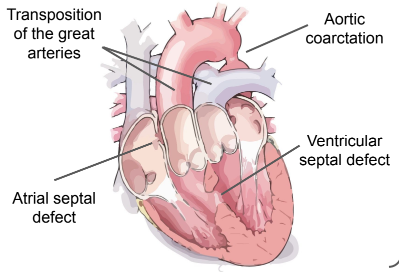
CHD Fetal Brain Analysis using Non-linear Transformations and Combined qMRI Features

Samantha A. Esparza-Esparza [Samantha.EsparzaEsparza@childrens.harvard.edu]
Milton O. Candela-Leal [Milton.CandelaLeal@childrens.harvard.edu]

Introduction

Congenital Heart Disease (CHD)

Examples



- Reduced oxygen supply
- Altered blood flow dynamics
- Neuroinflamation

[1]

Early detection relevance

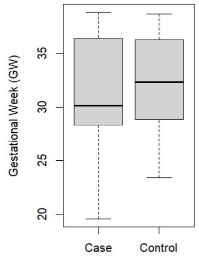
Enhanced Prenatal Counseling

Potential for discovery of **unique biomarkers** indicative of CHD impact on fetal brain

Non-linear Combined qMRI Features Generation and Selection

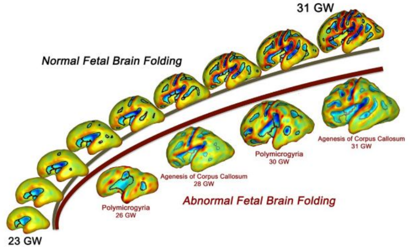
Cohort composed of:

- 96 CHD MRI (mean = 31.38 weeks, sd = 4.80)
- 62 TD MRI (mean = 32.38 weeks, sd = 4.17)



Age-adjusted qMRI features

As Gestational Weeks (GW) is a cofounding variable in fetal MRI



[2]

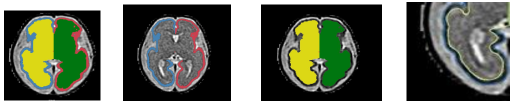
[1] Khalil et al ., *Ultrasound Obstet Gynecol.*, 2016
[2] Im., Advances in Magnetic Resonance Technology and Applications., 2021

Data pre-processing

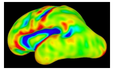
Subjects

96 Congenital Heart Disease (CHD)
62 Typically Developed (TD)

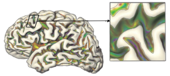
Base Features



Area | Volume CP | Volume In | Gyrification Index



Depth



Absolute Mean Curvature (aMC)

* Each feature has Left and Right hemisphere, and Average
18 Base Features

Gestational Weeks (GW) as a cofounding variable

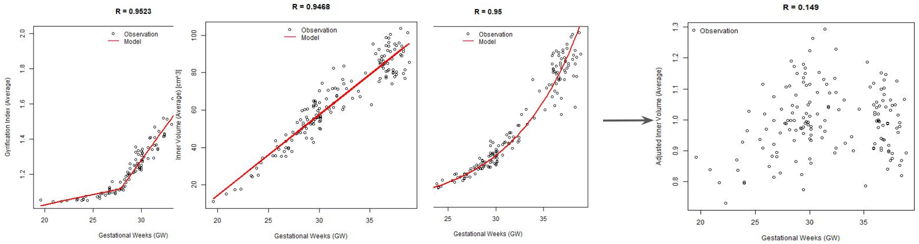
(Cofounding variable)
Gestational Weeks (GW)

(Independent Variable)
qMRI measurements

(Dependent Variable)
CHD

Covariance adjusted feature via age distribution

Leaky ReLU | Linear | Exponential



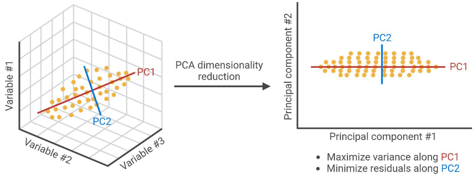
Non-adjusted (high GW correlation) | Adjusted (low GW correlation)

Min-max feature normalization

$$X_{std} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad [1]$$

PCA*

*For comparison purposes



Feature Generation & Selection

[1] Candela-Leal et al., *Appl. Sci.*, 2022
[2] Blanco-Rios and Candela-Leal et al., *Front. Hum. Neurosci.*, 2024

Feature Generation & Selection

Non-linear feature transformations [1]

Name	Quadratic	Cube	Inverse	Logarithm	Inter-hemisphere	Intra-hemisphere	Difference
Description	F^2	F^3	$\frac{1}{F+1}$	$ln\|F+1\|$	$\frac{F_L}{F_R+1} \,\|\|\, \frac{F_R}{F_L+1}$	$\frac{F_{VolCP_L}}{F_{VolIn_L}+1}, \frac{F_{VolIn_L}}{F_{VolCP_L}+1}, \frac{F_{VolCP_R}}{F_{VolIn_R}+1} \,\|\|\, \frac{F_{VolIn_R}}{F_{VolCP_R}}$	$F_L - F_R \,\|\|\, F_R - F_L$
Left/Right Hemisphere	×	×	×	×	×	×	×
Average	×	×	×	×			
Volume CP or In only						×	

+ 100 NL Features

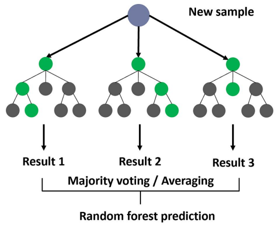
Combined Feature Creations

$Division \rightarrow 13,806$ [2]	$Multiplication \rightarrow 6,647$ [3]
Pairwise combinations with order	Non-repeating pairwise combinations with repetition
$(n \cdot (n-1))$	$\frac{(n \cdot (n-1))}{2}$

+ 20,453 Combined Features

= 20,571 Total Features

Random Forest (RF) Feature Selection [4]



Best 14 Features

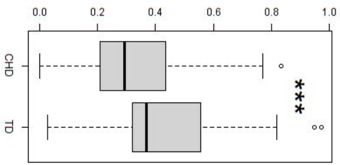
kNN model fitting [5]

[1] Aguilar-Herrera et al., *ML-DT Edu. Innovation Workshop*, 2021
[2] Ramírez-Moreno et al., *Int. J. Environ. Res. Public Health*, 2021
[3] Olivas-Martínez et al., *ML-DT Edu. Innovation Workshop*, 2021
[4] Candela-Leal et al., *IEOM-NA VI*, 2021
[5] Shon et al., *Int. J. Environ. Res. Public Health*, 2018

Results

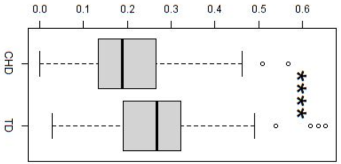
Examples from top 5 features from RRFI

$$(\ln|Area_R + 1|) \cdot \frac{GI_R}{GI_L+1}$$



Wilcoxon, p = 0.00028

$$\frac{(VolCP_R)^3}{aMC_R - aMC_L}$$



Wilcoxon, p = 0.00002

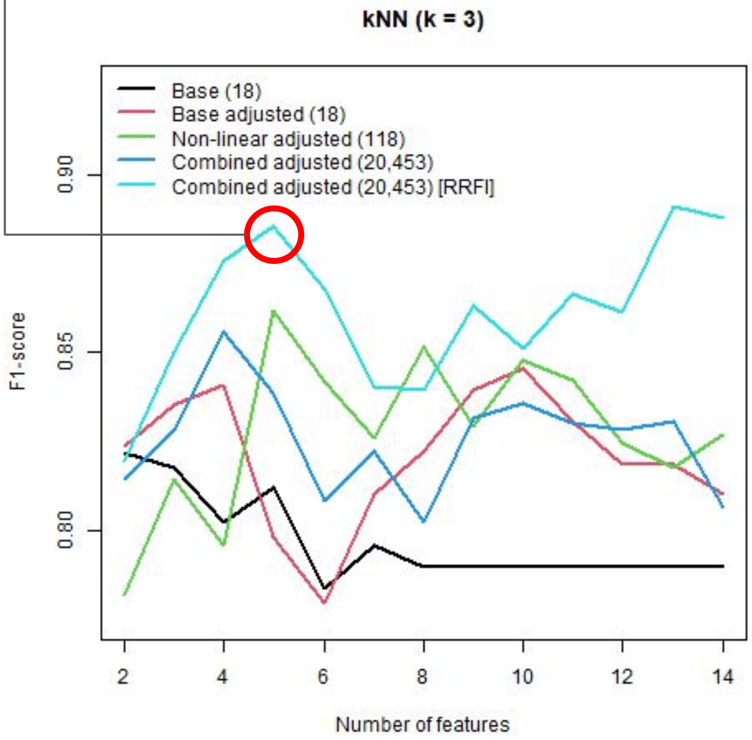
Algorithm 1 recursive_random_forest_imp(X, Y, n_f)

Input X source features.
Input Y target features.
Input n_f maximum number of best source features.
Require: $n_f \geq 1$

$m \leftarrow dim(X, 1)$
while $m > n_f$ **do**
  $I \leftarrow$ random_forest(X, Y).importance
  Sort I in descending order
  $I_{\text{best}} \leftarrow I[:(m - n_f)]$
  $X \leftarrow X[:, I_{\text{best}}]$
  $m \leftarrow dim(X, 1)$
end while

F1-score = Harmonic mean of Precision and Recall $= 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$ [1]

RRFI f1-score = 0.88 using top 5 features

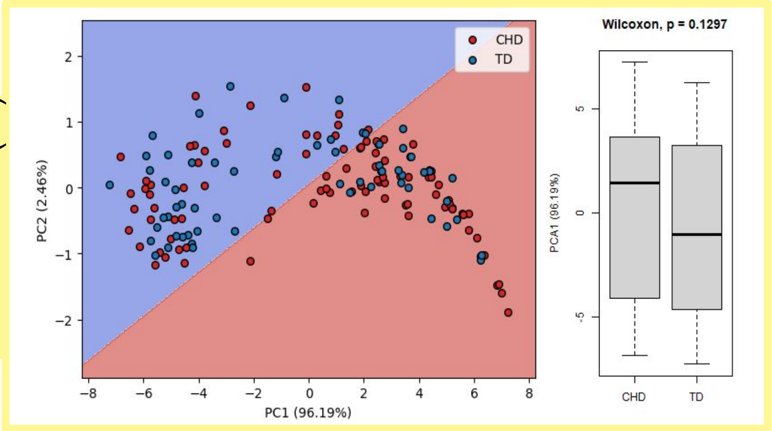


[1] Blanco-Rios and Candela-Leal et al., *Front. Hum. Neurosci.*, 2024

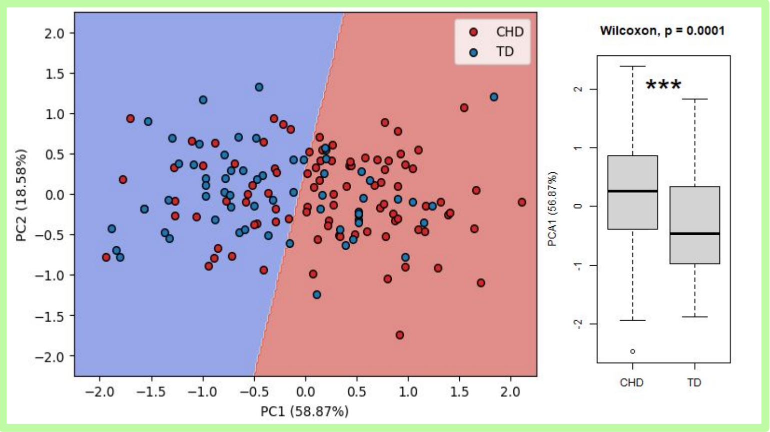
Discussion

Linear weighted SVM surface plot [1]

Base features (18)



Adjusted features (18)



Insights

- 0.88 f1-score using best 5 BoS features in kNN (k = 3)
- F1-score = 0.69, increased by 0.07 when adjusting features to GW
- Statistically significant PC1 at differentiating between CHD and TD

Next steps

- Including data from other control protocols (placenta, normative)
- Further implement cross-validations and stratified training/test split
- Use the framework in other abnormal brain conditions (ventriculomegaly, cerebral palsy)

[1] Candela-Leal et al., *IEEE-EMBS BSN*, 2023