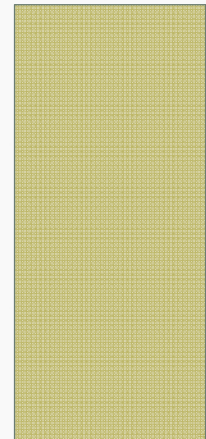


DECIDING WHAT EVIDENCE TO INCLUDE

REBECCA A. MAYNARD

NOVEMBER 18, 2014
WORKSHOP ON STANDARDS FOR BENEFIT-COST ANALYSIS OF PREVENTIVE INTERVENTIONS FOR
CHILDREN, YOUTH AND FAMILIES
NATIONAL ACADEMY OF SCIENCES



FOUR KEY CONSIDERATIONS

1. Overall relevance of the study
2. Relevance of particular impact estimates reported
3. Causal validity of the impact estimates
4. Adequacy of reporting on impact estimates

1. OVERALL RELEVANCE OF THE STUDY

- Intervention
 - Is the intervention relevant to policy or practice decision under consideration?
- Counterfactual
 - Was the intervention tested against a relevant counterfactual condition?
- Context
 - Was the study conducted in a location and in a time period that is relevant?
- Setting
 - Do the study findings pertain to a relevant setting option?

2. RELEVANCE OF IMPACT ESTIMATES (IN RELEVANT STUDIES)

- Reference period for impact estimates
 - Match for “investor’s”?
- Relevance of particular estimates to the mission
 - Match to the programmatic or policy goals?
 - Especially important when considering RD designs:
 - Is the policy/practice question about overall or marginal treatment impacts?
- Primary versus secondary outcomes of the study
 - Likelihood of “selective” reporting?

CHALLENGES TO CAUSAL VALIDITY

- Studies that began as a well executed RCT
 - Attrition
 - Differential attrition
 - Measurement short-falls
 - Analysis and reporting shortfalls
- Matched comparison groups (e.g., PSM)
 - Untestable assumptions
- RDD
 - Restricted range impact estimates (LATE)

3. CAUSAL VALIDITY OF IMPACT ESTIMATES

- Method of creating intervention v. contrast groups
 - Randomized controlled trial (RCT) or other matching technique?
- Comparability of analysis samples
 - For RCTs, was there substantial sample loss? Was it differential between the Intervention and comparison groups?
 - For other designs, are “comparability” standards met?
- Analysis methods
 - Appropriate to the sample design?
 - Adequately addresses selection issues?
 - Strategies differ depending on the issue (e.g., nonresponse, RD design, self-selection to participate)

4. ADEQUACY OF REPORTING

- Meaningful metrics for outcome measure
 - Uniform measures (e.g., nationally normed tests) versus study specific effect sizes
- Contextual information
 - Characteristics of the study sample, intervention characteristics, intervention context, etc.
- Implementation information (e.g., fidelity, dose)

RECOMMENDATIONS

- Build on existing standards
 - What Works Clearinghouse, Campbell Collaboration, Blueprints, Campbell Collaboration, etc.
- Share coded studies
 - Individual studies have relevance a range of B-C analyses
 - Relevant impact estimates and costs will vary by use
- Take account of confidence intervals
 - Consider intervals other than 95%

STEPS IN THE WWC REVIEW PROCESS

1. Develop the review protocol
 - Ensures alignment & consistency
2. Identify the relevant literature
 - Promotes consistency & completeness
3. Screen and review the studies
 - Uses a consistent format & facilitates re-use
4. Summarize the findings
 - Consistent, re-usable format
5. Archive in a shareable format
 - Data sharing agreements & technical guidelines

IMPLICATIONS OF RECOMMENDATIONS

- Some studies on a topic will be excluded from a B-C analysis
 - Reasons may vary—sample, analysis, reporting
- Not all evidence within “accepted studies” will be deemed relevant for a particular B-C analysis
 - Some outcomes or time periods of analysis may not be relevant for any number of reasons
- Not all outcomes within a study will receive the same treatment
 - The reasons may pertain to properties of the sample design, the analysis methods, or something else— all effect sizes are not created equal
- Evidence for different periods of follow-up may be treated differently in a B-C analysis
 - This may be because of differences in the study sample, relevance of the time period, or something else

PREPARE DATA FOR REPEATED ANALYSIS

TAILORED TO THE QUESTION, SETTING, AND AUDIENCE