Design of a Community Randomized HIV Prevention Trial in Botswana

Rui Wang, Ph.D.

Department of Population Medicine
Harvard Medical School and Harvard Pilgrim HealthCare Institute

November 3rd, 2016
CRT: Randomized in Groups or Clusters

Simple RCT

Cluster RCT

Individual randomization

Cluster randomization

Intervention group  Control group  Intervention group  Control group

Villages

Schools

Hospitals
CRTs Well-suited to Study HIV Prevention Strategies

- Direct effect: prevention of HIV in the individuals receiving the intervention
- Indirect effect: prevention of HIV in individuals ‘connected’ to those receiving the intervention – driven by sexual network
The Botswana Combination Prevention Project (BCPP)

- Combination prevention: enhanced HIV testing and counseling (HTC), prevention of mother-to-child transmission, enhanced linkage to care, and male circumcision (MC)
- Outcome: HIV incidence, from cohorts (20% of the population) followed longitudinally, over 3 year period
The Botswana Combination Prevention Project (BCPP)

- **Sponsor:** Centers for Disease Control and Prevention
  
  Botswana Harvard Aids Institute  
  Botswana Ministry of Health  
  Harvard T.H. Chan School of Public Health  
  Tebelopecle Voluntary Counseling and Testing Center

- **Collaborators:**

- **Principal Investigators:**
  
  Janet S Moore, PhD  
  Myron Essex, DVM, PhD  
  Refeletswe Lebelonyane, MD  
  Shenaaz El Halabi, MPH  
  Joseph Makhema, MBChB, FRCP  
  Shahin Lockman, MD, MS  
  Eric Tchetgen Tchetgen, PhD  
  Molly Pretorius Holme, MS  
  Pam Bachanas, PhD  
  Tafireyi Marukutira, MD

  Centers for Disease Control and Prevention  
  Harvard School of Public Health  
  Botswana Ministry of Health  
  Botswana Ministry of Health  
  Botswana Harvard Aids Institute  
  Harvard T.H. Chan School of Public Health  
  Harvard T.H. Chan School of Public Health  
  Harvard T.H. Chan School of Public Health  
  Centers for Disease Control and Prevention  
  Centers for Disease Control and Prevention

- **ClinicalTrials.gov Identifier:** NCT01965470
Botswana Communities with Sizes 3,000 – 15,000
Sample Size Calculation for CRT

- Must take into account possible correlation of outcomes within randomized units

- For continuous, binary, or count endpoints:
  - Formulas based on intraclass correlation ($\rho$) (Donner and Klar 2000) or coefficient of variation ($k$) (Hayes and Bennett 1999)
  - Estimation of power through simulations (Nicholas et al., 2012): using a generalized linear mixed model framework as the data generating model

- For survival endpoints: Xie and Waksman (2003), Antje Jahn-Elimermacher et al. (2013)
Challenges in the Design of HIV Prevention Trials

- Require information on the magnitude of intervention effects and the HIV incidence rate in the control group

- Accurate estimates of $\rho$ or $k$ are difficult to obtain
  - Exam the required sample size for various plausible values
  - Bayesian approach to incorporate the use of prior opinion
  - Recalculation of sample size using an internal pilot study

- Cross-contamination of intervention and control clusters
  - Expect a considerable fraction of sexual relationships to be with members of other communities
  - Outcomes depend on the treatment of other clusters due to mixing
  - Ensuring the clusters are sufficiently distant from each other by some metric (for example, geographic) can help, but not feasible in the BCPP
Design of the BCPP

- Develop an agent-based network/epidemic simulation model to
  - Simulate the intervention effect
  - Simulate the intraclass correlation (or the coefficient of variation)
  - Incorporate contamination between clusters
  - Assess the impact of contamination on intervention effect
Model Intervention Impact on HIV Spread over 3 Years

- Generate sexual networks then propagate disease spread on these networks

- Community characteristics
  - Sexual network characteristics (including mixing between communities)
  - Varying coverage level for different prevention modalities
  - Population sizes

- Individual characteristics
  - Disease progression
  - Transmission risk
  - Condom use
  - Linkage to care
  - Circumcision status
Model Overview

**Input**
- Degree Distribution (Likoma Island)
- Mixing between Communities
- Relationship Durations (Mochudi)

**Initial Conditions**
- HIV Prevalence
- % on ART
- % of males circumcised
- Condom usage
- VL/CD4 distr.

**Biology**
- Viral Load / CD4 Trajectories
- Transmission Probabilities
- Transmission reductions

**Intervention**
- % annual testing
- % circumcision
- Linkage to Care

**Output**
- Annual Incidence
Network Construction

- Bipartite graph (Relationship only between genders)
- Two arms (control and treatment)
- Degree (number of partnerships) distribution based on data from Likoma Island
- Permit incorporation of user-specified uncertainty associated with network properties
Network Construction

- Use a Metropolis-Hastings algorithm: constrains the degree distribution by proposing only networks with the prescribed degree distribution, and the accept-reject probability ensures the proportion of mixing is consistent with the target.

From Static Network to Dynamic Network

- Relationship durations, \( d \), are drawn from a survival distribution estimated from self-reported data from the Mochudi study.

- A start date is drawn from a uniform distribution on the interval from start of study minus \( d \) to end of study.
Simulation of the Disease Epidemic: Initial Conditions

- HIV prevalence: 24.8%
- VL/CD4 distribution: based on data from a household survey in Mochudi
- % on ART among eligible subjects (CD4 < 350): 60.9%
- Condom use: 40%
- % of males circumcised: 12.7%
Simulation of the Disease Epidemic: Transmission

- Disease progression based on data from Botswana/Durban incidence cohort

- Impact of viral load category on transmission risk: Quinn et al., 2000; sensitivity analysis: Attia et al., 2009, and Lingappa et al., 2010

- Reduction in transmission risks associated with knowing infection status: 30%

- Reduction in transmission risks associated with condom use: 85%

- Reduction in HIV acquisition risks for circumcision: 60%
Evaluation of Intervention Effect

- Randomly pick 20% of the population in each community to form the incidence cohort.

- Incidence cohort are tested annually for HIV infection.

- Outside of incidence cohort are tested with probabilities set to be the specified coverage levels for testing.

<table>
<thead>
<tr>
<th></th>
<th>Enhanced Standard of Care</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HTC</td>
<td>MC</td>
</tr>
<tr>
<td>Baseline</td>
<td>37%</td>
<td>12.7%</td>
</tr>
<tr>
<td>End of year 1</td>
<td>37%</td>
<td>31.4%</td>
</tr>
<tr>
<td>End of year 2</td>
<td>45%</td>
<td>50.0%</td>
</tr>
<tr>
<td>End of year 3</td>
<td>52%</td>
<td>60.0%</td>
</tr>
</tbody>
</table>
Difference in Cumulative Incidences and Mixing Levels

Reflecting the effect of cross-contamination on the randomized effects
Projected 3-Year Cumulative HIV Incidence

Mixing levels in the range of 15% to 25%, based on results from 1500 pairs of communities.

<table>
<thead>
<tr>
<th></th>
<th>Standard of Care</th>
<th>Intervention</th>
<th>% Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cumulative Incidence</td>
<td>Cumulative Incidence</td>
<td></td>
</tr>
<tr>
<td>End of Year 1</td>
<td>1.74%</td>
<td>1.42%</td>
<td>18.4%</td>
</tr>
<tr>
<td>End of Year 2</td>
<td>2.98%</td>
<td>1.99%</td>
<td>33.2%</td>
</tr>
<tr>
<td>End of Year 3</td>
<td>3.93%</td>
<td>2.34%</td>
<td>40.5%</td>
</tr>
</tbody>
</table>
Simulated Coefficient of Variation

- Assign both communities to SOC, calculate a coefficient of variation for each pair, then take the average across many pairs.

- The simulated $k$ is 0.08.

- Provide a lower bound: all clusters are assumed to have the same population sizes, initial conditions, and rates of disease progression for infected subjects.

- Consider a range of values from 0.08 to 0.35.
Sample Size and Study Power

\[ c = 2 + \left( z_{\alpha/2} + z_\beta \right)^2 \frac{\pi_0 (1 - \pi_0) / m + \pi_1 (1 - \pi_1) / m + k_m^2 (\pi_0^2 + \pi_1^2)}{(\pi_0 - \pi_1)^2} \],

- \( c \): number of clusters per treatment arm
- \( \pi_0 \): proportion of subjects who reach endpoint in SOC arm
- \( \pi_1 \): proportion of subjects who reach endpoint in intervention arm
- \( m \): number of sampled individuals within each cluster
- \( z_{\alpha/2} \) and \( z_\beta \): usual upper tail normal probabilities
- \( k_m \): coefficient of variation in true proportions between clusters within matched pairs in the absence of intervention
Sample Size Needed with 90% Power

- Detect anticipated differences 3.93% vs. 2.34%
- 15 pairs of clusters; 500 cluster members:
  - 99% power for $k = 0.08$
  - 84% power for $k = 0.35$
### Sensitivity Analyses: Varying Rates of HTC, MC and Linkage to care

<table>
<thead>
<tr>
<th></th>
<th>Setting 1</th>
<th>Setting 2</th>
<th>Setting 3</th>
<th>Setting 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MC&lt;sup&gt;1&lt;/sup&gt;</td>
<td>HTC&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Linkage to Care</td>
<td>Varying all three</td>
</tr>
<tr>
<td>Baseline</td>
<td>SOC&lt;sup&gt;3&lt;/sup&gt;</td>
<td>Intervention</td>
<td>SOC&lt;sup&gt;3&lt;/sup&gt;</td>
<td>Intervention</td>
</tr>
<tr>
<td>Hold</td>
<td>12.7%</td>
<td>12.7%</td>
<td>37%</td>
<td>37%</td>
</tr>
<tr>
<td>End of Year 1</td>
<td>31.4%</td>
<td>46.4%</td>
<td>37%</td>
<td>70%</td>
</tr>
<tr>
<td>End of Year 2</td>
<td>31.4%</td>
<td>46.4%</td>
<td>37%</td>
<td>70%</td>
</tr>
<tr>
<td>End of Year 3</td>
<td>31.4%</td>
<td>46.4%</td>
<td>37%</td>
<td>70%</td>
</tr>
<tr>
<td>3-Year Cumulative Incidence</td>
<td>4.07%</td>
<td>2.42%</td>
<td>4.06%</td>
<td>2.59%</td>
</tr>
<tr>
<td>Power k=0.3</td>
<td>91%</td>
<td>82%</td>
<td>89%</td>
<td>87%</td>
</tr>
</tbody>
</table>

The planned sample size achieves >80% power for all the settings considered here for a $k$ value as large as 0.3.
Sensitivity Analyses: Varying Projected Treatment Effects and Rates of Losses to Follow-up

For the planned sample size and a $k$ of 0.25, the study has $>80\%$ power to detect a reduction of 34% in the cumulative incidence even with 20% losses to follow-up
Selected Communities in the BCPP
Modeling Assumptions

- Does not incorporate different types of sexual relationships (e.g., regular or casual), with different frequencies of sex and probability of condom usage

- Does not target concurrency metrics

- Assigns initial infection status randomly among the population. Did not take into account potential correlation between HIV status and network properties

- Assumes independence of knowledge of HIV infection status and sexual practice

- Closed cohort
Empirical data were limited to inform the choice of input parameters
- Sexual network based on data from Likoma island
- Disease progression from the Botswana/Durban incidence cohort \( (n = 77) \)

Model estimates for incidence of the SOC communities similar to the UNAIDS estimates

Makes use of information from a wide variety of sources regarding biology and behavior information

Models/Parameters can be updated reflecting study experience, reflecting actual coverage levels and changes in Botswana national treatment guidelines over time
Acknowledgment

- Victor De Gruttola, Max Essex, Ravi Goyal, Quanhong Lei
- Colleagues on the Botswana study team
- Funding sources: NIH R37 AI51164, R01 AI24643, R01 AI083036
- Thank you for your attention