

XGBOOSTING THE SIGNAL

Machine Learning Approach for Informant-Sensitive Predictions of OCD Symptoms in Children Based on Brain Morphology

Student

Margrete Heimvik

Supervisors

Øystein Sørensen
Ina Drabløs



LCBC
LIVSLØPSENDRINGER
i hjerne og kognisjon



**UNIVERSITY
OF OSLO**

AGENDA

Background

Understanding OCD in Adolescence: Neurobiology, Informant Discrepancies, and Machine Learning Approaches

The Present Study

Research Questions

Methods

Sample Characteristics and Modeling Strategy

Results

Key Findings and Predictive Performance

Discussion

Interpreting the Results and Future Directions for Improvement

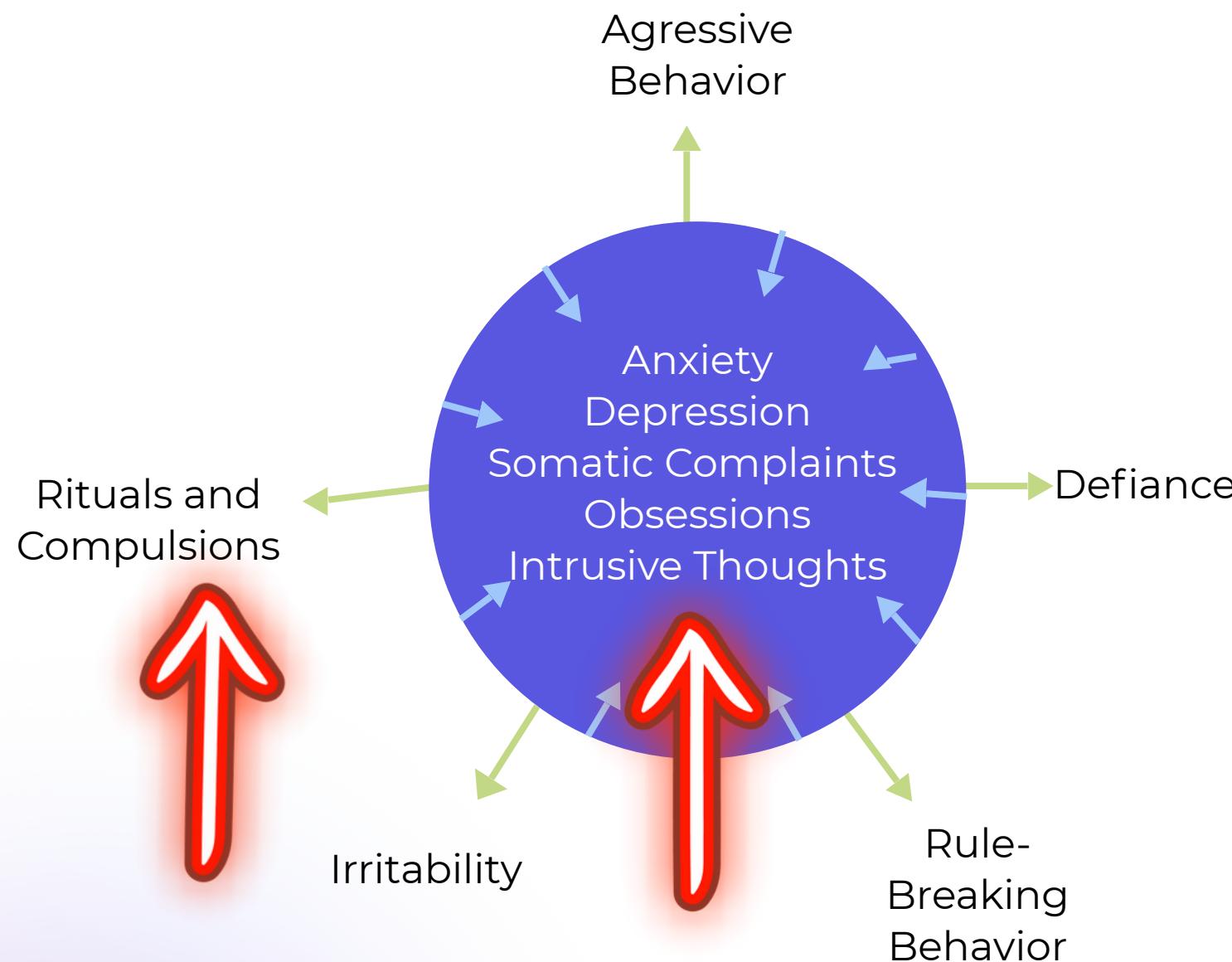
Obsessive-Compulsive Disorder



- OCD is a chronic neuropsychiatric disorder, affecting 2–3% of the global population and 1–4% of youth.
- Onset often **occurs in childhood or adolescence**, a critical period of brain development.
- Adolescent symptoms are heterogeneous, involving both **internalizing and externalizing** symptoms.

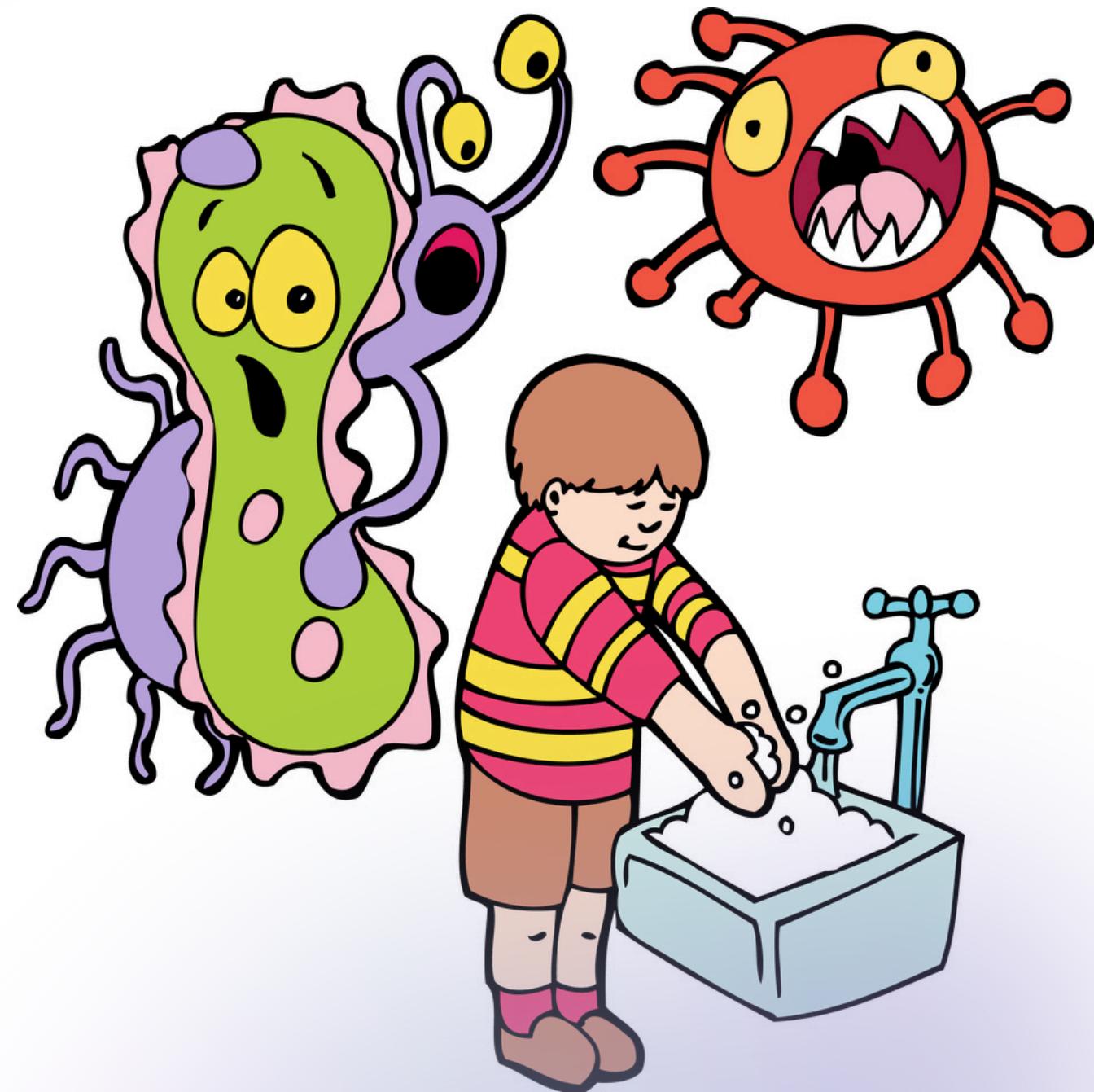
de Mathis (2013), Nazeer (2020), Casey (2008), Saad (2017)

Symptom Complexity & Diagnostic Challenges



- **Obsessions** are intrusive, anxiety-provoking thoughts, urges, or images = Internalizing
- **Compulsions** are repetitive behaviors or mental acts performed to reduce anxiety or prevent feared outcomes = Externalizing
- Considerations: High rates of comorbidity, **informant discrepancies**

Informant Discrepancies



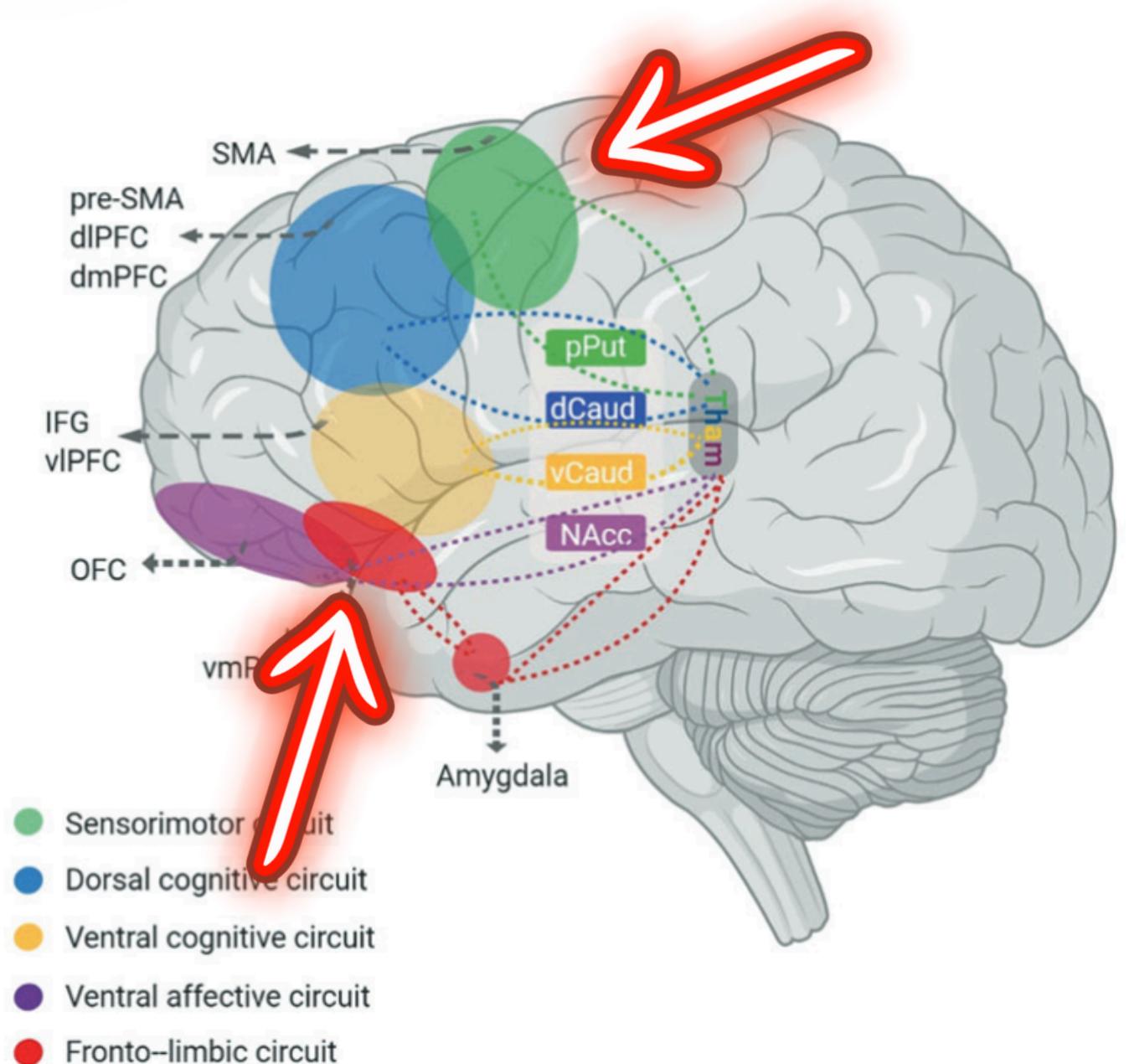
- Children and parents often **report symptoms differently**
- Clinical decisions often rely on adult informants

Neurobiological Findings

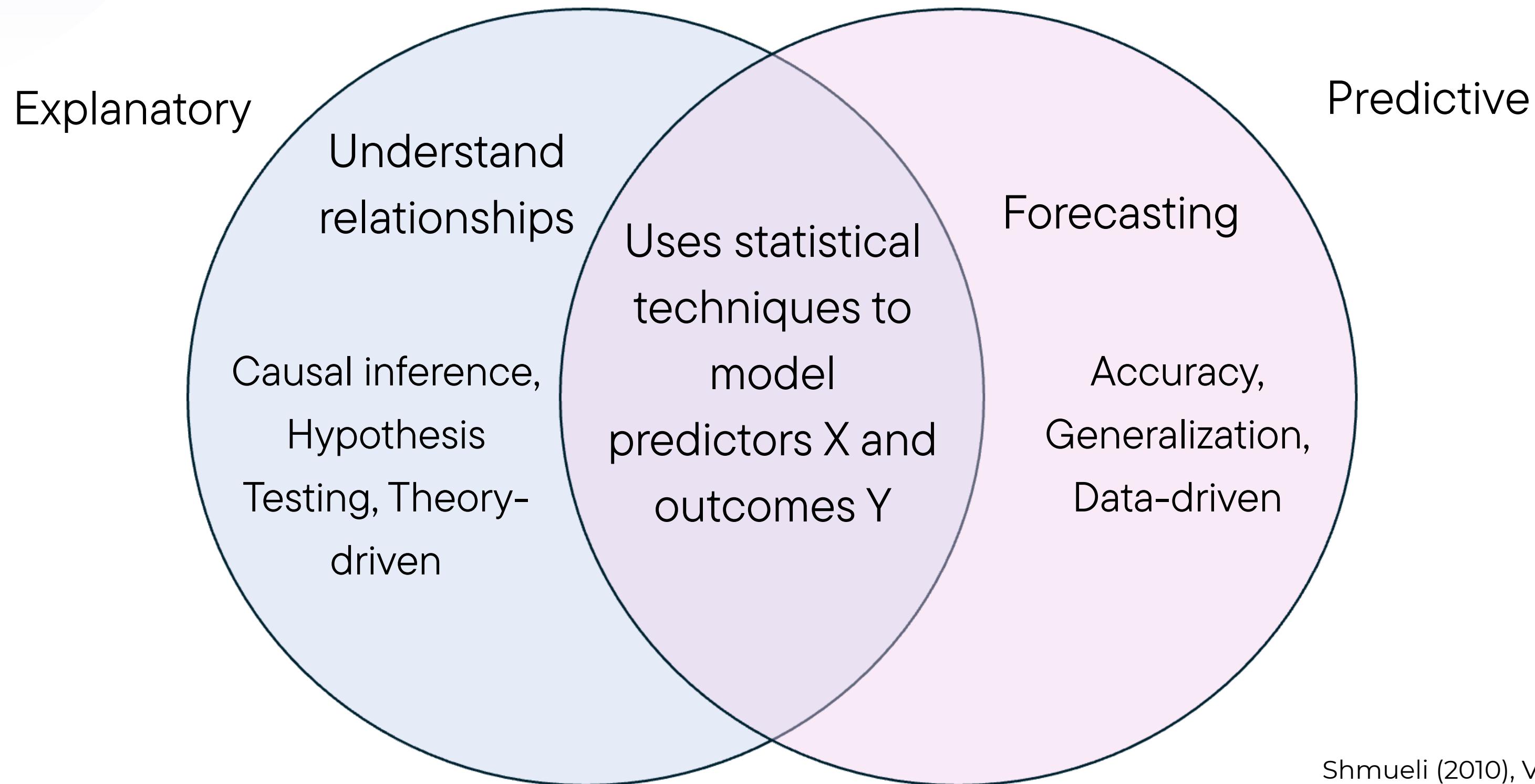
- Childhood = a key period for brain maturation
- Involves dysfunction in the CSTC circuit
- Linked to **impaired behavior regulation and motor control**
- Alterations mainly seen in **unmedicated patients**

Neurobiological Findings

- Brain **structural abnormalities** in adolescents with OCD
- Suggests a **developmental shift** in both brain structure and symptom expression:
 - Youth: sensory-driven compulsions
 - Adults: fear-based, rigid behaviors
- Findings remain **mixed and inconclusive**



Statistical Modelling



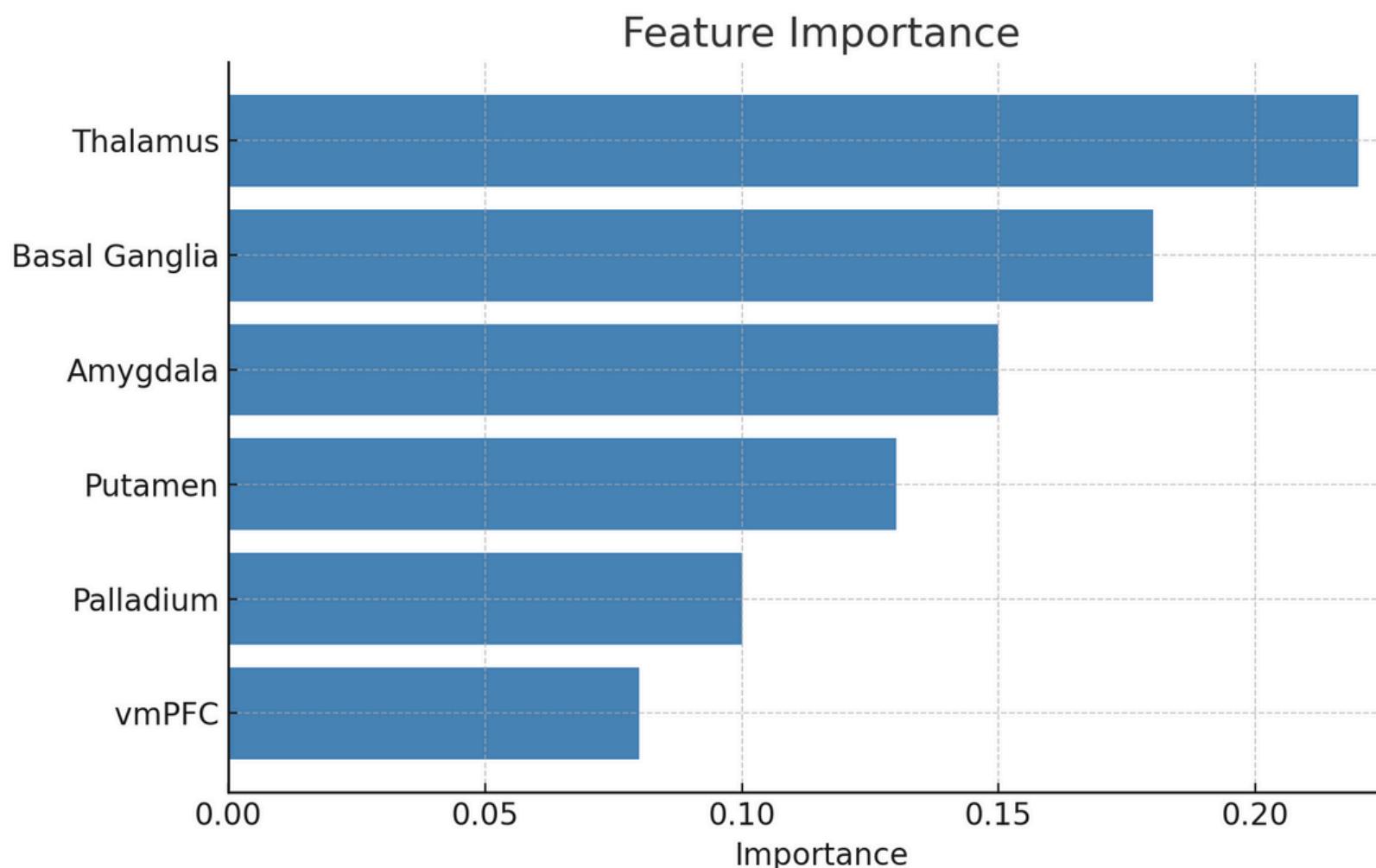
Shmueli (2010), Vapnik (1999)

Statistical Modelling

Predictive

- Predictive modelling is **function estimation** problem.
- The goal is to estimate a function that maps inputs X to an output Y in a way that **minimizes prediction error**.

Why Extreme Gradient Boosting?



- Learns complex, nonlinear relationships without strong model assumptions
- **Strong predictive performance** across research and industry applications
- **Interpretable** using tools like feature importance and partial dependence plots

THE PRESENT STUDY

Can structural brain features predict the severity of OCD-related internalizing symptoms in children?

Does predictive accuracy differ between child-reported and parent-reported internalizing symptoms?

METHODS



Methods

Materials

ABCD Study

Brain Regions of Interest (ROI)

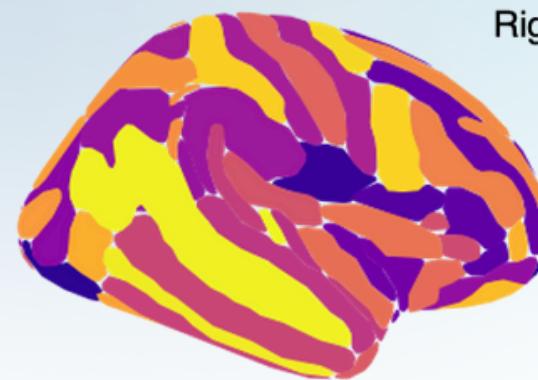
- Destrieux Atlas: 148 cortical regions
- ASEG Atlas: 46 subcortical regions

Internalizing Symptoms

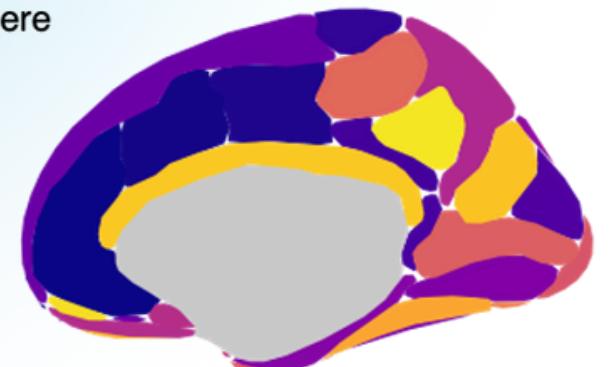
- ASEBA Instruments: CBCL (parent-report), BPM (self-report)

Psychiatric Diagnosis

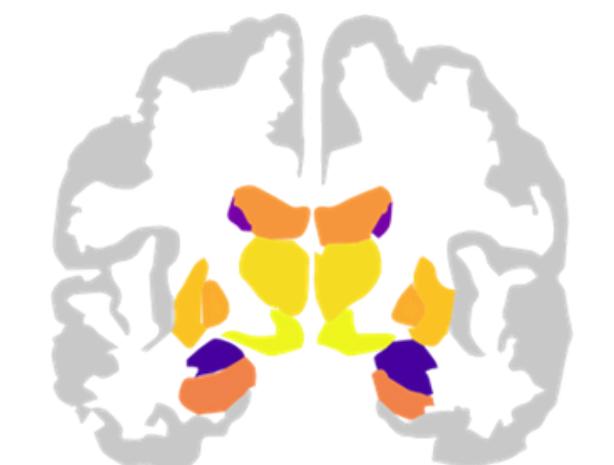
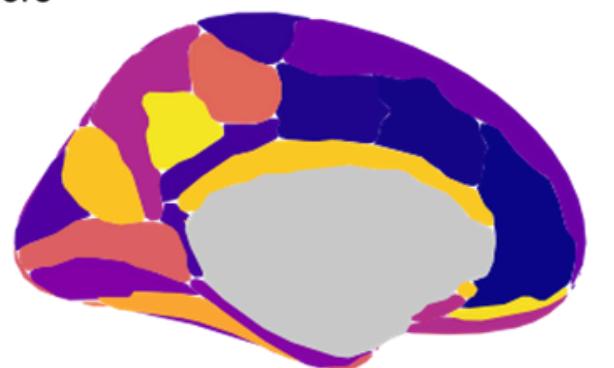
- KSADS-COMP: Computerized DSM-5 diagnostic tool



Right Hemisphere

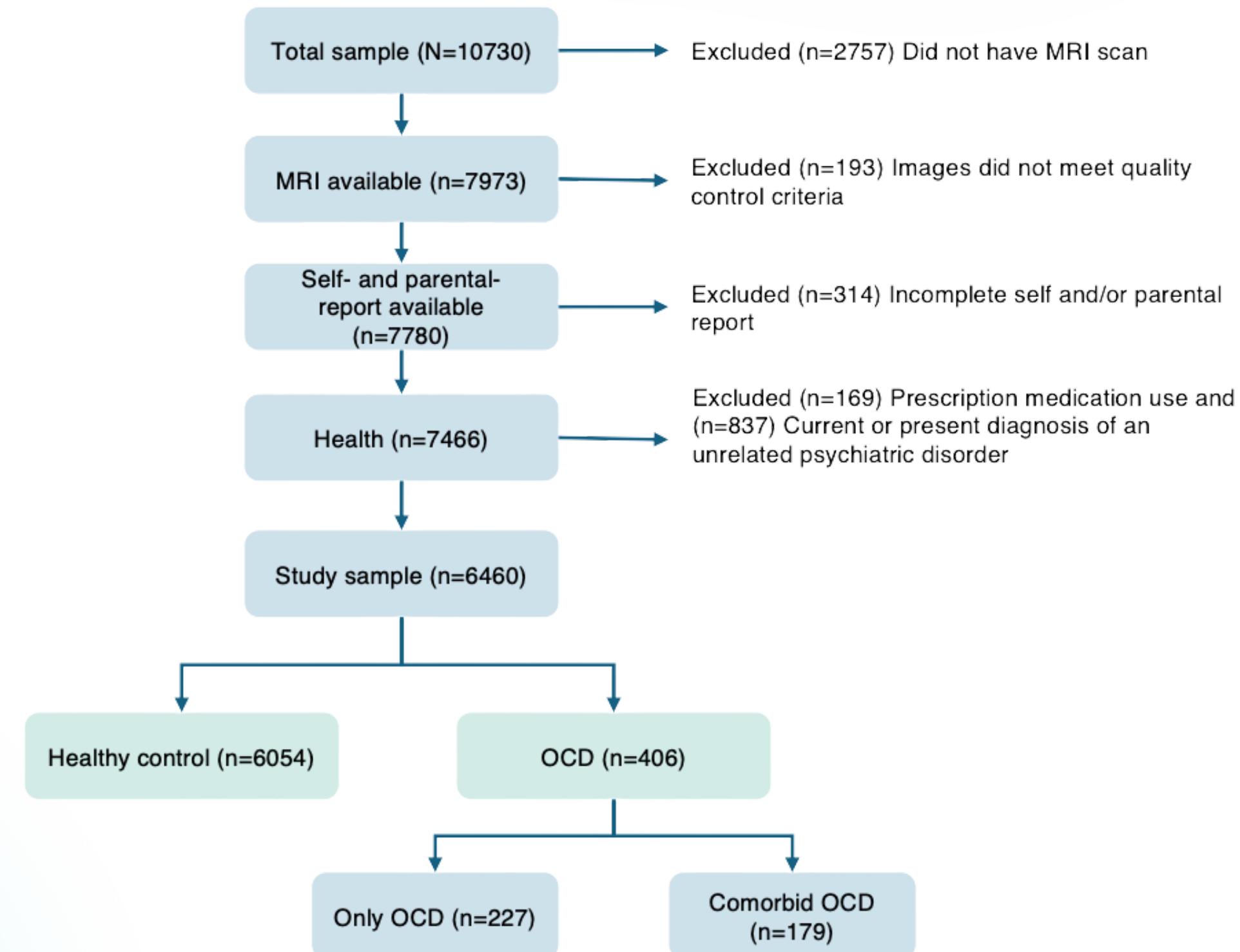


Left Hemisphere



Methods

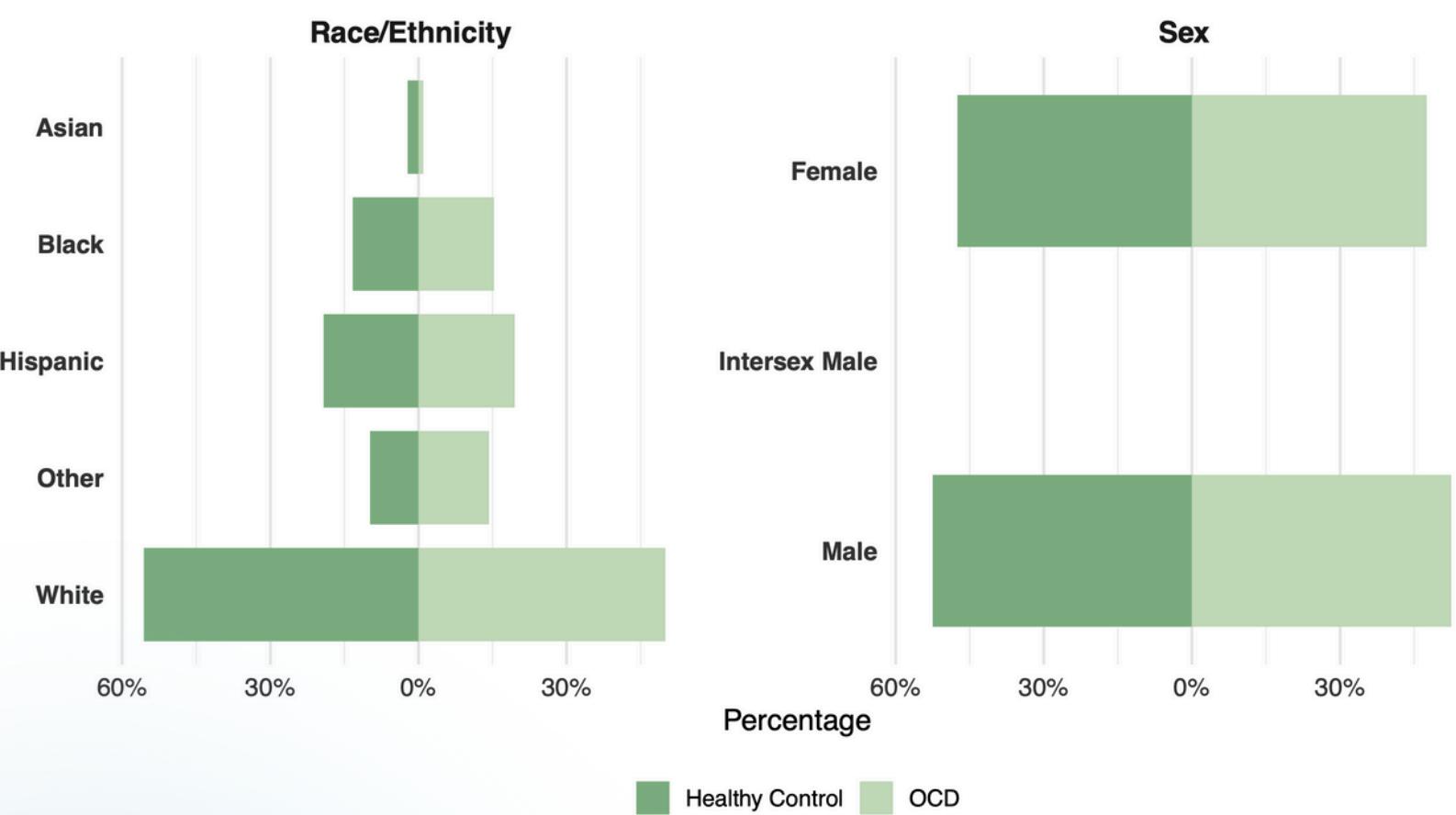
Sample Demographics



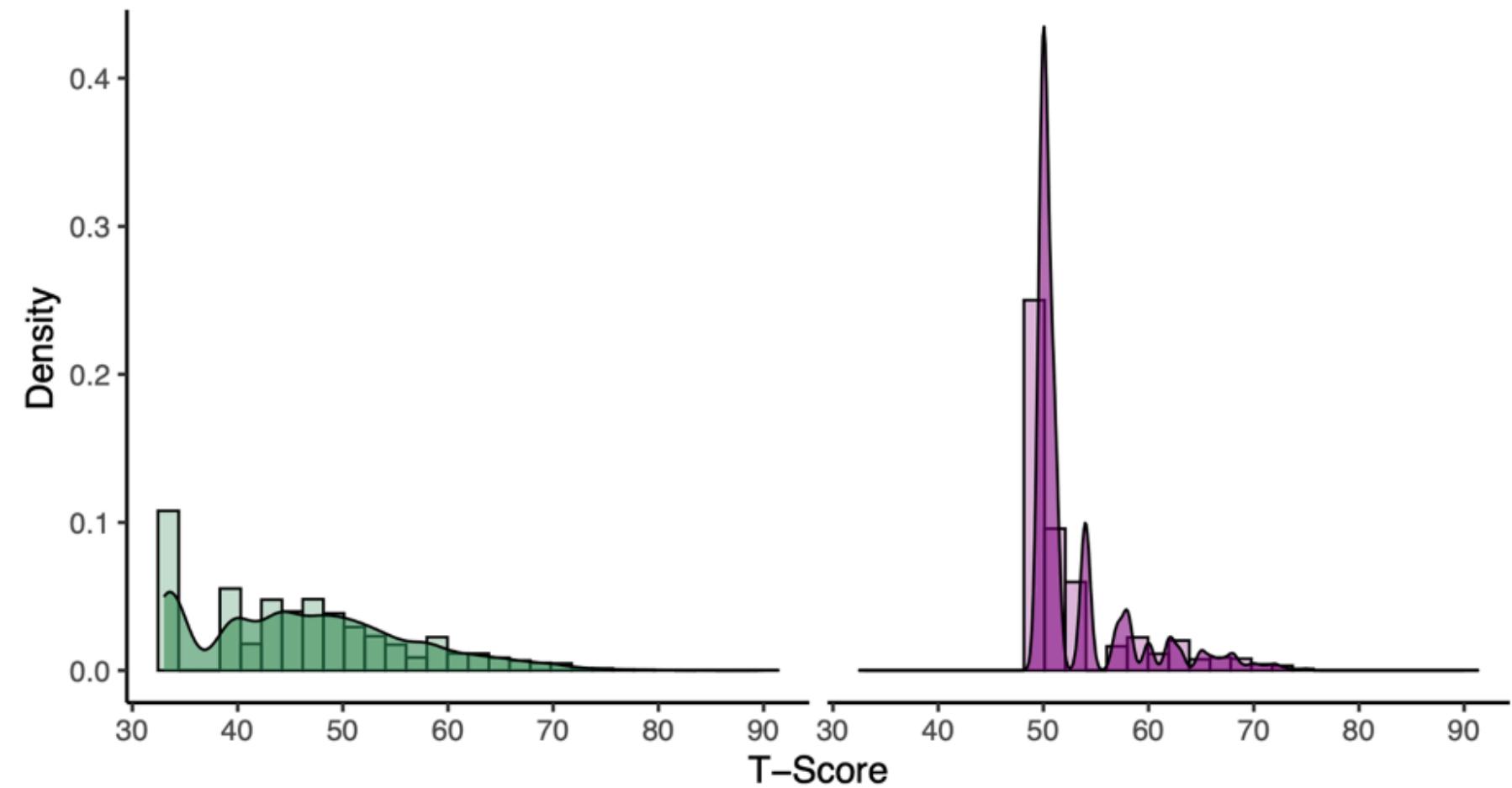
Methods

Sample Demographics

Demographics



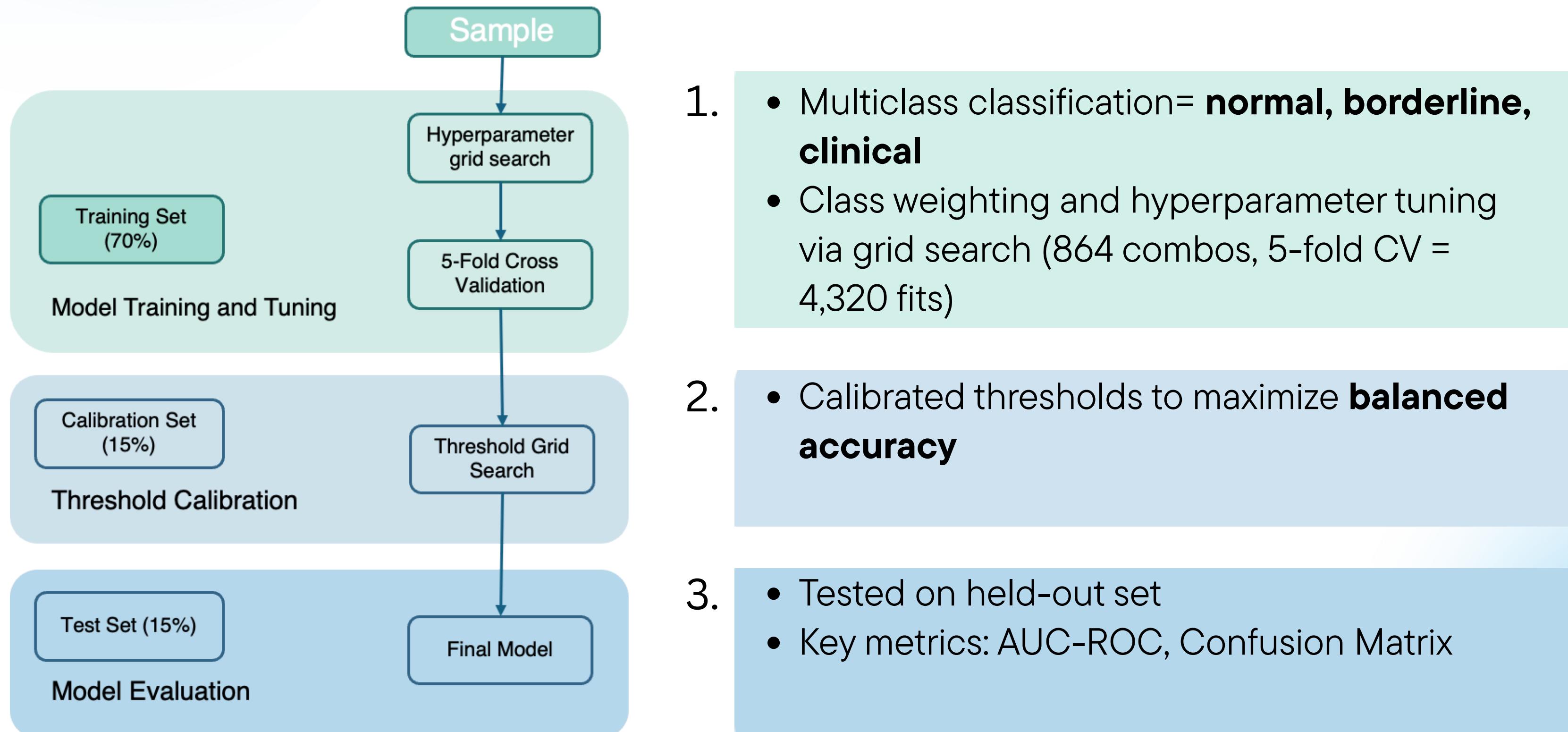
Internalizing Score Distribution



Parent-
Reported

Child-
Reported

Modelling Approach



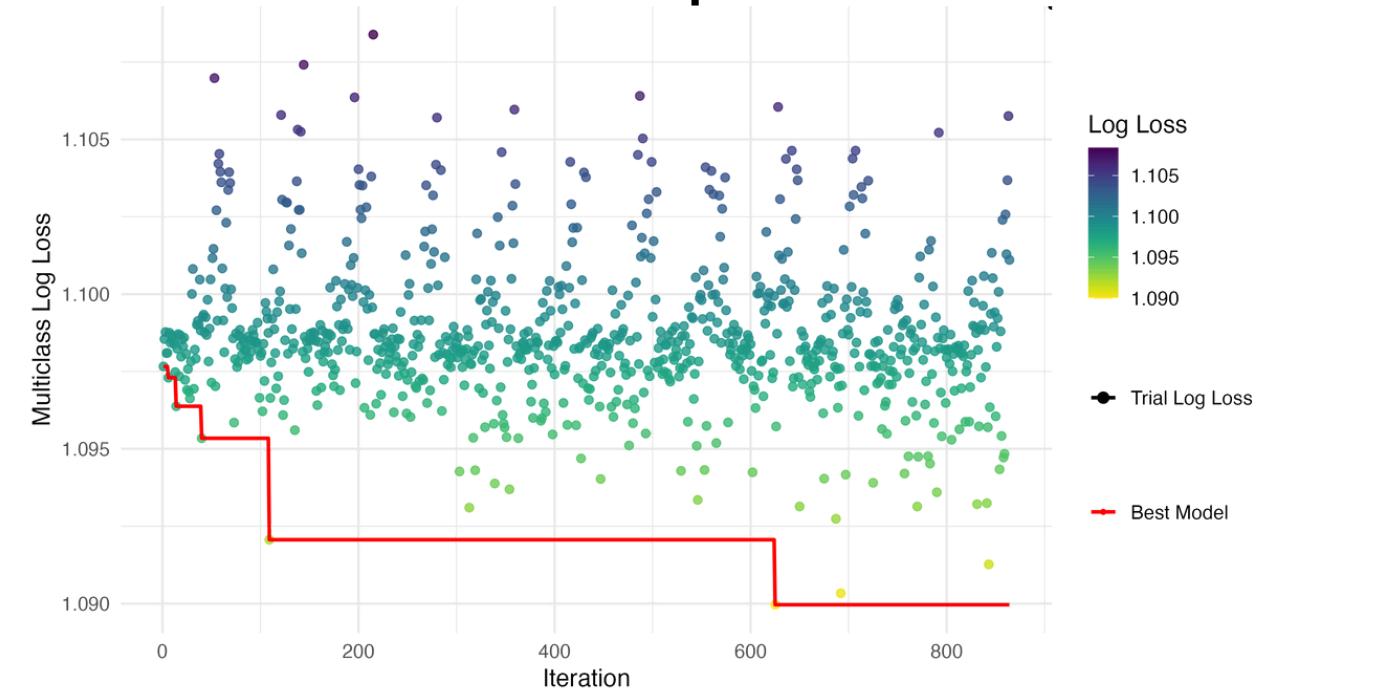
RESULTS



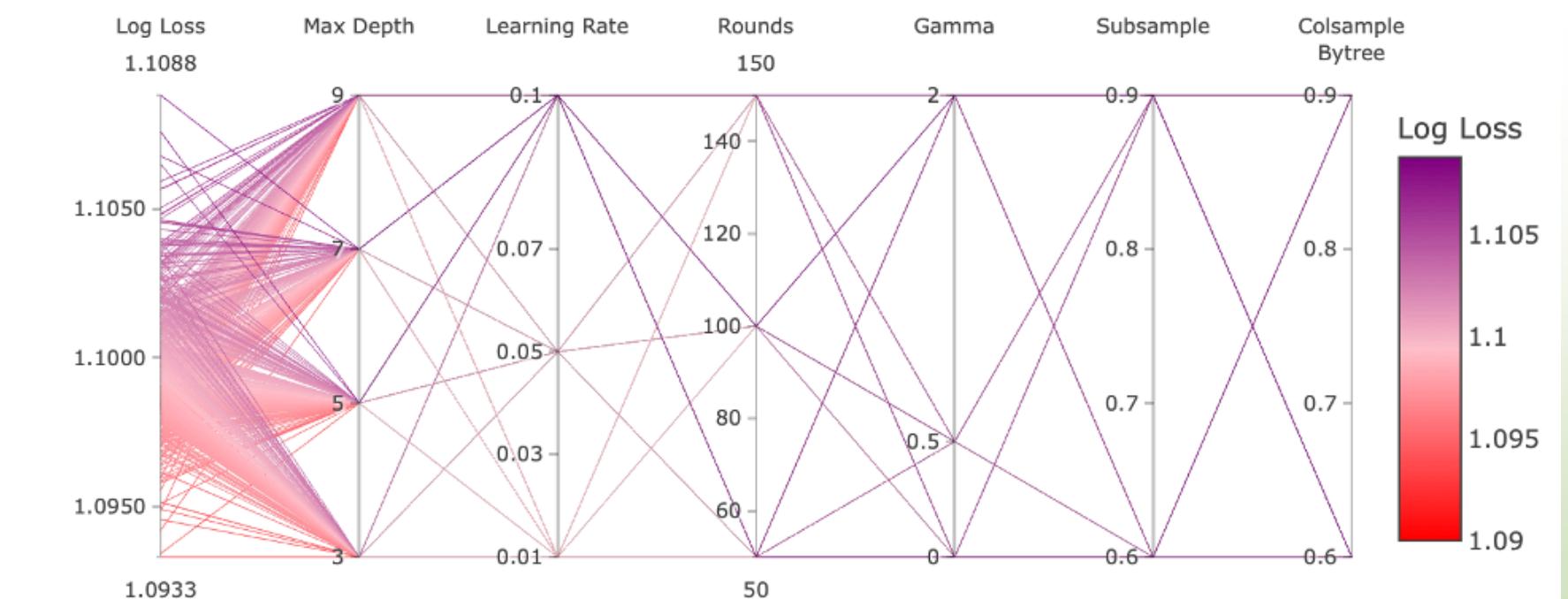
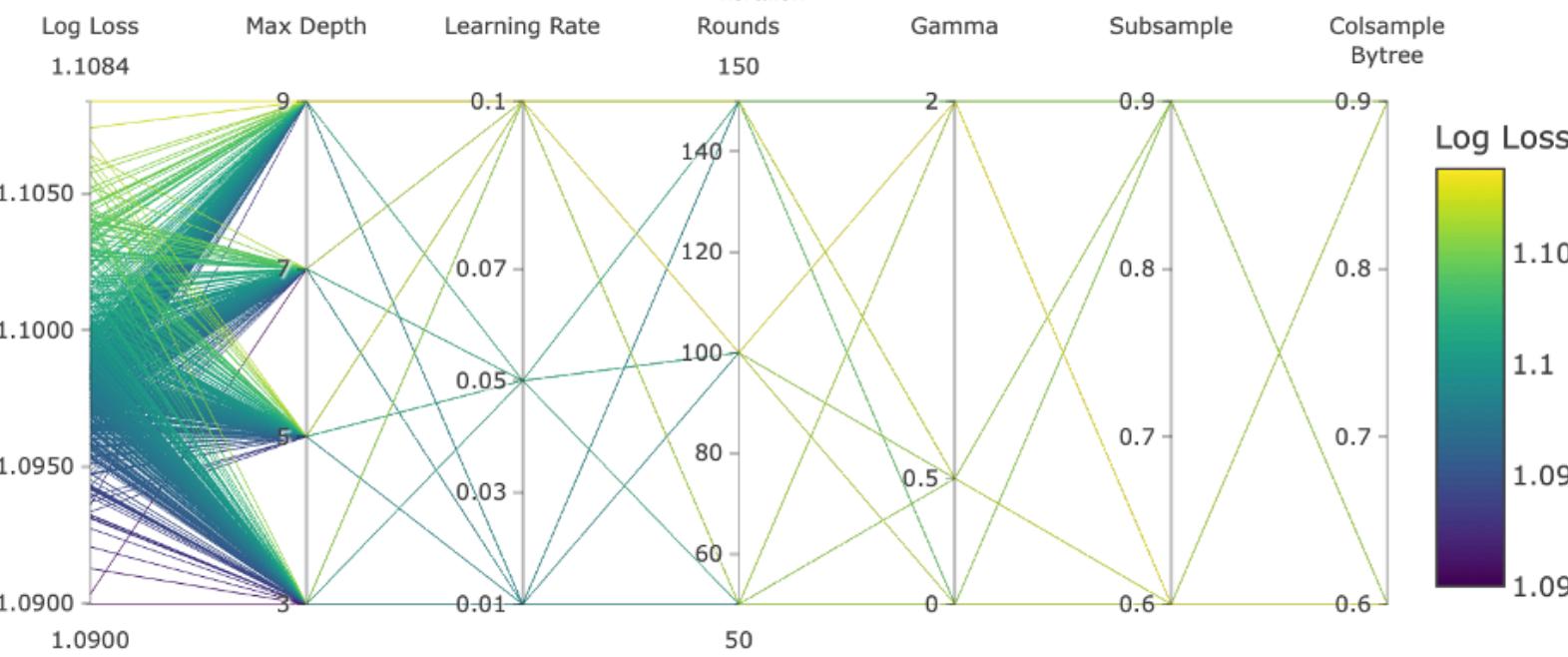
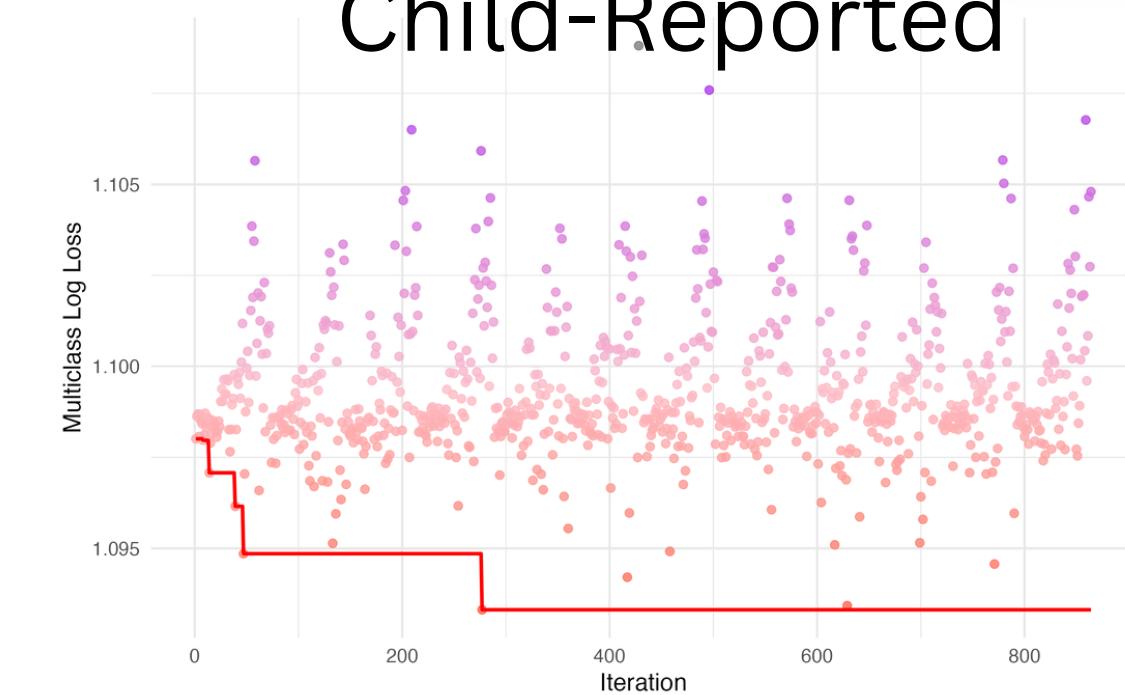
Results

Model Training and Tuning

Parent-Reported



Child-Reported



Results

Performance Metrics

	Parent-Reported		Child-Reported	
Model Configuration	Tuned + Argmax	Tuned + Threshold-Cal.	Tuned + Argmax	Tuned + Threshold-Cal.
Overall Performance				
Accuracy	0.95	0.68	0.93	0.92
95% CI	0.96	0.64–0.71	0.95	0.91–0.94
Cohen's Kappa	-0.0	-0.03	0.0	-0.02
Class Balance				
Sensitivity (Clinical)	0.0	0.083	0.0	0.0
Specificity (Clinical)	1.0	0.86	1.0	1.0
Class-Level Discrimination				
F1 Score (Clinical)	0.0	0.03	0.0	0.0

- High accuracy is **misleading**
- None of the models reliably detected clinical cases, indicating failure to identify at-risk individuals.

Results

Performance Metrics

	Parent-Reported		Child-Reported	
Model Configuration	Tuned + Argmax	Tuned + Threshold-Cal.	Tuned + Argmax	Tuned + Threshold-Cal.
Overall Performance				
Accuracy	0.95	0.68	0.93	0.92
95% CI	0.93–0.96	0.64–0.71	0.92–0.95	0.91–0.94
Cohen's Kappa	-0.00	-0.03	0.0	-0.02
Class Balance				
Sensitivity (Clinical)	0.0	0.083	0.0	0.0
Specificity (Clinical)	1.0	1.0	1.0	1.0
Class-Level Discrimination				
F1 Score (Clinical)	0.0	0.05	0.0	0.0

- High accuracy is **misleading**
- None of the models reliably detected clinical cases, indicating failure to identify at-risk individuals.

Results

Performance Metrics

	Parent-Reported		Child-Reported	
Model Configuration	Tuned + Argmax	Tuned + Threshold-Cal.	Tuned + Argmax	Tuned + Threshold-Cal.
Overall Performance				
Accuracy	0.95	0.68	0.93	0.92
95% CI	0.93–0.96	0.64–0.71	0.92–0.95	0.91–0.94
Cohen's Kappa	-0.00	-0.03	0.0	-0.02
Class Balance				
Sensitivity (Clinical)	0.0	0.083	0.0	0.0
Specificity (Clinical)	1.0	0.86	1.0	1.0
Class-Level Discrimination				
F1 Score (Clinical)	0.0	0.0	0.0	0.0

- High accuracy is **misleading**
- None of the models reliably detected clinical cases, indicating failure to identify at-risk individuals.

Results

Performance Metrics

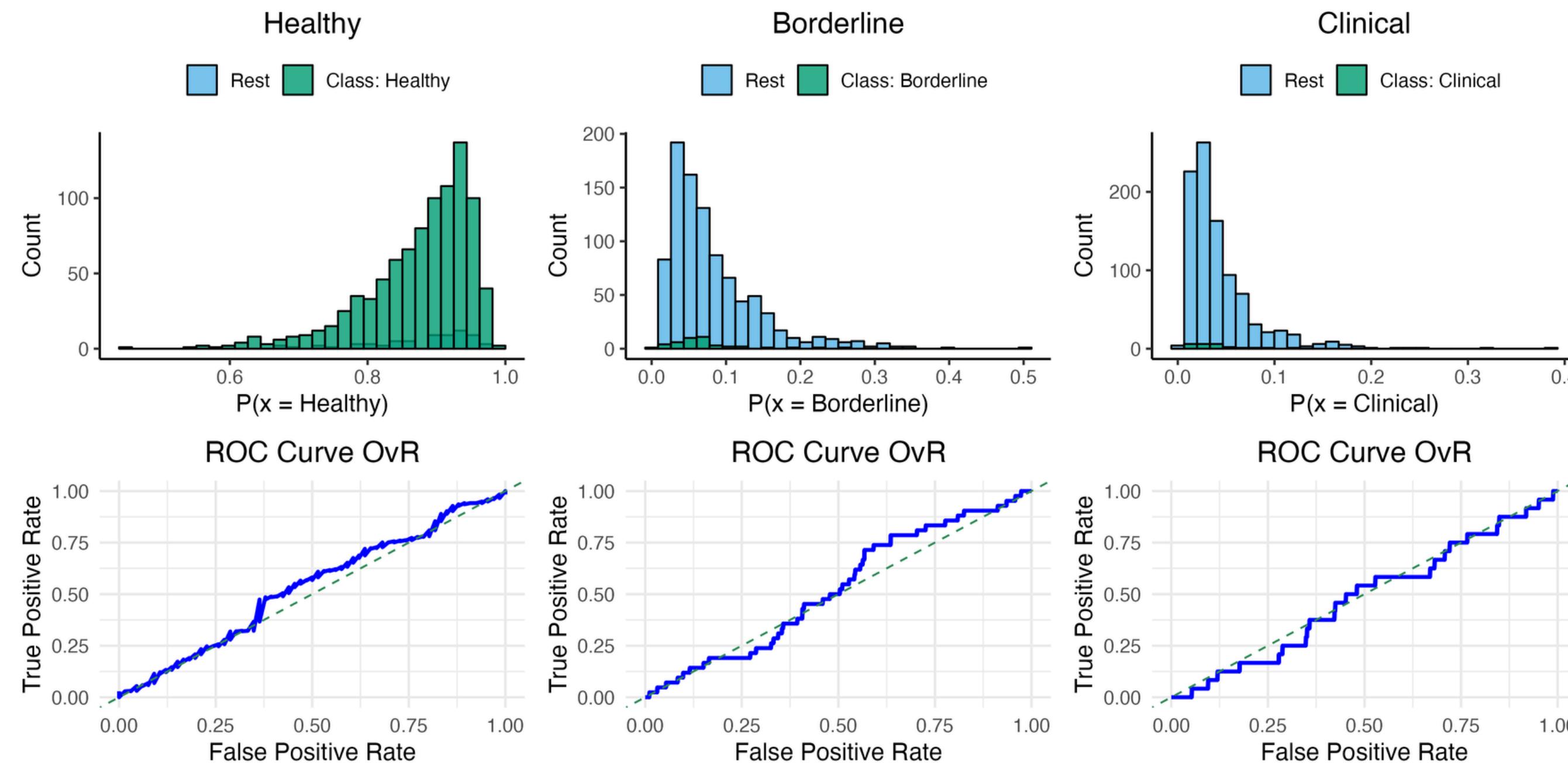
	Parent-Reported		Child-Reported	
Model Configuration	Tuned + Argmax	Tuned + Threshold-Cal.	Tuned + Argmax	Tuned + Threshold-Cal.
Overall Performance				
Accuracy	0.95	0.68	0.93	0.92
95% CI	0.93–0.96	0.64–0.71	0.92–0.95	0.91–0.94
Cohen's Kappa	-0.00	-0.03	0.0	-0.02
Class Balance				
Sensitivity (Clinical)	0.0	0.083	0.0	0.0
Specificity (Clinical)	1.0	0.86	1.0	1.0
Class-Level Discrimination				
F1 Score (Clinical)	0.0	0.03	0.0	0.0

- High accuracy is **misleading**
- None of the models reliably detected clinical cases, indicating failure to identify at-risk individuals.

Results

Model Performance by Class

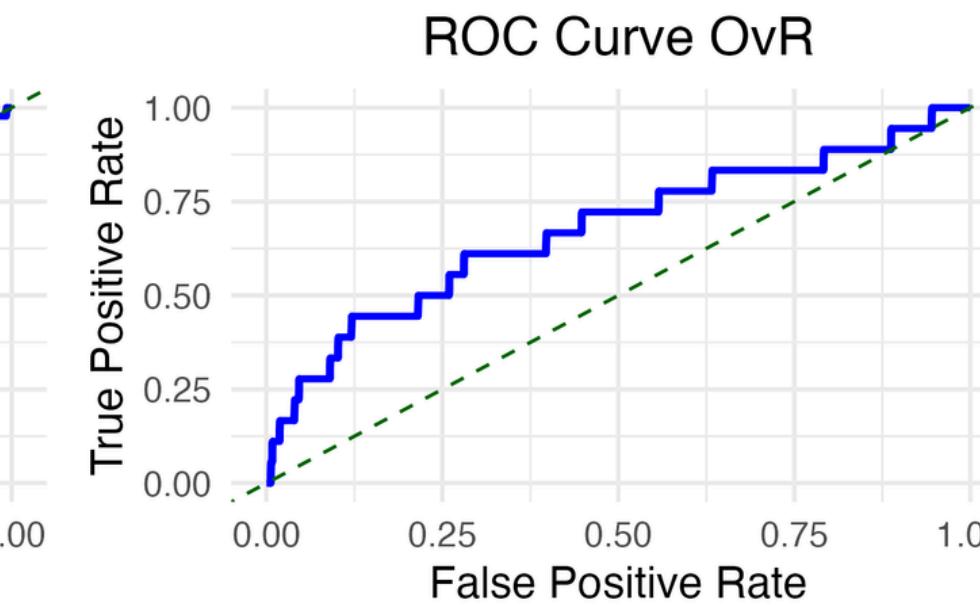
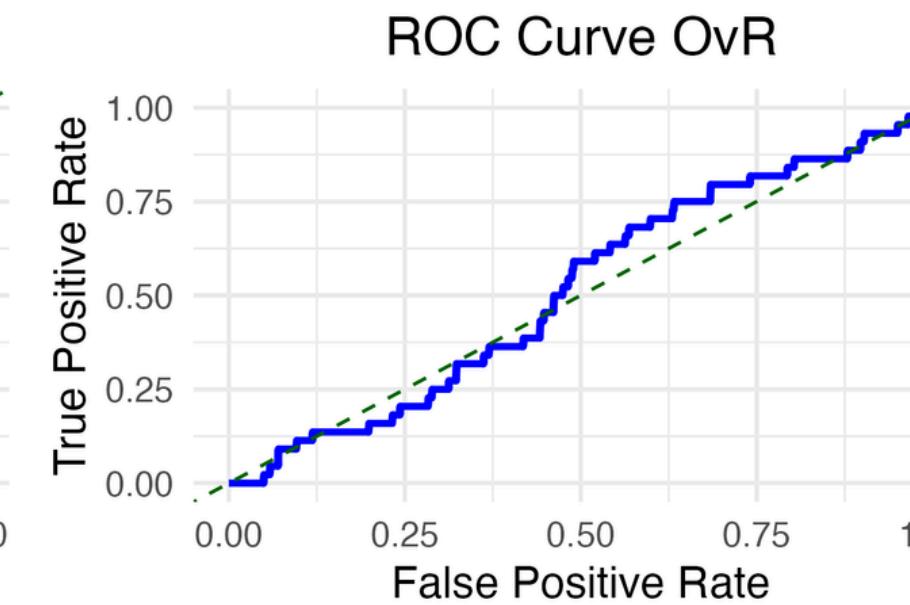
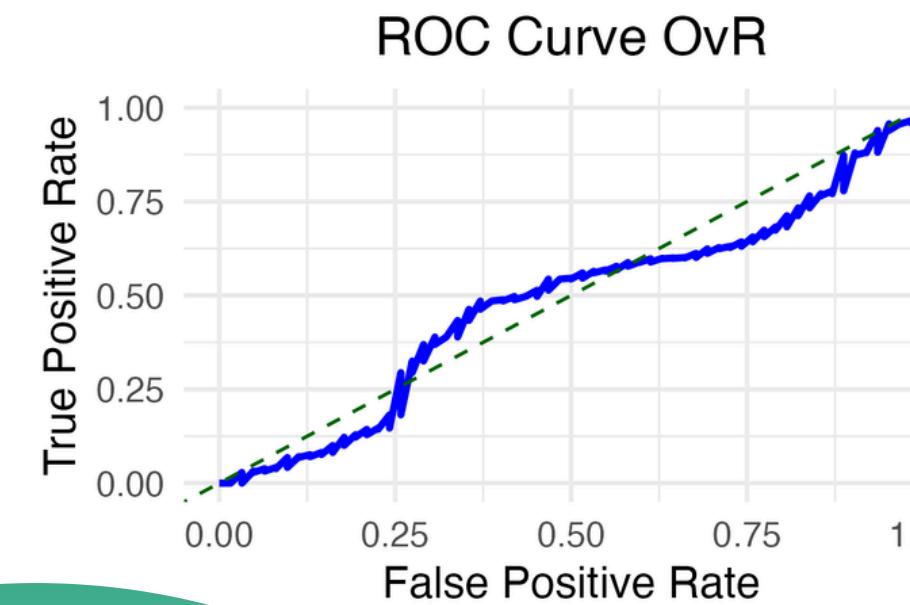
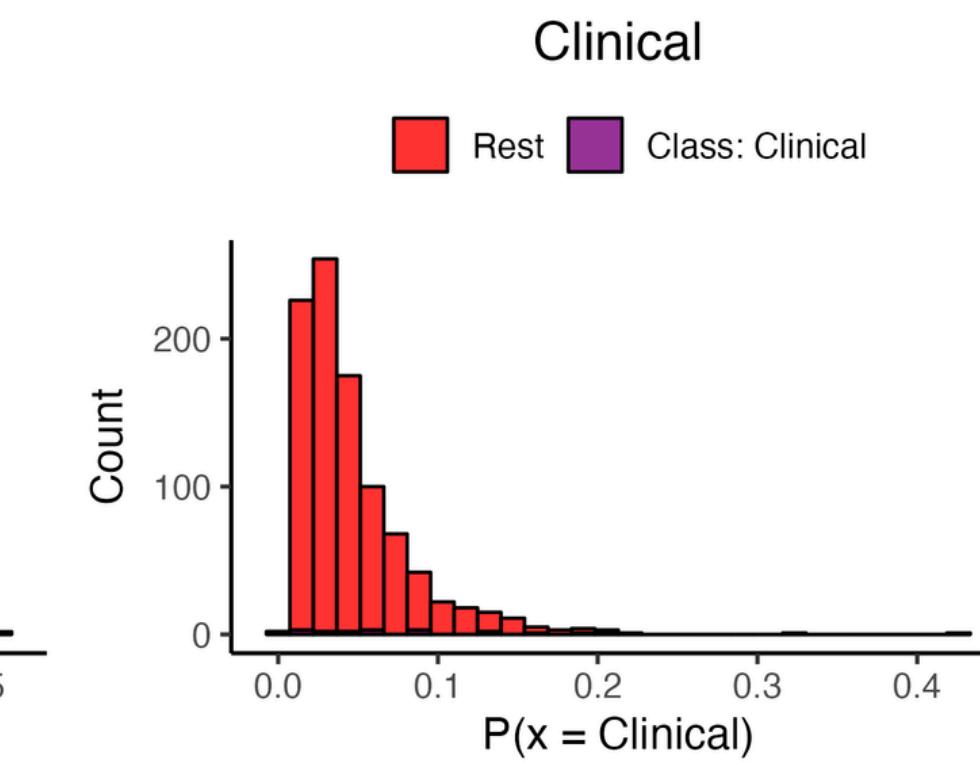
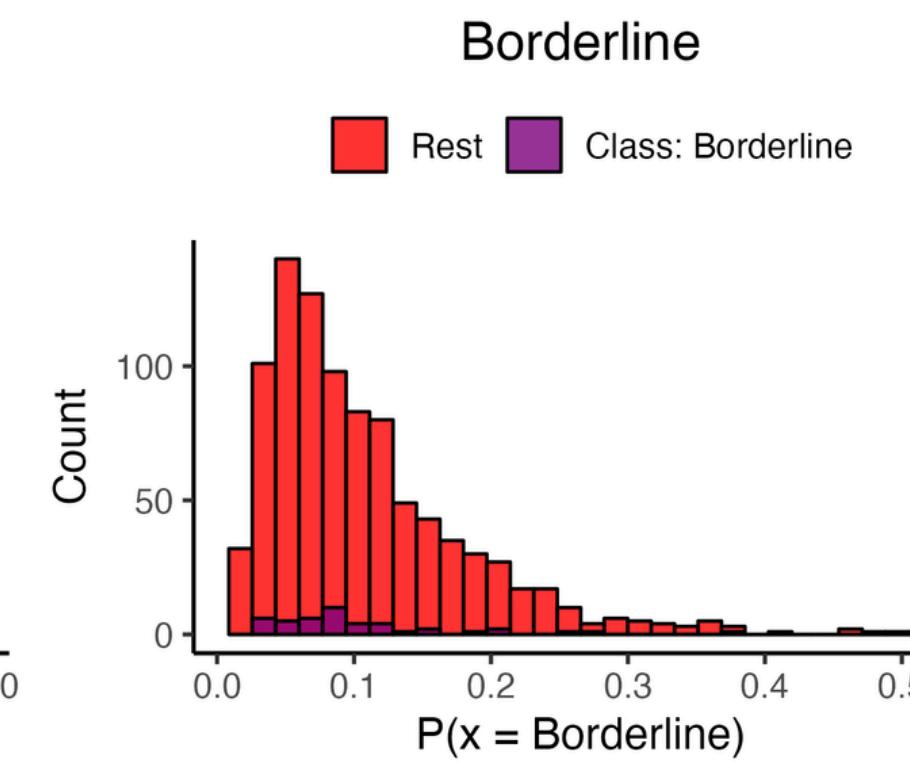
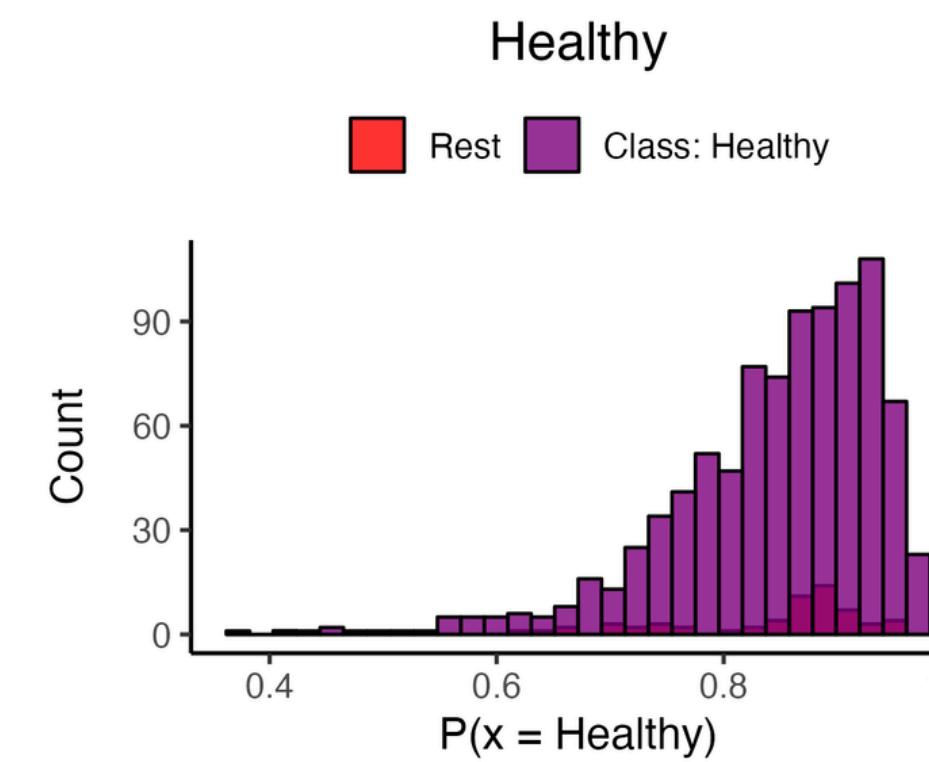
Parent-Reported



Results

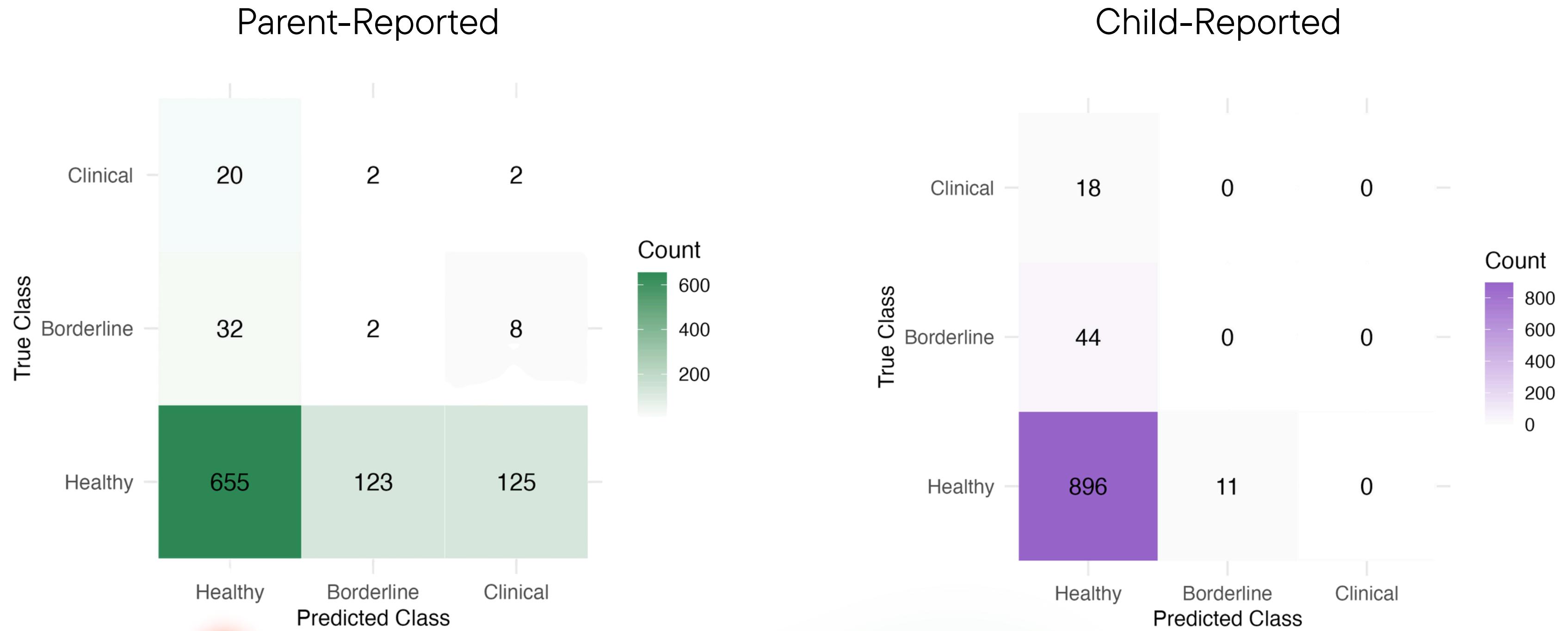
Model Performance by Class

Child-Reported



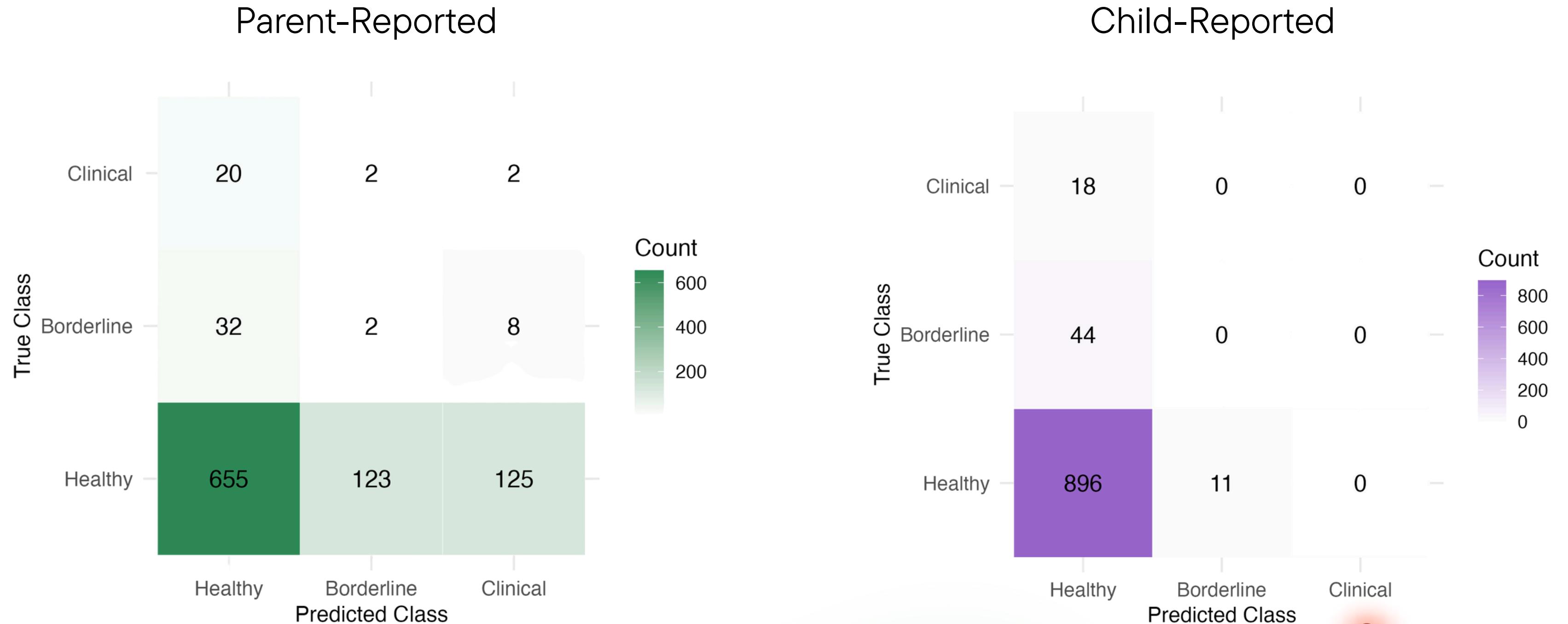
Results

Confusion Matrix



Results

Confusion Matrix



DISCUSSION



Summary

Can structural brain features predict the severity of OCD-related internalizing symptoms in children?

Does predictive accuracy differ between child-reported and parent-reported internalizing symptoms?

- No significant predictive signal found in either model.
- High accuracy was misleading, did not translate to relevant true prediction.

Why Did The Models Struggle?

The Problem of Imbalanced Data

Too few clinical cases.

Insufficient Sensitivity of Neuroimaging-Based Input Features

Noisy, indirect, adult-based regions.

Insufficient Sensitivity of Symptom Measure

Mild or hidden symptoms.

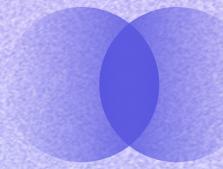
Lack of contextual and behavioral data

No behavioral or environmental features included.

Take Home Message

The absence of strong predictive results invites us to think more critically, not only about what we measure, but also how we model it, and what kind of signal we expect to find using rigorous, generalizable methods.

QUESTIONS?



Email
hello@reallygreatsite.com

Social Media
[@reallygreatsite](https://www.twitter.com/reallygreatsite)

Call us
123-456-7890

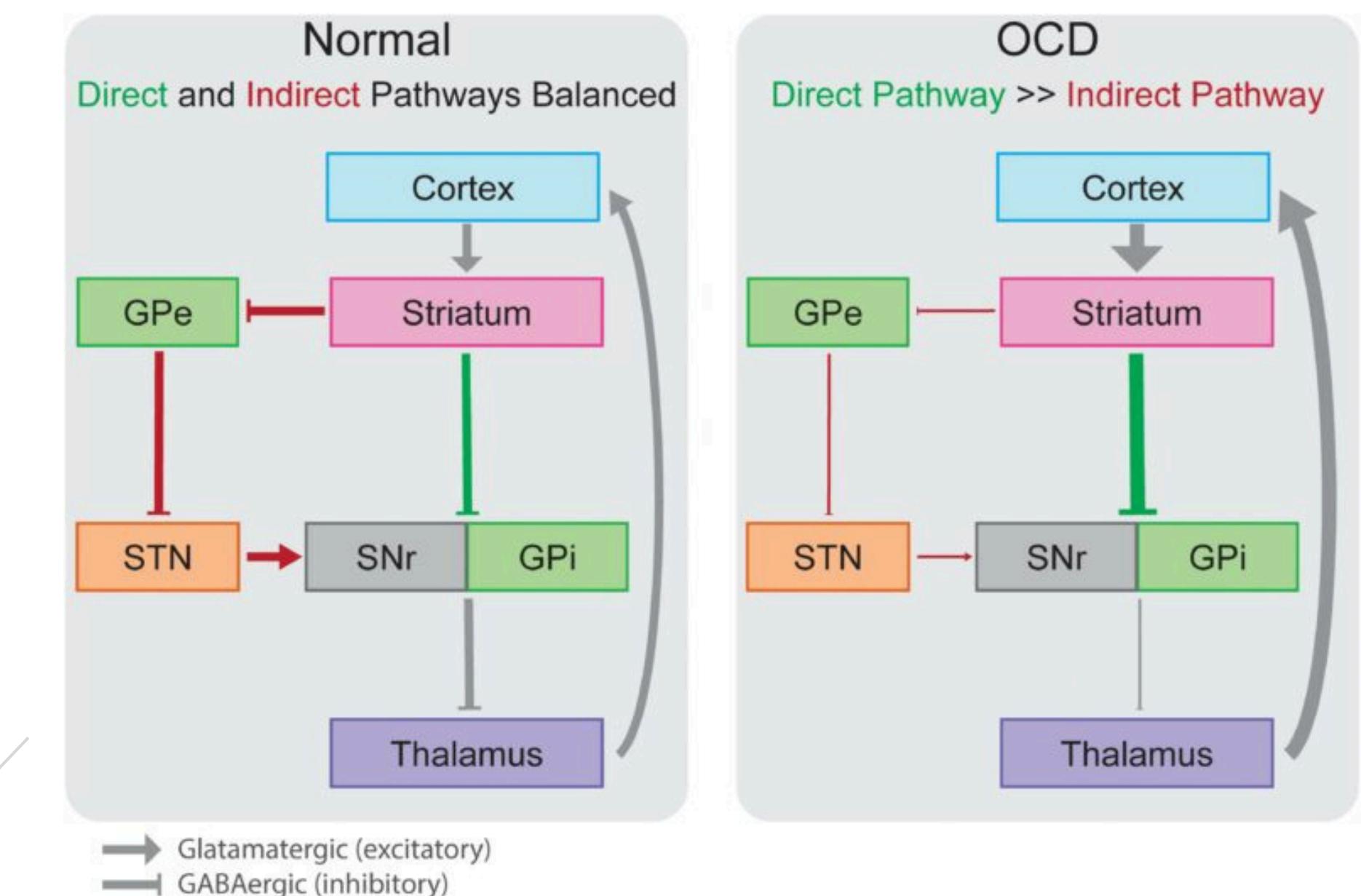
Excluded Medication Use

```
antipsychotics <- c("Risperidone", "Aripiprazole", "Olanzapine", "Quetiapine", "Haloperidol",  
"Ziprasidone", "Clozapine", "Lurasidone", "Paliperidone")
```

```
antidepressants <- c("Fluoxetine", "Sertraline", "Citalopram", "Escitalopram", "Paroxetine",  
"Bupropion", "Duloxetine", "Venlafaxine", "Desvenlafaxine", "Amitriptyline")
```

BRAIN DEVELOPMENT & CORE CIRCUITS

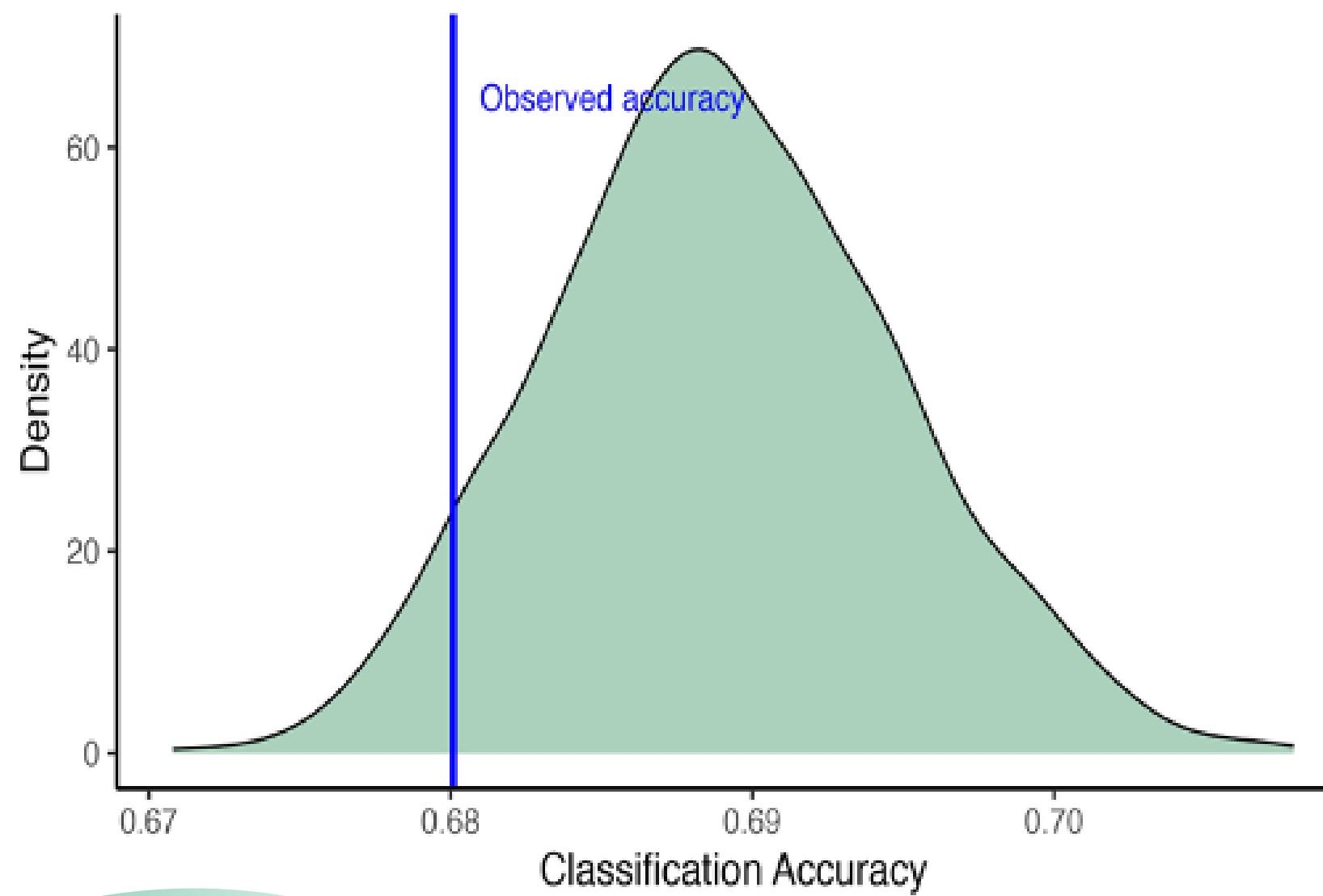
- OCD often emerges during late childhood, a key period for brain maturation
- Involves dysfunction in the CSTC circuit (basal ganglia, thalamus, prefrontal cortex)
- Linked to impaired behavior regulation and control
- Structural findings in youth:
 - Larger thalamus
 - Thinner prefrontal & parietal cortex
- Alterations mainly seen in unmedicated patients



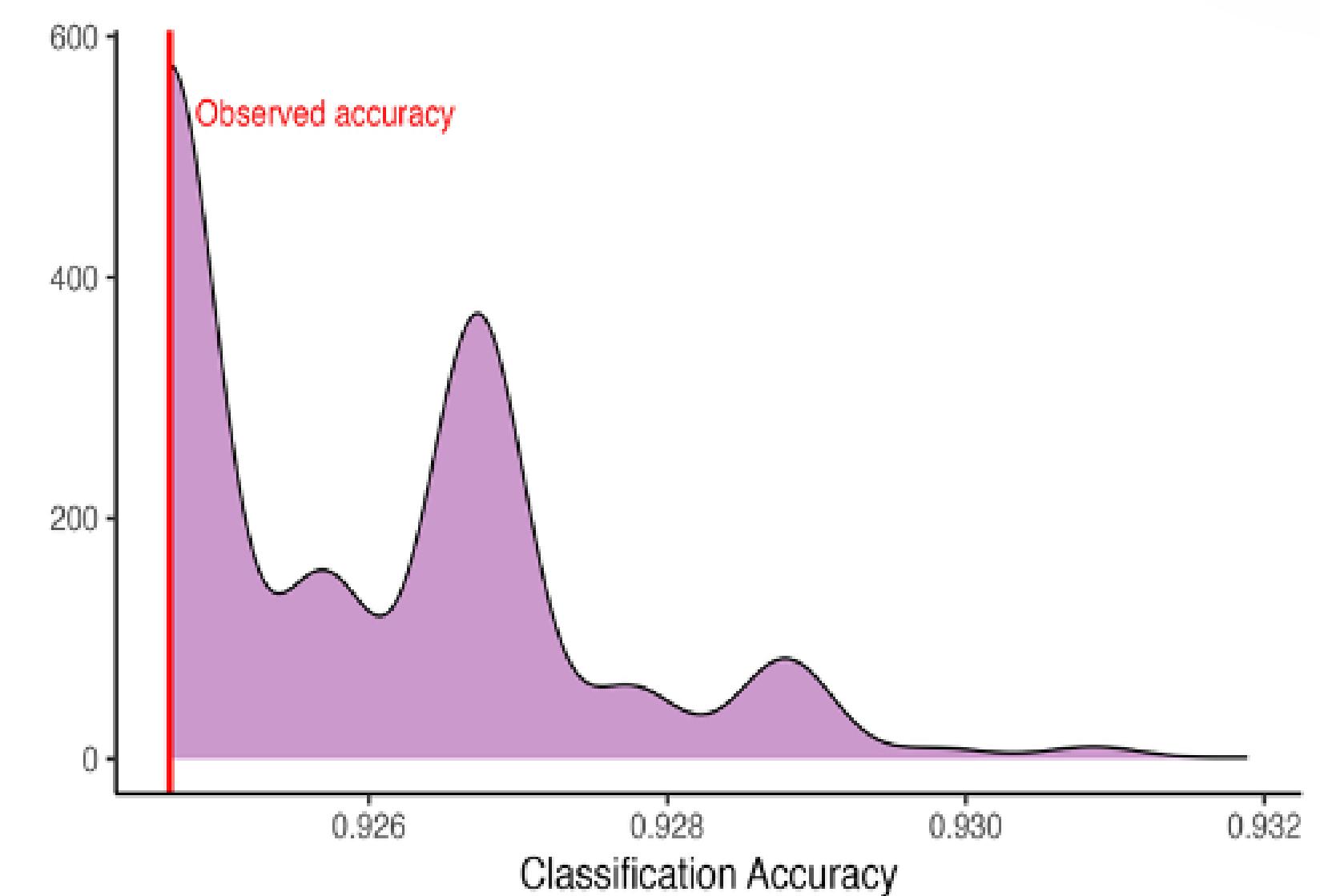
Results

Sanity Check: Permutation Test

Parent-Reported (CBCL)



Child-Reported (BPM)



	CBCL				BPM			
Model Configuration	Default	With Weights	Tuned + Argmax	Tuned + Threshold-Cal.	Default	With Weights	Tuned + Argmax	Tuned + Threshold-Cal.
Overall Performance								
Accuracy	0.3333	0.3323	0.9505	0.6801	0.3137	0.3158	0.934	0.9247
95% CI	0.3037–0.364	0.3027–0.363	0.9349–0.9633	0.6497–0.7094	0.2846–0.344	0.2866–0.3461	0.9164–0.9488	0.9062–0.9405
Cohen's Kappa	-0.0074	0.0224	-0.0016	-0.0289	-0.0136	-0.009	0	-0.0168
Class Balance								
Balanced Accuracy	0.4942	0.5687	0.4997	0.465	0.4745	0.4795	0.5	0.4966
Sensitivity (Borderline)	0.2857	0.4524	0	0.048	0.2273	0.2955	0	0
Sensitivity (Clinical)	0.4166	0.5417	0	0.083	0.3889	0.3334	0	0
Specificity (Borderline)	0.6505	0.644	0.9989	0.865	0.6605	0.6681	1	0.9881
Specificity (Clinical)	0.6878	0.6804	1	0.859	0.6572	0.6509	1	1
Class-Level Discrimination								
Precision (Borderline)	0.0357	0.0544	0	0.016	0.0309	0.0407	NaN	0
Precision (Clinical)	0.0327	0.0413	NaN	0.015	0.021	0.0178	NaN	NaN
F1 Score (Borderline)	0.0634	0.0967	0	0.0236	0.0538	0.0723	0	0
F1 Score (Clinical)	0.061	0.0763	0	0.0248	0.0382	0.0334	0	0