randomized controlled trials (cRCTs) have become an increasingly common method for evaluating interventions for infectious diseases, including vaccines. Compared to individually-randomized controlled trials, cRCTs may offer logistical, operational, and acceptability advantages (1), and allow the measurement of the combined direct and indirect effects of vaccination, which are often of great relevance to policy-makers (2). The statistical theory of cRCT design has largely focused on the effect of intracluster (i.e. within-cluster) correlation on statistical power (3-5). Within-cluster correlation arises because members of the same cluster share characteristics which are correlated with the outcome, and means that increasing the number of individuals within a cluster provides less information than would adding the same number of individuals to a new cluster. When the trial outcome is an infectious disease, correlation arises also because each case in a cluster can transmit infection to other cluster members. Two rules of thumb have prevailed regarding the effect of clustering on trial power: (1) a cRCT is always less efficient than an iRCT, meaning that a greater trial participant sample size is necessary to achieve the same statistical power to detect a given effect size; and (2) for a given sample size, a cRCT is more statistically efficient when it comprises more clusters with fewer individuals rather than fewer clusters with more individuals (2, 6). In this paper we use a disease transmission model to examine these two rules in the context of vaccine trials during an emerging epidemic, in which the ability of a cRCT to measure direct and indirect effects of vaccination counteracts some of the loss of efficiency brought about by clustering.

We first compare the statistical power of an iRCT with that of a cRCT in the same population, during an epidemic of a directly-transmitted disease, across a broad range of realistic parameters. We consider a cRCT design that measures the total vaccine effect (sum of the direct and indirect

effect) because it is the most statistically efficient of the common cRCT designs (2). Second, we restrict our attention to cRCTs only and consider two decisions an investigator must consider when balancing the number of clusters with the size of a cluster, for a trial of a given sample size (Fig. 1). Given recruitment of a specified number of participants from a set of communities, the first decision (*enrollment proportion*) concerns whether it is more efficient to enroll a larger proportion of each community from fewer communities, or to enroll a smaller proportion from a larger number of communities. The second decision (*community size*) concerns whether it is more efficient to recruit individuals from a smaller number of large communities or recruit from a larger number of small communities.

With regard to enrollment proportion, recruiting a higher proportion of each community leads to higher vaccine coverage in communities receiving vaccination and thus more indirect protection to individuals therein. The greater overall protection may lead to increased statistical efficiency. With regard to community size, larger communities may experience an increased rate of introduction into the community if, for example, disease importations are proportional to the number of travelers to and from the community, which likely scales with community size. Both the increased indirect protection and the increased importation rate may increase statistical efficiency because they increase the effect size and average number of cases in the trial population, respectively, both of which increase power. These effects may thus partially counterbalance the loss of efficiency that is known to accompany having fewer, larger clusters (due to within-cluster correlation). We use a network transmission model to estimate the magnitude of each effect, and the extent to which these effects mediate the relative efficiency of iRCTs and cRCTs.

METHODS

Population structure

The population is divided into two distinct groups: a main population in which a major epidemic is progressing, and a smaller population made up of multiple small *communities*. The trial takes place in a subset of these communities. The communities are represented with a stochastic block network model (7), in which contacts between individuals within the same block are far more common than those between blocks. A connection between individuals in the network represents a single infectious contact per day, and the number of connections for a given individual (degree) is Poisson-distributed. Communities are the unit of randomization for cRCT designs. Ideally for the trial, there should be no direct contact between communities to minimize spillover effects, so in the baseline scenario we assume that this is the case.

Transmission models

To balance realism with computational feasibility we implement the transmission models differently in the main population and in the communities. In particular we use a deterministic compartmental model in the main population and a stochastic compartmental model for transmission within the communities, as described below.

The main population model is a standard deterministic susceptible-exposed-infectious-removed (SEIR) compartmental model, with three exposed and three infectious compartments to yield gamma-distributed incubation and infectious periods. We assumed a time-varying transmission rate in the main population, so that the importation rate into the communities is proportional to

the prevalence of infection in the main population, and disease natural history parameters representative of the 2014-2015 Ebola epidemic in Liberia (8).

The disease model in the communities is a stochastic susceptible-exposed-infectious-removed-vaccinated (SEIRV) model. Each susceptible individual has a daily hazard of becoming infected and moving into the exposed compartment from two sources: the daily hazard of infection from each infectious *neighbor* is β, and the daily hazard of infection for an individual in community i from the main population is $F_i I$, where I is the number of infectious individuals in the main population and F_i is a proportionality constant reflecting the degree of contact between the main population and the ith community. The rate of introduction into the study communities varies with the size of the community. In particular, larger communities experience more disease importation events, with community importation rate M_i increasing with $\sqrt{N_i}$, where N_i is the size of the ith community (9). See Supplementary Material for more details on case importation rate.

We assume that the incubation and latent periods are concurrent, meaning that symptom onset occurs when infectiousness begins. Once infected, individuals spend a number of days in the exposed compartment drawn from a gamma distribution with mean 9.7 days and SD 5.5 days before moving into the infectious compartment (10). They spend a number of days in the infectious compartment drawn from an independent gamma distribution with mean 5 and SD 4.7 based on data on the time to hospitalization (10), after which they move into the removed compartment. For simplicity, we assume no *post mortem* transmission, meaning that whether an individual dies or recovers does not affect the estimated efficacy or power of the trial.

Vaccine trial design

For both designs, on a fixed calendar day a number of communities are enrolled and a target proportion is enrolled at random from the susceptible and exposed individuals therein to form that community's *study cluster*. In the iRCT, half the individuals in each study cluster are randomized to vaccination with the other half to control (in this model, assumed to be placebo control). In the cRCT all individuals in half the study clusters are assigned to vaccination, while those in the other half of the study clusters are assigned to control. In this design, all enrolled individuals in clusters assigned to vaccination are vaccinated.

Once enrolled, individuals are followed for a number of days and time from enrollment to symptom onset is recorded. Individuals who never develop symptoms are censored at the end of the study, and there are no other sources of censoring. The vaccine is multiplicative leaky (11), reducing susceptibility to infection by a factor (1-VE) and having no effect on those who are already exposed or infectious when vaccinated, and no effect on the progression or infectiousness of vaccinated individuals who become infected. We assume the protective efficacy of the vaccine starts on the day of vaccination.

Statistical analysis

Statistical analysis of the trial is based on time to symptom onset. For the iRCT a Cox proportional hazards analysis is performed to estimate the direct effect of the vaccine, stratifying by community. We determine statistical significance at the α=5% level using a two-tailed Wald test, and for each combination of parameters we simulate 500 trials, estimating the power as the proportion of trials that find the vaccine significantly efficacious. We examine the median vaccine effect estimate across the 500 simulated vaccine effect estimates when comparing across

designs or parameter values. To estimate the Type I error of a trial we repeat the above process with the true vaccine efficacy set to 0 simulating 1000 trials. To measure the magnitude of clustering in the cRCT we report the design effect, defined as $DE = \rho(m - 1)$, where ρ is the intracluster correlation coefficient (ICC) calculated using (12) and m is the average size of a study cluster. The design effect increases with ICC, as subjects in the same cluster are more similar, and with the size of each cluster, as there are more individuals in the trial who are similar to each other. The ICC is a measure of between-cluster variance relative to total variance in the outcome: if between-cluster variance is large relative to within-cluster variance, the ICC is large and individuals in the same cluster provide little information relative to individuals in different clusters.

In this cRCT design, a Cox proportional hazards model with treatment assignment as the explanatory variable estimates the total effect of vaccination (equal to the sum of direct and indirect effect). To ensure we use a cRCT analysis that balances power and Type I error when comparing cRCT power to that of an iRCT, we try several methods to account for clustering when determining statistical significance within the cRCT design: namely, a Cox proportional hazards model with Gaussian- or gamma-distributed shared frailty, and a Cox proportional hazards model with robust standard error estimate. The estimation of power, vaccine effect, and Type I error are performed as above. For all analyses, cases who developed symptoms within 10 days after vaccination (i.e. the average incubation/latent period) are excluded to avoid those who were already infected when vaccinated.

Choice of parameters

Table 1 shows the parameters used in the model, their meanings, values under baseline assumptions, range explored (where applicable), and references or justifications. See Supplementary Material for parameters values for the disease model in the main population.

RESULTS

Comparison of iRCT and cRCT

Across a broad range of parameters, including population structure, trial design and vaccine efficacy parameters, we found that individually-randomized vaccine trials almost always were more efficient than cluster-randomized trials in the same population, despite the larger effect size being measured in cluster-randomized trials. Figure 2 illustrates the relative power of iRCT versus cRCT designs versus R_0, and highlights three findings. Firstly, the cRCT generally yields greater effect size estimates than the iRCT, because it measures the total (direct + indirect) vaccine effect rather than solely direct effects (Fig. 2A). Secondly, the design effect is large and increases with increasing R_0 (Fig. 2B), because large R_0 leads to more outbreaks within communities, which increases between-cluster variance and thus the ICC (see Fig. S1). Therefore, the efficiency that the cRCT gains by measuring a larger effect is more than compensated by loss of efficiency due to within-cluster correlation. These two points explain why cRCT power first increases and then decreases with increasing R_0. As R_0 increases past a certain threshold, the effect of clustering begins to dominate the effect of increased incidence in the study population, and the trial loses rather than gains power from the increased transmission.

Finally, the estimate of vaccine effect from a Cox proportional hazards model with robust standard errors (robust SE model) decreases drastically as R_0 increases. Because communities are disconnected, outbreaks are confined to communities and as a result the hazard of infection varies significantly between study clusters. Uncontrolled heterogeneity in exposure to infection leads to underestimation of vaccine efficacy (13, 14), which cannot be adjusted for in the robust SE model though, as discussed below, can be addressed in a gamma frailty model.

From Figure 2 it is clear that the power of the cRCT is strongly affected by the design effect (Fig. 2C), and that the difference in power between the cRCT and iRCT is smaller when there is low R_0. This observation held when other parameters were varied, including trial start day, vaccine efficacy, importation rate, and population structure. In the setting of low R_0, epidemics will die out stochastically in most clusters experiencing one or more case importation. The cluster-level attack rates are thus close to zero and the between-cluster variance is small. These results can be viewed in R script app.R in the Supplementary Data.

Varying community enrollment proportion in a cRCT

Restricting attention to cRCTs, Figure 3 displays the vaccine effect estimate (Fig. 3A), design effect (Fig. 3B), and power (Fig. 3C) for a cRCT across varying *community enrollment proportions* (holding community sizes constant, but varying number of communities). As expected the estimate of total vaccine effect increases with increasing proportion enrolled because it increases vaccine coverage and, consequently, the effect of herd immunity in vaccinated clusters. However, the increased effect size is counterbalanced by increases in the design effect (driven by larger clusters). Thus, for all values of R_0 displayed except R_0=3 there is no clear trend in power with the community enrollment proportion. Only for R_0=3 do the

simulations follow the trend normally expected for cRCTs in which the use of more, smaller clusters increases trial power.

Varying size of enrolled communities in a cRCT

Figure 4 displays the attack rate in the study population (Fig. 4A), design effect (Fig. 4B) and power (Fig. 4C) for a cRCT with varying *size of enrolled communities*, holding the total number of participants constant, when community case importation rate scales with the square root of community size. As the community size increases, the per-community importation rate increases and with it the proportion of communities that receive an importation. When R_0 is greater than 1 and many importations lead to outbreaks, fewer communities means that a higher proportion of communities experience outbreaks. In effect, with fewer communities the small communities are linked up, meaning that an outbreak in one can spread to other small communities, boosting the attack rate. On the other hand, having many small communities means that outbreaks are limited by the size of the communities and the overall attack rate is lower. However, in this case the increased attack rate with fewer communities is not large enough to offset the greater design effect and thus power decreases with increasing community size, consistent with the standard expectation for cRCTs.

When community case importation rate scales linearly with community size, there is a greater increase in attack rate when there are fewer large communities (Fig. 4D). In this case, even though the design effect increases with community size (Fig. 4E), the higher attack rate offsets the increased design effect and power does not change appreciably with size of enrolled communities when transmission is moderate (Fig. 4F).

Analysis methods for a cRCT

In answering our primary research questions, we explored a range of analysis methods for the cRCT. We found that a Cox proportional hazards model with Gaussian-distributed frailty had significantly elevated Type I error and low power to detect vaccine effect. Fortunately, the two most common approaches to analyzing clustered survival data, a Cox proportional hazards model with gamma-distributed frailty or robust standard error estimation, were the best methods in terms of power and validity. The robust SE analysis has the advantage of estimating one parameter fewer than the gamma-frailty model so it has higher power when transmission is low. However, the model ignores population structure in its estimate of the vaccine effect, leading to a downward bias that is particularly apparent when R_0 is high, as seen in Figure 2. The gamma-frailty model is not susceptible to this bias.

DISCUSSION

Statistical comparisons of cRCTs versus iRCTs usually focus on within-cluster correlation and the design effect, but neglect other ways in which the unit of randomization affects cRCT power. We show that investigators should additionally consider how a cRCT's ability to measure both indirect and direct effects can partially compensate for the loss in power due to clustering. Although we do not find convincing evidence that the cRCT is ever substantially more efficient than an iRCT in the same trial population, in low transmission settings the difference in efficiency between them may be small.

Analyses focused on optimizing cRCT power have similarly focused on clustering, while neglecting other effects. For instance, when designing a cRCT it is generally held that, for a given sample size, enrolling greater number of smaller clusters is more statistically efficient than

fewer clusters of greater size (2, 6). Using greater numbers of clusters certainly reduces the design effect, which consequently contributes to increased power. However, enrolling a greater proportion of a community also increases indirect protection to all participants in that cluster, thus increasing the measured total effect size of the vaccine in the trial. To our knowledge, this fact has been alluded to but the effect on power has never been explicitly quantified (6, 15). Here we have shown that the ability to measure indirect effects can offset the loss of efficiency due to the increased design effect. Another counter-intuitive finding arises from the fact that, because larger communities experience a greater influx of transmission imported from elsewhere, enrolling fewer but larger communities may yield a greater attack rate, and thereby partly or fully compensate for the loss in power due to the design effect. This result is dependent on the relationship between case importation rate and community size, meaning that it is only relevant to settings in which the pathogen is not endemic to trial communities and the probability of pathogen introduction into a community is relatively low, and will differ by disease and population setting.

Our findings also highlighted the importance of identifying statistical analyses that can adequately account for the effect of heterogeneity between study clusters while maintaining the nominal false positive rate and maximizing power. We limited the analysis methods to those that are widely available in software packages and found that a Cox PH model with gamma-distributed frailty performs best overall, although when transmission is low a Cox PH model with robust SE has higher power and produces valid estimates of the total vaccine effect.

The results presented here are part of a body of work demonstrating the utility of simulation when considering the design of vaccine trials for infectious diseases (16, 17). It is only by including transmission dynamics in model that we are able to quantify the relative strength of

clustering and indirect protection in affecting trial power. Our study is intended to explore these effects more generally, but we expect our findings to be relevant to investigators considering cRCT design, whether or not they develop a full-fledged trial simulation study during the planning phase. Theoretical work on trial design can help prepare stakeholders to rapidly design trials in the face of unexpected epidemics of emerging pathogens.

Statistical efficiency is only one of many factors that must be taken into consideration when planning a vaccine trial. Considerations of logistics, ethics, acceptability or the particular research question of interest may, in certain contexts, hold priority. Indeed, in this work we have focused on a standard cRCT design in which communities randomized to vaccination have all enrolled participants vaccinated (2) as used in many vaccine trials (18, 19). This approach yields the largest effect size; but designs that only vaccinate a proportion of enrollees in vaccinated communities have other advantages such as portraying trial vaccine coverages that may more realistically reflect vaccine rollout.

There are at least two sources of intercluster variation (and thus intracluster correlation) in a cRCT for an infectious disease: transmission within clusters, which makes infection risk more similar within clusters than between, and shared characteristics of individuals within a cluster. When R_0 is large enough, any outbreak that takes off will infect many individuals in a community and will be close to its deterministic final size (for large enough community sizes), while the outcomes of individuals in a cluster with no cases are perfectly correlated, leading to ICCs approaching 1. Clustering can also arise due to shared characteristics in a community, such as proximity to a large city or sanitation. Clustering of cases in the study population increases the ICC and thus the design effect, which can be used to calculate the necessary sample size for future trials. Given these different sources of clustering, and the fact that we observed ICCs

ranging from 0.05 to 0.8 in our simulations, it is critically important that ICCs are reported by study investigators when presenting the results of a cRCT (20).

The model presented here neglects some important aspects of a realistic population in which a trial is conducted. For example, we do not consider the second source of clustering described above (i.e. shared characteristics). More broadly, individuals and communities in the model do not vary in any characteristics other than degree, whereas real populations would vary in age structure, proximity to the epicenter of the epidemic, and other variables that would predict disease incidence. By ignoring these characteristics we underestimate the extent of clustering in a cRCT. In this way, our findings might overstate the power of the cRCT, such that our conclusion that the cRCT rarely, if ever, is more efficient than the iRCT is robust.

The communities represented in the model do not reflect all possible communities considered for enrollment into a cRCT. We have conceptualized the population structure as being a number of small groups separated in space so that there is minimal spillover between communities, but in a real population the structure is likely to be less distinct. We have not considered permanent or temporary migration of people into, out of, and between communities, nor have we considered secondary structure within communities to represent individual households. Moreover, real-life degree distributions have a heavier tail than the Poisson-like degree distribution chosen. While we did not explore the full range of degree distributions, we performed a sensitivity analysis in with a power law degree distributions and our results were robust to this assumption.

We find that the usual rule that increased community size leads to decreased power is strongly dependent on the relationship between community size and rate of importation, as shown in Figure 4. The pattern of importation frequency proportional to the square root of community size

is based on a finding for measles (9), but it is likely that for some diseases the community-level importation rate is independent of community size, in which case the increased design effect would entirely dictate the loss of power as community size increases. Our conclusions should thus be considered in light of the context of each specific disease and population.

We have not explored the full range of cluster-randomized designs, nor the full suite of possible methods to analyzing such trials. Future work should might explore the effects discussed here in the context of stratified or matched cluster-randomization designs, which can improve the efficiency and Type I error of cRCTs. We restricted ourselves to methods readily available, but some work has been on development of GEE methods for survival data (21), which could provide a useful extension to the robust SE method that is unhindered by the bias seen here.

The indirect effect of vaccination should be considered along with clustering in calculating the power of a cluster-randomized trial and in comparing different trial designs for interventions against infectious diseases. Using a mathematical model we show that it does not always increase power to enroll more small clusters into a cRCT, when doing so reduces the indirect protection to vaccinated individuals or the importation of cases into the study population. However, although cRCTs measure a greater vaccine effect than iRCTs, we found that iRCTs are generally more efficient, though power is comparable in low-transmission settings.

References and Notes:

1. Smith PG, Morrow R. *Field Trials of Health Interventions in Developing Countries: A Toolbox*. London: Macmillan; 1996.
2. Halloran ME, Longini IM, Struchiner CJ. *Design and analysis of vaccine studies*. New York: Springer; 2010.

3. Donner A, Birkett N, Buck C. Randomization by cluster: Sample size requirements and analysis. *Am J Epidemiol* 1981;114(6):906-14.

4. Campbell MJ, Donner A, Klar N. Developments in cluster randomized trials and Statistics in Medicine. *Stat Med* 2007;26(1):2-19.

5. Rutterford C, Copas A, Eldridge S. Methods for sample size determination in cluster randomized trials. *Int J Epidemiol* 2015;44(3):1051-67.

6. Hayes RJ, Alexander NDE, Bennett S, et al. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat Meth Med Res* 2000;9:95-116.

7. Karrer B, Newman ME. Stochastic blockmodels and community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 2011;83(1 Pt 2):016107.

8. Kucharski AJ, Camacho A, Flasche S, et al. Measuring the impact of Ebola control measures in Sierra Leone. *PNAS* 2015;112(46):14366-71.

9. Keeling MJ, Rohani P. The Importance of Imports. In: Keeling MJ, Rohani P, eds. *Modeling Infectious Diseases in Humans and Animals*. Princeton: Princeton University Press, 2007:209-12.

10. Team WHOER. Ebola virus disease in West Africa--the first 9 months of the epidemic and forward projections. *N Engl J Med* 2014;371(16):1481-95.

11. Smith PG, Rodrigues LC, Fine PEM. Assessment of the protective efficacy of vaccines against common disease using case-control and cohort studies. *Int J Epidemiol* 1984;13(1):87-93.

12. Shoukri MM, Donner A, El-Dali A. Covariate-adjusted confidence interval for the intraclass correlation coefficient. *Contemp Clin Trials* 2013;36(1):244-53.

13. White MT, Griffin JT, Drakeley CJ, et al. Heterogeneity in malaria exposure and vaccine response. *Malaria J* 2010;9(82).

14. O'Hagan JJ, Hernan MA, Walensky RP, et al. Apparent declining efficacy in randomized trials: examples of the Thai RV144 HIV vaccine and South African CAPRISA 004 microbicide trials. *AIDS* 2012;26(2):123-6.

15. Halloran ME, Longini IM, Jr., Cowart DM, et al. Community interventions and the epidemic prevention potential. *Vaccine* 2002;20:3254-62.

16. Hitchings MD, Grais RF, Lipsitch M. Using simulation to aid trial design: Ring-vaccination trials. *PLoS Negl Trop Dis* 2017;11(3):e0005470.

17. Bellan SE, Pulliam JRC, Pearson CAB, et al. Statistical power and validity of Ebola vaccine trials in Sierra Leone: a simulation study of trial design and analysis. *The Lancet Infectious Diseases* 2015;15(6):703-10.

18. O'Brien KL, Moulton LH, Reid R, et al. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. *The Lancet* 2003;362(9381):355-61.

19. Sur D, Ochiai L, Bhattacharya SK, et al. A Cluster-Randomized Effectiveness Trial of Vi Typhoid Vaccine in India. *N Engl J Med* 2009;361(4):335-44.

20. Henao-Restrepo AM, Camacho A, Longini IM, et al. Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). *The Lancet* 2017;389(10068):505-18.

21. Zhang X. Generalized estimating equations for clustered survival data. *Retrospective These and Dissertations* 2006;3063.

22. Meyers LA, Pourbohloul B, Newman ME, et al. Network theory and SARS: predicting outbreak diversity. *J Theor Biol* 2005;232(1):71-81.

Figures legends:

Fig. 1. Schematic of an iRCT (top) and a cRCT (bottom). Study clusters (solid outlined) are enrolled from communities. In the iRCT, individuals within each cluster are randomized to vaccine (green) or control (blue). In the cRCT, half the clusters are randomized to vaccine and half to control. In the cRCT design, there are two ways to balance cluster size and number: (**1**) among communities of a certain size, vary the enrollment proportion and the number of communities enrolled, and (**2**) change the size of communities recruited into the trial, keeping the percentage of each community enrolled constant.

Fig. 2. Comparison of vaccine effect estimates and power of individually- and cluster-randomized controlled trials. Vaccine effect estimates (**A**), design effect (**B**), and power (**C**) from an individually-randomized controlled trial (iRCT) and from a cluster-randomized controlled trial (cRCT) analyzed either using a shared gamma frailty model or using a Cox PH model with robust standard error estimates (robust SE). The incidence rate of importations into an average community is 0.5 cases/year, and the vaccine efficacy is 60%.

Fig. 3. Relationship between power and community enrollment proportion for a cRCT. Vaccine effect estimates (**A**), design effect (**B**), and power (**C**) from a cRCT versus the percentage of individuals enrolled from each community, with total sample size held constant and assuming a vaccine efficacy of 60%.

Fig. 4. Relationship between power and size of enrolled communities for a cRCT. Attack rates in the trial population (**A** and **D**), design effects (**B** and **E**), and power (**C** and **F**) for cluster-randomized vaccine efficacy trials versus the size of the communities recruited, with total

sample size held constant. In the left-hand column, community case importation rate is proportional to the square root of community size, and in the right-hand column it is proportional to the community size. Panels **A** and **D** shows results for R_0=3 only. All results shown here assume 60% community enrollment.

Table 1. Model Parameter Names, Values and Ranges Varied Across, Meanings and References or Justifications.

Parameter	Meaning	Value/Range	Reference
R_0	Average number of secondary infections generated by an infected individual	0.6-3	Wide range spanning most emerging infectious diseases. Calculated for network models using (22).
Mean (latent)	Mean latent period length (days)	9.7	(10)
SD (latent)	Standard deviation of latent period length (days)	5.5	(10)
Mean (infectious)	Mean infectious period length (days)	5.0	Time to hospitalization (10).
SD (infectious)	Standard deviation of infectious period length (days)	4.7	(10)
VE	Individual vaccine efficacy	0.6 (0.4-0.8)	Baseline assumption.
N_i	Size of community *i*	100 (50-200)	Assumption that some unit of this size exists in the population.

M_i	Importation rate into communities	0.025 (0.0125-0.05)$\sqrt{N_i}$ cases/1,000 person-years	Based on a calculation for measles (9), with the magnitude of the rate chosen so that there is on average 0.5 importations into a community of size 100 over a two-year epidemic.
Within-community degree	Average total number of contacts of an individual within the same community	14.85 (14.83-14.85)	Based on *Ebola, ça suffit!* trial (20) (ring size of 90, <20% of which were primary contacts).
Between-community degree	Average total number of contacts of an individual from outside their community	0 (0-0.02)	Assumption that communities disconnected to minimize spillover effect. A range was explored to represent one or two contacts outside each community.
Trial size	Average number of individuals enrolled	4,000	Assumption to achieve reasonable power for chosen parameters.

Trial start day	First day of enrollment, vaccination and start of follow-up, relative to the first day of the epidemic in the main population	150 (100-250)	Assumed the trial starts before the peak of the epidemic in the main population and that the trial team is ready to go when epidemic starts.
Trial length	Length of follow-up after trial start (days)	140 (70-210)	Assumption to achieve reasonable power for chosen parameters.

List of supplementary material

Supplementary Materials and Methods.

Figure S1. Relationship between R_0 and distribution of cluster-level attack rates.

Code for running epidemic model. simulate_trial.R.

Figures:

Figure 1:

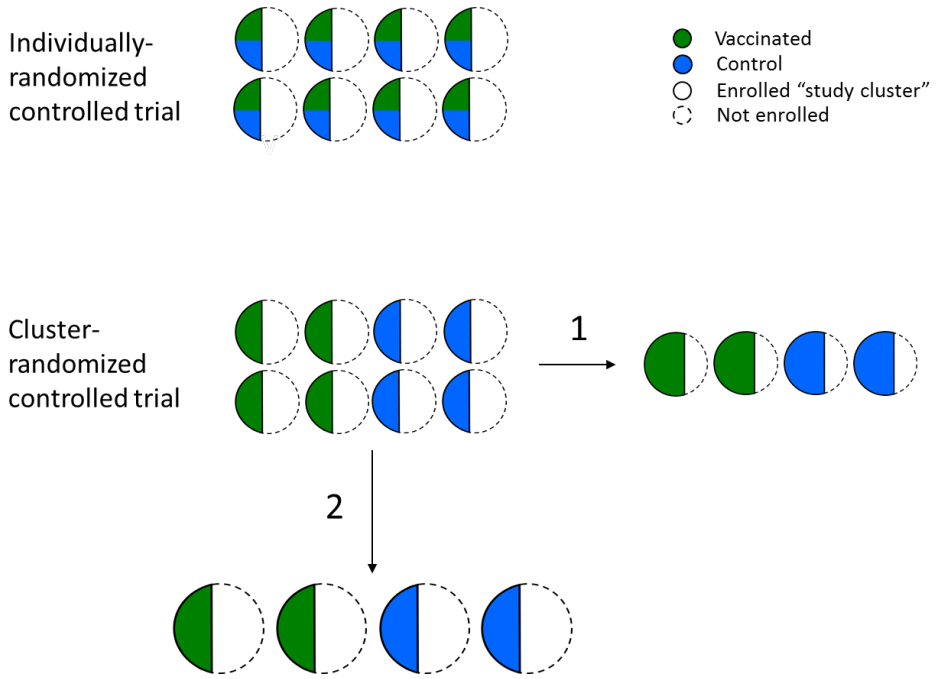


Figure 2:

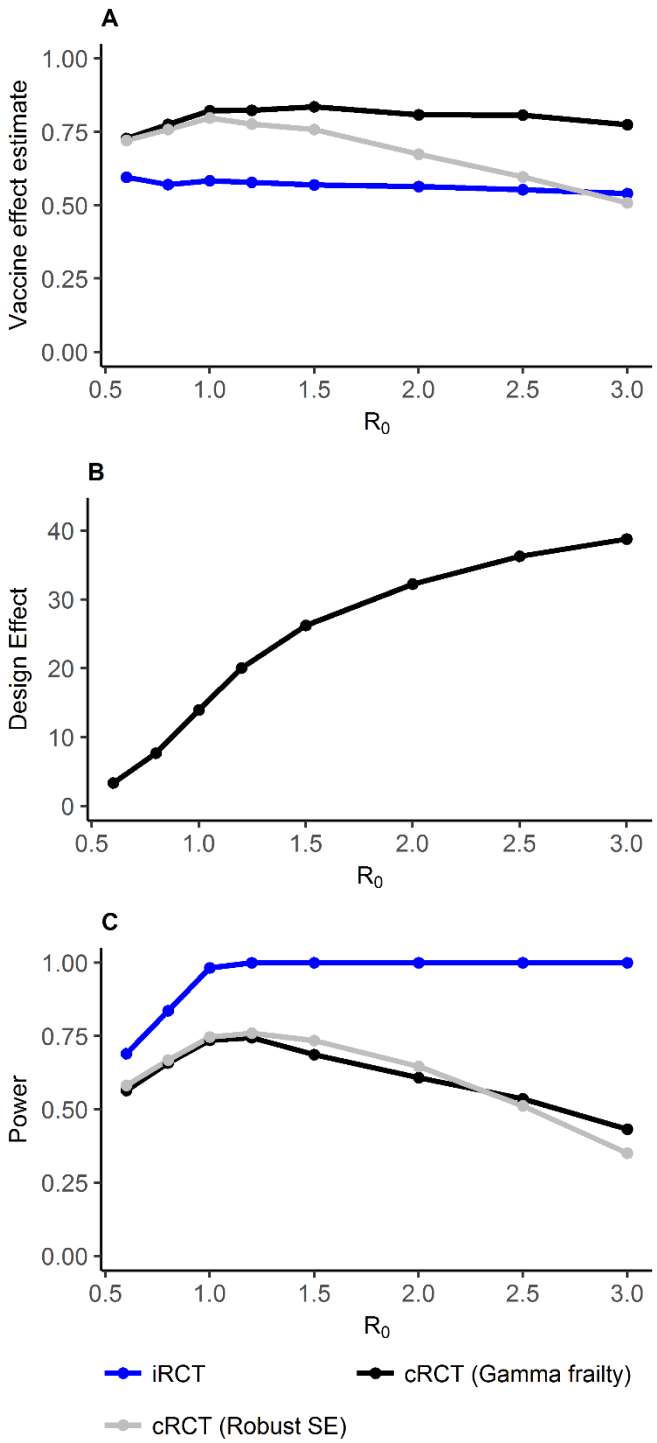
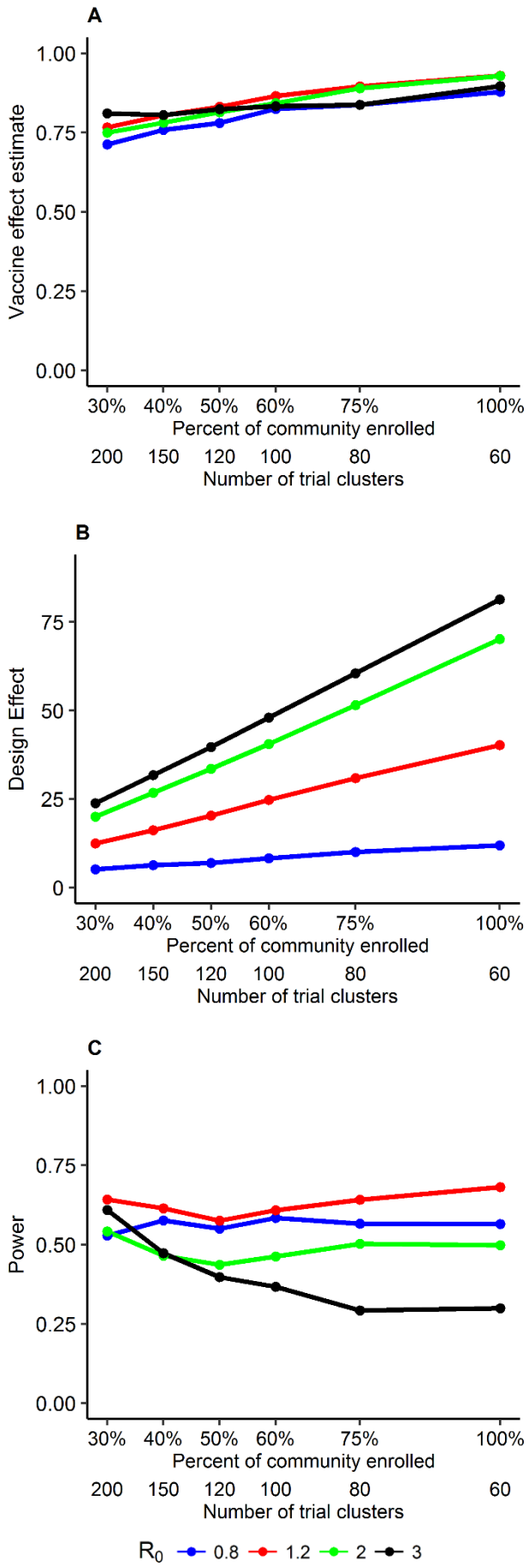

A
Vaccine effect estimate
1.00
0.75
0.50
0.25
0.00
0.5
1.0
1.5
2.0
2.5
3.0
R0
B
Design Effect
40
30
20
10
0
0.5
1.0
1.5
2.0
2.5
3.0
R0
C
Power
1.00
0.75
0.50
0.25
0.00
0.5
1.0
1.5
2.0
2.5
3.0
R0
iRCT
cRCT (Gamma frailty)
cRCT (Robust SE)

Figure 3:


A
Vaccine effect estimate
1.00
0.75
0.50
0.25
0.00
30% 40% 50% 60% 75% 100%
Percent of community enrolled
200 150 120 100 80 60
Number of trial clusters
B
Design Effect
75
50
25
0
30% 40% 50% 60% 75% 100%
Percent of community enrolled
200 150 120 100 80 60
Number of trial clusters
C
Power
1.00
0.75
0.50
0.25
0.00
30% 40% 50% 60% 75% 100%
Percent of community enrolled
200 150 120 100 80 60
Number of trial clusters
R0 0.8 1.2 2 3

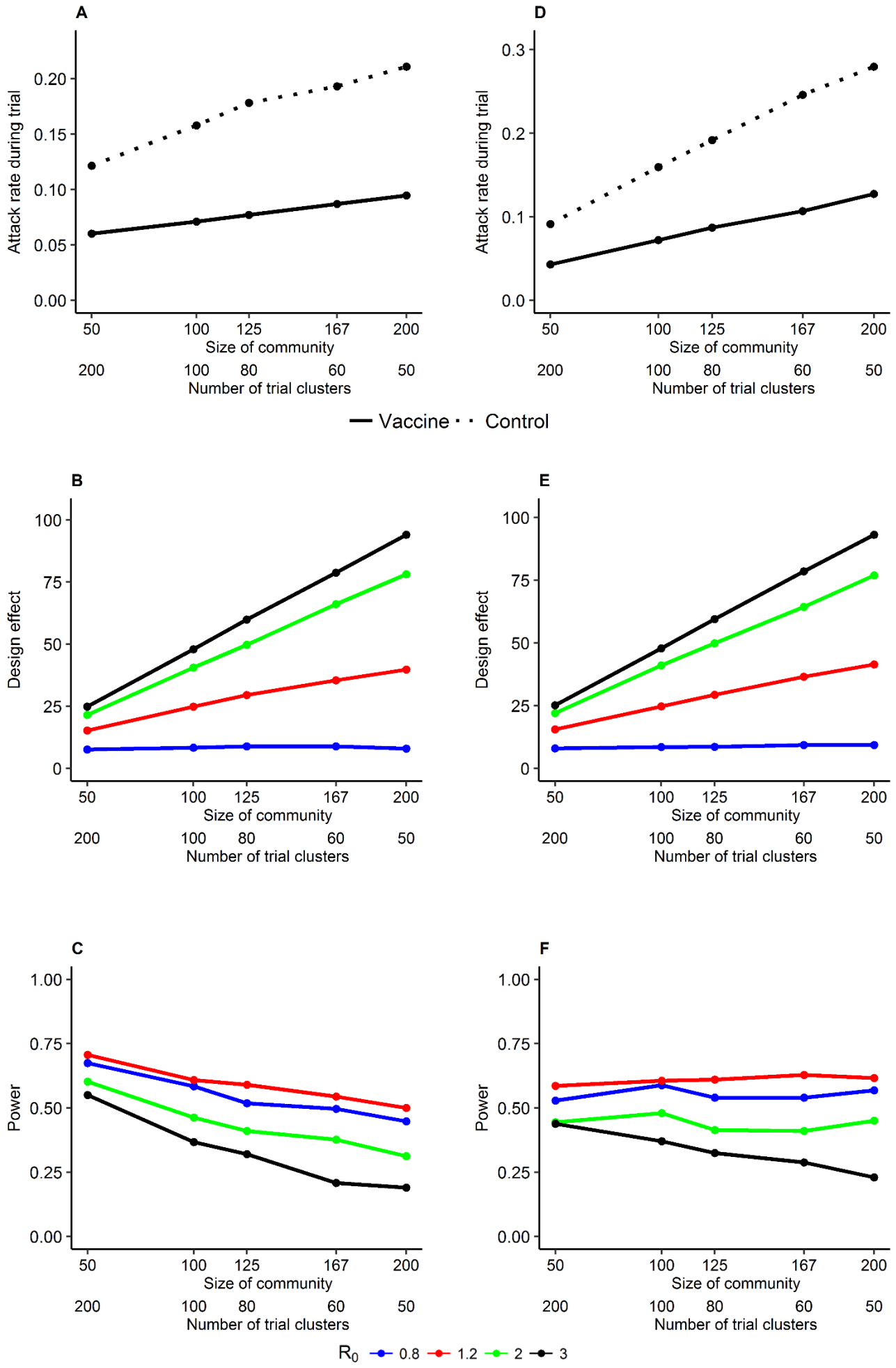

A
Attack rate during trial
0.20
0.15
0.10
0.05
0.00
50
100
125
167
200
Size of community
200
100
80
60
50
Number of trial clusters
D
Attack rate during trial
0.3
0.2
0.1
0.0
Vaccine
Control
B
Design effect
100
75
50
25
0
E
C
Power
1.00
0.75
0.50
0.25
0.00
F
R_0
0.8
1.2
2
3

Figure 4 (previous page):

Supplementary Materials:

Methods

The hazard rate of introduction into the study population is time-varying with the progression of the epidemic in the main population, and we calibrate the constant of proportionality in each cluster F_i using an assumed incidence rate of importation events, M_i cases/person-year. The formula that connects these two quantities is $F_i = -\frac{\ln(1-M_i*T)}{f/\mu}$, where f is the final size of the epidemic in the main population, μ is the mean infectious period, and T is the length of the epidemic in years. We model the relationship between importation rate and community size in two ways. Firstly, for community i we assume $M_i = a\sqrt{N_i}$, where N_i is the community size (*15*), and the *per capita* importation rate in community i is $\frac{a}{\sqrt{N_i}}$, where the constant a determines the magnitude of the importation rate. Secondly, we assume $M_i = a'N_i$, so that the *per capita* importation rate in community i is a'. The values for a and a' were chosen so that a community of size N_i=100 had on average between 0.25 and 1 introductions over the course of a two-year epidemic.

The transmission rate β in the main population varied with time using the formula $\beta(t) = \hat{\beta}(1 - \frac{\alpha_2}{1+e^{-\alpha_1(t-\alpha_\tau)}})$. Parameters were chosen to give a reasonable fit to weekly Ebola incidence data from Liberia. Specifically, $\hat{\beta}$ = 0.94, α_1 = 0.19, α_2 = 0.6, α_τ = 27.79. The average incubation/latent period is 7.14 days and the average infectious period is 3 days.

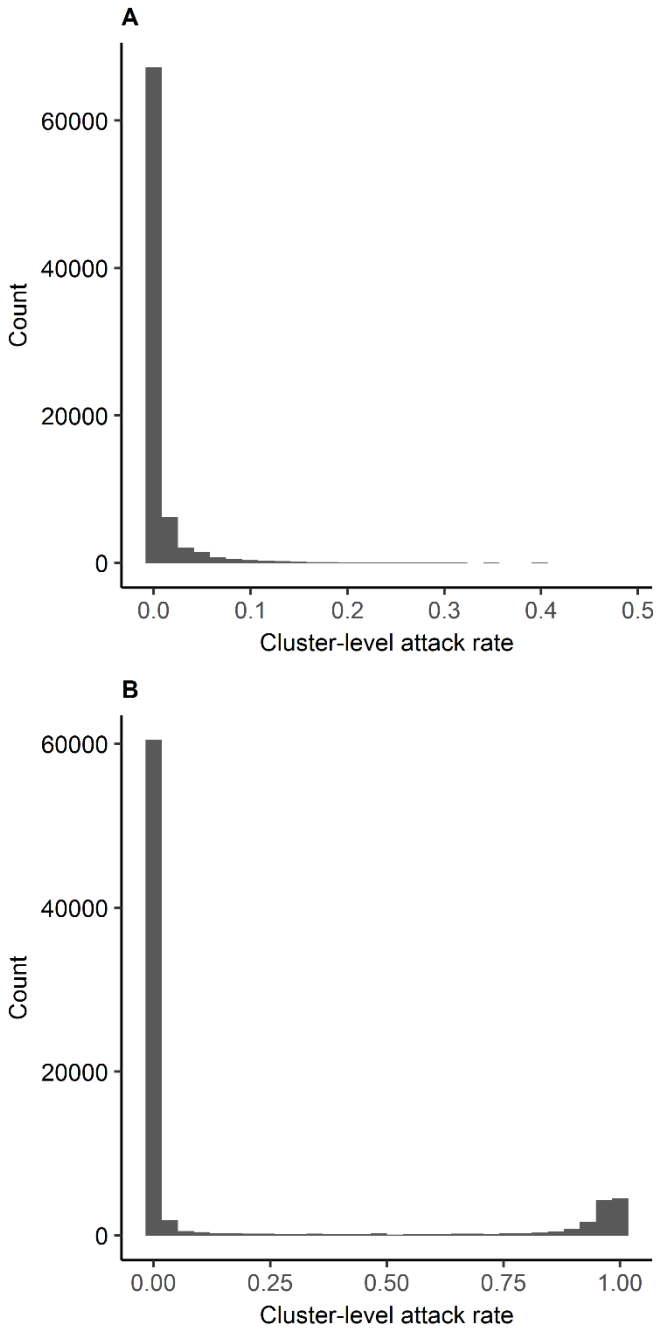



Fig. S1. Relationship between R_0 and distribution of cluster-level attack rates. Histogram of cluster-level attack rate for R_0=0.6 (**A**) and R_0=3 (**B**).