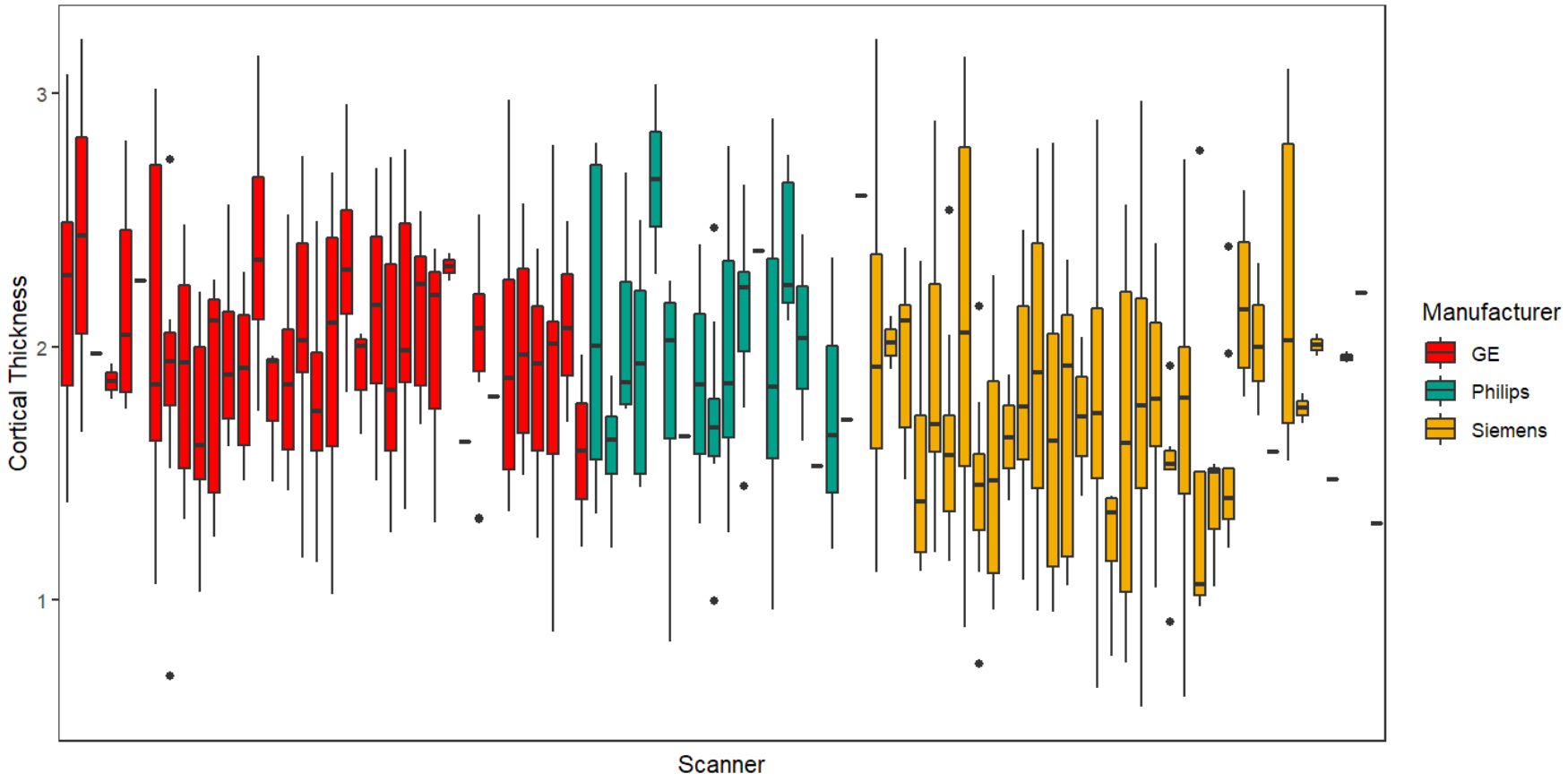
PENNSIVE

Removal of Scanner Effects in Covariance

Andrew Chen
Advised by Haochang Shou and Taki Shinohara

Heterogeneity across Scanners



Current Harmonization Method

Step 1: Assume site effect exists in mean and variance of observations

$$y_{ijv} = \alpha_v + \mathbf{x}'_{ij}\beta_v + \gamma_{iv} + \delta_{iv}e_{ijv}$$

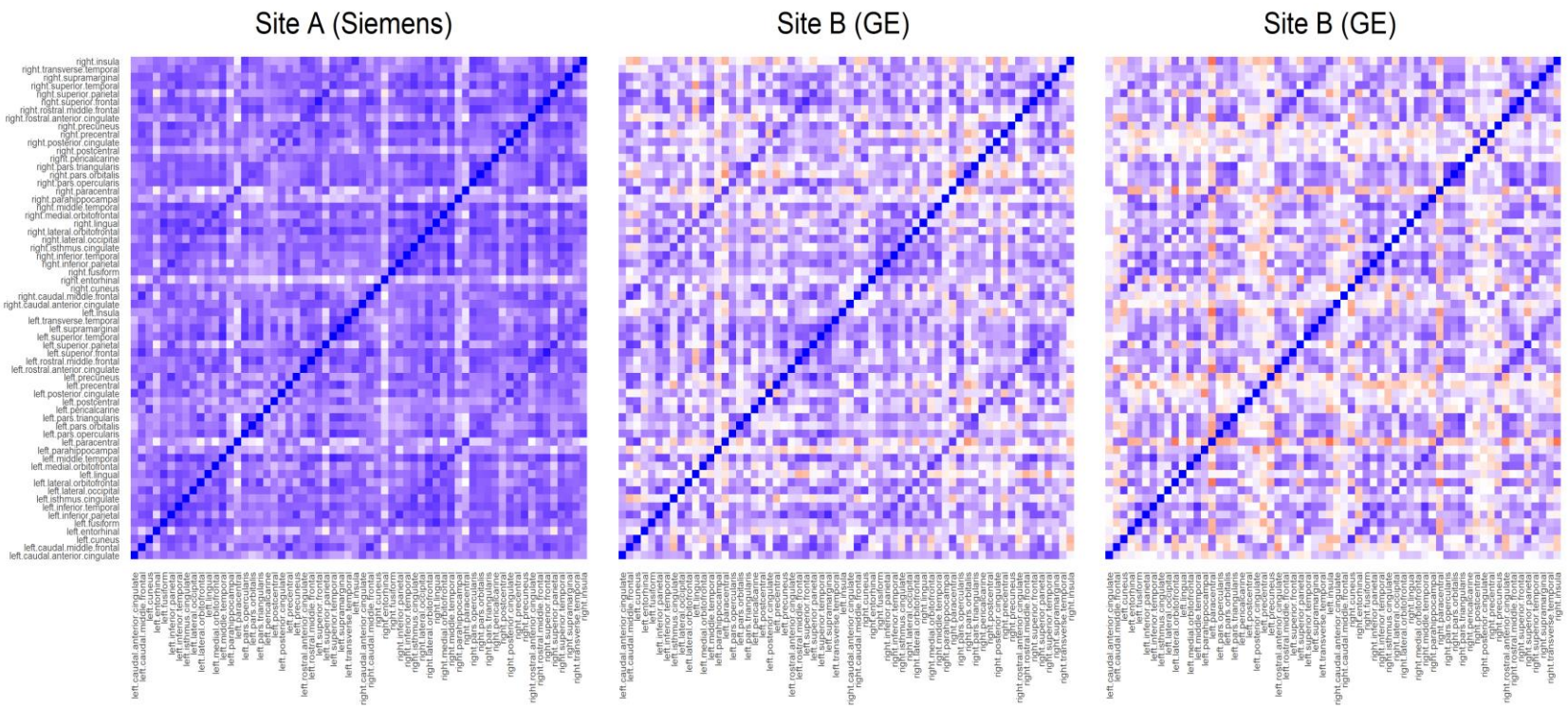
Step 2: Obtain empirical Bayes point estimates by imposing a common prior across features for both mean and variance site effect

Step 3: Remove site effect while retaining covariate effect

$$y_{ijv}^{ComBat} = \frac{y_{ijv} - \hat{\alpha}_v - \mathbf{x}'_{ij}\hat{\beta}_v - \gamma^*_{iv}}{\delta^*_{iv}} + \hat{\alpha}_v + \mathbf{x}'_{ij}\hat{\beta}_v$$

(Fortin et al., 2017, 2018; Johnson et al., 2007)

Covariance Differences across Sites



Proposed Harmonization: CovBat

Step 1: Apply ComBat to harmonize mean and variance across sites, then residualize on intercept and covariates

Step 2: Perform principal component analysis on full data to obtain

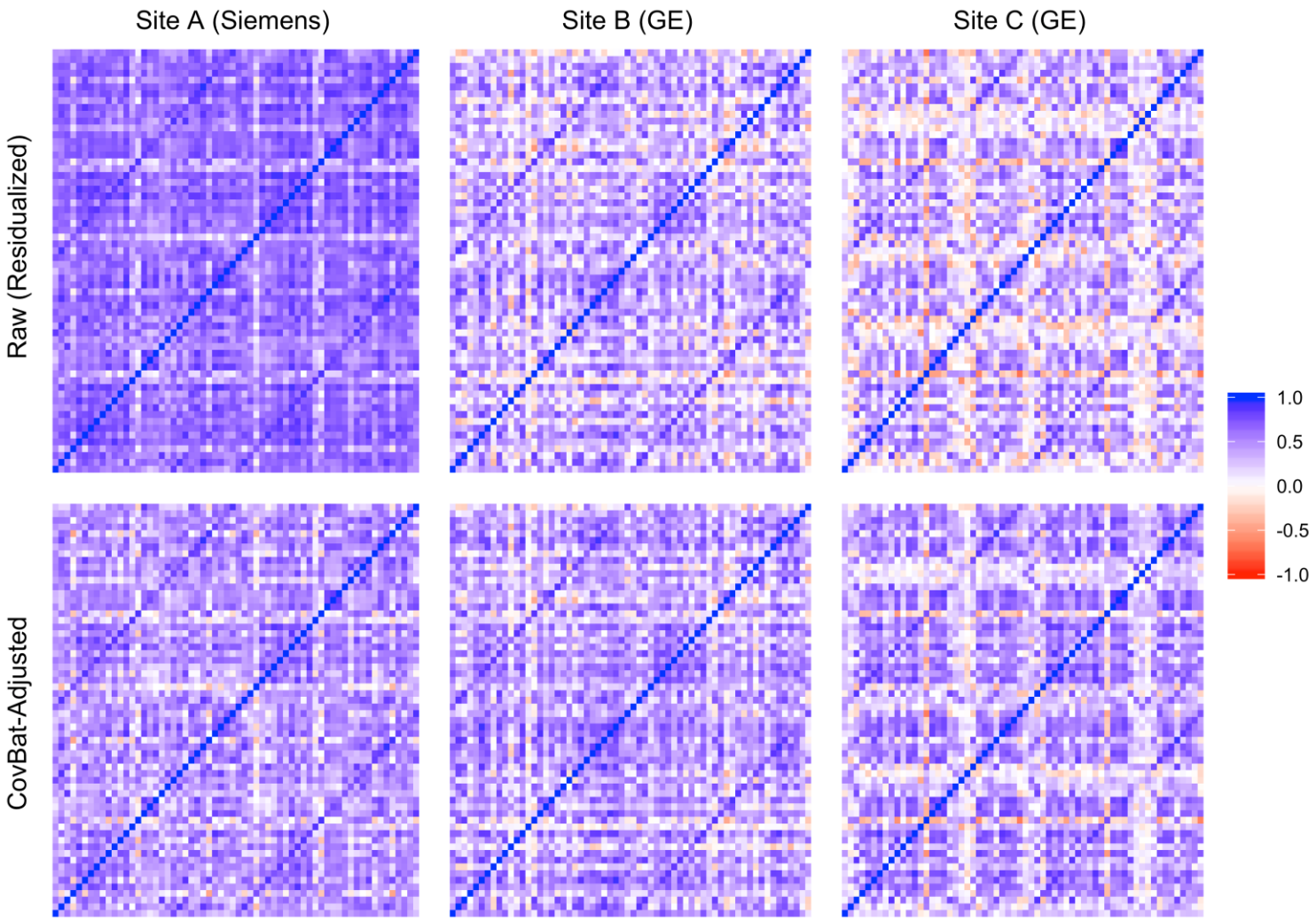
$$\Sigma_i = \sum_{k=1}^{q} \lambda_{ik} \phi_k \phi_k' \qquad e_{ij}^{ComBat} = \sum_{k=1}^{q} \xi_{ijk} \phi_k$$

Step 3: Harmonize mean and variance of first K PC scores and reintroduce intercept and covariates

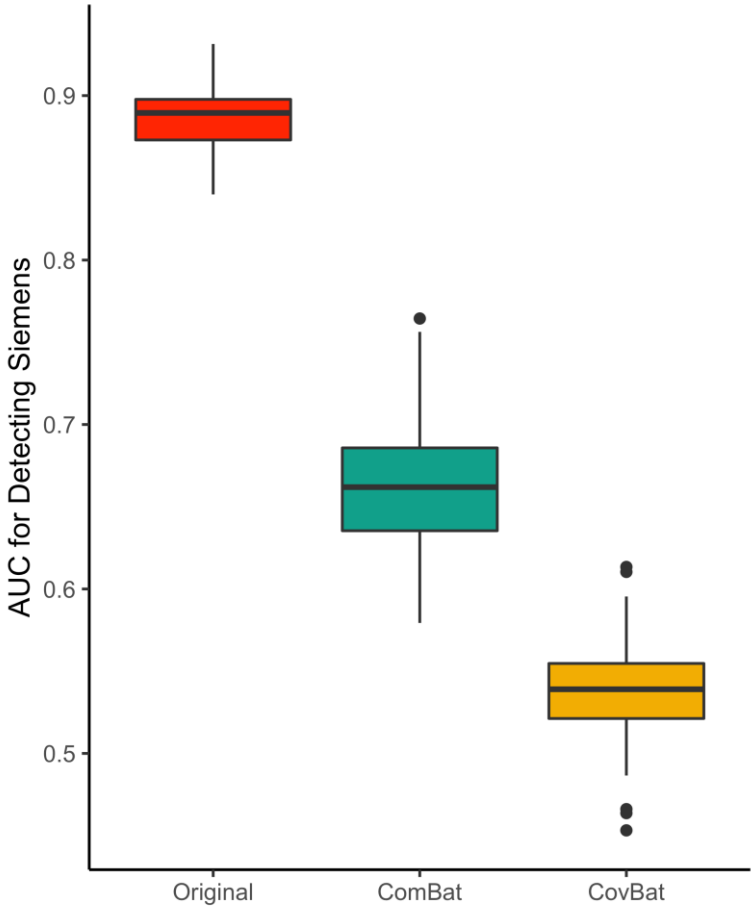
$$e_{ij}^{CovBat} = \sum_{k=1}^{K} \xi_{ijk}^{CovBat} \phi_k + \sum_{l=K+1}^{q} \xi_{ijl} \phi_l$$

$$y_{ijv}^{CovBat} = e_{ijv}^{CovBat} + \hat{\alpha}_v + \mathbf{x}_{ij}' \hat{\beta}_v$$

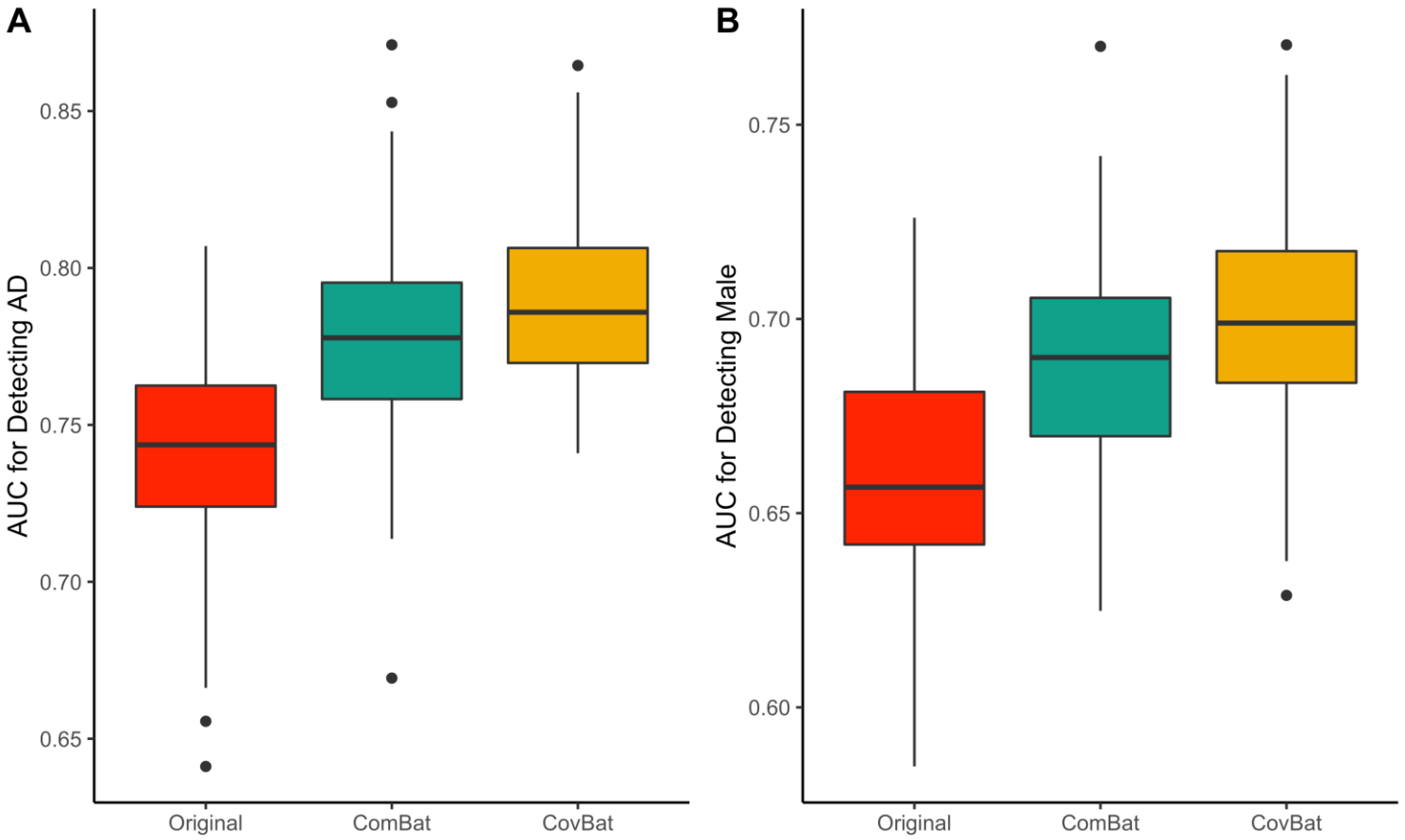
CovBat Harmonizes Covariance



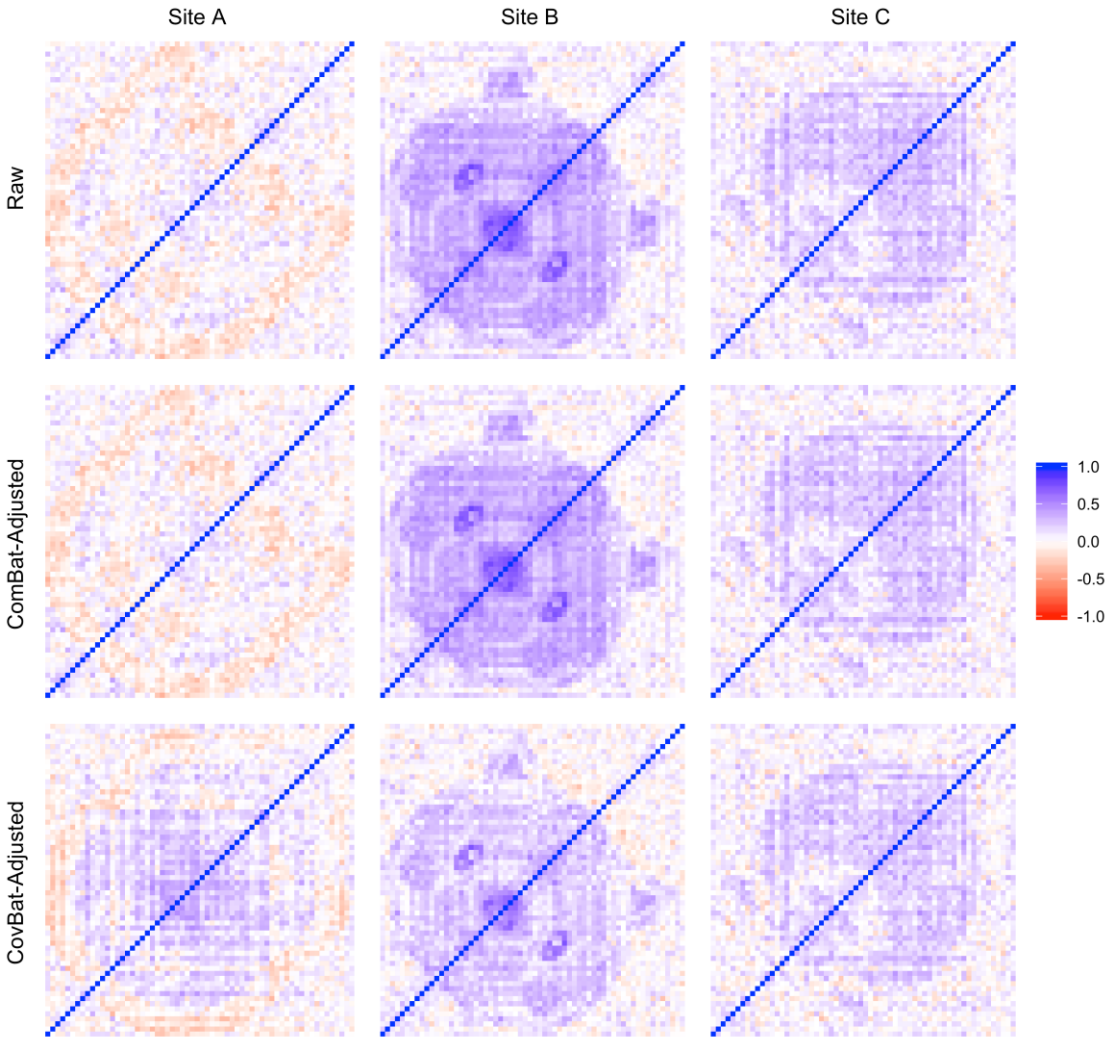
Obscures Detection of Scanner



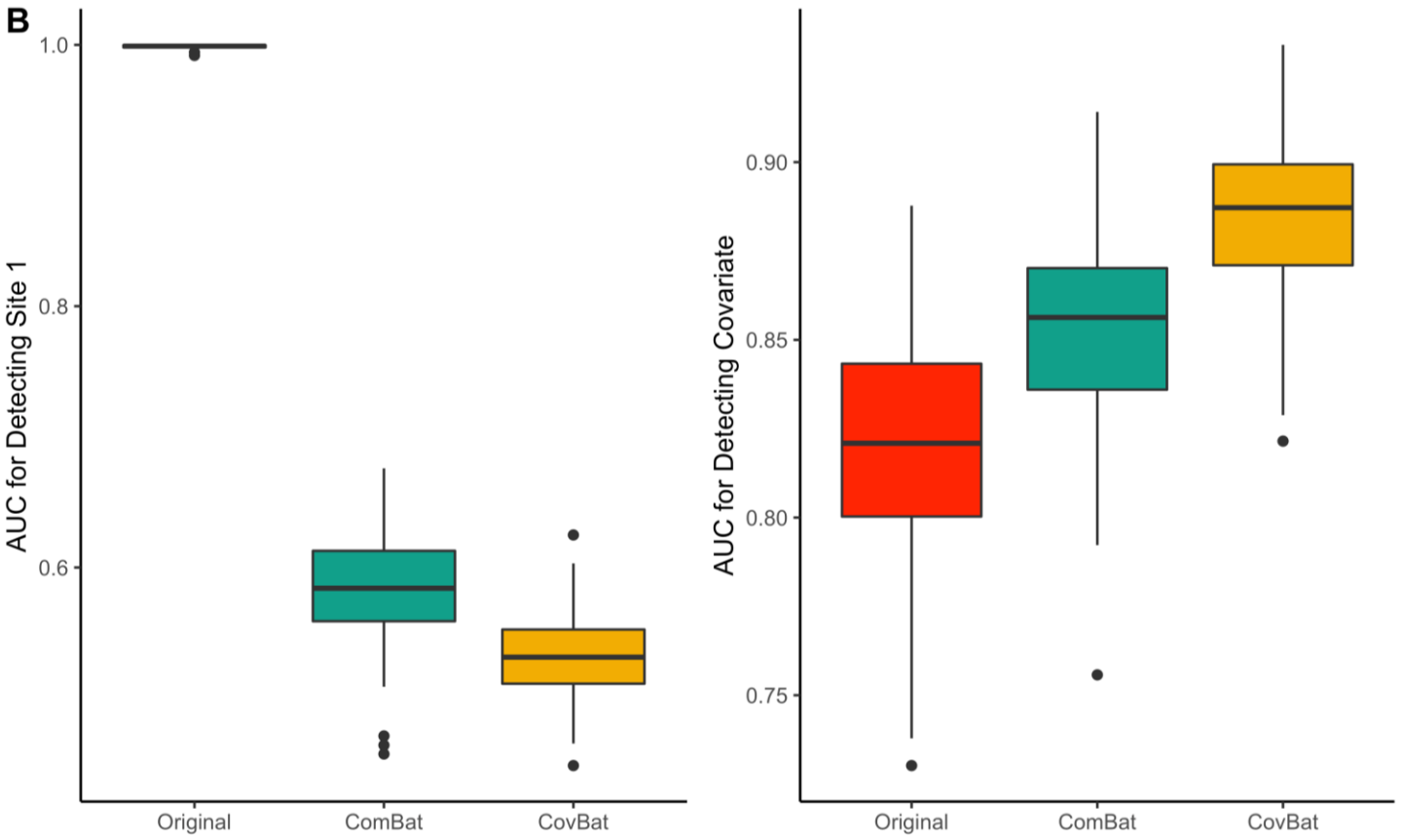
Improves Detection of Covariates



Simulation Results



Simulation Results



Acknowledgements

Preprocessing and formatting:
Joanne C. Beer
Nicholas J. Tustison
Philip A. Cook

Advisors:
Russell T. Shinohara
Haochang Shou

ComBat in neuroimaging:
Jean-Philippe Fortin

Team:
PennSIVE

References

WE Johnson, C Li, A Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007)

JP Fortin, et al., Harmonization of multi-site diffusion tensor imaging data. *NeuroImage* **161**, 149–170 (2017).

JP Fortin, et al., Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* 167, 104–120 (2018)