Study Designs for Implementation and Hybrid Trials

C Hendricks Brown
Hendricks.brown@northwestern.edu
Director, Center for Prevention Implementation Methodology for Drug Abuse and HIV (Ce-PIM, NIDA P30DA027828)
Funding Acknowledgments

NIDA: Center for Prevention Implementation Methodology for Drug Abuse and HIV Ce-PIM (P30DA027828, Brown, Mustanski PD)

NIDA: Measuring Sustainment in Prevention Programs and Initiatives (R34DA037516, Palinkas PI)

NIAID: Third Coast Center for AIDS Research (P30AI117943, D’Aquila, Moskowitz, Bouris PD)

CDC: Chicago Center for Youth Violence Prevention (U01CE002712, Gorman-Smith, Brown PD)

NIMH: A Pragmatic Trial of Two Strategies for Implementing an Effective eHealth HIV Prevention Program (R01 MH118213 Mustanski PI)
Coauthors

Greg Aarons, UCSD
John Landsverk, OSLC
Larry Palinkas, USC
Carlos Gallo, Northwestern
Brian Mustanski, Northwestern
JD Smith, Northwestern
Nanette Benbow, Northwestern
Wouter Vermeer, Northwestern
Moira McNulty, U Chicago

Ce-PIM Collaborators
Outline

1. Purposes of Study Designs for Pre-Implementation, Implementation, and Hybrid Effectiveness/Implementation Designs

2. To Randomize or Not, and What to Randomize
   A. Non-Randomized Evaluations
      Within Site Designs
   B. Randomized Trials
      Between Site Designs
      Rollout Designs – Within and Between Site Designs

3. Statistical Power Considerations
   Types of Assignments, (Hypotheses, Random Effects to Consider)

4. Brief Mention
   A. Designs for “Doing Better” -- Engineering Approaches
   B. Designs for “Scaling Out”

5. Summary
Could a program work?

Does a program work?

Making a program work

Effectiveness studies

Preparation

Exploration

Implementation

Sustainment

Implementation Research

Implementation Practice

Real-world relevance

Local knowledge

Generalizable knowledge

Time

Traditional Translational Pipeline

Brown et al., ARPH 2017
Implementation Research Has a Different Emphasis Than Other Types of Research

Effectiveness

System to Support Adoption and Delivery w Fidelity

Intervention

Evaluate Health Outcomes

Implementation

System to Support Adoption and Delivery with Fidelity

Intervention

Carries over into what to Measure, what to Model, What to Test or Evaluate

Evaluate Quality, Quantity, Speed of Delivery

Effectiveness vs. Implementation

ISC3I
1. Purposes of Study Designs for Pre-Implementation, Implementation, and Hybrid Effectiveness/Implementation Designs

• **Pre-Implementation:**
  Examine Determinants (Barriers and Facilitators) to Customize Implementation Strategies

• **Implementation:**
  A. Is implementation working **HERE**, can it be improved? **Local Knowledge**
     Monitor Implementation to Provide **Actionable Feedback**
  
  B. Research Evaluation of the Implementation Strategy itself – **Generalizable Knowledge**
     Randomized Implementation Trial (designed for generalization)
     Is the Monitoring and Feedback Process Translatable to Other Situations?

• **Hybrid Effectiveness/Implementation Designs**
  Full or Partial Evaluation of Effectiveness and/or Implementation Questions
Three Types of Hybrid Effectiveness/Implementation Designs

Hybrid Type I: Test effectiveness, observe/gather information on implementation

Hybrid Type II: Test effectiveness, Test/Study implementation strategy

Hybrid Type III: Test implementation strategy, observe/gather information on effectiveness

Illustrations of Implementation Research Questions

• What are the primary barriers and facilitators of implementing PrEP in Local Health Departments?

• Under what conditions does implementation Strategy A work better, faster, more efficiently than Strategy B?

• What are the characteristics of research/service/community partnerships that are sustained over time?

Proctor et al. 2012
Implementation is Necessarily...

1. **Multilevel**
e.g., organizational leaders, “implementation agencies”, “Intervention agents”

2. **Systems Oriented**
Interaction of multiple components
e.g., service providers / community partners collaboration

3. **Dynamic**
Implementation process goes through stages
e.g., Stages of Implementation Completion* (Saldana, 2014 Imp Sci)

**Dialectic Process**

**Shared Implementation Mediation Model vs Equifinality**
Aarons et al., IS 2017

**Implications for Design:**
- *group-level assignment(s)* at organization or community level(s)
- Implementation Process is measurable as well as Implementation Outcomes
2. To Randomize or Not, and What to Randomize

Randomize to Improve Generalizability and Protection Against External Factors
   CAL-OH Randomized Head-to-Head Implementation Trial of MTFC – Brown et al., Imp Sci 2014

**Equipoise** no longer holds in implementation of EBPs
   Communities won’t stand for typical controls – Brown et al., Adm Pol MH 2012
   Randomize the ORDER of Implementation – Rollout Designs

Don’t reject randomization only on considerations of power alone – Brown et al., 2006 Clinical Trials, Brown et al., 2009 ARPH -- “Cumulative Designs”
2A. Non-Randomized Evaluations

A Single Site or All Sites Receive the Same Implementation (At the Same Time)
2A. Within Site Non-Randomized Designs

• Post Design
• Pre-Post Design
Post Design of an Implementation Strategy to Adopt a Novel Clinical/Preventive Intervention

TIME

Start                                               End

Clinical/Preventive Intervention Absent Present

Rapid Oral HIV Testing in Dental Practices, recommended by CDC Implementation Strategy

Full Time HIV Counselor

HIV Testing Rate
New HIV Positives Identified

Oni et al. 2010
http://online.liebertpub.com/doi/pdf/10.1089/apc.2010.0159
Pre-Post Design of an Implementation Strategy of an Existing Clinical/Preventive Intervention when delivered differently

Boeke et al., BMC Health Services Research, 2018

Changes in Patient linkage and retention in care for HIV in 20 Uganda health facilities

Implementation of training lay health workers in best practices for patient follow-up and counseling

Improved retention but not linkage
2B. Randomized Implementation Designs

Between Site Designs
Between and Within Site Designs == Rollout Designs
Between Site Design: **Head-to-Head Trial of Two Implementation Strategies of the Same Clinical/Preventive Intervention**

For Implementation, the **Program Delivery System**, rather than the Clinical/Preventive Intervention, is in the Foreground.

Between Site Implementation Designs

• Keep It Up! eHealth Randomized Implementation Trial – Mustanski et al.
• Compares 2 alternative strategies for large-scale implementation of KIU!
  • counties randomized to receive the same KIU! but through
    • CBOs
    • receive Direct to Consumer

RE-AIM
Primary Outcome: Index measuring Public Health Impact
Reach into County (C) involving MSM’s race, age prevalence of HIV, $P_{ic}$
Individual’s HIV risk reduction from Baseline: $R_{ic1} - R_{ic0}$ (CAS, STIs, PrEP usage)
Effectiveness in Reducing Risk: Between Site Comparison of
$$\text{PHI}_C = \sum_i P_{ic} \times (R_{ic1} - R_{ic0}).$$

Examine Variation in adoptions, adaptation, sustainment
Designs for Between Site Implementation Trials Involving Single, Stable, clinical/preventive interventions through the Traditional Translation Pipeline

• Post Design of an Implementation Strategy to Adopt a Novel Clinical/Preventive Intervention
• Pre-Post Design of an Implementation Strategy of a Clinical/Preventive Intervention that is already being delivered
• New Versus Usual Care Non-Randomized Implementation Designs
• New Versus Usual Care Randomized Implementation Designs
• Head-to-Head Randomized Implementation Trial Design
• Incomplete Block Designs for Comparison of Implementation Strategies
• Factorial Designs for Implementation
• Doubly Randomized, Two-Level Nested or Split Plot Designs for Testing Two Nested Implementation Factors
• Diffusion Designs for an Effective a Clinical/Preventive Practice, Program or Policy
• Dissemination Designs for an Effective a Clinical/Preventive Practice, Program or Policy
• Dosage Implementation Randomized Trial Design
Between and Within Site Randomized Implementation Trials == Rollout Design

- Sites get randomized twice
  - Which of (>=) 2 new implementation strategies they receive
  - What time they receive the new implementation strategy
51 Counties Randomized to both Implementation (Community Development Team CDT or Standard) and Timing

Randomize 51 Counties in CA and OH to Implementation Strategy and Time (Cohort) Randomized Roll-Out Design*

Brown et al., 2014 Imp Science
Wyman et al., 2015 Prev Sci
Landsverk et al., 2017 D&I Book
Stages of Implementation Completion (SIC)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Stage</th>
<th>Involvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-implementation</td>
<td>1</td>
<td>Engagement</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Consideration of feasibility</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Readiness planning</td>
</tr>
<tr>
<td>Implementation</td>
<td>4</td>
<td>Staff hired and trained</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Adherence monitoring processes in place</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Services and consultation begin</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Ongoing services, consultation, fidelity monitoring and feedback</td>
</tr>
<tr>
<td>Competency</td>
<td>8</td>
<td>Certification/competency</td>
</tr>
</tbody>
</table>

SIC: Milestone Attainment, Speed, Quality, and Quantity of Delivery

Chamberlain et al., 2011
Saldana et al., 2012, 2015, 2016, 2019
Between and Within Site Implementation Trial
Variations in Site & Time
Stepped Wedge : Steps

Stepped Wedge Design
11 Steps and 11 time periods

Time

Clinics (1 per time period)

Steps
Stepped Wedge Designs have Steps and 2 Full Wedges
Lots of Different Steps ...
There are **Lots of Different** Rollout Designs
Start Measuring Outcomes For All Sites at the Same Time: Useful When Blue = Implement as Usual

Hybrid Design measuring Effectiveness
Period Prior to Implementation Has No Information on Implementation Outcomes
Typically Little Additional Cost To Extend Sustainment Period
3. Statistical Power Considerations

Bad News: Power heavily dependent on number of Units Randomized

Good News: Small number of Rollout Times generally needed for high power, even with Learning Collaboratives involving non-independence
20 Independent Sites vs 4 Clusters of 5 Sites
Brown et al., 2006 Clin Trials
Stay Tuned to Ce-PIM for Rollout Design Power Program
4A. Designs for “Doing Better” -- Engineering Approaches and Simulation

Statistical Control Charts
SMART Designs
Monitoring:

**Primary Outcome** – Problematic if a “Low Base rate” outcome such as HIV incidence

**Key Hypothesized Change Factors**

Example: Training Staff to Recognized Suicidal Youth: Attitudes and Self-Reported Behaviors
Monitoring a Low Baserate Primary Effectiveness Outcome
Effect Size Attitudes Changed through QPR Training  Wyman et al., 2008

<table>
<thead>
<tr>
<th>Improvements from Training and Time Effect Size</th>
<th>Null</th>
<th>Low</th>
<th>Med</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of Warning Signs and QPR behaviors</td>
<td></td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attitudes about Suicide Prevention</td>
<td></td>
<td></td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>Self-Evaluation of Suicide Prevention Knowledge</td>
<td></td>
<td></td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>Knowledge of Clinical Resources</td>
<td></td>
<td></td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Efficacy to Perform Gatekeeper Role</td>
<td></td>
<td></td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>Reluctance to engage with suicidal students</td>
<td></td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Control Chart
Self Efficacy for Gate

Benneyan et al., 2003
Quality Improvement and Optimization: Taking an engineering perspective

• Working systematically toward development of an implementation strategy that meets specific criteria (e.g., cost, effectiveness)

Produce LOCAL KNOWLEDGE


SMART Designs for Adaptive Implementation Strategies
Kilbourne et al., Impl Sci 2014
http://www.implementationscience.com/content/pdf/s13012-014-0163-3.pdf

158 community outpatient clinics using Re-Engage to implement evidence-based programs to address mood disorders.

<table>
<thead>
<tr>
<th></th>
<th>6 Months</th>
<th>12 Months</th>
<th>18 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Re-Engage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Responsive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enhanced</td>
<td>Standard</td>
<td>Enhanced</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>Enhanced</td>
<td>Standard</td>
<td></td>
</tr>
</tbody>
</table>

“SMART” = Sequential Multiple Assignment Randomized Trial
4B. “Scaling Out” rather than “Scaling Up”
Aarons et al. Imp Sci 2017

Scaling Up: Delivery of an EBP to Same Population through Same Delivery System where evidence was attained

Scaling Out - “Borrowing Strength” when Adapting to Different Populations and/or Different Delivery Systems

When is “Off-Label Usage” Likely to Work?
Three Options for Evaluation of Scaling Out

• Don’t Bother to Evaluate and Hope for the Best

• Start Evaluation Completely Over
  • Time-intensive
  • Costly

• Borrow Strength from Existing Evaluations
  • Utilizing evidence from previous trials in combination with new evidence from scale-out trial
  • Accelerate & expand benefit
Evaluation Guided by How a Program Should Work: A Conceptual Theory of Mediation

Minimal monitoring of how a program should work involves repeated checking of fidelity and participation.

Could monitor OUTCOMES repeatedly, modifying implementation if outcome is not moving as it should.
Assuming a conceptual theory of mediation holds for scaling-out


Fonner et al., 2016, *AIDS*
Assuming a conceptual theory of mediation holds for scaling-out


<table>
<thead>
<tr>
<th>PrEP Adherence</th>
<th>RR of HIV Infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (&gt;70%)</td>
<td>0.30</td>
</tr>
<tr>
<td>Moderate (41-70%)</td>
<td>0.55</td>
</tr>
<tr>
<td>Low (&lt;= 40%)</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Fonner et al., 2016, *AIDS*
What Level of Evidence is Required in Each Component when EBI is Scaled Out: RE-AIM

<table>
<thead>
<tr>
<th>Level of Evidence</th>
<th>Implementation Fidelity</th>
<th>Intervention Fidelity</th>
<th>Reach &amp; Exposure</th>
<th>Adoption</th>
<th>Sustainment</th>
<th>Effect on Health Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>0: Minimal or no new empirical evidence</td>
<td>Not measured</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Proxy empirical evidence</td>
<td></td>
<td>Staff Self-Efficacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2: Direct empirical evidence</td>
<td></td>
<td>Milestone Attainment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: Full randomized hybrid trial</td>
<td></td>
<td>SIC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary: A Bestiary of Implementation Designs

Randomized?
Multi-Level?
Roll-Out?
Evolving Intervention?
Effectiveness?
Adapt to New Population?
Adapt to New Delivery System? (see references)
Implementation Designs: Community and Organizations are Much More Involved in Design Decisions and their Ownership

• Ethical responsibility
  Implementation ➔ longer in Equipoise so Controls

• Moral responsibility
  Social Justice: Implementation Science should improve:
  Health Equity
  Health Service Equity

Scientific Equity
  Brown et al., JAIDS 2013, Adm Pol MH 2012
  Perrino et al., Prev Sci 2015
  McNulty et al., Ethnicity & Disease, 2019
References on Designs


More on the Ce-PIM Website: [http://cepim.northwestern.edu/publications](http://cepim.northwestern.edu/publications)