Title?

Table of Contents

[Introduction 3](#_Toc192525368)

[Obsessive-Compulsive Disorder 3](#_Toc192525369)

[Informant Discrepancies 3](#_Toc192525370)

[MRI 5](#_Toc192525371)

[Neurobiological Theories of OCD 6](#_Toc192525372)

[Fronto-Limbic Circuit 7](#_Toc192525373)

[Sensorimotor Circuit 7](#_Toc192525374)

[Ventral Cognitive Circuit 7](#_Toc192525375)

[Ventral Affective Circuit 8](#_Toc192525376)

[Dorsal Cognitive Circuit 8](#_Toc192525377)

[Adolescent OCD 8](#_Toc192525378)

[Structural abnormalities 8](#_Toc192525379)

[Machine Learning 9](#_Toc192525380)

[eXtreme Gradient Boosting 10](#_Toc192525381)

[The present study (?) 11](#_Toc192525382)

[Research question 11](#_Toc192525383)

[Hypothesis 11](#_Toc192525384)

[Methods 12](#_Toc192525385)

[Data Source and Collection Procedures 12](#_Toc192525386)

[Data acquisition 12](#_Toc192525387)

[Questionaries 12](#_Toc192525388)

[Kiddie Affective 12](#_Toc192525389)

[Achenbach System and Empirically Based Assessment (ASEBA) 12](#_Toc192525390)

[The Obsessive-Compulsive Symptom (OCS) Scale (?) 13](#_Toc192525391)

[Structural MRI 13](#_Toc192525392)

[Sample 14](#_Toc192525393)

[Statistical analyses/Preliminary analyses(?) 14](#_Toc192525394)

[Modelling approach(?) 14](#_Toc192525395)

[References 15](#_Toc192525396)

Introduction

Obsessive-Compulsive Disorder

Obsessive-compulsive disorder (OCD) is recognized as a prevalent and persistent neuropsychiatric condition, impacting an estimated 2% to 3% of individuals worldwide (de Mathis et al., 2013). The disorder commonly arises in early life and is characterized by the presence of compulsions – ritualized behavioral or mental acts, and obsessions – intrusive and unwanted thoughts and worries (Karno et al., 1988). OCD is unique among mental illnesses in that it exhibits both externalizing and internalizing symptoms (Guzick et al., 2019). Externalizing features, like compulsivity and repetitive actions, are often outwardly disruptive and align with disorders such as attention-deficit/hyperactivity disorder (ADHD) and disruptive behavior disorders. In contrast, internalizing aspects, including anxiety, concerns, and obsessions, cause internal distress and align with conditions like depressive and anxiety disorders, often leading to avoidance or withdrawal (T. M. Achenbach, 2001). Understanding OCD within this dual framework enhances our grasp of its complexity and informs more effective therapeutic strategies. This framework is therapeutically beneficial and supported by empirical research (Kessler et al., 2011; Slade & Watson, 2006).

Informant Discrepancies

Traditionally, clinicians have depended on parents to provide comprehensive information about how an illness and its treatment affect their children. This reliance stems from the perception that children may not possess the cognitive and linguistic skills required to understand and respond to surveys accurately (ref). Consider the case of Liam, a 12-year-old having battled severe OCD for several years. After starting therapy, he was showing signs of improvement. According to Liam, he felt he was making excellent progress. He reduced his handwashing rituals from every hour to three times a day and started joining some family meals. He was also beginning to meet his friends for short walks around the neighborhood. However, his parents observed a different reality. While Liam had made some progress, he often became trapped in lengthy rituals that caused him significant distress. He had yet to return to school full-time, attending only partial days if he went at all. Though he started venturing out with friends, it was only to familiar, controlled environments. His parents continued to monitor his progress closely, supporting him in his journey while remaining aware of the continuous obstacles that his OCD presented.

This vignette demonstrates the importance of recognizing that the child's perspective is distinct but equally valid. However, the insights provided by Liam can differ significantly from those of his mother, highlighting the potential discrepancies in information regardless of whether the goal is clinical assessment or research. Mental health problems can vary across different contexts (Bauducco et al., 2024; Beesdo et al., 2009). Children and adolescents may exhibit mental health concerns in specific environments, such as at home or school, while appearing unaffected in others, like during peer interactions. These contextual variations are evident across various domains, including conduct problems, attention, hyperactivity, and anxiety (Beesdo et al., 2009). This divergence in perspectives introduces complexities in clinical practice, research, and theory regarding child psychiatry and psychopathology (Chen et al., 2017; Salbach-Andrae et al., 2009). These differences have been thoroughly examined and will be covered in more detail below. Furthermore, while the use of multiple informants in mental health assessment is thought to enhance our understanding of the psychological functioning of children, particularly in the infant population, we are still in the process of discovering how to effectively utilize this wealth of information (Reyes, 2013).

The issue of informant discrepancies is particularly pertinent when interpreting study findings in the field of developmental psychopathology. A significant portion of the evidence about prevalence rates of psychological disorders, classification of diagnosis, and effectiveness of interventions for children is derived from reports by multiple informants (Weisz et al., 2005). For instance, depending on whether parent or teacher assessments are used to categorize the child's problem or if both are taken into account at the same time, the prevalence rates of conduct and oppositional defiant disorders in community samples vary from 1.6% to 10.2% (Offord et al., 1996). Furthermore, depending on the informant, it is typical to find inconsistent results from controlled studies evaluating psychological therapies (De Los Reyes & Kazdin, 2005). Understanding these discrepancies is crucial for accurately assessing intervention outcomes and advancing research in developmental psychopathology.

The phenomenon of informant discrepancy has been recognized for nearly 70 years, dating back to Lapouse and Monk's work in 1958. Achenbach, McConaughy, and Howell (1987) conducted a seminal analysis of 119 studies investigating these informant inconsistencies. Their key findings included: (a) reports of the same behavior by different informants generally show low to moderate agreement; (b) the reports of two informants observing children in the same setting are more similar than those of two informants observing children in different settings; (c) there is greater agreement between informants' reports for younger children compared to older ones; and (d) reports of externalizing behaviors like aggression show higher consistency than those of internalizing behaviors such as anxiety. They concluded by stating, "Different informants are needed for different situations. . . there is no royal road or preeminent gold standard for phenomena that are inevitably affected by assessment procedures and other situational variables" (p. 227–228).

Consequently, the primary objectives of the informant discrepancies research summarized by Achenbach et al. (1987) were to outline the extent of informant discrepancies, identify the informant pairs (e.g., parent and child, teacher and parent) with the greatest discrepancies, and pinpoint the behavioral domains where these discrepancies were most pronounced. A prominent finding indicated discrepancies and varying accuracy in symptom reporting, with no clear consensus.

Agreement as a test of validity for multiple informant (?) Child reports internalized symptoms more accurately, while parents tend to be more precise in identifying externalized symptoms (Silverman & Eisen, 1992).

MRI

Advances in neuroimaging, particularly magnetic resonance imaging (MRI), have elucidated the brain's role in OCD, pointing to abnormalities within the cortico-striato-thalamo-cortical (CSTC) circuit and other key regions (de Wit et al., 2014; Hu et al., 2017; Picó-Pérez et al., 2020). Such findings suggest that neuroimaging biomarkers hold promise for enhancing diagnostic accuracy and understanding the neurobiological underpinnings of OCD.

Neurobiological Theories of OCD

A diagram of a brain

AI-generated content may be incorrect.

OCD is a clinically and etiologically highly heterogeneous disorder with multiple overlapping symptom dimensions (Bragdon & Coles, 2017). These symptoms are mediated by partially distinct neural systems (van den Heuvel et al., 2009). The CSTC model is the most widely accepted explanation for the neurobiological underpinnings of OCD (Graybiel & Rauch, 2000; van den Heuvel et al., 2016). The CTSC model describes how loops between the cortex, striatum, thalamus, and back to the cortex regulate thought patterns and behaviors (Brennan & Rauch, 2017). The neurocircuit-based model of OCD builds upon the traditional CSTC model by incorporating additional brain circuits that contribute to the diverse symptom profiles of the disorder (Shephard et al., 2021). While the CSTC model primarily explains compulsions as failures in inhibitory control, the neurocircuit-based approach expands this by considering how emotional regulation, habit formation, sensory processing, and reward sensitivity play crucial roles in OCD. This model identifies five key circuits: fronto-limbic, sensorimotor, ventral cognitive, ventral affective, and dorsal cognitive circuits, each associated with specific symptom dimensions. It is important to note that these neurocircuits have more complex functions beyond the specific neurocognitive alterations discussed in this short summary. Additionally, these circuits are highly interconnected rather than functioning as isolated systems, despite how they may be presented in the following sections.

Fronto-Limbic Circuit

The fronto-limbic circuit, involving the amygdala and ventromedial prefrontal cortex (vmPFC), regulates fear and emotional responses (Kohn et al., 2014). In OCD, hyperactivity in the amygdala strengthens feelings of fear and anxiety, leading to excessive worry and intrusive thoughts, while impaired top-down regulation by the vmPFC makes it difficult to control these emotions (Milad et al., 2013). Dysregulated fear reactions to intrusive thoughts, controlled by this circuitry, may be the initial cause of obsessions for some patients. Furthermore, studies indicate alterations in amygdala-prefrontal connectivity during the onset of symptoms (Paul et al., 2019). The connectivity within this neural pathway has been found to be predictive of therapy outcomes for OCD in youth (Cyr et al., 2021).

Sensorimotor Circuit

The sensorimotor circuit, which includes the supplementary motor area (SMA), putamen, pallidum, and thalamus, governs motor behaviors and sensory integration (van den Heuvel et al., 2016). Dysfunction in this circuit explains why some OCD symptoms stem from sensory-driven urges, such as "not-just-right" feelings, averse or uncomfortable sensations that drive compulsions like excessive touching or arranging objects (Stern et al., 2025). This is particularly evident in compulsions related to cleanliness, where patients may feel a tactile sensation of dirtiness, prompting excessive washing or cleaning. This circuit is also implicated in habit formation, where repeated compulsions become automatic and disconnected from their original triggers, leading to rigid, motor-driven rituals that persist even when anxiety is no longer present (Gillan & Robbins, 2014).

Ventral Cognitive Circuit

The ventral cognitive circuit includes prefrontal areas such as the inferior frontal gyrus (IFG) and ventrolateral prefrontal cortex, alongside subcortical regions like the subthalamic nucleus (STN), ventral caudate, and thalamus, which are integral to self-regulation functions. The IFG and the ventral caudate function as a "braking system" for response inhibition, facilitating the ability to suppress inappropriate reactions (Shephard et al., 2021). This means dysfunction here prevents individuals from stopping compulsions even when they recognize them as irrational (van den Heuvel et al., 2016).

Ventral Affective Circuit

The ventral affective circuit including the orbitofrontal cortex (OFC) and nucleus accumbens (NAcc), is responsible for reward processing and motivation (van den Heuvel et al., 2016). In some OCD cases, compulsions may not just alleviate anxiety but become rewarding behaviors themselves, reinforcing habitual and compulsive loops. Consistent with these clinical findings, studies have reported heightened connectivity between the NAcc and other reward-processing regions, such as the OFC during resting-state brain activity, with this increased connectivity correlating with the severity of OCD symptoms (Xie et al., 2017).

Dorsal Cognitive Circuit

Lastly, the dorsal cognitive circuit, involving the dorsolateral prefrontal cortex (dlPFC) and dorsomedial prefrontal cortex (dmPFC), is essential for executive functioning and cognitive flexibility (van den Heuvel et al., 2016). Due to this circuit's broad and fundamental role, specific dorsal cognitive dysfunctions are not linked to distinct OCD symptom profiles but rather general impairments that interact with other neurocircuit disruptions (Shephard et al., 2021). When this circuit is impaired, individuals struggle with rigid thinking, difficulty shifting attention, and poor emotional regulation, which can contribute to the persistence of obsessions and repetitive behaviors (Van Schalkwyk et al., 2016).

Adolescent OCD

This expanded model is especially relevant for adolescents because their brains are still developing, particularly in areas like the prefrontal cortex, which governs impulse control and emotional regulation (ref). By mapping specific OCD symptoms to these circuits, the neurocircuit-based model provides a more comprehensive understanding of the disorder and allows for the development of targeted interventions (Shephard et al., 2021).

Structural abnormalities

Many studies highlight structural differences in the brains of individuals with OCD, particularly in regions like the thalamus, parietal cortices, striatal regions, and fronto-parietal areas (van den Heuvel et al., 2022; Wang et al., 2022; Wu et al., 2022). Some studies note mixed findings and developmental differences, while others focus more on surface shape alterations, specifically in pediatric samples. Multiple studies agree that adolescents with OCD tend to show a reduction in cortical thickness or volume in parietal and frontal regions (Pagliaccio et al., 2021; Wu et al., 2022). This is particularly noted in areas such as the inferior and superior parietal cortices and certain frontal gyri. There is general agreement that the thalamus may show increased volume in adolescents with OCD; however, the degree of enlargement and the specific subnuclei involved can vary across studies (van den Heuvel et al., 2022). There are consistent findings of structural alterations in several subcortical regions, including the caudate nucleus, putamen, and pallidum, which are implicated in OCD (Wang et al., 2022).

Wang et al. (2022) identified specific structural changes in subcortical brain regions in individuals diagnosed with OCD, noting variations in the NAcc, amygdala, and pallidum.

Importantly, they found structural-developmental variations in the pallidum and accumbens associated with the presence of an OCD diagnosis. Adolescents show more pronounced structural deviations in the NAcc and pallidum surfaces than adults, which relates to the sensorimotor circuit and ventral affective circuit. They noted that the NAcc could potentially serve as a biomarker for OCD development. Additionally, they found an inward deformation of the amygdala, most substantially in adults, which correlated with symptom severity. This aligns with the fronto-limbic circuit, emphasizing the role of fear and emotional dysregulation in OCD. These findings suggest that OCD is not only a disorder of habit formation (as CSTC emphasizes) but also involves dysfunctional emotional regulation and altered motivation systems. Furthermore, these findings in pediatric OCD suggest that OCD symptoms in youth may be more habit-driven and sensory-based, aligning with the sensorimotor circuit in the expanded model. Meanwhile, the amygdala findings in adults suggest a shift toward fear-based compulsions consistent with the fronto-limbic circuit. This supports a developmental progression in OCD, where younger individuals may struggle more with automatic, sensory-driven compulsions, whereas adults may experience more emotional dysregulation and cognitive rigidity (Wang et al., 2022).

Statistical Learning

Statistical learning is fundamentally connected to generalization, aiming to identify patterns within a dataset that can be applied to new, unseen data for accurate predictions. This is the core of supervised learning, which focuses on establishing relationships between a response variable Y and a group of predictor variables X. Shmueli (2011) highlights the difference between explanatory modeling, which seeks to understand the causal relationship between X and Y, and predictive modeling, which primarily aims to forecast Y, using X as tools to achieve accurate prediction. Predictive modeling does not delve into causal theory; rather, it prioritizes accuracy in predictions using various methodologies (Shmueli, 2011). There are two main approaches to predictive modeling: probabilistic models, which assume a specific statistical form for the data, and supervised learning models, which focus on creating functions that map inputs to outputs without necessarily assuming a data generation process. The accuracy of predictions is gauged by a loss function that measures the discrepancy between the predicted and the true or observed outcomes, crucial for refining the model to make accurate predictions in real-world applications. This thesis will focus on predictive modeling of adolescent OCD symptomology as reported by adults vs. self. By using sMRI data as predictor variables, the study will apply the supervised learning algorithm XGBoost to try achieve accurate predictions.

There are multiple approaches one might take towards constructing a predictive model. One approach is to build a probabilistic model of the data generating process, often referred to as a statistical model. Once a probabilistic model is built, one might use this to make predictions. Another approach to the predictive modeling task is using the framework of statistical learning theory, which provides the theoretical basis for many modern machine learning algorithms (von Luxburg and Schoelkopf, 2008). In this framework, the task of building predictive models is referred to as supervised learning. The models built in this framework are simply predictive functions designed to make accurate predictions on new, unobserved data from PY,X . These models need not be probabilistic models of the data generating process, but as we will see, they often end up having a probabilistic interpretation either way.

In the context of neuroscience, the ability of neuroimaging to identify the neurological underpinnings of OCD has sparked interest in using MRI indices to distinguish between healthy people and OCD patients. In the most common machine learning paradigm, a computer program learns to associate brain imaging data with specific classifications, such as diagnostic groups. This approach can help identify key predictive features that differentiate between these categories (Enrico et al., 2021). Effectively, machine learning aims to uncover patterns within data without requiring explicit instructions and minimal assumptions (Nosari et al., 2024).

Learning methods

XGBoost models are Tree based models are a type of supervised learning algorithms that use a branching tree structure.

The basic premise of all tree-based classification algorithms is that they learn a sequence of questions that separates cases into different classes. Each question has a binary answer, and cases will be sent down the left or right branch depending on which criteria they meet. There can be branches within branches; and once the model is learned, it can be graphically represented as a tree. (Chapter 7. Classifying with Decision Trees, n.d.)

Figure 7.1. The structure of a decision tree. The root node is the node that contains all the data prior to splitting. Nodes are split by a splitting criterion into two branches, each of which leads to another node. Nodes that do not split any further are called leaves.

Notice that our model has a branching, tree-like structure, where each question splits the data into two branches. Each branch can lead to additional questions, which have branches of their own. The question parts of the tree are called nodes, and the very first question/node is called the root node. Nodes have one branch leading to them and two branches leading away from them. Nodes at the end of a series of questions are called leaf nodes or leaves. Leaf nodes have a single branch leading to them but no branches leading away from them. When a case finds its way down the tree into a leaf node, it progresses no further and is classified as the majority class within that leaf.

At each stage of the tree-building process, the rpart algorithm considers all of the predictor variables and selects the predictor that does the best job of discriminating the classes. It starts at the root and then, at each branch, looks again for the next feature that will best discriminate the classes of the cases that took that branch. But how does rpart decide on the best feature at each split? This can be done a few different ways, and rpart offers two approaches: the difference in entropy (called the information gain) and the difference in Gini index (called the Gini gain).

 Entropy and the Gini index are two ways of trying to measure the same thing: impurity. Impurity is a measure of how heterogeneous the classes are within a node.

|  |
| --- |
|  |

Note

If a node contains only a single class (which would make it a leaf), it would be said to be pure.

By estimating the impurity (with whichever method you choose) that would result from using each predictor variable for the next split, the algorithm can choose the feature that will result in the smallest impurity. Put another way, the algorithm chooses the feature that will result in subsequent nodes that are as homogeneous as possible.

Gradient boosting is a powerful machine learning technique introduced by Friedman (2001). The technique was motivated as being a gradient descent method in function space, capable of fitting generic nonparametric predictive models. Gradient boosting has been particularly successful when applied to tree models, in which case it fits additive tree models. Friedman devised a special enhancement for this case (Friedman, 2001, 2002). We will refer to this method as MART (Multiple Additive Regression Trees), but it is also known as GBRT (Gradient Boosted Regression Trees) and GBM (Gradient Boosting Machine). More recently, a new tree boosting method has come to stage and quickly gained popularity. It goes by the name XGBoost (Chen and Guestrin, 2016), and while it is conceptually similar to Friedmans tree boosting method MART, it also differs in multiple ways.

eXtreme Gradient Boosting

A diagram of a flowchart

AI-generated content may be incorrect.

Gradient tree boosting is a tree-based model widely used for predictive modeling tasks. The objective is to determine a function (F(x)) that maps input variables (x={x\_1, ..., x\_n}) to an output variable (y). This is done by minimizing a specified loss function (L(y,F(x))) to achieve the most accurate approximation of the function mapping (Friedman, 2001). As an ensemble learning technique, it enhances prediction accuracy by aggregating multiple weak learners. Boosting constructs trees sequentially, where each tree utilizes information from trees built previously (James et al., 2021). In the regression setting, boosting incrementally refines the model by fitting trees to the residuals rather than directly to the response variable. Small trees are used to target residuals, gradually improving areas where the model performs poorly.

XGBoost is an implementation of gradient boosting, which frames the boosting challenge as an optimization problem (Ren et al., 2019). When using XGBoost for regression tasks, several key parameters can be fine-tuned to optimize model performance (XGBoost Developers, 2022). It's important to note that this overview is not exhaustive of all tuning parameters available in XGBoost; rather, it is a short summary of some of the most impactful parameters to consider when optimizing regression models. The learning rate, or eta, determines how quickly the model learns patterns, with smaller values allowing for more cautious learning to reduce overconfidence and potential overfitting. Max\_depth sets how complex each decision tree can be by limiting the number of splits, with deeper trees capturing more intricate patterns but risking overfitting. Min\_child\_weight requires a certain amount of data in a leaf node before further splitting, promoting simpler, less complex trees. Colsample\_bytree decides how many features the model considers when building each tree, allowing for randomness that can enhance generalization. Gamma establishes the required improvement for a split to occur, encouraging simpler models by preventing unnecessary splits. Finally, subsample controls the fraction of the training data used per tree, adding variability and reducing overfitting by preventing the model from relying too heavily on specific data subsets. Together, these parameters help balance the model's ability to learn complex patterns with its ability to generalize well to new, unseen data.

XGBoost can model complex interactions within datasets, providing a robust framework for integrating diverse data types, such as neuroimaging data with behavioral and demographic variables.

The present study (?)

Despite advances in neuroimaging that emphasize the potential of structural MRI to uncover neurobiological correlates of OCD, discrepancies between parental and self-reports of OCD symptoms remain a significant challenge. Traditional diagnostic methods often overlook critical nuances in symptom severity and informant variability. Considering these challenges, this study aims to address the research question: "To what extent does structural brain data explain the variation in OCD symptoms as reported by youths versus their parents?" The study will use the ABCD dataset to explore whether the machine learning model XGBoost can detect established patterns of informant discrepancies commonly observed in psychiatric assessments.

Research question

Focus on exploratory model: To what extent does structural brain data explain the variation in OCD symptoms as reported by youths versus their parents?

Focus on predictive model: Can structural brain data be used to predict the level of OCD symptoms reported by youths and parents?

Hypothesis

There will be a significant difference in the prediction accuracy of structural brain data between self-reported and parent-reported OCD symptoms in adolescents with OCD, with an expectation of higher accuracy for self-reported symptoms for the internalizing dimension and higher accuracy for parent-reported symptoms for the externalizing dimension.

Methods

Data Source and Collection Procedures

The Adolescent Brain and Cognitive Development (ABCD) Study is a comprehensive decade-long research initiative in the United States, tracking children from ages 9-10 through late adolescence and into early adulthood. It conducts annual lab-based evaluations and biannual imaging scans to assess various mental and physical health metrics (Saragosa-Harris et al., 2022; Barch et al., 2018). The ABCD Study is designed to enhance our understanding of the behavioral, genetic, neurobiological, and environmental factors influencing health and risk factors for physical and mental health problems during adolescence. The study includes 12,000 children at baseline, recruited from 21 research sites across the United States (Karcher & Barch, 2021). To ensure the cohort is diverse and representative, the ABCD Study employs a multi-stage probability sampling technique, along with weighting methods and stratified sampling within specific regions to minimize selection bias.

Data acquisition

Questionaries

Kiddie Affective

Demographic variables

Achenbach System and Empirically Based Assessment (ASEBA)

Child Behavior Checklist (CBCL).

The CBCL is a component of the ASEBA and is used to assess children's behavioral, emotional, or social problems (Achenbach, 2001). It is a 112-item parent-reported survey, which uses a 3-point Likert scale for responses: "Very True," "Somewhat True," or "Not True," where parents are asked to rate each item based on their child's behavior "now or within the past six months." It consists of several dimensions categorized into Syndrome Scales and DSM-Oriented Scales. The eight syndrome scales are established through factor analysis (ref). They encompass clusters of common behaviors or symptoms, including (1) Anxious/Depressed, (2) Withdrawn/Depressed, (3) Somatic Complaints, (4) Social Problems, (5) Thought Problems, (6) Attention Problems, (7) Rule-Breaking Behavior, (8) Aggressive Behavior. Meanwhile, the more recently developed seven DSM-Oriented Scales align with diagnostic categories outlined in the Diagnostic and Statistical Manual of Mental Disorders. They include; (1) OCD Problems, (2) Depressive Problems, (3) Anxiety Problems, (4) Somatic Problems, (5) Attention Deficit/Hyperactivity Problems, (6) Oppositional Defiant Problems, (7) Conduct Problems (American Psychiatric Association, 2013; Nelson et al., 2001). Furthermore, these scales are grouped into three high-level domains known as (1) Internalizing Problems (which combines Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints), (2) Externalizing Problems (which combines Rule-Breaking Behavior and Aggressive Behavior), and a (3) Total Problems score that sums all problem items. These dimensions offer a detailed assessment of a child's emotional, social, and behavioral functioning, aiding in identifying areas that may benefit from therapeutic or educational interventions.

The Obsessive-Compulsive Symptom (OCS) Scale (?)

**Discrepancy (?)**

small discrepancy = child agrees with parent about OCS score (probably)

high discrepancy = child probably disagrees with the parent about the OCD score

Brief Problem Monitor (BPM).

The BPM is a 19-item self-reported survey used to assess children's behavioral and emotional functioning and their responses to interventions (RTIs). It also uses a 3-point Likert scale for responses: "Very True", "Somewhat True," or "Not True." children are instructed to rate each item based on their behavior "currently or within the past six months." The BPM measures four scales, including Internalizing, Attention Problems, Externalizing, and Total Problems scales, paralleling the items and scales found on the more comprehensive CBCL/6-18 (Achenbach et al., 2017).

Structural MRI

High-resolution T1-weighted and T2-weighted 3D structural images were acquired using Siemens, Philips, and GE 3T MRI scanners. Preprocessing includes correcting for bias field, distortion, and resampling (Hagler et al., 2019). Images were corrected for gradient nonlinearity distortions (Jovicich et al., 2006), and T2w images were registered to T1w images using a mutual information-based approach (Wells et al., 1996). Intensity non-uniformity was corrected through tissue segmentation and sparse spatial smoothing. All images were resampled to 1 mm isotropic resolution and rigidly aligned to a standard atlas space. Regions of interest (ROIs) were defined using the Destrieux atlas-based classification (Destrieux et al., 2010). This atlas uses a sulco-gyral classification, distinguishing between exposed gyri and buried sulci based on mean curvature and convexity, thus providing 74 bilateral regions (148 total). While it excludes subcortical structures, it is widely used in structural MRI studies to analyze cortical volume, thickness, and sulcal depth in neurodevelopmental and neurodegenerative research.

Also, subcortical atlas

Sample

Statistical analyses/Preliminary analyses(?)

Modelling approach(?)

References

Achenbach, McConaughy, S., Ivanova, M., & Rescorla, L. (2017). Manual for the aseba brief problem monitor for ages 6–18 (bpm/6–18). *Burlington: University of Vermont Research Center for Children, Youth, and Families*.

Achenbach, T. M. (2001). *Manual for the ASEBA school-age forms & profiles: Child behavior checklist for ages 6-18, teacher’s report form, youth self-report: An integrated system of multi-informant assessment*. ASEBA.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Association.

Bauducco, S., Gardner, L. A., Smout, S., Champion, K. E., Chapman, C., Gamble, A., Teesson, M., Gradisar, M., & Newton, N. C. (2024). Adolescents’ trajectories of depression and anxiety symptoms prior to and during the COVID-19 pandemic and their association with healthy sleep patterns. *Scientific Reports*, *14*, 10764. https://doi.org/10.1038/s41598-024-60974-y

Beesdo, K., Knappe, S., & Pine, D. S. (2009). Anxiety and Anxiety Disorders in Children and Adolescents: Developmental Issues and Implications for DSM-V. *The Psychiatric Clinics of North America*, *32*(3), 483. https://doi.org/10.1016/j.psc.2009.06.002

Bragdon, L. B., & Coles, M. E. (2017). Examining heterogeneity of obsessive-compulsive disorder: Evidence for subgroups based on motivations. *Journal of Anxiety Disorders*, *45*, 64–71. https://doi.org/10.1016/j.janxdis.2016.12.002

Brennan, B. P., & Rauch, S. L. (2017). Functional Neuroimaging Studies in Obsessive-Compulsive Disorder: Overview and Synthesis. In C. Pittenger & C. Pittenger (Eds.), *Obsessive-compulsive Disorder: Phenomenology, Pathophysiology, and Treatment* (p. 0). Oxford University Press. https://doi.org/10.1093/med/9780190228163.003.0021

*Chapter 7. Classifying with decision trees*. (n.d.). Retrieved March 17, 2025, from https://learning.oreilly.com/library/view/machine-learning-with/9781617296574/OEBPS/Text/kindle\_split\_017.html

Chen, Y.-Y., Ho, S.-Y., Lee, P.-C., Wu, C.-K., & Gau, S. S.-F. (2017). Parent-child discrepancies in the report of adolescent emotional and behavioral problems in Taiwan. *PLOS ONE*, *12*(6), e0178863. https://doi.org/10.1371/journal.pone.0178863

Cyr, M., Pagliaccio, D., Yanes-Lukin, P., Goldberg, P., Fontaine, M., Rynn, M. A., & Marsh, R. (2021). Altered fronto-amygdalar functional connectivity predicts response to cognitive behavioral therapy in pediatric obsessive-compulsive disorder. *Depression and Anxiety*, *38*(8), 836–845. https://doi.org/10.1002/da.23187

De Los Reyes, A., & Kazdin, A. E. (2005). Informant Discrepancies in the Assessment of Childhood Psychopathology: A Critical Review, Theoretical Framework, and Recommendations for Further Study. *Psychological Bulletin*, *131*(4), 483–509. https://doi.org/10.1037/0033-2909.131.4.483

de Mathis, M. A., Diniz, J. B., Hounie, A. G., Shavitt, R. G., Fossaluza, V., Ferrão, Y., Leckman, J. F., de Bragança Pereira, C., do Rosario, M. C., & Miguel, E. C. (2013). Trajectory in obsessive-compulsive disorder comorbidities. *Eur Neuropsychopharmacol*, *23*(7), 594–601. https://doi.org/10.1016/j.euroneuro.2012.08.006

de Wit, S. J., Alonso, P., Schweren, L., Mataix-Cols, D., Lochner, C., Menchón, J. M., Stein, D. J., Fouche, J.-P., Soriano-Mas, C., Sato, J. R., Hoexter, M. Q., Denys, D., Nakamae, T., Nishida, S., Kwon, J. S., Jang, J. H., Busatto, G. F., Cardoner, N., Cath, D. C., … van den Heuvel, O. A. (2014). Multicenter Voxel-Based Morphometry Mega-Analysis of Structural Brain Scans in Obsessive-Compulsive Disorder. *American Journal of Psychiatry*, *171*(3), 340–349. https://doi.org/10.1176/appi.ajp.2013.13040574

Enrico, P., Delvecchio, G., Turtulici, N., Pigoni, A., Villa, F. M., Perlini, C., Rossetti, M. G., Bellani, M., Lasalvia, A., Bonetto, C., Scocco, P., D’Agostino, A., Torresani, S., Imbesi, M., Bellini, F., Veronese, A., Bocchio-Chiavetto, L., Gennarelli, M., Balestrieri, M., … Brambilla, P. (2021). Classification of Psychoses Based on Immunological Features: A Machine Learning Study in a Large Cohort of First-Episode and Chronic Patients. *Schizophrenia Bulletin*, *47*(4), 1141–1155. https://doi.org/10.1093/schbul/sbaa190

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, *29*(5), 1189–1232.

Gillan, C. M., & Robbins, T. W. (2014). Goal-directed learning and obsessive–compulsive disorder. *Phil. Trans. R. Soc. B*, *369*(1655), 20130475–20130475. https://doi.org/10.1098/rstb.2013.0475

Graybiel, A. M., & Rauch, S. L. (2000). Toward a Neurobiology of Obsessive-Compulsive Disorder. *Neuron*, *28*(2), 343–347. https://doi.org/10.1016/S0896-6273(00)00113-6

Guzick, A. G., Cooke, D. L., McNamara, J. P. H., Reid, A. M., Graziano, P. A., Lewin, A. B., Murphy, T. K., Goodman, W. K., Storch, E. A., & Geffken, G. R. (2019). Parents’ Perceptions of Internalizing and Externalizing Features in Childhood OCD. *Child Psychiatry Hum Dev*, *50*(4), 692–701. https://doi.org/10.1007/s10578-019-00873-w

Hagler, D. J., Hatton, SeanN., Cornejo, M. D., Makowski, C., Fair, D. A., Dick, A. S., Sutherland, M. T., Casey, B. J., Barch, D. M., Harms, M. P., Watts, R., Bjork, J. M., Garavan, H. P., Hilmer, L., Pung, C. J., Sicat, C. S., Kuperman, J., Bartsch, H., Xue, F., … Dale, A. M. (2019). Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage*, *202*, 116091. https://doi.org/10.1016/j.neuroimage.2019.116091

Hu, X., Du, M., Chen, L., Li, L., Zhou, M., Zhang, L., Liu, Q., Lu, L., Mreedha, K., Huang, X., & Gong, Q. (2017). Meta-analytic investigations of common and distinct grey matter alterations in youths and adults with obsessive-compulsive disorder. *Neuroscience & Biobehavioral Reviews*, *78*, 91–103. https://doi.org/10.1016/j.neubiorev.2017.04.012

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R*. Springer US. https://doi.org/10.1007/978-1-0716-1418-1

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., MacFall, J., Fischl, B., & Dale, A. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, *30*(2), 436–443. https://doi.org/10.1016/j.neuroimage.2005.09.046

Karno, M., Golding, J. M., Sorenson, S. B., & Burnam, M. A. (1988). The Epidemiology of Obsessive-Compulsive Disorder in Five US Communities. *Arch Gen Psychiatry*, *45*(12), 1094–1099. https://doi.org/10.1001/archpsyc.1988.01800360042006

Kessler, R. C., Ormel, J., Petukhova, M., McLaughlin, K. A., Green, J. G., Russo, L. J., Stein, D. J., Zaslavsky, A. M., Aguilar-Gaxiola, S., Alonso, J., Andrade, L., Benjet, C., de Girolamo, G., de Graaf, R., Demyttenaere, K., Fayyad, J., Haro, J. M., Hu, C. yi, Karam, A., … Üstün, T. B. (2011). Development of Lifetime Comorbidity in the World Health Organization World Mental Health Surveys. *Arch Gen Psychiatry*, *68*(1), 90–100. https://doi.org/10.1001/archgenpsychiatry.2010.180

Kohn, N., Eickhoff, S. B., Scheller, M., Laird, A. R., Fox, P. T., & Habel, U. (2014). Neural network of cognitive emotion regulation—An ALE meta-analysis and MACM analysis. *Neuroimage*, *87*, 345–355. https://doi.org/10.1016/j.neuroimage.2013.11.001

Nelson, E. C., Hanna, G. L., Hudziak, J. J., Botteron, K. N., Heath, A. C., & Todd, R. D. (2001). Obsessive-Compulsive Scale of the Child Behavior Checklist: Specificity, Sensitivity, and Predictive Power. *Pediatrics*, *108*(1), e14–e14. https://doi.org/10.1542/peds.108.1.e14

Nosari, G., Delvecchio, G., Diwadkar, V. A., & Brambilla, P. (2024). Brain Imaging in Psychiatry. In A. Tasman, M. B. Riba, R. D. Alarcón, C. A. Alfonso, S. Kanba, D. Lecic-Tosevski, D. M. Ndetei, C. H. Ng, & T. G. Schulze (Eds.), *Tasman’s Psychiatry* (pp. 1285–1316). Springer International Publishing. https://doi.org/10.1007/978-3-030-51366-5\_115

Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J. E., Sanford, M., & Lipman, E. L. (1996). Integrating Assessment Data from Multiple Informants. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(8), 1078–1085. https://doi.org/10.1097/00004583-199608000-00019

Pagliaccio, D., Durham, K., Fitzgerald, K. D., & Marsh, R. (2021). Obsessive-Compulsive Symptoms Among Children in the Adolescent Brain and Cognitive Development Study: Clinical, Cognitive, and Brain Connectivity Correlates. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *6*(4), 399–409. https://doi.org/10.1016/j.bpsc.2020.10.019

Paul, S., Beucke, J. C., Kaufmann, C., Mersov, A., Heinzel, S., Kathmann, N., & Simon, D. (2019). Amygdala–prefrontal connectivity during appraisal of symptom-related stimuli in obsessive–compulsive disorder. *Psychol. Med*, *49*(2), 278–286. https://doi.org/10.1017/S003329171800079X

Picó-Pérez, M., Moreira, P. S., de Melo Ferreira, V., Radua, J., Mataix-Cols, D., Sousa, N., Soriano-Mas, C., & Morgado, P. (2020). Modality-specific overlaps in brain structure and function in obsessive-compulsive disorder: Multimodal meta-analysis of case-control MRI studies. *Neuroscience & Biobehavioral Reviews*, *112*, 83–94. https://doi.org/10.1016/j.neubiorev.2020.01.033

Ren, H., Wang, X., Wang, S., & Zhang, Z. (2019). Predict Fluid Intelligence of Adolescent Using Ensemble Learning. In K. M. Pohl, W. K. Thompson, E. Adeli, & M. G. Linguraru (Eds.), *Adolescent Brain Cognitive Development Neurocognitive Prediction* (pp. 66–73). Springer International Publishing.

Reyes, A. D. L. (2013). Strategic objectives for improving understanding of informant discrepancies in developmental psychopathology research. *Development and Psychopathology*, *25*(3), 669–682. https://doi.org/10.1017/S0954579413000096

Salbach-Andrae, H., Lenz, K., & Lehmkuhl, U. (2009). Patterns of agreement among parent, teacher and youth ratings in a referred sample. *European Psychiatry*, *24*(5), 345–351. https://doi.org/10.1016/j.eurpsy.2008.07.008

Shephard, E., Stern, E. R., van den Heuvel, O. A., Costa, D. L. C., Batistuzzo, M. C., Godoy, P. B. G., Lopes, A. C., Brunoni, A. R., Hoexter, M. Q., Shavitt, R. G., Reddy, Y. C. J., Lochner, C., Stein, D. J., Simpson, H. B., & Miguel, E. C. (2021). Toward a neurocircuit-based taxonomy to guide treatment of obsessive–compulsive disorder. *Mol Psychiatry*, *26*(9), 4583–4604. https://doi.org/10.1038/s41380-020-01007-8

Shmueli, G. (2011, January 5). *To Explain or to Predict?* arXiv.Org. https://doi.org/10.1214/10-STS330

Silverman, W. K., & Eisen, A. R. (1992). Age Differences in the Reliability of Parent and Child Reports of Child Anxious Symptomatology Using a Structured Interview. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*(1), 117–124. https://doi.org/10.1097/00004583-199201000-00018

Slade, T., & Watson, D. (2006). The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychol. Med*, *36*(11), 1593–1600. https://doi.org/10.1017/S0033291706008452

Stern, E. R., Collins, K. A., Bragdon, L. B., Eng, G. K., Recchia, N., Coffey, B. J., Leibu, E., Murrough, J. W., Tobe, R. H., Iosifescu, D. V., Burdick, K. E., & Goodman, W. K. (2025). Randomized Controlled Trial of the Effects of High-Dose Ondansetron on Clinical Symptoms and Brain Connectivity in Obsessive-Compulsive and Tic Disorders. *American Journal of Psychiatry*, *182*(3), 285–296. https://doi.org/10.1176/appi.ajp.20240294

van den Heuvel, O. A., Boedhoe, P. S. W., Bertolin, S., Bruin, W. B., Francks, C., Ivanov, I., Jahanshad, N., Kong, X.-Z., Kwon, J. S., O’Neill, J., Paus, T., Patel, Y., Piras, F., Schmaal, L., Soriano-Mas, C., Spalletta, G., van Wingen, G. A., Yun, J.-Y., Vriend, C., … ENIGMA-OCD working group. (2022). An overview of the first 5 years of the ENIGMA obsessive-compulsive disorder working group: The power of worldwide collaboration. *Human Brain Mapping*, *43*(1), 23–36. https://doi.org/10.1002/hbm.24972

van den Heuvel, O. A., Remijnse, P. L., Mataix-Cols, D., Vrenken, H., Groenewegen, H. J., Uylings, H. B. M., van Balkom, A. J. L. M., & Veltman, D. J. (2009). The major symptom dimensions of obsessive-compulsive disorder are mediated by partially distinct neural systems. *Brain: A Journal of Neurology*, *132*(Pt 4), 853–868. https://doi.org/10.1093/brain/awn267

van den Heuvel, O. A., van Wingen, G., Soriano-Mas, C., Alonso, P., Chamberlain, S. R., Nakamae, T., Denys, D., Goudriaan, A. E., & Veltman, D. J. (2016). Brain circuitry of compulsivity. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *26*(5), 810–827. https://doi.org/10.1016/j.euroneuro.2015.12.005

Van Schalkwyk, G. I., Bhalla, I. P., Griepp, M., Kelmendi, B., Davidson, L., & Pittenger, C. (2016). Toward understanding the heterogeneity in obsessive-compulsive disorder: Evidence from narratives in adult patients. *Aust N Z J Psychiatry*, *50*(1), 74–81. https://doi.org/10.1177/0004867415579919

Wang, Z., Fontaine, M., Cyr, M., Rynn, M. A., Simpson, H. B., Marsh, R., & Pagliaccio, D. (2022). Subcortical shape in pediatric and adult obsessive-compulsive disorder. *Depression and Anxiety*, *39*(6), 504–514. https://doi.org/10.1002/da.23261

Weisz, J. R., Doss, A. J., & Hawley, K. M. (2005). Youth Psychotherapy Outcome Research: A Review and Critique of the Evidence Base. *Annual Review of Psychology*, *56*(Volume 56, 2005), 337–363. https://doi.org/10.1146/annurev.psych.55.090902.141449

Wells, W. M., Viola, P., Atsumi, H., Nakajima, S., & Kikinis, R. (1996). Multi-modal volume registration by maximization of mutual information. *Med Image Anal*, *1*(1), 35–51. https://doi.org/10.1016/S1361-8415(01)80004-9

Wu, X., Yu, G., Zhang, K., Feng, J., Zhang, J., Sahakian, B. J., & Robbins, T. W. (2022). Symptom-Based Profiling and Multimodal Neuroimaging of a Large Preteenage Population Identifies Distinct Obsessive-Compulsive Disorder–like Subtypes With Neurocognitive Differences. *Biological Psychiatry : Cognitive Neuroscience and Neuroimaging*, *7*(11), 1078–1089. https://doi.org/10.1016/j.bpsc.2021.06.011

XGBoost Developers. (2022). *XGBoost Parameters—Xgboost 3.1.0-dev documentation*. https://xgboost.readthedocs.io/en/latest/parameter.html#general-parameters

Xie, C., Ma, L., Jiang, N., Huang, R., Li, L., Gong, L., He, C., Xiao, C., Liu, W., Xu, S., & Zhang, Z. (2017). Imbalanced functional link between reward circuits and the cognitive control system in patients with obsessive-compulsive disorder. *Brain Imaging Behav*, *11*(4), 1099–1109. https://doi.org/10.1007/s11682-016-9585-7