

# Meta-Analysis of Rare Adverse Events in Randomized Clinical Trials: Bayesian and Frequentist Methods

Gary Rosner<sup>1</sup>   Hwanhee Hong<sup>2</sup>   Chenguang Wang<sup>1</sup>

<sup>1</sup>Division of Biostatistics and Bioinformatics  
Sidney Kimmel Comprehensive Cancer Center  
Johns Hopkins University School of Medicine  
<sup>2</sup>Department of Biostatistics and Bioinformatics  
Duke University School of Medicine

NJ Chapter of the ASA 40th Spring Symposium:  
Clinical Trials for Rare Diseases and Pediatric Populations  
Challenges in Design and Analysis  
28 June 2019



# Introduction

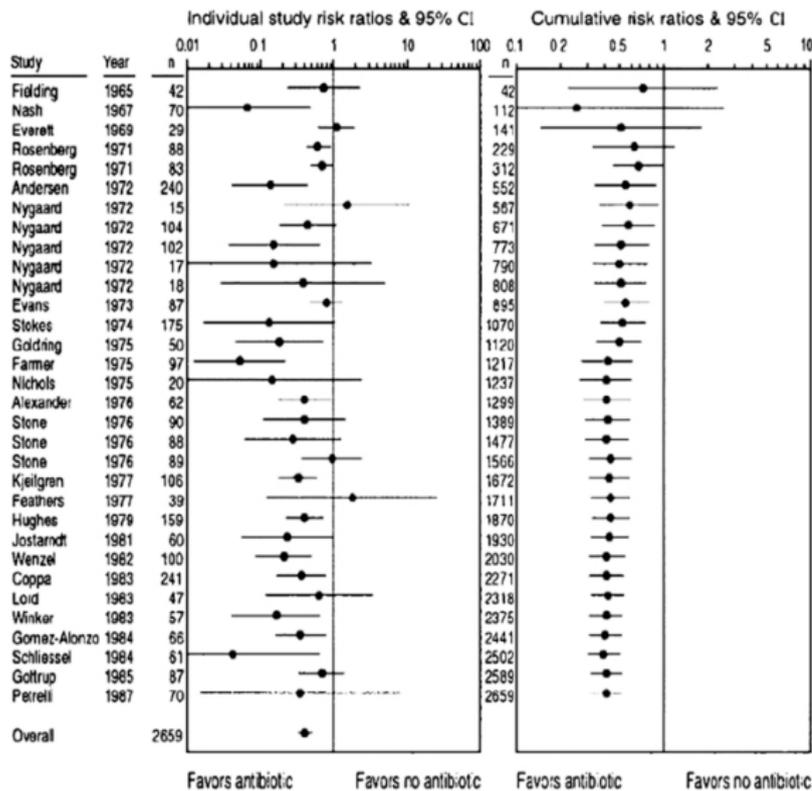
- Regulatory approval of a drug or device involves an assessment of benefits AND risks of adverse events associated with the therapeutic agent or device
- For a rare adverse event, one or two RCTs may not be able to provide enough of a signal of the outcome
- Meta-analysis pools information on events over trials (e.g., Phase 3 RCTs or surveillance studies) to increase the sample size (and statistical information)
  - Larger sample size should improve chance of seeing rare adverse effects
- Some high profile situations have suggested increased risks of uncommon serious AEs and generated controversy
- Many papers compare performance of many frequentist meta-analysis methods in the setting of rare events, but only a few examined Bayesian approaches

# Objective

- We summarize the current state of meta-analysis with rare events.
- We compare both frequentist and Bayesian meta-analytic approaches via extensive simulation studies.
- We will
  - Review of meta-analysis models
  - Data analysis using Nissen and Wolskis rosiglitazone data
  - Simulation study

# Meta-Analysis

- Meta-analysis is a statistical technique for combining the findings from independent studies to assess the clinical effectiveness or safety of healthcare interventions.
- This approach provides a quantitative (statistical) estimate of net benefit and effect heterogeneity aggregated over all the included studies.



# Vioxx (Rofecoxib) and Cardiovascular Side Effects

## Risk of Cardiovascular Events Associated With Selective COX-2 Inhibitors

Debabrata Mukherjee, MD

Steven E. Nissen, MD

Eric J. Topol, MD

**A**SPIRIN AND NONSTEROIDAL anti-inflammatory agents (NSAIDs) have proven anal-

Atherosclerosis is a process with inflammatory features and selective cyclooxygenase 2 (COX-2) inhibitors may potentially have antiatherogenic effects by virtue of inhibiting inflammation. However, by decreasing latory and antiaggregatory prostacyclin production, COX-2 antagonists lead to increased prothrombotic activity. To define the cardiovascular effects of COX-2 inhibitors when used for arthritis and musculoskeletal



<https://www.wsj.com/articles/SB109654671320932405>

## Merck Pulls Vioxx From Market After Link To Heart Problems

Drug's Demise Raises Concerns About Company's Future; Loss of \$2.5

By Barbara Martinez, Anna Wilde Mathews, Joann S. Lublin and Winslow Staff Reporters of The Wall Street Journal

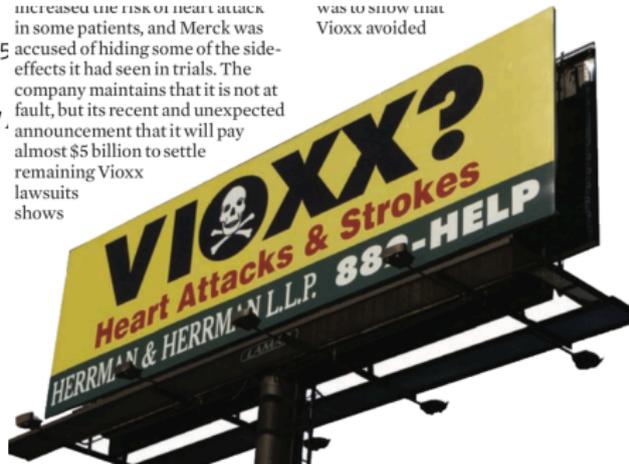
Updated Oct. 1, 2004 12:01 am ET

increased the risk of heart attack in some patients, and Merck was accused of hiding some of the side effects it had seen in trials. The company maintains that it is not at fault, but its recent and unexpected announcement that it will pay almost \$5 billion to settle remaining Vioxx lawsuits

voluntarily withdrew Vioxx from the market in Research published in the medical journal *Lancet* ites that 88,000 Americans had heart attacks from Vioxx, and 38,000 of them died.

*n McDermaid/EPA/Corbis*

was to show that Vioxx avoided



# Avandia (Rosiglitazone) and Cardiovascular Side Effects



Health » Food | Fitness | Wellness | Parenting | Live Longer

## Avandia and diabetes: Was revolution worth the risks?

By **Caleb Hellerman**, CNN  
July 8, 2010 1:18 p.m. EDT



The New York Times

Prescriptions  
Making Sense of the Health Care Debate

## F.D.A. Panel Votes to Restrict Avandia

By **Gardiner Harris** July 14, 2010 8:59 am

A majority of the advisory panel of the Food and Drug Administration voted today to restrict the sales of Avandia, a controversial diabetes drug, because of its potential risk for causing heart attacks. The 33-member advisory committee was deeply divided. Twelve voted to remove Avandia from the market altogether; 10 for continued sale but with new label revisions and possible restrictions; 7 to add more warnings and 3 for no change at all. But the votes can also be viewed as a decision by a majority to continue allowing sales of Avandia, with more restrictions. A

## The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

JUNE 14, 2007

VOL. 356 NO. 24

## Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes

Steven E. Nissen, M.D., and Kathy Wolski, M.P.H.

REVIEW ARTICLE

ONLINE FIRST

## Rosiglitazone Revisited

An Updated Meta-analysis of Risk for Myocardial Infarction and Cardiovascular Mortality

Steven E. Nissen, MD; Kathy Wolski, MPH

ARCH INTERN MED/VOL 170 (NO. 14), JULY 26, 2010

I want the FDA to order the company to withdraw the drug. The evidence is there.

—Dr. Steven Nissen, Cleveland Clinic

Let's treat our patients instead of sensationalizing half-truths and conclusions that are not yet formed.

## Epidemiology and Prevention

### Efficacy and Safety of Varenicline for Smoking Cessation in Patients With Cardiovascular Disease

#### A Randomized Trial

Nancy A. Rigotti, MD; Andrew L. Pipe, CM, MD; Neal L. Benowitz, MD; Carmen Arteaga, PhD;  
Dahlia Garza, MD; Serena Tonstad, MD, PhD, MPH

- Varenicline: First approved nicotinic receptor partial agonist for smoking cessation
- RCT in patients with CV disease (Rigotti et al. *Circulation*. 2010;121:221-229)
  - Conclusions: Varenicline effective for smoking cessation in smokers with CV disease. "It was well tolerated and did not increase cardiovascular events or mortality"

# Meta-analysis around the same time

(Singh S et al. *CMAJ*. 2011;183:1359-66.)

CMAJ

RESEARCH

## **Risk of serious adverse cardiovascular events associated with varenicline: a systematic review and meta-analysis**

Sonal Singh MD MPH, Yoon K. Loke MBBS MD, John G. Spangler MD MPH, Curt D. Furberg MD PhD

- 14 double-blind RCTs involving 8216 participants
- Selected pubs that reported CV events as “serious adverse events associated with the use of varenicline.”
- Conclusion: “Varenicline was associated with a significantly increased risk of serious adverse cardiovascular events compared with placebo”
- 1.06% [52/4908] varenicline group v. 0.82% [27/3308] placebo group
- Peto OR 1.72, 95% CI (1.09, 2.71); I<sup>2</sup> = 0%



# Varenicline (cont'd): Network Meta-Analysis

- Network meta-analysis exploring CV AEs with pharmacotherapies for smoking cessation (Mills et al. *Circulation*. 2014;129:28-41)

## Epidemiology and Prevention

### Cardiovascular Events Associated With Smoking Cessation Pharmacotherapies A Network Meta-Analysis

Edward J. Mills, PhD, MSc; Kristian Thorlund, PhD, MStat; Shawn Eapen, MSc;  
Ping Wu, MBBS, MSc; Judith J. Prochaska, PhD, MPH

- 10 electronic databases and accessed internal FDA reports
- Included any RCT of 3 approved trts (nicotine replacement therapy, bupropion, & varenicline) that reported CV disease outcomes
- 63 eligible RCTs involving 21 nicotine replacement therapy RCTs, 28 bupropion RCTs, and 18 varenicline RCTs. They found
  - “...no increase in the risk of all cardiovascular disease events with bupropion” (RR, 0.98; 95% CI, 0.54–1.73)
  - “or varenicline (RR, 1.30; 95% CI, 0.79–2.23)
  - Re: major adverse CV events: “...found a protective effect with bupropion (RR, 0.45; 95% CI, 0.21–0.85) and no clear evidence of harm with varenicline (RR, 1.34; 95% CI, 0.66–2.66) or nicotine replacement therapy (RR, 1.95; 95% CI, 0.26–4.30).”
  - Conclusion: “Smoking cessation therapies do not appear to raise the risk of serious cardiovascular disease events.”

# What to Do with Meta-Analyses with Rare Events?

## Issues in Meta-Analysis with Rare Events<sup>1</sup>

- Sparsity of data with zero-event trials
- Insufficient statistical power in random effects models
- Potential dominance of a few large trials

Several authors argue that traditional meta-analysis methods may be ill-defined or have poor performance properties with rare events

- Stoto MA. *Drug Safety*, 2015;38(3):233-243
- Bennetts M et al. *Research Synthesis Methods*, 2017;8(1):64-78
- Shuster JJ and Walker MA. *Statistics in Medicine*, 2016;35:2467-2478
- Bradburn MJ et al. *Statistics in Medicine*, 2007;26(1):53-77
- Chakravarty AG and Levenson M. *Quantitative Evaluation of Safety in Drug Development: Design, Analysis and Reporting* 67 2014;67:237
- Lane PW. *Statistical Methods in Medical Research*, 2013;22(2):117-132
- Sweeting MJ et al. *Statistics in Medicine*, 2004;23(9):1351-75

# Some Alternative Meta-Analytic Methods for Rare Events

- Poisson random-effects models using a likelihood-based approach
  - Cai T et al. *Statistics in Medicine*, 2010;29(20):2078-2089
- Combining confidence intervals
  - Tian L et al. *Biostatistics*, 2009;10(2):275-281
- Using an arcsine difference
  - Rücker et al. *Statistics in Medicine*, 2009;28(5):721-738

## A common limitation of these simulation studies

They generated simulated data sets assuming no heterogeneity in treatment effect across trials

# Avandia (Rosiglitazone)

- We use the Nissen and Wolski data published in 2010.
- NW collected the data to investigate the effect of rosiglitazone on an increase in the risk of myocardial infarction (MI) or cardiovascular causes death (CV death) compared to the control group.
- The total number of trials is 56
  - 15 trials did not report any MI events.
  - 29 trials did not report any CV death events.
- Control groups in each trial vary widely.



Table 1. Rosiglitazone Clinical Trials Included in the Meta-analysis

Clinical Trial No.	Phase	No. of Weeks	Rosiglitazone Therapy		Control Therapy	
			Agent	No.	Agent	No.
<b>Trials Included in Original Registration Package</b>						
49653/011	3	24	Rosiglitazone	357	Placebo	176
49653/020	3	52	Rosiglitazone	391	Glyburide	207
49653/024	3	26	Rosiglitazone	774	Placebo	185
49653/093	3	26	Rosiglitazone±metformin	213	Metformin	109
49653/094	3	26	Rosiglitazone and metformin	232	Metformin	116
<b>Subtotal</b>				<b>1967</b>		<b>793</b>
<b>Additional Phase 2, 3, and 4 Efficacy Trials</b>						
100684	4	52	Rosiglitazone and glyburide	43	Glyburide	47
49653/143	4	24	Rosiglitazone and glyburide	121	Glyburide	124
49653/211	4	52	Rosiglitazone and usual care	110	Usual care	114
49653/284	4	24	Rosiglitazone and metformin	382	Metformin	384
712753/008	4	48	Rosiglitazone and metformin	284	Metformin	135
AVM100264	4	52	Rosiglitazone and metformin	294	Metformin and sulfonylurea	302
BRL 49653C/185	4	32	Rosiglitazone±metformin	563	Usual care±metformin	142
BRL 49653/334	4	52	Rosiglitazone	278	Placebo	279
BRL 49653/347	4	24	Rosiglitazone and insulin	418	Insulin	212
49653/015	3	24	Rosiglitazone and sulfonylurea	395	Sulfonylurea	198
49653/079	3	26	Rosiglitazone±glyburide	203	Glyburide	106
49653/080	3	156	Rosiglitazone	104	Glyburide	99
49653/082	3	26	Rosiglitazone and insulin	212	Insulin	107
49653/085	3	26	Rosiglitazone and insulin	138	Insulin	139
49653/095	3	26	Rosiglitazone and insulin	196	Insulin	96
49653/097	3	156	Rosiglitazone	122	Glyburide	120
49653/125	3	26	Rosiglitazone and sulfonylurea	175	Sulfonylurea	173
49653/127	3	26	Rosiglitazone and glyburide	56	Glyburide	58
49653/128	3	28	Rosiglitazone	39	Placebo	38
49653/134	3	28	Rosiglitazone	561	Placebo	276
49653/135	3	104	Rosiglitazone and glypizide	116	Glypizide	111

# FDA, EMA, WHO Classifications According to Frequency of Adverse Drug Reactions

Characterization	Frequency
Very common	$\geq 10\%$
Common (frequent)	$\geq 1\%$ to $< 10\%$
Uncommon (infrequent)	$\geq 0.1\%$ to $< 1\%$
Rare	$\geq 0.01\%$ to $< 0.1\%$
Very rare	$< 0.01\%$

# Literature Review

- We reviewed about 20 papers.
  - Most papers (Diamond et al. (2007); Friedrich et al. (2009); Bohning et al. (2015); Lane (2013)) re-analyzed the Nissen and Wolski data using various existing (mostly frequentist) meta-analysis models.
  - Sweeting et al. (2004) and Bradburn et al. (2007) show extensive simulation studies to compare many frequentist methods; datasets generated only under a fixed-effects model framework
    - Sweeting et al. (2004): MH with modification (+0.5) gives least biased results when sample sizes in treated and control groups are unbalanced.
    - Bradburn et al. (2007): Peto's estimator provides least biased results when event rates  $< 0.01$
  - There are a few frequentist method papers
    - Cai et al. (2010) use Poisson regression
    - Tian et al. (2009) combine confidence intervals on risk difference scale
  - Shuster et al. (2007, 2016) propose a new ratio estimator approach forcing all studies to receive the same weight.

# Assumptions in a Meta-Analysis

- One can make two basic model assumptions in a meta-analysis
  - **Common Treatment Effect (CTE)**: treatment effects are constant across trials (a.k.a. fixed-effects model)
  - **Heterogeneous Treatment Effect (HTE)**: treatment effects are heterogeneous across trials (a.k.a. random-effects model)

# Methods We Compare

## Frequentist

- CTE
  - Naïve
  - Peto
  - Mantel-Haenszel (MH)
  - MH-Data Modification (DM)
  - Inverse-Variance-DM
  - SGS-Unweighted
  - SGS-Weighted
- HTE
  - DerSimonian & Laird IV-DM
  - Simple Average
    - incl. between-study heterogeneity as in DerSimonian & Laird ( $\tau_{DL}^2$ )

## Bayesian

- CTE
  - Logistic Regression
  - Beta Hyperprior
- HTE
  - Logistic Regression
  - Logistic w/ Shrinkage prior
  - Arm-based model
  - Beta Hyperprior

Table: Generic 2x2 table for study  $i$

Group	No. of Events	No. of Nonevents	Total
Control	$y_{i1}$	$n_{i1} - y_{i1}$	$n_{i1}$
Treatment	$y_{i2}$	$n_{i2} - y_{i2}$	$n_{i2}$
Total	$y_{i\cdot}$	$n_{i\cdot} - y_{i\cdot}$	$n_{i\cdot}$

- $i$  indexes study
- Assume outcomes follow binomial distributions

$$y_{i1} \mid p_{i1} \sim \text{Bin}(n_{i1}, p_{i1}) \text{ and } y_{i2} \mid p_{i2} \sim \text{Bin}(n_{i2}, p_{i2})$$

# Data Modification (DM) for Cells with Zeros

If study has at least one cell with 0 events, add a constant (usually 0.5) to each cell in the study's 2x2 table

	No. of events	No. of nonevents	Total
Control group	$y_{i1} + 0.5$	$n_{i1} - y_{i1} + 0.5$	$n_{i1} + 1$
Treated group	$y_{i2} + 0.5$	$n_{i2} - y_{i2} + 0.5$	$n_{i2} + 1$
Total	$y_{i\cdot} + 1$	$n_{i\cdot} - y_{i\cdot} + 1$	$n_{i\cdot} + 2$

- Adding 0.5 to each cell removes 1<sup>st</sup>-order bias
  - (Cox & Snell, 1985; Bhaumik et al., 2012)

For our estimators...

- Mantel-Haenszel estimator with (MH(DM)) this modification
- Simple average (SA) adds 0.5 to each cell across all studies
- Inverse variance estimators include modification (CTE-IV(DM) and HTE-DL(DM))

# Naïve Pooling

Table: 2x2 table after pooling across studies

	No. of events	No. of nonevents	Total
Control group	$\sum_i y_{i1}$	$\sum_i (n_{i1} - y_{i1})$	$\sum_i n_{i1}$
Treated group	$\sum_i y_{i2}$	$\sum_i (n_{i2} - y_{i2})$	$\sum_i n_{i2}$
Total	$\sum_i y_{i\cdot}$	$\sum_i (n_{i\cdot} - y_{i\cdot})$	$\sum_i n_{i\cdot}$

- We first estimate  $\hat{p}_1 = \frac{\sum_i y_{i1}}{\sum_i n_{i1}}$  and  $\hat{p}_2 = \frac{\sum_i y_{i2}}{\sum_i n_{i2}}$
- Then,

$$\widehat{LOR}_{naive} = \log \left[ \frac{\hat{p}_2(1 - \hat{p}_1)}{\hat{p}_1(1 - \hat{p}_2)} \right]$$

with

$$se(\widehat{LOR}_{naive}) = \sqrt{\frac{1}{\sum_i y_{i1}} + \frac{1}{\sum_i n_{i1} - y_{i1}} + \frac{1}{\sum_i y_{i2}} + \frac{1}{\sum_i n_{i2} - y_{i2}}}$$

# Peto's Method

The Peto\* estimator of LOR and its variance are

$$\widehat{LOR}_{Peto} = \frac{\sum_i O_i - \sum_i E_i}{\sum_i V_i}$$
$$\widehat{Var}(\widehat{LOR}_{Peto}) = 1 / \sum_i V_i,$$

where

- $O_i = y_{i2}$
- $E_i = \frac{y_{i \cdot} \cdot n_{i2}}{n_{i \cdot}}$
- $V_i = \frac{y_{i \cdot} \cdot n_{i1} \cdot n_{i2} \cdot (n_{i \cdot} - y_{i \cdot})}{n_{i \cdot}^2 \cdot (n_{i \cdot} - 1)}$
- $n_{i \cdot} = n_{i1} + n_{i2}$
- $y_{i \cdot} = y_{i1} + y_{i2}$

\* Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: An overview of the randomised trials. *Progress in Cardiovascular Diseases* 1985;27(5):335-371.

# Mantel-Haenszel Estimator

The MH method pools odds ratios across studies. The MH estimator of LOR and its approximate variance proposed by Robins et al.\*:

$$\widehat{LOR}_{MH} = \log \left[ \frac{\sum_i \frac{y_{i2}(n_{i1}-y_{i1})}{n_{i.}}}{\sum_i \frac{y_{i1}(n_{i2}-y_{i2})}{n_{i.}}} \right]$$

$$\widehat{Var}(\widehat{LOR}_{MH}) = \left[ \frac{\sum_i P_i R_i}{2(\sum_i R_i)^2} + \frac{\sum_i (P_i S_i + Q_i R_i)}{2(\sum_i R_i)(\sum_i S_i)} + \frac{\sum_i Q_i S_i}{2(\sum_i S_i)^2} \right],$$

where

- $R_i = \frac{y_{i2}(n_{i1}-y_{i1})}{n_{i.}}$ ,  $S_i = \frac{y_{i1}(n_{i2}-y_{i2})}{n_{i.}}$
- $P_i = \frac{y_{i2}+n_{i1}-y_{i1}}{n_{i.}}$ ,  $Q_i = \frac{y_{i1}+n_{i2}-y_{i2}}{n_{i.}}$

Recall: MH(DM) adds 0.5 to cells if study has a 0

\*Robins J, Breslow N, Greenland S. Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics* 1986;**42**:311-323.

# Inverse-Variance Estimator (IV)

DerSimonian and Laird (1986) proposed a weighted estimator.

- Study-specific estimator:  $LOR_i = \log \left( \frac{y_{i2}(n_{i1}-y_{i1})}{y_{i1}(n_{i2}-y_{i2})} \right)$
- Study-specific weights  $w_i^*$  combine within- & between-study variances

$$w_i^* = \frac{1}{v_i^*} = \frac{1}{\frac{1}{w_i} + \tau^2}$$

- $1/w_i$  is within-study variance (i.e., variance of  $LOR_i$  for study  $i$ )
- $\tau^2$  is between-study variance
- For CTE, we set  $\tau^2 = 0$ , denoted by CTE-IV(DM).
- For HTE, we estimate  $\tau^2$ , denoted by HTE-DL(DM).

Recall: The inverse-variance estimators add 0.5 to cells if study has a 0

## Inverse-Variance Estimator (IV) cont'd

- Estimate  $\tau^2$  as

$$\hat{\tau}^2 = \begin{cases} \frac{Q-df}{C}, & \text{if } Q > df \\ 0, & \text{if } Q \leq df, \end{cases}$$

- $df$  is the number of studies minus 1
- $Q = \sum_i w_i (LOR_i - \widehat{LOR})^2$
- $\widehat{LOR} = \frac{\sum_i w_i LOR_i}{\sum_i w_i}$
- $C = \sum_i w_i - \frac{\sum_i w_i^2}{\sum_i w_i}$

### The IV estimator and its variance:

$$\widehat{LOR}_{IV} = \frac{\sum_i w_i^* LOR_i}{\sum_i w_i^*}$$
$$\widehat{Var}(\widehat{LOR}_{IV}) = 1 / \sum_i w_i^*$$

# Shuster and colleagues\*

- In a series of papers, Shuster and colleagues argue that most meta-analyses use incorrect estimators.
- Most MA estimation assumes *effects at random*
  - Treatment effect and study design are independent
  - **Consequence:** Assumes weights (e.g., sample size, variance, etc.) not correlated with trt effect
- Alternative framework assumes *studies at random*
  - Treatment effect and study design are **not** independent
  - **Consequence:** Weights **are** correlated with treatment effect
- Let  $\hat{p}_{jk} = \sum_i Y_{ijk} / n_{jk}$ ,  $k = 1, 2; j = 1, \dots, M$
- Unweighted estimator of relative risk:  $RR = \frac{\sum_j \hat{p}_{j2}}{\sum_j \hat{p}_{j1}}$
- Weighted estimator of relative risk:  $RR = \frac{\sum_j n_j \hat{p}_{j2}}{\sum_j n_j \hat{p}_{j1}}$

\*E.g., Shuster, Guo, Skyler. *Res Syn Meth*, 2012;3:30-50.

# Shuster et al. (cont'd): Estimators of Odds Ratio (OR)

- Unweighted estimator of OR (SGS-Unwgt):
  - Let  $\hat{\pi}_k = \sum_j \hat{p}_{jk} / M$ 
    - Trt-specific estimate averaged across studies
  - $\text{SGS-Unwgt} = \frac{\hat{\pi}_2(1 - \hat{\pi}_1)}{\hat{\pi}_1(1 - \hat{\pi}_2)}$
  - SGS\* give formula for standard error of  $\text{Log}(\text{SGS-Unwgt})$
  - For confidence intervals: non-central  
 $t_{M-1}(\cdot \mid \text{Log}(\text{SGS-Unwgt}), \text{StErr-Unwgt})$
- Weighted estimator of OR (SGS-Wgt):
  - Study-specific weight =  $U_j = (n_{j1} + n_{j2})/2$
  - Let  $A_{jk} = U_j \hat{p}_{jk}$  and  $B_{jk} = U_j(1 - \hat{p}_{jk})$ .
  - Let  $\bar{A}_k = \sum_j A_{jk} / M$  and  $\bar{B}_k = \sum_j B_{jk} / M$
  - $\text{SGS-Wgt} = \frac{\bar{A}_2 \bar{B}_1}{\bar{A}_1 \bar{B}_2}$
  - SGS\* give formula for standard error of  $\text{Log}(\text{SGS-Wgt})$
  - For confidence intervals: non-central  
 $t_{M-2}(\cdot; \log(\text{SGS-Unwgt}), \text{StErr-Wgt})$

# Simple Average (SA) Estimator

- Bhaumik et al. (2012): Compute average of study-specific LORs after adding 0.5 to *all* cells in each study's 2x2 table
  - The study-specific  $LOR_i$  (trt 2 vs trt 1):

$$\widehat{LOR}_i = \log\left(\frac{y_{i2} + 0.5}{n_{i2} - y_{i2} + 0.5}\right) - \log\left(\frac{y_{i1} + 0.5}{n_{i1} - y_{i1} + 0.5}\right)$$

- The SA estimator and its variance are defined as

$$\widehat{LOR}_{SA} = \frac{\sum_i \widehat{LOR}_{i,1/2}}{M}$$
$$\widehat{Var}(\widehat{LOR}_{SA}) = \frac{\sum_i \hat{\sigma}_i^2(\hat{\tau}^2)}{M^2}$$

- $M$  is the number of studies
- $\hat{\sigma}_i^2(\hat{\tau}^2) = [n_{i1}\hat{p}_{i1}(1 - \hat{p}_{i1})]^{-1} + [n_{i2}\hat{p}_{i2}(1 - \hat{p}_{i2})]^{-1} + \hat{\tau}^2$
- $\hat{p}_{ik} = \frac{y_{ik} + 0.5}{n_{ik} + 1}$  for  $k = 1, 2$
- $\hat{\tau}^2$ , inter-study var, as for IV estimator (DerSimonian & Laird)

# Note: Compare Shuster's Estimator to Simple Average

- Shuster's estimators:
  - ① Calculate study-specific risk,  $\hat{p}_{k_i}$ ,  $k = 1, 2$
  - ② Average them,  $E[p_{k_i}]$
  - ③ Calculate  $LOR$
- SA
  - ① Add 0.5 to **all** cells (treatment groups and studies)
  - ② Calculate study-specific  $LOR_i$  and then
  - ③ Average them  $E(LOR_i)$

Shuster's and SA estimates provide different results

# Bayesian Approaches

We consider three different Bayesian hierarchical meta-analysis models

- 1 Logistic regression
  - 2 Arm-based approach
  - 3 Adopting a beta hyperprior
- Also consider constant (**CTE**) and heterogeneous (**HTE**) treatment effects with Bayesian models
    - For logistic regression and beta hyperprior models, we consider **CTE** and **HTE** models
    - We assume only **HTE** for arm-based models

# Bayesian Logistic Regression

Assume the following for binomial parameter  $p_{ik}$  for study  $i$  trt  $k$

$$\text{CTE model: } \text{logit}(p_{ik}) = \mu_i + d * I(k = 2)$$

$$\text{HTE model: } \text{logit}(p_{ik}) = \mu_i + \delta_i * I(k = 2)$$

## CTE

- $\mu_i \sim N(0, 100)$ : baseline log odds for study  $i$ 's control group
- $d \sim N(0, 100)$ : Assumed common log odds ratio

HTE: Distinguish types of heterogeneity

- HTE-Logit assumes  $\delta_i | d, \tau \sim N(d, \tau^2)$ ,  $d \sim N(0, 100)$ ,  $\tau \sim \text{Unif}(0, 2)$ 
  - $\tau$  measures the heterogeneity of LOR across studies
- HTE-LogitSh adds shrinkage prior for baseline effect  $\mu_i \sim N(m_\mu, \tau_\mu^2)$ 
  - $m_\mu \sim N(0, 100)$  and  $\tau_\mu \sim \text{Unif}(0, 2)$

# Bayesian Arm-Based Model

- Model each arm's risk within each study separately
  - Based on Hong et al. (2016) *arm-based* parameterization in network meta-analysis
- We simplify this model to fit an arm-based meta-analysis

$$\text{logit}(p_{ik}) = \theta_k + \eta_{ik}$$

- $\theta_k = \log$  odds in arm  $k = 1, 2$
- $\eta_{ik}$  = random to allow trt-specific risk heterogeneity across studies
- We assume  $(\eta_{i1}, \eta_{i2})^T \sim N_2((0, 0)^T, \Sigma)$
- We use priors  $\theta_k \sim N(0, 10^2)$  and  $\Sigma^{-1} \sim \text{Wishart}(\Omega, 2)$ 
  - $\Omega = 0.02 \times I_2$
- Log odds ratio can be obtained using  $\theta_2 - \theta_1$

# Bayesian CTE with Beta Priors

- Consider beta prior distributions
- For CTE, likelihood is rewritten as

$$y_{i1} \mid p_1 \sim \text{Bin}(n_{i1}, p_1) \text{ and } y_{i2} \mid p_2 \sim \text{Bin}(n_{i2}, p_2)$$

- Priors, such as

$$p_1 \sim \text{Beta}(\alpha, \beta) \text{ and } p_2 \sim \text{Beta}(\alpha, \beta)$$

- $\alpha$  &  $\beta$  are prespecified.
- We used  $\alpha = \beta = 1$

# Bayesian: HTE with Beta Hyperprior

- For HTE, likelihood is

$$y_{i1} \mid p_{i1} \sim \text{Bin}(n_{i1}, p_{i1}) \text{ and } y_{i2} \mid p_{i2} \sim \text{Bin}(n_{i2}, p_{i2})$$

- Consider beta priors but different parameterization

$$p_{i1} \sim \text{Beta}(U_1 V_1, (1 - U_1) V_1) \text{ and } p_{i2} \sim \text{Beta}(U_2 V_2, (1 - U_2) V_2)$$

- Note:  $E[p_{ik}] = U_k$  and  $\text{var}[p_{ik}] = U_k(1 - U_k)/(V_k + 1)$
- We assign Beta(1,1) hyperpriors to  $U_k$  and vague gamma hyperpriors to  $V_k$ ,  $k = 1, 2$

- 
- If  $X \sim \text{Beta}(\mu\phi, (1 - \mu)\phi)$ ,  $E[X \mid \mu, \phi] = \mu$  and  $\text{var}[X \mid \mu, \phi] = \frac{\mu(1-\mu)}{\phi+1}$
  - In terms of  $\text{Beta}(\alpha, \beta)$  parameterization,  $\mu = \frac{\alpha}{\alpha+\beta}$  and  $\phi = \alpha + \beta$

# Simulation Studies of 15 Models: 9 Freq & 6 Bayesian

- Compare (1) bias, (2) MSE, & (3) coverage prob of LOR estimates
- Each simulated meta-analysis dataset has 30 studies
- Sample sizes: Control group's  $n_1 \sim Unif(50, 1000)$  &  $n_2 = n_1$
- Risk for trt  $k = 1, 2$  in study  $i$ :
  - $p_{ik} \sim Unif(p_k(1 - 0.5D), p_k(1 + 0.5D))$
  - $p_k$  depends on scenario
  - $D$  controls variability of risks:  $D = 0, 1, 2$
- We consider the following risk parameters,  $p_k$

Null				Alternative			
$p_1$	$p_2$	OR	LOR	$p_1$	$p_2$	OR	LOR
0.002	0.002	1	0	0.002	0.004	2.00	0.70
0.005	0.005	1	0	0.005	0.01	2.01	0.70
0.05	0.05	1	0	0.05	0.1	2.11	0.74

- We consider 18 scenarios, 10,000 simulated meta-analyses per scenario

# Methods We Compare [Abbreviation]

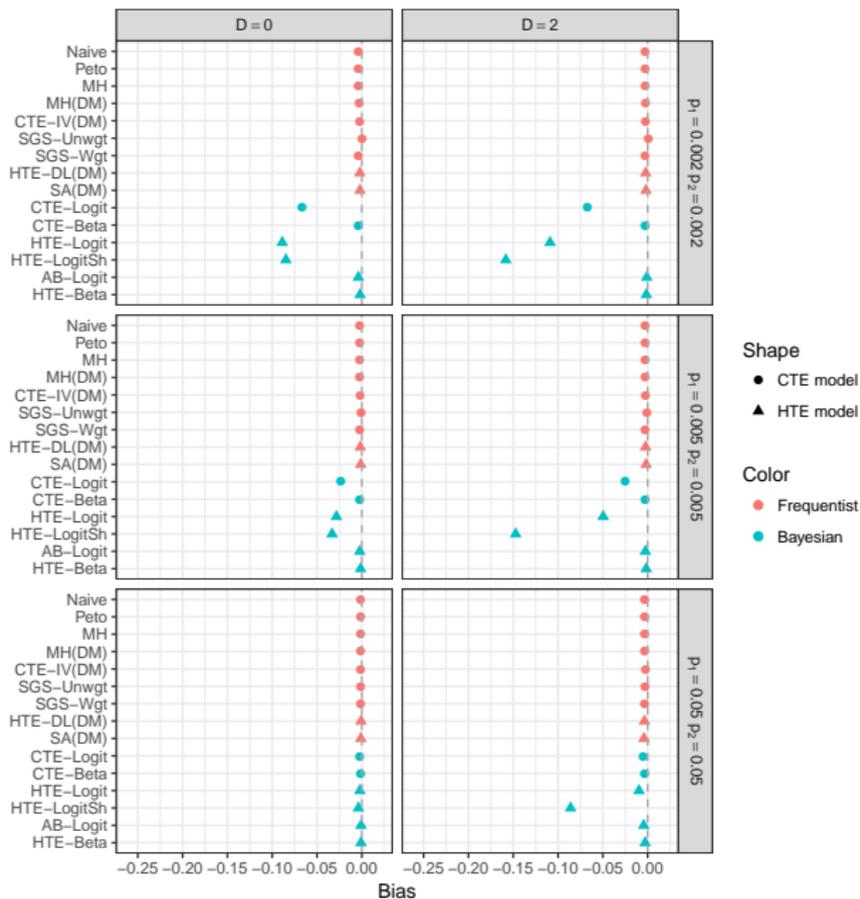
## Frequentist

- CTE
  - Naïve [Naive]
  - Peto [Peto]
  - Mantel-Haenszel (MH) [MH]
  - MH-Data Modification (DM) [MH(DM)]
  - Inverse Variance-DM [CTE-IV(DM)]
  - SGS-Unweighted [SGS-Unwgt]
  - SGS-Weighted [SGS-Wgt]
- HTE
  - DerSimonian & Laird IV-DM [HTE-DL(DM)]
  - Simple Average with  $\tau_{DL}$  [SA(DM)]

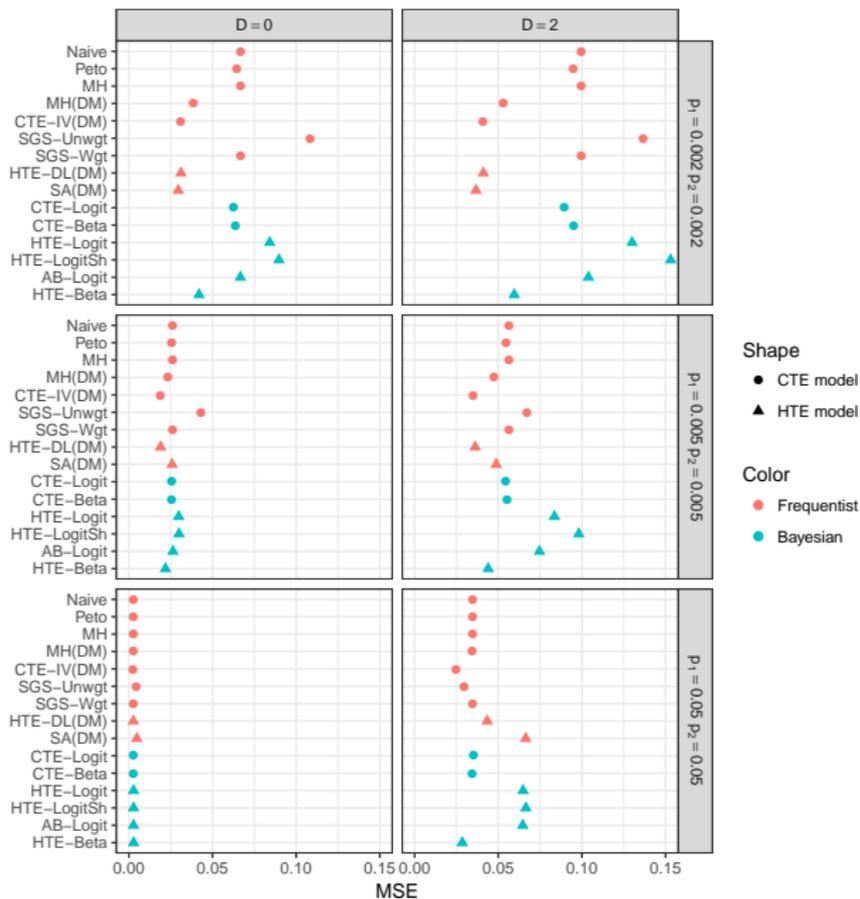
## Bayesian

- CTE
  - Logistic Regression [CTE-Logit]
  - Beta Hyperprior [CTE-Beta]
- HTE
  - Logistic Regression [HTE-Logit]
  - Logistic w/ Shrinkage prior [HTE-LogitSh]
  - Arm-based model [AB-Logit]
  - Beta Hyperprior [HTE-Beta]

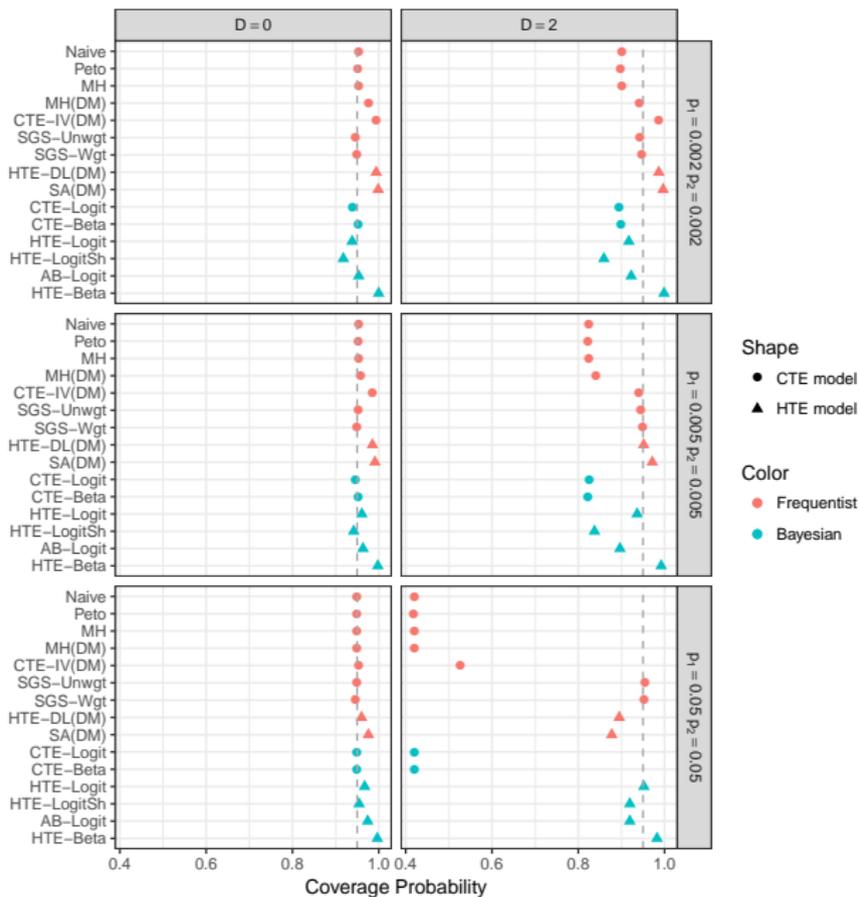
# Results for Log Odds Ratio: Null Case - Bias



# Results for Log Odds Ratio: Null Case - MSE



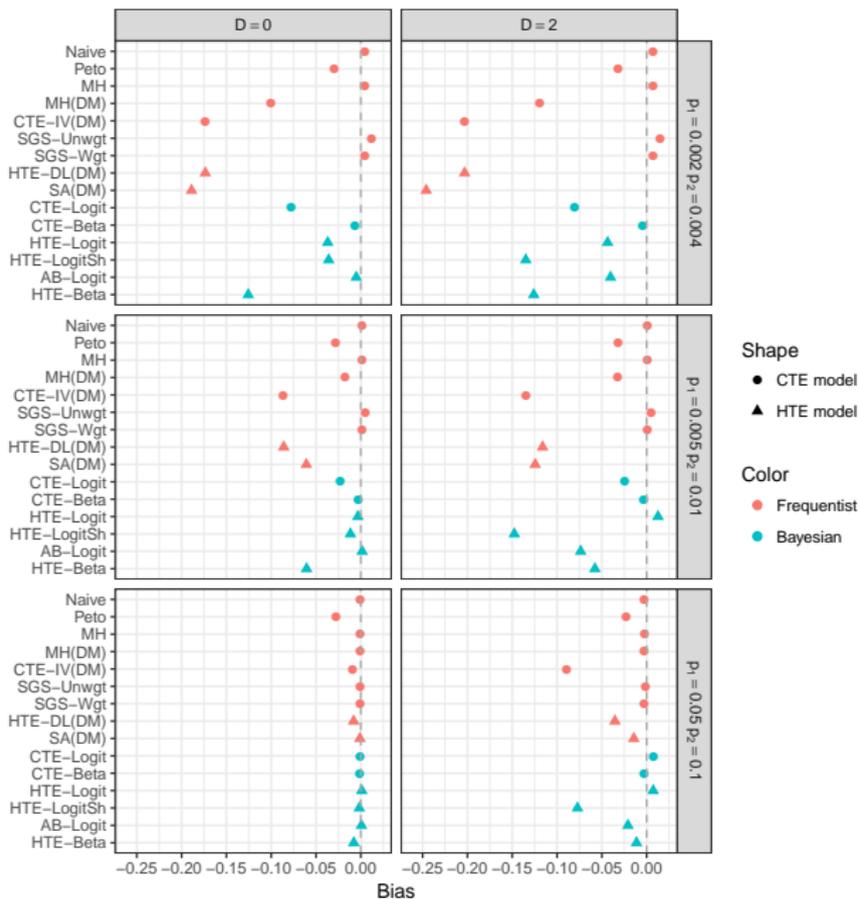
# Results for Log Odds Ratio: Null Case - Coverage Prob



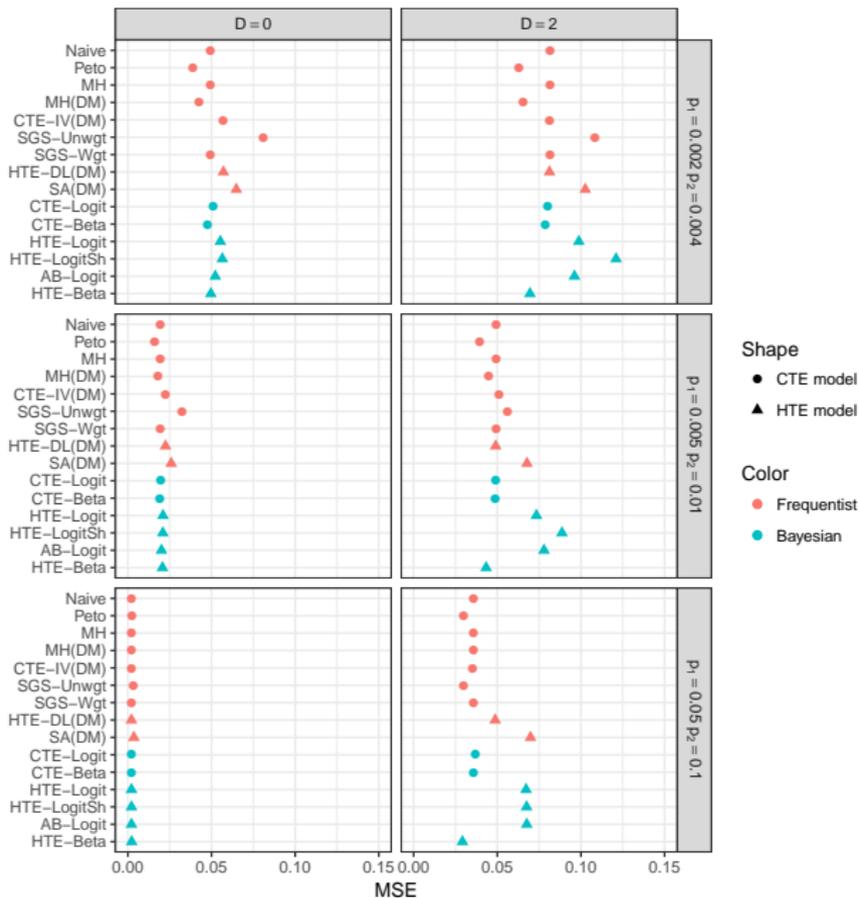
## Simulation Results: Summary for Null Case

- All frequentist estimators (which are moment-based estimators) and three Bayesian approaches modeling trt-specific risks (CTE-Beta, AB-Logit, and HTE-Beta) showed little to no bias
  - Bayesian models with priors on study-specific effects did OK
- CTE-Logit, HTE-Logit, and HTE-LogitSh showed relatively large biases
  - Data-generating mechanism and model structure do not agree
- MSE decreased as risks increased (i.e. more events)
- When between-study variability existed ( $D=2$ ), SGS-Unwgt, HTE-Logit, HTE-LogitSh, and AB-Logit tended to have large MSEs
- $D=2$  case: CTE models (except SGS estimators) had poor coverage. Got worse as risks grew larger

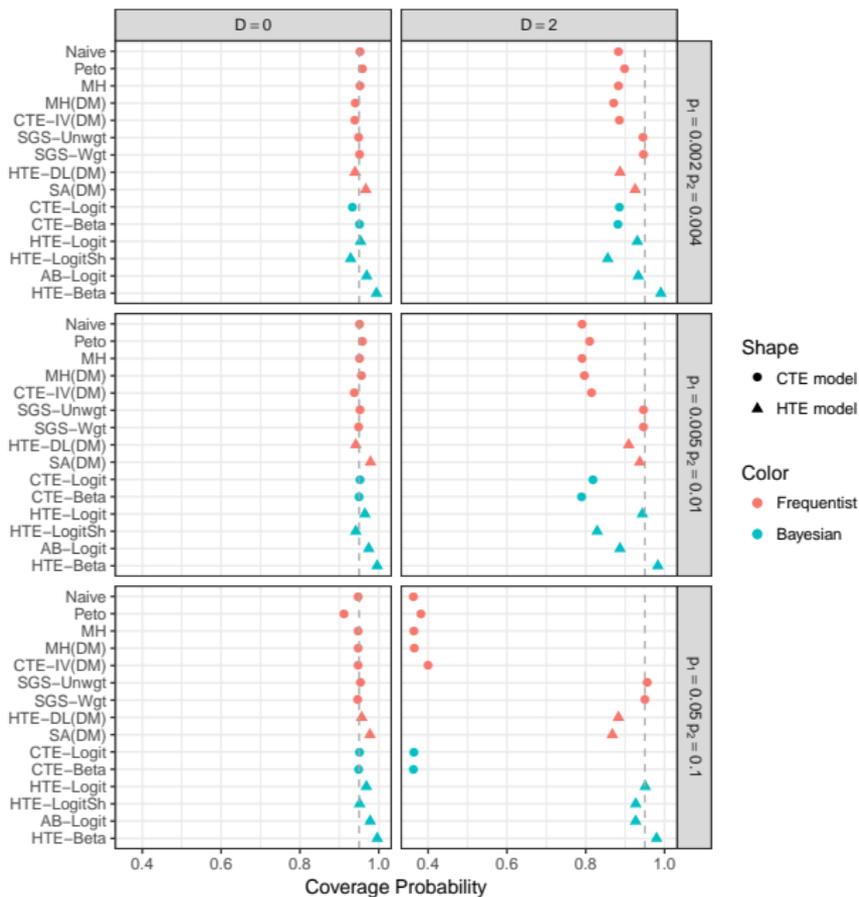
# Results for Log Odds Ratio: Alt Case - Bias



# Results for Log Odds Ratio: Alt Case - MSE



# Results for Log Odds Ratio: Alt Case - Coverage Prob



## Simulation Results: Summary of Non-Null Scenarios

- Naïve, MH, SGS, and HTE-Logit had small biases with  $D=0$  and  $D=2$
- SGS-Unwgt had large MSE in lowest risk scenario ( $p_1 = 0.002, p_2 = 0.004$ )
- When  $D=2$ , SGS, HTE-Logit, and HTE-Beta gave coverage probability close to nominal level (0.95) across all scenarios

# Nissen & Wolski (NW) Data: Description of Studies

- MI:
  - 159 events in 19,509 among pts in rosiglitazone-containing arms
    - Pooled risk=0.00815
  - 136 events among 16,022 pts in arms w/o rosiglitazone
    - Pooled risk=0.00849
  - Naïve pooled estimate of LOR: -0.041 (95% interval: -0.271, 0.189)
- CV death:
  - 105 events among 19,509 pts in rosiglitazone-containing arms
    - Pooled risk=0.00538
  - 100 events among 16,022 pts in arms w/o rosiglitazone
    - Pooled risk=0.00624
  - Naïve pooled estimate of LOR: -0.149 (95% interval: -0.424, 0.126)

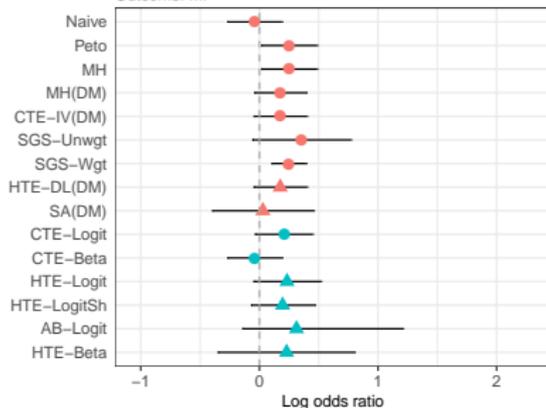
# Summary of Data: Studies with Zeros

	Outcome	No. of studies	0-0 studies	0 in control	0 in trt	Non-zero studies	Corr ( $p_{c_j}, n_{c_j}$ )	Corr ( $p_{a_j}, n_{a_j}$ )	$E(p_{c_j})$	$E(p_{a_j})$
Data1	MI	56	15	35	21	15	0.312	0.181	0.004	0.006
Data2	MI	55	15	35	21	14	0.154	-0.003	0.004	0.006
Data3	MI	42	9	29	13	9	-0.024	-0.156	0.003	0.006
Data4	MI	13	3	8	4	4	0.243	-0.147	0.002	0.005
Data1	CV	56	29	44	33	8	0.318	0.202	0.002	0.003
Data2	CV	55	29	44	33	7	0.042	-0.030	0.002	0.003
Data3	CV	42	20	35	23	4	0.035	-0.063	0.002	0.003
Data4	CV	13	4	10	6	1	0.563	0.104	0.001	0.002

# NW Data Analysis: Log Odds Ratio - MI

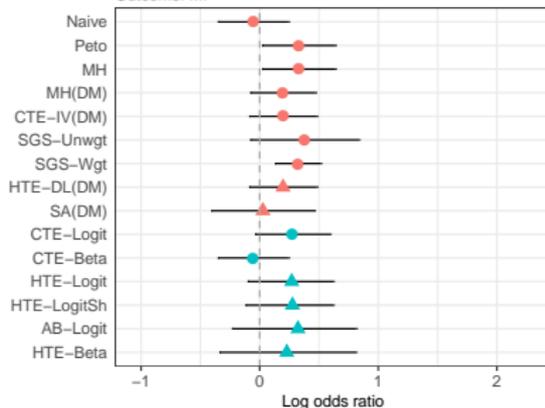
All studies (56 studies)

Outcome: MI



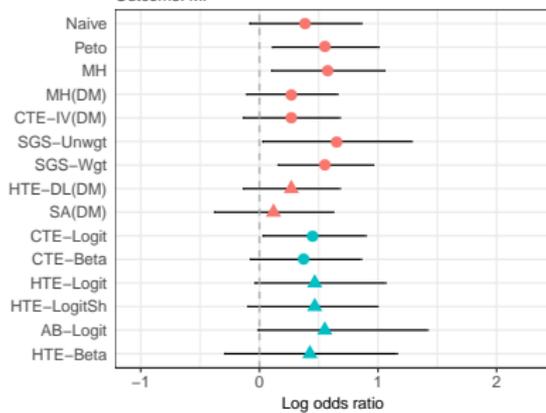
All studies except RECORD trial (55 studies)

Outcome: MI



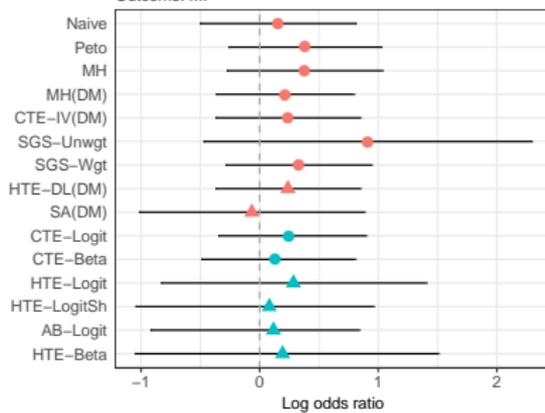
Trials comparing rosiglitazone+X vs. X alone (42 studies)

Outcome: MI



Trials comparing rosiglitazone vs. placebo (13 studies)

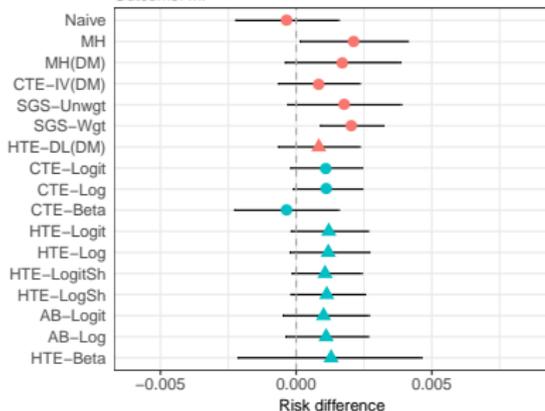
Outcome: MI



# NW Data Analysis: Risk Difference - MI

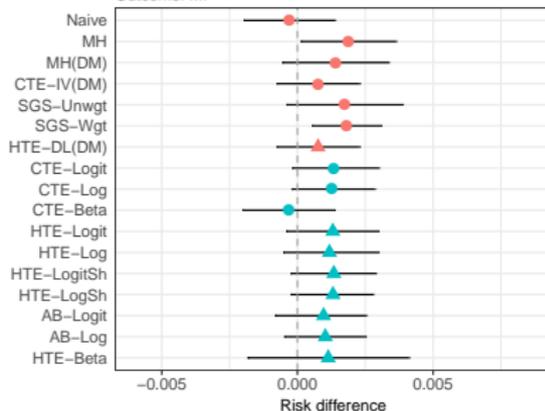
All studies (56 studies)

Outcome: MI



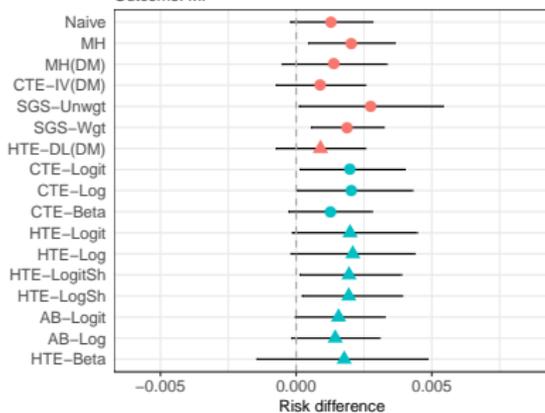
All studies except RECORD trial (55 studies)

Outcome: MI



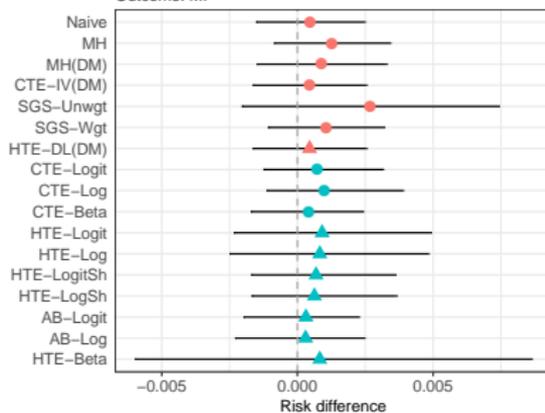
Trials comparing rosiglitazone+X vs. X alone (42 studies)

Outcome: MI



Trials comparing rosiglitazone vs. placebo (13 studies)

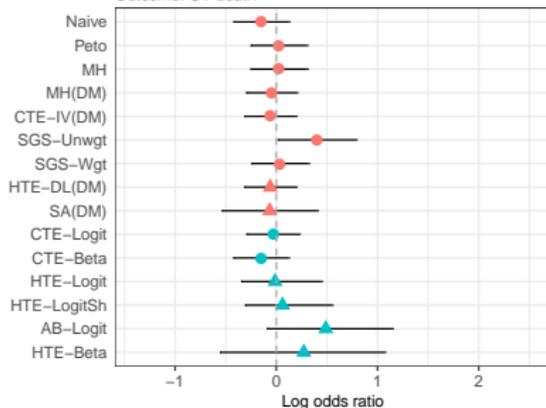
Outcome: MI



# NW Data Analysis: Log Odds Ratio - CV death

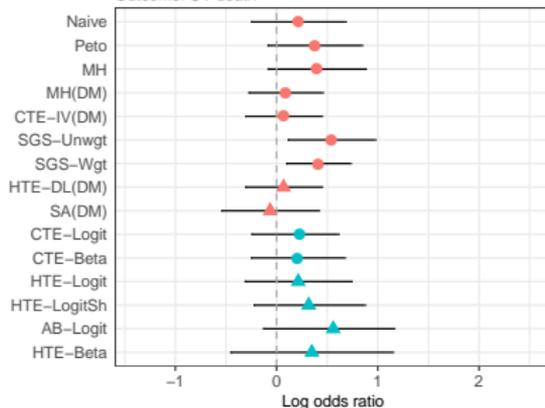
All studies (56 studies)

Outcome: CV death



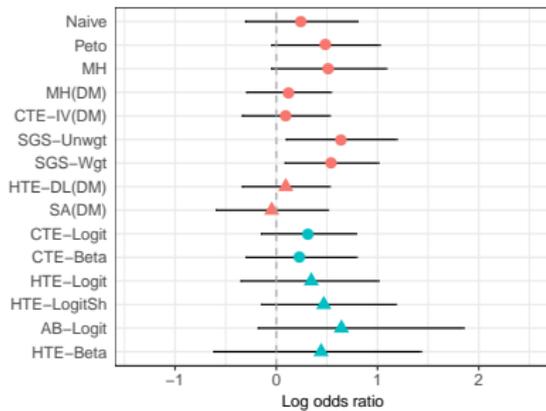
All studies except RECORD trial (55 studies)

Outcome: CV death



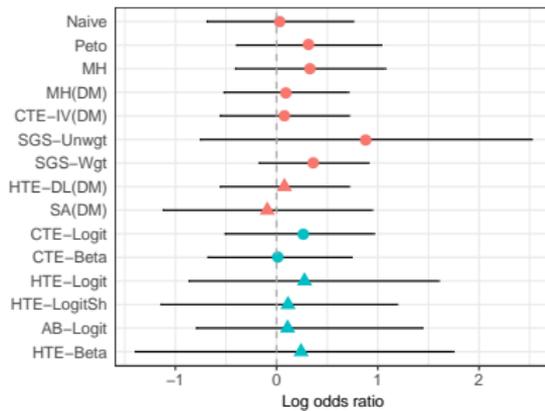
Trials comparing rosiglitazone+X vs. X alone (42 studies)

Outcome: CV death



Trials comparing rosiglitazone vs. placebo (13 studies)

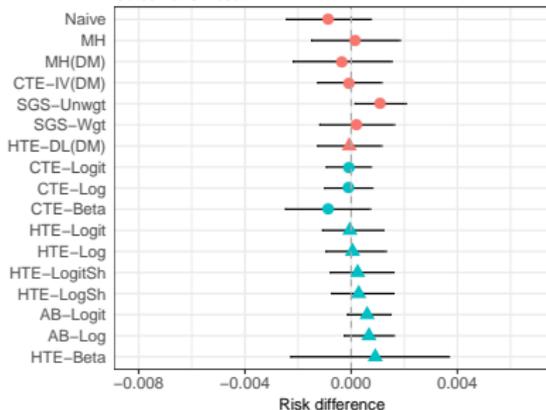
Outcome: CV death



# NW Data Analysis: Risk Difference - CV death

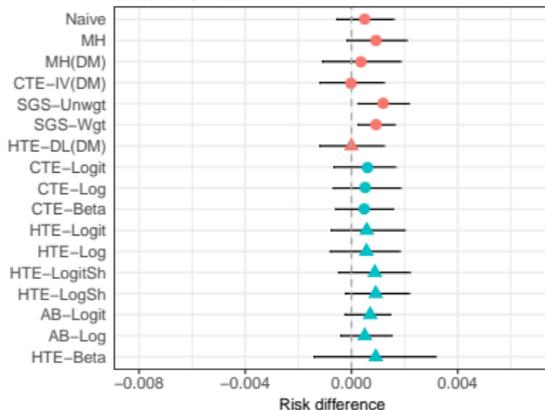
All studies (56 studies)

Outcome: CV death



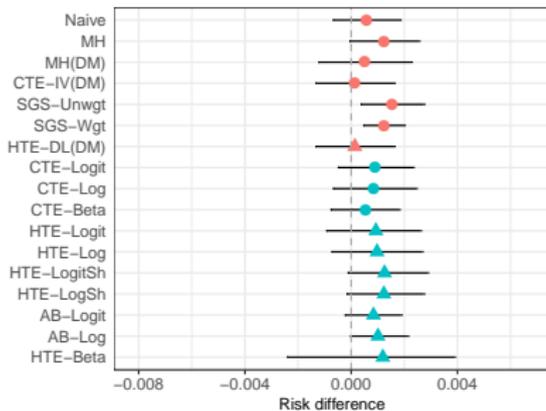
All studies except RECORD trial (55 studies)

Outcome: CV death



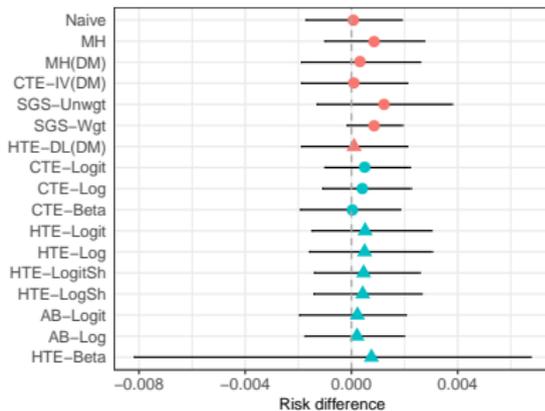
Trials comparing rosiglitazone+X vs. X alone (42 studies)

Outcome: CV death



Trials comparing rosiglitazone vs. placebo (13 studies)

Outcome: CV death



# NW Data Analysis: Summary (MI outcome, LOR)

- Naive and CTE-Beta provided similar results
  - They ignore heterogeneity
- CTE-IV(DM) and HTE-DL(DM) provided exactly the same results because  $\hat{\tau} = 0$  in HTE-DL(DM)
- When all studies were included, posterior means of  $\tau$  in HTE-Logit and HTE-LogitSh are 0.12 and 0.14.
- SGS-Wgt estimates always provided much narrower 95% confidence intervals than SGS-Unwgt
  - This trend was not observed in our simulations. It may be due to correlation between study design (sample size) and risk induced in the NW data.
- SGS-Unwgt very different from SGS-Wgt when only look at placebo-controlled RCTs
- SGS-Unwgt and SA provided very different estimates.

# Conclusion

- The best way to estimate treatment effects correctly with rare events is to conduct large RCTs (and combine information from them).
- Simulation results suggest that one interpret associations carefully in rare (very low frequency) event settings.
- When the outcome is really uncommon, the normality random effect assumption might not hold.
- Effect estimates vary when applying different models; inference can differ, too.
- Future work, we will relax the normality assumption. In addition, we will consider nonparametric approaches.

# References:

- Bhaumik DK, Amatya A, Normand S-L, et al. Meta-analysis of rare binary adverse event data. *J Am Statist Assoc* 2012;107, 555-567.
- Bohning D, Mylona K, Kimber A. Meta-analysis of clinical trials with rare events. *Biom J* 2015;57(4), pp.633-648.
- Bradburn MJ, Deeks JJ, Berlin JA, Localio AR. Much ado about nothing: A comparison of the performance of meta-analytical methods with rare events. *Stat Med*, 2007;26(1), pp.53-77.
- Cai T, Parast L, Ryan L. Meta-analysis for rare events. *Stat Med*, 2010;29(20), pp.2078-2089.
- Chakravarty AG, Levenson M. Regulatory Issues in Meta-Analysis of Safety Data. *Quantitative Evaluation of Safety in Drug Development: Design, Analysis and Reporting* 67 2014;67: p.237.
- Diamond GA, Bax L, Kaul S. Uncertain effects of rosiglitazone on the risk for myocardial infarction and cardiovascular death. *Ann Intern Med*, 2007;147(8), pp.578-581.
- Friedrich JO, Adhikari NK, Beyene J. Inclusion of zero total event trials in meta-analyses maintains analytic consistency and incorporates all available data. *BMC Med Res Methodol* 2007;7(1):1.
- Friedrich JO, Beyene J, Adhikari NK. Rosiglitazone: Can meta-analysis accurately estimate excess cardiovascular risk given the available data? Re-analysis of randomized trials using various methodologic approaches. *BMC Res Notes*, 2009;2(1), p.5.
- Hong H, Chu H, Zhang J, Carlin BP. A Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons. *Res Synth Methods*, 2016;7(1):6-22.
- Lane PW. Meta-analysis of incidence of rare events. *Stat Methods Med Res*, 2013;22(2), pp.117-132.
- Liao JG. A hierarchical Bayesian model for combining multiple  $2 \times 2$  tables using conditional likelihoods. *Ann Intern Med*, 1999;55(1), pp.268-272.
- Marchenko O, Jiang Q, Chakravarty A, et al. Evaluation and Review of Strategies to Assess Cardiovascular Risk in Clinical Trials in Patients with Type 2 Diabetes Mellitus. *Stat Biopharm Res*, 2015;7(4), pp.253-266.
- Nissen SE, Wolski K. Rosiglitazone revisited: An updated meta-analysis of risk for myocardial infarction and cardiovascular mortality. *Arch Intern Med* 2010;170(14), pp.1191-1201.
- Rücker G, Schwarzer G, Carpenter J, Olkin I. Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Stat Med* 2009;28(5):721-38.
- Shuster JJ, Jones LS, Salmon DA. Fixed vs random effects meta-analysis in rare event studies: the rosiglitazone link with myocardial infarction and cardiac death. *Stat Med* 2007;26: 4375-4385.
- Shuster JJ, Walker MA. Low event rate meta-analyses of clinical trials: Implementing good practices. *Stat Med* 2016;35: 2467-2478.
- Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004;23(9):1351-75.
- Tian L, Cai T, Pfeffer MA, et al. Exact and efficient inference procedure for meta-analysis and its application to the analysis of independent  $2 \times 2$  tables with all available data but without artificial continuity correction. *Biostatistics* 2009;10(2), pp.275-281.