MDPI

Article

Brain morphometry and cognition as predictors of irritable bowel syndrome (2024-10-22) Submission deadline: 2024-11-30

Julie Billing [1,‡], Ben René Bjørsvik [2,3‡], Daniela M. Pfabigan [1] 0000-0002-4043-1695, [2] 0000-0002-6819-6164, Birgitte Berentsen [5,6] 0000-0003-3574-7078, Gülen Arslan Lied [5,6] 0000-0002-1827-5008, Elisabeth K Steinsvik [5] 0000-0002-8371-1988, Trygve Hausken [5] 0000-0001-7080-8396, [4‡], Astri J. Lundervold [1,‡] 0000-0002-6819-6164, Arvid Lundervold [2,3‡] 0000-0002-0032-4182*

1 Department of Biological and Medical Psychology, University of Bergen, Norway
2 Department of Biomedicine, University of Bergen, Bergen, Norway
3 Medical-AI, Mohn Medical Imaging and Visualization Center, Department of Radiology, Haukeland University Hospital, Bergen, Norway
4 Department of Clinical Medicine, University of Bergen, Bergen 5021, Norway
5 National Center for Functional Gastrointestinal Disorders, Department of Medicine, Haukeland University Hospital, Bergen 5021, Norway
6 Department of Biomedicine, University of Bergen, Bergen, Norway
5 Medical-AI, Mohn Medical Imaging and Visualization Center, Department of Radiology, Haukeland University Hospital, Bergen, Norway
* Correspondence:Arvid.lundervold@uib.no
† Current address: Affiliation.
‡ These authors contributed equally to this work.

Abstract: *Background* Irritable bowel syndrome (IBS) is a disorder within the brain-gut-axis-related diseases, characterized by abdominal pain, bloating, and altered bowel habits. The prevalence is high, 5-10% globally, with a female-to-male ratio of $\sim$ 2:1. A recent study (2022) by Skrobisz and collaborators reported that brain morphometric measures, in particular (eTIV-)normalized volume of the left thalamus, could distinguish between patients with IBS and healthy controls (HCs). *Objectives*. We aim to replicate the findings of Skrobisz et al. and extend their study by including information about sex and cognitive function and incorporating multivariate statistics and prediction models in a machine-learning framework. *Methods*. A sample of 78 participants, 49 with IBS and 29 HCs underwent a multiparametric MRI examination, performed the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), and reported age and sex. An automated brain segmentation and parcellation method (Freesurfer versions FS 6.0 and 7.4.1) was executed to obtain a collection of predefined brain morphometric measures from each individual. Univariate and multivariate statistical methods, including classification models from machine learning, were then used to analyze the degree of brain differences between the IBS and the HC group and the prediction accuracy based on MRI-derived measures alone and when included together with measures of cognitive function. With the inclusion of more women than men in the current study, we also investigated the value of morphometric measures in predicting sex. Analyses of feature importance were included to investigate the relative weight of the included measures. *Results* The findings reported by Skrobisz and collaborators, using FS 6.0, could not be replicated in our sample. Results produced by a more recent version of the MRI data suggested a software-dependent bias and complex non-biological variation. The multivariate analyses confirmed poor separation between IBS and HCs in morphometric space. The memory-related RBANS scores were significantly lower in the IBS than in the HC group, and adding information about the RBANS indexes improved the prediction of IBS versus HC, with the strongest importance of left hippocampus, verbal index, putamen, and the central region of the corpus callosum. The brain morphometric patterns did also distinguish well between women and men though, with the strongest importance of eTIV, left hippocampus, and corpus callosum volumes. *Discussion* The strong sex-related differences in specific brain regions, like the hippocampus and various white matter areas, align with existing knowledge. The improved distinction between IBS and HCs when brain morphometric measures were added by cognitive

Citation: Billing, J.; Bjørsvik, B.R.; Pfabigan, S.M.; Lundervold, A.J.; Lundervold, A. Brain morphometry, gender and cognition in IBS. *Diagnostics* **2024**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

indexes calls for future studies of the brain-gut relationship in IBS that incorporate additional MRI modalities (e.g., diffusion MRI and functional MRI) and a wider range of clinical variables.

Keywords: Irritable bowel syndrome; structural MRI; brain morphometry, cognition; supervised classification; machine learning

Introduction

Irritable bowel syndrome (IBS) is an impairing gastrointestinal (GI) disorder [1] affecting approximately 10% of the world population [1]. The syndrome is characterized by recurrent abdominal pain related to defecation, altered stool frequency, or changes in stool form [2], and can be divided into different clinical subtypes related to bowel habit abnormalities [3] and symptom severity [4]. The heterogeneity within the group is large, including patients with deliberating effect on general health status to more moderate symptoms [4]. The frequency of IBS is higher among women than men has been explained by influences from hormonal differences, as well as healthcare-seeking behavior, and cultural factors [5–9].

A bidirectional association between somatic symptoms of IBS and symptoms of psychological distress is also well documented [10], where symptoms of anxiety and depression both may be a consequence of GI symptoms and factors that may intensify or increase the frequency of abdomal pain [11]. The impact of impaired cognitive function has also been increasingly recognized as significant [12,13], at least in a subgroup of patients with IBS [10,14], probably linked to severity of IBS symptoms [15] It is essential both through its impact on daily life functioning and on adherence to treatment programs that put a heavy load on memory and executive function (see e.g., [16]). A close connection between brain structures and cognition is well documented, and it has been speculated that altered brain networks more directly influence several cognitive modalities [17] and that brain volumes associated with IBS are positively associated with cognitive performance on tests of language skills and memory function [15].

This complex symptom profile is captured by models describing a bi-directional connection between the gut and the brain [18]. In this model, the symptoms of IBS are related to a disruption of the gut-brain axis, where a myriad of sensory signals are integrated into and interpreted in the brain. Magnetic resonance imaging (MRI) is commonly used to investigate the brain's involvement in the processing of gut-related signals. Studies have most consistently identified structural and functional features of the insula, anterior cingulate cortex, thalamus, basal ganglia, and prefrontal cortex that distinguish patients with IBS from healthy controls (HCs) (see [19] for an overview). The discriminatory role of thalamic volumes has been noted in several studies [20–24], and its role as a relay center for sensory information has been interpreted as a key factor underlying the abnormal processing of gut-related pain in patients with IBS. Others have noted structural changes in the anterior cingulate cortex [22] and hippocampus, likely through their involvement in the hypothalamic-pituitary-adrenal (HPA) axis, which controls the body's stress response [25]. Furthermore, the insula and cingulate cortex have been described as key regions due to their connectivity with subcortical structures, particularly the thalamus and basal ganglia. In the study by Li et al. [15], negative and positive associations with IBS symptoms were found for different brain regions, including the globus pallidus, caudate, and putamen, while negative associations included brain regions like anterior cingulate cortex, dorsolateral prefrontal cortex, anterior and midcingulate cortices, anterior insula, hippocampus, parahippocampal cortex, thalamus, precentral gyrus, and supplementary motor area.

Overall, results in studies focusing on specific morphometric structures have been subtle and conflicting. With its complex symptom pattern, several regions are expected to be central to the development and maintenance of IBS; mediating the integration of visceral signals and responses to pain. Furthermore, altered connectivity in neural networks is

likely to contribute to the heightened pain sensitivity and dysregulated emotional responses seen in IBS patients. Despite expectations of such a widespread involvement of neural networks, there are still arguments for further morphological studies. Structural changes in specific brain regions may underlie the brain-gut axis dysregulation and could offer valuable biomarkers for diagnosis and treatment. Additionally, a morphological analysis may reveal the structural basis of cognitive or emotional symptoms linked to IBS. Given the complexity of IBS, there are thus strong arguments to include a wide range of brain regions in further studies on IBS. By revealing morphometric patterns, such results may help us to better understand individual variability among patients and guide future research on how these regions connect within larger neural networks.

A recent study by Skrobisz et al. [24] responded to this call by including a wide range of morphometric measures to discriminate between HCs and patients with digestive disorders, including IBS. Analyzed with Freesurfer version 6.0 (FSv6.0), a subtle discriminative effect between IBS and HC was found for a thalamic measure. This inspired the present study to replicate and extend the findings reported by [24] in a larger group of patients with IBS in four different directions:

(i) Replicate the findings reported by [24] by analyzing a data-set including a larger number of patients with IBS, and distinguish software-related from biological differences by comparing results using two different Freesurfer versions: the one used by Skrobisz (Fv6.0) and the more recent FSv7.4.1.

(ii) Add multivariate analytic approaches to enhance the sensitivity of detecting significant patterns and effects. We will evaluate relationships between morphometric brain structures based on covariance patterns and by applying supervised machine learning techniques to predict clinical outcomes (IBS versus HC).

(iii) Explore the add-on effect on prediction accuracy by adding cognitive performance to the morphometric measures as predictors of clinical outcomes (IBS versus HC). We assume that combining brain morphology with cognitive features would give a more comprehensive view of how IBS manifests and can provide a better understanding of how central nervous system dysfunction contributes to IBS symptoms.

We expect that these extensions will contribute with results of importance to further studies on broader brain network dysfunctions in patients with IBS [25–27].

Materials and Methods

We have discussed differences between the two sexes/genders in the current manuscript. In that there is still little knowledge regarding IBS associated problems of LGBTQ+, biological and behavioral differences are discussed mainly between the two sexes in this manuscript. We follow the recommendations from the APA: https://apastyle.apa.org/style-grammar-guidelines/bias-free-language/gender" to use "male" and "female" as adjectives (e.g., a male participant, a female experimenter) when appropriate and relevant. Use "male" and "female" as nouns only when the age range is broad or ambiguous or to identify a transgender person's sex assignment at birth (e.g., "person assigned female at birth" is correct, not "person assigned girl at birth"). Otherwise, avoid using "male" and "female" as nouns and instead use the specific nouns for people of different ages (e.g., women)."

Participants

The study is part of the Bergen Brain-Gut project, conducted at Haukeland University Hospital in Norway in 2020 - 2022 (see Berentsen et al. (2020) [28] for the protocol). The current study includes participants with IBS (n=49) and healthy controls (n=29), at least 18 years old. They were mainly recruited through media and flyers, with some patients recruited at the outpatient clinic at the hospital. A nurse contacted all participants for screening according to the inclusion and exclusion criteria (see Table 1) before they took part in an examination including a range of GI-related measures, psychometric tests, and a multiparametric MRI examination.

In the present study, we included participants who had responses on all key measures selected for this study, i.e., no missing values (see description in the Method section) and with artifact-free MRI data that could be processed with standard brain segmentation pipelines.

Inclusion criteria	Exclusion criteria
Rome-IV criteria: Recurrent abdominal pain average at least 1 day/week during the last 3 and months, and associated alterations in bowel habits at least 6 months before diagnosis. Other causes are excluded.	Pharmacological treatment affecting GI tract, including medication for anxiety and depression, diabetes, coeliac disease, IBS, Polycystic ovary syndrome, active Helicobacter pylori infection, Parkinson's disease, amyotrophic lateral sclerosis, or Psychiatric disorders.
Normal diet at least 3 weeks before inclusion	Treated with antibiotics for the last 3 months
IBS score equal to or above 175	Diets such as vegetarian or vegan
	Use of probiotics or low-FODMAP diet within the last 3 weeks
	Previous intestinal surgery, except appendectomy
	Metallic implants, claustrophobia, incompatible with MRI
	Travel outside Europe last 3 weeks
	Plan to travel in the near future
	Pregnancy

Table 1. Exclusion and inclusion criteria for the IBS patients. Source: Retrieved from [28]

Measures

Age and sex (not genetically verified) were self-reported by the participants at baseline.

The IBS-Severity Scoring system (IBS-SSS)

The IBS-Severity Scoring system is a questionnaire used to assess the severity of GI-related IBS symptoms [29]. The questionnaire includes five items related to abdominal pain, distention, bowel habits, and quality of life. The maximum score for each question is 100. A sum of scores < 75 is used to define "no or minimal problems", and the scores in the ranges $[75, 175)$, $[175, 300]$, and > 300 as "mild", "moderate", and "severe" IBS symptoms, respectively [29]. In the present study, an IBS-SSS score ≥ 175 was used as the inclusion criteria for the IBS group. Almost all HCs obtained an IBS-SSS score at the lowest level (< 75), with some reporting a score within the mild level ($[75, 175)$).

Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)

All participants performed the Norwegian A version of RBANS, administrated by a nurse trained by a clinical neuropsychologist, following the test manual's instructions [30]. The test battery comprises 10 subtests, which are combined into five index scores and a total score. These scores are expressed both as raw and as age-corrected scaled scores. The scaled scores have a mean value of 100 and a standard deviation of 15 and are based on performance in a normative group matched to population statistics of 2012 in Norway, Sweden, and Denmark. We used these scaled scores on each of the five RBANS indices for a pairwise correlation analysis between brain morphometric measures and cognitive performance.

3T MRI scanning protocol

Multiparametric MRI recordings were acquired on a Siemens Biograph mMR PET/MRI scanner with a 3 Tesla magnet. The following measurement techniques were applied at each MR examination (TA denotes the time of acquisition [min:s]): 3D T1-weighted (TA=5:35), T2-weighted (TA=5:12), GRE field mapping (TA=0:54), resting state fMRI (ep2d BOLD motion corrected, TA=9:48), and diffusion MRI (ep2d multi-directional (n=30) diffusion weighting with three b-values, TA=8:34). Total examination time was about 45 min. For this

brain morphometric study, we used the 3D T1-weighted recordings only (a 3D T1 MPRAGE sequence, with 192 sagittal slices, TR=2500 ms, TE=2.26 ms, TI=900 ms, and voxel size $1 \times 1 \times 1$ mm^3). See Fig. 1 for an example of a recorded 3D T1-w image and Fig. 2 for its semantic segmentation used to derive morphometric measures to represent the subject.

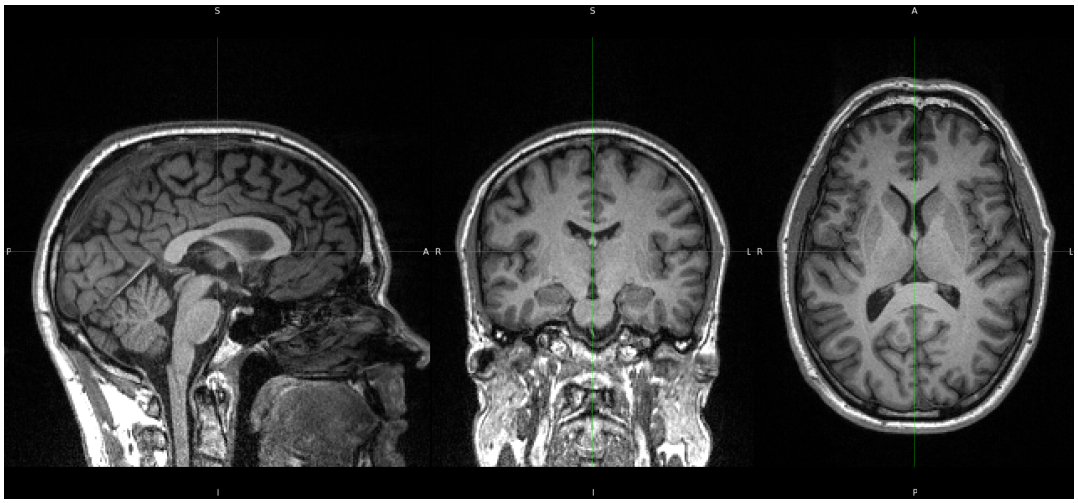

Figure 1. *3D T1-weighted MPRAGE recording from BGA_046*. Panels left to right: Sagittal, Coronal, Axial section, respectively.

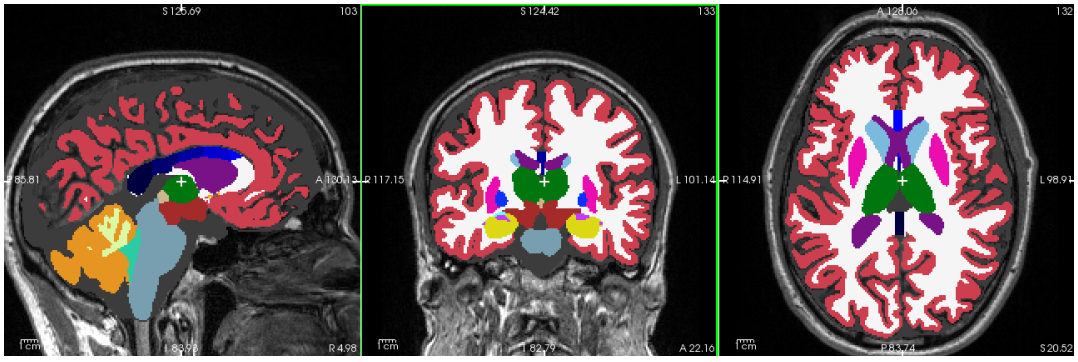

Figure 2. *The color-coded ASEG segmentation mask by FreeSurfer 7.4.1 overlaid on 3D T1-w MPRAGE from BGA_046*. Panels left to right: Sagittal, Coronal, Axial section, respectively. The white cross is located in the medial part of Left-Thalamus. Thalamus: green, Hippocampus: yellow, Caudate: light blue, Putamen: pink, Pallidum: purple, Cortex: red, White-Matter: white.

Brain segmentation - Freesurfer versions and improvements

For the neuroimaging analysis we employed *Freesurfer* (https://freesurfer.net), an open-source software suite widely used for processing and analyzing brain MRI images, being initially released in 1999 [31]. Specifically, we utilized the current Freesurfer version 7.4.1, which offers several improvements over the earlier 6.0 version, being the widely used version when Skrobisz et al. [24] published their study. Version 7.0 (released July 2020) enhanced the semantic segmentation with improved accuracy of subcortical regions. Relevant to our IBS investigations, version 7.4.1 (released June 2023), provides even better segmentation accuracy, particularly in regions such as the hippocampus and the amygdala. Furthermore, it offers improved compatibility with multi-modal imaging data, facilitating the integration of structural and functional MRI analyses. It also implements a refined longitudinal processing stream, valuable for tracking brain changes over time in IBS patients.

These improvements in version 7.4.1 were particularly relevant for our study, as they allowed for a more precise analysis of brain regions involved in pain processing, cognition, emotional regulation, and other possible gut-brain interactions associated with IBS. However, to assess and replicate the findings reported by Skrobisz et al. [24], we also processed all our T1-w recordings with Freesurfer 6.0 to document any undesirable effects of segmentation method and Freesurfer version in the

comparisons [32–37]. To be more specific, we used the brain regions provided by FreeSurfer's `aseg` mask reported in [24]) (see description in Table A1).

Data analysis

All statistical analyses and machine learning procedures were performed using Python (version 3.10). Further algorithmic and computational details are provided in our *Jupyter notebooks* available on the GitHub repository (https://arvidl.github.com/ibs-brain-morphometry) accompanying the present project. A p-value < 0.05 was considered statistically significant for all relevant statistical tests, and Bonferroni correction was included in cases of multiple comparisons. The non-parametric effect size measure Cliff's Depta was used to quantify the differences between the IBS and HC groups. The strength is interpreted as negligible (.00-.14), small (.15-.33), medium (.34-.47) and large effect (.48-1.00). We used Spearman's rank correlation, a non-parametric measure of the strength and direction of association between two variables. The strength is interpreted as weak (.20-.39), moderate (.40-.59), strong (.60-79), and very strong (.80-1.00).

SHOULD the A-E be included in the final manuscript?

Here, we present our analysis methods to investigate the following research questions.

A Is it possible to replicate the morphometric findings in Skrobisz et al. [24] regarding IBS vs. HC, using the same FreSurfer-derived features and the same FreeSurfer version?

- **(i)** By employing a feature-by-feature (univariate) comparison?

B Are there differences in morphometric feature values between FreeSurfer 6.0.1 and FreeSurfer 7.4.1 applied to the same set ($n = 78$) of 3D T1-w recordings in our cohort?

- **(i)** By employing a feature-by-feature comparison?
- **(ii)** Employing a multivariate comparison, incorporating covariance structures in the morphometric features?

C Is it possible to separate IBS individuals from HC based on morphometric features?

- **(i)** By employing a feature-by-feature comparison (FS 7.4.1)?
- **(ii)** Employing a multivariate comparison, incorporating covariance structures in the morphometric features?
- **(iii)** By predicting IBS vs. HC from the morphometric features using a machine learning framework (ML)?
- **(iv)** identifying feature importance of the morphometric measures in the model with the best prediction?

E Would adding cognitive performance as a predictor improve the accuracy of separating IBS from HC?

- **(i)** By employing a feature-by-feature comparison?
- **(ii)** Employing a multivariate comparison, incorporating covariance structures in the cognitive features?
- **(iii)** By predicting IBS vs. HC from the morphometric and cognitive features using a machine learning framework (ML)?
- **(iv)** identifying feature importance of the morphometric and cognitive measures included in the model with the best prediction?

Since our research questions address both univariate and not-so-common multivariate statistical methods, including training and testing predictive models in a machine learning framework, we find it expedient to describe and motivate these in more detail. As an initial step we performed simple explorative data analysis of all numerical features in our sample, similar to [24], and cross-tabulation of categorical variables (*Group*: `HC`, `IBS`; *Sex*: `F`, `M`). In the separation tasks **A(i)** - **E(i)** we performed t-tests for each measure, applied Bonferroni correction for multiple comparisons, and identified significant differences before and after correction. We used non-parametric tests (Mann-Whitney U) if the data were not normally distributed. Also, effect sizes (Cohen's d, using pooled variance), of Cliff's delta were computed to provide additional insight into the magnitude of differences.

In addition, we applied permutation testing, a non-parametric method that does not rely on assumptions about the underlying distribution of the data. In short, it calculates an observed test

statistic based on the actual group assignments (the sum of squared differences between group means across all variables). The test then randomly reassigns the group labels (e.g., HC and IBS, or F and M) to the data points many times (in our case 1000 times). For each permutation, it recalculates the test statistic. This creates a null distribution of test statistics that would be expected if there were no real differences between the groups. A p-value is calculated as the proportion of permuted test statistics that are as extreme as or more extreme than the observed test statistic. If the observed statistic is unusual compared to the permuted statistics (resulting in a low p-value), it suggests a significant difference between groups regarding the selected features. This permutation testing has several advantages: (i) It is robust to non-normality and can handle multivariate data (e.g., tasks **B(ii) - E(ii)**); (ii) it accounts for multiple comparisons inherently, and (iii) it is suitable for small sample sizes.

In tasks **B(ii) - E(ii)** we checked for multivariate normality using two common tests, Mardia's test (checking for skewness and curtosis) and Henze-Zirkler's test (a more powerful omnibus test), as implemented in *stats* in the *Scipy* Python library. For both tests, a p-value < 0.05 suggests rejection of the null hypothesis of multivariate normality. We also calculated (in Python) Partial η^2 (Eta-squared) as an effect size measure (e.g. [38]). This represents the proportion of variance in the dependent variables (morphometric measures) that is associated with the group membership (e.g., HC vs IBS, F vs M). Partial η^2 ranges from 0 to 1, where 0 indicates no effect and 1 indicates a perfect effect. Common guidelines for interpreting Partial η^2 are small effect: 0.01 to 0.06; Medium effect: 0.06 to 0.14; Large effect: 0.14 and above [39]. For all the comparison tasks **B(ii) - E(ii)**, we anticipated mild non-normality of single features. In this context, we choose *Generalized Mahalanobis distance* [40] as an effect size measure appropriate for data in a higher-dimensional feature space. Generalized Mahalanobis distance provides a single measure of the separation between two multivariate distributions, taking into account both the differences in means and the covariance structure of the data and has the following properties, relevant to our tasks: (i) it accounts for the covariance structure of the data, unlike Euclidean distance; (ii) it is scale-invariant, meaning it is not affected by the scale of individual variables; (iii) in cases where the pooled covariance matrix (weighted average of the covariance matrices of the two distributions) might be singular or near-singular and it's inverse not computable or numerically unstable, it uses the Moore-Penrose pseudoinverse instead of the regular inverse covariance Σ^{-1}; (iv) from the squared Mahalanobis distance it is possible to calculate p-values using the chi-square distribution; (v) in case of non-normality of single features, instead of relying on the chi-square distribution for p-values, we can use permutation testing to assess significance; (v) the Generalized Mahalanobis distance measure assumes that the covariance structure is similar between the two groups. If this assumption is violated, interpretations should be made with caution.

Prediction of class belonging using machine learning

In tasks **C(iii) - E(iii)** we applied a comprehensive machine learning framework, utilizing morphometric features derived from FreeSurfer (`aseg`) to develop predictive models for three distinct classification tasks. We employed *PyCaret* (https://pycaret.org), an open-source, low-code machine learning library in Python, to develop and evaluate our classification models.

Model Development

Our model development process began with data preparation, where we divided our dataset into training (70%) and testing (30%) sets using stratified sampling to maintain class distributions across both sets. This stratification ensured that the proportions of different classes (e.g., IBS vs. HC, female vs. male, cognitive function levels) were approximately equal in both the training and testing sets, which is crucial for maintaining the integrity of our model evaluation.

We then made use of the comprehensive suite of machine learning algorithms implemented in PyCaret to evaluate multiple classification models. These included, but were not limited to, Logistic Regression, Random Forest, Support Vector Machines (SVM), Gradient Boosting Machines (such as XGBoost and LightGBM), and K-Nearest Neighbors (KNN). The diversity of these algorithms allowed us to explore different approaches to classification, from linear models to complex ensemble methods.

To ensure robust model performance and mitigate overfitting, we employed 10-fold cross-validation. This process involved dividing the training set into 10 equal subsets, training the model on 9 subsets, and validating on the remaining subset. This procedure was repeated 10 times, with each subset serving as the validation set once. Performance metrics were then averaged across all 10 iterations, providing a more reliable estimate of model performance.

We further refined our models through hyperparameter tuning, utilizing PyCaret's built-in optimization functionality. This process employed techniques such as random search or grid search with cross-validation to identify the optimal combination of model parameters. The specific hyperparameters tuned varied depending on the algorithm but generally included parameters such as the number of trees in Random Forest, the regularization strength in Logistic Regression, or the kernel type in SVM.

Model Evaluation

We assessed the performance of each classifier using a comprehensive set of metrics to provide a holistic view of model performance. Accuracy, which represents the proportion of correct predictions among the total number of cases examined, served as our primary metric. However, recognizing that accuracy alone can be misleading, especially in cases of class imbalance, we supplemented this with additional metrics.

The F1 score, calculated as the harmonic mean of precision and recall, provided a balanced measure of the model's performance, particularly useful in scenarios where false positives and false negatives have similar costs. We also computed the Receiver Operating Characteristic - Area Under Curve (ROC-AUC), which quantifies the model's ability to distinguish between classes. This metric is particularly valuable for imbalanced datasets and provides insight into the model's performance across various classification thresholds.

To visualize the models' performance, we generated confusion matrices for each classification task. These matrices offer a representation of true positives, true negatives, false positives, and false negatives, allowing for a nuanced understanding of where each model excels or struggles in its predictions.

For the three-class cognitive function classification task, we employed macro-averaged versions of these metrics. This approach calculates the metric independently for each class and then takes the average, treating all classes equally. This is particularly important in multi-class problems where class imbalance may be present.

All performance metrics were calculated on both the cross-validated training set and the held-out test set. This dual evaluation allowed us to assess both the model's ability to learn from the training data and its generalization to unseen data. We also calculated 95% confidence intervals for all relevant metrics to provide a measure of the precision of our estimates.

Feature Importance Analysis

To gain insights into the contribution of different morphometric features to our models, we conducted feature importance analyses using two complementary approaches: permutation importance and SHAP (SHapley Additive exPlanations) values. Permutation importance provides a measure of a feature's importance by randomly shuffling its values and observing the consequent decrease in model performance. This process was repeated multiple times for each feature, and the average decrease in performance was used as the importance score. Features that, when permuted, led to a larger decrease in performance were deemed more important. This method has the advantage of being model-agnostic, allowing for consistent comparison across different types of models. SHAP values, based on coalitional game theory, offer a unified measure of feature importance that provides both global and local interpretability [41]. For global interpretability, we aggregated SHAP values across all instances to identify the overall importance of each feature. For local interpretability, we examined SHAP values for individual predictions, allowing us to understand how each feature contributed to specific classifications. This dual perspective provided a nuanced understanding of feature importance, capturing both overall trends and instance-specific explanations.

We applied these feature importance analyses to our best-performing models for each classification task. The results were visualized using summary plots, which display the distribution of SHAP values for each feature, and feature importance rankings based on both permutation importance and mean absolute SHAP values. These visualizations allowed for easy identification of the most influential morphometric features for each classification task. By combining these two approaches to feature importance, we aimed to provide a robust and comprehensive understanding of which brain regions, as captured by FreeSurfer morphometric features, were most predictive in our classification tasks. This analysis not only enhanced the interpretability of our models but also provided valuable insights into the neurobiological underpinnings of IBS, sex/gender differences, and cognitive function.

Results

Characteristics of the sample

The median ages of the IBS and HC groups were in the early thirties, with a wider inter-quartile range in the HC than the IBS group. More women than men were included in the IBS (n = 38/49) and the HC group (n = 20/29). IBS-SSS scores, with imputation for three participants in each group, were much higher in the former group, as expected from the inclusion criteria used in the present study.

Group	Age Median (IQR)	IBS_SSS Median (IQR)	Gender F/M (%)	N	Missing Age	Missing IBS_SSS
HC	33.0 (23.0)	21.0 (30.0)	69.0% (F)/31.0% (M)	29	0	3
IBS	34.0 (14.0)	264.0 (95.0)	77.6% (F)/22.4% (M)	49	0	3

Table 2. Demographic and clinical characteristics. Source: `al-38-BGA-ML-X-Morph-Cognition-y-IBS.ipynb`

Morphometric measures in [24] and the FSv6.0 measures in the present study

Table 1 shows the mean and standard deviations on the selected volume measures in the study by Skrobisz et al. and the results in the present study, analyzed by FSv6.0 and FSv7.4.1, respectively. Overall, the results corresponded well when comparing results from the FSv6.0, with some noteworthy differences when analyzed with v7.4.1. *MUST BE DESCRIBED*.The correspondence between the two Freesurfer versions in the present study is illustrated in Figure 2.

Table 2: Add OR REPLACE the volume measures of FSv7.4.1 to FSv6.0. the measures in Skrobisz's study.

Table 3. Table tab:brain-regions-full: Comparison of Brain Region Volumes in IBS Patients and Healthy Controls

Brain Region	Skrobisz Cohort (FS 6.0)				Bergen Cohort (FS 7.4.1)			
	HC (N=19)		IBS (N=20)		HC (N=29)		IBS (N=49)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
BrainSegVol	0.75340	0.01784	0.74913	0.01647	0.79889	0.03855	0.79087	0.02974
BrainSegVolNotVent	0.74137	0.01880	0.73857	0.01836	0.78583	0.03869	0.77607	0.03147
CC_Anterior	0.00062	0.00010	0.00061	0.00008	0.00059	0.00010	0.00062	0.00009
CC_Central	0.00039	0.00009	0.00043	0.00008	0.00036	0.00008	0.00036	0.00009
CC_Mid_Anterior	0.00041	0.00009	0.00044	0.00013	0.00036	0.00007	0.00038	0.00010
CC_Mid_Posterior	0.00038	0.00007	0.00036	0.00007	0.00035	0.00007	0.00037	0.00008
CC_Posterior	0.00065	0.00013	0.00065	0.00010	0.00063	0.00010	0.00067	0.00011
CSF	0.00061	0.00009	0.00060	0.00012	0.00071	0.00013	0.00074	0.00013
CerebralWhiteMatterVol	0.30205	0.01500	0.30133	0.01461	0.30679	0.01944	0.30158	0.01899
CortexVol	0.30829	0.01298	0.30780	0.01715	0.34118	0.02108	0.33789	0.01968
Left-Accumbens-area	0.00031	0.00005	0.00034	0.00006	0.00044	0.00007	0.00043	0.00007
Left-Amygdala	0.00118	0.00013	0.00113	0.00015	0.00125	0.00014	0.00120	0.00011
Left-Caudate	0.00239	0.00025	0.00228	0.00021	0.00251	0.00031	0.00247	0.00032
Left-Cerebellum-Cortex	0.03628	0.00302	0.03553	0.00256	0.03744	0.00363	0.03723	0.00357
Left-Cerebellum-White-Matter	0.00992	0.00113	0.00971	0.00107	0.01089	0.00108	0.01069	0.00104
Left-Hippocampus	0.00270	0.00021	0.00272	0.00020	0.00297	0.00027	0.00293	0.00024
Left-Pallidum	0.00140	0.00012	0.00135	0.00010	0.00142	0.00016	0.00137	0.00010
Left-Putamen	0.00336	0.00033	0.00324	0.00028	0.00370	0.00042	0.00362	0.00037
Left-Thalamus	0.00511	0.00037	0.00500	0.00024	0.00538	0.00052	0.00526	0.00047
Right-Accumbens-area	0.00034	0.00004	0.00036	0.00005	0.00050	0.00006	0.00051	0.00006
Right-Amygdala	0.00125	0.00012	0.00120	0.00012	0.00133	0.00012	0.00132	0.00012
Right-Caudate	0.00244	0.00024	0.00236	0.00024	0.00261	0.00031	0.00258	0.00030
Right-Cerebellum-Cortex	0.03652	0.00321	0.03616	0.00264	0.03847	0.00372	0.03818	0.00365
Right-Cerebellum-White-Matter	0.00908	0.00106	0.00922	0.00100	0.01022	0.00101	0.01023	0.00095
Right-Hippocampus	0.00282	0.00022	0.00285	0.00021	0.00310	0.00025	0.00303	0.00024
Right-Pallidum	0.00136	0.00012	0.00133	0.00010	0.00135	0.00016	0.00133	0.00012
Right-Putamen	0.00336	0.00030	0.00330	0.00028	0.00379	0.00042	0.00377	0.00038
Right-Thalamus	0.00488	0.00030	0.00475	0.00024	0.00547	0.00047	0.00534	0.00046
SubCortGrayVol	0.03930	0.00194	0.03855	0.00162	0.04292	0.00319	0.04221	0.00274
TotalGrayVol	0.42105	0.01376	0.41884	0.01868	0.45999	0.02859	0.45568	0.02574
WM-hypointensities	0.00047	0.00015	0.00048	0.00013	0.00076	0.00064	0.00061	0.00027
lhCerebralWhiteMatterVol	0.15101	0.00748	0.15058	0.00742	0.15387	0.00961	0.15155	0.00931
lhCortexVol	0.15339	0.00620	0.15313	0.00871	0.17057	0.01050	0.16899	0.01011
rhCerebralWhiteMatterVol	0.15103	0.00757	0.15075	0.00727	0.15292	0.00990	0.15003	0.00975
rhCortexVol	0.15490	0.00690	0.15467	0.00859	0.17061	0.01067	0.16890	0.00965

Note: All volumes are normalized to estimated total intracranial volume (eTIV).
HC = Healthy Controls; IBS = Irritable Bowel Syndrome; SD = Standard Deviation.

Generated in `al-22-BGA-brain-morphometry-comparisons.ipynb`

Univariate comparisons between HC and IBS in the FSv7.4.1

Non-parametric tests with Bonferroni correction showed no statistically significant differences between the HC and the IBS group. Analyses with sex as the grouping factor showed a significantly smaller value in women than men in the eTIV ($p < .001$). The difference in the left hippocampus was close to statistical significance after Bonferroni correction ($p = .0005$). All the other differences were statistically non-significant.

`Must be reproduced in the notebook.`

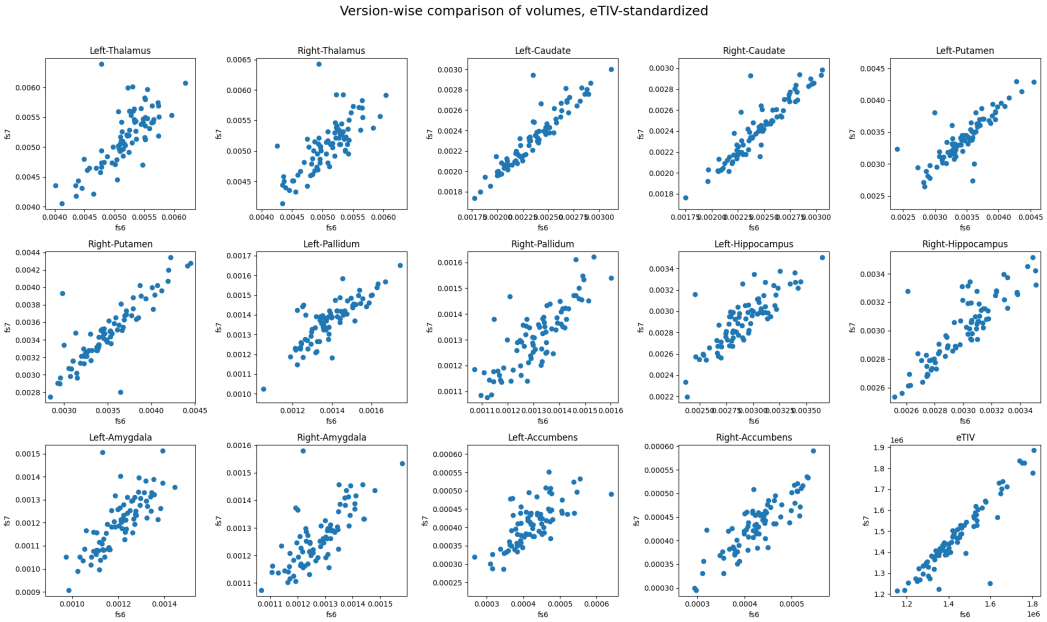

Figure 3. Comparisons between Fressurfer v6.0 and v7.4.1 on a selection of variables

Multivariate analyses: IBS versus HC

A generalized Mahalanobis distance, a partial η^2 (Eta-squared), and a permutation test were computed to reveal how distinct the two groups (IBS and HC) are in a multidimensional space. Analog to the univariate analyses, no significant differences were observed: general Mahalanobis distance (= 2.0427), Partial η^2 (= 0.013), the permutation test (p = .45 (4515). From a clinical perspective, these results indicate that the brain morphometry of IBS patients, as measured by these variables, does not differ substantially from that of healthy controls.

Prediction of IBS versus HC from morphometric measures

The training set included 54 participants (43 females), 34 from the IBS group. Among the 24 participants allocated to the test set (15 women), 15 participants were from the IBS group.

Extreme Gradient Boosting (xgboost) was selected as the best model, with an accuracy of 0.7200 in the train-set, with a reduction (accuracy = 0.5000) in the test set. The confusion matrix, Figure 4 (b) shows that the majority of the IBS patients in the test set (73%) were correctly classified (11/15), while all but one HC was misclassified as IBS. The misclassified IBS patients were somewhat older than the correctly classified patients, with a higher and more distributed IBS-SSS score.

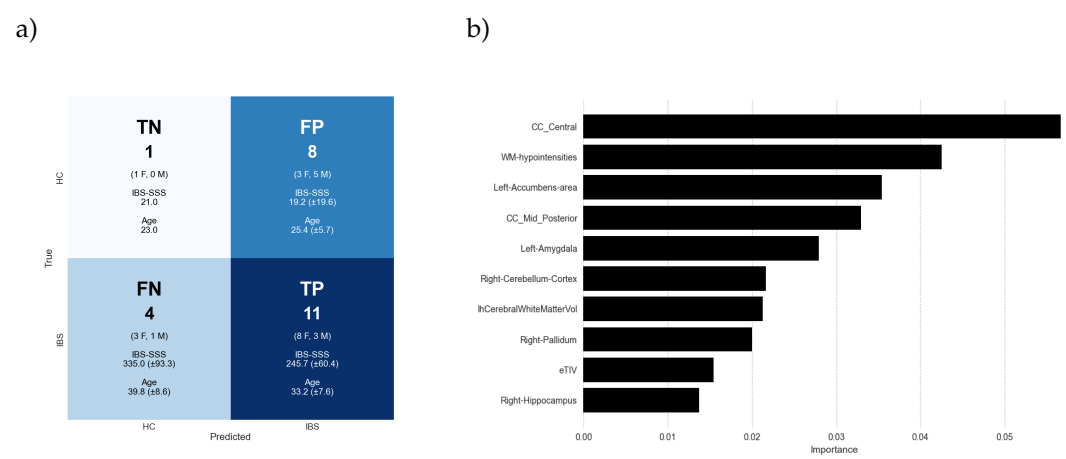

Figure 4. (a) Confusion matrix (XGBOOST) predicting IBS versus HC from brain morphometry.(b) Top 10 feature importance (permutation importance, 100 repeats) with Group (IBS vs HC) as the outcome variable and morphometric features as predictors

Figure 4 displays the permutation importance of the various brain features for a list of the 10 most important features, as indicated by the importance values on the y-axis. Positive features suggest that variations in these regions contribute most significantly to the model's predictive power. Two sections of the corpus callosum have a strong importance in the model, together with the accumbens and amygdala in the left hemisphere, and the cerebellum cortex, cerebral white matter, pallidum, and hippocampus in the right hemisphere. We also find white matter hypointensities and eTIV on the list of important features.

Univariate analysis of the cognitive features

Table 3 shows that the HC group obtained significantly higher scores on the full-scale RBANS measure, even after Bonferroni correction, with a medium effect size. Non-parametric comparisons were used to take into account non-normality distributions for several index scores. Scores for the Visuospatial Index and the Attention Index were non-significant, with negligible effect sizes. For the two memory indexes, the Immediate Memory and Recall Indexes, the HC group scored significantly higher than the IBS group. The statistical significance was retained only for the Recall index, with a small effect size.

Variable	HC	IBS	p-value	Cliff's delta
Fullscale_RBANS	103.0 (93.0-108.0)	91.0 (85.0-100.0)	0.002	0.213
Memory_Index	100.0 (86.0-109.0)	86.0 (81.0-105.0)	0.031	0.147
Visuoaspatial_Index	97.0 (90.0-107.0)	96.0 (90.0-105.0)	0.763	0.021
Verbal skills Index	105.0 (95.0-113.0)	95.0 (89.0-111.0)	0.087	0.116
Attention_Index	98.0 (89.0-108.0)	97.0 (83.0-101.0)	0.118	0.107
Recall_Index	107.0 (92.0-113.0)	95.0 (85.0-100.0)	0.006	0.186

Table 4. A non-parametric analysis comparing cognitive features in the IBS and HC groups. Cliff's delta is used to estimate effect sizes.

Correlations between morphometric and cognitive features

Figure 5 shows a heatmap to illustrate correlations between the included morphometric volume measures, the RBANS indexes, and the Fullscale RBANS score. The correlations between the Fullscale score and each of the five indexes are moderate to strong for all expect the Verbal skills index. Furthermore, there is a strong correlation between the two memory scores (Immediate memory and Recall). Weak correlations were found between the Recall index and hippocampus, (Amygdla, Pallidum, and Accumbance. correlations

The correlations between the morphometric features are found to be moderate to weak, with the strongest positive correlations when a given volume measure is measured both in the right and left hemispheres, between subcortical parts of the brain, and between all and the total gray volume measure.

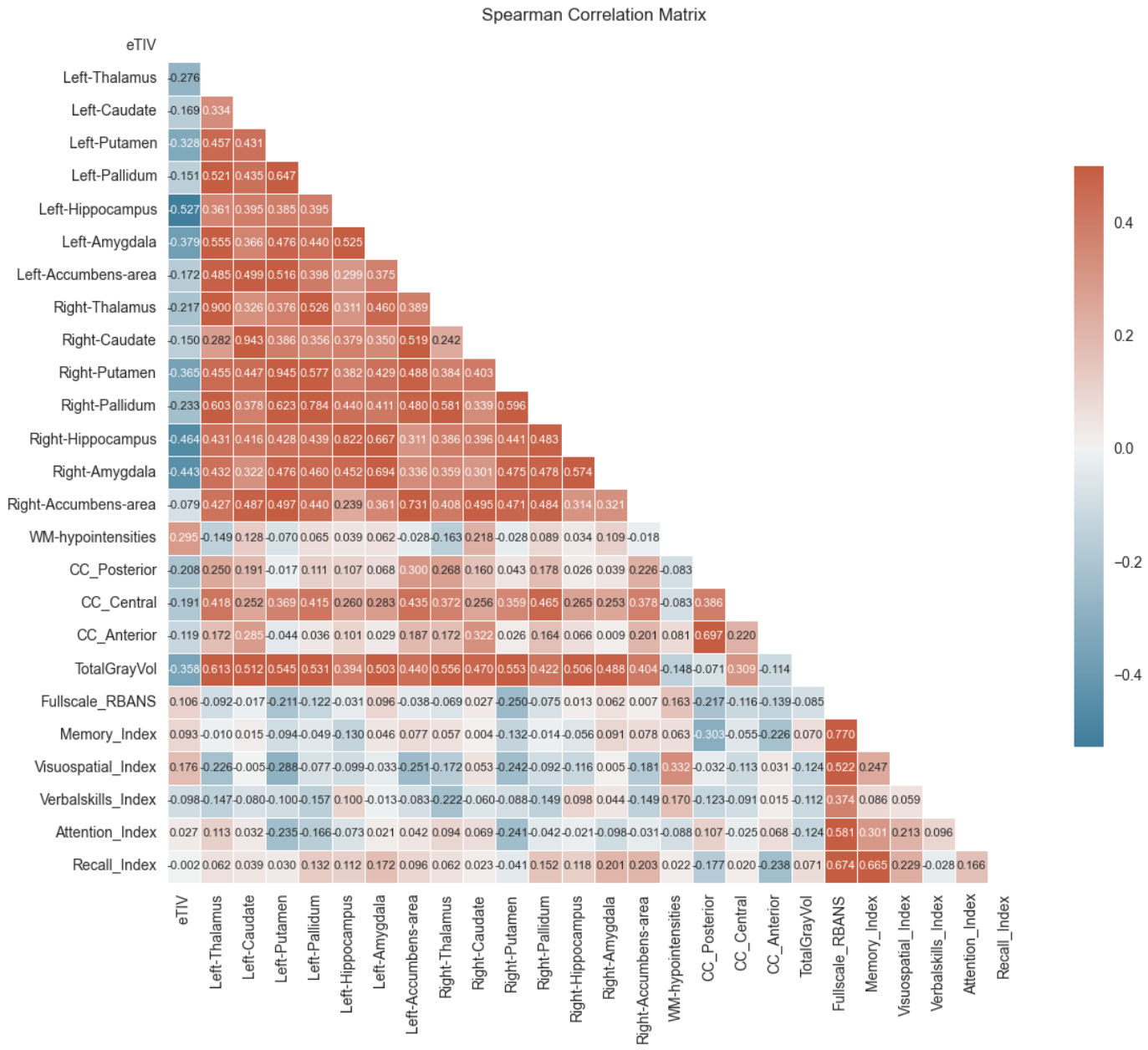

Figure 5. Pairwise correlations between the morphometric and cognitive variables

Prediction of IBS versus HC from morphometric and cognitive measures

The Extreme Gradient Boosting (xgboost) was also selected as the best model when cognitive features (the five indexes) were included as predictors together with the morphometric volume measures. An accuracy of 0.6500 in the train set improved to 0.6667 in the test set, and the confusion matrix shown in Figure 6 (a) indicates an improved generalization when cognitive features were included. The Figure shows that 93% of the IBS patients (14/15) were correctly classified (one

misclassified man in the older age range). However, the specificity was lower, in that 78%, which included all men in the HC group, were misclassified as IBS patients.

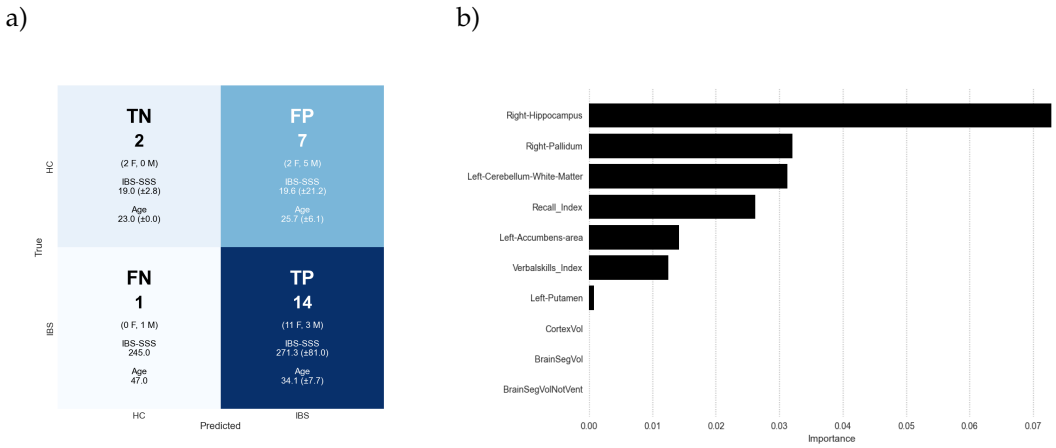

Figure 6. (a) Confusion matrix (XGBOOST) predicting IBS versus HC from brain morphometry and cognition. (b) Top 10 feature importance (permutation importance, 100 repeats) with Group (IBS vs HC) as the outcome variable and morphometric and cognitive features as predictors

Figure 6 (b) displays the permutation importance of the various morphometric and cognitive features. At the top, we find Hippocampus and Pallidum in the right, followed by Cerebellum White Matter and Accumbens in the left hemisphere. Interestingly, both the Recall and Verbal Skills Indexes from RBANS are found among the six features given importance in the model. Zero importance was given to the three features printed at the bottom of the figure.

The SHAP values shown in Figure 7 show how different brain morphometric features when the cognitive features are included (a) or excluded (b). When both morphometry and cognition were included, the Verbal Skills and the Recall indexes from RBANS are identified among the most differentiating factors, supporting an important role of cognitive function in patients with IB, represented by the blue points pushing them towards low values on the indexes. This is also true for the hippocampus in the right hemisphere, obtaining the best differentiation between the two groups, followed by the caudate, putamen, and pallidum (right hemisphere). For all these three morphometrics, higher volumes (red) push towards one group, while lower volumes (blue) lean towards the other group. From the unvariate analysis, we know that the blue represents lower results in the IBS group. Attention should also be given to cerebral and cerebellar white matter volumes.

The right hippocampus and caudate, putamen, pallidum, and amygdala are also found to differentiate well between the IBS and HC groups based on the morphometric features. This analysis also included accumbens, corpus callosum_central, and not at least cerebellum white matter. By this, the hippocampus, caudate, and putamen in the right hemisphere, and cerebellum white matter in the left hemisphere turned out to be the most important morphometric features independent of the inclusion of the cognitive features. The most interesting finding was the strength of two of the cognitive features, the Verbal skills and Recall indexes based on performance on a set of tests from RBANS.

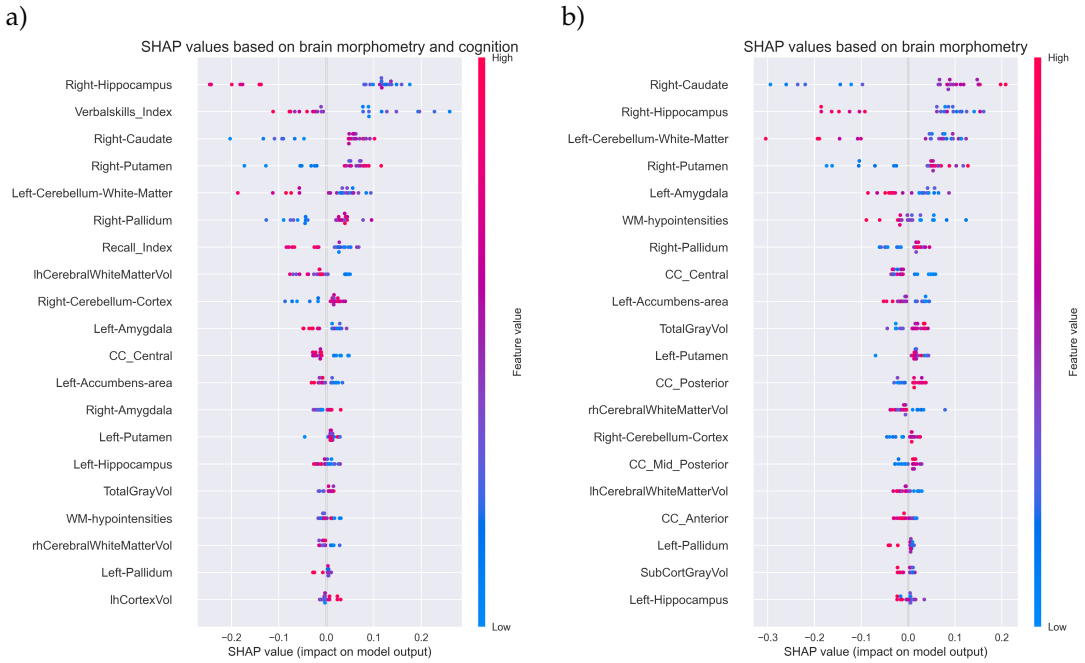

Figure 7. (a) SHAP values morphometry with cognition (b) SHAP values morphometry without cognition

Note: Y-axis (Features): This lists the brain morphometric features (e.g., left hippocampus, cerebral white matter volume, etc.) used by the model. X-axis (SHAP values): The SHAP values indicate the impact each feature has on the model's output, which is the probability of classifying someone as male or female. Values to the right of 0 indicate a positive contribution (towards one class, likely HCs in this case). Values to the left indicate a negative contribution (likely towards IBS). Color gradient (Feature Value): This represents the actual feature value: Red/pink indicates a high feature value. Blue/purple indicates a low feature value.

1. Discussion

Summary of the results

The group difference in thalamus volumes between the IBS and HC group reported by Skrobisz et al. [24] was confirmed in the present dataset, neither when using the FV6.0 or FV7.4.1.

The multivariate separation of participants with IBS and in the IBS and HC, as well as the overall accuracy of predicting IBS versus HS, was found to be weak in our study. This indicates significant morphometric overlap between the two groups and aligns with the heterogeneous nature of IBS and the challenge of identifying consistent biomarkers of the condition [17]. These findings highlight the complexity of IBS, where morphometric features alone may not provide sufficient discriminatory power to reliably separate individuals with IBS from healthy individuals. However, when cognitive features were included alongside morphometric features as predictors, the ability to correctly classify patients with IBS improved. Notably, only one patient in the test set was misclassified as an HC. Furthermore, although the overall prediction accuracy remains modest, we believe that the feature importance analysis offers insights for future research directions. The analyses revealed that a wide range of morphometric features, including cortical, subcortical, and white matter structures, were important in predicting IBS versus HC. This broad involvement of multiple brain systems supports the notion that IBS may be linked to widespread neuroanatomical alterations. When cognitive features were included as predictors, we could show that two of these cognitive measures emerged as particularly important predictors together with the hippocampus and other subcortical structures. Importantly, the discriminative power of these features in differentiating IBS from HC was confirmed by a SHAP analysis, a method based on graph theory. This consistency across multiple analyses underscores the potential of combining morphometric and cognitive features to identify patients with IBS. While the study design does not allow for firm conclusions regarding brain function, the combination of findings suggests that these key features, both cognitive and morphological, are likely involved in emotional regulation, stress and pain responses, and cognitive performance.

Morphometric findnings

The strong importance of subcortical structures was a primary finding in the present study, particularly the basal ganglia. This is in accordance with previous findings REFS. The basal ganglia, including putamen and caudate, are involved in motor control but also play a role in reward processing and habitual responses. From this, one should also expect involvement in the anticipation and modulation of pain. In IBS, changes in areas of the basal ganglia might contribute to the abnormal processing of pain and discomfort, influencing both the severity and perception of symptoms. The involvement of amygdala's morphology may be important due to its connectivity with other regions of the brain, such as the prefrontal cortex and insula, which are involved in the modulation of pain and emotional responses. This could potentially exacerbate the experience of visceral pain and the emotional disturbances associated with IBS.

Our findings of altered accumbens volumes are also intriguing, as they suggest potential alterations in reward processing and motivation in IBS. This could relate to altered pain processing and visceral hypersensitivity often observed in IBS patients. Dysregulated reward processing may also contribute to amplifying the emotional burden of chronic pain and discomfort in IBS.

A strong importance was given to the hippocampus in the right hemisphere, both without and with the inclusion of cognitive features. Hippocampus has been linked to memory processes. MORE. The inclusion of verbal skills as the second feature after the hippocampus in the predictive mode was more surprising. There is a growing awareness of the relationship between memory and verbal skills, in that the latter may be of importance to the ability to store, retrieve, and organize information, in a context of semantic memory tasks, storytelling, and episodic memory. Most studies on this relation have, however, focused on the hippocampus in the left hemisphere. When we find the right hippocampus, we have probably touched upon the collaborative role of the right and left hippocampus. Although shown to be more active during tasks involving spatial navigation, visual memory, and contextual recall, most memory tasks require both verbal and non-verbal memory, particularly memory tasks requiring complex contextual integration. Furthermore, damage to the right hippocampus is shown to lead to problems with narrative storytelling and difficulties in retrieving non-verbal contextual details. See [42]. Furthermore, more recent studies have shown a more differentiated picture of the subfields of the hippocampus than a separation between left and right [43]. We will therefore conclude that morphometric measures of hippocampus seem to play a key role in characteristics of the symptomatology of IBS and other DGBis.

The present study did not confirm the importance of Thalamus in separating individuals with IBS from healthy controls, as previously reported [24]. This replication failure might be explained by methodological considerations. In contrast to Skrobisz and colleagues [24], the current study investigated a considerably larger sample of exclusively IBS patients and compared them to matched healthy controls, which renders the findings more reliable. Furthermore, the results of Skrobisz were much stronger for patients with IBD than IBS. Even without support by the present study, one should Still, we assume that the thalamus is an important structure when it comes to understanding IBS and other disorders of the gut-brain interaction. A study reporting on the microstructural organization of the thalamus using diffusion tensor imaging found lower fractional anisotropy and higher mean diffusivity in the thalamus in IBS patients, indicating less directional integrity or organization within the fibers, as well as a loss of tissue integrity in the thalamus of IBS patients [44]. Other studies have shown the importance of including different imaging and clinical variables in studies of IBS. A study be Batt et

al. [45] which aimed to predict improvement in IBS-SSS score in women, used a multidisciplinary phenotypic approach including patient data that included multi-modal neuroimaging, behavioral testing, and clinical self-report questionnaires. From this, they hypothesized that an integrated neuropsychosocial signature may enable the classification of IBS patients into those who improve over time and those who do not.

Sjekk også Hellgren et al. [46]: A Swedish study found that a significant number of patients had long-term neurocognitive abnormalities months after recovering from COVID-19. The presence of white matter lesions in the brain was also a common finding. They used RBANS and found below cut-off performance (2 SD) for most patients where immediate memory and delayed memory (no control group).

Cognitive function in patients with IBS

In the present study, we used a multidimensional approach to predicting IBS by incorporating cognitive performance measures alongside morphological brain imaging findings. We found that the inclusion of cognitive measures improved the prediction accuracy for IBS, likely because these measures helped to explain the functional correlates of the brain morphometry. Cognitive assessments are better suited to capturing the impact of psychological and emotional dysregulation, thereby increasing the overall accuracy of predictive models by addressing both the neurological and behavioral components of IBS.

By integrating brain structural data with results from performance on cognitive tests, we contributed by showing links between brain morphology and functional manifestations, and by this offering a model how future studies including Larger data sets can create more sensitive and holistic representation of the brain-gut axis. On particular, cognitive measures may reflect function in brain regions crucial to pain processing, cognitive control, and emotion regulation, all of which are areas where impairment represents key symptoms in patients with IBS. Thus, this integrative approach may not only improve the predictive capability of statistical models but also underscore the role of functional brain regions in mediating IBS symptoms.

In future studies, the inclusion of larger and more diverse datasets incorporating both cognitive and neuroimaging measures could lead to the development of more sensitive, and clinically meaningful predictive models. Such models may capture the complexity of the brain-gut axis, and further substantiate the notion that IBS is not merely a gastrointestinal condition, but a disorder intricately connected to neural and behavioral networks involved in pain, stress responses, and emotional regulation.

Additionally, by integrating more features of psychological distress into neurobiological models, we may improve clinical interventions, targeting both the cognitive-emotional-social behavioral and neurobiological aspects of IBS, potentially leading to personalized treatment strategies that take into account the wide range of dysregulations often observed in patients with IBS.

In the present study, we observed the importance of Hippocampus volume in separating IBS from HC. With a close link to the cognitive function [47], this may explain why indexes reflecting verbal skills and memory processing were identified as strong together with a volume measure of hippocampus.

Relation to Brain-Gut Axis and Recent IBS Research

The present findings align with the complex nature of the brain-gut axis [18,48], and its critical role in IBS. The involvement of structures such as the hippocampus and other subcortical regions in both cognitive and emotional regulation may reflect the psychological and neurological impact on IBS symptomatology. This may also provide a perspective on how the brain may influence the physiological GI-related manifestations of IBS and vice versa. Although the combination of morphometric and cognitive data may contribute to unraveling the complex neurobiological underpinnings of IBS, there is definitely a call for future studies, integrating data presented in the present study with functional neuroimaging, gut microbiome data, and a wider range of clinical symptoms to develop a more comprehensive understanding of IBS pathophysiology similar to the recent study by Li et al. [15]. Additionally, longitudinal studies will be crucial to determine whether these brain changes are a cause or consequence of IBS and to track their evolution over time. Our findings support the growing body of evidence suggesting that IBS involves alterations in brain structure and function. The observed differences in cortical volumes, and subcortical structures (particularly the accumbens and amygdala), align with the current understanding of IBS as a disorder of brain-gut interactions.

Recent studies that highlighted the role of the brain-gut axis in IBS?

- reports of alterations in brain networks involved in sensory processing and salience detection in IBS patients, which aligns with our findings of differences in accumbens and amygdala volumes.
- [17] emphasized the role of altered brain-gut interactions in IBS, which is supported by our observation of widespread structural changes across multiple brain systems.
- studies highlighting sex-specific differences in brain responses to visceral stimuli in IBS, which is consistent with the sex differences we observed in brain morphometry.
- studies reviewing the complex interplay between gut microbiota, brain structure, and function in IBS, supporting our findings of multi-system involvement.

Our exploratory data analysis of the FS_741 cohort reveals significant differences in brain morphometry between IBS patients and healthy controls, as well as important sex differences. These findings support the current understanding of IBS as a disorder involving altered brain-gut interactions, they highlight the complexity of the condition, and they emphasize a sex-specific view on brain-gut interactions. The last point is supported by a proposal by [49] suggesting that the origin and maintenance of IBS symptoms are driven by enhanced sensory sensitivity in women compared to men. Morphometric differences between the sexes might constitute one contributing factor to the proposed differences in sensory sensitivity in IBS patients.

Future research should focus on integrating these structural findings with functional neuroimaging, gut microbiome data, and clinical symptoms to develop a more comprehensive understanding of IBS pathophysiology similar to the recent study by Li et al. [15] who included a very large grpiå of åarticipants from the UK Biobank. Additionally, longitudinal studies will be crucial to determine whether these brain changes are a cause or consequence of IBS and to track their evolution over time.

Methodological considerations

In the present explorative study we included multivariate analyses to capture the complexity and interactions between multiple brain regions and cognitive measures. It should also be more coherent to the multidimensional nature of brain structure and function than results from univariate analyses. The sample size was a severe restriction of the present study, in that multivariate models generally require larger sample sizes. To reduce the risk of spurious and non-generalizable results in the present study, we included a 10 fold validation method and testing of a machine learning model trained on a hold-out part of the data-set.

Longitudinal studies are also essential. By combining brain structure with cognitive profiles in a longitudinal analysis can help predict the trajectory of the disorder (e.g., whether it will worsen or improve) and provide a better basis for personalized treatment. This can be especially useful for developing multimodal treatments that address both physical and psychological symptoms, such as the dietarian LowFODmap intervention.

1.1. Strengths and limiatations

The sample size is an important limitation of the present study. Interestingly, still some of our results, e.g., the importance of subcortical structures like putamen, pallidum, caudatus and hippocampus was also shown in the large study presented by Li et al. [**?**]

The sample size of the present study is too small to enable sex/gender based analyses. such studies are definitely called for, not at least because there is particularily little knowledge regarding the medical problems of LGBTQ+. Recently, several studies have pointed to sex-gender differences in health care, not only in the way symptoms occur but also in risk factors and the way women and men respond to a given treatment [50]. It is therefore important to take into account gender in the prevention, management, and examination of diseases. It is for example known that IBS with constipation (IBS-C) is significantly more common in women, whereas IBS with diarrhea (IBS-D) is more common in men, and that this can be explained by a wide range of biological and environmental factors [51]. In a book on sex/gender-specific medicine, NayongKim refers to a wide range of psychosocial factors when explaining gender-related factors associated with a disorder like iBS [52]. Others have pointed to an effect of sex hormones, where especially estrogen and progesterone may affect bowel function and gastrointestinal transit time, as well as how pain stimuli are processed in the central nervous system [53]. It has even been shown that the gastrointestinal tract microbiota may differ in women and men [54].

The severity of IBS symptoms should have been included. A limitation of the present study. In the paper by Li et al. [15], individuals with more severe IBS symptoms exhibited poorer cognitive performance, and lower brain volumes in regions related to emotion and cognition.

2. Conclusions

In summary, research into brain morphology in IBS shows the importance of accounting for sex. These findings help to elucidate the neurological underpinnings of sex differences in the gut-brain axis dysfunction in IBS, particularly concerning pain perception, emotional regulation, and stress response.

`TO BE CONTINUED`

References

1. Black, C.J.; Ford, A.C. Global burden of irritable bowel syndrome: trends, predictions and risk factors. *Nature Reviews Gastroenterology & hepatology* **2020**, *17*, 473–486. https://doi.org/10.1038/s41575-020-0286-8.
2. Lovell, R.M.; Ford, A.C. Global prevalence of and risk factors for irritable bowel syndrome: a meta-analysis. *Clinical Gastroenterology and Hepatology* **2012**, *10*, 712–721. https://doi.org/10.1016/j.cgh.2012.02.029.
3. Bonetto, S.; Fagoonee, S.; Battaglia, E.; Grassini, M.; Saracco, G.M.; Pellicano, R. Recent advances in the treatment of irritable bowel syndrome. *Polish Archives of Internal Medicine* **2021**, *131*, 709–715.
4. Drossman, D.A.; Tack, J. Rome Foundation clinical diagnostic criteria for disorders of gut-brain interaction. *Gastroenterology* **2022**, *162*, 675–679. https://doi.org/10.1053/j.gastro.2021.11.019.
5. Heitkemper, M.M.; Cain, K.C.; Jarrett, M.E.; Burr, R.L.; Hertig, V.; Bond, E.F. Symptoms across the menstrual cycle in women with irritable bowel syndrome. *Official journal of the American College of Gastroenterology | ACG* **2003**, *98*, 420–430.
6. Meleine, M.; Matricon, J. Gender-related differences in irritable bowel syndrome: potential mechanisms of sex hormones. *World Journal of Gastroenterology: WJG* **2014**, *20*, 6725.
7. Kim, Y.S.; Kim, N. Sex-gender differences in irritable bowel syndrome. *Journal of neurogastroenterology and motility* **2018**, *24*, 544.
8. Lovell, R.M.; Ford, A.C. Effect of gender on prevalence of irritable bowel syndrome in the community: systematic review and meta-analysis. *Official journal of the American College of Gastroenterology | ACG* **2012**, *107*, 991–1000.
9. Toner, B.B.; Akman, D. Gender role and irritable bowel syndrome: literature review and hypothesis. *Official journal of the American College of Gastroenterology | ACG* **2000**, *95*, 11–16.
10. Lundervold, A.J.; Billing, J.E.; Berentsen, B.; Lied, G.A.; Steinsvik, E.K.; Hausken, T.; Lundervold, A. Decoding IBS: a machine learning approach to psychological distress and gut-brain interaction. *BMC gastroenterology* **2024**, *24*, 267.
11. Shiha, M.G.; Aziz, I. Physical and psychological comorbidities associated with irritable bowel syndrome. *Alimentary Pharmacology & Therapeutics* **2021**, *54*, S12–S23. https://doi.org/10.1111/apt.16589.
12. Lam, N.C.Y.; Yeung, H.Y.; Li, W.K.; Lo, H.Y.; Yuen, C.F.; Chang, R.C.C.; Ho, Y.S. Cognitive impairment in irritable bowel syndrome (IBS): a systematic review. *Brain Research* **2019**, *1719*, 274–284. https://doi.org/https://doi.org/10.1016/j.brainres.2019.05.036.
13. Wong, K.M.F.; Mak, A.D.P.; Yuen, S.Y.; Leung, O.N.W.; Ma, D.Y.; Chan, Y.; Cheong, P.K.; Lui, R.; Wong, S.H.; Wu, J.C.Y. Nature and specificity of altered cognitive functioning in IBS. *Neurogastroenterology & Motility* **2019**, *31*, e13696. https://doi.org/10.1111/nmo.13696.
14. Billing, J.; Berentsen, B.; Lundervold, A.; Hillestad, E.M.; Lied, G.A.; Hausken, T.; Lundervold, A.J. Cognitive function in patients with irritable bowel syndrome: impairment is common and only weakly correlated with depression/anxiety and severity of gastrointestinal symptoms. *Scandinavian Journal of Gastroenterology* **2023**, pp. 1–9. https://doi.org/10.1080/00365521.2023.2256916.
15. Li, Z.; Ma, Q.; Deng, Y.; Rolls, E.T.; Shen, C.; Li, Y.; Zhang, W.; Xiang, S.; Langley, C.; Sahakian, B.J.; et al. Irritable Bowel Syndrome Is Associated With Brain Health by Neuroimaging, Behavioral, Biochemical, and Genetic Analyses. *Biological Psychiatry* **2024**, *95*, 1122–1132.
16. Lundervold, A.J.; Hillestad, E.M.; Lied, G.A.; Billing, J.; Johnsen, T.E.; Steinsvik, E.K.; Hausken, T.; Berentsen, B.; Lundervold, A. Assessment of Self-Reported Executive Function in Patients with Irritable Bowel Syndrome Using a Machine-Learning Framework. *Journal of Clinical Medicine* **2023**, *12*, 3771. https://doi.org/10.3390/jcm12113771.
17. Mayer, E.A.; Labus, J.S.; Tillisch, K.; Cole, S.W.; Baldi, P. Towards a systems view of IBS. *Nature Reviews Gastroenterology & Hepatology* **2015**, *12*, 592–605.
18. Mayer, E.A.; Nance, K.; Chen, S. The Gut-Brain Axis. *Annual Review of Medicine* **2022**, *73*, 439–453. https://doi.org/10.1146/annurev-med-042320-014032.
19. Nan, J.; Yang, W.; Meng, P.; Huang, W.; Zheng, Q.; Xia, Y.; Liu, F. Changes of the postcentral cortex in irritable bowel syndrome patients. *Brain Imaging and Behavior* **2020**, *14*, 1566–1576.
20. Bhatt, R.R.; Gupta, A.; Labus, J.S.; Zeltzer, L.K.; Tsao, J.C.; Shulman, R.J.; Tillisch, K. Altered brain structure and functional connectivity and its relation to pain perception in girls with irritable bowel syndrome. *Psychosomatic medicine* **2019**, *81*, 146–154.

21. Seminowicz, D.A.; Labus, J.S.; Bueller, J.A.; Tillisch, K.; Naliboff, B.D.; Bushnell, M.C.; Mayer, E.A. Regional gray matter density changes in brains of patients with irritable bowel syndrome. *Gastroenterology* **2010**, *139*, 48–57.
22. Mertz, H.; Morgan, V.; Tanner, G.; Pickens, D.; Price, R.; Shyr, Y.; Kessler, R. Regional cerebral activation in irritable bowel syndrome and control subjects with painful and nonpainful rectal distention. *Gastroenterology* **2000**, *118*, 842–848.
23. Blankstein, U.; Chen, J.; Diamant, N.E.; Davis, K.D. Altered brain structure in irritable bowel syndrome: potential contributions of pre-existing and disease-driven factors. *Gastroenterology* **2010**, *138*, 1783–1789.
24. Skrobisz, K.; Piotrowicz, G.; Rudnik, A.; Naumczyk, P.; Sabisz, A.; Markiet, K.; Szurowska, E. Evaluation of subcortical structure volumes in patients with non-specific digestive diseases. *Diagnostics* **2022**, *12*, 2199. https://doi.org/10.3390/diagnostics12092199.
25. Wolff, M.; Vann, S.D. The cognitive thalamus as a gateway to mental representations. *Journal of Neuroscience* **2019**, *39*, 3–14.
26. Parnaudeau, S.; O'neill, P.K.; Bolkan, S.S.; Ward, R.D.; Abbas, A.I.; Roth, B.L.; Balsam, P.D.; Gordon, J.A.; Kellendonk, C. Inhibition of mediodorsal thalamus disrupts thalamofrontal connectivity and cognition. *Neuron* **2013**, *77*, 1151–1162.
27. Wright, N.F.; Vann, S.D.; Aggleton, J.P.; Nelson, A.J. A critical role for the anterior thalamus in directing attention to task-relevant stimuli. *Journal of Neuroscience* **2015**, *35*, 5480–5488.
28. Berentsen, B.; Nagaraja, B.H.; Teige, E.P.; Lied, G.A.; Lundervold, A.J.; Lundervold, K.; Steinsvik, E.K.; Hillestad, E.R.; Valeur, J.; Brønstad, I.; et al. Study protocol of the Bergen brain-gut-microbiota-axis study: A prospective case-report characterization and dietary intervention study to evaluate the effects of microbiota alterations on cognition and anatomical and functional brain connectivity in patients with irritable bowel syndrome. *Medicine* **2020**, *99*, e21950. https://doi.org/doi:10.1097/MD.0000000000021950.
29. Francis, C.Y.; Morris, J.; Whorwell, P.J. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. *Alimentary pharmacology & therapeutics* **1997**, *11*, 395–402. https://doi.org/10.1046/j.1365-2036.1997.142318000.x.
30. Randolph, C. *Repeatable battery for the assessment of neuropsychological status. Norwegian manual*; NL:Pearson, 2013.
31. Fischl, B. FreeSurfer. *Neuroimage* **2012**, *62*, 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021.
32. Klauschen, F.; Goldman, A.; Barra, V.; Meyer-Lindenberg, A.; Lundervold, A. Evaluation of automated brain MR image segmentation and volumetry methods. *Human Brain MApping* **2009**, *30*, 1310–1327. https://doi.org/https://doi.org/10.1002/hbm.20599.
33. Jovicich, J.; Czanner, S.; Han, X.; Salat, D.; van der Kouwe, A.; Quinn, B.; Pacheco, J.; Albert, M.; Killiany, R.; Blacker, D.; et al. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *Neuroimage* **2009**, *46*, 177–192. https://doi.org/https://doi.org/10.1016/j.neuroimage.2009.02.010.
34. Gronenschild, E.H.; Habets, P.; Jacobs, H.I.; Mengelers, R.; Rozendaal, N.; Van Os, J.; Marcelis, M. The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements. *PloS one* **2012**, *7*, e38234. https://doi.org/https://doi.org/10.1371/journal.pone.0038234.
35. Glatard, T.; Lewis, L.B.; Ferreira da Silva, R.; Adalat, R.; Beck, N.; Lepage, C.; Rioux, P.; Rousseau, M.E.; Sherif, T.; Deelman, E.; et al. Reproducibility of neuroimaging analyses across operating systems. *Frontiers in neuroinformatics* **2015**, *9*, 12. https://doi.org/https://doi.org/10.3389/fninf.2015.00012.
36. Knussmann, G.N.; Anderson, J.S.; Prigge, M.B.; Dean III, D.C.; Lange, N.; Bigler, E.D.; Alexander, A.L.; Lainhart, J.E.; Zielinski, B.A.; King, J.B. Test-retest reliability of FreeSurfer-derived volume, area and cortical thickness from MPRAGE and MP2RAGE brain MRI images. *Neuroimage: Reports* **2022**, *2*, 100086. https://doi.org/10.1016/j.ynirp.2022.100086.
37. Debiasi, G.; Mazzonetto, I.; Bertoldo, A. The effect of processing pipelines, input images and age on automatic cortical morphology estimates. *Computer Methods and Programs in Biomedicine* **2023**, *242*, 107825. https://doi.org/https://doi.org/10.1016/j.cmpb.2023.107825.
38. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology* **2013**, *4*, 863. https://doi.org/https://doi.org/10.3389/fpsyg.2013.00863.

39. Kirk, R.E. Practical significance: A concept whose time has come. *Educational and psychological measurement* **1996**, *56*, 746–759.
40. Xiang, S.; Nie, F.; Zhang, C. Learning a Mahalanobis distance metric for data clustering and classification. *Pattern recognition* **2008**, *41*, 3600–3612.
41. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.I. From local explanations to global understanding with explainable AI for trees. *Nature machine intelligence* **2020**, *2*, 56–67.
42. Burgess, N.; Maguire, E.A.; O'Keefe, J. The human hippocampus and spatial and episodic memory. *Neuron* **2002**, *35*, 625–641.
43. Fritch, H.A.; Spets, D.S.; Slotnick, S.D. Functional connectivity with the anterior and posterior hippocampus during spatial memory. *Hippocampus* **2021**, *31*, 658–668.
44. Ellingson, B.M.; Mayer, E.; Harris, R.J.; Ashe-McNally, C.; Naliboff, B.D.; Labus, J.S.; Tillisch, K. Diffusion tensor imaging detects microstructural reorganization in the brain associated with chronic irritable bowel syndrome. *PAIN®* **2013**, *154*, 1528–1541.
45. Bhatt, R.R.; Gupta, A.; Labus, J.S.; Liu, C.; Vora, P.P.; Stains, J.; Naliboff, B.D.; Mayer, E.A. A neuropsychosocial signature predicts longitudinal symptom changes in women with irritable bowel syndrome. *Molecular psychiatry* **2022**, *27*, 1774–1791.
46. Hellgren, L.; Thornberg, U.B.; Samuelsson, K.; Levi, R.; Divanoglou, A.; Blystad, I. Brain MRI and neuropsychological findings at long-term follow-up after COVID-19 hospitalisation: an observational cohort study. *BMJ open* **2021**, *11*, e055164.
47. Lisman, J.; Buzsáki, G.; Eichenbaum, H.; Nadel, L.; Ranganath, C.; Redish, A.D. Viewpoints: how the hippocampus contributes to memory, navigation and cognition. *Nature neuroscience* **2017**, *20*, 1434–1447.
48. Zhao, M.; Hao, Z.; Li, M.; Xi, H.; Hu, S.; Wen, J.; Gao, Y.; Antwi, C.O.; Jia, X.; Yu, Y.; et al. Functional changes of default mode network and structural alterations of gray matter in patients with irritable bowel syndrome: a meta-analysis of whole-brain studies. *Frontiers in Neuroscience* **2023**, *17*, 1236069. https://doi.org/10.3389/fnins.2023.1236069.
49. Labus, J.S.; Wang, C.; Mayer, E.A.; Gupta, A.; Oughourlian, T.; Kilpatrick, L.; Tillisch, K.; Chang, L.; Naliboff, B.; Ellingson, B.M. Sex-specific brain microstructural reorganization in irritable bowel syndrome. *Pain* **2023**, *164*, 292–304.
50. van Kessel, L.; Teunissen, D.; Lagro-Janssen, T. Sex-gender differences in the effectiveness of treatment of irritable bowel syndrome: A systematic review. *International Journal of General Medicine* **2021**, pp. 867–884.
51. Lee, O.Y.; Mayer, E.A.; Schmulson, M.; Chang, L.; Naliboff, B. Gender-related differences in IBS symptoms. *Official journal of the American College of Gastroenterology | ACG* **2001**, *96*, 2184–2193.
52. Kim, N. Sex/Gender Differences in the Diseases. In *Sex/Gender-Specific Medicine in Clinical Areas*; Springer, 2024; pp. 25–40.
53. Chang, L.; Heitkemper, M.M. Gender differences in irritable bowel syndrome. *Gastroenterology* **2002**, *123*, 1686–1701.
54. Vemuri, R.; Sylvia, K.E.; Klein, S.L.; Forster, S.C.; Plebanski, M.; Eri, R.; Flanagan, K.L. The microgenderome revealed: sex differences in bidirectional interactions between the microbiota, hormones, immunity and disease susceptibility. In Proceedings of the Seminars in immunopathology. Springer, 2019, Vol. 41, pp. 265–275.

Author Contributions: "Conceptualization of the present study, J.B., B.R.B., A.L., A.J.L.; methodology, J.B., B.R.B., A.L., A.J.L.; formal analysis, B.R.B., A.L.; data collection G.A.L., E.S. T. H., B.B., and A.J.L.; writing original draft preparation J.B., A.J.L.; review and editing J. B., G.A.L., E.S. T.H., B.B. A.J.L., A.L., project administration, B.B.; funding acquisition, T.H., A.L. All authors have read and agreed to the published version of the manuscript.".

Funding: This research was funded by Research Council of Norway (grant ID FRIMED-BIO276010) and Helse Vest's Research Funding (grant ID HV912243) and by the Trond Mohn Research Foundation, grant number BFS2018TMT0, and from The Research Council of Norway, project number 294594.

Institutional Review Board Statement: The B-BGM project was approved by the Southeast Regional Ethical Committees (REC) for medical and health research ethics in Norway (REK2015-01621). All participants provided written consent to participate, and the project was conducted following the ethical requirements from the Declaration of Helsinki. The project is registered at www.clinicaltrials.gov (#NCT04296552).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The implementation of the complete workflow, the setup of the corresponding `conda` environment, the cleaned input dataset in `.csv` format, and code for all tables and figures in the Results section are available as *Jupyter notebooks* at https://github.com/arvidl/ibs-brain.

Acknowledgments: We sincerely thank all patients and healthy volunteers for their participation in the Bergen Brain-Gut Microbiota (B-BGM) project. We also thank all the present and previous members of the B-BGM project.

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability Statement: We encourage all authors of articles published in MDPI journals to share their research data. In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Where no new data were created, or where data is unavailable due to privacy or ethical restrictions, a statement is still required. Suggested Data Availability Statements are available in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics.

Abbreviations

The following abbreviations are used in the manuscript:

B-BGM project	Bergen Brain-Gut Microbiota project
CBT	Cognitive Behavior Therapy
CM	Confusion matrix
Cohen's d	Effect size
DGBI	Disorders of the gut-brain interaction
GI	Gastrointestinal
HC	Healthy Control
IBS	Irritable bowel syndrome
IBS-SSS	IBS Severity Scoring System
MRI	Magnetic Resonance Imaging
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status
RF	Random Forest
SD	Standard deviation
SHAP	SHapley Additive exPlanations

Appendix A

Appendix A.1

Table A1. FreeSurfer segmented brain regions (`aseg`) with short descriptions of functional anatomy

Region	Description
eTIV	Estimated Total Intracranial Volume
Left-Cerebellum-White-Matter	White matter in the left cerebellum
Left-Cerebellum-Cortex	Gray matter (cortex) in the left cerebellum
Left-Thalamus	Left thalamus. *Thalamus* is a relay center for sensory and motor signals. In IBS, thalamic activity will contribute to pain perception and visceral hypersensitivity
Left-Caudate	Left caudate nucleus. *Nucleus caudatus* is involved in motor and motility control and learning
Left-Putamen	Left putamen. *Putamen* is part of the basal ganglia involved in motor control and may contribute to habitual responses to gastrointestinal discomfort
Left-Pallidum	Left globus pallidus. *Globus pallidus* is involved in regulating voluntary movement and gut motility patterns
Left-Hippocampus	Left hippocampus. *Hippocampus* is crucial for memory formation and spatial navigation, and in IBS, involved in contextual fear learning related to gastrointestinal symptoms
Left-Amygdala	Left amygdala. *Amygdala* is involved in processing emotions, fear, and anxiety
Left-Accumbens-area	Left nucleus accumbens. *Nucleus accumbens* is involved in reward and motivation, stress responsivity, and pain modulation
CSF	Cerebrospinal Fluid
Right-Cerebellum-White-Matter	White matter in the right cerebellum
Right-Cerebellum-Cortex	Gray matter (cortex) in the right cerebellum
Right-Thalamus	Right thalamus
Right-Caudate	Right caudate nucleus
Right-Putamen	Right putamen
Right-Pallidum	Right globus pallidus
Right-Hippocampus	Right hippocampus
Right-Amygdala	Right amygdala
Right-Accumbens-area	Right nucleus accumbens
WM-hypointensities	White matter hypointensities (dark on T1-w sequences), can be associated with small vessel disease, demyelination, inflammation, fluid accumulation
CC_Posterior	Posterior part of the corpus callosum
CC_Mid_Posterior	Mid-posterior part of the corpus callosum
CC_Central	Central part of the corpus callosum
CC_Mid_Anterior	Mid-anterior part of the corpus callosum
CC_Anterior	Anterior part of the corpus callosum
BrainSegVol	Total volume of brain segmentation
BrainSegVolNotVent	Brain segmentation volume without ventricles
lhCortexVol	Volume of the left hemisphere cortex
rhCortexVol	Volume of the right hemisphere cortex
CortexVol	Total cortical volume (left + right)
lhCerebralWhiteMatterVol	Volume of left hemisphere cerebral white matter
rhCerebralWhiteMatterVol	Volume of right hemisphere cerebral white matter
CerebralWhiteMatterVol	Total cerebral white matter volume (left + right)
SubCortGrayVol	Volume of subcortical gray matter
TotalGrayVol	Total gray matter volume

Appendix B

All appendix sections must be cited in the main text. In the appendices, Figures, Tables, etc. should be labeled, starting with "A"—e.g., Figure A1, Figure A2, etc.