Sample Size Estimates for Cluster Randomized Trials in Infection Control and Antimicrobial Stewardship

Natalia Blanco[1], Anthony D. Harris[1], Laurence S. Magder[1], Kelly M. Hatfield[2], John A. Jernigan[2], Sujan C. Reddy[2], Lisa Pineles[1], Eli Perencevich[3], Lyndsay M. O'Hara[1] for the CDC Prevention Epicenters Program

University of Maryland School of Medicine

Background

- Cluster randomized trials (CRTs) are used frequently in the field of infection control and antimicrobial stewardship because randomization at the patient level is often not feasible due to contamination, ethical, or logistical issues.
- The correlation that exists among individual patients in a cluster must be accounted for when estimating sample size, yet many studies neglect to consider or report the intracluster correlation coefficient (ICC) or the coefficient of variation (CV).
- The aim of this study was to estimate the sample sizes needed to adequately power studies of hospital-level interventions to reduce rates of healthcare-associated infections.

Methods

- Parameters needed were estimated using national rates from the National Healthcare Safety Network (NHSN) for methicillin-resistant *Staphylococcus aureus* (MRSA) bacteremia, central-line associated bloodstream infections (CLABSI), catheter-associated urinary tract infections (CAUTI), *C. difficile* infections (CDI) and variation between hospitals in these rates (Table 1).
- Minimum number of clusters or hospitals that would be needed in a study to have good power for detecting an impact of the intervention.

Methods (Cont.)

- These calculations were based on the assumption that hospitals were uniform and moderate in size and were studied for one year.

Table 1. Model parameters

Outcomes*	Average Rate (per patient day)	SD of hospital-specific acquisition rates (per patient day)	CV of hospital-specific acquisition rates¥	No. of Hospitals	Average cluster size∞
MRSA	0.000057	0.000031	0.54	706	243.2
CAUTI	0.001053	0.00092	0.87	827	29.2
CLABSI	0.000985	0.00058	0.62	760	12.7
C. *difficile*	0.000709	0.00029	0.41	1088	146.5

*CAUTI, MRSA, C. *difficile* estimates are derived from NHSN data, 2016; CLABSI estimates are derived from NHSN data, 2015.
¥ The coefficients of variation (CVs) were calculated by dividing the standard of deviation of hospital-specific acquisition rates by the average rate.
∞Average cluster size was defined as the number of patient- or device-days per cluster. This was calculated from the NHSN data by dividing the total number of patient- or device-days by the total numbers hospitals and total study days.

Results

- To study an intervention leading to a 50% decrease in daily rates and using the CVs calculated from NHSN, 44 average-sized hospitals for MRSA bacteremia are needed, 68 for CAUTI, 18 for CDI, and 54 for CLABSI to have a statistically significant decrease with a type I error rate of 0.05 and a type II error rate of 0.8.
- If a 10% decrease in rates is expected instead, 1418, 2410, 558, and 1732 hospitals respectively are needed to have a statistically significant decrease with a type I error rate of 0.05 and a type II error rate of 0.8.

Table 2. Example of sample size estimates for one healthcare-associated infection outcome

	MRSA Bacteremia											
Anticipated incidence rate (intervention) (effect size)	$2.85x10^{-5}$ (50%)	$3.99x10^{-5}$ (30%)	$5.13x10^{-5}$ (10%)	$2.85x10^{-5}$ (50%)	$3.99x10^{-5}$ (30%)	$5.13x10^{-5}$ (10%)	$2.85x10^{-5}$ (50%)	$3.99x10^{-5}$ (30%)	$5.13x10^{-5}$ (10%)	$2.85x10^{-5}$ (50%)	$3.99x10^{-5}$ (30%)	$5.13x10^{-5}$ (10%)
Anticipated incidence rate (control)	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$	$5.7x10^{-5}$
Cluster size	243.2	243.2	243.2	243.2	243.2	243.2	243.2	243.2	243.2	243.2	243.2	243.2
CV	0.44	0.44	0.44	0.54	0.54	0.54	0.64	0.64	0.64	0.54	0.54	0.54
Power	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9
TOTAL SAMPLE SIZE (hospitals)	36	110	1138	44	136	1418	54	166	1752	58	180	1896

Anticipated incidence rate (intervention) (Ie) = The anticipated daily average incidence rate in the experimental group with the outcome
Anticipated incidence rate (control) (Ic) = The anticipated daily average incidence rate in the control group with the outcome.
Effect size= % difference between the anticipated daily average incidence rate in the experimental group with the outcome and the anticipated daily average incidence rate in the control group with the outcome
Cluster size (m)= The anticipated average (or actual) cluster size
CV = The coefficient of variation of cluster-specific incidence rates, assumed constant over both the treatment and control groups. The bolded CV was estimated from Hospital Compare data.
Power = The desired level of power, recall power = 1 - type II error
Other parameters included (and held constant) but not shown in table:
Planned follow-up time (t) for the study (days)= 365
Allocation Ratio (AR) of patients in intervention group to control group=1
Alpha or the desired type I error rate= 0.05

Conclusions

- Sample size estimates for CRTs are most influenced by the CV and the expected effect size. Given the large sample size requirements, it is likely that many CRTs in hospital epidemiology are under-powered.
- We hope that these findings lead to more definitive CRTs in the field of hospital epidemiology that are properly powered and more studies reporting their ICC or CV.

Contact information:
Lyndsay O'Hara
University of Maryland
655 W. Baltimore St.
Baltimore, MD 21201
lohara@som.umaryland.edu

Author affiliations:
1. University of Maryland School of Medicine, Baltimore, Maryland, USA.
2. Division of Healthcare Quality Promotion, Centers for Disease Prevention and Control
3. Department of Internal Medicine, Carver College of Medicine, University of Iowa