Leveraging Contact Network Structure for Cluster Randomized Trial Design

Guy Harling[1,2], Rui Wang[2,3], JP Onnela[2], Victor DeGruttola[2]

[1.] Department of Global Health & Population, Harvard T. H. Chan School of Public Health [2.] Department of Biostatistics , Harvard T. H. Chan School of Public Health
[3.] Brigham and Women's Hospital, Boston MA

Background

- Cluster randomized trials (CRT) for Ebola are in development
- Standard (e.g. Parallel or Stepped Wedge) randomization methods do not take account of the connections between clusters
- In an acute epidemic setting, there is an urgent need to achieve control of the infection, as well as to evaluate intervention efficacy

Aim

- To develop new CRT study designs that reduce the number of new infections more rapidly than standard designs, while still allowing for the evaluation of treatment effectiveness

Methods

Step 1: Simulate an undirected, individual-level contact network for a multiple cluster setting; hold mean contact number constant, but vary within- & between-cluster contacts

Step 2: Simulate Ebola epidemics on these networks using an agent-based SEIHFR state-transition model (*Legrand et al Epidemiol Infect 2007*)

Step 3: Simulate the impact of various vaccination designs based on randomized assignment of clusters

Step 4: Repeat the entire process (networks generation and epidemic spreading) for each study design, discarding runs where epidemics die out pre-intervention

Step 5: Calculate outcomes under each design for epidemic impact and power to estimate a treatment effect

Number of communities	20 more-connected (MC); 20 less-connected (LC)
Size of each community	100
Within-community ties (mean; SD)	MC: (5.0 0). LC: (4.5; 0)
Between-community ties (mean; SD)	MC: (0.5; 0.5). LC: (1.0; 0.5)
Disease transmission parameters	From Legrand et al. Epidemiol Infect. 2007
Number of initial infections	4
Vaccine uptake, effectiveness	90%, 85%
No. communities vaccinated per week	2

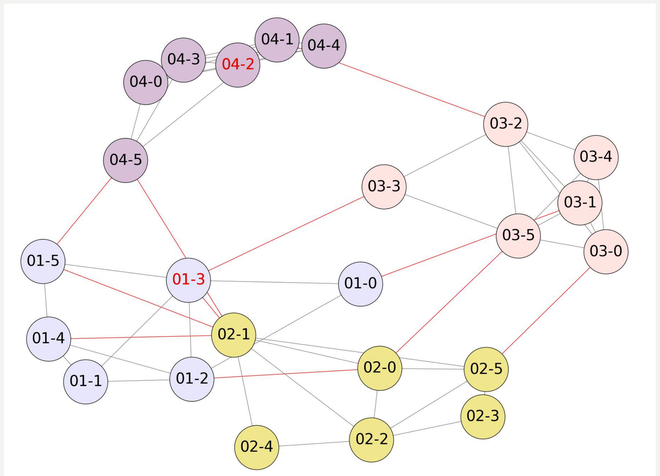
Baseline parameter values

Potential data sources

Using these designs would require connectivity data. Potential sources include: (i) contact tracing data; (ii) transport pattern data; (iii) cellphone data records; (iv) community interviews

Acknowledgments

This research was supported by R37 AI51164 and R01 AI24643 from the National Institutes of Health



A schematic of a 4-cluster network
(Blue lines: within-cluster ties; Red lines: between-cluster ties)

Vaccination designs

All proposed designs are derived from the Stepped Wedge

Static network approaches:
At start of trial, order clusters from most to least connected

1. Strict Hierarchy (without randomization):
   Treat in order from most to least connected

   Pro: Likely to be fastest control strategy
   Con: Confounding of treatment effectiveness measures

2. Fuzzy Hierarchy (with randomization):
   t=0: Randomize 1st & 2nd most-connected clusters
   t=1: Randomize the untreated cluster from t=0, and the 3rd most-connected cluster
   Continue until all clusters randomized

   Pro: Ability to make pairwise comparisons
   Con: Limited follow-up time on control clusters

3. Parallel Pairs (with randomization):
   t=0: Randomize 1st & 2nd most-connected clusters
   t=1: Randomize 3rd and 4th most-connected clusters
   Once all clusters randomized, treat control clusters in order

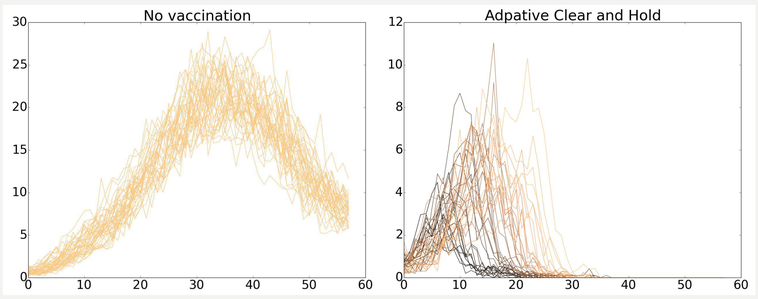
   Pro: Provides longer follow-up time on controls
   Con: Likely to reduce speed of epidemic control

Adaptive network approaches:
At each timepoint, re-calculate connectivity, excl. vaccinated clusters This method can be applied to designs 1 & 2 above

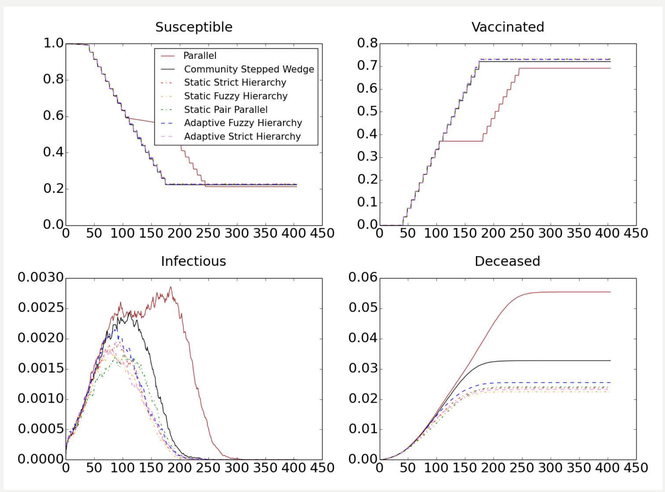
Pro: Likely to increase speed of epidemic control
Con: Requires data on complete contact patterns between clusters

Connectivity can be measured by either local or global network properties. Here we use *absolute cluster out-degree*, the number of connections leaving the cluster.

Results



Mean cluster incidence/1000 person-weeks (100 runs)
(*Darker lines represent clusters which received vaccination sooner*)



Mean temporal disease dynamics by day (100 runs)

	Weeks to $R_e < 1$		Cumulative Incidence (%)	Days to Last Infectious
No vaccine	33.5 (14-41)		49.9 (44.2-54.3)	405 (404-405)
Parallel CRT (10 wk delay)	13 (9-20.25)		9.5 (4.6-14.1)	274 (249-289)
Stepped Wedge	12 (9-16.25)		4.7 (2.9-7.4)	207 (186-221)
Network-based designs:				
Static Strict Hierarchy	11 (8-13)	**	3.6 (2.2-5.5) **	189 (160-208) ***
Static Fuzzy Hierarchy	10 (8-13)	***	3.2 (2.0-5.2) ***	178 (162-197) ***
Static Pair Randomization	11.5 (8-15)		3.7 (1.9-5.4) **	193 (169-213) ***
Adaptive Strict Hierarchy	11 (9-13.25)	**	3.8 (2.2-6.2) *	192 (169-207) **
Adaptive Fuzzy Hierarchy	10 (8-13)	**	2.9 (2.0-5.4) **	186 (161-208) ***

Significantly lower than the stepped wedge design based on $\chi^2_{(1)}$ tests: * $\alpha<0.05$; ** $\alpha<0.01$; *** $\alpha<0.001$.
Simulation ran 405 days.

Population level outcome measures (median & IQR)

Next Steps

1. Explore parameter space for between-cluster contact variability and real-world patterns of cross-cluster connectivity
2. Identify most useful metrics for describing intervention impact

Conclusion

Randomization for CRTs based on cluster-level network properties may provide more rapid epidemic control than standard designs, as well as allowing inference on treatment effectiveness

HARVARD | SCHOOL OF PUBLIC HEALTH

gharling@hsph.harvard.edu | www.guyharling.com