An alternative approach to outliers in meta-analysis

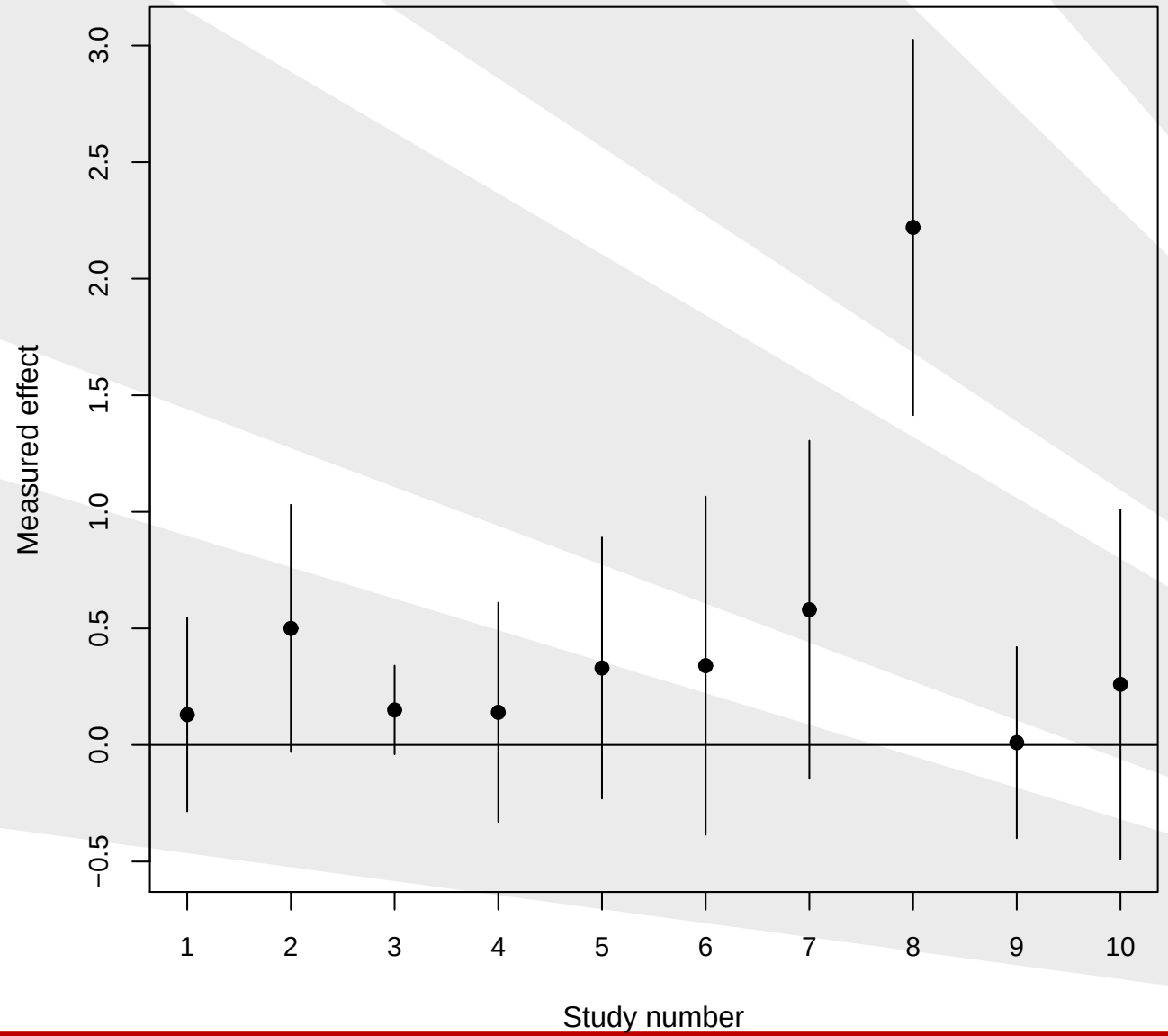
Freedom Gumedze
Department of Statistical Sciences
University of Cape Town

European GenStat and ASReml Applied Statistics Conference 2010
Rothamsted Research, 14 July 2010

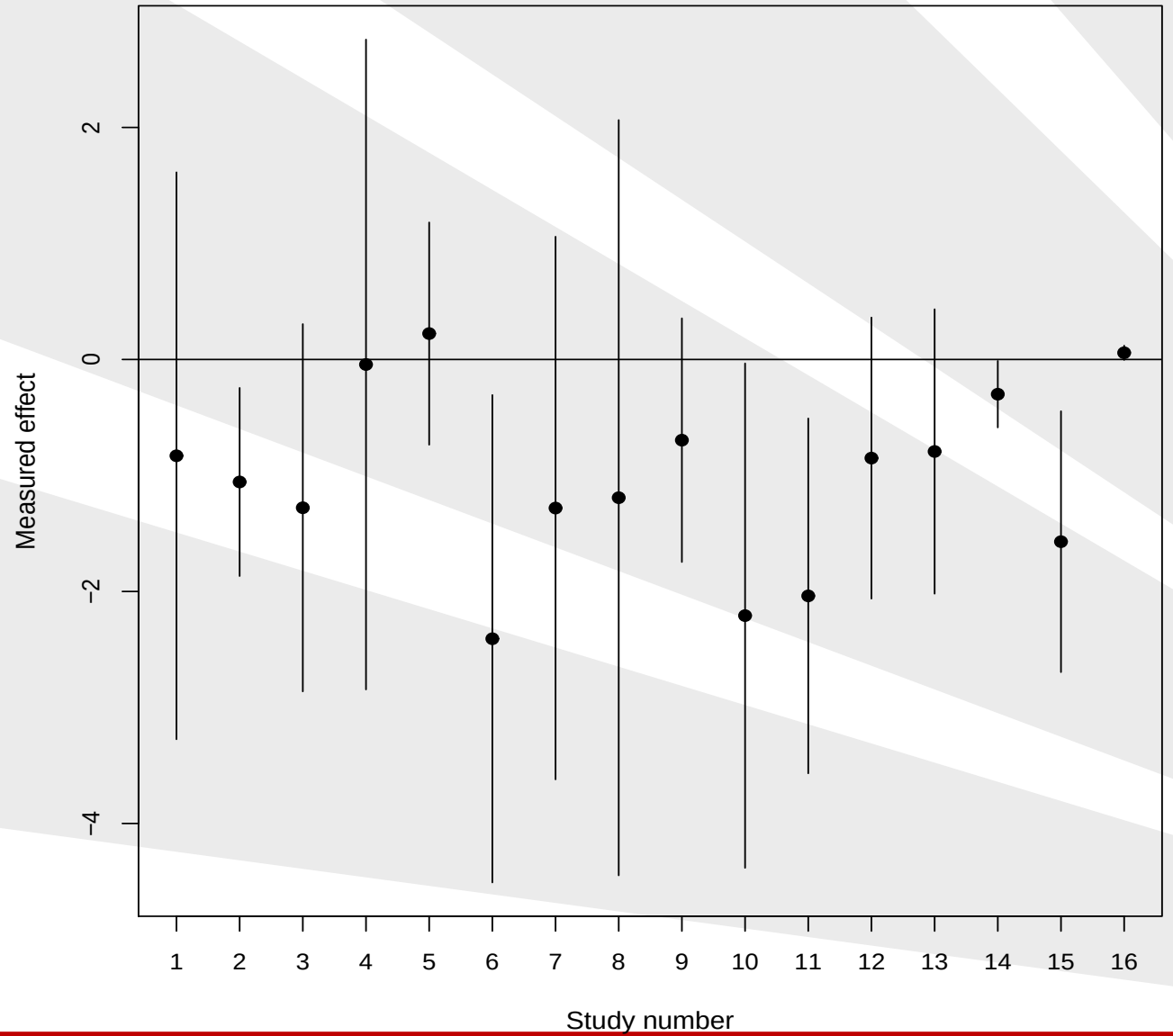
Outline

- Motivating examples
- Related work
- Random effect variance shift outlier model (RVSOM) in meta-analysis
- Variance parameter estimation and inference under the RVSOM
- Illustration of RVSOM
- Conclusions and remarks

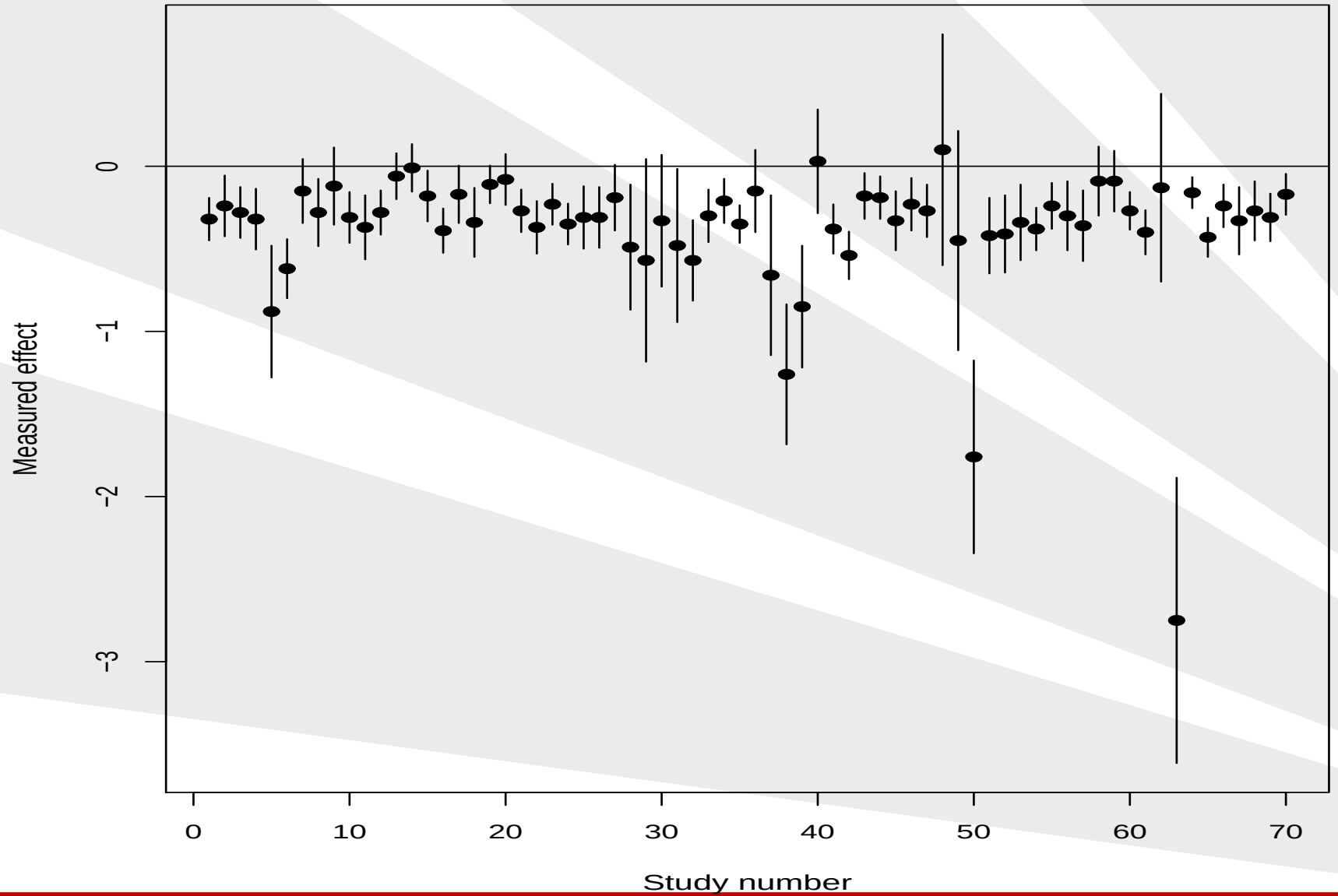
The CDP-choline data



Intravenous magnesium in acute myocardial infarction data



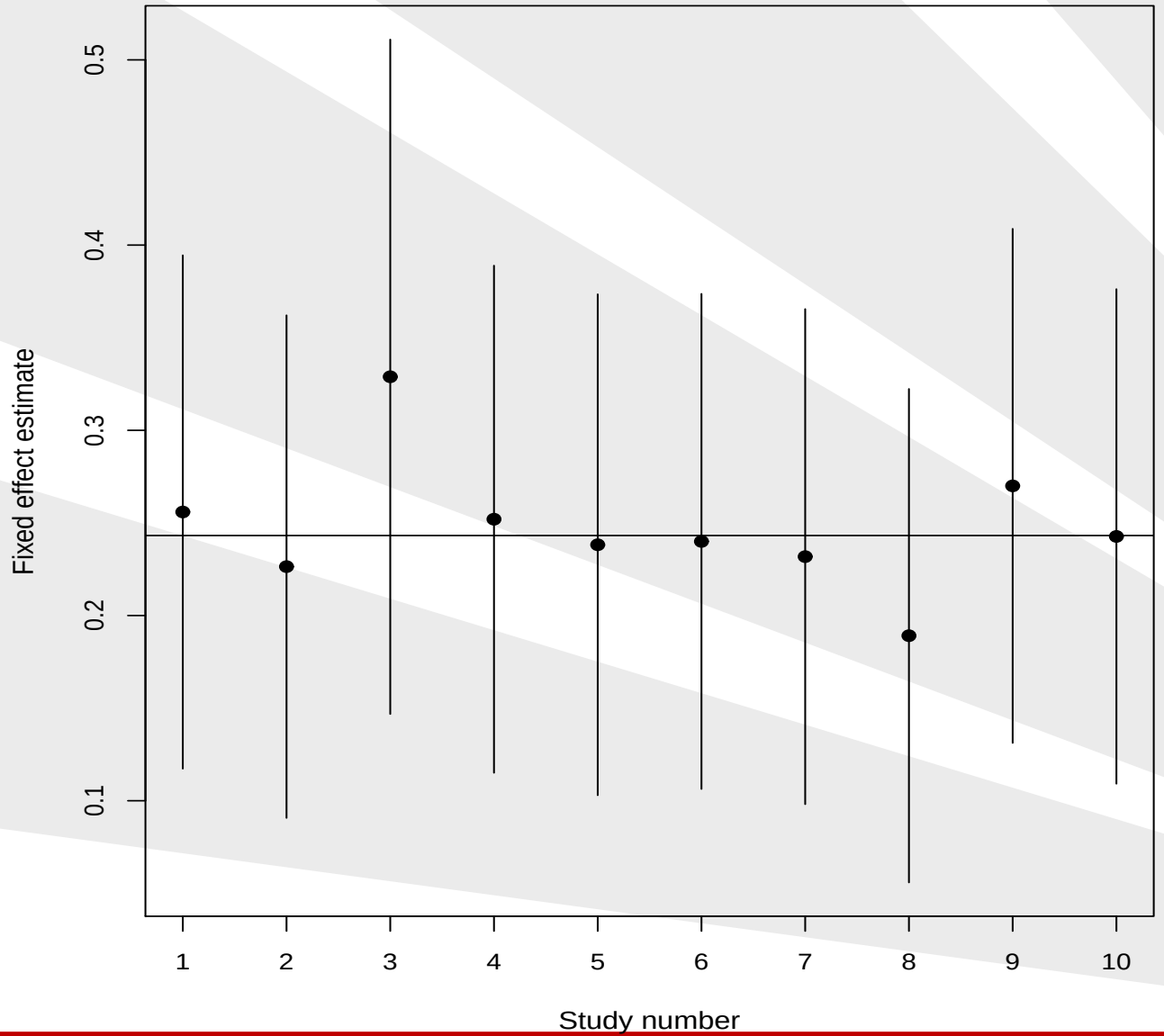
Fluoride toothpaste for preventing dental caries data



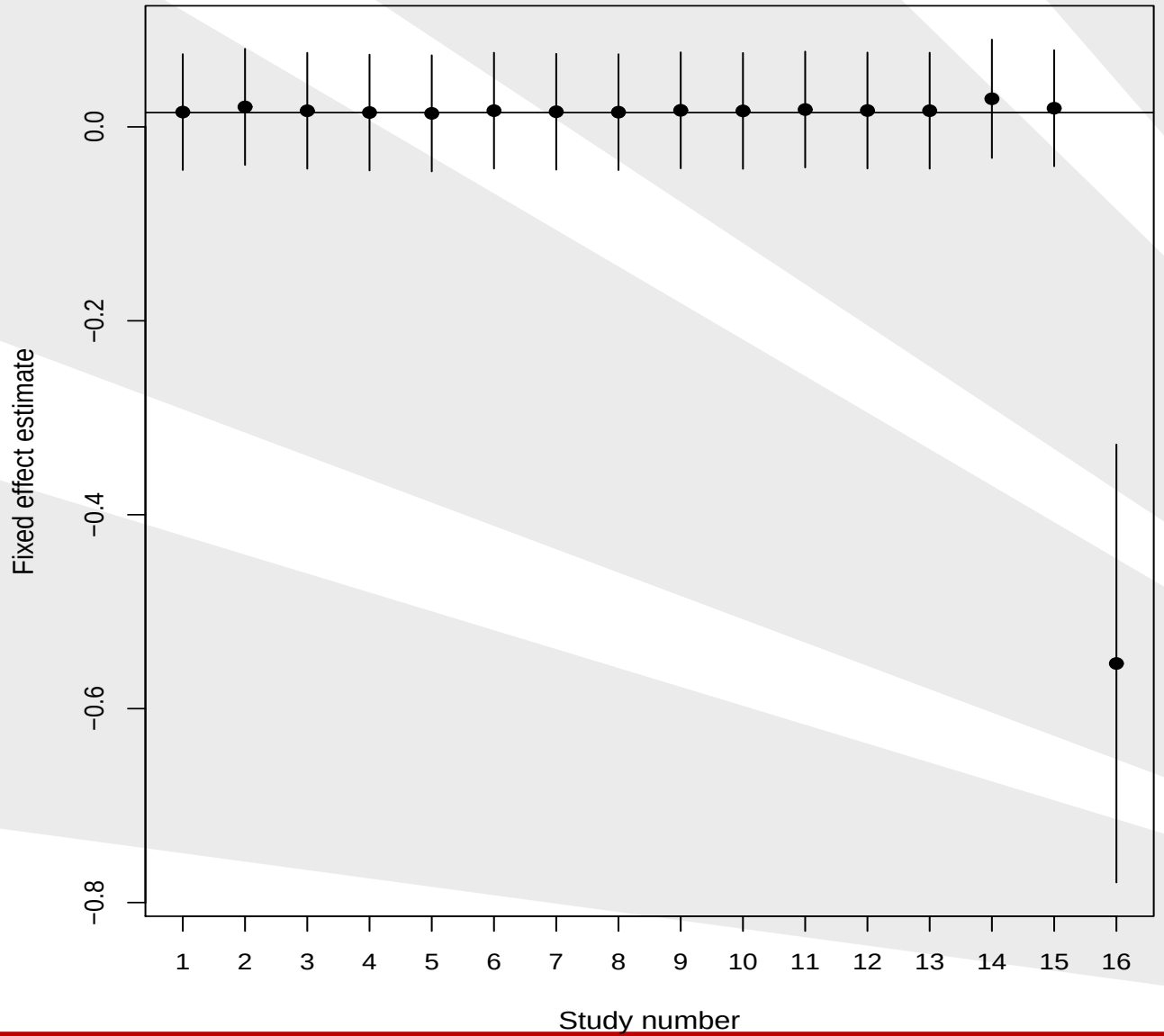
Related work

- Tobias (Stata Tech. Bull., 1999)
- Lee and Thompson (Stat. Med, 2008)
- Baker and Jackson (Health Care Manag. Science, 2008).

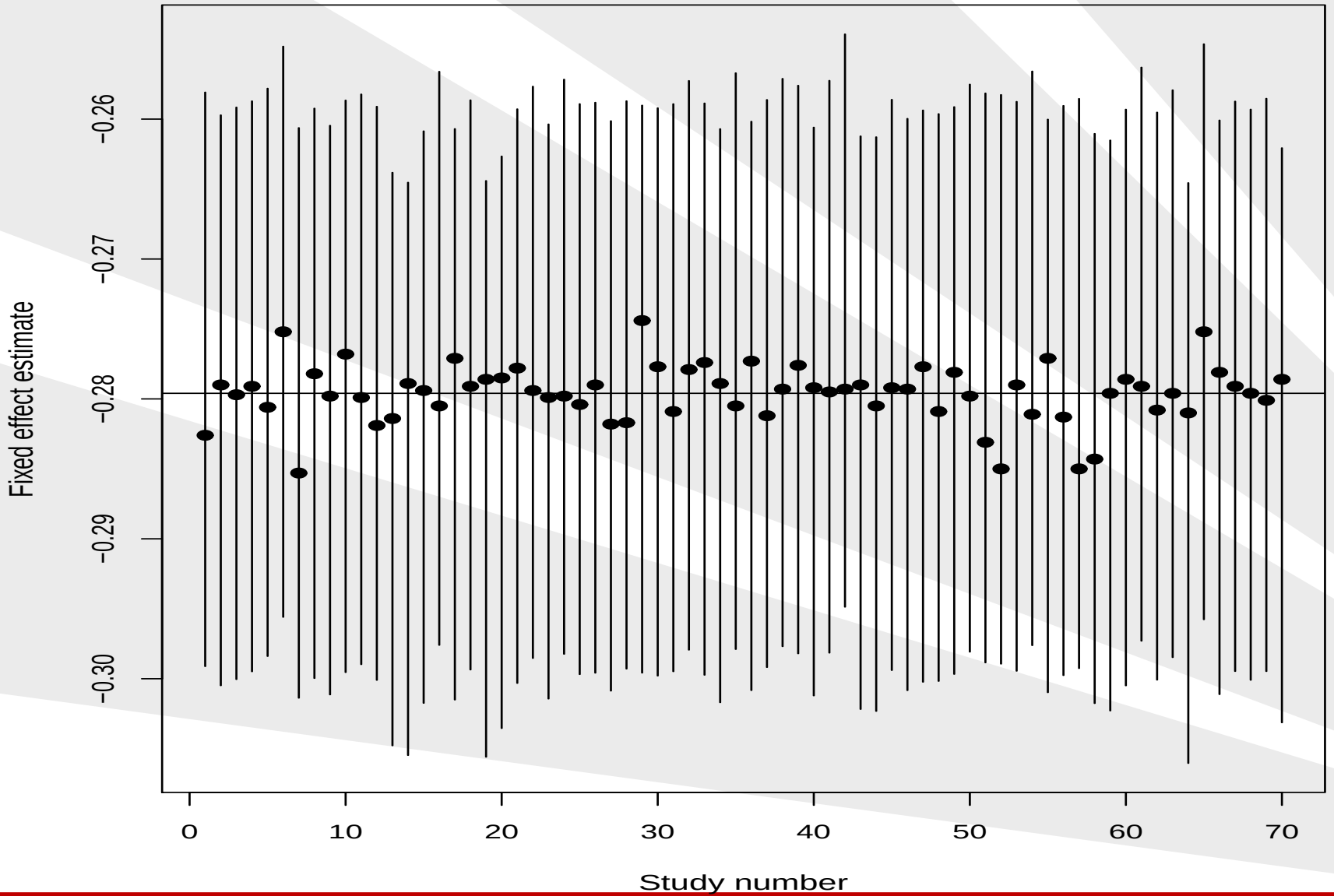
The CDP-choline data



Intravenous magnesium in acute myocardial infarction data



Fluoride toothpaste for preventing dental caries data



A random effect variance shift outlier model

- Random effects model for meta-analytic data is

$$\begin{aligned} \boldsymbol{y} &= \mu \mathbf{1}_n + \boldsymbol{u} + \boldsymbol{e} \,, \\ &\sim N\left(\mu \mathbf{1}_n, \sigma_s^2 \boldsymbol{I}_n + \boldsymbol{R}\right), \end{aligned} \tag{1}$$

  where $\boldsymbol{R} = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$.

- A RVSOM for the ith observation in the random effects model for meta-analytic data takes the form

$$\begin{aligned} \boldsymbol{y} &= \mu \mathbf{1}_n + \delta_i \boldsymbol{d}_i + \boldsymbol{u} + \boldsymbol{e} \\ &\sim N\left(\mu \mathbf{1}_n, \omega_i^2 \boldsymbol{d}_i \boldsymbol{d}_i' + \sigma_s^2 \boldsymbol{I}_n + \boldsymbol{R}\right), \end{aligned} \tag{2}$$

  where: $\boldsymbol{d}_i$ is a unit vector of length

$$\delta_i \sim N\left(0, \omega_i^2\right).$$

Variance parameter estimation and hypothesis testing in a RVSOM

- Set of parameters to be estimated: $\mu, \omega_i^2, \sigma_s^2$.

  $\Rightarrow$ Variance parameters ω_i^2, σ_s^2 are estimated iteratively using REML.

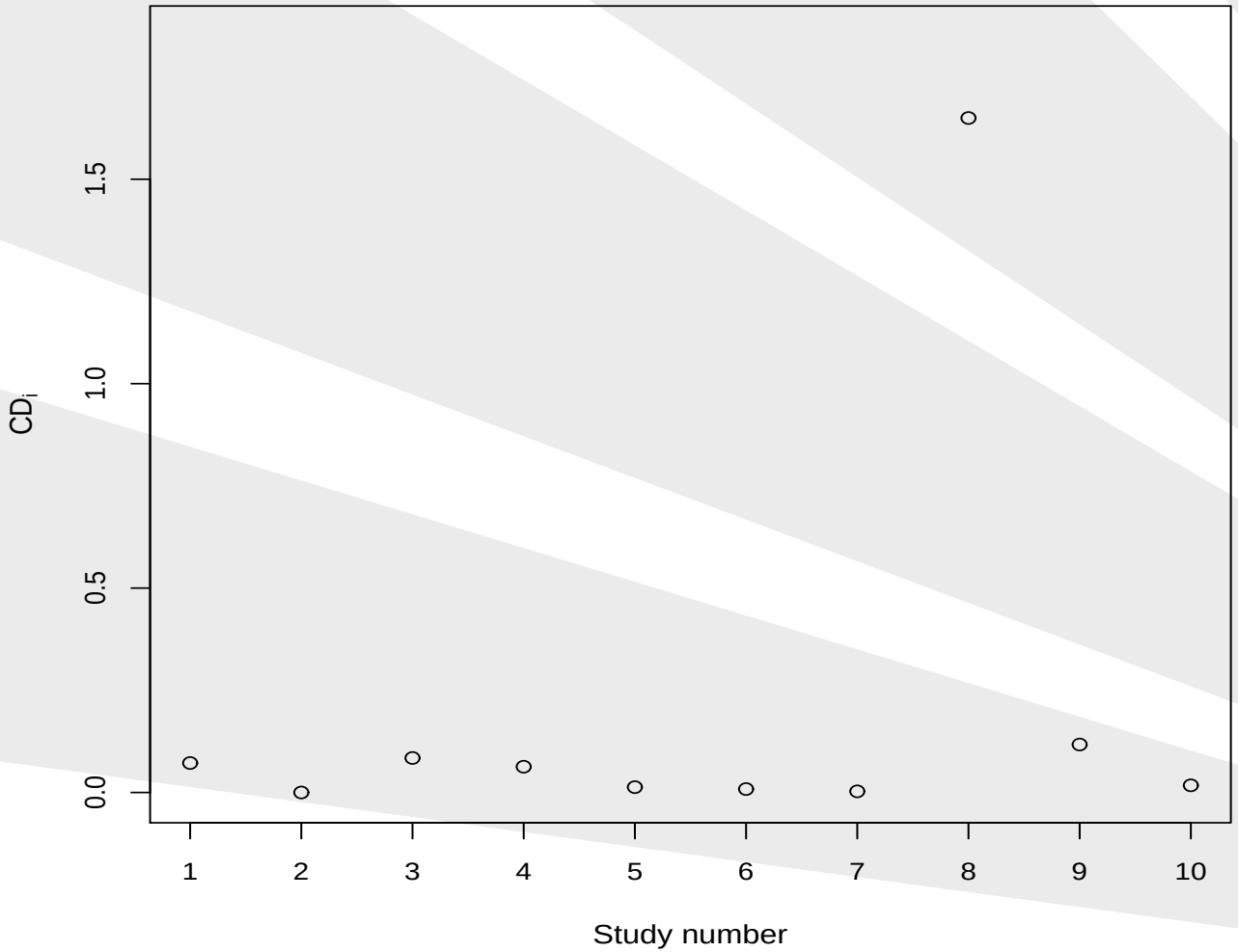
- Likelihood ratio test statistic (LRT):

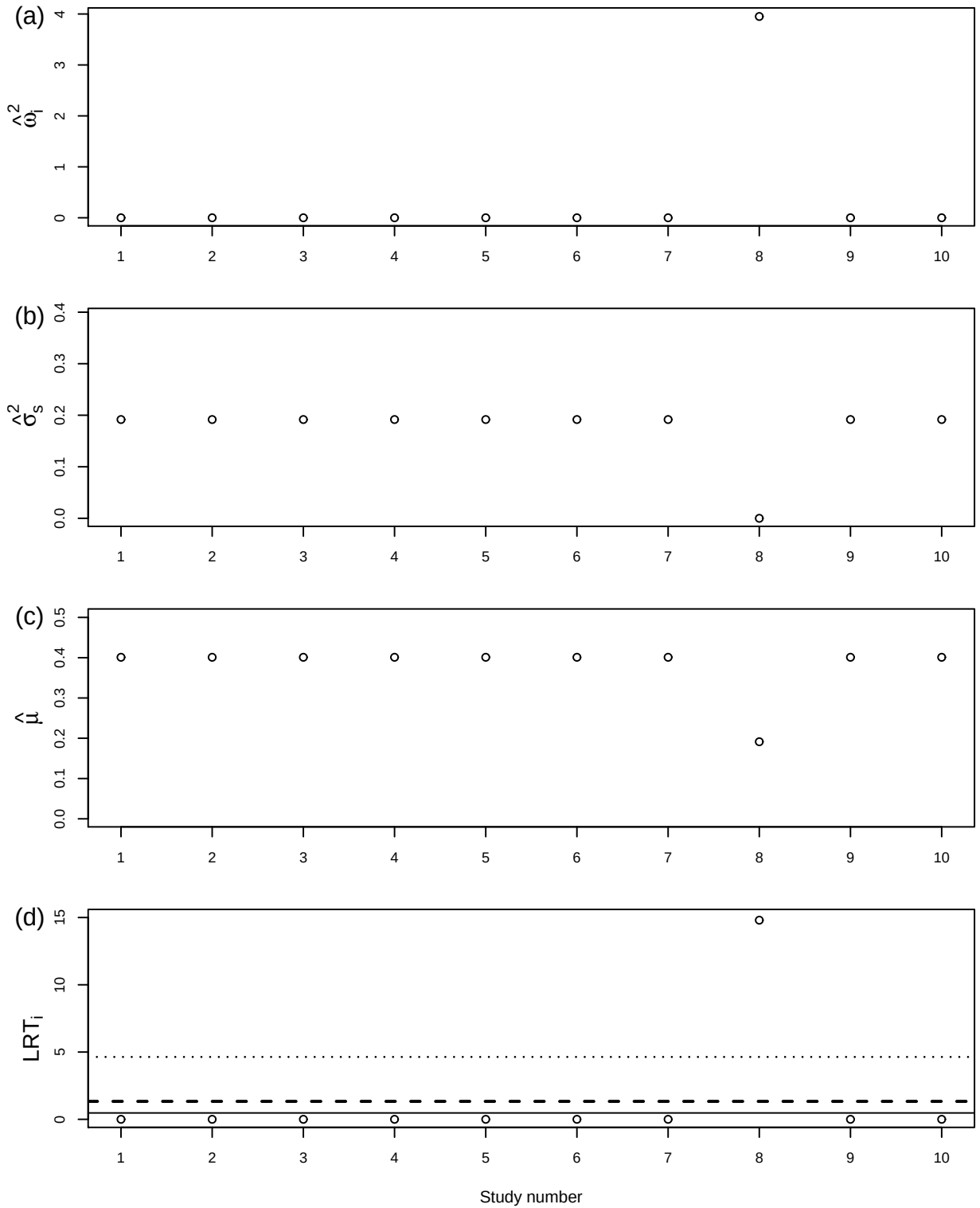
$$LRT_i = \begin{cases} -2\left\{RL(\hat{\sigma}_s^2; \boldsymbol{y}) - RL_{(i)}(\hat{\omega}_i^2, \hat{\sigma}_s^2; \boldsymbol{y})\right\} & \hat{\omega}_i^2 > 0 \\ 0 & \text{otherwise.} \end{cases}$$

- Sampling distribution of LRT via parametric bootstrap

  $\Rightarrow$ parametric bootstrap also deals with problem of multiple testing.

The CDP-choline data

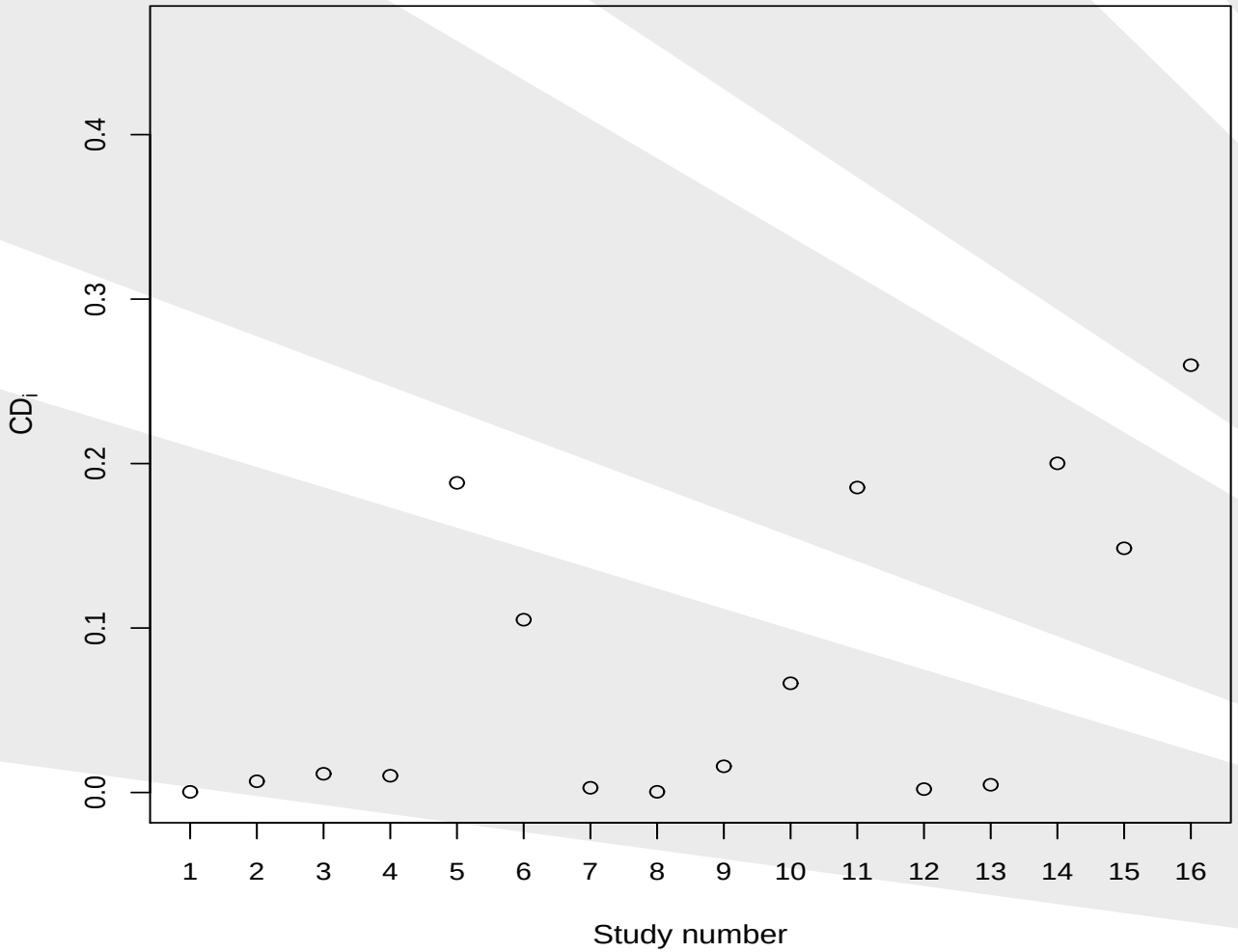


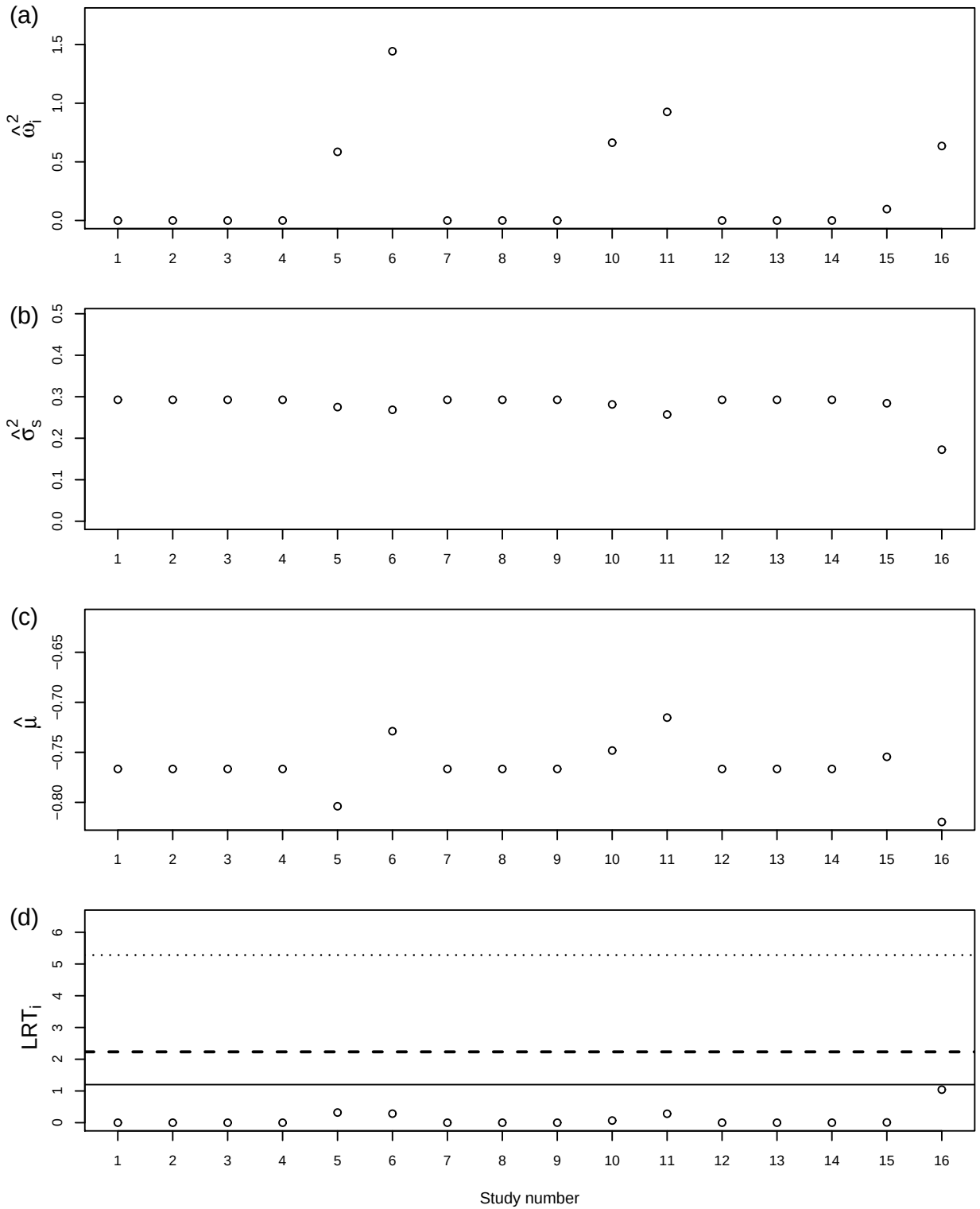

(a)
$\hat{\omega}_i^2$
0
1
2
3
4
1
2
3
4
5
6
7
8
9
10
(b)
$\hat{\sigma}_s^2$
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
6
7
8
9
10
(c)
$\hat{\mu}$
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
6
7
8
9
10
(d)
LRT_i
0
5
10
15
1
2
3
4
5
6
7
8
9
10
Study number

The CDP-choline data

	M_0		M_1		$M_4^\dagger$	
Para.	Est.	95% CI	Est.	95% CI	Est.	95% CI
μ	0.401	(0.08;0.72)	0.191	(0.058;0.324)	0.189	(0.056;0.32
σ_s^2	0.192	-	6.4×10^{-8}	-	6.4×10^{-8}	-
ω_8^2	-	-	3.951	-	-	-

Intravenous magnesium in acute myocardial infarction data

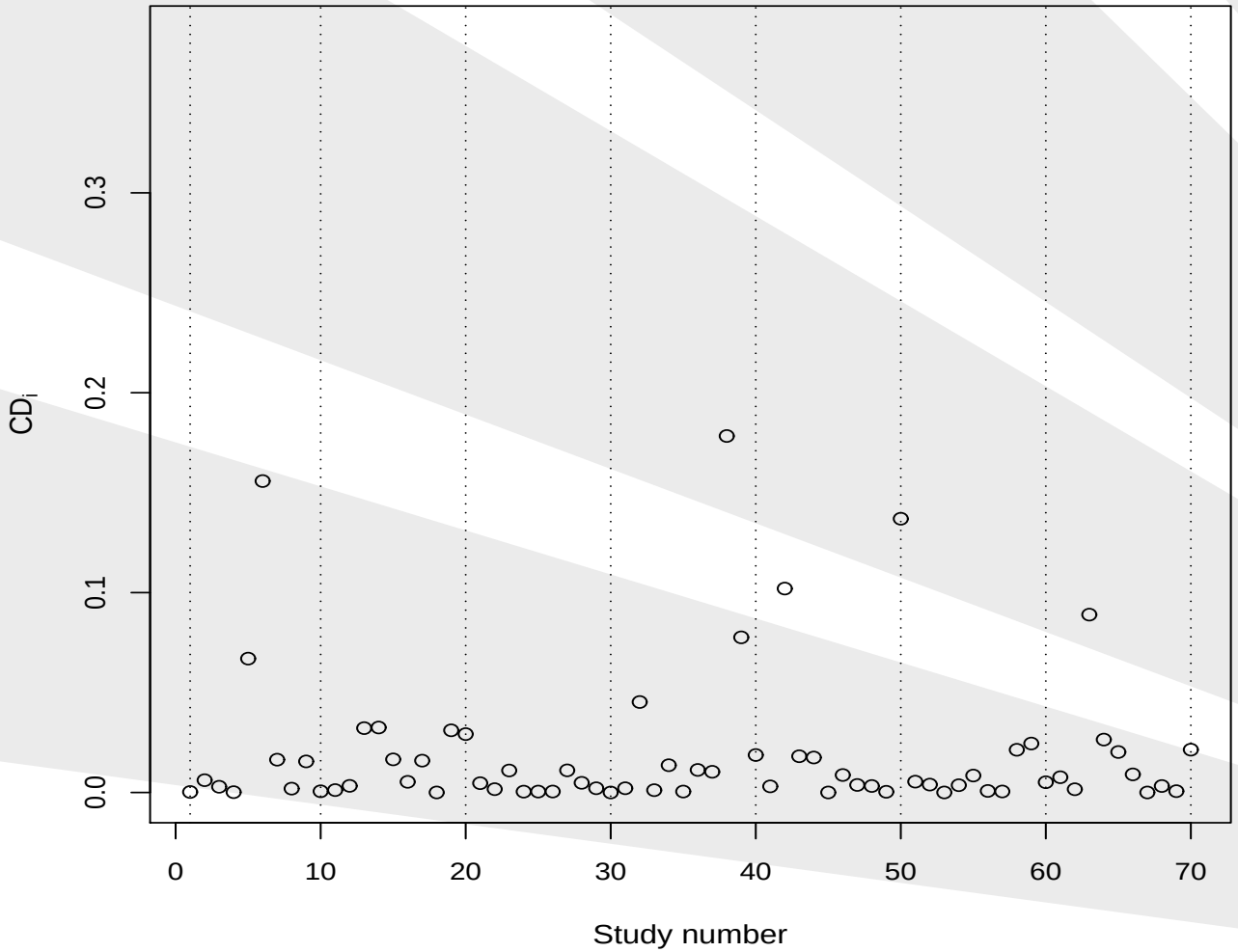


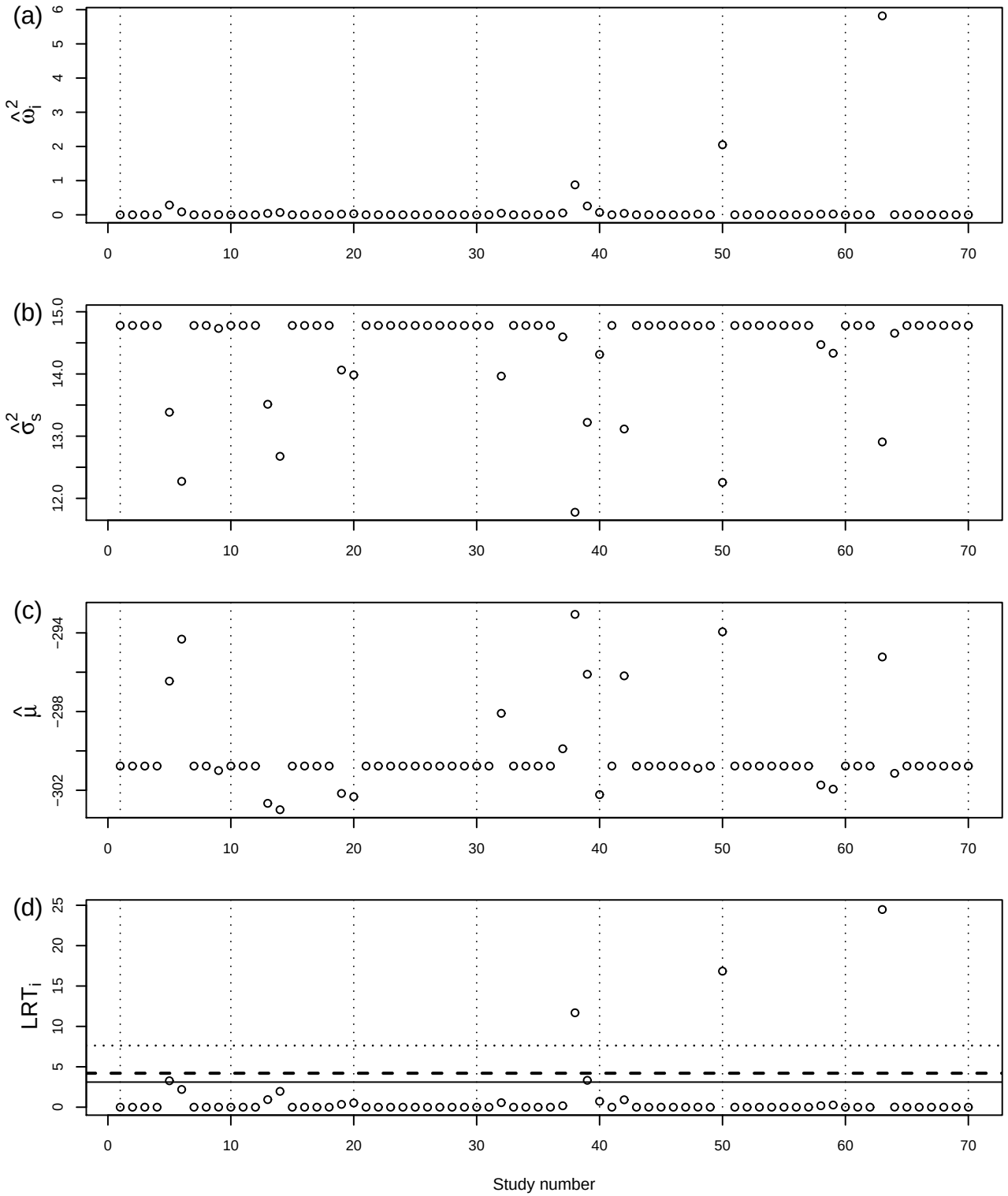

(a)
$\hat{\omega}_i^2$
0.0
0.5
1.0
1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
(b)
$\hat{\sigma}_s^2$
0.0
0.1
0.2
0.3
0.4
0.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
(c)
$\hat{\mu}$
−0.80
−0.75
−0.70
−0.65
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
(d)
LRT_i
0
1
2
3
4
5
6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Study number

Intravenous magnesium in acute myocardial infarction data

	M_0		M_1		$M_4^\dagger$	
Para.	Est.	95% CI	Est.	95% CI	Est.	95% CI
μ	-0.766	(-1.18;-0.35)	-0.820	(-1.21;-0.43)	-0.875	(-1.28;-0.47)
σ_s^2	0.293	-	0.172	-	0.191	-
ω_{16}^2	-	-	0.695	-	-	-

Fluoride toothpaste for preventing dental caries data




(a)
$\hat{\omega}_i^2$
0
1
2
3
4
5
6
0
10
20
30
40
50
60
70
(b)
$\hat{\sigma}_s^2$
12.0
13.0
14.0
15.0
0
10
20
30
40
50
60
70
(c)
$\hat{\mu}$
−294
−298
−302
0
10
20
30
40
50
60
70
(d)
LRT_i
0
5
10
15
20
25
0
10
20
30
40
50
60
70
Study number

Fluoride toothpaste for preventing dental caries data

	M_0		M_1		$M_4^\dagger$	
Para.	Est.	95% CI	Est.	95% CI	Est.	95% CI
μ	-0.3008	(-0.33;-0.27)	-0.284	(-0.32;-0.25)	-0.283	(-0.31;-0.25)
σ_s^2	0.015	-	0.009	-	0.009	-
ω_{38}^2	-	-	0.897	-	-	-
ω_{50}^2	-	-	2.082	-	-	-
ω_{63}^2	-	-	5.879	-	-	-

Conclusion and remarks

- A RVSOM downweights outliers but does not eliminate them from the analysis.
- The LRT gives an objective measure for detecting outliers in meta-analytic data.
- Variance shift outlier model (VSOM) under fixed effects model in meta-analysis.
- RVSOM may not be appropriate when there are *many* outliers are detected.
  - Use heavy-tailed distributions for the random effect (Baker and Jackson, 2008).
- Computation of thresholds for LRTs.