Ratio of Mean (RoM) as an Effect Measure in Meta-Analysis of Continuous Outcome

Cardiff SMG short course
March 4-5, 2010

Jan O. Friedrich, Neill Adhikari, Joseph Beyene
University of Toronto and McMaster University

Objectives

- Briefly review conventional mean difference methods for pooling continuous variables in meta-analyses
- Introduce the Ratio of Means (RoM) as a clinically interpretable alternative
  - Compare performance characteristics of RoM to mean difference methods using simulation
  - Compare treatment effects and heterogeneity using an empirical study of a large number of clinically diverse Cochrane meta-analyses pooling continuous outcomes

Conventional Effect Measures in Meta-Analyses of Continuous Outcomes

Continuous Outcomes

- Difference Methods Traditionally Used
  - Mean Difference (MD)
  - Standardised Mean Difference (SMD)
    - mean difference expressed in pooled standard deviation units

Continuous Outcomes - MD

- Mean Difference (MD)
  - Advantages:
    - Easy to interpret
  - Disadvantages:
    - Requires variable to be reported in identical units in all studies

Continuous Outcomes - SMD

- Standardised Mean Difference (SMD)
  - Advantages:
    - Allows pooling of studies when outcomes are measured in different units
    - Allows comparisons of effect sizes across different interventions
      - 0.2 (small), 0.5 (medium), 0.8 (large)
  - Disadvantages:
    - Variance dependent on the actual value of the SMD leading to “negative bias” (towards no treatment effect) and decreased heterogeneity due to lower weighting of extreme values
    - Interpretations requires knowledge of the pooled standard deviation, a quantity generally unknown to clinicians

Conventional Effect Measures - Summary

	Variable Type	
	Binary	Continuous
Difference	RD	MD, SMD
Ratio	RR, OR	?

Ratio of Means (RoM)

An alternative effect measure for continuous outcomes

Ratio of Means (RoM)

RoM = $\frac{\text{mean}_{exp}}{\text{mean}_{control}}$.

– approximate variance can be obtained using the delta method after logarithmic transformation

$$Var[\ln(RoM)] = \frac{1}{n_{\exp}}\left(\frac{sd_{\exp}}{mean_{\exp}}\right)^2 + \frac{1}{n_{contr}}\left(\frac{sd_{contr}}{mean_{contr}}\right)^2$$
$$= \frac{CV^2_{\exp}}{n_{\exp}} + \frac{CV^2_{contr}}{n_{contr}}$$

– apply the generic inverse variance method using the estimate and its standard error (i.e. ln(RoM) and SE[ln(RoM)]) for each study

Ratio of Means (RoM)

- Potential Advantages
  - Like SMD
    - Allows pooling of studies when outcomes are measured in different units
    - Allows comparisons of effect sizes across different interventions
  - Unlike SMD
    - Does not require knowledge of pooled SD
  - Has a form similar to RR, an effect measure understood by clinicians

Ratio of Means (RoM)

- Potential Disadvantages
  - Requires both experimental and control values to have same sign
  - Variance equation approximated using first-order terms (higher order terms excluded)

Motivating Example: MD, SMD and RoM

- Renal physiological outcomes of low-dose dopamine

Effect Measure (*p-value*)/I^2

	MD	SMD	RoM
Urine Output	--	0.5 *(<.001)*	1.24*(<.001)*
		71%	77%
Serum Creatinine	-3.5 μM *(.01)*	-0.3 *(.04)*	0.96 *(.01)*
	73%	79%	73%

Friedrich JO, Adhikari N, Herridge MS, Beyene J. Ann Intern Med. 2005;142:510-24.

Continuous Outcomes – Simulation

Friedrich, Adhikari, Beyene: *BMC Med Res Method* 2008,**8**:32

Continuous Outcomes – Simulation

- Data sets were simulated and meta-analyses were carried out using all three continuous outcome effect measures (MD, SMD, RoM)
  - 10,000 simulated data sets per scenario
  - Random effects model (inverse variance weighting)
- *Parameters Varied:*
  - Number of Patients Per Trial 10, 100
  - Number of Trials 5, 10, 30
  - Standard Deviation (SD) 10, 40, 70 % of mean (set it equal between control and experimental groups)
  - Effect Size 0.2, 0.5, 0.8 (SD units)
  - Heterogeneity (τ) 0, 0.5 (SD units)

Continuous Outcomes – Simulation

Baseline Scenario: stdev 40% mean, effect size 0.5, no heterogeneity

		Bias (% mean)			***%Coverage***		
Pts	Studies	MD	SMD	RoM	MD	SMD	RoM
10	5	0	-4	0	95	97	95
	10	0	-5	0	95	97	95
	30	0	-5	0	94	96	96
100	5	0	0	0	96	97	96
	10	0	0	0	96	96	96
	30	0	0	0	96	96	96

Continuous Outcomes – Simulation

Baseline Scenario: stdev 40% mean, effect size 0.5, no heterogeneity

		Bias (% mean)			***%Power***		
Pts	Studies	MD	SMD	RoM	MD	SMD	RoM
10	5	0	-4	0	64	57	62
	10	0	-5	0	91	89	90
	30	0	-5	0	100	100	100
100	5	0	0	0	100	100	100
	10	0	0	0	100	100	100
	30	0	0	0	100	100	100

Continuous Outcomes – Simulation

Broadened Mean Scenario: stdev 70% mean, effect size 0.5, no heterogeneity

		Bias (% mean)			%Coverage		
Pts	Studies	MD	SMD	RoM	MD	SMD	RoM
10	5	0	-4	-1	95	97	95
	10	0	-5	-2	95	97	95
	30	0	-5	-2	94	96	**92**
100	5	0	0	0	96	97	96
	10	0	0	0	96	96	96
	30	0	0	0	96	96	95

Continuous Outcomes – Simulation

Broadened Mean Scenario with Heterogeneity (τ=0.5): stdev 70% mean, effect size 0.5

		Bias (% mean)			%Coverage		
Pts	Studies	MD	SMD	RoM	MD	SMD	RoM
10	5	0	-5	-1	90	92	91
	10	0	-6	-2	92	93	92
	30	0	-6	-3	94	93	91
100	5	0	0	2	88	88	87
	10	0	0	1	92	92	90
	30	0	0	1	94	94	92

Continuous Outcomes – Simulation

Broadened Mean Scenario with Heterogeneity (τ=0.5): stdev 70% mean, effect size 0.5

		Bias (% mean)			%Coverage		
Pts	Studies	MD	SMD	RoM	MD	SMD	RoM
10	5	0	-5	-1	90	92	91
	10	0	-6	-2	92	93	92
	30	0	-6	-3	94	93	91
100	5	0	0	2	88	88	87
	10	0	0	1	92	92	90
	30	0	0	1	94	94	92

Continuous Outcomes – Simulation

Broadened Mean Scenario with Heterogeneity (τ=0.5): stdev 70% mean, effect size 0.5

		Bias (% mean)			Heterogen (I^2)		
Pts	Studies	MD	SMD	RoM	MD	SMD	RoM
10	5	0	-5	-1	59	48	47
	10	0	-6	-2	60	47	47
	30	0	-6	-3	60	47	48
100	5	0	0	2	93	92	91
	10	0	0	1	93	92	91
	30	0	0	1	93	92	91

Summary of Simulation Results of Continuous Outcomes

- Bias:
  - MD: Low Bias (<1.5%) for all scenarios
  - SMD: Negative bias with small studies (5-6%)
  - RoM: 1) Negative bias with small studies and large within-study standard deviations (3-4%)

    2) Positive bias with large studies and increasing heterogeneity (1-2%)

Summary of Simulation Results of Continuous Outcomes

- Coverage:
  - Relatively similar between methods
  - Close to expected 95% for scenarios without heterogeneity
  - Decreases to low 90% or high 80% range when heterogeneity is introduced

Summary of Simulation Results of Continuous Outcomes

- Statistical Power:
  - As expected, increases with increasing effect size, number of patients, and number or trials, and decreases when heterogeneity is introduced
  - Relatively similar for the three methods in most scenarios
    - Decreased statistical power of SMD or RoM in scenarios where they exhibit negative bias, compared to MD
    - However, the effect of these biases is relatively small so that the differences in statistical power between the effect measures are less than 5 percentage points

Summary of Simulation Results of Continuous Outcomes

- Heterogeneity:
  - In scenarios where SMD and RoM are biased, heterogeneity, expressed as I^2, is lower compared to MD, which is relatively free of bias.
    - This occurs because bias decreases the weighting of the extreme values, decreasing heterogeneity
  - In the scenarios exhibiting less bias, heterogeneity among all methods is more similar

RoM – Summary and Conclusions from Simulation

- RoM
  - Appears to exhibit comparable statistical performance characteristics in terms of bias, coverage, power and heterogeneity, compared to MD and SMD

Continuous Outcomes – Empirical Study

- Compare treatment effects and heterogeneity for MD, SMD, and RoM in a large sample of clinically diverse meta-analyses pooling continuous outcomes

Friedrich, Adhikari, Beyene: Journal of Clinical Epidemiology (revised manuscript submitted)

Continuous Outcomes – Empirical

- Methods:
  - Searched the Cochrane Database of Systematic Reviews for all reviews containing :
    - "wmd" or "weighted mean difference", or
    - "smd" or "standardis(z)ed mean difference"

    in the title, abstract, or keywords
  - Included reviews containing at least one meta-analysis of at least 5 trials and reporting a continuous outcome (MD or SMD)

Continuous Outcomes – Empirical

- Methods:
  - conducted meta-analysis using
    - RoM
    - SMD
    - MD (if possible, i.e. identical units in all trials)

    using inverse variance weighting and random-effects models

Continuous Outcomes – Empirical

- Methods:
  - Differences in p-values
    - treatment effects
    - heterogeneity (Cochran's Q statistic)

    tested using the sign test
  - Pairwise differences between methods
    - treatment effect (threshold p-value of 0.05)
    - heterogeneity (threshold p-value of 0.10 for Q)

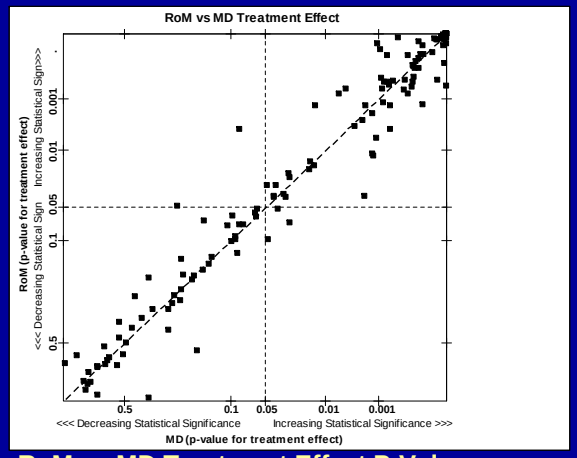
    assessed with Exact tests.

Empirical Study – Results (Search)

- 897/5053 (18%) mentioned WMD and/or SMD in title, abstract or key words
  - 232/897 (26%) included
  - 665/897 (74%) excluded
    - 628 (70%) less than 5 trials
    - 37 (4%) mixture of negative and positive values
- 143/232 (62%) used MD
- 89/232 (38%) used SMD

Empirical Study - Results (Treatment Effects)

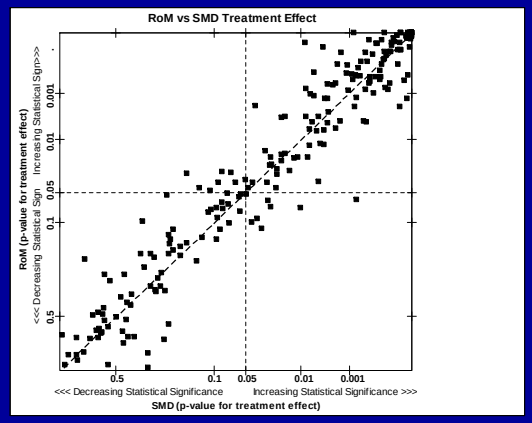
- Median SMD 0.33 (IQR 0.16-0.62)
- Median RoM change away from unity 14% (IQR 7-31%)
- There was no meta-analysis in which two effect measures were both statistically significant and in opposite directions.
  - only 1/143 (MD) and 3/232 (SMD) meta-analyses gave pooled results in the opposite direction to RoM (all p-values >0.3)



RoM vs MD Treatment Effect P-Values

Similar treatment effect p-values (Sign Test p=0.49)

Similar discordant pairs (2 vs 3 [p=1.00])



RoM vs SMD Treatment Effect P-values

Similar treatment effect p-values (Sign Test p=0.21)

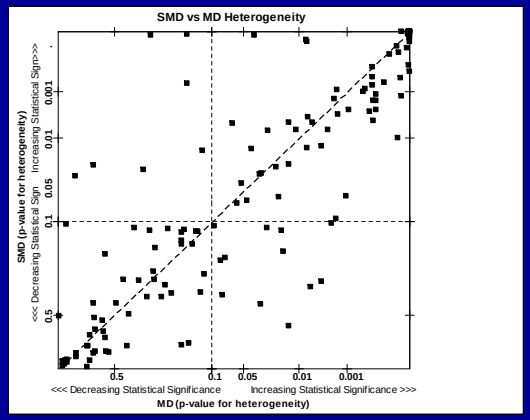
Similar discordant pairs (7 vs 8 [p=1.00])

Empirical Study - Results (Treatment Effects)

	RoM vs MD (n=143)	RoM vs SMD (n=232)	SMD vs MD (n=143)
Median Difference in p-values			
Median Diff	$+7(10^{-16})$	$-3(10^{-10})$	$+2(10^{-12})$
IQR	(-0.001,0.002)	(-0.004,0.003)	(-0.001,0.003)
Sign Test p=	0.49	0.21	0.31
Discordant Pairs			
Number	5 (3%)	15 (6%)	7 (5%)
Distribution	2 vs 3 (p=1.00)	7 vs 8 (p=1.00)	4 vs 3 (p=1.00)

Empirical Study - Results (Heterogeneity)

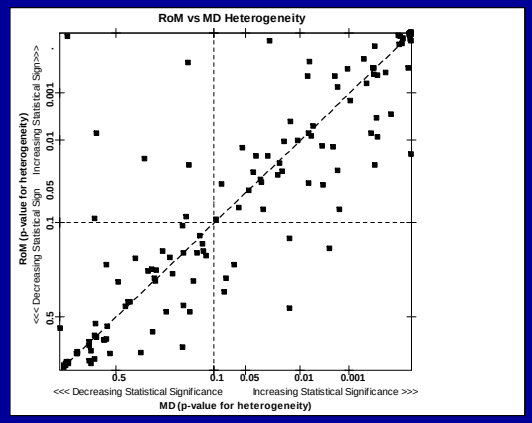
- The percentage of meta-analyses with statistically significant heterogeneity relatively similar
  - Q-statistic p<0.10
    - 61% RoM
    - 58% MD
    - 56% SMD
  - I^2 statistic > 25%
    - 69% RoM
    - 70% MD
    - 66% SMD



SMD vs MD Heterogeneity Q-Stat. P-value

SMD demonstrates lower heterogeneity p-values (p=0.004 by Sign Test)

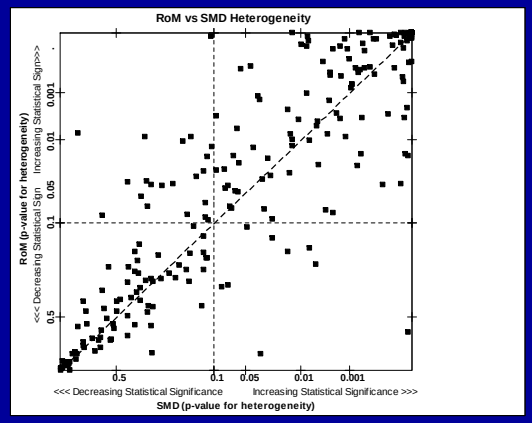
No statistically significant directional assymmetry in discordant pairs (7 vs 12 [p=0.36])



RoM vs MD Heterogeneity Q-Stat. P-values

RoM demonstrates lower heterogeneity p-values (Sign Test p=0.007)

Similar discordant pairs (7 vs 6 [p=1.00])



RoM vs SMD Heterogeneity Q-Stat. P-value

SMD demonstrates lower heterogeneity p-values (Sign Test p=0.005)

SMD demonstrates a lower number of discordant pairs (21 vs 9 [p=0.04])

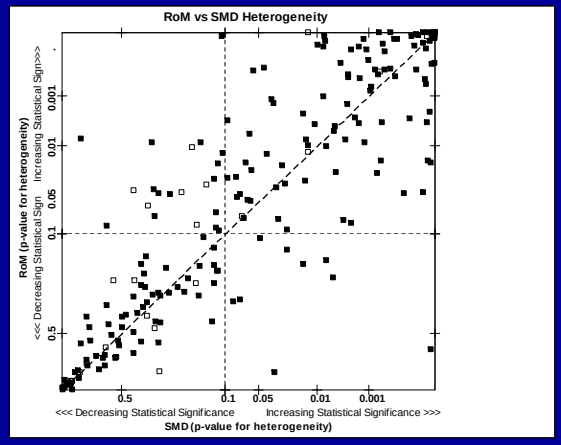
Empirical Study - Results (Heterogeneity)

	RoM vs MD (n=143)	RoM vs SMD (n=232)	SMD vs MD (n=143)
Median Difference in p-values			
Median Diff	$+4(10^{-7})$	$-5(10^{-7})$	$+9(10^{-7})$
IQR	(-0.001,0.019)	(-0.027,0.008)	(-0.010,0.024)
Sign Test p=	0.007	0.005	0.004
Discordant Pairs			
Number	13 (9%)	30 (13%)	19 (13%)
Distribution	7 vs 6 (p=1.00)	21 vs 9 (p=0.04)	7 vs 12 (p=0.36)

Empirical Study - Results (Heterogeneity): RoM vs SMD

Mean patients/trial arm:	Small Trials ≤15 (n=24)	Large Trials >15 (n=208)
Discordant Pairs		
Number	6 (25%)	24 (12%)
Distribution	6 vs 0 (p=0.04)	15 vs 9 (0.31)

This effect is consistent with the known bias of SMD to no effect and less heterogeneity in smaller trials.

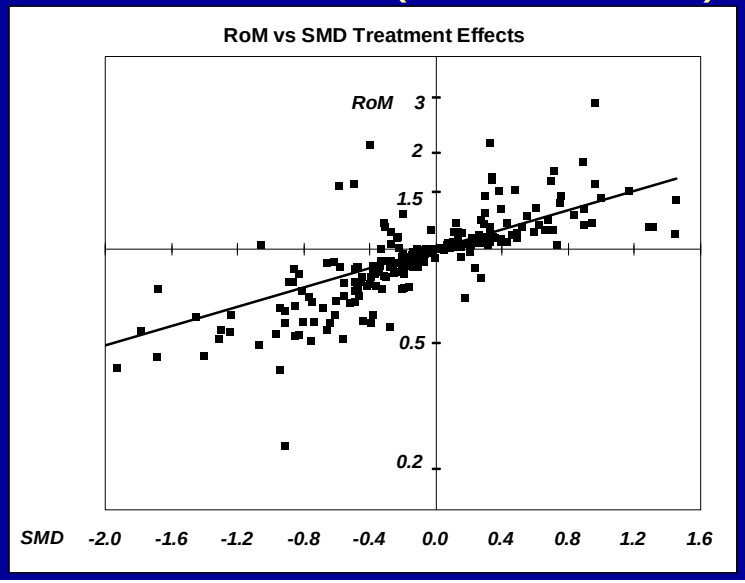


RoM vs SMD Heterogeneity Q-Stat. P-value

SMD demonstrates lower heterogeneity p-values (Sign Test p=0.005)

SMD demonstrates a lower number of discordant pairs (21 vs 9 [p=0.04])

SMD vs. ROM (treat effects)



SMD vs. RoM

- Linear regression without intercept
- ln(RoM) = 0.352 * SMD.
- SMDs of 0.2, 0.5, and 0.8, correspond to increases in RoM of approximately 7%, 19%, and 33%, respectively

RoM – Summary of Empirical Study

- RoM
  - demonstrated similar treatment effect estimates compared to MD and SMD
  - RoM demonstrated less heterogeneity than MD, but more than SMD (SMD demonstrated less heterogeneity than MD)
    - Considering statistically significant heterogeneity (discordant pairs), fewer meta-analyses showed heterogeneity only with SMD compared to RoM and this effect appeared to be restricted to the small trial meta-analyses

RoM – Conclusions based on Empirical Study

- Similar treatment effects among RoM, MD, and SMD
- Some differences in heterogeneity
  - difficult to separate out true differences from the influence of known biases towards no effect and decreased heterogeneity for SMD and RoM under certain conditions

RoM – Overall Summary and Conclusions

- RoM provides the option of using a ratio effect measure in addition to the traditionally used difference methods (MD and SMD)
  - Simulation suggests RoM exhibits comparable performance characteristics in terms of bias, coverage, power and heterogeneity
  - Empirical data suggests RoM yields similar pooled treatment effect estimates and heterogeneity
- RoM should be considered as an alternative effect measure for analysing continuous outcomes in meta-analysis