PRAGMATIC TRIAL DESIGN IN COMMUNITY SETTINGS

Ying Zhang
Department of Biostatistics
College of Public Health
University of Nebraska Medical Center

November 4, 2021
GP IDeA-CTR Community Engaged Research Institute
OUTLINE

• Explanatory Randomized Clinical Trials (E-RCT) vs Pragmatic Randomized Clinical Trials (P-RCT)
• PRECIS-2 tool for P-RCT
• Statistical Considerations for P-RCT
• Challenges of P-RCT in Community Settings
• Study Designs with Cluster Randomization
## E-RCT

- Originally designed for a new drug development-Phase III clinical trial
  - Efficacy
  - Safety
- Patient population is very **HOMOGENEOUS!**
  - Inclusion criteria
  - Exclusion criteria
- Double-blind randomization
  - Randomly assign treatment and placebo to trial participants
    - Investigators blinded
    - Participants blinded
  - Remove potential bias in assessing efficacy
  - Causal inference for treatment efficacy
• The study is often carried out in multiple sites
  ➢ Sites with many years experience in conducting clinical studies
  ➢ Centralized study protocol
    ✓ Participant's recruitment plan
    ✓ Informed consent
    ✓ Clinical visit schedules
    ✓ Testable scientific hypotheses
    ✓ Clearly defined study endpoint(s)
    ✓ Data collection via standard electronic case report forms (eCRFs)
    ✓ Protocol deviation
  ➢ Comprehensive statistical analysis plan (SAP)
    ✓ Sample size and power
    ✓ Interim analysis plan - efficacy and futility
    ✓ Safety monitoring plan
• The study is usually conducted in an artificial environment (Ideal Scenario by design) following Good Clinical Practice Guidelines
  ➢ Focusing on maximizing internal validity (prevention of bias)
  ➢ Does not necessarily apply to real-world scenario

• Limitations of E-RCT
  ➢ Slow – A systematic review suggests that as many as 50% of trials fail to recruit to the target and of those that are able to do so, 50% exceed their planned recruitment period (McDonald et al., 2006)
  ➢ Expensive – Median cost of $21.4 M from 726 pharmaceutical interventional trials (Martin et al., 2017)
  ➢ Findings are usually lack of real-world evidence for treatment efficacy
    ✓ E-RCTs usually study the effectiveness of a single treatment versus one placebo, delivered to carefully selected populations under ideal conditions and comparative effectiveness between multiple treatments cannot be derived, which makes it difficult to translate results to the real-world setting
    ✓ Findings from E-RCT can rarely yield implication in practice
    ✓ A study shows only 14% research findings in clinical research have led to widespread changes in day-to-day health care (Balas et al., 2000)
P-RCT

- P-RCTs aim to provide information on the relative merits of real-world clinical alternatives in routine care (Dal-Ré et al., 2018)
  - Focusing on maximizing external validity – generalizability of the results to many real-world settings
  - Preserving as much internal validity as possible

- Core characteristics of P-RCT
  - Questions from and important to stakeholders
  - Diverse representative populations
  - Multiple, heterogeneous settings
  - Multiple outcomes important to decision and policy makers
  - Comparison conditions are real-world alternatives, not a placebo or no treatment

- P-RCT is the main study design for conducting comparative effectiveness research (CER)
• Differences between E-RCT and P-RCT

Questions to be asked before participating in any clinical trial

<table>
<thead>
<tr>
<th></th>
<th>E-RCT</th>
<th>P-RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is it important to us?</td>
<td></td>
<td>You and you colleagues help formulate the research Questions</td>
</tr>
<tr>
<td>Can we do it here?</td>
<td></td>
<td>The study is built around your normal health care operations</td>
</tr>
<tr>
<td>Will it take us more times?</td>
<td></td>
<td>Flexible study protocols minimize intrusion in your daily workflow</td>
</tr>
<tr>
<td>Will it help our patients?</td>
<td></td>
<td>The study’s explicit goal is to give you evidence that improves patient care &amp; clinical decision making</td>
</tr>
</tbody>
</table>
Study elements need to be determined when performing any clinical trial.

<table>
<thead>
<tr>
<th>Goals</th>
<th>E-RCT</th>
<th>P-RCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>To determine causes and effects of treatment (Causal Inference)</td>
<td>To improve practice and inform clinical &amp; policy decisions</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td>Test the intervention against placebo using rigid study protocols &amp; minimal variation</td>
<td>Tests two or more real-world treatments using flexible protocols &amp; local customization</td>
</tr>
<tr>
<td>Participants</td>
<td>Highly defined &amp; carefully selected</td>
<td>More representative because eligibility criteria are less strict</td>
</tr>
<tr>
<td>Measures</td>
<td>Require data collection outside routine clinical care</td>
<td>Brief and designed so data can be easily collected in clinical settings</td>
</tr>
<tr>
<td>Results</td>
<td>Rarely relevant to everyday practice</td>
<td>Useful in everyday practice, especially clinical decision making</td>
</tr>
</tbody>
</table>
Because the value of “real-world” evidence is highly appreciated, labeling a RCT as “pragmatic” is almost a badge of honor (Dal-Ré et al., 2018)

- 615 trials were labelled “pragmatic” from 1977-2017
- 89 of the 615 trials were RCT on medicines
- 32 of the 89 RCTs (36%) should not be termed as “pragmatic”
  - 5 were for investigational medicines prior to licensing
  - 16 were single-center
  - 4 used multiple placebo
  - 14 used a single placebo

RCT should be considered as a continuum and E-RCT and P-RCT are currently widely regarded as the two extremes in this continuum.
• PRECIS stands for **pragmatic-explanatory continuum Indicator Summary** originally developed in 2005-2008 by 25 international trialists and methodologists (Thorpe et al., 2009)
  - 10 domains
  - Cited over 300 times by end of 2014 since 2009
  - Some weaknesses identified
    - No rating scale
    - Duplication in some domains
    - Need better guidance
    - Not validated

• PRECIS-2 was developed in 2015 with the help of over 80 international trialists, clinicians, and policy makers (Loudon et al., 2015)
  - 9 domains
  - Rating scale – 5-point Likert continuum
  - Validated
  - Detailed guidance provided
• PRECIS-2 Wheel, adapted from Loudon et al. (2015)

- **1- Very explanatory**
- **2- Rather explanatory**
- **3- Equally pragmatic and explanatory**
- **4- Rather pragmatic**
- **5- Very pragmatic**
• **Eligibility criteria**—Who is selected to participate in the trial?
  
  ➢ A highly pragmatic approach to eligibility criteria would be to include in the trial anyone with the condition of interest who is likely to be a candidate for the intervention if it was being provided in usual care for the condition.
  
  ➢ PRECIS-2 Score
    
    ✓ **Inclusion criteria**—All patients age ≥ 15 with suspected asthma were included in the study, based on a crude clinical diagnosis of asthma, as the standard respiratory testing machinery was not available in a rural setting.
    
    ✓ **Inclusion criteria**—Patients with confirmed diagnosis: ≥ 16 years of age with unilateral facial nerve weakness of no identifiable cause who presented to primary care or an emergency department and could be referred to a collaborating otorhinolaryngologist < 72 hours after the onset of symptom; excluding patients with pregnancy, breast feeding, uncontrolled diabetes, peptic ulcer disease ...

• **Recruitment**—How are participants recruited into the trial?
  
  ➢ The most extreme pragmatic approach to recruitment would be to do this in usual care so that only the people who attend a clinic with the condition of interest are recruited after they present on their own behalf without any overt recruitment effort.
  
  ➢ PRECIS-2 Score
    
    ✓ For a diabetes control program, investigators search medical record systems for eligibility participants and them mailing invitation letters.
    
    ✓ Investigators solicitate people who visit usual care clinical with BMI ≥ 25 to join a weight loss program.
• **Setting** – Where is the trial being done?
  ➢ The most extremely pragmatic approach to setting would be to do the trial in an identical setting to which investigators intend the results to be applied.
  ➢ **PRECIS-2 Score**
    ✓ A clinical trial of manual physical therapy versus corticosteroid injection to treat shoulder impingement in a specialized center, Madigan Army Medical Center, USA.
    ✓ A clinical trial of studying ibuprofen, paracetamol, and steam for patients with respiratory tract infection was taken placed in primary care where patients usually go for advice and treatment of the common cold in UK.

• **Organization** – What expertise and resource are needed to deliver the intervention?
  ➢ A highly pragmatic design would aim to slot the intervention into the usual organization of care for the condition of interest, making use of no more than the existing healthcare staff and resources in the setting
  ➢ **PRECIS-2 Score**
    ✓ A clinical trial for early lens extraction and intraocular lens implantation to treat glaucoma was conducted in a usual clinic with same number of staff and fully qualified ophthalmologists and without additional training and resources.
    ✓ In 1994, a network of acute respiratory distress (ARDS) was established consisting of 10 academic centers and 75 intensive care units, recruiting additional research staffs devoted to the study, and using equipment that was planned for at the trial design stage.
• **Flexibility (delivery)** – How should the intervention be delivered?
  ➢ The most pragmatic design approach to delivery flexibility would leave the details of how to implement the intervention up to providers, in other words, what happens in usual care.
  ➢ PRECIS-2 Score
    ✓ In a cognitive behavioral therapy (CBT) trial for depression, the therapy was delivered by selected experts who received regular training and supervision and no specific protocol for timing or co-interventions was developed.
    ✓ In an elective caesarean section syntocinon infusion trial, a detailed protocol was developed with protocol violation recorded in self-reported case form, and specific direction for co-interventions and complication or side effects management.

• **Flexibility (adherence)** – What measures are in place to ensure participants adhere to the intervention?
  ➢ A highly pragmatic design approach would allow for full flexibility in how end user recipients engage with the intervention.
  ➢ PRECIS-2 Score
    ✓ In a music therapy trial to support communication with autistic children, the therapy sessions were all individuals based on interaction with child and allowed for range of responses to the intervention.
    ✓ In a drug trial, it might withdraw participants if they failed to take more than 90% of their medication.
• **Follow-up** – How closely are participants followed up?
  - The most pragmatic position with regard to follow-up would be to have no more follow-up of recipients than would be the case in usual care.
  - **PRECIS-2 Score**
    - In a clinical trial of perioperative β blockade for patients undergoing infrarenal vascular surgery, clinical follow-up was conducted until patient left hospital (discharge or death) or until 30 days after surgery; more extensive data on ECG were collected; and unscheduled follow-up could be triggered by cardiovascular events.
    - In a self management course for chronic musculoskeletal pain, participants were twice a year to complete a postal validated questionnaire regarding two outcomes “pain related disability” and “pain intensity”, no additional follow-up was required for data collection.

• **Primary outcome** – How relevant is it to participants?
  - A pragmatic approach would be to select an outcome that is obvious importance from the patient’s perspective and would be also relevant to commissioners of care, the people who decide whether to implement the intervention based on its results.
• **Primary analysis** – To what extent are all data included?

  - The pragmatic approach to the analysis would typically be an intention-to-treat analysis using all available data.
  - The most explanatory approach would be to use the “as treated analysis” or “by protocol analysis” as the primary analysis, in which only those patients who actually received and did not receive the intervention would be analyzed in the intervention and control groups, respectively, irrespective of their initial randomized group allocation.
• Guidance in using PRECIS-2: four steps

➢ Step 1 – What design are you taking?
  ✓ Aiming to take an explanatory approach to answer question, “Can this intervention work under ideal conditions?”
  ✓ Aiming to take a pragmatic approach to answer the question, “Does this intervention work under usual condition?”

➢ Step 2 – Consider your trial design choices for each of the PRECIS-2 domains

➢ Step 3 – Score 1 to 5 for the choice made in Step 2 and mark on the PRECIS-2 wheel

➢ Step 4 – Review the PRECIS-2 wheel
  ✓ Review the score from the 9 domains for your design choice on the PRECIS-2 wheel to see whether they will produce a trial that support the overall aim you identified for your trial in Step 1.
  ✓ Go back to Step 2 and modify your design choices if necessary

• PRECIS-2 - Home Page for training
STATISTICAL CONSIDERATIONS FOR P-RCT

When planning P-RCT, several important design and analysis aspects should be carefully considered (Gamerman et al., 2019)

- Randomization
  - Patient level – standard in clinical trials in general
  - Cluster level – clinic, doctor, community site etc.
  - Cohort multiple control randomized studies
    - A cohort of patients is followed over time
    - A subgroup of the cohort is selected and then a random group of patients within the specific subgroup is invited to participate in a trial and to receive a new treatment
    - The patients in the random group are then compared with the patients from the subgroup who were not selected and were treated with standard of care
    - The process can be repeated over time for conducting different trials within changing subgroups within the cohort
    - This type of studies allows brand new therapies to be investigated in timely manner once they become available
    - Represent real-world population subgroups
• **Blinding** - Blinding is a useful strategy that can minimize the impact of subjective factors on the trial outcome. However, in many P-RCT, particularly in community settings, study participants or investigators cannot be blinded, several strategies may be used to minimize the potential subjective bias.

  ➢ Only objective endpoints should be used for primary and key secondary endpoints
  ➢ When event- or outcome-based endpoints are used, adjudication by blinded medical experts is recommended
  ➢ The statistical/data analyst may be blinded during the trial conduct, during the statistical analysis plan development, and beyond for as long as possible
  ➢ **Cluster randomization** may be considered and can help if blinding of individuals is not reasonable
• **Capture of Endpoint** – In P-RCT with a standard clinical care, clinical endpoints may not be manually recorded by study investigators

  - The endpoint of interest may not be precisely or fully observed
  - The endpoint information would be missing if a patient does not return to the healthcare system during the study visit times
  - EHR often cannot be automatically annotated precisely
    - Diagnostic codes are often imprecise with varying degree of accuracy
    - Outcome information may be documented in the narrative notes
  - Natural language processing (NLP) is an imperfect tool for extracting information related to a condition of interest
  - The naïve use of imprecise and partially missing endpoint information may lead to a bias in the treatment effect assessment
• **Data Augmentation**
  - Mobile devices or online portals can be used to capture the endpoint as patient-reported outcomes, if missing data due to lack of information in the EHR occur.
  - Second sampling – a small number of local sites to additional collect endpoint data using traditional mechanism as in standard RCT as validation data for assessing and correcting for potential biases induced by mission or imprecise data on the endpoint, some fine methodologies need to be developed.
  - To overcome the challenges of imprecise endpoint information extraction and the infeasibility of large-scale manual chart review, one employs machine learning algorithms to predict the outcome using partially labeled data.

• **Statistical Analysis**
  - Due to the potential lack of blinding, patients may have different patterns of missingness in their endpoint across different treatment groups, standard complete case analysis may result in biased treatment effect estimates despite the randomization. Causal inference tools such as propensity score adjustment, marginal structural modeling, and double robust methods should be considered.
  - Due to the imprecision in the endpoint derived from various algorithms, the inference tools that incorporate measurement errors in the outcomes should be employed to correct for the potential bias.
A pragmatic randomized clinic trial of weight loss maintenance in adults (WILMA) was designed in England in 2010 aiming to deliver the intervention in non-National Health Service community settings (NHS), to an obese adult population who had already achieved a 5% weight loss with the scientific promises that the level of weight loss has been shown to be associated with improved cardiovascular disease risk factors. It was a three-arms P-RCT

1. Intensive intervention arm – 12-month aggressive intervention
2. Less intensive intervention arm – 12-month normal intervention
3. Control arm – information packet and usual care

The follow-up assessments were planned for 6, 12, 24, and 36 months post randomization. However, the recruitment targets were not achieved, the originally designed trial was closed in Sept. 2012 and converted to a feasibility trial, which was completely closed in January 2014.
Study Design

• Sample size: 950 adults

• Age: 18-70, very pragmatic!

• Recruitment period: 18 months from October 2010 to April 2012

• Inclusion criteria: BMI (≥30 kg/m²) in the past 12 months and intentional weight loss (at least 5% of their body weight) during the same period, pragmatic.

• Recruitment routes
  - Route 1 - achieved a 5% weight loss and were able to provide independent verification of this
  - Route 2 – no yet achieved 5% weight loss, but
    - Attend a screening meeting with a research staff or
    - Self-screen by providing verified evidence of starting weight and subsequent 5% loss.
Recruitment Strategy – Community Setting

- General practitioner practice (GP)
- Exercise on Referral Schemes (Government funded schemes)
- Slimming world (SW, commercial weight loss program)
- Gyms and other community settings
- Identify participants through patient identification centers (PICs)
- Well-trained recruiters
  - Training sessions
  - Study materials
  - Telephone contact and support
  - Newsletters
- Monetary incentive for GP
- Press release and new items in SW magazine and on Websites
- Advertisements and flyers in Gyms and other community settings
Challenges in Implementation

- Delays in opening for recruitment
  - Difficulties in obtaining consensus on a standard treatment for obesity
  - A great deal of regional variation in the interpretation of attribution and liability of excess treatment cost (ETC)-Complex interventions often incur costs that cannot be clearly assigned, which are not easily reconciled
  - PIC classification in some regions in England was not accepted by comprehensive local research networks and full site approval had to be obtained in these areas
  - There were discrepancies in the interpretation of the roles and responsibilities of research network staff and their capacity to support the study, which largely impacted on the feasibility of delivery and cost of the study
• Recruitment

➢ Time delayed
➢ Recruitment route was not happened as expected

<table>
<thead>
<tr>
<th>Recruiter</th>
<th>Route 1 EOI</th>
<th>Route 1 Recruited</th>
<th>Route 2 EOI</th>
<th>Route 2 Recruited</th>
<th>Total EOI</th>
<th>Total Recruited</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP/nurse</td>
<td>91</td>
<td>51 (56.0%)</td>
<td>830</td>
<td>15 (1.8%)</td>
<td>921</td>
<td>66 (7.2%)</td>
</tr>
<tr>
<td>SW</td>
<td>65</td>
<td>47 (72.3%)</td>
<td>17</td>
<td>0 (0.0%)</td>
<td>82</td>
<td>47 (57.3%)</td>
</tr>
<tr>
<td>Exercise on referral</td>
<td>24</td>
<td>19 (79.2%)</td>
<td>133</td>
<td>3 (2.3%)</td>
<td>157</td>
<td>22 (14.0%)</td>
</tr>
<tr>
<td>Other/advertising</td>
<td>61</td>
<td>34 (55.7%)</td>
<td>63</td>
<td>1 (1.6%)</td>
<td>124</td>
<td>35 (28.2%)</td>
</tr>
<tr>
<td>Total</td>
<td>241</td>
<td>151 (62.7%)</td>
<td>1043</td>
<td>19 (1.8%)</td>
<td>1284</td>
<td>170 (13.2%)</td>
</tr>
</tbody>
</table>

✓ GP practices could reach a large number of individuals but unrefined database search tool and/or a lack of information to assess eligibility meant that the majority fell into Route 2

✓ For exercise on referral schemes, although many who attend were overweight or obese, they were not necessarily interested in weight loss
Lack of PIC activity

✓ Engagement from exercise on referral and SW staff was generally quite low despite encouragement from management
✓ Drop out of practitioners required recruitment of staff and training and refresher training added to the resource burden to the team

• Revised recruitment plan – plan divided into sections for maximizing the impact/success of recruitment via PICs, advertising and other routes

➢ High priorities

✓ Targeting only those who were more likely to be able to provide the information of 5% weight loss
✓ Sustaining motivation of PIC staff by increasing use of newsletters and telephone contact
✓ Introducing compensation and financial incentives for exercise on referral and SW staff by for top recruiters
✓ Making SW high priority PICs over GP practices and exercise on referral, since their clients proved to be more likely to have lost weight and, crucially, be able to provide evidence
✓ Widespread use of posters and email advertisement as a low cost, time efficient and potentially far-reaching option

• Medium priorities

✓ Social media utilization
✓ Presentation of the trial at dietetics meetings, to gym and fitness club managers
Lessons Learned

• Never underestimate the amount of time required to gain governance approvals, cost and network support on recruitment and its impact on the workload of the trial manager and study team.
• Develop a detailed recruitment strategy early on and consider recruitment from different sources.
• Conduct a feasibility study.
STUDY DESIGNS WITH CLUSTER RANDOMIZATION

• Cluster randomized trial (CRT) is a common pragmatic trial design in community settings when individual randomization is not possible for various reasons
  ➢ Intervention can only be administrated on a facility-wide or community-wide scale
  ➢ Individuals within a cluster are correlated and the intra-cluster correlation needs to be incorporated into power calculation and data analysis

• CRT has several variations in design (Hemming and Taljaard, 2016)
  ➢ Parallel CRT: Immediate randomized with intervention
  ➢ CRT-BA: Parallel CRT with a control phase before randomization
  ➢ Stepped Wedge-CRT (SW-CRT): All clusters start with control and gradually roll into intervention
Three Types of CRT

<table>
<thead>
<tr>
<th>Site</th>
<th>Year in study</th>
<th>Site</th>
<th>Year in study</th>
<th>Site</th>
<th>Year in study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>1</td>
<td>O X</td>
<td>1</td>
<td>O X X X X X</td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td>2</td>
<td>O X</td>
<td>2</td>
<td>O X X X X X</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>3</td>
<td>O X</td>
<td>3</td>
<td>O O X X X X</td>
</tr>
<tr>
<td>4</td>
<td>O</td>
<td>4</td>
<td>O O</td>
<td>4</td>
<td>O O X X X X</td>
</tr>
<tr>
<td>5</td>
<td>O</td>
<td>5</td>
<td>O O</td>
<td>5</td>
<td>O O O X X X</td>
</tr>
<tr>
<td>6</td>
<td>O</td>
<td>6</td>
<td>O O</td>
<td>6</td>
<td>O O O X X X</td>
</tr>
</tbody>
</table>

Intervention schedules for parallel-CRT, CRT-BA and SW-CRT

- “O” represents control of existing intervention
- “X” represents intervention phase
Features of the CRTs

• The parallel CRT is the commonly adopted CRTs
  ➢ It is a natural design for measuring community effect
  ➢ It is less efficient than individual-based RCT because of intra-cluster correlation

• The CRT-BA allows to measure “change from baseline” to remove possible confounding effects in assessing intervention
  ➢ It can be regarded as a uni-directional crossover trial
  ➢ Maybe unethical

• The SW-CRT removes the ethical concern of CRT
  • It remains the essence of uni-directional crossover design
  • It is more practically feasible in terms of implementation of cluster randomization
Sample Size Determination for CRTs

Main idea: The standard approach in sample size calculation for CRTs consists of:

- Calculate the sample size that would be needed if individual-level RCT were to be conducted ($N_I$)
- The total number of sample size for a CRT is:
  \[ N_T = N_I \times \text{DE} \]
  where DE is the “design effect” depended on the type of CRT (Woertman et al., 2013; Biao et al., 2015; Hemming and Taljaard, 2016)
**Design Effect**

- **Parallel –CRT:**
  
  \[ DE_1 = 1 + (N - 1) \times \rho \]

- **CRT-BA:**
  
  \[ DE_2 = 2 \times [1 + (N - 1) \times \rho] \times \left[1 - \left(\frac{N \times \rho}{1 + (N - 1) \times \rho}\right)^2\right] \]

- **SW-CRT:**
  
  \[ DE_3 = J \times \frac{1+\rho \times (JN-1)}{1+\rho \times \left[\frac{(J+1) \times N}{2}\right] - 1} \times \frac{3 \times (1-\rho)}{2 \times (J-1 - \frac{1}{J-1})} \]

where

- \( \rho \): Intra-class correlation (ICC)
- \( N \): Cluster size at each step
- \( J \): Number of steps in SW-CRT
Sample Size Algorithm for SW-CRT

- Given ICC ($\rho$), number of steps ($J$), and the cluster size ($N$), the number of clusters ($I$) can be easily determined by

$$I = \frac{N_I \times DE}{JN}$$

- Given ICC ($\rho$), number of steps ($J$), and the number of clusters ($I$), the cluster size ($N$) needs an iterative algorithm to determine (Woertman et al., 2013)

  - Make an initial guess of $N$, $N^*$
  - Calculate the number of clusters $I^*$
    - If $I^* < I$, $N^*$ needs to be decreased
    - If $I^* > I$, $N^*$ needs to be increased
  - This process should be repeated until a cluster size $N$ is found for which $I^*$ is close to equal to $I$; but not larger than $I$
Statistical Properties of SW-CRT

• Because of positive intra-cluster correlation (ICC), the total sample size of SW-CRT is generally larger than that of an individual-level RCT

• When ICC is small, parallel-CRT tends to be more efficient

• When ICC is small, but the cluster size is large, SW-CRT is almost as efficient as parallel-CRT

• When ICC is large (greater than 0.1), SW-CRT tends to be more efficient

• SW-CRT is relatively insensitive to ICC

• Optimal power will be achieved when only one cluster switches to the intervention at each step

• Failure to account for a potential time effect will artificially and grossly overestimate the power of a study with SW-CRT
CASE STUDIES

Study 1: Make Play Safe: A Virtual Really App Intervention to Improve Concussion Recognition and Reporting among Athletes Ages 9-12. (R01 Application)

- The major goal of this study is to test the efficacy of Make Play Safe (MPS), an innovative, theory-driven, two-component intervention aimed to improve concussion recognition and reporting among boys’ and girls’ soccer athletes ages 9-12 using virtual reality technology in conjunction with strategies that promote parent/coach-child communication about concussion.

Figure 3. Concussion Symptom Display

Figure 5. Make Play Safe
Study Aims

1. **Determine the efficacy of the intervention on soccer athletes’ concussion symptom recognition and reporting intentions.** *Hypothesis:* Intervention athletes will exhibit greater change scores on concussion symptom recognition-related attitudes, perceived behavioral control, and reporting intentions from pre- to post-season and will retain higher scores six months after the season ends than Control athletes.

2. **Determine the efficacy of the intervention on soccer athletes’ concussion report.** *Hypothesis:* Intervention athletes will have a higher rate of reporting a suspected or actual concussion (per 1,000 athlete exposures) to an appropriate adult throughout the soccer season than Control athletes.

3. **Determine if the efficacy of the intervention is mediated by the frequency of parents’/coaches’ communication.** *Hypothesis:* Intervention parents/coaches will demonstrate more frequent communication with athletes than Control parents/coaches, and more frequent child-parent/coach communication will be associated with greater changes in athletes’ concussion symptom recognition and reporting behaviors.
Study Design

- Two-arm, cluster RCT to test the efficacy of MPS with two-component intervention
  - A virtual reality (VR) app concussion education session
  - A weekly message notification program for parents and coaches

- Boy and girl soccer athletes ages 9-12 and their parents and coaches

- Athletes recruited from 8 youth soccer leagues in central Ohio
  - Two leagues will be enrolled at the beginning of each season and randomized into either the intervention or control group with randomization taken placed at the league level
  - Eight teams will be randomly selected from each participating league using a stratified random block of sex (boys vs. girls) and age group (9-10 vs. 11-2 years)

- Intervention athletes will receive the MPS VR session at the beginning of the soccer season
- Control athletes will receive the concussion education materials currently provided by their soccer league (standard care)

- Parallel RCT!
Data Collection Schedule

T1: Pre-season survey

Randomization

Intervention (8 teams/season)
- A VR session
- Weekly message notifications

Control (8 teams/season)

Bi-weekly survey

Primary Outcome 2:
Concussion reporting

Primary Outcome 1:
Concussion symptom recognition

T2: Post-season survey

T3: 6-month after the season ends

Figure 1. Overview of Study Design
Power Analysis

➢ **Primary endpoint** – Change in concussion knowledge score pre- and post-season

➢ **Pilot Study**
  ✓ Intervention increases the change in the score by 7.1%
  ✓ Standard deviation 8.4%
  ✓ Standardize effect size 0.85

➢ **Assumption**
  ✓ The intracluster correlation coefficient (ICC) < 0.9
  ✓ Attrition < 20%
  ✓ Each team is made by 15 athletes

➢ With 64 teams (960 athletes) in a cluster RCT, the study is **powered at 0.94** to establish the intervention efficacy if the standardized effect size as observed in the pilot study sustains in the study
Study 2: Improving update of HIV interventions in SIY in East Africa (R01 Application)

- **Aim:** to effectively link the street-involved youth (SIY) in East Africa who are HIV positive to initiate antiretroviral treatment (ART)
- **Intervention:** to use peer navigators (PN) for bridging the gap between the SIY population and health facilities by providing a supportive and knowledge resource for HIV testing and treatment to SIY
- **Study sites (clusters)**
  - Mid-sized towns: Mbarara and Bungoma
  - Cities: Kisumu and Mwanza
  - All in medium-high HIV prevalence areas on or near Lake Victoria
**Study Design:** Stepped-Wedge RCT

<table>
<thead>
<tr>
<th>Site</th>
<th>1-10</th>
<th>11-20</th>
<th>21-30</th>
<th>31-40</th>
<th>41-50</th>
<th>51-60</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>2</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>O</td>
</tr>
</tbody>
</table>

- **O:** Non-intervention period
- **X:** Intervention period
- A 10-month period for intervention sustainability study
Design Parameters

- Pilot study conducted for HIV in SIY in Eldoret showed
  - Prevalence of HIV positive among SIY is about 9.7%
  - Baseline rate of initiating ART among HIV positive SIY is 35%
  - Post PN intervention rate of initiating ART among HIV positive SIY is 60%
- Cluster size determination at each step with a normal test at significance level 0.05 powered at 0.9

<table>
<thead>
<tr>
<th>ICC</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td># of HIV+</td>
<td>18</td>
<td>17</td>
<td>15</td>
<td>13</td>
<td>11</td>
<td>9</td>
<td>7</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td># of SIY</td>
<td>186</td>
<td>155</td>
<td>155</td>
<td>135</td>
<td>114</td>
<td>93</td>
<td>73</td>
<td>42</td>
<td>21</td>
</tr>
</tbody>
</table>
REFERENCES


