

Difference in mean -0.07(-0.07 to -0.07)

RCT Gold Standard (Caironi et al. 2014) -0.00(-0.05 to 0.05)

### Propensity Score Matching

Agg=['first'], Est=Regularized LR -0.02(-0.04 to -0.01)

Agg=['first', 'last'], Est=Regularized LR 0.01(-0.00 to 0.04)

Agg=['first'], Est=Forests -0.04(-0.05 to -0.01)

Agg=['first', 'last'], Est=Forests 0.67( 0.67 to 1.48)

Outlier ►

### Inverse Propensity Weighting

Agg=['first'], Est=Regularized LR -0.03(-0.06 to 0.01)

Agg=['first', 'last'], Est=Regularized LR -0.03(-0.06 to 0.00)

Agg=['first'], Est=Forests -0.03(-0.05 to -0.02)

Agg=['first', 'last'], Est=Forests -0.03(-0.04 to -0.02)

### Outcome model (Tlearner)

Agg=['first'], Est=Regularized LR -0.05(-0.12 to 0.01)

Agg=['first', 'last'], Est=Regularized LR -0.05(-0.12 to 0.02)

Agg=['first'], Est=Forests -0.01(-0.21 to 0.18)

Agg=['first', 'last'], Est=Forests -0.01(-0.21 to 0.18)

### Double Machine Learning

Agg=['first'], Est=Regularized LR -0.07(-0.08 to -0.05)

Agg=['first', 'last'], Est=Regularized LR -0.06(-0.07 to -0.04)

Agg=['first'], Est=Forests -0.02(-0.03 to -0.01)

Agg=['first', 'last'], Est=Forests -0.01(-0.02 to -0.00)

### Doubly Robust (AIPW)

Agg=['first'], Est=Regularized LR -0.08(-0.14 to -0.01)

Agg=['first', 'last'], Est=Regularized LR -0.08(-0.15 to -0.02)

Agg=['first'], Est=Forests -0.01(-0.02 to 0.00)

Agg=['first', 'last'], Est=Forests -0.00(-0.01 to 0.01)

-0.15 -0.10 -0.05 0.00 0.05 0.10

ATE on 28-day mortality