

# Introduction to Meta-analysis

Jim Derzon  
Battelle Memorial Institute  
Presented at the AEA/CDC Summer Institute, 2008

---

---

---

---

---

---

---

---

## Why, and What is, Meta-analysis

- Meta-analysis is a quantitative summary of evidence across diverse studies of the same outcome, relationship, or finding
- Why do meta-analysis?
  - Resolve controversy
  - Identify best programs or practices
  - Document the breadth and depth of knowledge
    - Generalizability
  - Identify research gaps
  - Identify important moderators of success

7/11/2008 Derzon, AEA/CDC Summer 2008 2

---

---

---

---

---

---

---

---

## Meta-analysis

- **Meta-analysis is a quantitative assessment and summary of group-wise evidence**
- Meta-analysis is characterized by
  - A systematic, detailed, and organized approach that
  - Makes explicit the domain of research covered,
  - The nature and quality of the information extracted from that research, and
  - The analytic techniques and results upon which interpretation is based.
- **Meta-analysis is a science that draws on qualitative, quantitative, and statistical theory to provide a concise and transparent synthesis of the results of the included studies**

7/11/2008 Derzon, AEA/CDC Summer 2008 3

---

---

---

---

---

---

---

---

## What's it Good For?

- Evidence of main effects
  - Central tendency research
    - Prevalence rates
  - Pre-post contrasts
    - Growth rates
  - Group contrasts
    - Experimentally created groups
      - RCTs, quasi-experimental groups
    - Comparing naturally occurring groups
      - With and without, Exposed and not-exposed
  - Association between variables
    - Measurement research
      - Validity generalization
    - Individual differences research
      - Correlations between constructs
- What's it not good for?
  - Conditioned findings
  - Qualitative evidence

7/11/2008

Derzon, AEA/CDC Summer 2008

4

---

---

---

---

---

---

---

---

## The Heart of Meta-analysis: The Effect Size

- Is the **"dependent variable"**
  - Standardizes findings across studies for direct comparison
- Can be any standardized index... but
  - Must be comparable across studies (generally requires standardization)
  - Must represent the magnitude and direction of the relationship of interest
  - Must be independent of sample size
- Different meta-analyses may use different effect size indices

7/11/2008

Derzon, AEA/CDC Summer 2008

5

---

---

---

---

---

---

---

---

## Alternatives

- Literature review
  - Mulrow (1987) – Only 1/50 reviews published 1985-1986 met eight criteria for identifying, assessing and synthesizing information.
  - Typically focus on statistical significance
- Expert review
  - Antman et al (1992): expert recommendations typically years behind the evidence (or years ahead as it were...)
- Systematic review
  - Sister to meta-analysis, often narrative in form, not typically centered on effect size estimating

7/11/2008

Derzon, AEA/CDC Summer 2008

6

---

---

---

---

---

---

---

---

## Meta-analysis v. Systematic review

- Not mutually exclusive...
- Systematic review allows complex findings
- Systematic review tends to focus on statistical significance testing
  - Significance testing is not well suited to this task
    - Highly dependent on sample size
    - Null finding does not carry the same "weight" as a significant finding
- Meta-analysis changes the focus to the **direction and magnitude** of the effects across studies
  - Isn't this what we are interested in anyway?
  - Direction and magnitude represented by effect size

7/11/2008

Derzon, AEA/CDC Summer 2008

7

---

---

---

---

---

---

---

---

## Prologue

- Belief in the aggregate
  - Laplace (c.1820)
    - Measurement contains error, Bell Curve
  - Karl Pearson (c. 1890)
    - Truth in the distribution (mean, sd, symmetry, kurtosis)
    - Documented the "randomness of observation"
  - RA Fisher (c. 1930)
    - Reality in the fit of observed data to "true" mathematical models
    - ANOVA, degrees of freedom, error
- **Reality in data distributions, not in datum**

7/11/2008

Derzon, AEA/CDC Summer 2008

8

---

---

---

---

---

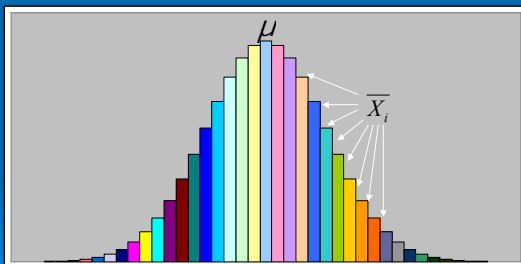
---

---

---

## The Normal Distribution

$$X_{\text{bar}} = X_{ij} + b + e$$



7/11/2008

Derzon, AEA/CDC Summer 2008

9

---

---

---

---

---

---

---

---

## Birth of Meta-analysis

- Lind (1747) Summarized existing literature on scurvy
  - "It is no easy matter to root out prejudices... before the subject could be set in a clear and proper light, it was necessary to remove a great deal of rubbish"
- Cochran (1953)
  - Discusses a method of averaging means across independent studies
  - Introduces inverse variance weighting and homogeneity testing
- Does psychotherapy work?
- Eysenck (1952) concludes no favorable effects of psychotherapy
  - 20 years of research and hundreds of studies failed to resolve the debate
- Glass & Smith (1978) statistically aggregate 375 psychotherapy outcome studies & conclude psychotherapy works
- Glass called his method "meta-analysis"

7/11/2008

Derzon, AEA/CDC Summer 2008

10

---

---

---

---

---

---

---

---

---

---

## Stages in Meta-analysis

- Specify topic
- Study eligibility
- Search – document identification
- Coding (aka data abstraction)
- Data preparation
- Synthesis
- Results and interpretation

7/11/2008

Derzon, AEA/CDC Summer 2008

11

---

---

---

---

---

---

---

---

---

---

## Topic Specification

- **Must be precise:**
  - Guides the search and selection of studies
  - Guides inclusion/exclusion criteria
  - Guides coding strategy
  - Guides analysis
- **Must fit the literature & resources**
  - Humpty Dumpty
  - Goldilocks and the Three Bears

7/11/2008

Derzon, AEA/CDC Summer 2008

12

---

---

---

---

---

---

---

---

---

---

## Examples

- How well do antecedent behaviors, characteristics, and experiences predict later violent behavior?
- How effective are school-based programs in reducing violent or antisocial behaviors?
- What components of mass-media interventions are associated with reduced substance use?

7/11/2008

Derzon, AEA/CDC Summer 2008

13

---

---

---

---

---

---

---

---

## Eligibility criteria

- **Must be explicit**
  - Supported by definitions
- Define the scope of the meta-analysis
  - What studies will contribute
  - What findings are eligible
  - To whom the findings apply
  - Under what conditions the evidence is valid
- **Criteria are refined as you interact with the literature – Humpty Dumpty, Goldilocks**

7/11/2008

Derzon, AEA/CDC Summer 2008

14

---

---

---

---

---

---

---

---

## Include or exclude low quality studies?

- Reality in data distributions, not in datum
  - All findings contain error and bias
- **Quality is multi-dimensional**
  - Some dimensions produce bias
  - Others increase error variance
  - Quality is dependent on type of evidence and topic
- **Quality dimensions are continuums, not a dichotomies**
- Quality as a criterion
  - **Strike a balance that you are comfortable defending**
  - Too restrictive = little data and reduced generalizability
  - Too inclusive = questionable confidence in results
- Quality as empirical variables
  - Measure what you can't control
  - Test what you measure
  - Use statistical adjustment or exclude on observed impact

7/11/2008

Derzon, AEA/CDC Summer 2008

15

---

---

---

---

---

---

---

---



## Abstract Review

- Be efficient!
  - Download citations & abstracts into a sortable database (by journal and year – Retrieval)
  - Display abstracts within a set field
- Be liberal
  - Better to be retrieve too much than miss something important
  - Yes, No, Maybe...
- Update search if indicated
  - Search ends when all you find is what you've found

7/11/2008

Derzon, AEA/CDC Summer 2008

19

---

---

---

---

---

---

---

---

---

---

## Coding: What?

- Four functions of meta-analysis data
  - Management variables
    - Linking data across data files, coder information, dates
  - Descriptive variables
    - Who, what, where, when, & how
    - Can be nominal (categorical), ordinal, interval, or ratio – no direction required
  - Explanatory variables
    - What affects study findings (moderators of effectiveness)
    - Can also be descriptive, but direction typically required
  - Effect size variables
    - Bare bones meta-analysis
      - Source, Effect size, n(s), & weight (IV, DV if multi-faceted)
- Refine value options as necessary – **Humpty Dumpty, Goldilocks**

7/11/2008

Derzon, AEA/CDC Summer 2008

20

---

---

---

---

---

---

---

---

---

---

## Coding: Who?

- What's a study? It is the unit of analysis
  - Not an analysis, not a document...
  - A study is the **Study-Sample**
    - A unique confluence of participants and status
      - A unique comparison in time
        - E.g., RCT, Quasi-experiment
      - A unique sample subjected to a unique practice
        - E.g., Pre-post
      - A unique sample in time (e.g., cross-sectional)
      - A unique sample over time (e.g., longitudinal)
  - The study ensures statistical independence

7/11/2008

Derzon, AEA/CDC Summer 2008

21

---

---

---

---

---

---

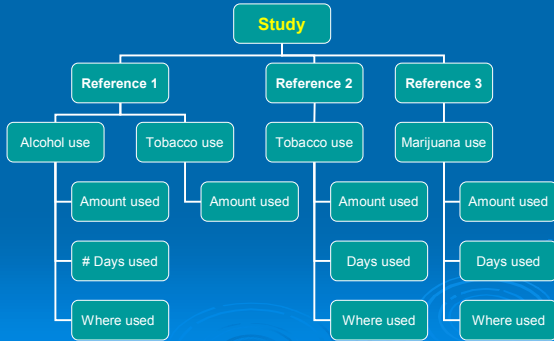
---

---

---

---

## The view from a single study-sample



7/11/2008

Derzon, AEA/CDC Summer 2008

22

---

---

---

---

---

---

---

---

## Coding: How?

- Meta-analysis data occur at different levels
  - Study, bibliography, practice/intervention, measures/outcomes, effect size
- Use relational files
  - Meta-analytic data coded at different levels
    - Reduces coder burden
    - Improves coder accuracy
    - Eases corrections
  - Use linking variables (aka “key” variables)
- Coding manuals are not optional
- Reconciliation/reliability check is not optional

7/11/2008

Derzon, AEA/CDC Summer 2008

23

---

---

---

---

---

---

---

---

## Example of coding items

- Characteristics of subjects that may affect results
  - Including demographics, recruitment, assignment, retention, risk status
- Characteristics of setting
  - Any exogenous features that may affect results
- Characteristics of intervention
  - Implementation issues
  - Route, dose, timing, duration
- Characteristics of measurement
  - What measured, how measured, reliability & sensitivity
  - Quality of the evidence
    - Statistical bias
    - Reporting bias
- Findings: Effect size, sample size(s), weight

7/11/2008

Derzon, AEA/CDC Summer 2008

24

---

---

---

---

---

---

---

---

## Coding Effect Sizes

### > Any standardized index of outcome

- Findings reported using different metrics, effect sizes encode them to a common index

- **Main effects**

### > Common effect sizes

- One variable

- Proportions (logit transformed)

$$\text{LogBS} = \text{Log}_e \left( \frac{Pa(1-Pc)}{Pc(1-Pa)} \right) * .551$$

- Two variables

- Pre-post contrasts
- Gain scores
- Group difference scores

$$d = \frac{TXij - CXij}{SD_{pooled}}$$

- Association between variables

- Correlations

$$ES_{r,r} = .5 \log_e \left[ \frac{1+r}{1-r} \right]$$

7/11/2008

Derzon, AEA/CDC Summer 2008

25

---

---

---

---

---

---

---

---

---

---

## Data Preparation: Mechanical

### > Most effect size metrics need tweaking

- Standardized difference (d-statistic)
  - Correction for pretest differences?
  - Small sample correction?
- Odds ratio
  - Adding a constant when cell values = zero
- Correlations
  - Z-transformation
  - Tetrachoric correction?

### > Some distributions need tweaking

- Winsorizing weights and effect sizes?

### > Weight effect sizes by their inverse variance

- Different sample sizes, different standard errors

7/11/2008

Derzon, AEA/CDC Summer 2008

26

---

---

---

---

---

---

---

---

---

---

## Data Preparation: Conceptual Representing a Study's Result

### > Studies report multiple findings

- Different measures, same construct
- Different constructs, same notion
- Different notions, same study

### > Constructs are not 'real'

- Built according to purpose; not empirically derived

### > Options

- Average measures within construct
- Select "best" measure
- Randomly select measure
- Or some combination thereof

7/11/2008

Derzon, AEA/CDC Summer 2008

27

---

---

---

---

---

---

---

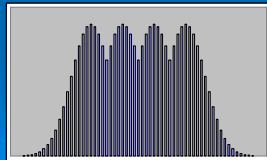
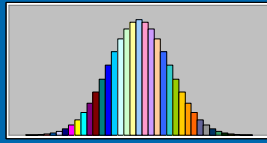
---

---

---

## Synthesis: Evaluating the Grand-Mean

- Data weighted by inverse variance
  - Or data not weighted by inverse variance?
- Fixed or random effects?
  - Depends on what you believe
- Homogeneity
  - Does the distribution represent a single population parameter?
  - What to do?
    - Winsorize? Trim?
    - Goldilocks?



7/11/2008

Derzon, AEA/CDC Summer 2008

28

---

---

---

---

---

---

---

---

---

---

## Other Interpretation Issues

- Strong meta-analysis will not 'correct' weak literatures
  - Sensitivity testing may improve understanding
  - Meta-regression (ANCOVA-analogue) may help
- Are substantive findings confounded with methodological problems?
- Report and appreciate variability
- All findings are good!
  - Research gaps & generalizability of results
- What is a "big" effect size?

7/11/2008

Derzon, AEA/CDC Summer 2008

29

---

---

---

---

---

---

---

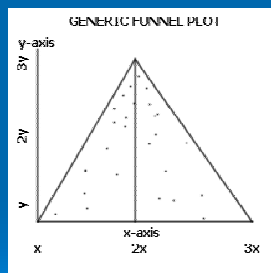
---

---

---

## Publication Bias?

- Funnel plots
  - A funnel plot is a scatterplot of treatment effect against a measure of study size.
- Failsafe-N
  - Number of zero effect sizes studies necessary to obtain a non-significant effect size



7/11/2008

Derzon, AEA/CDC Summer 2008

30

---

---

---

---

---

---

---

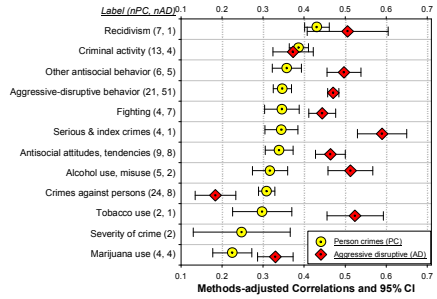
---

---

---

## Dot-error Bar Plots

### Longitudinal relationships between antisocial behaviors with aggressive, disruptive behavior and crimes against persons



7/11/2008

Derzon, AEA/CDC Summer 2008

31

---

---

---

---

---

---

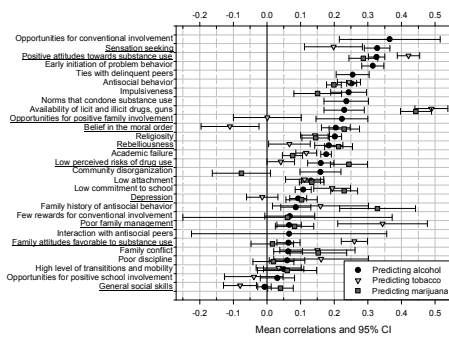
---

---

---

---

## 29 Commonly Discussed Risk Factors and Their Relation to Substance Abuse



7/11

32

---

---

---

---

---

---

---

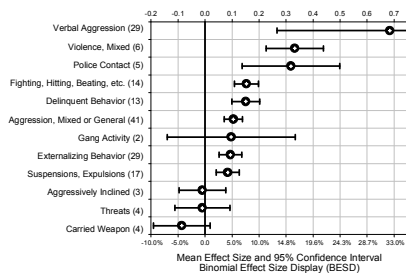
---

---

---

## School-Based Interventions for Violent & Antisocial Behaviors: Outcomes

### Antisocial Outcomes



7/11/2008

Derzon, AEA/CDC Summer 2008

33

---

---

---

---

---

---

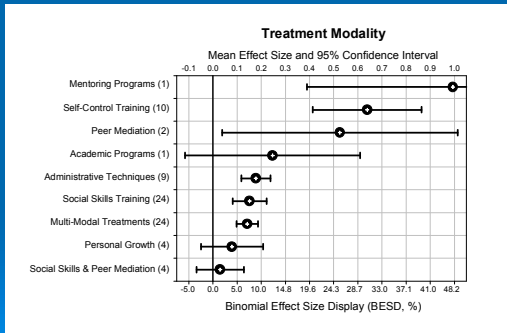
---

---

---

---

## School-Based Interventions for Violent & Antisocial Behaviors: Practices



7/11/2008

Derzon, AEA/CDC Summer 2008

34

---

---

---

---

---

---

---

---

---

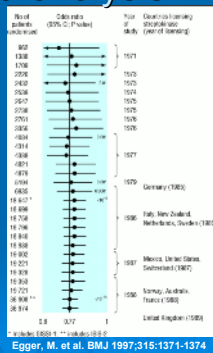
---

---

---

## Cumulative meta-analysis

- Each subsequent result is added to prior results
- Total mortality results from randomized controlled trials of intravenous streptokinase in myocardial infarction



7/11/2008

Derzon, AEA/CDC Summer 2008

35

---

---

---

---

---

---

---

---

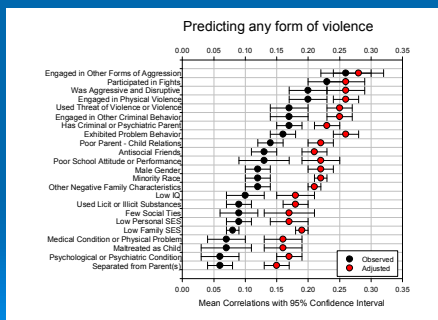
---

---

---

---

## Innovative Meta-analysis: Correcting for Methods Bias, Correlations



7/11/2008

Derzon, AEA/CDC Summer 2008

36

---

---

---

---

---

---

---

---

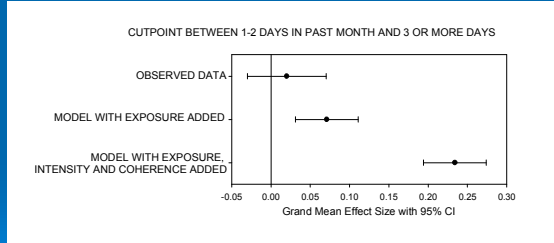
---

---

---

---

# Innovative Meta-analysis: Correcting for Methods & Practice Bias: Interventions



7/11/2008

Derzon, AEA/CDC Summer 2008

37

---

---

---

---

---

---

---

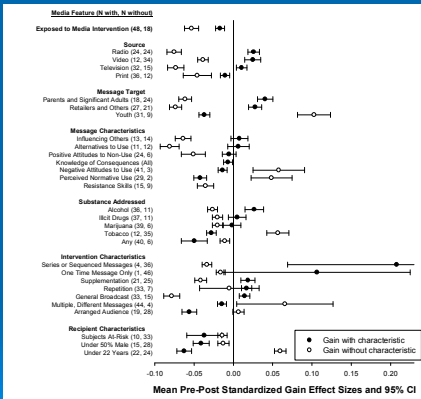
---

---

---

## Innovative Meta-analysis: Pre-Post With and Without Design Features

### Mass Media Effects on Substance Using Behaviors



7/11/2008

Figure 1: Behavior

---

---

---

---

---

---

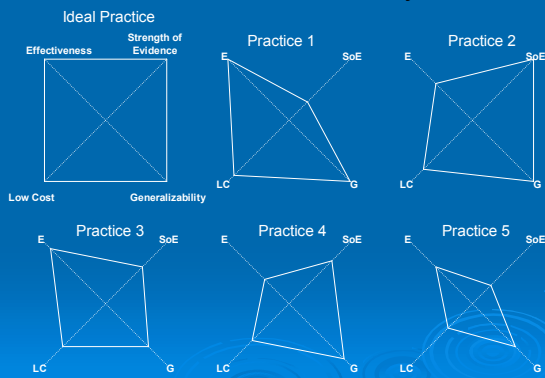
---

---

---

---

## The Future of Meta-analysis?



7/11/2008

Derzon, AEA/CDC Summer 2008

39

---

---

---

---

---

---

---

---

---

---

## Recap

- Meta-analysis is a quantitative assessment and summary of group-wise evidence
- The effect size is the heart of meta-analysis
- Reality in data distributions, not in datum
- The study-sample is the unit of analysis
- Meta-analysis draws on qualitative, quantitative, and statistical theory to provide a concise and transparent synthesis of the results of the included studies

7/11/2008

Derzon, AEA/CDC Summer 2008

40

---

---

---

---

---

---

---

---