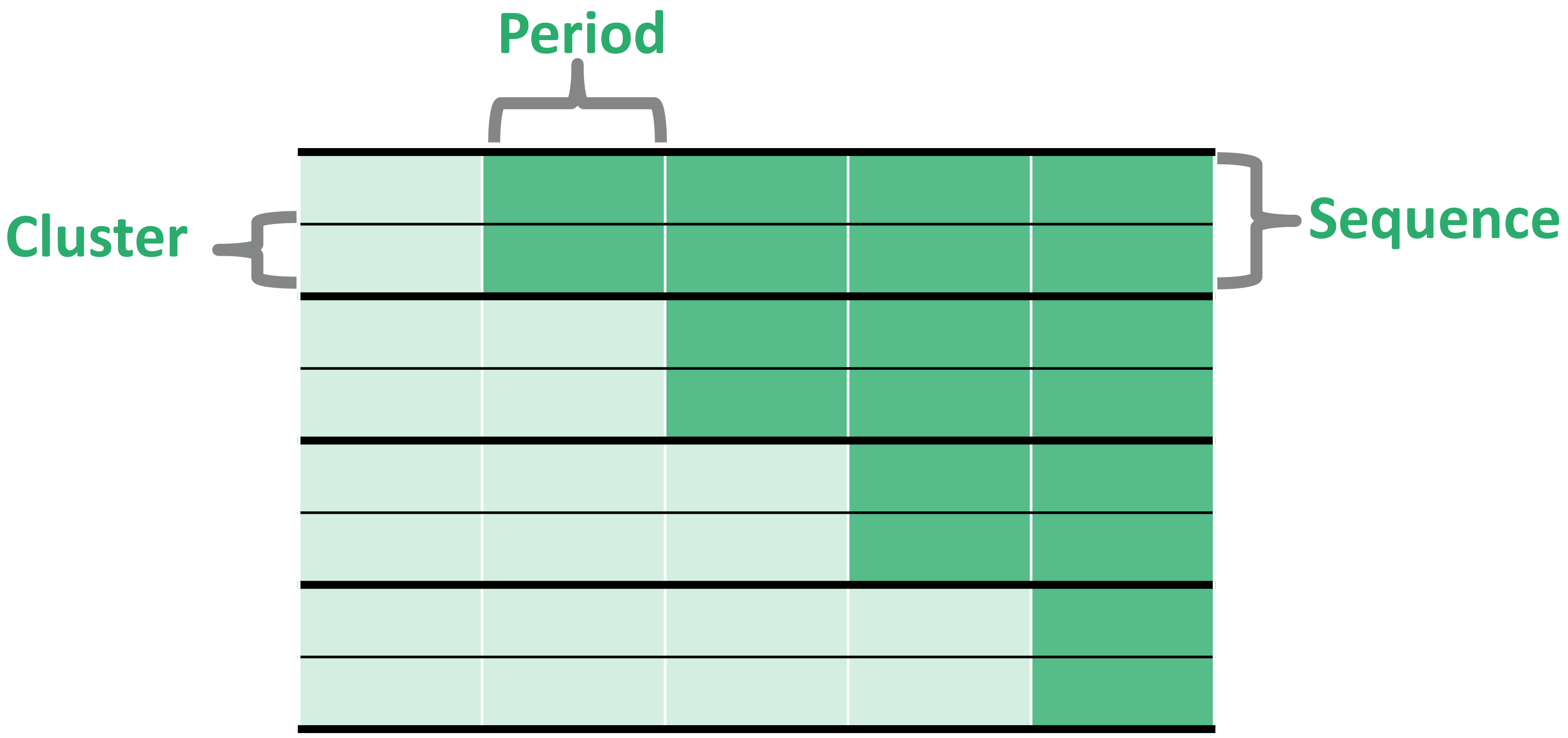
Advice for using Generalised Estimating Equations in a Stepped Wedge Trial

Jennifer Thompson

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE

Background



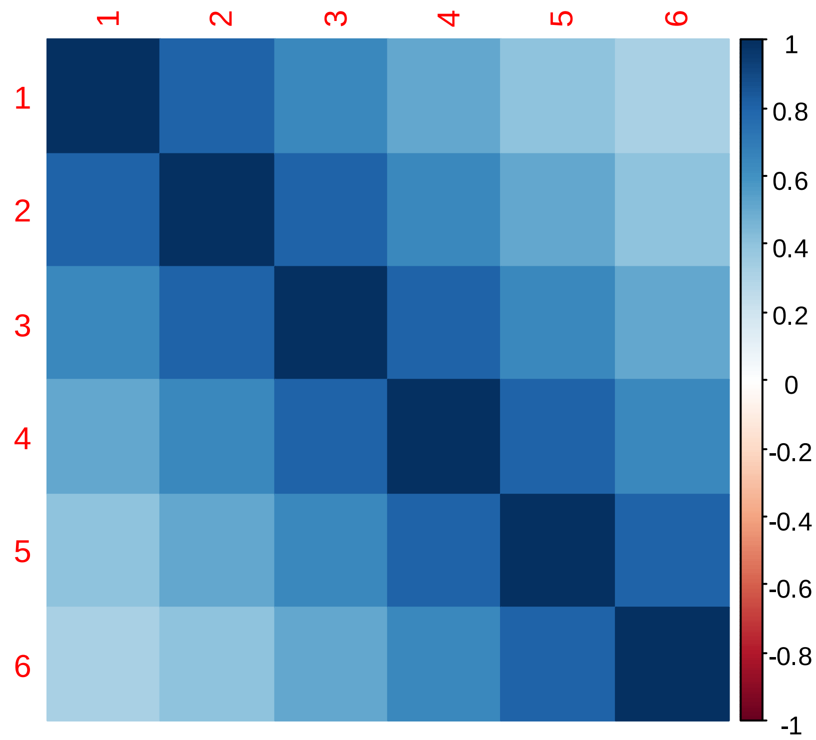
Background

The analysis of a stepped wedge trial must account for:

- Clustering
- Secular trends
- Correlations over time

Simulation study: Methods

Autoregressive correlation



Background

Commonly used methods:

- Generalised linear mixed models
- Generalised estimating equations with robust standard errors

Most research to date has focused on generalized linear mixed models

Growing evidence that these require correct specification of correlation structure to give unbiased standard errors [1,2,3]

1. Thompson, J.A., et al., *Bias and inference from misspecified mixed-effect models in stepped wedge trial analysis.* Statistics in Medicine, 2017. **36**(23): p. 3670-3682.

2. Bellan, S.E., et al., *Statistical power and validity of Ebola vaccine trials in Sierra Leone: a simulation study of trial design and analysis.* The Lancet Infectious Disease, 2015. **15**(6): p. 703-10.

3. Ji, X., et al., *Randomization inference for stepped-wedge cluster-randomised trials: An application to community-based health insurance.* Annals of Applied Statistics, 2017. **11**(1): p. 1-20.

Another way?

Generalised estimating equations with robust standard errors are known to be robust to misspecified correlation structures[1]

- Not known the degree to which this will hold for stepped wedge trial.

Require a large number of clusters

- Standard errors too small with a small number of clusters[2]
- Small sample corrections available[2]
- How well do corrections perform for stepped wedge trials?

1. Liang, K.Y. and S.L. Zeger, *Longitudinal Data-Analysis Using Generalized Linear-Models.* Biometrika, 1986. **73**(1): p. 13-22.
2. Fay, M.P. and B.I. Graubard, *Small-sample adjustments for Wald-type tests using sandwich estimators.* Biometrics, 2001. **57**(4): p. 1198-206.

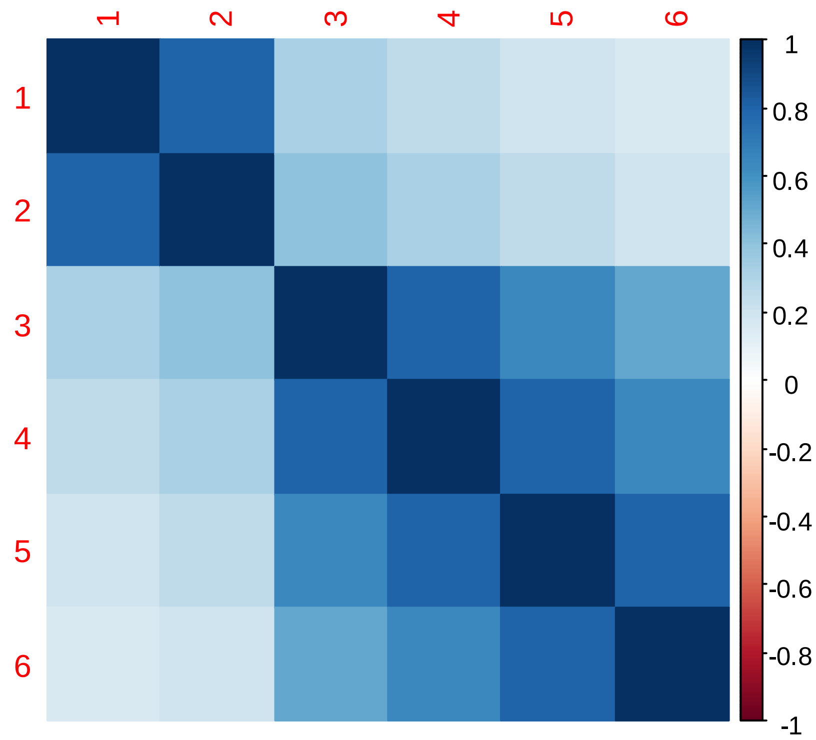
Simulation study: Methods

The data:

- Binary outcome
- Linear change over time
- Within-period ICC: 0.01-0.1
- Several different correlation structures

Simulation study: Methods

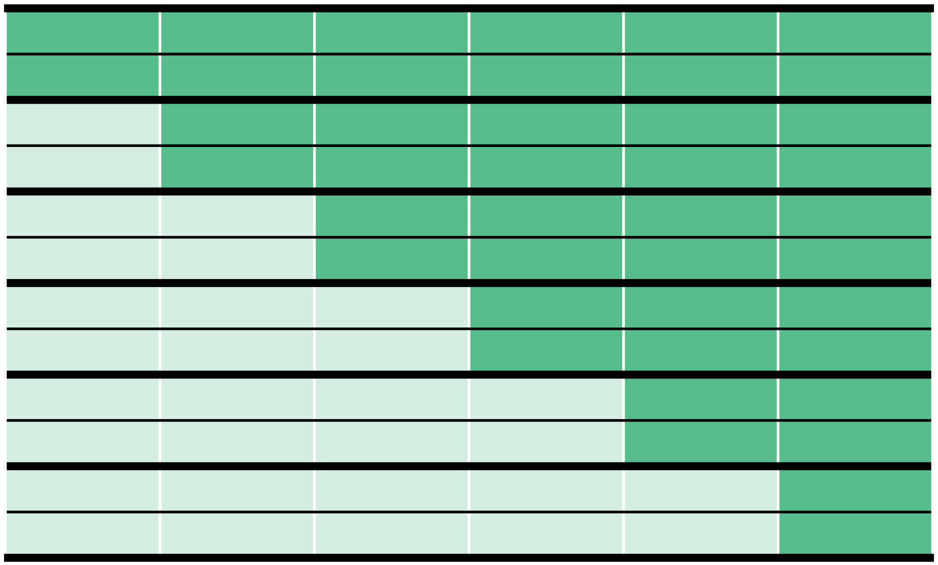
Autoregressive + Less correlation between intervention and control



Simulation study: Methods

The trial design:

- 6 sequences
- 6, 18, 48, 60 clusters
- 24, 60 observations per cluster
- Common or varying cluster-size



Simulation study: Methods

1000 simulations

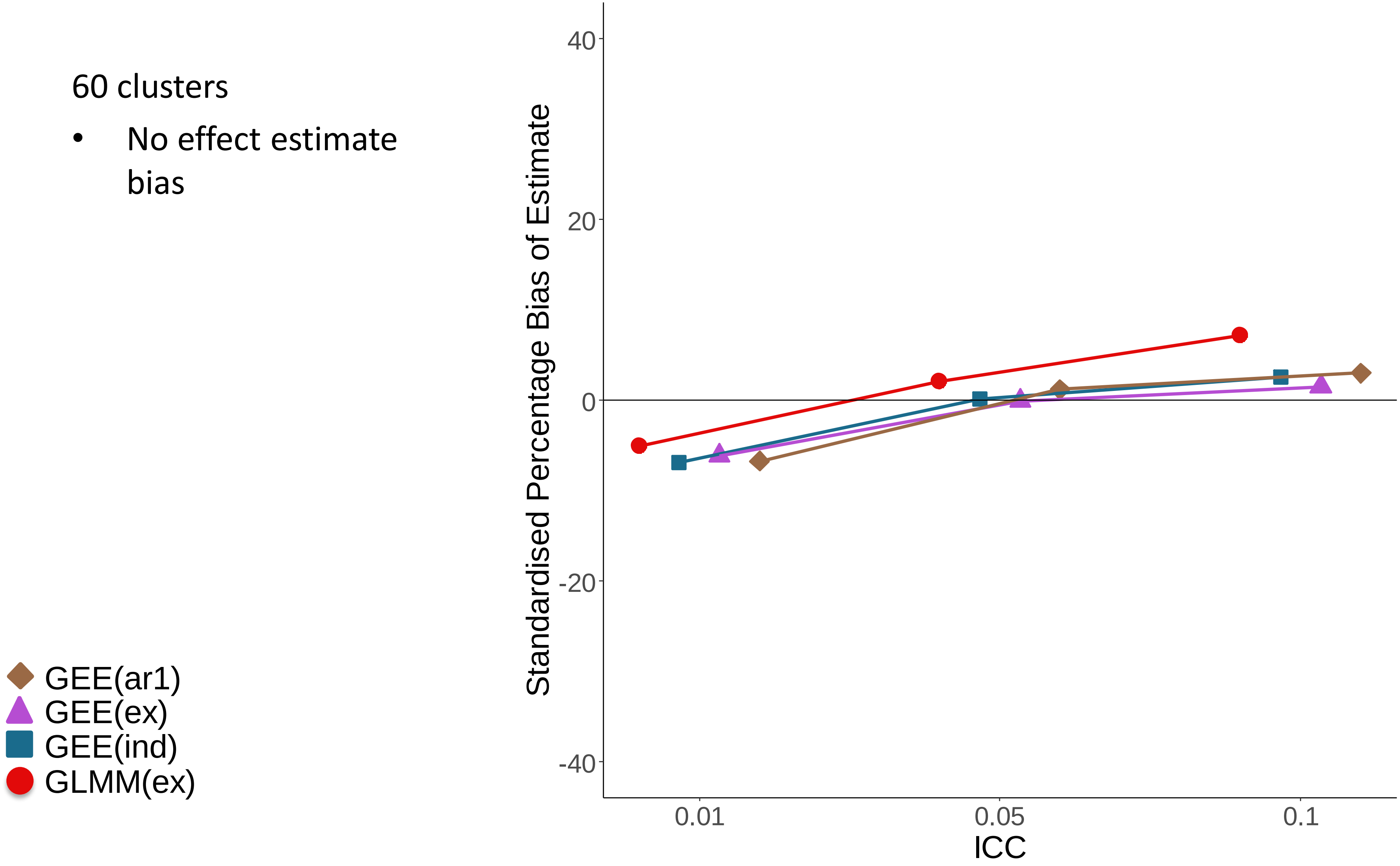
Analysis methods:

- Generalised estimating equations (robust standard errors)
  - Working correlation matrix:
    - Independent
    - Exchangeable
    - Autoregressive
  - With and without Fay and Graubard small sample correction
- Generalised linear mixed model (Hussey and Hughes)
  - random intercept for cluster
  - Assumes exchangeability within clusters
- All adjust for time as a categorical variable

Simulation study: Results

60 clusters

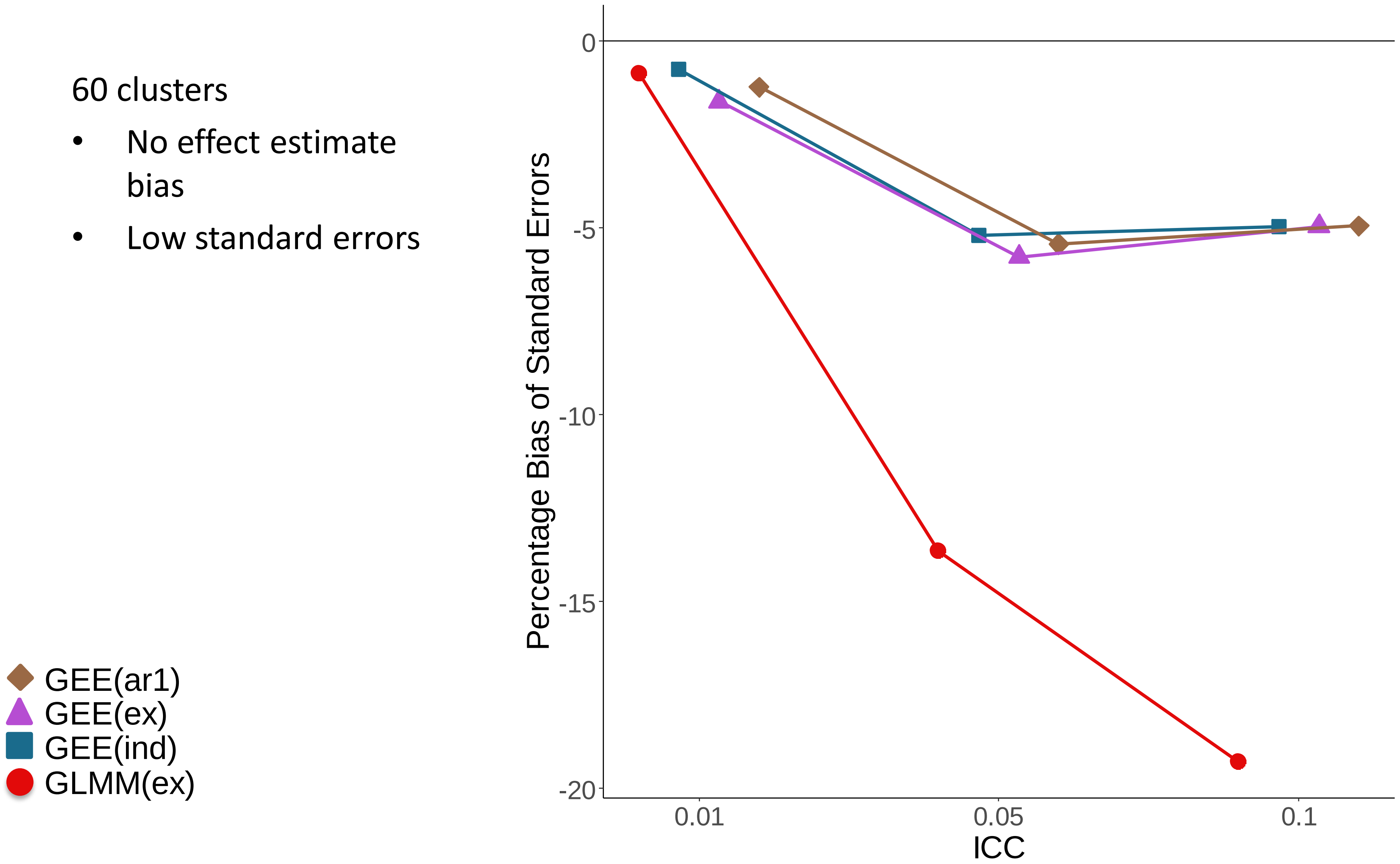
- No effect estimate bias



Simulation study: Results

60 clusters

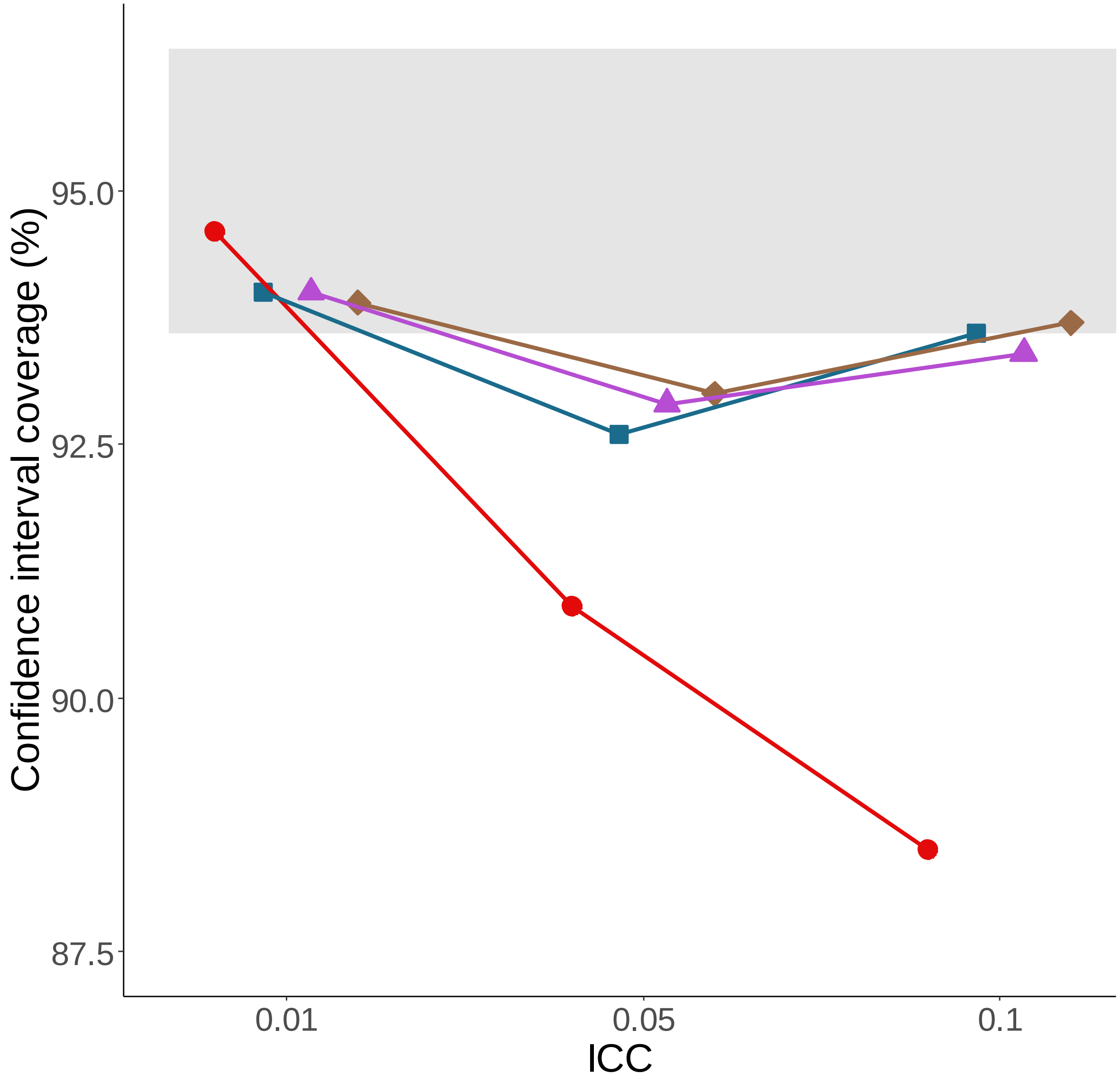
- No effect estimate bias
- Low standard errors



Simulation study: Results

60 clusters

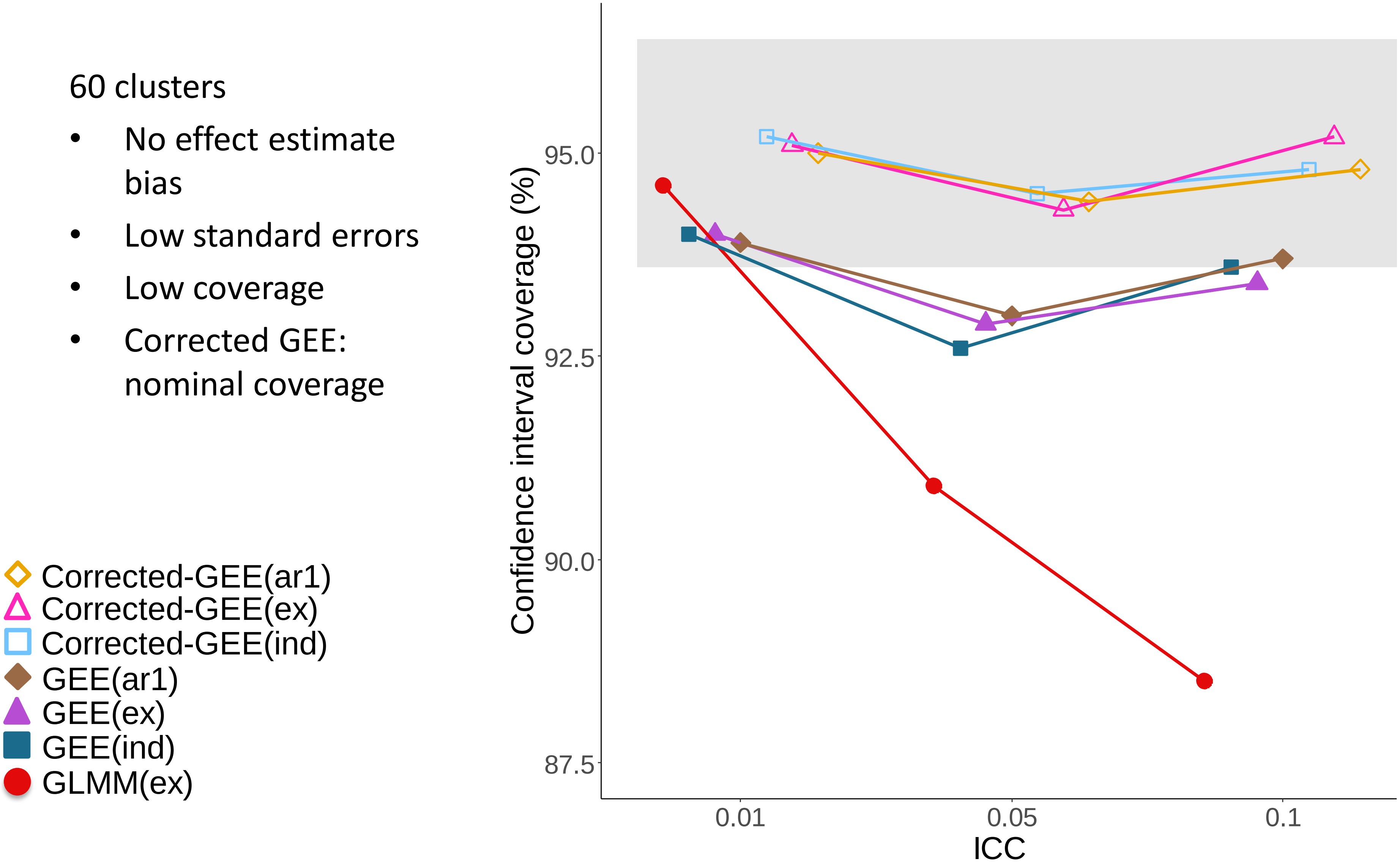
- No effect estimate bias
- Low standard errors
- Low coverage



Simulation study: Results

60 clusters

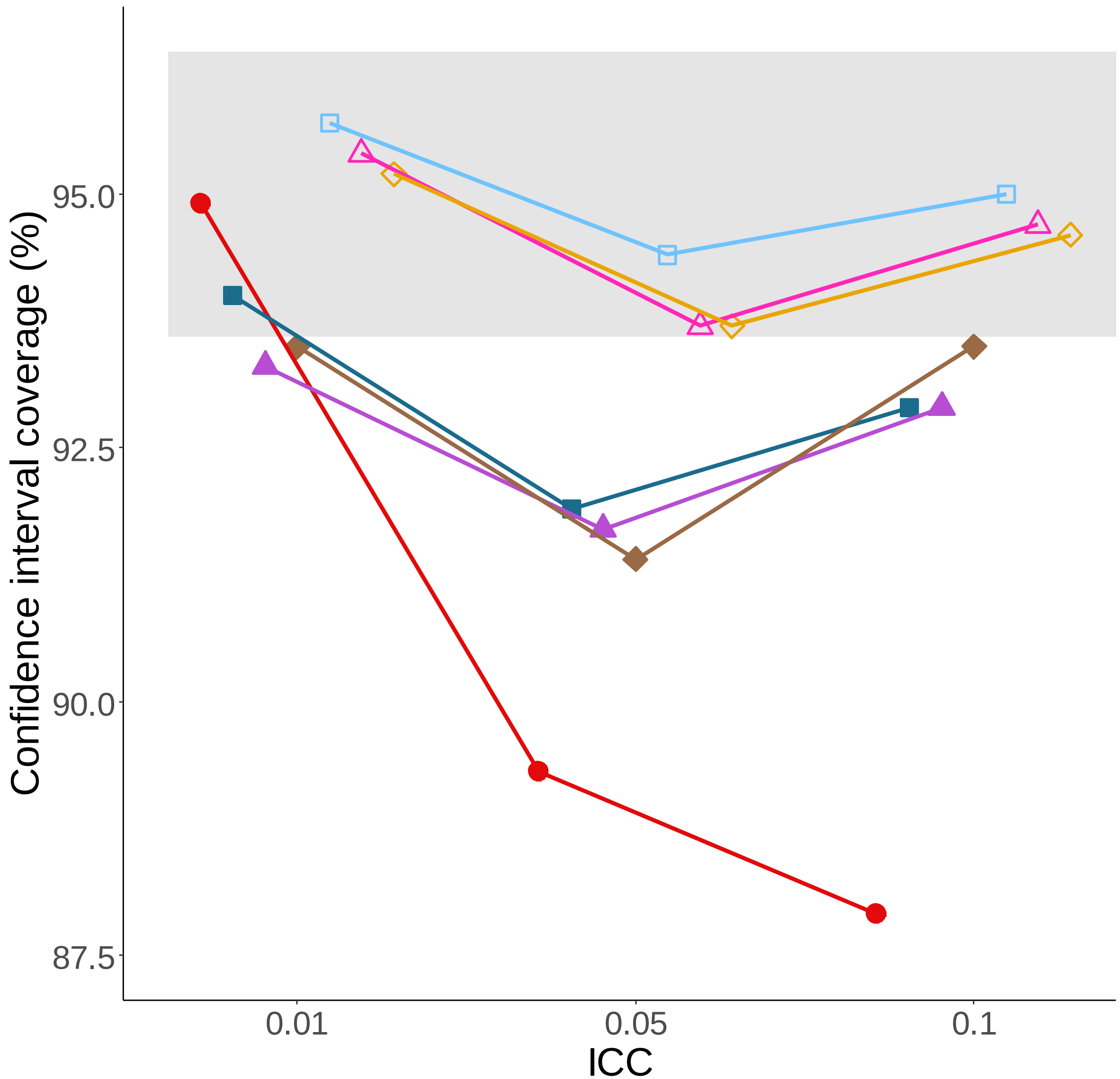
- No effect estimate bias
- Low standard errors
- Low coverage
- Corrected GEE: nominal coverage



Simulation study: Results

48 clusters

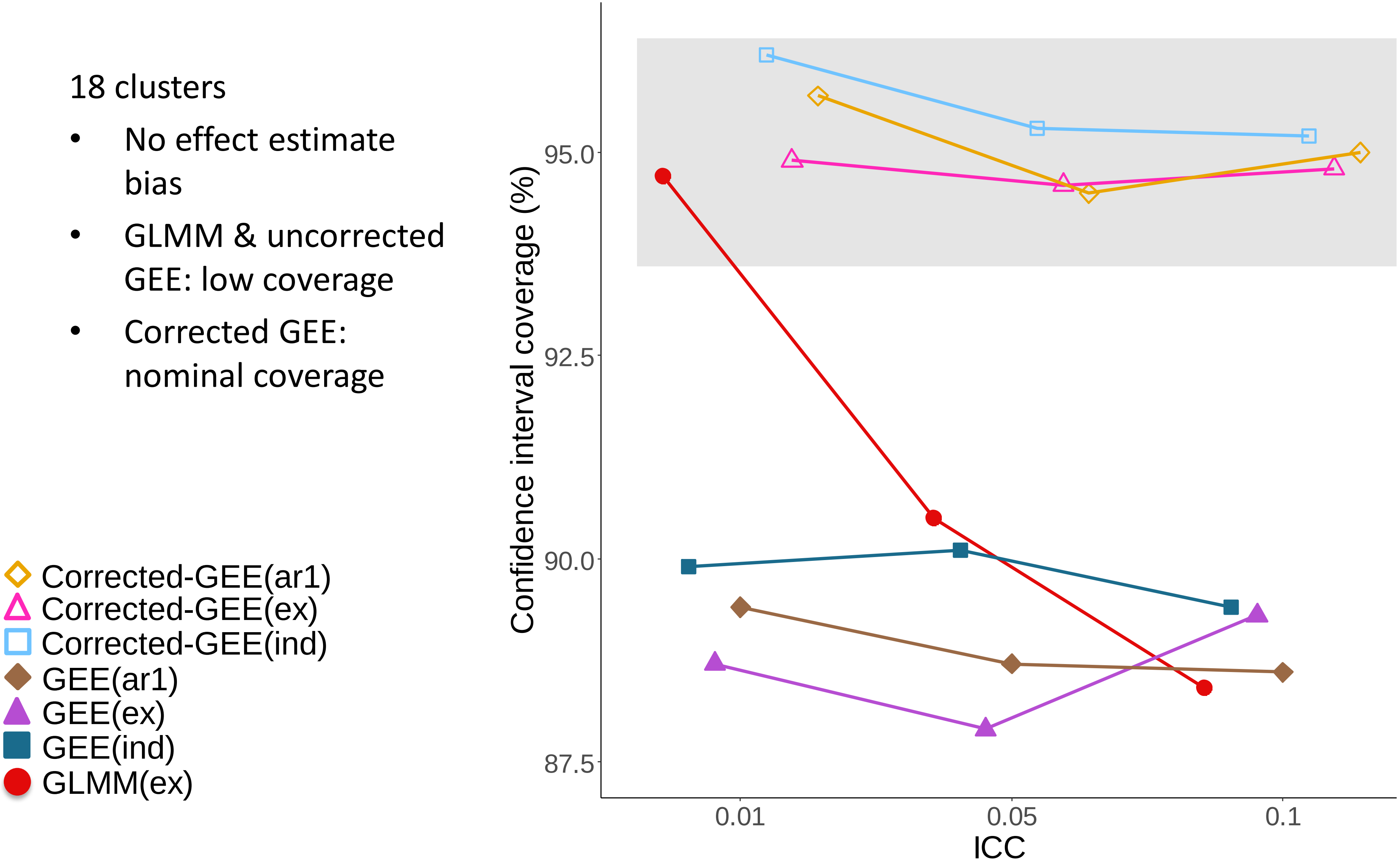
- No effect estimate bias
- GLMM & uncorrected GEE: low coverage
- Corrected GEE: nominal coverage



Simulation study: Results

18 clusters

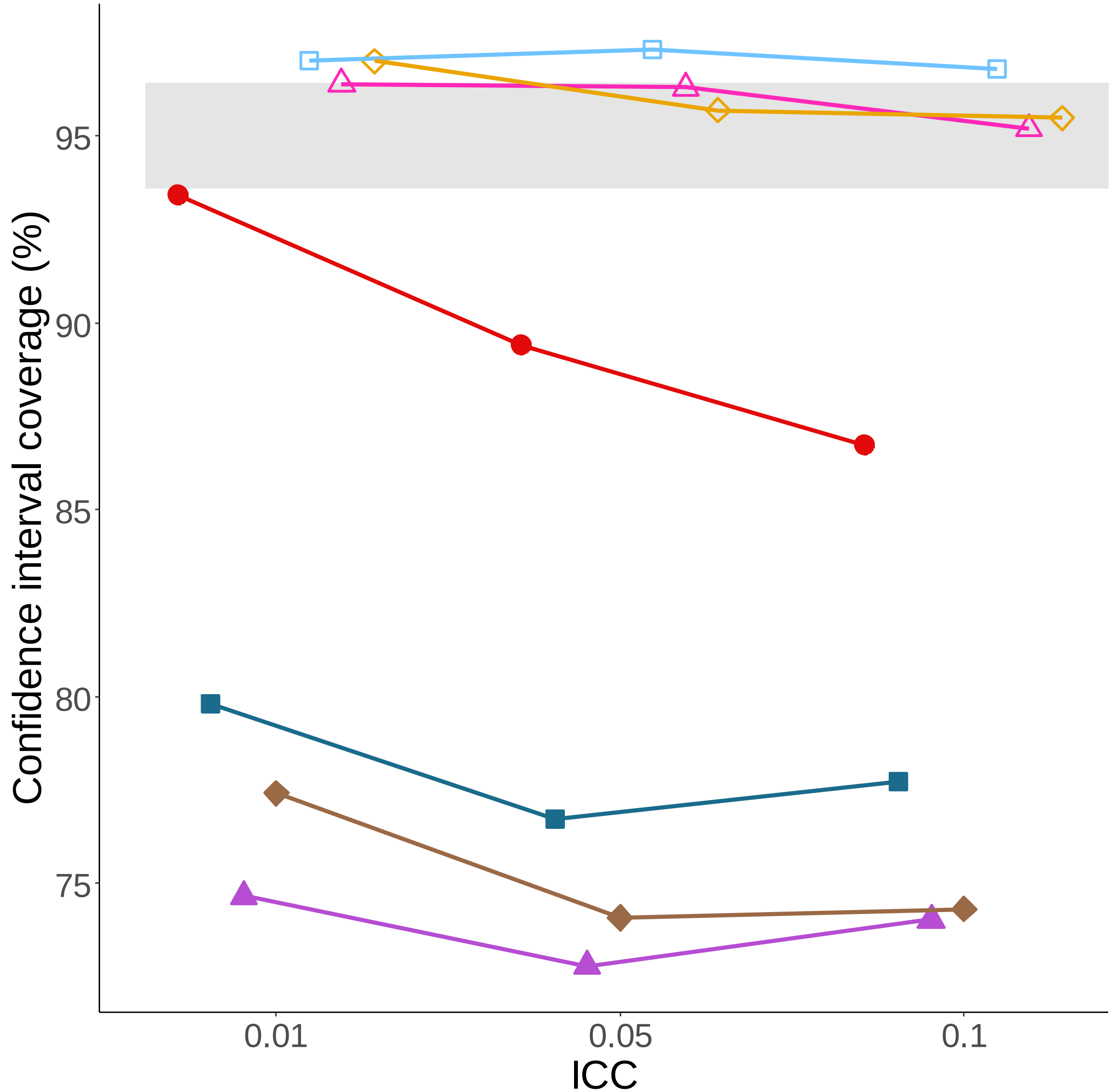
- No effect estimate bias
- GLMM & uncorrected GEE: low coverage
- Corrected GEE: nominal coverage



Simulation study: Results

6 clusters

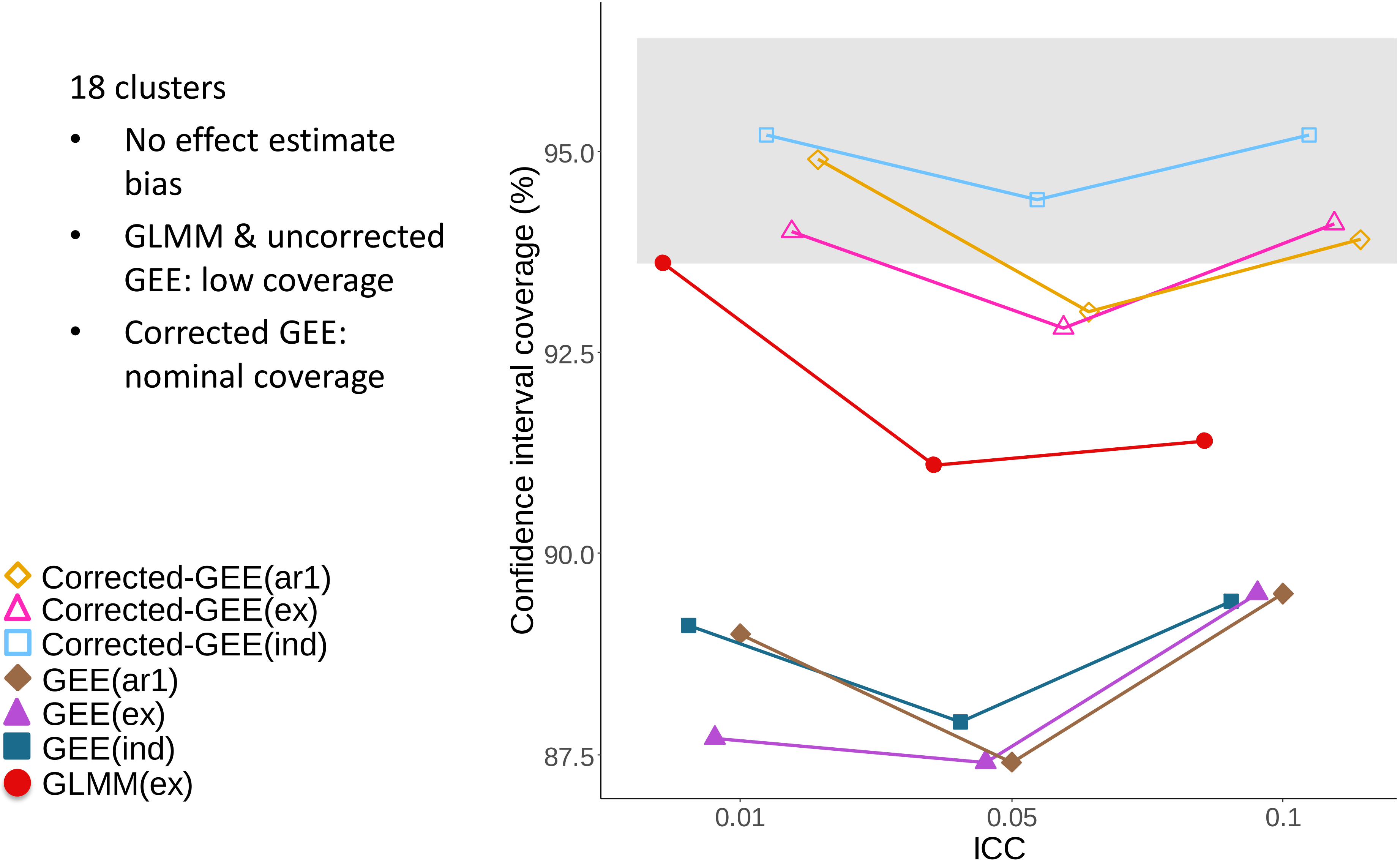
- No effect estimate bias
- GLMM & uncorrected GEE: low coverage
- Corrected GEE: conservative coverage



Simulation study: Results

18 clusters

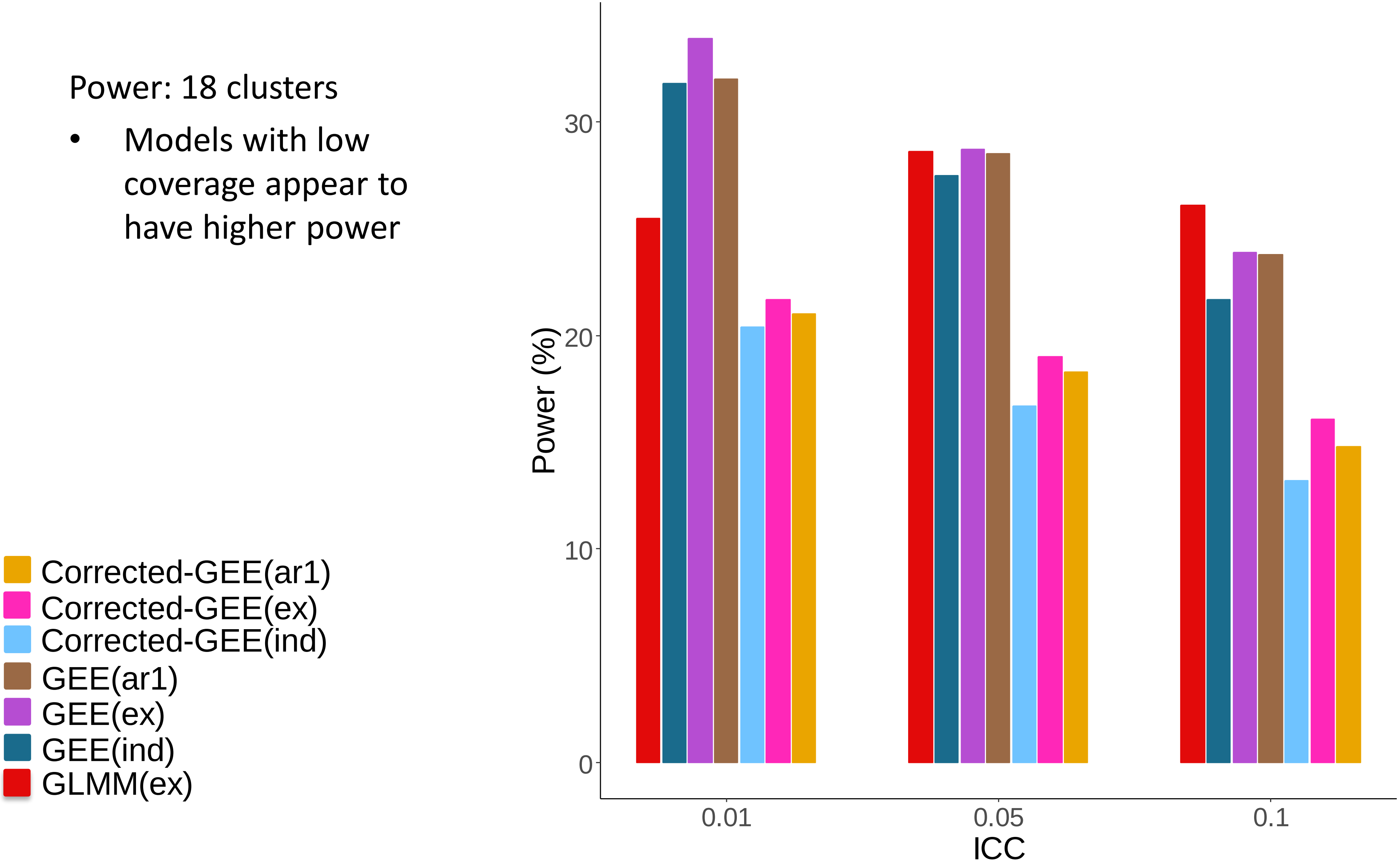
- No effect estimate bias
- GLMM & uncorrected GEE: low coverage
- Corrected GEE: nominal coverage



Simulation study: Results

Power: 18 clusters

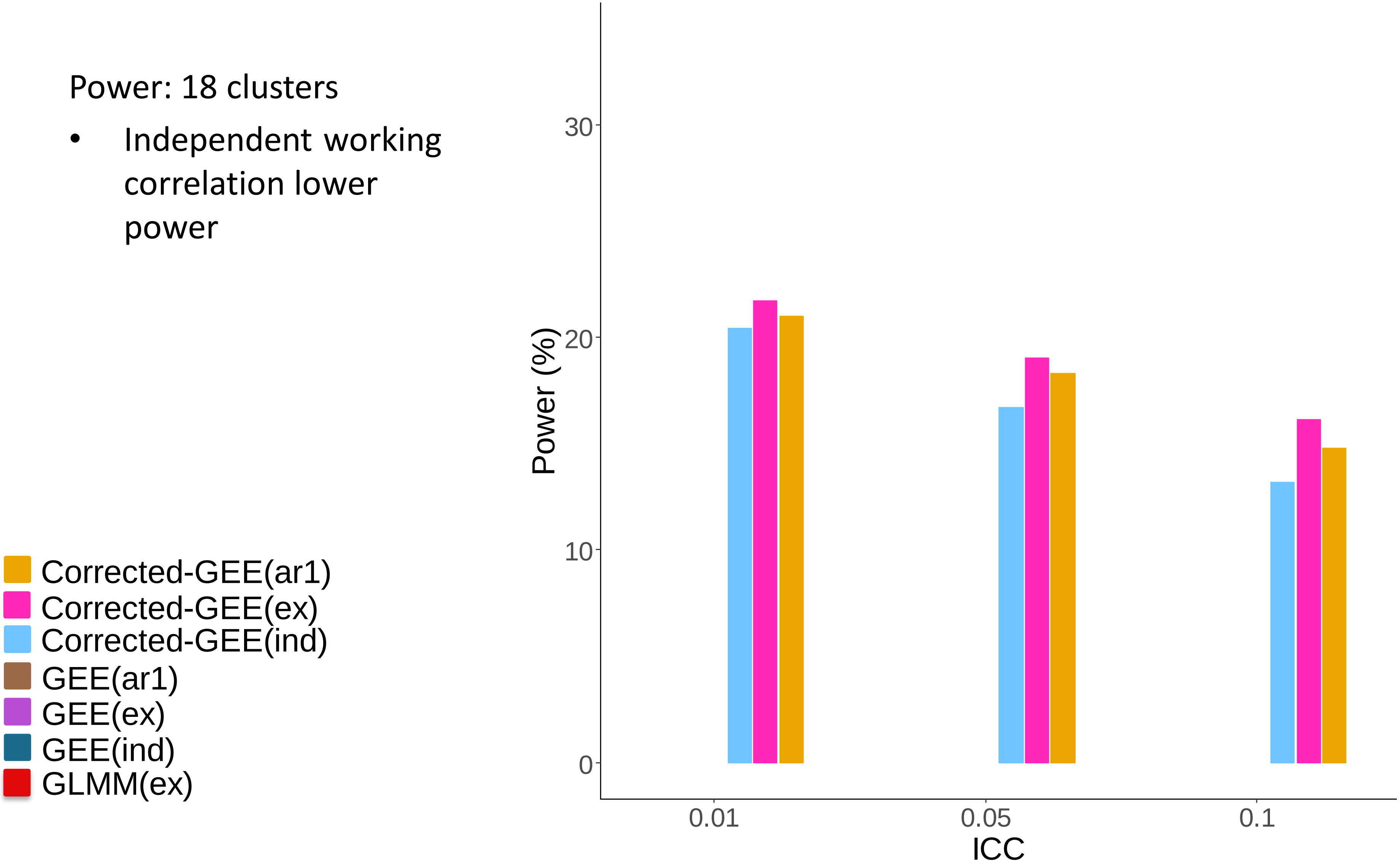
- Models with low coverage appear to have higher power



Simulation study: Results

Power: 18 clusters

- Independent working correlation lower power



Simulation study: Conclusions

- Confirmed GLMM sensitive to misspecification of the correlation structure
- GEE not affected by misspecification of the correlation structure
- Uncorrected GEE low coverage, sometimes with as many as 60 clusters
- Fay and Graubard[1] correction worked well
- Little difference between working correlation matrix
    - Suggestions that Independent most robust

1. Fay, M.P. and B.I. Graubard, *Small-sample adjustments for Wald-type tests using sandwich estimators.* Biometrics, 2001. **57**(4): p. 1198-206.

Acknowledgements:

Katherine Fielding, Richard Hayes

Thank you

jennifer.thompson@lshtm.ac.uk