Randomized experiments vs. Propensity scores matching: a Meta-analysis.

Antonio Olmos
Priyalatha Govindasamy
Research Methods and Statistics
University of Denver

American Evaluation Association Conference
Denver, CO. 2014

Background

- There is a substantial increase in the use of Propensity score in the field of Medicine
- Especially in the field of Cardiovascular
- Propensity scores with observational design is used as the alternate to Randomized Experiments

Introduction

What is Propensity scores

- This single propensity score is referred to as the probability of receiving the treatment or not given the observed covariates (Rubin, 1997)

- Propensity scores represent a single score of the relationship between the multiple observed covariates in the assignment of people into treatment or control groups (Stone & Tang, 2013).

- Propensity score matching in observational data creates matched treatment and control groups that are as similar as possible based on a wide range of observed covariates.

Introduction (continue…)

The role of Propensity Scores

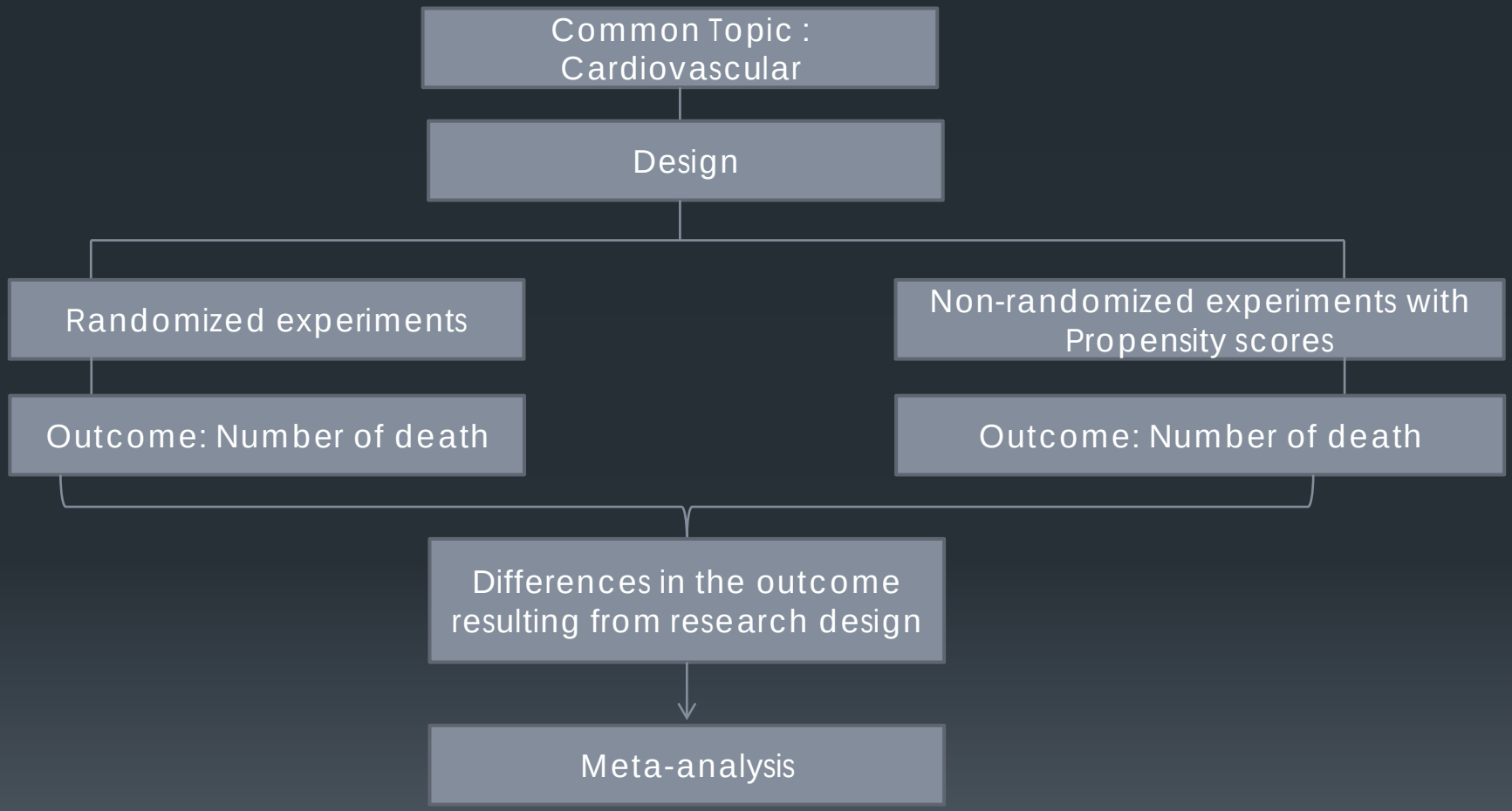
Upon computing and matching the groups on Propensity scores, the only differences between the treatment and control group should be the reflection of whether the groups have received treatment or not.

At this point, researchers could conclude that any observed differences in the outcome is the result of the treatment.

The use of Propensity scores in correcting the selection biases and creating group equivalence on the observed covariates has gain increased popularity among researchers.

Objective of the study

Purpose of the study



Topic: Coronary Bypass Artery Grafting procedures

Research question

- Meta-analysis approach was used to quantitatively evaluate the agreement between randomized designs and observational studies using Propensity scores methods in Coronary Bypass Artery Grafting procedures.

- A systematic comparison of the On-pump versus Off-pump Coronary Artery Grafting procedure was performed to access the effect of the surgical procedures.

- Mortality was recorded as the primary outcome estimated from the effect of the two surgical procedures.

Research question was:

Is there any difference in the mortality rates for the clinical procedure when comparing randomized versus observational designs using Propensity scores?

Definition

On-pump Coronary Artery Bypass

This is a conventional bypass grafting surgical procedure where surgeons perform delicate anastomoses on an arrested heart under optimal visualization with the help of CPB circuit (Lamy et al., 2012).

Off-pump Coronary Artery Bypass

It's a design with new heart stabilizer that allows surgeon to perform cardiopulmonary artery bypass grafting on a beating heart without using CPB circuit.

Methodology

Approach: Meta-analysis

Meta-analysis is a technique for combining (synthesizing) a collection of primary studies to estimate a general recommendation or conclusion

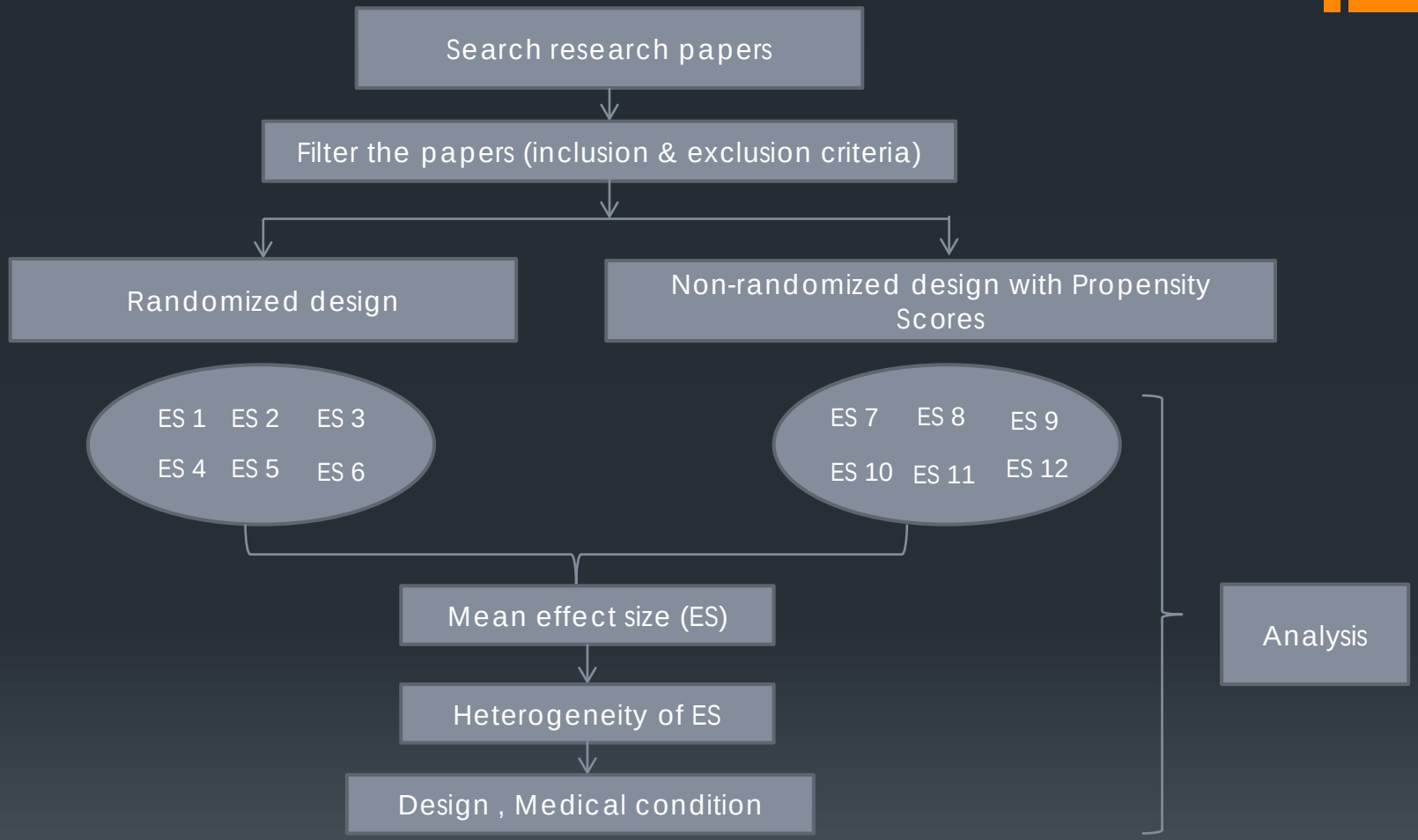
Effect size: is the magnitude of effect for a statistical test

How does meta-analysis work?

Effect size usually quantifies the degree of difference between or among groups or the strength of association between variables such as a group-membership variable and an outcome variable

Focuses on the direction and magnitude of effects across studies,

Meta-analysis process



Methodology (continue)

Process

Inclusion criteria:

(1) Published and unpublished studies in English on Off–pump vs On-pump Coronary Bypass Artery Grafting,
(2) Randomized control trial study,
(3) Non-randomized (observational) study using Propensity score,
(4) Numerical values reporting the mortality rates.

Exclusion criteria:

(1) Studies without mortality rates and failures in reporting the study designs

Variables

The outcome variable in the study is the mortality rates of patients who had Off-pump versus On-pump surgical procedures.

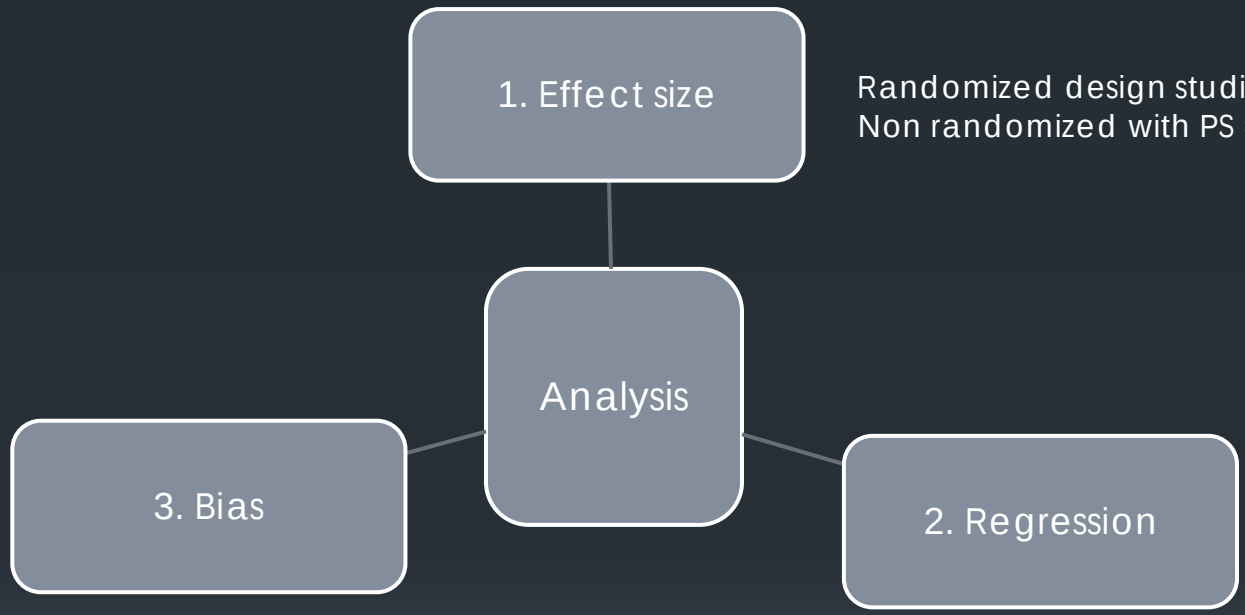
Outcome (death) observed in counts is converted into odds ratio.

Two moderators were identified and coded in the study.

1. Design (0 = Randomized control trial, 1= Observational study with PS)
2. Condition (0 = No medical condition, 1= With medical condition)

Methodology (continue)

Analysis



1. Presences of bias
2. Artifact of bias
3. Impact of bias

1. Design
2. Medical condition
3. Design + Medical condition

Software for analysis : R, packages (meta, metafor)

Results

1. **Effect size**

Effect size = -0.1059 (-0.1956 – 0.0161) , p = 0.021

A significant mean effect size indicated that the two surgical procedures reported differences in the number of deaths.

However, the test of heterogeneity for the fixed model was significant

Q(df = 20) = 78.5769, p-val < .0001

There was variability in the effect sizes

Therefore, takes into account the variability and uses the Random effect model

Result (continue…)

Random effect model

Re-estimate the mean effect size and use adjusted inverse variance weight that incorporates the random variance component

Effect size = -0.2594 (-0.5170 – -0.0017), p = 0.0485

A significant mean effect size indicated that the two surgical procedures reported differences in the number of deaths.

However, the test of heterogeneity for the fixed model was significant

Q(df = 20) = 78.5769, p-val < .0001

Results (continue)

2. Regression

Test of moderator

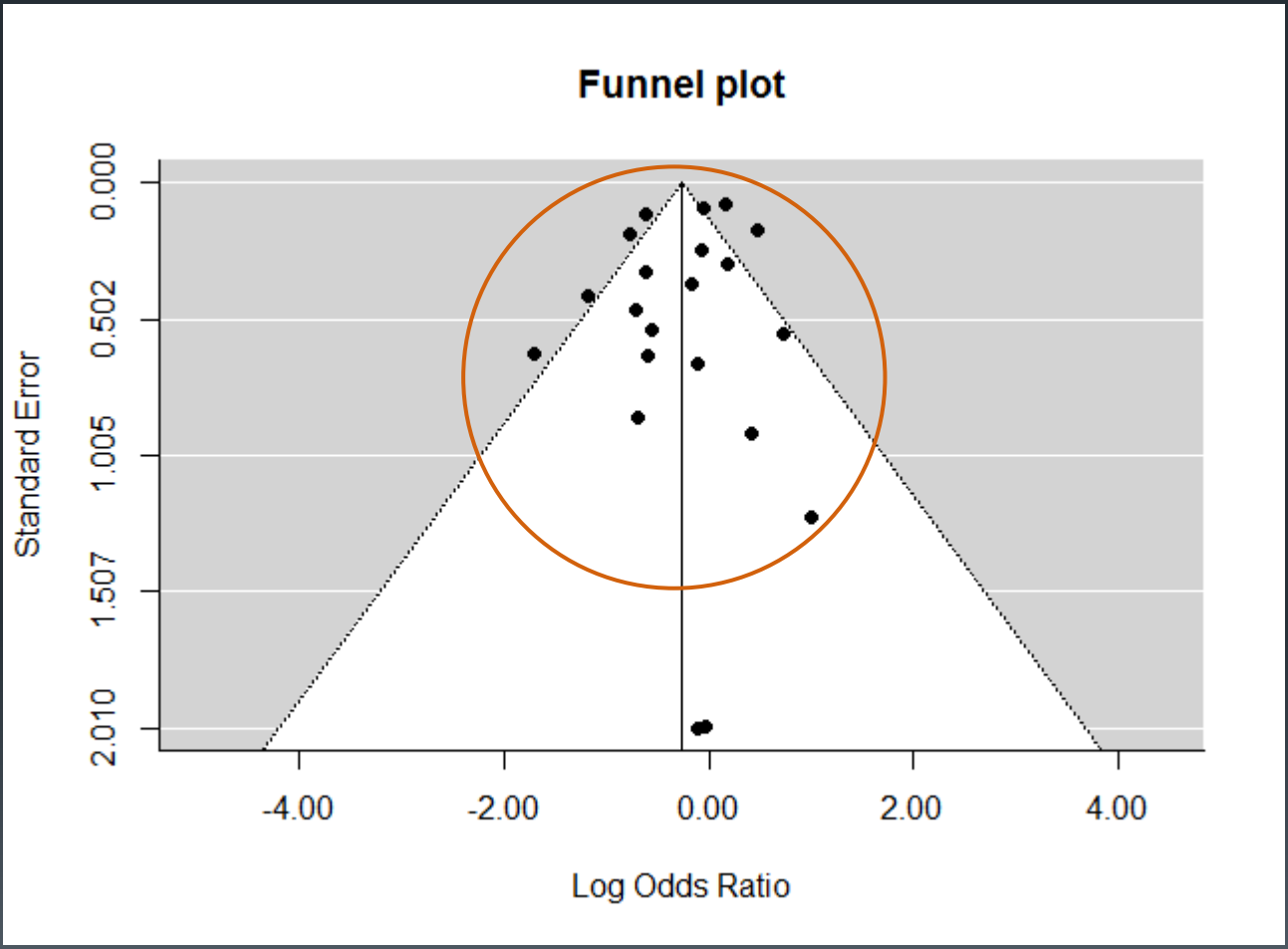
Model	Moderator	Test of Moderator
1	Design	QM(df=1) = 2.61, p = 0.106
2	Medical condition	QM(df=1) = 25.27, p <.001
3	Design + Medical condition	QM(df=2) = 36.53, p <.001

Test of heterogeneity

Model	Moderator	Test of Heterogeneity	Variance unexplained
1	Design	QE(df=19) = 77.31, p <.001	R^2 = 96.63%
2	Medical condition	QE(df=18) = 19.62, p = 0.355	R^2 = 9.92%
3	Design + Medical condition	QE(df=17) = 16.60, p = 0.482	R^2 = 1.26%

Result (continue...)

3a. Presences of bias



Funnel plots shows that smaller error variances between the studies as they were pointing towards the tip of the funnel plot

Result (Continue…)

ii. Presences of bias

1. Regression

Sample size of the study was regressed on the effect sizes

Is to examine influences of sample sizes in effect sizes.

Results from the linear regression reported sample size as a non-significant predictor of the effect size, b=-2.25, t(19) 0.42, p=0.631.

2. Rank correlation

- Kendall's tau = 0.0476, p = 0.7884

3. Egger regression

- test for funnel plot asymmetry: z = -0.3233, p = 0.7465

The non-significant result shows no influence of sample size in the estimated effect size.

Results (continue….)

3b. Potential Artifact biases

Fail-safe N Calculation Using the Orwin Approach

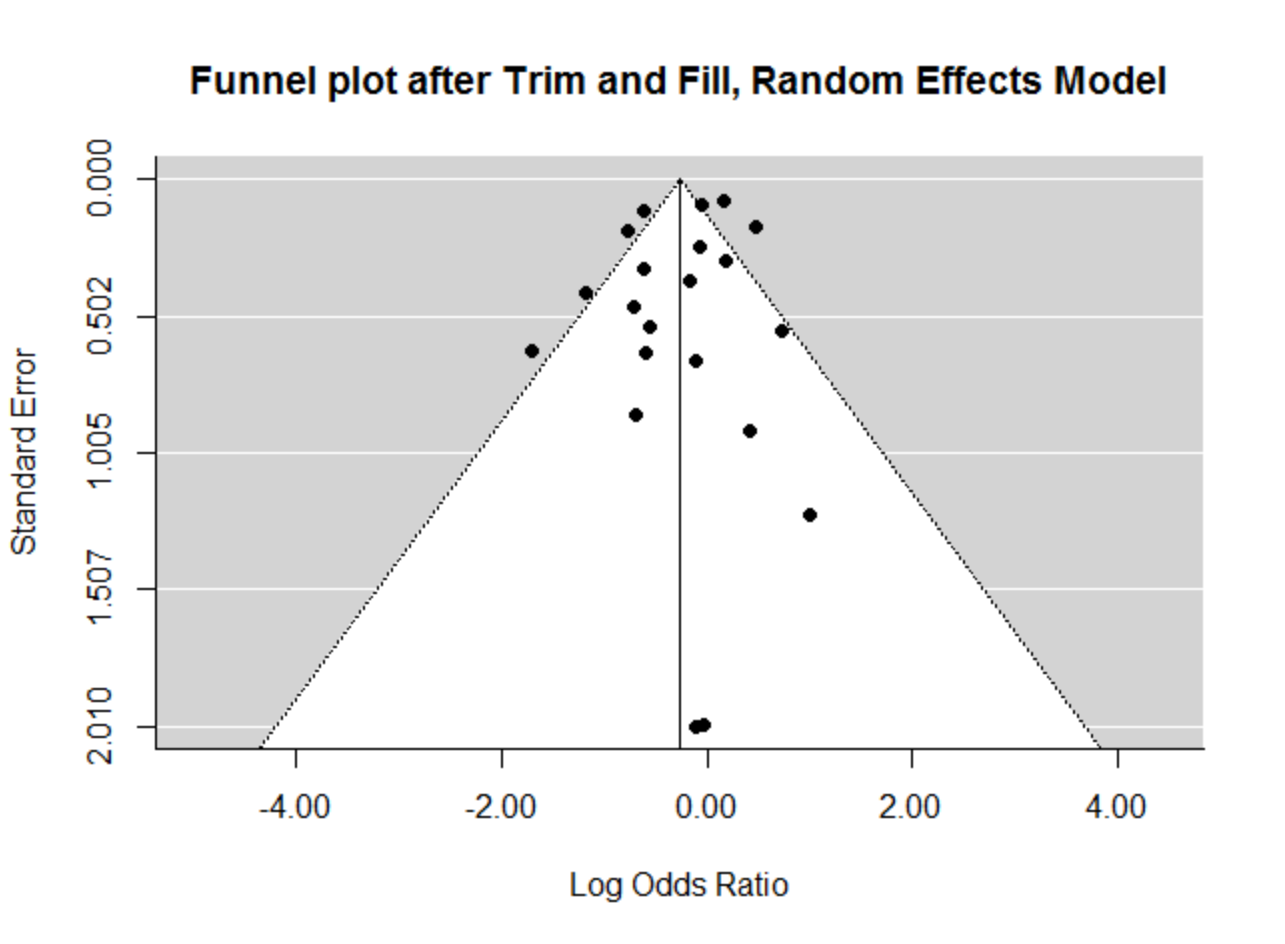
- Average Effect Size: -0.2370
- Target Effect Size: -0.1185
- **Fail-safe N: 21**

Fail-safe N Calculation Using the Orwin Approach

- Average Effect Size: -0.2370
- Target Effect Size: 0.1000
- **Fail-safe N: 0**

Results

3c. Impact of bias on the effect size estimates



The Trim and fill approach (Card, 2012) shows no need for the reduction or inclusion of new studies to the current list. This result was consistent with earlier results intended to test the presence of biases.

Discussion

- In this study, we conducted a meta-analysis to compare the outcomes of 21 clinical trials where either a randomized control trial or an observational study using propensity scores was used. The studies compared the mortality rates for Off-pump versus On-pump surgical procedures.

- The reported effect size was -0.259 (log odds-ratio), thus meaning that the odds of dying, OR = 0.772 (OR = $e^{-.259}$) are higher for patients undergoing On-pump procedure.

- Individuals receiving the On-pump procedures are 1.30 (i.e., 1/0.772= 1.30) times more likely to die than patients who received the Off-pump procedure.

Discussion (continue…)

- No statistically significant differences were found between the randomized and observational design with Propensity score in the estimated effect sizes.

- Finding appears to support the claim that Observational/quasi-experimental designs with Propensity Scores produce similar results to those using randomized control studies.

Discussion (continue…)

- No differences in the outcome between the two designs highlights the success of the Propensity score method to eliminate the selection biases that has been randomized designs greatest strength.

- This success provides more credit to the design and gives confidence to the researcher about implementing this approach as a viable alternative to randomized control studies.

Limitations

- The result observed in this study, the findings should be evaluated with caution.
- This is due to the fact that the sample used in this study is relatively small.

Issues in meta-analysis

- Difficulty in studies meeting the inclusion criteria.
- Repetition of the sample (eg, different follow up duration, slicing the data into specific characteristics)
- No clear presentation on the use of Propensity scores
- Matching techniques are not clear.

Conclusion

- The use of non-Randomized design using Propensity score is comparable to a Randomized design
- Hope to provide confidence in the use of Propensity scores
- According to Wolfe (2000), observational designs using Propensity scores are considered as a practical design of lower cost compared to randomized control studies.
- New opportunities in the field of research

Question

Thank you

Antonio Olmos, Ph.D
Priyalatha Govindasamy, M.A

Research Methods and Statistics
Morgridge College of Education
University of Denver

Antonio.olmos@du.edu
gpriyalatha@hotmail.com