Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Journal Club

"Choice of Mean Difference Estimation for Meta-analysis of Continuous Outcomes:

Final values, Change scores, and ANCOVA"

Commentator

Sunatee Sa-nguansai, MD

M.Sc. Medical Epidemiology Student

Academic Year 2022

Topic Overview

- Introduction
- Meta-analysis of change score: Recommendation from Cochrane Handbook for Systematic Reviews of Interventions

Introduction

Meta-analysis of Continuous Outcomes

- **The term 'continuous'** in statistics conventionally refers to a variable that can take any value in a specified range.
- When dealing with numerical data, number may be measured and reported to an arbitrary number of decimal places.
  - Example: weight, area, and volume
- Meta-analyses of continuous outcomes are recognized to be more challenging than those with binary outcomes.

Wisdom of the Land

Effect Size Measurement

- In meta-analyses of continuous outcomes, the common effect size measurements are as follow:
  - Mean difference (MD)
  - Standardized mean difference (SDM)
  - Ratio of means (RoM)
- The choice of the two most commonly used effect size measures (MD and SDM) is mainly determined by the scale of measurement.
- If the scale of measurement is similar as in HbA1c, the MD can be used to aggregate effect sizes across studies.

Pooling Effect size

- To pool the effect sizes of continuous outcomes, it is important to consider whether the baseline MD of the outcome data between the intervention and control groups is adequately balanced.
- **Baseline imbalance can result from:**
  - Chance especially in small trials
  - Selection bias due to inadequate allocation concealment
  - Poor randomization
- Therefore, in meta-analysis, it is important to consider accounting for baseline imbalance and pre/post correlation.
- However, the meta-analyses available have not taken into account a specific methodological challenge posed by baseline imbalances between groups.

Choice of Mean Difference Estimation

- **Choice of mean difference estimation:**
  - Simple analysis of change scores (CS)
  - Simple analysis of final values (FV)
  - Analyses of Covariance (ANCOVA)
- Meta-analyses of RCTs reporting continuous outcomes with high baseline imbalances require adjustments using an ANCOVA procedure effect size estimator at a pooled and individual study levels.
- ANCOVA produces a relatively more precise effect size estimate than CS and FV.

Choice of Mean Difference Estimation

- There is methodological guidance available to employ ANCOVA effect size.
- **Limitation of ANCOVA for meta-analysis:**
  - The unavailability of summary data from RCTs
  - Absence of individual participant data (IPD)
  - The complexity of the ANCOVA methodology

Choice of Mean Difference Estimation

- Three commonly used estimators are the final value (FV), change score (CS), and analysis of covariance (ANCOVA) estimator

$$\hat{\theta}_{FV} = \bar{y}_{\text{int}} - \bar{y}_{ctrl}$$

$$\hat{\theta}_{CS} = (\bar{y}_{\text{int}} - \bar{y}_{ctrl}) - (\bar{x}_{\text{int}} - \bar{x}_{ctrl})$$

$$\hat{\theta}_{ANCOVA} = (\bar{y}_{\text{int}} - \bar{y}_{ctrl}) - \beta(\bar{x}_{\text{int}} - \bar{x}_{ctrl})$$

where $\beta = \rho \frac{\sigma_y}{\sigma_x}$

Some observations from comparing the analytical methods

Estimates of intervention effect:

- For a particular data set, the three estimators can produce different estimates of intervention effect.
  - When the correlation (btw baseline value and final value) is close to 0, $\hat{\theta}_{ANCOVA} \approx \hat{\theta}_{FV}$
  - When the correlation (btw baseline value and final value) is close to 1, $\hat{\theta}_{ANCOVA} \approx \hat{\theta}_{CS}$
- Over the data sets (varying correlation), the ANCOVA estimate varies, however, a simple analysis of either change scores (CS) or final values (FV) dose not.

Some observations from comparing the analytical methods

Standard errors:

- The standard error (SE) of the FV estimator remains the same over all data sets.
- As the correlation increases, the SE of the CS estimator decreases.
  - **When the correlation < 0.5,** the SE of CS estimator is > SE of FV estimator.
  - **When the correlation > 0.5,** the SE of CS estimator is < SE of FV estimator.
- For a particular data set, the SE of the ANCOVA estimator is smaller compared to the SEs of FV and CS.

Relationship Between the Estimators

$$\hat{\theta}_{ANCOVA} = (\bar{y}_{\text{int}} - \bar{y}_{ctrl}) - \beta(\bar{x}_{\text{int}} - \bar{x}_{ctrl})$$

- **Assuming $\sigma_y^2 = \sigma_x^2$:**
  - When the correlation is close to 0, $\beta \approx 0$; $\hat{\theta}_{\mathbf{ANCOVA}} \approx \hat{\theta}_{\mathbf{FV}}$
  - When the correlation is close to 1, $\beta \approx 1$; $\hat{\theta}_{\mathbf{ANCOVA}} \approx \hat{\theta}_{\mathbf{CS}}$
- When there is minimal baseline imbalance, the three methods produce similar estimates since

Statistical properties of the estimators: Unconditional inference

- How do the estimators perform over hypothetical repetitions of RCTs where baseline imbalance randomly varies?
- **Bias:**
  - FV, CS, and ANCOVA are unbiased.
  - Type I error rates will be designated.
- **Precision:**
  - ANCOVA is (generally) more efficient compared with FV, or a CS.

Statistical properties of the estimators: Conditional inference

- How do the estimators perform over hypothetical repetitions of RCTs with the same baseline imbalance?
- **Bias:**
  - ANCOVA is conditionally unbiased.
  - FV and CS is conditionally biased.
- **Precision:**
  - ANCOVA is (generally) more efficient compared with FV, or a CS.
- **ANCOVA is the preferred** analytical method both conditionally and unconditionally.

Meta-analysis in the real world

- **In publication of trials:**
  - Generally only one type of analysis will be reported.
  - Less frequently, an alternative analysis will be provided, or enough information to allow an alternative analysis to be performed.
- The analytical method used may (is likely to) vary across trials.

Meta-analysis Options: A proposed hierarchy of options

1. Obtain individual patient data for each trial and re-analyze these data.
2. Pool using only ANCOVA results. For each trial, recreate the ANCOVA estimate from the summary statistics provided. This will generally involve imputing only the correlation.
3. Pool using results from only one analytical method (Change scores or Final values).
4. Pool using a mix of results from different analytical methods.

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Meta-analysis of Change Score:

Recommendation from Cochrane Handbook for Systematic Reviews of Interventions

Meta-analysis of Change Score

- An analysis based on changes from baseline (CS) will be more efficient and powerful than comparison of final values (FV).
  - Remove between-person variability from the analysis
- However, calculation of a change score (CS) requires measurement of the outcome twice.
- In practice may be less efficient for outcomes which are unstable or difficult to measure precisely.
- Change-from-baseline outcomes (CS) may also be preferred if they have a less skewed distribution than final measurement outcomes (FV).
- Although sometimes used as a device to 'correct' for unlucky randomization, this practice is not recommended.

The Preferred Statistical Approach

- **The preferred statistical approach** to accounting for baseline measurements of the outcome variable:
  - A regression model by include the baseline outcome measurement as a covariate in the model
  - ANCOVA
- These analyses produce an 'adjusted' estimate of the treatment effect together with its standard error.
  - The most precise
  - The least bias
- They should be included in the analysis when they are available.
- They can only be included in a meta-analysis using the generic inverse-variance method, since means and SD are not available for each intervention group separately.

Wisdom of the Land

Change Score vs Final value

- **In practice**, an author is likely to discover that the studies included in a review may include a mixture of CS and FV.
- However, mixing of outcomes is not a problem when it comes to meta-analysis of mean differences.
- There is no statistical reason why studies with CS should not be combined in a meta-analysis with studies with FV when using the (unstandardized) mean difference method.

Change Score vs Final value

- **In a RCT**, mean differences based on CS can usually be assumed to be addressing exactly the same underlying intervention effects as analyses based on FV.
- The difference in FV will on average be the same as the difference in CS.
- If the use of CS does increase precision, the studies presenting CS will appropriately be given higher weights in the analysis than they would have received if FV had been used, as they will have smaller SD.

Pooling Change Score and Final Value

- When combining the data, authors must be careful to use the appropriate means and SD (either of CS or FV) for each study.
- **Since the mean and SD for the two types of outcome may differ substantially**, it may be advisable to place them in separate subgroups to avoid confusion for the reader, but the results of the subgroups can legitimately be pooled together.

Pooling Change Score and Final Value

- However, CS and FV should not be combined together as SMD, since the difference in SD reflects not differences in measurement scale, but differences in the reliability of the measurements.
- **A common practical problem** associated with including CS is that SD of changes is not reported. Imputations of SD is needed.

Summary

Meta-analysis Options:
A proposed hierarchy of options

1. Obtain individual patient data for each trial and re-analyze these data.
2. Pool using only ANCOVA results. For each trial, recreate the ANCOVA estimate from the summary statistics provided. This will generally involve imputing only the correlation.
3. Pool using results from only one analytical method (Change scores or Final values).
4. Pool using a mix of results from different analytical methods.

References

- Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0

- Cochrane Statistical Methods Group: Short Courses SMG Training Course 2010 Slides

  (https://methods.cochrane.org/statistics/)

… Thank you …