

Research Meeting

Identifying an unexpected but robust racial confound in functional connectivity and update on dictionary learning

Anton Orlichenko

March 2023



Outline

1 Robustness of Race Confound

- Background and Objective
- BSNIP Dataset
- Experiments
- Results

2 Dictionary Learning

- Iterative Dictionary
- Subgroup Selection for SZ Identifies Race Confound

3 Future Goals



Robustness of Race Confound



Background On Previous Race Confound Work

- Previous studies have shown that ML models may sometimes trivially detect race in medical images¹
- Work has shown that factors external to medical images such as race, socioeconomic status, or stress are important in ML models²
- In general racial bias can make its way into healthcare algorithms³

¹Gichoya et al. Lancet 2022 DOI: 10.1016/S2589-7500(22)00063-2

²Pierson et al. Nature Medicine 2021 <https://doi.org/10.1038/s41591-020-01192-7>

³Obermeyer et al. Science 2019 DOI: 10.1126/science.aax2342



Objective

Transfer of race-FC correlation

- We have shown that race is one of the most prominent FC signals in the PNC dataset.
- Is this signal found in other datasets, and does it take the same form?



Validation Dataset

Bipolar and Schizophrenia Network for Intermediate Phenotypes (BSNIP)⁴

- 933 patients (schizophrenia, bipolar, schizoaffective), 1059 relatives, and 459 normal controls
- 6 different study sites
- Used 387 African Americans (AA) and 778 Caucasians (CA) for whom we had fMRI scans
- Same Power template parcellation but different preprocessing



⁴ Tamminga et al 2014 doi: 10.1093/schbul/sbt179

Experiments

- **PNC → BSNIP** Trained model on PNC dataset, evaluated race prediction accuracy on BSNIP
 - Since AA and EA demographics were nearly balanced, only stratification was used
- **BSNIP → PNC** Trained on BSNIP, evaluated on PNC
 - Trained on a balanced 664-subject dataset (holdout used for within-BSNIP evaluation)
- Simple logistic regression model with $C=1000$
- 20 bootstrapping repetitions



Results

Average 68% accuracy PNC→BSNIP, 66% accuracy BSNIP→PNC

Trained on PNC		Trained on BSNIP	
Evaluation Group	Accuracy	Evaluation Group	Accuracy
PNC (all, n=733)	85±3%	BSNIP (all, n=1165)	79±4%
BSNIP AA (n=387)	76±5%	PNC EA (n=407)	90±3%
BSNIP CA (n=778)	64±5%	PNC AA (n=326)	38±7%

All predictions are better than the null model, except for identification of the AA group in the PNC dataset by a model trained on BSNIP.

20% holdout used for within-dataset validation (not full n)

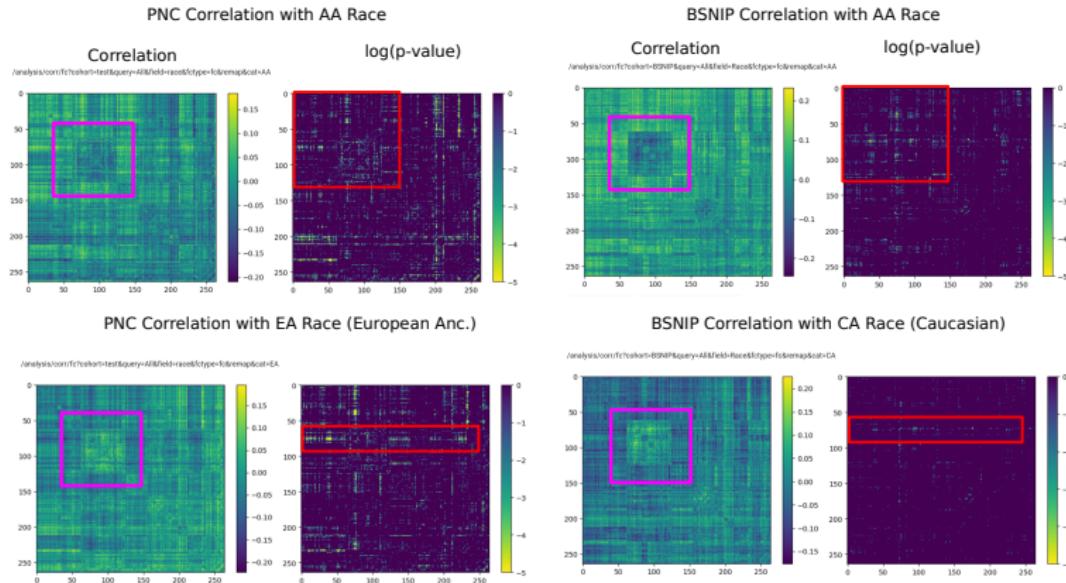


Discussion

- Note that PNC consisted of healthy adolescents 8-23 years old
- BSNIP consisted of young to older adults split into patients, patient relatives, and healthy subjects



Correlation



Correlation of FC with race in the PNC and BSNIP datasets. The overall FC in the Default Mode Network (DMN), highlighted in pink, seems to be highly predictive of race.



Dictionary Learning



Iterative Dictionary

Idea

Components for PCA are high-dimensional, information-dense, and may not be reproducible between populations

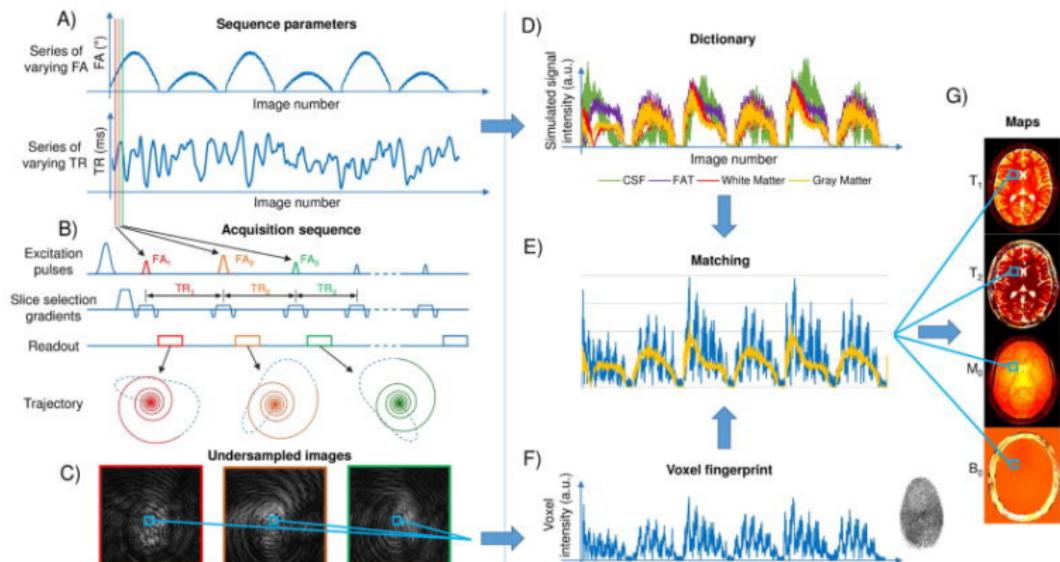
Goal

Create a robust and reproducible dictionary for diagnosis and correlation with phenotypes, demographics, or genomics



MRI Fingerprinting

One of the most successful examples of dictionary learning (although slightly unrelated to our work) is **MRI fingerprinting**⁵



⁵ Panda et al. 2017 doi: 10.1016/j.cobme.2017.11.001

Model

We estimate a basis of rank-1 or rank-2 FC matrices using an iterative algorithm based on gradient descent and SVD

Algorithm 1 An iterative algorithm to estimate low rank basis for FC

$$\mathbf{R}^{(0)} \leftarrow \mathbf{X}$$

$$i \leftarrow 0$$

while $i \neq 20$ **do**

$$\underset{\mathbf{A}, \mathbf{w}}{\text{minimize}} \quad \|\mathbf{R}^{(i)} - w_{\text{sub}}(\mathbf{A}\mathbf{A}^T)\|_F \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{A}^T\|_F = 1$$

$$\mathbf{A}^{(i)} \leftarrow \mathbf{A}$$

$$\mathbf{R}^{(i+1)} \leftarrow \mathbf{R}^{(i)} - w_{\text{sub}}(\mathbf{A}\mathbf{A}^T)$$

$$i \leftarrow i + 1$$

end while

Weights can then be estimated by any method iteratively

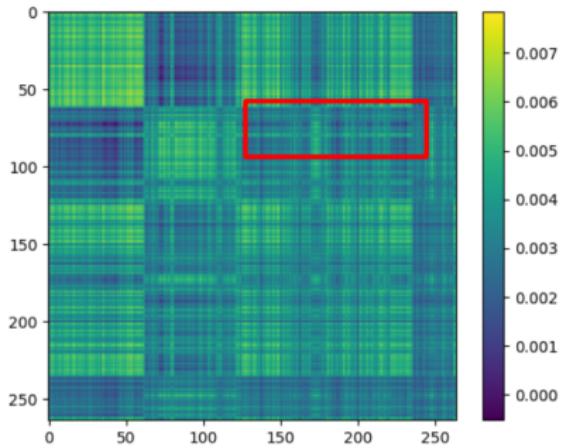


Low Rank Basis vs PCA

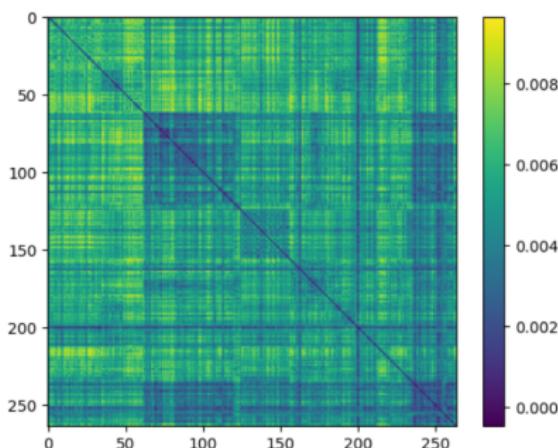
The Low Rank components are less noisy than PCA components. We hope they may better decompose network effects.

SZP + NC ($n=441$)

Iterative Rank-2 Basis
Component 0



PCA Component 0



Subgroup Selection in BSNIP

While investigating feature decomposition, we came across the following result

Preview of Findings

Identifying subgroups of schizophrenia patients (SZP) improves classification accuracy



Experiment

- 199 SZP subjects and 242 Normal Controls (NC)
- Same preprocessing, Power atlas
- KMeans clustering on SZP subjects with $N_{cluster} = 2$ identified
(Group 1: 130, Group 2: 69) subject subgroups
- Performed classification on Group 1, Group 2, or all SZP mixed with NC using Logistic Regression
- 20 bootstrap repetitions on balanced (69 SZP, 69 NC) samples of the cohort, 20% holdout



Results

Sample	Accuracy
All SZP+NC	71±8%
Group 1 SZP+NC	81±6%
Group 2 SZP+NC	75±9%

By segregating subgroups, we achieve better predictive accuracy in both subgroups

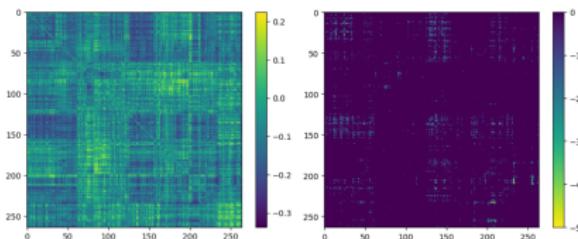


Subtype FC Correlation

Correlation of FC with SZ diagnosis in SZ versus normal controls.

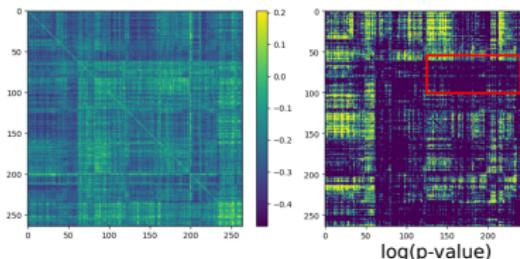
All SZP vs NC log(p-value)

/analysis/corr/fc?cohort=BSNIP&query=DXGROUP_1 == 'SZP' or DXGROUP_1 == 'NC' & field=DXGROUP_1&fc_type=fcl&remap&cat=SZP



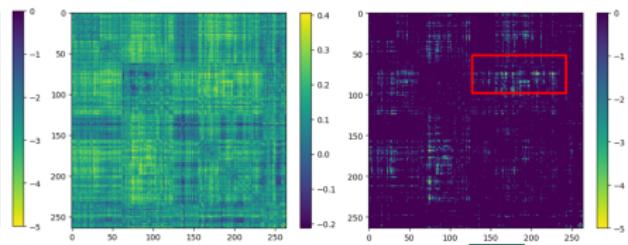
SZ Subtype 1 vs NC

/analysis/corr/fc?cohort=BSNIP&query=sz_subtype == '1' or DXGROUP_1 == 'NC' & field=DXGROUP_1&fc_type=fcl&remap&cat=SZP



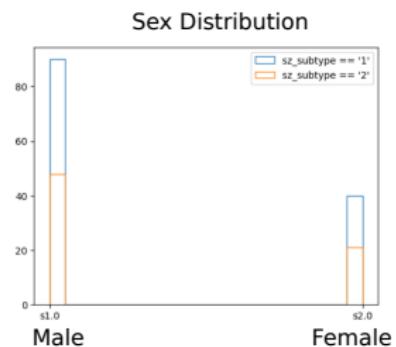
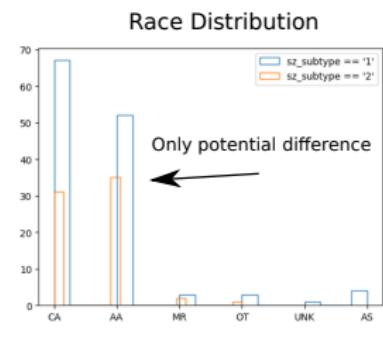
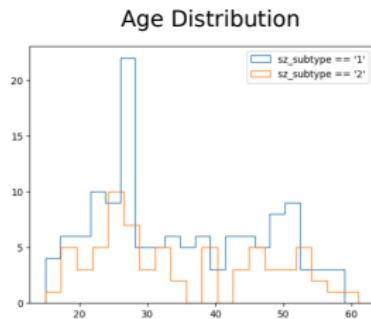
SZ Subtype 2 vs NC

/analysis/corr/fc?cohort=BSNIP&query=sz_subtype == '2' or DXGROUP_1 == 'NC' & field=DXGROUP_1&fc_type=fcl&remap&cat=SZP



Feature Correlation

Is there any explanation for subgroups?



No difference in medications or continuous psychiatric evaluation metrics

Future Goals



Future Goals

- Confirm validity of race signal in different study groups or patient populations
- Optimize iterative dictionary for subgroup selection
- Identify interaction between subgroup categories (e.g., age vs race, age vs sex, race vs SZ diagnosis, etc.)

