Title:

FEMA: Fast and efficient mixed-effects algorithm for population-scale whole brain imaging data

Chun Chieh Fan[*+1, 2, 3], Clare E Palmer[*4], John Iverson[5, 6], Diliana Pecheva[2,3], Wesley K. Thompson[1,3], Donald Hagler[2, 3], Terry L. Jernigan[4, 7, 8], Anders M. Dale[+2,3,7,9]

* Equal contribution

+ Corresponding authors

1. Population Neuroscience and Genetics Lab, University of California San Diego, USA.

2. Center for Multimodal Imaging and Genetics, University of California San Diego, USA

3. Department of Radiology, School of Medicine, University of California San Diego, USA

4. Center for Human Development, University of California San Diego, USA

5. Institute for Neural Computation, University of California San Diego, USA

6. The Swartz Center for Computational Neuroscience, University of California San Diego, USA

7. Department of Cognitive Science, University of California San Diego, USA

8. Department of Psychiatry, University of California San Diego, USA

9. Department of Neuroscience, University of California San Diego, USA

Abstract:

Despite the importance and versatility of linear mixed effects models (LME), they have seldom been used in whole brain imaging analyses due to the computational requirement. Here, we introduce a fast and efficient mixed-effects algorithm (FEMA) that makes whole brain voxelwise imaging LME analyses possible. We demonstrate the equivalency of statistical power and control of type I errors between FEMA and classical LME, whilst showing an order of magnitude improvement in the speed of FEMA compared to classical LME. By applying FEMA on diffusion images and resting state functional connectivity matrices from the ABCD Study^{SM} release 4.0 data, we show voxelwise annualized changes in fractional anisotropy (FA) and functional connectomes in early adolescence, highlighting a critical time of establishing associations among cortical and subcortical regions.

Introduction:

Linear mixed effects (LME) models have become an important tool for analyzing large-scale population data. The flexibility of random effects in LME has allowed researchers to account for repeated measures, explicit/implicit family relatedness, and longitudinal study design (Beck et al., 2021; Bernal-Rusiel et al., 2013; Chen et al., 2013; Dick et al., 2021). Complex study designs and the heterogeneity in the study population can be parameterized by the random covariance in the LME leading to an increase in statistical power and reduction of inferential biases (Bernal-Rusiel et al., 2013). However, despite the versatility of the LME, there are limited utilizations of the LME in vertex and volxe-wise neuroimaging analyses, due to computational demands.

The classical LME requires extensive computational resources. The numeric solution for LME is typically obtained through Restricted Maximum Likelihood (REML) through which the parametrized study covariance is iteratively inverted in order to identify the random effects that maximize the likelihood of the model (Bernal-Rusiel et al., 2013; Chen et al., 2013; Guillaume et al., 2014). Because of this procedure, the computational complexity scales polynomially with the number of data entries, making the classical solutions infeasible in large scale population data. It is even more challenging for whole brain analyses, as the dependent variable can be tens of thousands of voxels, surface vertices, or connectivities across region-of-interests (ROI). Even a procedure that takes a minute to fit a single LME analysis at a single voxel would require days to complete for the whole-brain. There is a need for a fast and efficient LME algorithm that is capable of handling large-scale imaging data. The arrival of the Adolescent Brain Cognitive Development (ABCD) Study ^{SM} (Casey et al., 2018) makes this unmet need more salient, as best practice's for statistical analysis are to control for study heterogeneity, ascertained relatedness due to the inclusion of multiple siblings and twins in the sample, and in order to model longitudinal measures (Dick et al., 2021; Smith and Nichols, 2018).

Previous attempts on extending LME to be applicable in large-scale neuroimaging data have relied on a divide-and-conquer approach (Chen et al., 2017; Guillaume et al., 2014). The LME toolbox for longitudinal analyses, released with Freesurfer, starts cortical surface vertexwise LME by segmenting vertices into a smaller set of ROIs that shared similar random effects and are geodesically close, thus reducing the number of outcome variables and the covariance matrix required for inverting (Guillaume et al., 2014). It currently lacks support for complex study design, such as inclusion of family members, and imaging measures beyond cortical surfaces. The LME tool implemented in AFNI utilizes the forking mechanism in R while calling the LME solver in R (from the lme4 package) to derive the results (Chen et al., 2017). It quickly becomes infeasible for large-scale population imaging data with thousands of dependent variables. The core of both Freesurfer LME and AFNI LME is REML which still needs polynomial time to estimate the parameterized covariances.

Here, we present a fast and resource efficient LME algorithm, called Fast and Efficient Mixed-effects Algorithm (FEMA), that can resource-efficiently perform whole brain analyses on complex large-scale imaging data. By utilizing the method-of-moments estimator, effect binning, and sparsity of the random effect design matrix, FEMA can finish whole brain voxelwise LME analyses within minutes, comparing to days using REML solvers, without the need for large memories or parallelization. Using both realistically simulated data and real data from the ABCD Study, we have demonstrated that the results from FEMA are consistent with classical REML results while the required resources and computational times are dramatically reduced, allowing researchers to quickly examine the relationships between neuroimaging metrics and variables of interest while considering repeated measures, family structures, and complex study designs.

Material and Methods

Figure 1 illustrates the algorithmic design of the FEMA. FEMA involves three main steps. First, FEMA sets up the model as a classical LME with random intercept that characterizes the within/between individual variations. Second, the random effects are estimated through moment-estimators instead of classical maximum likelihood approach, utilizing the benefit of large sample size and sparse indicator matrices. In the third step, the random effects are binned into finite discrete configurations, reducing the need for inverting parameterized covariance matrices, and then the estimated covariance matrices are used in weighted least squares to obtain the unbiased estimation on the fixed effects of variables of interest. In the following section, we elaborate each of these three steps.

Model Set Up

As in most of the classical analytic approaches in brain mapping, we assume an additive linear relationship between imaging measures (Y) and covariates of interest (X). In matrix form, it is:

$$Y = XB + \mu$$

where

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1j} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nj} \end{bmatrix}$$

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1c} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nc} \end{bmatrix}$$

for n subjects, j voxels/vertices, and c covariates of interest. In terms of massive univariate analyses, the fixed effect regression coefficients take the matrix form as:

$$B = \begin{bmatrix} \beta_{11} & \cdots & \beta_{j1} \\ \vdots & \ddots & \vdots \\ \beta_{1c} & \cdots & \beta_{jc} \end{bmatrix}$$

Meanwhile, the random intercept, μ, captures the correlations among individuals attributable to characteristics of the data entries:

$$\mu_j \sim N(0, V_j) \text{ for j imaging measures.}$$

For example, given the longitudinal design and inclusion of family members, the V can be parameterized as the variance attributable to living in the same family (F), repeated measures from the same subject (S), and random errors (E):

$$V = \sigma_F^2 Z_F Z_F{}' + \sigma_S^2 Z_S Z_S{}' + \sigma_E^2 I$$

as Z is the indicator matrix specifying the membership of each observation, whether it comes from the same person (repeated measures, S) or from the same family members (family effect, F).

The main computational bottleneck here is the estimation of V for every imaging outcome measure, as the computational complexity is at least $O(n^2)$ if REML is used for each imaging measure. When the number of imaging measures, j, is large, such as voxelwise analyses, solving REML for every single measure would be infeasible. In the following section, we would discuss how our algorithm overcomes those two computational issues.

<u>Estimating the variance components</u>

To overcome the computational bottleneck of estimating parameterized covariance matrix, V, we implemented the method of moments estimator to obtain the values of σ_F^2, σ_S^2, which is fast and unbiased with some sacrifice on statistical efficiency (Ge et al., 2016; Haseman and Elston, 1972). It can be obtained as:

$$E[y_m y_l] = \hat{\sigma}^2 z_m z_l$$

for m, l pairs in a random effect cluster, indexed by Z. Because of the matrix formation with the moment estimator, we can simultaneously obtain the variance parameters for all voxels/vertices.

Estimating fixed effects for the variables of interest

After deriving the parameterized covariance, the estimation of B takes the form as weighted least squares

$$\widehat{B_j} = (X'V_k{}^{-1}X)^{-1}X'V_k{}^{-1}Y_j$$

with variance as

$$var(B_j) = (X'V_k{}^{-1}X)^{-1}$$

To further speed up the process, instead of inverting j parameterized covariance, V, we grouped the imaging measures based on the values of estimated random effects. In series of simulations, we would show how a finite number of gridded bins would be sufficient to capture the variance components for all input imaging measures while dramatically reducing the required computational time.

Simulations

To evaluate the validity of our algorithm, we performed both forward simulations with data generated from random effects models and empirical covariance structure from realistic data, i.e. ABCD study. Here, we randomly sampled X and family indices from the real ABCD data. We then forward simulated

the imaging measures Y, based on the sampled X and variance components defined by family indices. We randomly generated the true values of variance parameters and the fixed effects coefficients to see how different variance component values would impact the validity and accuracy of the estimation of the parameters of interest. In each simulation scenario, we compared our FEMA with REML based LME (REML-LME) in terms of: 1) validity and reliability of the fixed effects estimation; 2) the computational cost of the algorithms; and, 3) the impact of model misspecification.

Empirical Application

To demonstrate the utility of FEMA, we applied our method to ABCD data release 4.0 to examine longitudinal changes in imaging measures during adolescence. We investigated how voxelwise diffusion imaging measures and resting-state connectivity measures changed over time as ABCD participants grew older. In our analyses, the family IDs and repeated scans were treated as indicators for random intercepts, capturing the variance attributable to shared family relatedness and the within-person longitudinal stabilities (Dick et al., 2021). The confounding variables include age-at-enrollment, sex, population structure (first 10 genetic principal components), MRI scanner, MRI software versions, parental incomes, and parental education levels. The variable of interest was the difference in age between the first imaging scan and the second imaging scan, allowing us to examine how the imaging measures changed over development. Given both imaging modalities have more than 100K measurements per scan, it is infeasible to apply classical LME for those datasets. Therefore, we only compared FEMA with naive applications of generalized linear models (GLM), showcasing the benefit of having a fast and efficient LME solver.

Results:

Simulations

Comparing to REML-LME, FEMA dramatically reduced the computational time without using parallel processing (Figure 2a and Figure 2b). With data entries equivalent to ABCD release 4.0 (n = 18K), FEMA obtained results from 5000 voxels within seconds, whereas the REML solver required more than 2 hours to finish just 2000 voxels (Figure 2b). Meanwhile, the statistical power was equivalent between FEMA and REML-LME across different sizes of input observations (Figure 2c). The type I errors were well controlled for both FEMA and REML-LME regardless of the underlying true number of random effects configurations (Figure 2d).

Indeed, when we deliberately simulated the scenario of having the number of random effects mis-specified, such as there are 100 true random effects across 5000 voxels, but the FEMA specified the number of bins wrong, the estimated fixed effects were still robust (Figure 2e). On the other hand, since the computational time of FEMA only scales linearly with number of k bins (Figure 2f), users can increase the resolutions of random effects without worrying about the feasibility of analyzing whole-brain data.

Empirical results - developmental changes on fractional anisotropy across voxels

To showcase the utility of FEMA, we performed voxelwise analysis on the diffusion images of ABCD release 4.0 data. We examined the rate of change in fractional anisotropy (FA) while controlling for age at recruitment, sex, scanner, schousehold income, highest parental education, and first 10 genetic principal components. The subject-wise repeated measures (S) and family relatedness (F) were specified as random intercepts in FEMA. Despite large numbers of voxels (100x100x130), eligible observations (n=13,428 entries), clusters of families (N = 8197), and repeated imaging scans (5,022 second scans), FEMA finished the analyses in 60 seconds without the use of parallel computing. The estimated random

effects indicate modest family effects across voxels and large variations in longitudinal stability (Figure 3A). Removing the F and S from the random errors results in improvement of statistical power, and the voxels with high family effects or longitudinal stability benefit the most (Figure 2B). The random effects have distinct anatomical distributions despite no spatial priors being used in the estimation process (Figure 3C for the variance in longitudinal stability, S, and Figure 3D for the variance in family effects, F). For the fixed effects, the annualized rate of change in FA was most evident within the cingulum and fibers right beneath the cortical ribbon (Figure 3E). Major fiber bundles such as corticostraite tracts instead showed no evident associations when thresholding at a conservative Bonferroni threshold ($p = 0.01/\text{total number of voxels} = 7.7e\text{-}9$).

Empirical results - developmental changes on resting state functional connectivies across cortical and subcortical regions

The versatility of FEMA also allows us to examine the annualized rate of change in resting state connectivities (Gordon's parcellations and subcortical regions, Figure 4). With the same model set up as the voxelwise analysis on FA, we examined how the resting state connectivities across cortical and subcortical regions changed over time. FEMA showed that the resting state networks had relatively low family effects and modest longitudinal stability (Figure 4A). Nevertheless, the inference on the fixed effect of annualized rate of change still benefited from having the S and F partitioned out from the residual errors (Figure 4B). Both S and F seems to follow the modularization of the connectivity groups (Figure 4C and 4D) (Baria et al., 2011; Gordon et al., 2016; Gordon et al., 2020). The rate of change was most evident within each module (Figure 4E), as increased connectivities within each network module were statistically significant after Bonferroni correction.

Discussion

The aim of FEMA is to make whole brain LME analyses feasible. We have shown that while the statistical power and type I errors are equivalent between FEMA and REML-LME, the computational times were dramatically reduced in FEMA, making it possible to perform whole brain imaging analyses quickly and efficiently regardless of imaging modality. FEMA is robust to model mis-specification and when there is strong evidence for family effects and longitudinal stabilities, FEMA can greatly improve power for statistical inference by partitioning this variability from the residual errors.

Our empirical investigations of age associations show an interesting pattern of how the human brain changes throughout early adolescent development in terms of tissue microstructure (Figure 3) and resting-state functional connectivity networks (Figure 4). For both imaging modalities, there were widespread changes with age across the brain highlighting the benefit of using a voxelwise analysis.

Using FEMA, we were able to map the estimated random effects across the brain. Longitudinal stability showed larger effects compared to family for both modalities. For FA, voxels nearer the cortical surface showed greater longitudinal stability than voxels more inferior in the brain and across subcortical regions. This potentially highlights how the microstructure of the cortical sheet may act as a unique 'fingerprint' afforded to each subject that is stable over time. Subcortical voxels showed greater family effects, which may highlight that FA in these regions is either highly heritable or is related to experiencing a common environment. With this model design it is difficult to tease apart these factors as our measure of family is based only on a measure of whether individuals belong to the same family unit. Future work will aim to incorporate genetic relatedness and other environmental data to better partition these random effects. The resting-state functional connectivity networks showed a similar difference in random effect sizes as FA, i.e. higher longitudinal stability than family effects, across multiple networks. The dorsal attention, fronto-parietal and visual attention networks appeared to show

the greatest longitudinal stability highlighting that connectivity within these networks may be particularly reliable within an individual over early adolescence.

By making FEMA publicly available, we hope to enable imaging researchers to exploit the versatility of LME. The distributions of the random effects can provide insight into how family effects associated with imaging measures and how reliable the imaging measurements are over time within a subject. The improved statistical power due to partitioning the variance components can greatly enhance our capacity for detecting important longitudinal effects. Moreover, FEMA will be implemented within the Data Exploration and Analysis Portal (DEAP) in order to make large-scale, whole-brain analyses accessible to all researchers regardless of access to extensive computational resources. This contributes to an ongoing effort to make science and ABCD data analysis more equitable.

Code Availability

FEMA will be publicly available on github (https://github.com/cmig-research-group). The development version can be obtained by sending the request to authors.

Data Availability

Adolescent data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development Study (ABCD Study®) (https://abcdstudy.org), held in the NIMH Data Archive (NDA).

Acknowledgement

This work was supported by grant R01MH122688, RF1MH120025, and R01MH118281 funded by the National Institute for Mental Health (NIMH). The ABCD Study is supported by the National Institutes

of Health and additional federal partners under award numbers: U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, and U24DA041147.

Conflict of Interest

Dr. Dale is a Founder of and holds equity in CorTechs Labs, Inc, and serves on its Scientific Advisory Board. He is a member of the Scientific Advisory Board of Human Longevity, Inc. and receives funding through research agreements with General Electric Healthcare and Medtronic, Inc. The terms of these arrangements have been reviewed and approved by UCSD in accordance with its conflict of interest policies. The other authors declare no competing interests.

Figure Legends

Figure 1. Estimation process of FEMA. First penal illustrates the example of model set up, assuming three random effects from family (F), repeated measures (S), and random errors (E). The second step involves in estimating the random effects through moment estimator. After binning voxels/images into k groups, the fixed effects are derived via weighted least square (Panel III).

Figure 2. Results from simulations. Blue lines are results from REML-LME. Red lines are results from FEMA. A. CPU time as a function of sample size, estimating 2000 voxels. B. CPU time as a function of number of voxels, holding sample size constant at 16000. C. Statistical power of detecting effect has Cohen's d greater than 0.005. D. Type I errors given different number of random effect binning. E. Mean square errors of the fixed effects as a function of number of random effect binning. F. CPU time as a function of effect binning.

Figure 3. Rate of changes on fractional anisotropy (FA) among ABCD participants. A. Random effects distribution of FA across all voxels. B. Power differences between FEMA and naive GLM given different levels of significance threshold (Bonferroni correction for P_{adjust} = 1e-2, 1e-4, 1e-6, and 1e-8) as a function of partitioned random effects (F + S). C. Spatial distribution of S. D. Spatial distribution of F. E. Voxelwise results of significant rate of changes on FA.

Figure 4. Rate of changes on resting state connectivities among ABCD participants. A. Random effects distribution across all pairwise connectivites. B. Power differences between FEMA and naive GLM given different levels of significance threshold (Bonferroni correction for P_{adjust} = 1e-2, 1e-4, 1e-6, and

1e-8) as a function of partitioned random effects (F + S). C. Distribution of S given all pairwise connectivities between Gordon's parcels and subcortical regions. The region-of-interests (ROIs) were sorted according to the network modules. D. Distribution of F given all pairwise connectivities between Gordon's parcels and subcortical regions.. E. Significant rate of changes on resting state connectivities.

Figure 1.

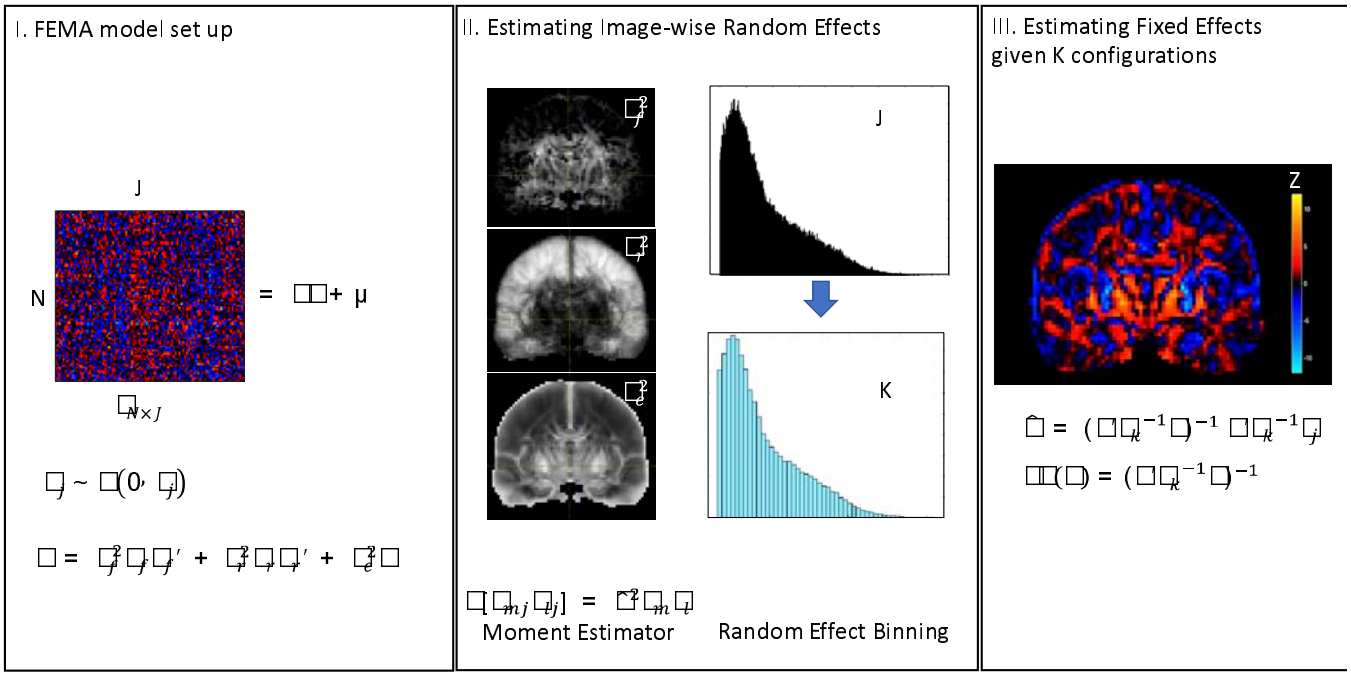

I. FEMA model set up
J
N
= □□+ μ
□N×J
□j ~ □(0, □j)
□ = □j2□j□j′ + □i2□i□i′ + □e2□
II. Estimating Image-wise Random Effects
□j2
□i2
□e2
J
K
□[□mj□lj] = □2□m□l
Moment Estimator
Random Effect Binning
III. Estimating Fixed Effects given K configurations
Z
□̂ = (□□k−1□)−1 □□k−1□j
□□(□) = (□□k−1□)−1

Figure 2.

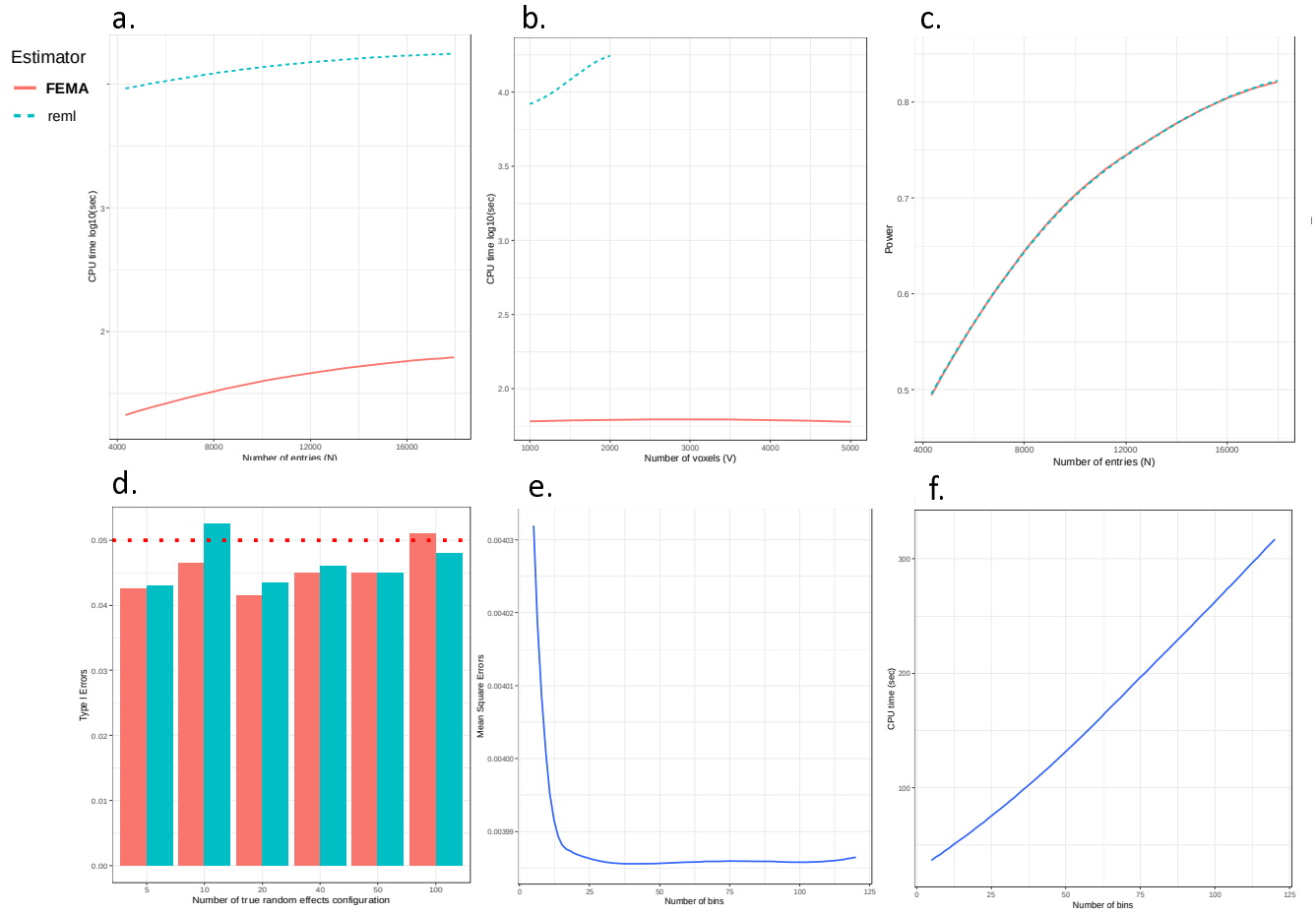

a.
Estimator
FEMA
reml
CPU time log10(sec)
Number of entries (N)
b.
CPU time log10(sec)
Number of voxels (V)
c.
Power
Number of entries (N)
d.
Type I Errors
Number of true random effects configuration
e.
Mean Square Errors
Number of bins
f.
CPU time (sec)
Number of bins

Figure 3

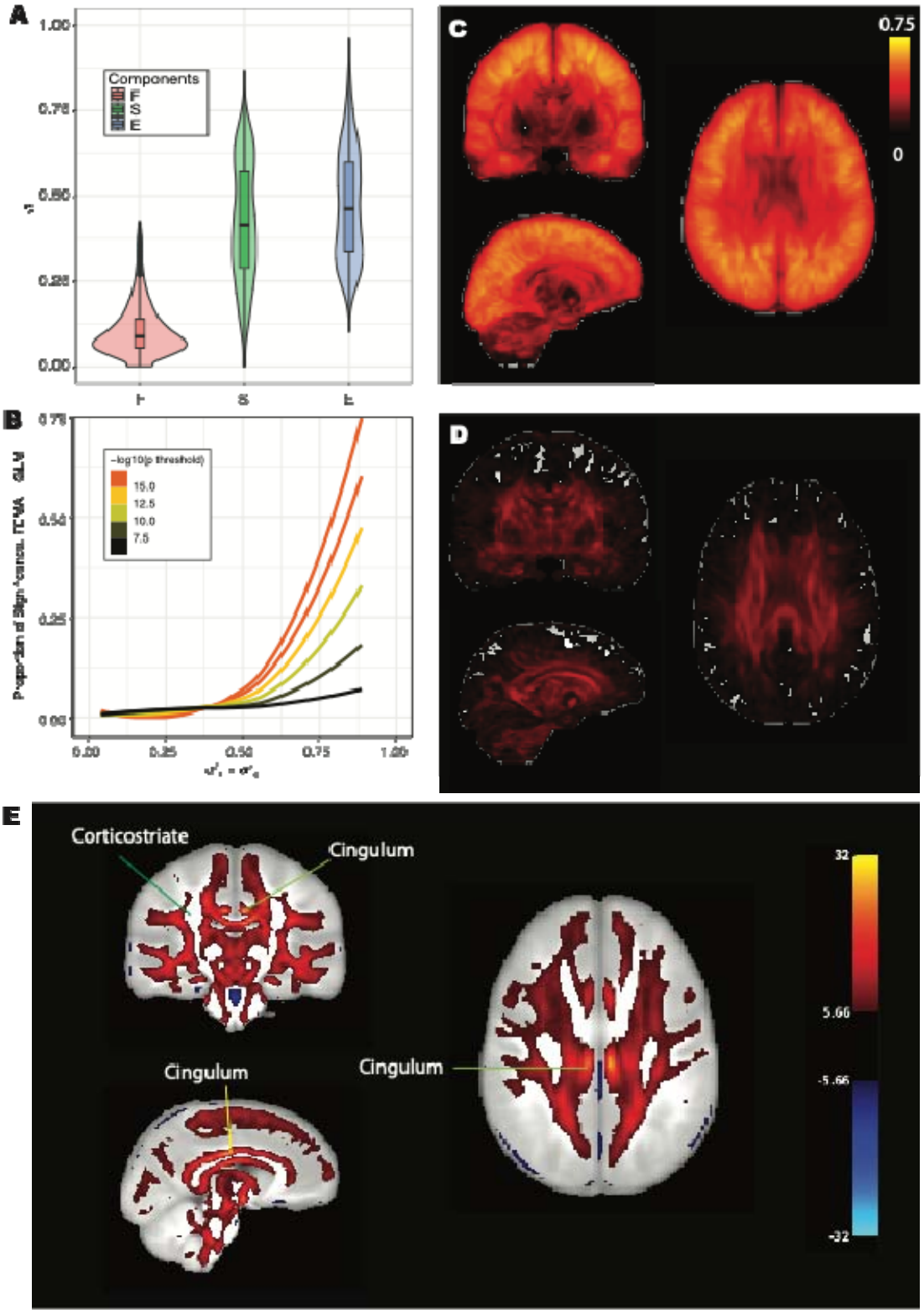
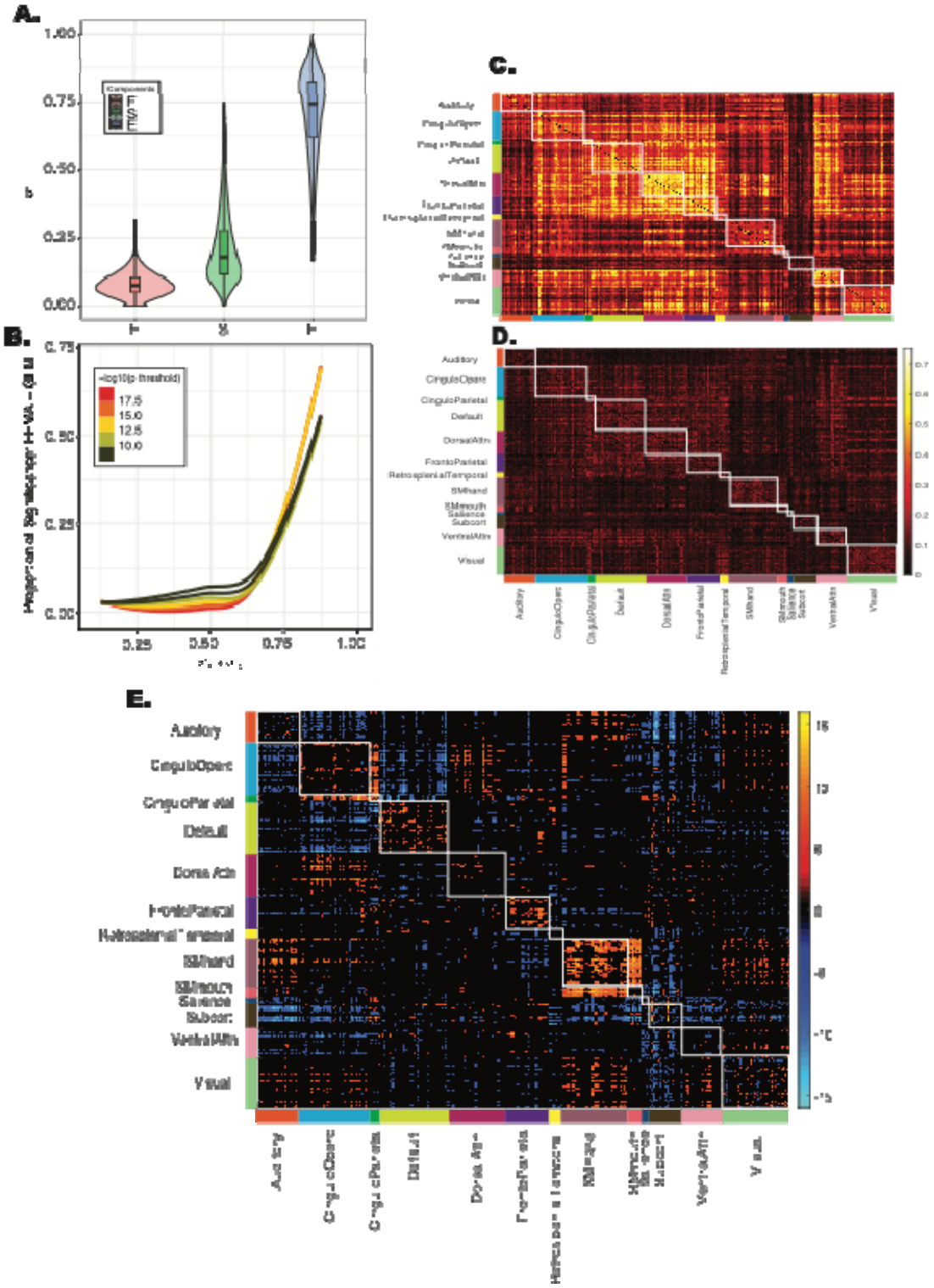

A
Components
F
S
E
1.00
0.75
0.50
0.25
0.00
F
S
E
B
-log10(p threshold)
15.0
12.5
10.0
7.5
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
C
0.75
0
D
E
Corticostriate
Cingulum
Cingulum
Cingulum
32
5.66
-5.66
-32

Figure 4


A.
B.
C.
D.
E.
Auditory
CingulOpere
CingulOParietal
Default
DorsalAttn
FrontoParietal
RetrosplenialTemporal
SMhand
SMmouth
Salience
Subcort
VentralAttn
Visual

References

Baria, A.T., Baliki, M.N., Parrish, T., Apkarian, A.V., 2011. Anatomical and functional assemblies of brain BOLD oscillations. J Neurosci 31, 7910-7919.

Beck, D., de Lange, A.G., Maximov, I.I., Richard, G., Andreassen, O.A., Nordvik, J.E., Westlye, L.T., 2021. White matter microstructure across the adult lifespan: A mixed longitudinal and cross-sectional study using advanced diffusion models and brain-age prediction. Neuroimage 224, 117441.

Bernal-Rusiel, J.L., Greve, D.N., Reuter, M., Fischl, B., Sabuncu, M.R., Initiative, A.s.D.N., 2013. Statistical analysis of longitudinal neuroimage data with Linear Mixed Effects models. Neuroimage 66, 249-260.

Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarco, D.V., Garavan, H., Orr, C.A., Wager, T.D., Banich, M.T., Speer, N.K., Sutherland, M.T., Riedel, M.C., Dick, A.S., Bjork, J.M., Thomas, K.M., Chaarani, B., Mejia, M.H., Hagler, D.J., Jr., Daniela Cornejo, M., Sicat, C.S., Harms, M.P., Dosenbach, N.U.F., Rosenberg, M., Earl, E., Bartsch, H., Watts, R., Polimeni, J.R., Kuperman, J.M., Fair, D.A., Dale, A.M., Workgroup, A.I.A., 2018. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Dev Cogn Neurosci 32, 43-54.

Chen, G., Saad, Z.S., Britton, J.C., Pine, D.S., Cox, R.W., 2013. Linear mixed-effects modeling approach to FMRI group analysis. Neuroimage 73, 176-190.

Chen, G., Taylor, P.A., Shin, Y.W., Reynolds, R.C., Cox, R.W., 2017. Untangling the relatedness among correlations, Part II: Inter-subject correlation group analysis through linear mixed-effects modeling. Neuroimage 147, 825-840.

Dick, A.S., Lopez, D.A., Watts, A.L., Heeringa, S., Reuter, C., Bartsch, H., Fan, C.C., Kennedy, D.N., Palmer, C., Marshall, A., Haist, F., Hawes, S., Nichols, T.E., Barch, D.M., Jernigan, T.L., Garavan, H., Grant, S., Pariyadath, V., Hoffman, E., Neale, M., Stuart, E.A., Paulus, M.P., Sher, K.J., Thompson, W.K., 2021.

Meaningful associations in the adolescent brain cognitive development study. Neuroimage 239, 118262.

Ge, T., Reuter, M., Winkler, A.M., Holmes, A.J., Lee, P.H., Tirrell, L.S., Roffman, J.L., Buckner, R.L., Smoller, J.W., Sabuncu, M.R., 2016. Multidimensional heritability analysis of neuroanatomical shape. Nat Commun 7, 13291.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2016. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cereb Cortex 26, 288-303.

Gordon, E.M., Laumann, T.O., Marek, S., Raut, R.V., Gratton, C., Newbold, D.J., Greene, D.J., Coalson, R.S., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Dosenbach, N.U.F., Nelson, S.M., 2020. Default-mode network streams for coupling to language and control systems. Proc Natl Acad Sci U S A 117, 17308-17319.

Guillaume, B., Hua, X., Thompson, P.M., Waldorp, L., Nichols, T.E., Initiative, A.s.D.N., 2014. Fast and accurate modelling of longitudinal and repeated measures neuroimaging data. Neuroimage 94, 287-302.

Haseman, J.K., Elston, R.C., 1972. The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2, 3-19.

Smith, S.M., Nichols, T.E., 2018. Statistical Challenges in "Big Data" Human Neuroimaging. Neuron 97, 263-268.