External Validity: Why it Matters

Helen I. Meissner, Ph.D.
Office of Behavioral and Social Sciences, NIH
And
Cynthia A. Vinson, MPA
Division of Cancer Control and Population Sciences, NCI
OUTLINE OF TALK

• Background & Definitions
• Approaches to maximizing external validity
• Reporting external validity findings
Definitions

• **Internal Validity** – identifies causal relationships … in this study, the intervention made a difference in the outcome.

• **External Validity** – findings are true beyond the controlled limits of the study. “To what populations, settings, treatment variables and measurement variables can this effect be generalized?” (Campbell & Stanley, 1963)

Internal vs. External Validity

• What are the trade-offs of in maximizing internal or external validity?
“Where did the field get the idea that evidence of an intervention’s efficacy from carefully controlled trials could be generalized as THE best practice for widely varied populations and settings?”

L.W. Green

Green LW. From research to "best practices" in other settings and populations

Am J Health Behav 2001; 25:165-78
Maximizing External Validity
Chen HT. The Bottom-up approach to integrative validity: A new perspective for program evaluation *Evaluation and Program Planning* 2010; 33:205-14
Integrative validity model: “intervention . . . must be scientifically credible as well as relevant and useful in practice”

- **Internal validity**: extent to which evaluation provides objective evidence that intervention causally affects outcome

- **External validity**: Credible evidence that findings of effectiveness can be transferable from research setting to real-world setting

- **Viable validity**: credible evidence on the intervention’s real-world viability (e.g., whether it is practical, affordable, helpful, etc.)

Closing the Gap Between Research and Practice

• Bottom-Up Approach increases focus on external validity
  – Addresses how to make research relevant/generalizable/applicable in real medical and public health practice settings and policies
  – Addresses feasibility and cost-effectiveness when delivered in actual practice settings and with diverse populations
Question

• How might you use these approaches in developing your own studies?
External Validity

• A framework for closing the gap between research and practice
Purposes of RE-AIM

- To broaden the criteria used to evaluate programs to include elements of external validity
- To evaluate issues relevant to program adoption, implementation, and sustainability
- To help close the gap between research studies and practice by:
  - Suggesting standard reporting criteria
  - Informing design of interventions
  - Providing guides for program planners
Goal of RE-AIM Evaluation

Determine characteristics of interventions that can:

- Reach large numbers of people, especially those who can most benefit
- Be widely adopted by different settings
- Be consistently implemented by staff members with moderate levels of training and expertise
- Produce replicable and long-lasting effects (and minimal negative impacts) at reasonable cost

### Ultimate Impact of ‘The Magic Pill’

<table>
<thead>
<tr>
<th>Dissemination</th>
<th>Concept</th>
<th>% Impacted</th>
</tr>
</thead>
<tbody>
<tr>
<td>50% of Clinics Use</td>
<td>Adoption</td>
<td>50%</td>
</tr>
<tr>
<td>50% of Clinicians Prescribe</td>
<td>Adoption</td>
<td>25%</td>
</tr>
<tr>
<td>50% of Patients Accept Medication</td>
<td>Reach</td>
<td>12.5%</td>
</tr>
<tr>
<td>50% Follow Regimen Correctly</td>
<td>Implementation</td>
<td>6.2%</td>
</tr>
<tr>
<td><strong>50% of Those Taking Correctly Benefit</strong></td>
<td>Effectiveness</td>
<td>3.1%</td>
</tr>
<tr>
<td>50% Continue to Benefit After 6 Months</td>
<td>Maintenance</td>
<td>1.6%</td>
</tr>
</tbody>
</table>
The Moral of the Story

1. “Focus on the Denominator” (not just the numerator)

2. Each step of the dissemination sequence, or each “RE-AIM” dimension is important
## RE-AIM Guidelines for Developing, Selecting, and Evaluating Programs and Policies Intended to Have a Public Health Impact

<table>
<thead>
<tr>
<th>RE-AIM ELEMENT</th>
<th>GUIDELINES AND QUESTIONS TO ASK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>REACH</strong></td>
<td>Can the program attract large and representative percent of target population? Can the program reach those most in need and most often left out (i.e., the poor, low literacy and numeracy, complex patients)?</td>
</tr>
<tr>
<td>Percent and representativeness of participants</td>
<td></td>
</tr>
<tr>
<td><strong>EFFECTIVENESS</strong></td>
<td>Does the program produce robust effects across sub-populations? Does the program produce minimal negative side effects and increase quality of life or broader outcomes (i.e., social capital)?</td>
</tr>
<tr>
<td>Impact on key outcomes, quality of life, unanticipated outcomes and subgroups</td>
<td></td>
</tr>
</tbody>
</table>

### RE-AIM Guidelines for Developing, Selecting, and Evaluating Programs and Policies Intended to Have a Public Health Impact (Cont)

<table>
<thead>
<tr>
<th>RE-AIM ELEMENT</th>
<th>GUIDELINES AND QUESTIONS TO ASK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADOPTION</strong></td>
<td>Is the program feasible for majority of real-world settings (costs, expertise, time, resources, etc.)? Can it be adopted by low resource settings and typical staff serving high-risk populations?</td>
</tr>
<tr>
<td>Percent and representativeness of settings and staff that participate</td>
<td></td>
</tr>
<tr>
<td><strong>IMPLEMENTATION</strong></td>
<td>Can the program be consistently implemented across program elements, different staff, time, etc.? Are the costs—personnel, up front, marginal, scale up, equipment costs—reasonable to match effectiveness?</td>
</tr>
<tr>
<td>Consistency and cost of delivering program and adaptations made</td>
<td></td>
</tr>
</tbody>
</table>

### RE-AIM ELEMENT

<table>
<thead>
<tr>
<th>MAINTENANCE</th>
<th>GUIDELINES AND QUESTIONS TO ASK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long-term effects at individual and setting levels, modifications made</td>
<td>Does the program include principles to enhance long-term improvements (i.e., follow-up contact, community resources, peer support, ongoing feedback)? Can the settings sustain the program over time without added resources and leadership?</td>
</tr>
</tbody>
</table>

What Evidence is Needed?
EXTENDED CONSORT DIAGRAM

RE-AIM Issue

Content

Critical Considerations

**ADOP**

**REACH**

**IMPLEMENTATION**

**EFFICACY**

**MAINTENANCE**

a) Individual Level

b) Setting Level

*At each step, record qualitative and quantitative information and factors affecting each RE-AIM dimension and step in flowchart*
External Validity
Checklist for Researchers

1. _____ Record recruitment and/or selection procedures, participation rate, and representativeness at each of the following levels:
   a. Individuals, patients, citizens, or clients
   b. Intervention staff, or program delivery agents
   c. Delivery settings, work sites, health care clinics, schools

2. _____ Take note of any differences in delivery across:
   a. Settings, populations, and/or staff
   b. Program components
   c. Time, taking special care to note any modifications over time

3. _____ Record all impacts of intervention, including:
   a. Quality of life, or unintended adverse consequences
   b. Costs of implementation and/or program replication
   c. Moderator variables, especially those related to health disparities

4. _____ When conducting long-term follow-up report, pay attention to:
   a. Long-term effects on item #3 above
   b. Attrition at all levels in #1 above
   c. Institutionalization, modification, or discontinuance of the program

Reporting External Validity
Future Directions

• Document reliability of EV coding criteria
• Consider *summary metrics*, composite or overall EV quality scores
• Assistance to practitioners on how to combine with theory and local experience
• Evaluate which criteria most strongly related to long-term dissemination success
• Revise criteria based on lessons learned
Take Home Points

• Failure to focus on external validity is a major contributor to the disconnect between research and practice

• Need a broader approach to evaluating interventions that places appropriate focus on dimensions of external validity

• Reporting on external validity issues is needed to facilitate research into practice
Questions?