

Brain functional connectivity, but not neuroanatomy, captures the interrelationship between sex and gender in preadolescents

Athanasia Metoki^{a,*} , Roselyne J. Chauvin^a, Evan M. Gordon^b, Timothy O. Laumann^c, Benjamin P. Kay^a, Babatunde Adeyemo^{a,c}, Samuel R. Krimmel^a, Scott Marek^b, Anxu Wang^a, Andrew N. Van^{a,g}, Noah J. Baden^a, Vahdeta Suljic^a, Kristen M. Scheidter^a, Julia Monk^a, Forrest I. Whiting^a, Nadeshka J. Ramirez-Perez^d, Deanna M. Barch^{c,e}, Aristeidis Sotiras^{b,f,1}, Nico U.F. Dosenbach^{a,b,e,g,h,i,1}

^a Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA

^b Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO, USA

^c Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA

^d Department of Psychiatry, Harvard University, Cambridge, MA, USA

^e Department of Psychological & Brain Sciences, Washington University, St. Louis, MO, USA

^f Institute for Informatics, Data Science & Biostatistics, Washington University School of Medicine in St. Louis, St. Louis, MO, USA

^g Department of Biomedical Engineering, Washington University, St. Louis, MO, USA

^h Program in Occupational Therapy, Washington University, St. Louis, MO, USA

ⁱ Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA

ARTICLE INFO

Keywords:

Resting-state functional connectivity
Cortical thickness
Sex
Gender
Adolescence
Brain networks
Adolescent brain cognitive development study

ABSTRACT

Understanding sex differences in the adolescent brain is crucial, as they relate to sex-specific neurological and psychiatric conditions. Predicting sex from adolescent brain data may reveal how these differences influence neurodevelopment. Recently, attention has shifted toward socially-identified gender (distinct from sex assigned at birth) recognizing its explanatory power. This study evaluates whether resting-state functional connectivity (rsFC), cortical thickness, or cortical volume better predicts sex and sex/gender alignment (congruence between sex and gender) in preadolescents. Using Adolescent Brain Cognitive Development (ABCD) Study data and machine learning, rsFC predicted sex more accurately (85 %) than cortical thickness (76 %) and cortical volume (70 %). Brain regions most predictive of sex belonged to association (default mode, dorsal attention, parietal memory) and visual networks. The rsFC classifier trained on sex/gender aligned youth classified more accurately unseen youth with sex/gender alignment ($n = 2013$) than unalignment ($n = 1116$). The female rsFC sex profile was positively associated with sex/gender alignment, while in males, there was a negative association. However, neither brain modality predicted sex/gender alignment. These findings suggest that while rsFC predicts sex in the adolescent brain more accurately, it does not directly capture sex/gender alignment, underscoring the need for further investigation into the neural underpinnings of gender.

1. Introduction

The investigation of sex differences in the human brain has been a research interest for decades (Cosgrove et al., 2007; Giedd et al., 2012; Jäncke, 2018; Kaczurkin et al., 2018; Lenroot and Giedd, 2010; Luders and Kurth, 2020; Luders and Toga, 2010; Sacher et al., 2013) spanning studies focusing on adults (Choleris et al., 2018; Jäncke, 2018;

Kodiweera et al., 2016; Lotze et al., 2019; Ritchie et al., 2018), to those examining children and adolescents (Giedd et al., 2012; Gur and Gur, 2016; Herting and Sowell, 2017; Tunç et al., 2016; Vijayakumar et al., 2018), and even infants (Benavides et al., 2018; Fenske et al., 2023; Gilmore et al., 2007). Understanding brain sex differences is crucial, as many neurological and psychiatric conditions vary by sex in prevalence, onset, and symptomatology (Bao and Swaab, 2010; Baron-Cohen et al.,

* Correspondence to: Department of Neurology, Washington University School of Medicine in St. Louis, 660 S Euclid Ave, St. Louis, MO 63110, USA.

E-mail address: athanasiametoki@wustl.edu (A. Metoki).

¹ Co-senior authors.

2011; Paus et al., 2008; Rutter et al., 2003), possibly due to differences in brain structure or function. Adolescence is a critical developmental period marked by increasing differences between males and females in physical traits, behavior, and mental health risks. Identifying sex-specific brain development patterns during this time could illuminate the roots of distinct mental health vulnerabilities, guiding tailored interventions that address the unique needs of each sex.

Diverse neuroimaging modalities have been used to study brain sex differences, including anatomical (e.g., volume, surface, and white matter integrity) and functional (e.g., resting-state functional connectivity; rsFC) measures. Structural variations have been observed between males and females (for a comprehensive review see Kaczkurkin et al., 2018), in brain size (Ball et al., 2012; Bramen et al., 2011; De Bellis et al., 2001; Luders et al., 2005; Ritchie et al., 2018; Ruigrok et al., 2014), cortical and subcortical volume (Allen et al., 2003; Goldstein et al., 2001; Gur et al., 1999; Luders et al., 2005), cortical surface area (Luders et al., 2006; Nopoulos et al., 2000), hemispheric white matter connectivity (Ingalhalikar et al., 2014), and cerebral blood flow (Satterthwaite et al., 2014). Despite these structural distinctions, the existence and extent of functional brain organization sex differences remains unclear. While some studies have pinpointed specific networks where rsFC appears to vary between males and females (Biswal et al., 2010; Ritchie et al., 2018; L. Tian et al., 2011; Zuo et al., 2010), others have reported no discernible impact of sex in resting-state functional magnetic resonance imaging (fMRI) data (Weissman-Fogel et al., 2010). Moreover, evidence shows that sex differences in rsFC networks are affected by the menstrual cycle (Avila-Varela et al., 2024; Weis et al., 2019), indicating that rsFC variation may depend not only on sex but also on underlying factors like hormonal fluctuations.

Most studies on brain sex differences focus on sex as assigned at birth. However, an individual's socially-identified gender may also influence brain development, potentially in part through societal and cultural perceptions of and reactions to their gender. This, in turn, can contribute to gender behavioral differences independent of sex (Bale and Epperson, 2016; Kim and Nafziger, 2000). To explore whether brain factors play a role in gender identity development, researchers have compared cisgender and transgender adult individuals, with mixed findings. Some studies suggest that the brains of transgender individuals resemble those of their sex assigned at birth (Emory et al., 1991; Luders et al., 2009; Mueller, Landré, et al., 2017; Yokota et al., 2005), others align them with their gender identity (Mueller, Landré, et al., 2017; Zhou et al., 1995), while additional research shows an intermediate (Flint et al., 2020; Kranz et al., 2014; Kurth et al., 2022; Mueller, De Cuypere, et al., 2017; Mueller, Landré, et al., 2017; Rametti, Carrillo, Gómez-Gil, Junque, Segovia, et al., 2011; Rametti, Carrillo, Gómez-Gil, Junque, Zubiarré-Elorza, et al., 2011) or unique phenotype (Mueller et al., 2021). Overall, evidence on brain differences related to sex and gender identity remains inconclusive.

Previous research on brain sex differences has often relied on traditional regression methods that treat networks or regions-of-interest (ROI) independently. In contrast, multivariate pattern analysis can reveal more complex patterns without assuming independence, making advanced computational methods better suited for identifying brain sex differences. Machine learning approaches, such as classification, can assess how accurately sex can be predicted from neuroimaging data (Bzdok, 2017), yet few studies have applied these methods to examine sex differences in youth (Adeli et al., 2020; Brennan et al., 2021; Kurth et al., 2020; Sepehrband et al., 2018; Shanmugan et al., 2022). Existing research has typically focused on either brain structure (Adeli et al., 2020; Brennan et al., 2021; Kurth et al., 2020; Sepehrband et al., 2018) or function (Shanmugan et al., 2022), with no study to date comparing the predictive power of neuroanatomy versus functional connectivity in children or adolescents using a single dataset. Additionally, the potential benefits of combining both structural and functional data have not been explored in young populations, and most studies have used small samples spanning broad age ranges (Kurth et al., 2020; Sepehrband et al.,

2018; Shanmugan et al., 2022) rather than focusing on larger, age-specific samples. Only one study has examined the neurobiological underpinnings of sex and gender in children (Dhamala et al., 2024) using linear ridge regression, finding that while sex and gender are uniquely represented in brain patterns, gender is less distinctly captured in functional connectivity than sex. However, this study did not explore a key cross-comparison evaluating the extent of shared versus independent predictive power between sex and gender in the same functional connections or networks, leaving unresolved questions about whether these unique representations share predictive power or are entirely independent.

In the present study, we moved beyond previous work by evaluating whether rsFC, cortical thickness or cortical volume is more effective in predicting sex and gender, while also exploring the extent of their interrelationship within the brain of preadolescents. We leveraged neuroimaging data from the Adolescent Brain Cognitive Development® (ABCD; Garavan et al., 2018) Study ($n = 3129$, 9–10 years old) at baseline, and self- and parent-reported gender data at the one-year follow-up time point. We use the following terminology: “sex” to describe an individual's sex assigned at birth (i.e., male or female), usually based on physical anatomy and/or chromosomes at birth, and “gender” as the internal sense of oneself as boy, girl, or something else (Potter et al., 2021). We also use the term “sex/gender alignment” to refer to the congruence between an individual's sex and their gender. Our approach involved unimodal (i.e., rsFC, cortical thickness and cortical volume separately) and multimodal (i.e., rsFC/cortical thickness and rsFC/cortical volume combined) support vector machine (SVM) learning neuroimaging analyses to determine the predictability of sex and sex/gender alignment. Additionally, we tested whether a sex classifier trained on individuals with sex/gender alignment would perform comparably when predicting the sex of similarly aligned versus unaligned individuals. This was done to assess whether the concept of sex/gender alignment is distinguishable from sex in the brain. Lastly, we examined the relationship between the rsFC, cortical thickness and cortical volume sex classification scores (brain profile predictions) and sex/gender alignment scores to assess the consistency between the adolescents' brain profiles and their sex/gender alignment.

2. Methods

2.1. ABCD study

The ABCD is a ten-year, longitudinal study of 11,875 youth enrolled at ages 9–10 from 21 sites in the United States. Participants (youth and parents) were recruited through schools, with minimal exclusion criteria (Garavan et al., 2018). The ABCD study obtained centralized institutional review board (IRB) approval from the University of California, San Diego. Each of the 21 sites also obtained local IRB approval. Ethical regulations were followed during data collection and analysis. Parents or caregivers provided written informed consent, and children gave written assent. The racial demographics of the participants roughly match the racial composition of the 2023 American Community Survey (U.S. Census Bureau, 2023). Although the ABCD study utilizes a longitudinal design, the analyses conducted for the current study were cross-sectional.

2.2. Participants

This project used rsFC, cortical thickness, and cortical volume data from $n = 10,259$ available participants from the ABCD BIDS (Brain Imaging Data Structure) Community Collection (Feczko et al., 2021) (ABCD collection 3165; <https://github.com/ABCD-STUDY/nda-abcd-collection-3165>), baseline visit (9–10 y). Demographic and behavioral data were obtained from the ABCD 4.0 release, which included data from the one-year follow-up visit, marking the beginning of gender measures' collection. Following the ABCD consortium's

recommendations, we excluded participants scanned with Philips scanners due to incorrect preprocessing (source: <https://github.com/ABCD-STUDY/fMRI-cleanup>).

Head motion can systematically bias developmental studies (Power et al., 2012; Siegel et al., 2017), as well as those relating rsFC to behavior (Siegel et al., 2017). However, these systematic biases can be addressed through rigorous head motion correction (Power et al., 2014). The inclusion criteria for the current project required participants to have at least 600 frames (equivalent to 8 min) of low-motion (Fair et al., 2020) rsFC data (low-motion defined as having a filtered framewise displacement (FD) of less than 0.08 mm). See Casey et al. (2018) for broader ABCD inclusion criteria. Based on these criteria, and after excluding subjects with missing values in demographic, gender, or behavioral questions, the final ABCD sample consisted of rsFC, cortical thickness, cortical volume, demographic, gender, and behavioral data from a total of $n = 3129$ youth (Table 1).

2.3. MRI acquisition and image processing methods

2.3.1. MRI acquisition

Imaging for each ABCD youth was performed across 21 sites within the United States, and was harmonized across Siemens Prisma, Philips, and GE 3 T scanners. Details on image acquisition can be found elsewhere (Casey et al., 2018). Twenty minutes (4×5 min runs) of eyes-open (passive crosshair viewing) resting-state fMRI data were collected to ensure at least 5 min of low-motion data. All resting-state fMRI scans were acquired using a gradient-echo EPI sequence (TR = 800 ms, TE = 30 ms, flip angle = 90° , voxel size = 2.4 mm^3 , 60 slices). Head motion was monitored online using the Framewise Integrated Real-time MRI Monitor (FIRMM) software at Siemens sites (Dosenbach et al., 2017).

2.3.2. Image processing overview

The image processing stream has been detailed elsewhere (Feczko et al., 2021). Briefly, the ABCD pipeline comprises six stages: (1) Pre-FreeSurfer, which normalizes anatomical data; (2) FreeSurfer, which constructs cortical surfaces from the normalized anatomical data; (3) PostFreeSurfer, which converts outputs from FreeSurfer to CIFTIs and transforms the volumes to a standard volume space using ANTs nonlinear registration; (4) “Vol” stage, which performs the atlas transformation, mean field distortion correction, and resampling to 2 mm isotropic voxels in a single step using FSL’s applywarp tool; (5) “Surf” stage, which projects the volumetric functional data onto the surface;

Table 1

Demographics ($n = 3129$) and descriptive statistics. SD, standard deviation; PDS, pubertal development scale; FD framewise displacement.

	TOTAL ($n = 3129$)		
	Males $n = 1542$; 49.3 %	Females $n = 1587$; 50.7 %	
Mean Age in Months (SD) Range	132.4 (7.8) 116–148	132.3 (7.7) 117–149	$t = 0.53, p = 0.59$
Sex/Gender Alignment			
Aligned	1153 (74.8 %)	860 (54.2 %)	$\chi^2 = 144.40,$ $p < 0.001$
Unaligned	389 (25.2 %)	727 (45.8 %)	
PDS			
Prepuberty	468 (30.3 %)	225 (14.2 %)	$\chi^2 = 118.63,$ $p < 0.001$
Early puberty	712 (46.2 %)	331 (20.9 %)	
Mid puberty	342 (22.2 %)	854 (53.8 %)	$\chi^2 = 331.43,$ $p < 0.001$
Late puberty	20 (1.3 %)	177 (11.1 %)	
Residual In-Scanner Motion (Mean FD), mm	0.13 (0.08) 0.03–0.107	0.12 (0.07) 0.03–0.70	$t = 3.75,$ $p < 0.001$

and (6) “DCANBOLDproc”, which performs functional connectivity processing.

DCANBOLDproc includes a respiratory filter to improve FD estimates calculated in the “vol” stage. Temporal masks were created to flag motion-contaminated frames using the improved FD estimates (Power et al., 2012). Frames with FD > 0.30 mm were flagged as motion contaminated. After computing the temporal masks for high motion frame censoring, the data were processed with the following steps: (i) demeaning and detrending, (ii) interpolation across censored frames using least squares spectral estimation of the values at censored frames (Power et al., 2014) so that continuous data can be (iii) denoised via a general linear model (GLM) including whole brain, ventricular, and white matter signals, as well as their derivatives. Denoised data were then passed through (iv) a band-pass filter ($0.008 \text{ Hz} < f < 0.10 \text{ Hz}$) without re-introducing nuisance signals (Hallquist et al., 2013) or contaminating frames near high motion frames (Carp, 2013).

2.3.3. Functional connectivity, cortical thickness, and cortical volume metrics

An ROI method was employed, utilizing a total of 394 parcels, which included the 333 cortical parcels initially defined by Gordon et al. (2016) and an additional 61 subcortical parcels described by Seitzman et al. (2020). For each participant, the rsFC timecourse was extracted from these 394 ROIs after discarding frames with a filtered FD > 0.08 mm. The 333 cortical ROIs can be organized into separable brain networks (e.g., default-mode, fronto-parietal, salience, etc.). A correlation matrix was created through the computation of the correlation between the timecourse of each ROI and that of every other ROI, yielding a 394×394 correlation matrix for each subject. The correlations were normalized using Fisher’s r -to- z transform (Fisher, 1915). The correlation matrices were then combined across all subjects to form a $394 \times 394 \times 3129$ matrix, which was employed for subsequent analyses. rsFC analyses were run at the edge level (ROI-ROI pair; $n = 77,421$). Furthermore, cortical thickness and cortical volume was extracted from the 333 cortical ROIs for each participant. Cortical volume for each ROI was normalized by dividing each ROI’s volume by the subject’s intracranial volume to account for individual differences in brain size. Additional cortical thickness analyses at the vertex-level (59,412 cortical vertices for each participant) were also conducted in an effort to capture regional effects that may not align with ROI boundaries (Supplementary Results).

2.4. Sex and gender measures

We used the ABCD’s parent-reported data on youths’ sex, which reflect their sex assigned at birth. In the ABCD dataset, gender identity is assessed as part of the background items in the Kiddie Schedule of Affective Disorders and Schizophrenia (K-SADS; Kaufman et al., 1997). However, this item has been identified as developmentally inappropriate for young adolescents, as evidenced by two out of five youths aged 9–10 years not understanding what the question is asking (Dube et al., 2021; A. Potter et al., 2021). In addition, the ABCD study assesses gender with the Youth Self-Report and Parent-Report Gender Questionnaires. All participants included in analyses completed the Youth Self-Report Gender Questionnaire which comprised four questions assessing felt-gender, contentedness with assigned sex at birth, and gender expression (e.g., “How much do you feel like a <boy/girl>?” for the felt-gender construct. For the complete list of questions and constructs see Supplementary Table 1). Their parents/caregivers completed an adaptation of the Gender Identity Questionnaire (Elizabeth and Green, 1984; Johnson et al., 2004) which included 12 questions that measure sex-typed behavior during play and gender dysphoria (Potter et al., 2022). All items use a five-point scale with higher scores reflecting more congruence with sex assigned at birth. The current study computed and used the mean scores of both constructs separately, capturing gender dimensionally (Potter et al., 2022). Sex/gender alignment status was

either “aligned”, i.e., sex/gender alignment defined as scoring the highest possible achievable score in the Youth Self-Report Gender Questionnaire, or “unaligned”, i.e., sex/gender unalignment defined as scoring anything other than the highest possible score. Participants who provided a “Decline to answer” response to any question were not included in the analyses (see [Supplementary Fig. 1](#) for response frequencies on the Youth Self-Report and Parent-Report Gender Questionnaires).

2.5. Support vector machine learning

2.5.1. Support vector machine classification

Support vector machine (SVM) learning was used to examine the predictability of sex in adolescents based on neuroimaging measures. Five binary linear SVM classifiers were constructed, each using a different set of predictors: rsFC, cortical thickness, cortical volume, rsFC combined with cortical thickness, and rsFC combined with cortical volume. The target labels for all models were the sex categories of male and female. The classifiers were trained exclusively using youth with sex/gender alignment (highest possible attainable score in the Youth Self-Report Gender Questionnaire). The objective of training the models exclusively on participants with the most extreme sex/gender alignment was to determine whether the SVM models, which aimed to predict sex, could effectively capture the concept of sex/gender alignment. If the models show no significant difference in accuracy between predicting sex in aligned and unaligned participants, it would suggest that sex and sex/gender alignment are indistinguishable in rsFC and/or cortical thickness/cortical volume. As a form of validation, we conducted the opposite analyses too, of training our SVM models on unaligned participants and testing on the aligned ones. The aligned group data was

split to retain 20 % of subjects held out entirely for testing and 80 % for training (for the validation analyses, this split was applied in the unaligned group). Depending on the analysis, independent testing sets were composed either entirely of youth with sex/gender alignment (aligned group) or entirely of youth with sex/gender unalignment (unaligned group) ([Fig. 1](#)). To account for statistically significant sex differences in pubertal development stage and residual in-scanner motion (mean FD) (see 2.6 Other Statistical Analyses, 3. Results, & [Table 1](#)), these variables were regressed out separately from each feature in the training, cross-validated, and independent test datasets to prevent data leakage. To ensure robust model evaluation and prevent overfitting, the classification was evaluated by using a nested five-fold cross-validation (5F-CV) ([Poldrack et al., 2020](#)) in the training set, with the inner 5F-CV determining the optimal regularization coefficient C via grid search for the SVM binary classifier and the outer 5F-CV estimating the generalizability of the model ([Fig. 1](#)). Each feature was linearly scaled between zero and one across the outer 5F-CV training sets; these scaling parameters were then applied to scale the outer 5F-CV and independent testing sets ([Cui and Gong, 2018; Erus et al., 2015](#)). For additional SVM parameters see [Supplementary Methods](#).

Assigning weights to observations (single data points or instances within a dataset) is a valuable approach for addressing the issue of treating all observations equally. The weights indicate the importance or influence of each observation during the SVM model training. Due to an imbalanced distribution of males and females in the aligned group (see 3. Results & [Table 1](#)), which would lead to an imbalance in the SVMs’ aligned training groups, inverse proportion instance weights were calculated and assigned to help balance the contribution of each class to the models.

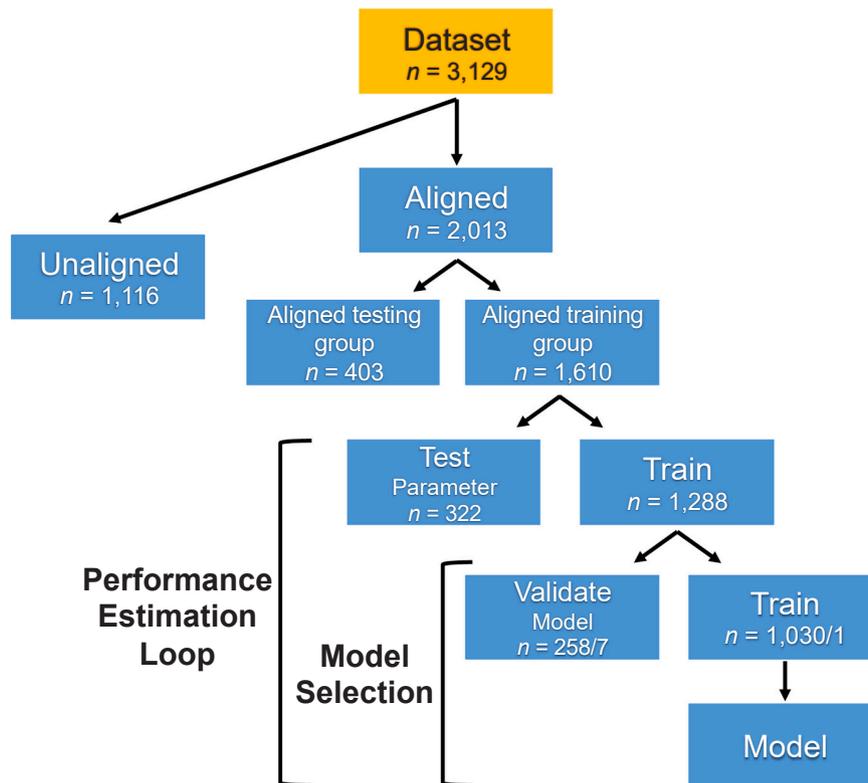


Fig. 1. Flowchart of the support vector machine (SVM) model construction. The dataset was split in two groups: Aligned (participants with sex/gender alignment) and unaligned (participants with sex/gender unalignment). The aligned group was split in an aligned training group (80 %) and an aligned hold-out testing group (20 %). A nested five-fold cross-validation (5F-CV) was employed, with the inner 5F-CV determining the optimal parameter C and the outer 5F-CV estimating the generalizability of the model. The final, optimal model was subsequently applied in the held-out aligned testing group and the unaligned group, and model performance was evaluated. The forward slash in the groupings denotes the different possible group splits (e.g. The performance estimation loop training group was split in three train/test group pairs with 1030/258 subjects and two train/test group pairs with 1031/257 subjects respectively).

2.5.2. Support vector regression

Linear support vector regression (SVR) was used to evaluate whether sex/gender alignment (a discrete variable derived from the Youth Self-Report and Parent-Report Gender Questionnaires) can be predicted by brain neuroanatomy or functional connectivity. Using the same parameters/choices as in the SVM classification models (Methods and [Supplementary Methods](#)), we ran a total of twelve models, utilizing different combinations of features and questionnaire types. For the rsFC analyses, two models were run: one exclusive to females and one exclusive to males, and we employed a data split of 20 % for hold-out testing and 80 % for training, in both the female and male datasets. Each of these models utilized the Youth Self-Report Gender Questionnaire as the target variable in one iteration and the Parent-Report Gender Questionnaire in another. Similarly, for the cortical thickness, cortical volume, and the combined rsFC/cortical thickness analyses, another set of models was executed, following the same structure as the rsFC analyses. The same grid search optimization of the regularization coefficient C using 5F-CV was performed as for the binary SVM classifiers ([Supplementary Methods](#)). Each feature was linearly scaled between zero and one across the outer 5F-CV training sets and these scaling parameters were then applied to scale the outer 5F-CV and the independent testing sets (Cui and Gong, 2018; Erus et al., 2015). The imbalance in the distribution of test scores in the Youth Self-Report and the Parent-Report Gender Questionnaires ([Supplementary Fig. 1](#)) led to assigning inverse proportion instance weights in the SVR models.

As a robustness check measure for the rsFC SVR analyses, we opted to run linear ridge regressions as performed by Dhamala et al. (2024) using the code they provided online (<https://zenodo.org/doi/10.5281/zenodo.10779163>). The only adjustment in the script was to use the updated “StandardScaler” to normalize the data, as the “normalize” parameter has been removed from the Scikit-learn (Pedregosa et al., 2011) “Ridge” regression command starting from version 1.0, recommending users handle normalization separately instead.

2.5.3. Assessment and significance of prediction performance

The performance evaluation of the SVM binary classification models encompassed a comprehensive set of metrics (Poldrack et al., 2020), including accuracy, sensitivity, specificity, the area under the receiver operating characteristic curve (AUC), and the Matthews correlation coefficient (MCC). As opposed to accuracy, which can produce over-optimistic inflated results (Akosa, 2017; Flach and Kull, 2015; Gu et al., 2009; Hand and Christen, 2018; Sokolova et al., 2006; Xi and Beer, 2018), the MCC treats both classes symmetrically, penalizing models that perform well on one class but poorly on another, thus providing a balanced evaluation (Chicco, Tötsch, et al., 2021; Chicco, Warrens, et al., 2021; Chicco and Jurman, 2020, 2022, 2023). It ranges from -1 to $+1$, with $+1$ indicating a perfect prediction, 0 a prediction no better than random chance, and -1 total disagreement between prediction and observation. This comprehensive suite of metrics ensured a thorough and nuanced evaluation of the binary classification models, capturing various aspects of their performance and facilitating informed decision-making.

The McNemar’s test (McNemar, 1947; Pembury Smith and Ruxton, 2020) was used to evaluate the differences in classification accuracies between the rsFC, cortical thickness, cortical volume, and combined rsFC/cortical thickness SVM binary classification models. A two-proportions z-test was conducted to compare the proportions of correct predictions between the aligned and the unaligned SVM models.

Permutation testing was used to evaluate whether the prediction performance of the SVM binary classification models was significantly better than expected by chance (Mourão-Miranda et al., 2005). The predictive framework was repeated 10,000 times. The coefficient of determination (R^2) (Poldrack et al., 2020) was used as a method of evaluation of the SVR models’ performance.

2.5.4. Interpreting model feature weights

The transformation proposed by Haufe et al. (2014) was utilized to transform feature weights derived from the linear SVM binary classification models with the aim of enhancing their interpretability and reliability (Chen et al., 2022; Y. Tian and Zalesky, 2021). To enhance interpretability for visualization purposes, the parcel-wise feature weights were consolidated at the network level by calculating their root mean square value and averaging across previously defined canonical functional networks (Gordon et al., 2016, 2023).

2.6. Other statistical analyses

To ensure that there were no confounders when running the SVM classification, we conducted a series of statistical tests to examine potential differences between males and females, allowing for a more accurate and unbiased analysis. Independent *t*-tests were performed to compare age and residual in-scanner motion (mean FD) between the two groups, using Student’s *t*-test when variances were equal and Welch’s *t*-test when they were not. Additionally, chi-squared tests were applied to assess the distribution of males and females across various categorical variables, including sex/gender alignment, pubertal development stage, and family income.

The National Academies of Science, Engineering, and Medicine (NASEM) issued a report providing recommendations on using race, ethnicity, ancestry, and other descriptors of population stratification (National Academies of Sciences et al., 2023). The NASEM report asserts that race does not have any biological basis and emphasizes that researchers should not “control for race” in brain-based association or prediction studies with non-brain variables. The current study follows these recommendations and therefore does not evaluate race and ethnicity in brain association and prediction models (Bird and Carlson, 2024; Schraiber and Edge, 2024).

To further confirm that the performance differences in our SVM models were driven by sex/gender alignment status, we conducted additional comparisons (Independent Student’s or Welch’s *t*-tests, depending on variance equality) between the pubertal development stage and the income-to-needs ratio (calculated by dividing participant-reported annual family income by the 2017 federal poverty level for a corresponding family size (U.S. Department of Health and Human Services, 2017)) in correctly classified subjects (true positives and true negatives) versus incorrectly classified subjects (false positives and false negatives). By examining pubertal development and the income-to-needs ratio, we aimed to determine whether the observed differences in classification performance could be attributed to factors beyond sex/gender alignment status.

Spearman’s correlations were used to examine the relationship between the rsFC, cortical thickness, and cortical volume sex classification scores (brain profile predictions) and sex/gender alignment scores in order to evaluate the consistency between the adolescents’ brain profiles and sex/gender alignment. We opted for a non-parametric approach due to the non-normal distribution of the sex/gender alignment scores ([Supplementary Fig. 1](#)).

Global signal regression (GSR) is a commonly used method for mitigating motion-related artifacts in functional connectivity analyses (Chang and Glover, 2009; Murphy et al., 2013; Power et al., 2014; Zhu et al., 2015). While some have raised concerns that GSR may complicate group comparisons due to differential re-referencing (Gotts et al., 2013; Saad et al., 2012), this interpretation remains debated in the field. To ensure that our findings were not dependent on this preprocessing choice, we repeated SVM classification group-comparison analyses without GSR.

3. Results

3.1. Youths' demographic characteristics

In this sample, $n = 3129$, males ($n = 1542$ [49.3 %]) and females ($n = 1587$ [50.7 %]) did not differ statistically in age (Table 1) or family income (Supplementary Table 2). Statistically significant sex differences were found in pubertal development scale scores for each puberty stage and residual in-scanner motion (mean FD of retained fMRI frames) (Table 1). Thus, pubertal development stage and mean FD, were used as covariates in all SVM analyses. Demographic details of the aligned and unaligned groups are provided in Supplementary Table 3.

3.2. Functional connectivity and neuroanatomical classification of sex

This study's first goal was to understand the way in which high dimensional patterns of brain functional connectivity and

neuroanatomy reflect sex. The rsFC sex classifier trained on the aligned participants was able to classify unseen aligned participants from the independent testing set as male or female with 85 % accuracy ($p < 0.001$; Fig. 2 A, 2B). Sensitivity, specificity, AUC, and MCC of the model were 0.83, 0.87, 0.92, and 0.70 respectively. Variations in the functional organization of the visual (visual and medial visual), default mode, dorsal attention, and parietal memory networks, in that order, contributed the most to the model and were therefore relatively more important in predicting participant sex (Fig. 2 C, 2D).

The cortical thickness sex classifier was trained on the same aligned participants as the rsFC classifier. It correctly separated aligned males from females with an accuracy of 76 % ($p < 0.001$; Fig. 2B). Sensitivity, specificity, AUC, and MCC of the model were 0.77, 0.74, 0.83, and 0.52 respectively. Variation in the anatomical organization of cortical parcels belonging to the frontoparietal, dorsal attention, visual (medial visual and visual), and action-mode networks, in that order, contributed the most to the model and were therefore relatively more important in

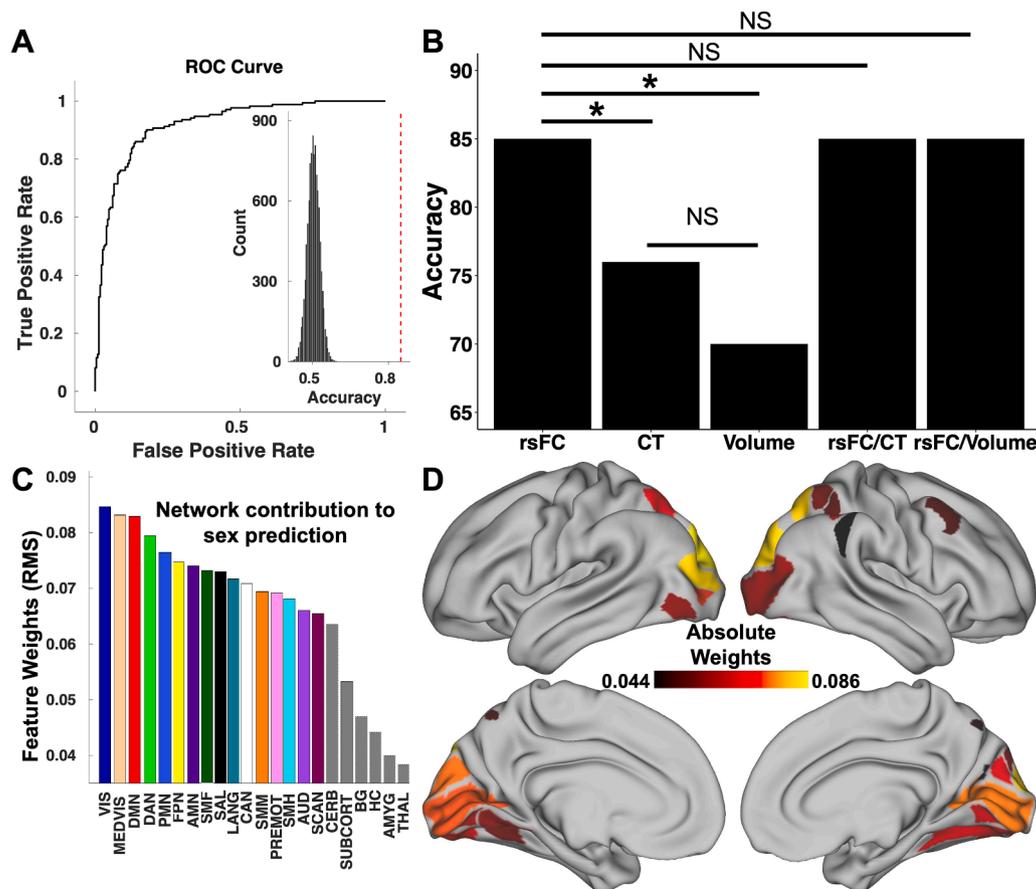


Fig. 2. Brain pattern analysis using support vector machine (SVM) learning predicts participant sex based on functional connectivity. (A) SVMs with nested five-fold cross-validation (5F-CV) were used to construct univariate and multivariate models that classified participants as male or female (sex assigned at birth). These models used resting-state functional connectivity (rsFC), cortical thickness, cortical volume, rsFC/cortical thickness, and rsFC/cortical volume combined as predictive features. The receiver operating characteristic (ROC) curve of the rsFC resulting model is depicted. The model classified participants as male or female with 85 % accuracy. Inset histogram shows distribution of permutated accuracies. The accuracy from real (nonpermutated) data is represented by the dashed red line. (B) McNemar's tests comparing the rsFC, cortical thickness, cortical volume, and combined rsFC/cortical thickness and rsFC/cortical volume sex classifiers revealed a statistically significant difference in performance between rsFC and cortical thickness ($\chi^2 = 7.36$, $p < 0.01$) and between rsFC and cortical volume ($\chi^2 = 19.67$, $p < 0.001$). However, the performance differences between cortical thickness and cortical volume, rsFC and rsFC/cortical thickness, and rsFC and rsFC/cortical volume sex classifiers were not significant. (C) To understand which networks contributed the most to the prediction, the parcel-wise feature weights' root mean square was calculated and then averaged for each network. The most important features in the rsFC model were found in the visual (visual and medial visual), default mode, dorsal attention, and parietal memory networks. (D) The top 10 % of cortical parcels in terms of feature importance in the rsFC SVM model. The parcel-wise feature weights' root mean square was calculated for each parcel. rsFC, resting-state functional connectivity; CT, cortical thickness; NS, not significant; DMN, default mode network; VIS, visual network; MEDVIS, medial visual network; FPN, frontoparietal network; DAN, dorsal attention network; LANG, language network; SAL, salience network; PMN, parietal memory network; AMN, action-mode network; PREMOT, premotor network, SMH, somatomotor hand network; SMM, somatomotor mouth network; SMF, somatomotor foot network; SCAN, somato-cognitive action network; AUD, auditory network; CAN, contextual association network; HC, hippocampus; AMG, amygdala; BG, basal ganglia; THAL, thalamus; CERB, cerebellum.

predicting participant sex.

Similarly, the cortical volume sex classifier was trained on the same aligned participants as the rsFC and cortical thickness classifiers. It correctly separated aligned males from females with an accuracy of 70 % ($p < 0.001$; Fig. 2B). Sensitivity, specificity, AUC, and MCC of the model were 0.68, 0.72, 0.79, and 0.39 respectively. Variation in the anatomical organization of cortical parcels belonging to the dorsal attention, frontoparietal, visual, premotor, and somatomotor mouth networks, in that order, contributed the most to the model and were therefore relatively more important in predicting participant sex.

The combined rsFC/cortical thickness sex classifier was similarly trained on the same set of aligned participants. It correctly separated aligned males from females with an accuracy of 85 % ($p < 0.001$; Fig. 2B). Sensitivity, specificity, AUC, and MCC of the model were 0.83, 0.87, 0.92, and 0.69 respectively. Variations in the organization of the medial visual, default mode, visual, dorsal attention, and frontoparietal networks, in that order, contributed the most to the model and were therefore relatively more important in predicting participant sex.

The combined rsFC/cortical volume, similarly, trained on the same set of aligned participants, correctly separated aligned males from females with an accuracy of 85 % ($p < 0.001$; Fig. 2B). Sensitivity, specificity, AUC, and MCC of the model were 0.84, 0.87, 0.92, and 0.70 respectively. Variations in the organization of the visual, dorsal attention, frontoparietal, action-mode, and default mode networks, in that order, contributed the most to the model and were therefore relatively more important in predicting participant sex.

The rsFC sex classifier exhibited significantly superior performance in predicting aligned individuals compared to the cortical thickness classifier (McNemar's test: $\chi^2 = 7.36$, $p < 0.01$; Fig. 2B) and the cortical volume classifier (McNemar's test: $\chi^2 = 19.67$, $p < 0.001$; Fig. 2B). While the cortical thickness sex classifier achieved numerically higher accuracy compared to the cortical volume classifier, that difference did not reach statistical significance (McNemar's test: $\chi^2 = 2.87$, $p = 0.09$; Fig. 2B). The combined rsFC/cortical thickness and rsFC cortical volume sex classifiers performed equally well as the rsFC classifier (both McNemar's tests: $\chi^2 < 0.001$, $p = 0.99$; Fig. 2B) and both sex classifiers performed significantly better than the cortical thickness (McNemar's test: $\chi^2 = 6.89$, $p < 0.001$; Fig. 2B) and cortical volume (McNemar's test: $\chi^2 = 20.45$, $p < 0.001$; Fig. 2B) sex classifiers respectively. Lastly, the combined rsFC/cortical thickness sex classifier did not perform significantly better than the combined rsFC/cortical volume sex classifier (McNemar's test: $\chi^2 < 0.01$, $p = 0.93$; Fig. 2B).

3.3. Sex classifiers' efficacy in predicting aligned and unaligned individuals

Although the rsFC sex classifier demonstrated superior predictive performance for sex within the aligned test group, surpassing both the cortical thickness and cortical volume classifiers, we examined all three – as well as the combined rsFC/cortical thickness and rsFC/cortical volume classifiers – to assess their efficacy in predicting the sex of individuals with unaligned sex/gender. The rsFC sex classifier trained on the aligned participants was able to classify unseen unaligned participants as male or female with 78 % accuracy ($p < 0.001$). Sensitivity, specificity, AUC, and MCC of the model were 0.84, 0.74, 0.88, and 0.56 respectively. The rsFC SVM model predicting the independent aligned group achieved statistically significantly higher prediction accuracy (85 %) than predicting the unaligned group (78 %; $z = 2.99$, $p < 0.01$). In both the aligned testing set and the unaligned group, we found no significant differences in pubertal development stage (aligned: $z = 0.60$, $p = 0.55$, unaligned: $z = 0.57$, $p = 0.57$) or income-to-needs (aligned: $t = -1.30$, $p = 0.19$, unaligned: $t = -0.49$, $p = 0.63$) between correctly and incorrectly classified subjects.

The cortical thickness sex classifier trained on the aligned participants was able to classify unseen unaligned participants as male or female with 72 % accuracy ($p < 0.001$; Sensitivity, specificity, AUC, MCC

= 0.80, 0.67, 0.81, and 0.45). Although the cortical thickness SVM model achieved higher numerical accuracy for the aligned independent testing set (76 %) compared to the unaligned group (72 %), this difference was not statistically significant ($z = 1.78$, $p = 0.08$).

Similarly, the cortical volume sex classifier trained on the aligned participants was able to classify unseen unaligned participants as male or female with 70 % accuracy ($p < 0.001$; Sensitivity, specificity, AUC, MCC = 0.82, 0.64, 0.80, and 0.44). The cortical volume SVM model had achieved the same numerical accuracy for the aligned independent testing set (70 %) and the unaligned group (70 %), with no statistically significant effect ($z = -0.20$, $p = 0.85$).

The combined rsFC/cortical thickness sex classifier trained on the aligned participants was able to classify unseen unaligned participants as male or female with 77 % accuracy ($p < 0.001$; Sensitivity, specificity, AUC, MCC = 0.85, 0.73, 0.88, and 0.56). Like the rsFC sex classifier, the rsFC/cortical thickness SVM model achieved a statistically significantly higher prediction accuracy for the aligned independent testing set (85 %) compared to the unaligned group (77 %; $z = 3.06$, $p < 0.01$).

Lastly, the combined rsFC/cortical volume sex classifier trained on the aligned participants was able to classify unseen unaligned participants as male or female with 78 % accuracy ($p < 0.001$; Sensitivity, specificity, AUC, MCC = 0.84, 0.74, 0.88, and 0.56). Similar to the rsFC sex classifier, the rsFC/cortical thickness SVM model achieved a statistically significantly higher prediction accuracy for the aligned independent testing set (85 %) compared to the unaligned group (78 %; $z = 3.10$, $p < 0.01$).

Just as for aligned individuals, the rsFC sex classifier exhibited significantly superior performance in predicting unaligned individuals compared to the cortical thickness (McNemar's test: $\chi^2 = 8.44$, $p < 0.01$) and cortical volume (McNemar's test: $\chi^2 = 12.19$, $p < 0.001$) sex classifiers. While the cortical thickness sex classifier achieved numerically higher accuracy compared to the cortical volume classifier in predicting unaligned individuals, that difference did not reach statistical significance (McNemar's test: $\chi^2 = 0.30$, $p = 0.58$). The combined rsFC/cortical thickness and rsFC cortical volume sex classifiers performed equally well as the rsFC classifier (McNemar's tests: $\chi^2 = 0.03$, $p = 0.86$ and $\chi^2 < 0.01$, $p = 0.96$) and both sex classifiers performed significantly better than the cortical thickness (McNemar's test: $\chi^2 = 7.20$, $p < 0.01$) and cortical volume (McNemar's test: $\chi^2 = 12.19$, $p < 0.001$) sex classifiers respectively. Lastly, the combined rsFC/cortical thickness sex classifier did not perform significantly better than the combined rsFC/cortical volume sex classifier (McNemar's test: $\chi^2 = 0.03$, $p = 0.86$).

Since combining rsFC and cortical thickness did not significantly improve the rsFC classification model, the separate rsFC, cortical thickness, and cortical volume models were used for the remainder of the classification-related analyses (see 3.4 Sex classification and sex/gender alignment correlation).

The validation analyses, in which rsFC, cortical thickness, cortical volume, and combined rsFC/cortical thickness and rsFC/cortical volume SVM models were trained on unaligned participants and tested on unaligned and aligned participants, revealed a pattern of results consistent with the main analyses (see Supplementary Results). The enhanced performance accuracy of the rsFC SVM models trained exclusively on aligned participants likely stems from the use of more precisely defined labels, which minimize label noise and enable the classifier to more effectively distinguish between sexes. In contrast, the smaller sample size in the unaligned training group ($n = 893$ vs $n = 1610$ aligned training group; Supplementary Fig. 2) likely reduced statistical power, potentially limiting the ability to detect significant differences in performance across models.

The rsFC sex classifier trained on the aligned participants using the data preprocessed without GSR was able to classify unseen aligned participants from the independent testing set as male or female with 83.4 % accuracy ($p < 0.001$). Sensitivity, specificity, AUC, and MCC of the model were 0.82, 0.85, 0.92, and 0.67 respectively. Variations in the

functional organization of the premotor, somato-cognitive action, auditory, and visual (visual and medial visual) networks, in that order, contributed the most to the model and were therefore relatively more important in predicting participant sex. The same classifier was able to classify unseen unaligned participants as male or female with 78.3 % accuracy ($p < 0.001$). Sensitivity, specificity, AUC, and MCC of the model were 0.85, 0.75, 0.88, and 0.57 respectively. The rsFC SVM model predicting the independent aligned group achieved statistically significantly higher prediction accuracy (83.4 %) than predicting the unaligned group (78.3 %; $z = 2.16, p = 0.03$).

3.4. Sex classification and sex/gender alignment correlation

The above findings indicate robust differences in brain functional connectivity (rsFC) and neuroanatomy (cortical thickness and cortical volume) between males and females. The relationship between the rsFC, cortical thickness, and cortical volume sex classification scores (brain profile predictions) and sex/gender alignment scores was then investigated to evaluate the consistency between the adolescents' brain profiles and sex/gender alignment. The classification scores derived from the rsFC, cortical thickness, and cortical volume sex classifiers serve as quantitative indicators of the degree to which a brain matches either female or male patterns. Higher scores suggest a stronger correspondence to the neural connectivity patterns associated with the sex commonly labeled as either female or male. A significant positive correlation was revealed in females between the rsFC classification testing set scores and the sex/gender alignment scores obtained from the Youth Self-Report Gender Questionnaire ($\rho = 0.18, p < 0.001$, Fig. 3 A). This suggests that the females with functional connectivity patterns more similar to the typical female pattern exhibit greater sex/gender alignment as reflected in higher scores on the Youth Self-Report Gender Questionnaire. In males, the inverse association was observed, with individuals with functional connectivity patterns more similar to the typical male pattern exhibiting lower sex/gender alignment ($\rho = -0.11, p < 0.01$; Fig. 3B). Nearly identical correlations between the non-GSR rsFC classification testing set scores and the sex/gender alignment scores were observed in both females ($\rho = 0.17, p < 0.001$) and males ($\rho = -0.11, p < 0.01$).

The correlation between cortical thickness classification and sex/gender alignment scores was not significant for females ($\rho = 0.04, p = 0.24$) but was significant for males ($\rho = -0.09, p = 0.03$) exhibiting

the same inverse relationship as with the rsFC brain patterns. Similarly, the correlation between cortical volume classification and sex/gender alignment scores was not significant for females ($\rho = 0.04, p = 0.24$) but was significant for males ($\rho = -0.09, p = 0.03$).

3.5. Functional connectivity and neuroanatomical prediction of sex/gender alignment

Next, we evaluated whether an adolescent's degree of sex/gender alignment can be predicted by brain neuroanatomy or functional connectivity using linear SVR models. Neither the functional connectivity (rsFC), neuroanatomy (cortical thickness and cortical volume), nor the combined rsFC/cortical thickness and rsFC/cortical volume sex/gender alignment SVR models successfully predicted the sex/gender alignment Youth Self-Report or Parent-Report Gender Questionnaire scores. Correlations between the original sex/gender alignment scores and predicted sex/gender alignment scores were not significant (all p -values > 0.05), except for the correlation between original parent-reported sex/gender alignment scores and predicted scores in males ($p = 0.03$). Importantly, all coefficients of determination were negative ($R^2 < 0$), including the exception ($R^2 = -0.41$), suggesting that the models performed worse than a null model predicting the mean of the dependent variable (sex/gender alignment scores) for all observations.

As a robustness check measure, linear ridge regressions, as performed by Dhamala et al. (2024) to explore functional brain correlates of sex and gender, were also tested. The findings were the same as in our rsFC SVR model. Correlations between the original Youth Self-Report and Parent-Report Gender Questionnaire gender scores and predicted gender scores were not significant (all p -values > 0.05) and all coefficients of determination were negative ($R^2 < 0$).

4. Discussion

In this study, we leveraged machine learning and a large sample of youths to determine whether variation in functional connectivity and neuroanatomy can predict sex (assigned at birth) and/or sex/gender alignment, and whether this variability is reflected in the degree of sex/gender alignment. Using complex univariate and multivariate patterns of rsFC, cortical thickness, and cortical volume we were able to predict an unseen participant's sex with a high degree of accuracy. We first demonstrated that rsFC is a better predictor of sex compared to cortical

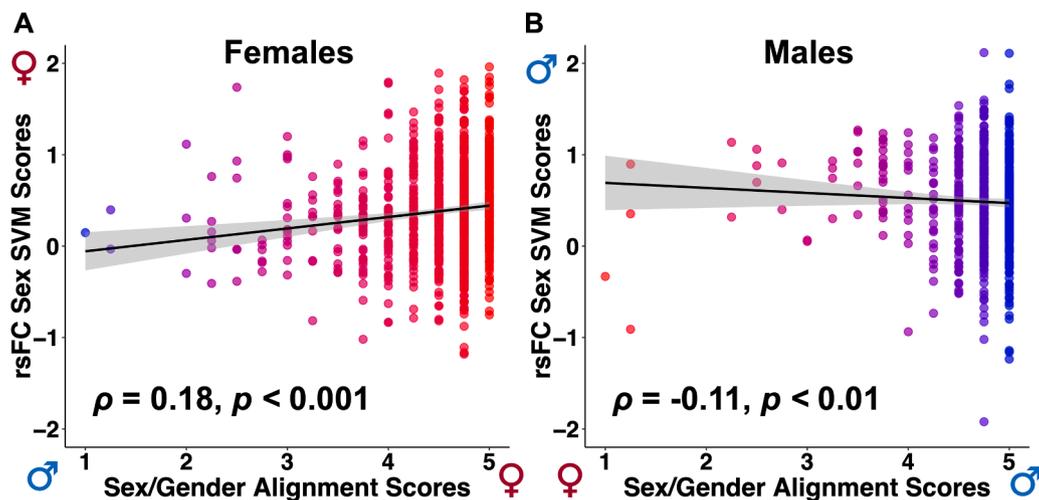


Fig. 3. Correlations between resting-state functional connectivity (rsFC) support vector machine (SVM) classification scores and sex/gender alignment scores. (A) A significant positive correlation was observed in females between the rsFC sex classification scores and the sex/gender alignment scores. (B) A significant negative correlation between the rsFC classification and sex/gender alignment scores was observed in males. Higher sex/gender alignment scores (x-axis) indicate greater sex/gender alignment. Higher SVM scores (y-axis) indicate a stronger correspondence to the neural connectivity patterns associated with the sex labeled as either female or male. Results in bold indicate statistical significance.

thickness and cortical volume. Combining rsFC with cortical thickness or rsFC with cortical volume did not enhance the rsFC model's performance. We then demonstrated that while sex differences in brain functional connectivity and neuroanatomy were most pronounced in visual networks, the majority occurred within association networks. Moreover, SVM analyses revealed that the rsFC of sex/gender aligned youth is more effective at predicting an unseen aligned participant's sex compared to predicting the sex of an unaligned participant. In contrast, the cortical thickness and cortical volume SVM classifiers failed to distinguish between aligned and unaligned participants. Furthermore, we found a positive relationship between the degree to which an individual's rsFC profile reflects female traits and sex/gender alignment in females, but a negative relationship in males. Lastly, we found that neither functional connectivity nor neuroanatomy predict sex/gender alignment.

4.1. rsFC outperforms cortical thickness and cortical volume as a predictor of sex

One of our principal findings is that rsFC demonstrates superior predictive power compared to cortical thickness and cortical volume in accurately classifying sex, although all modalities showed a good degree of prediction. While the cortical thickness sex classifier achieved numerically higher accuracy compared to the cortical volume classifier, that difference was not statistically significant. Prior studies have demonstrated similar accuracies using network-level resting-state fMRI data to discriminate between male and female youth (Shanmugan et al., 2022), achieving slightly higher accuracies when combined with anatomical data (Adeli et al., 2020; Brennan et al., 2021; Kurth et al., 2020; Sepehrband et al., 2018). In our analysis, however, combining rsFC and cortical thickness was better than cortical thickness alone, but did not show a significant improvement in prediction accuracy compared to the rsFC-only model. The same was observed with cortical volume, where combining rsFC with cortical volume did not significantly enhance prediction accuracy over the rsFC-only model. The superior predictive power of rsFC suggests that, even in a young cohort, neural activity patterns during rest capture subtle, sex-specific functional dynamics that are not evident in structural measures like cortical thickness. This highlights rsFC's potential as a more sensitive biomarker in neurodevelopmental studies. Additionally, the fact that rsFC outperforms cortical thickness and cortical volume as a predictor could imply that the information provided by cortical thickness and cortical volume overlaps with what is already captured by rsFC. As a result, cortical thickness and cortical volume may not offer additional or unique information to enhance prediction accuracy.

4.2. Sex differences in brain functional connectivity and neuroanatomy are greatest in association networks

Our findings reveal significant sex-related variability in brain functional connectivity and neuroanatomy, particularly within association networks, an observation consistent with prior research (Bijsterbosch et al., 2018; Cui et al., 2020; Gordon et al., 2017; Kong et al., 2019; Li et al., 2017; Shanmugan et al., 2022). The most pronounced differences in rsFC between sexes were found in visual (visual and medial visual), default mode, dorsal attention, and parietal memory networks (Filippi et al., 2013; Mathew et al., 2020; Murray et al., 2018; Shaqiri et al., 2018; Vanston and Strother, 2017; Weis et al., 2019). Studies suggest that men and women may process visual information differently (Mathew et al., 2020; Murray et al., 2018; Shaqiri et al., 2018; Vanston and Strother, 2017), though neuroimaging research in adults presents mixed results – some identify sex differences in sensory networks like the visual network (Biswal et al., 2010; Filippi et al., 2013), while others do not (Allen et al., 2011; Tomasi and Volkow, 2012). The functional organization of association networks, which are linked to emotional, social, and executive functions (Cui et al., 2020; Kong et al., 2019), also demonstrates variability tied to sex differences (Alarcón et al., 2020;

Andreato and Cahill, 2009; Cahill, 2010). For instance, variations in default mode network connectivity have been associated with sex hormones such as estrogen and progesterone (Engman et al., 2016; Hidalgo-Lopez et al., 2020; Petersen et al., 2014; Weis et al., 2019). Our findings further support the role of the dorsal attention and parietal memory networks in predicting sex differences, which also corresponds with prior research (Dumais et al., 2018; Filippi et al., 2013; Hill et al., 2014; Hjelmervik et al., 2014; Stumme et al., 2020; Zhang et al., 2018). These sex differences in rsFC could relate to variations in cognitive processes and behavioral outcomes observed between males and females, influencing domains such as decision-making, emotion regulation, and attention allocation (Stoica et al., 2021). Our results partially align with the work of Dhamala et al. (2024), who also utilized the ABCD dataset to explore functional brain correlates of sex and gender, finding that sex is strongly associated with connectivity patterns within and between the somatomotor, visual, control, and limbic networks (Yeo et al., 2011). However, it is also possible that societal mores that influence engagement with certain cognitive activities also shape brain functional connectivity, leading to patterns that differ between sexes that are not biologically determined.

The discrepancy in networks contributing most to classification between the GSR and non-GSR models (visual, medial visual, default mode, dorsal attention, and parietal memory networks with GSR; premotor, somato-cognitive action, auditory, visual, and medial visual without GSR) likely reflects the impact of GSR on the covariance structure of large-scale networks. By attenuating widespread global variance, GSR can enhance network-specific contrasts and increase the apparent contribution of association networks (e.g., default mode, dorsal attention, and parietal memory), whereas non-GSR data retain more global and sensorimotor-driven variance, increasing the apparent importance of sensory and motor networks (e.g., premotor, auditory, somato-cognitive action).

4.3. rsFC of sex/gender aligned youth is less accurate in predicting unaligned youths' sex

We found that that observed differences in rsFC between males and females may not entirely be driven by biology (sex) but also from the influence of social constructs (gender) (Eliot, 2011; Eliot et al., 2021). Using a machine learning neuroimaging approach, we showed that an SVM classifier trained to learn sex from rsFC data using only youth with sex/gender alignment, was significantly more accurate in predicting sex in unseen participants with sex/gender alignment than in those with sex/gender unalignment. We further demonstrated that the performance differences in our rsFC SVM model were driven by sex/gender alignment status, as there were no significant differences in pubertal development stage or income-to-needs ratio between correctly classified (true positives and true negatives) and incorrectly classified (false positives and false negatives) subjects. These findings suggest subtle yet detectable differences in rsFC patterns between individuals with aligned and unaligned sex and gender. Notably, cortical thickness and cortical volume SVM classifiers trained to learn sex using only youth with sex/gender alignment failed to differentiate effectively, as they were equally accurate in predicting sex in unseen participants with sex/gender alignment and those with unalignment. This might imply that gender has greater influence on functional connectivity brain patterns than on brain neuroanatomy (i.e., cortical thickness and cortical volume). These results could be considered distinct from research suggesting that the structural brain profile of hormonally untreated transgender individuals resembles that of their gender identity or is intermediate between sexes (Flint et al., 2020; Kranz et al., 2014; Kurth et al., 2022; Mueller, De Cuypere, et al., 2017; Mueller, Landré, et al., 2017; Rametti, Carrillo, Gómez-Gil, Junque, Segovia, et al., 2011; Rametti, Carrillo, Gómez-Gil, Junque, Zubiarré-Elorza, et al., 2011). However, those studies reported on white matter microstructure and gray matter volume of subcortical structures, and not cortical thickness or cortical volume.

The reduced accuracy of the rsFC classifier when applied to gender-diverse youth highlights that the neurobiological patterns linked to sex in youth with sex/gender alignment do not generalize well to those with sex/gender unalignment. This underscores the need to consider gender diversity in neuroscientific research and cautions against relying on brain-based classifiers trained exclusively on sex/gender aligned individuals.

4.4. Female youth with more sex-specific brain patterns exhibit greater sex/gender alignment

We found that females with higher rsFC sex SVM scores (greater brain “femaleness” on the “maleness-femaleness” continuum) also had higher sex/gender alignment scores on the Youth Self-Report Gender Questionnaire (Fig. 3 A). In males, however, there was a surprising negative correlation: males with higher rsFC sex SVM scores (greater brain “maleness”) showed lower sex/gender alignment (Fig. 3B). One possible explanation for this discrepancy may be linked to lower social tolerance for gender nonconformity in young males compared to females (Potter et al., 2021, 2022). The sex/gender alignment scores from the Youth Self-Report Gender Questionnaire rely on self-reported responses. Males, even from preadolescence, might feel pressured to answer gender-related questions less truthfully. Consequently, males with rsFC brain patterns less similar to the typical male pattern might still report high sex/gender alignment scores. Thus, the groups of males with higher sex/gender alignment scores may include participants whose sex and gender are not actually aligned. Two observations reinforce this possibility: 1) There were significantly more aligned males than females in the sample (1172 vs. 873; Table 1), and Supplementary Table 2) unaligned females had overall lower Youth Self-Report Gender Questionnaire scores compared to unaligned males (Supplementary Fig. 1), both patterns also observed in the entire ABCD dataset (Potter et al., 2022). A question is thus raised of whether the Youth Self-Report Gender Questionnaire is measuring gender equally effectively for both males and females.

While rsFC successfully captured the relationship between brain patterns and sex/gender alignment in both males and females, cortical thickness and cortical volume only showed this relationship in males, reflecting the same negative correlation. The lack of a significant correlation between the cortical thickness and cortical volume sex SVM scores and sex/gender alignment scores in females underscores the cortical thickness’s and cortical volume’s inability to effectively capture the relationship between adolescents’ brain profiles and sex/gender alignment.

4.5. Functional connectivity and neuroanatomy do not predict sex/gender alignment

Neither the rsFC, cortical thickness, cortical volume, nor combined rsFC/cortical thickness and rsFC/cortical volume SVR models were able to capture the variance in sex/gender alignment. A recent study by Dhamala et al. (2024) demonstrated that sex/gender alignment, as measured by the Parent-Report Gender Questionnaire, could be predicted using rsFC and linear ridge regression. Although both our study and theirs used the ABCD dataset, our findings do not replicate theirs. To assess the robustness of our findings, we also implemented linear ridge regression using online scripts provided by Dhamala et al. (2024). Similar to their approach, separate models were constructed for males and females. rsFC was used as the predictor and sex/gender alignment scores were used as labels. Our robustness check results further confirmed that rsFC is unable to predict sex/gender alignment, as measured by the Youth Self-Report and Parent-Report Gender Questionnaires. The divergence in findings could be attributed to several differences in our analyses including the use of different cortical (400 (Schaefer et al., 2018) vs. 333 (Gordon et al., 2016)) and noncortical (19 (Fischl et al., 2002) vs. 61 (Seitzman et al., 2020)) parcels, and choices

that led to different final ABCD samples (e.g., Dhamala et al. 2024 excluded individuals with less than 4 min of data – final sample $n = 4757$, while we excluded subjects with less than 8 min of data – final sample $n = 3129$). Additionally, they band-pass filtered at $0.009 \text{ Hz} \leq f \leq 0.08 \text{ Hz}$ while we did at $0.008 \text{ Hz} < f < 0.10 \text{ Hz}$.

Although rsFC effectively captures the relationship between sex and gender (significant accuracy difference when predicting aligned vs. unaligned individuals), the inability to straightforwardly predict sex/gender alignment from brain biomarkers suggests that gender may be a more complex construct that is not as clearly reflected in functional connectivity, cortical thickness or cortical volume patterns. Another possibility would be that linear regression models may not be able to capture potential non-linear gender effects. Additionally, the challenge of capturing variance in gender could be attributed to the overall scarcity of population-level variability in gender scores. It is possible that developmental factors play a role in this phenomenon, and the results might differ if we were examining a population at a more advanced pubertal stage. The inability to replicate findings from a study that reported sex/gender alignment prediction from rsFC using the same dataset suggests that more research is necessary to confirm the robustness and reliability of neural predictors of gender versus sex.

4.6. Limitations and future directions

This study has certain limitations that should be acknowledged. First, sex was assessed using a binary parent-report question. It is important to acknowledge that existing evidence and theory suggest that binary sex classification may be suboptimal, since human brains are largely composed of unique mosaics of female-typical and male-typical features (Fine, 2014; Joel, 2020). Second, the self-reported gender data displayed minimal variability, meaning that the self-reported gender data showed limited diversity in responses (Supplementary Fig. 1), with most participants reporting gender identities aligning with their sex assigned at birth. This lack of variability likely restricted the range of sex/gender alignment values in the dataset, which in turn may have reduced the ability to detect potential relationships between functional connectivity, neuroanatomy, and the degree of sex/gender alignment. Future work should include individuals with varying levels of gender nonconformity to enhance the ability to capture these relationships. Additionally, a deeper investigation into the relative contributions of puberty stage and rsFC to sex/gender alignment, as well as potential interactions with gender roles, would be an important and valuable direction for future research. This would help clarify how these factors interact and contribute to the broader understanding of sex/gender alignment and its neural correlates. Third, we chose to dichotomize sex/gender alignment and variable dichotomization has recognized pitfalls (Maccallum et al., 2002). Dichotomizing can obscure subtle gradations in sex/gender alignment and potentially misrepresent differences between groups. However, we only dichotomized sex/gender alignment for use in our SVM models, which aimed to predict sex. From a machine learning perspective, training SVM classifiers only on aligned individuals—who represent a more homogeneous subpopulation in terms of sex/gender alignment—reduces label noise and enables the classifier to more accurately distinguish between sexes. Establishing this baseline performance with optimized label quality provides a benchmark for evaluating classifier efficacy under more heterogeneous conditions. This approach was intended to explore whether brain structure or function could, indirectly, capture sex/gender alignment. Importantly, for our SVR analyses, where the objective was to directly predict sex/gender alignment, we utilized the continuous sex/gender alignment measure. Fourth, there is a possibility that the Youth Self-Report Gender Questionnaire may not be measuring gender equally effectively for both males and females. This is based on the different score distribution between males and females which could be due to greater social tolerance to nonconformity in young females compared to males (Potter et al., 2021, 2022). Future studies need to consider a measuring tool that captures gender identity,

dysphoria, and expression more effectively for both sexes. Fifth, gender is shaped significantly by social and cultural factors, which can influence how gender identity is expressed and understood across different societies. The ABCD dataset, being collected entirely in the United States, reflects only the experiences and environments of individuals within this specific cultural and societal context, which may not be representative of the world population (Ricard et al., 2022). Additionally, gender identity is often intertwined with a variety of societal factors such as socioeconomic status, exposure to adversity, and family dynamics (Barboza-Salerno and Meshelemiah, 2024; Carpenter et al., 2020; Price et al., 2023), all of which can contribute to differences in brain development. Because of these complexities, future research should explore whether the relationship between brain development and gender is consistent across different cultures, taking into account the multitude of factors (e.g., adversity, societal expectations, exposure to different cultural gender norms) that may mediate or moderate these associations. This will be crucial for determining whether observed differences in model performance across sexes are truly gender-related or whether they are driven by other social, environmental, and cultural factors. Lastly, this study provided a limited evaluation of subcortical and cerebellar networks. Subsequent analyses should focus on assessing functional and neuroanatomical sex differences in these networks, as they are as crucial as the cortex for understanding human behavior and function.

5. Conclusions

In summary, our study demonstrates that brain functional connectivity (i.e., rsFC) captures the relationship between sex and gender more effectively compared to cortical thickness and cortical volume, as well as the extent to which sex/gender alignment correlates with a brain sex profile. While we observed normative sex differences in functional connectivity, cortical thickness, and cortical volume in youth, functional connectivity proved to be a more sensitive predictor of sex, with association networks playing a central role in these differences. Additionally, we found limitations in using functional connectivity and neuroanatomy to predict sex/gender alignment, highlighting the complexity of these concepts. Our findings challenge binary models of brain-sex differences, advocating for more inclusive and nuanced approaches to studying the relationship between gender and neurobiology.

CRediT authorship contribution statement

Julia Monk: Writing – review & editing. **Roselyne J. Chauvin:** Writing – review & editing. **Kristen M. Scheidter:** Writing – review & editing. **Athanasia Metoki:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Vahdeta Suljic:** Writing – review & editing. **Aristeidis Sotiras:** Writing – review & editing, Supervision, Methodology. **Samuel R. Krimmel:** Writing – review & editing. **Deanna M. Barch:** Writing – review & editing, Supervision, Methodology. **Benjamin P. Kay:** Writing – review & editing. **Nadeshka J. Ramirez-Perez:** Writing – review & editing. **Timothy O. Laumann:** Writing – review & editing. **Forrest I. Whiting:** Writing – review & editing. **Evan M. Gordon:** Writing – review & editing. **Noah J. Baden:** Writing – review & editing. **Andrew N. Van:** Writing – review & editing. **Babatunde Adeyemo:** Validation. **Anxu Wang:** Writing – review & editing. **Nico U.F. Dosenbach:** Writing – review & editing, Supervision, Methodology. **Scott Marek:** Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests. E. M.G. may receive royalty income based on technology developed at

Washington University School of Medicine and licensed to Turing Medical Inc. N.U.F.D. has a financial interest in Turing Medical Inc. and may benefit financially if the company is successful in marketing Framework Integrated Real-Time Motion Monitoring (FIRMM) software products. N.U.F.D. may receive royalty income based on FIRMM technology developed at Washington University School of Medicine and Oregon Health and Sciences University and licensed to Turing Medical Inc. N.U.F.D. is a co-founder of Turing Medical Inc. TOL is a consultant for Turing Medical Inc. TOL holds a patent for taskless mapping of brain activity licensed to Sora Neurosciences and a patent for optimizing targets for neuromodulation, implant localization, and ablation is pending. These potential conflicts of interest have been reviewed and are managed by Washington University School of Medicine. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by NIH grants T32 MH100019 (A.M.), MH122066 (E.M.G.), MH121276 (E.M.G.), NS129521 (E.M.G.), MH129616 (T.O.L.), NS123345 (B.P.K.), MH121518 (S.M.), MH096773 (N.U.F.D.), MH122066 (N.U.F.D.), MH121276 (N.U.F.D.), MH124567 (N.U.F.D.), NS129521 (N.U.F.D.), and NS088590 (N.U.F.D.); by the National Spasmodic Dysphonia Association (E.M.G.); by the Taylor Family Institute for Innovative Psychiatric Research (T.O.L.); by the Intellectual and Developmental Disabilities Research Center (N.U.F.D.); by the Kiwanis Foundation (N.U.F.D.); by the Washington University Hope Center for Neurological Disorders (E.M.G. and N.U.F.D.); and by Mallinckrodt Institute of Radiology pilot funding (N.U.F.D.).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.dcn.2025.101624](https://doi.org/10.1016/j.dcn.2025.101624).

Data availability

The data used for this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (<https://abcdstudy.org>), held in the National Institute of Mental Health Data Archive. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award nos. U01DA041022, U01DA041025, U01DA041028, U01DA041048, U01DA041089, U01DA041093, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U01DA050987, U01DA050988, U01DA050989, U01DA051016, U01DA051018, U01DA051037, U01DA051038, U01DA051039, U24DA041123, U24DA041147. A listing of supporters is available at <https://abcdstudy.org/federal-partners/> and a full list of participating sites and study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this article. This manuscript reflects the views of the authors and may not reflect the opinions or views of the National Institutes of Health or ABCD consortium investigators. The ABCD data repository grows and changes over time. All ABCD Study data can be accessed via the National Institute of Mental Health Data Archive. All scripts and code used to generate the results have been archived and can be found on GitHub (https://github.com/nasiametoki/brain_sex_gender.git).

References

- Adeli, E., Zhao, Q., Zahr, N.M., Goldstone, A., Pfefferbaum, A., Sullivan, E.V., Pohl, M., 2020. Deep learning identifies morphological determinants of sex differences in the

- pre-adolescent brain. *NeuroImage* 223, 117293. <https://doi.org/10.1016/j.neuroimage.2020.117293>.
- Akosa, J., 2017. Predictive accuracy: a misleading performance measure for highly imbalanced data. *Proc. SAS Glob. Forum* 12, 1–4.
- Alarcón, G., Morgan, J.K., Allen, N.B., Sheeber, L., Silk, J.S., Forbes, E.E., 2020. Adolescent gender differences in neural reactivity to a friend's positive affect and real-world positive experiences in social contexts. *Dev. Cogn. Neurosci.* 43. <https://doi.org/10.1016/j.dcn.2020.100779>.
- Allen, E.A., Erhardt, E.B., Damaraju, E., Gruner, W., Segall, J.M., Silva, R.F., Havlicek, M., Rachakonda, S., Fries, J., Kalyanam, R., Michael, A.M., Caprihan, A., Turner, J.A., Eichele, T., Adelsheim, S., Bryan, A.D., Bustillo, J., Clark, V.P., Ewing, S.W.F., Calhoun, V.D., 2011. A baseline for the multivariate comparison of resting-state networks. *Front. Syst. Neurosci.* 5, 2093. <https://doi.org/10.3389/FNSYS.2011.00002/BIBTEX>.
- Allen, J.S., Damasio, H., Grabowski, T.J., Bruss, J., Zhang, W., 2003. Sexual dimorphism and asymmetries in the gray-white composition of the human cerebrum. *Neuroimage* 18, 880–894. [https://doi.org/10.1016/S1053-8119\(03\)00034-X](https://doi.org/10.1016/S1053-8119(03)00034-X).
- Andreano, J.M., Cahill, L., 2009. Sex influences on the neurobiology of learning and memory. *Learn. Mem.* 16 (4), 248–266. <https://doi.org/10.1101/LM.918309>.
- Avila-Varela, D.S., Hidalgo-Lopez, E., Dagnino, P.C., Acero-Pousa, I., del Agua, E., Deco, G., Pletzer, B., Eschrichs, A., 2024. Whole-brain dynamics across the menstrual cycle: the role of hormonal fluctuations and age in healthy women. *Npj Women's Health* 2 (1), 1–9. <https://doi.org/10.1038/s44294-024-00012-4>.
- Bale, T.L., Epperson, C.N., 2016. Sex as a biological variable: who, what, when, why, and how. *Neuropsychopharmacology* 42 (2), 386–396. <https://doi.org/10.1038/npp.2016.215>.
- Ball, W.S., Byars, A.W., Schapiro, M., Bommer, W., Carr, A., German, A., Dunn, S., Rivkin, M.J., Waber, D., Mulken, R., Vajapeyam, S., Chiverton, A., Davis, P., Koo, J., Marmor, J., Mrakotsky, C., Robertson, R., McAnulty, G., Brandt, M.E., O'Neill, J., 2012. Total and regional brain volumes in a Population-Based normative sample from 4 to 18 years: the NIH MRI study of normal brain development. *Cereb. Cortex* 22 (1), 1–12. <https://doi.org/10.1093/CERCOR/BHR018>.
- Bao, A.M., Swaab, D.F., 2010. Sex differences in the brain, behavior, and neuropsychiatric disorders. *Neuroscientist* 16 (5), 550–565. <https://doi.org/10.1177/1073858410377005>.
- Barboza-Salerno, G.E., Meshelemiah, J.C.A., 2024. Associations between early child adversity and lifetime suicide attempts among gender diverse individuals: a moderated mediation. *Child Abuse Negl.* 149, 106705. <https://doi.org/10.1016/J.CHIBU.2024.106705>.
- Baron-Cohen, S., Lombardo, M.V., Auyeung, B., Ashwin, E., Chakrabarti, B., Knickmeyer, R., 2011. Why are autism spectrum conditions more prevalent in males? *PLOS Biol.* 9 (6), e1001081. <https://doi.org/10.1371/JOURNAL.PBIO.1001081>.
- Benavides, A., Metzger, A., Tereshchenko, A., Conrad, A., Bell, E.F., Spencer, J., Ross-Sheehy, S., Georgieff, M., Magnotta, V., Nopoulos, P., 2018. Sex-specific alterations in preterm brain. *Pediatr. Res.* 85 (1), 55–62. <https://doi.org/10.1038/s41390-018-0187-5>.
- Bijsterbosch, J.D., Woolrich, M.W., Glasser, M.F., Robinson, E.C., Beckmann, C.F., Van Essen, D.C., Harrison, S.J., Smith, S.M., 2018. The relationship between spatial configuration and functional connectivity of brain regions. *ELife* 7. <https://doi.org/10.7554/ELIFE.32992>.
- Bird, K.A., Carlson, J., 2024. Typological thinking in human genomics research contributes to the production and prominence of scientific racism. *Front. Genet.* 15, 1345631. <https://doi.org/10.3389/FGENE.2024.1345631/BIBTEX>.
- Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.M., Ernst, M., Fair, D., Hampton, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kötter, R., Li, S.J., Milham, M.P., 2010. Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. USA* 107 (10), 4734–4739. https://doi.org/10.1073/PNAS.0911855107/SUPPL_FILE/PNAS.200911855SLPDP.
- Bramen, J.E., Hranilovich, J.A., Dahl, R.E., Forbes, E.E., Chen, J., Toga, A.W., Dinov, I.D., Worthman, C.M., Sowell, E.R., 2011. Puberty influences medial temporal lobe and cortical gray matter maturation differently in boys than girls matched for sexual maturity. *Cereb. Cortex* 21 (3), 636–646. <https://doi.org/10.1093/CERCOR/BHQ137>.
- Brennan, D., Wu, T., Fan, J., 2021. Morphometrical brain markers of sex difference. *Cereb. Cortex* 31 (8), 3641–3649. <https://doi.org/10.1093/CERCOR/BHAB037>.
- Bzdok, D., 2017. Classical statistics and statistical learning in imaging neuroscience. *Front. Neurosci.* 11, 543. <https://doi.org/10.3389/FNINS.2017.00543/BIBTEX>.
- Cahill, L., 2010. Sex influences on brain and emotional memory: the burden of proof has shifted. *Prog. Brain Res.* 186. <https://doi.org/10.1016/B978-0-444-53630-3.00003-8>.
- Carp, J., 2013. Optimizing the order of operations for movement scrubbing: comment on power et al. *NeuroImage* 76, 436–438. <https://doi.org/10.1016/j.neuroimage.2011.12.061>.
- Carpenter, C.S., Eppink, S.T., Gonzales, G., 2020. Transgender status, gender identity, and socioeconomic outcomes in the United States. *ILR Rev.* 73 (3), 573–599. https://doi.org/10.1177/0019793920902776/ASSET/IMAGES/LARGE/10.1177_0019793920902776-FIG1.JPEG.
- Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarcio, D.V., Garavan, H., Orr, C.A., Wager, T.D., Banich, M.T., Speer, N.K., Sutherland, M.T., Riedel, M.C., Dick, A.S., Bjork, J.M., Thomas, K.M., Dale, A.M., 2018. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. In: *Developmental Cognitive Neuroscience*, 32. Elsevier Ltd., pp. 43–54. <https://doi.org/10.1016/j.dcn.2018.03.001>.
- Chang, C., Glover, G.H., 2009. Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. *NeuroImage* 47 (4), 1448–1459. <https://doi.org/10.1016/J.NEUROIMAGE.2009.05.012>.
- Chen, J., Tam, A., Kebets, V., Orban, C., Ooi, L.Q.R., Asplund, C.L., Marek, S., Dosenbach, N.U.F., Eickhoff, S.B., Bzdok, D., Holmes, A.J., Yeo, B.T.T., 2022. Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. *Nat. Commun.* 13 (1), 1–17. <https://doi.org/10.1038/s41467-022-29766-8>.
- Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* 21 (1), 1–13. <https://doi.org/10.1186/S12864-019-6413-7/TABLES/5>.
- Chicco, D., Jurman, G., 2022. An invitation to greater use of Matthews correlation coefficient in robotics and artificial intelligence. *Front. Robot. AI* 9, 876814. <https://doi.org/10.3389/FROBT.2022.876814/BIBTEX>.
- Chicco, D., Jurman, G., 2023. The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Min.* 16 (1), 1–23. <https://doi.org/10.1186/S13040-023-00322-4/FIGURES/11>.
- Chicco, D., Warrens, M.J., Jurman, G., 2021. The Matthews correlation coefficient (MCC) is more informative than Cohen's kappa and brier score in binary classification assessment. *IEEE Access* 9, 78368–78381. <https://doi.org/10.1109/ACCESS.2021.3084050>.
- Chicco, D., Tötsch, N., Jurman, G., 2021. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Min.* 14 (1), 1–22. <https://doi.org/10.1186/S13040-021-00244-Z/TABLES/5>.
- Choleris, E., Galea, L.A.M., Sohrabji, F., Frick, K.M., 2018. Sex differences in the brain: implications for behavioral and biomedical research. In: *Neuroscience and Biobehavioral Reviews*, 85. Elsevier Ltd, pp. 126–145. <https://doi.org/10.1016/j.neubiorev.2017.07.005>.
- Cosgrove, K.P., Mazure, C.M., Staley, J.K., 2007. Evolving knowledge of sex differences in brain structure, function, and chemistry. *Biol. Psychiatry* 62 (8), 847–855. <https://doi.org/10.1016/j.biopsych.2007.03.001>.
- Cui, Z., Gong, G., 2018. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *Neuroimage* 178, 622–637. <https://doi.org/10.1016/j.neuroimage.2018.06.001>.
- Cui, Z., Li, H., Xia, C.H., Larsen, B., Adebimpe, A., Baum, G.L., Cieslak, M., Gur, R.E., Gur, R.C., Moore, T.M., Oathes, D.J., Alexander-Bloch, A.F., Raznahan, A., Roalf, D.R., Shinohara, R.T., Wolf, D.H., Davatzikos, C., Bassett, D.S., Fair, D.A., Satterthwaite, T.D., 2020. Individual variation in functional topography of association networks in youth. *Neuron* 106 (2), 340–353.e8. <https://doi.org/10.1016/j.neuron.2020.01.029>.
- De Bellis, M.D., Keshavan, M.S., Beers, S.R., Hall, J., Frustaci, K., Masalehdan, A., Noll, J., Boring, A.M., 2001. Sex differences in brain maturation during childhood and adolescence. *Cereb. Cortex* 11 (6), 552–557. <https://doi.org/10.1093/CERCOR/11.6.552>.
- Dhamala, E., Bassett, D.S., Yeo, B.T.T., Holmes, A.J., 2024. Functional brain networks are associated with both sex and gender in children. *Sci. Adv.* 10 (28), 4202. <https://doi.org/10.1126/SCIADV.ADN4202>.
- Dosenbach, N.U.F., Koller, J.M., Earl, E.A., Miranda-Dominguez, O., Klein, R.L., Van, A.N., Snyder, A.Z., Nagel, B.J., Nigg, J.T., Nguyen, A.L., Wesevich, V., Greene, D.J., Fair, D.A., 2017. Real-time motion analytics during brain MRI improve data quality and reduce costs. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2017.08.025>.
- Dube, S., Ivanova, M., Potter, A., 2021. "I Don't Understand": who is missed when we ask early adolescents, "Are you Transgender?" *Arch. Sex. Behav.* 50 (3), 741. <https://doi.org/10.1007/S10508-021-01986-X>.
- Dumais, K.M., Chernyak, S., Nickerson, L.D., Janes, A.C., 2018. Sex differences in default mode and dorsal attention network engagement. *PLOS ONE* 13 (6), e0199049. <https://doi.org/10.1371/JOURNAL.PONE.0199049>.
- Eliot, L., 2011. The trouble with sex differences. *Neuron* 72 (6), 895–898. <https://doi.org/10.1016/j.neuron.2011.12.001>.
- Eliot, L., Ahmed, A., Khan, H., Patel, J., 2021. Dump the "dimorphism": comprehensive synthesis of human brain studies reveals few Male-female differences beyond size. *Neurosci. Biobehav. Rev.* 125, 667–697. <https://doi.org/10.1016/J.NEUBIOREV.2021.02.026>.
- Elizabeth, P.H., Green, R., 1984. Childhood Sex-Role behaviors: similarities and differences in twins. *Acta Genet. Med. Et Gemellol. Twin Res.* 33 (2), 173–179. <https://doi.org/10.1017/S000156600007200>.
- Emory, L.E., Williams, D.H., Cole, C.M., Amparo, E.G., Meyer, W.J., 1991. Anatomic variation of the corpus callosum in persons with gender dysphoria. *Arch. Sex. Behav.* 20 (4), 409–417. <https://doi.org/10.1007/BF01542620/METRICS>.
- Engman, J., Linman, C., Van Dijk, K.R.A., Milad, M.R., 2016. Amygdala subnuclei resting-state functional connectivity sex and estrogen differences. *Psychoneuroendocrinology* 63, 34–42. <https://doi.org/10.1016/j.psyneuen.2015.09.012>.
- Erus, G., Battapady, H., Satterthwaite, T.D., Hakonarson, H., Gur, R.E., Davatzikos, C., Gur, R.C., 2015. Imaging patterns of brain development and their relationship to cognition. *Cereb. Cortex* 25 (6), 1676–1684. <https://doi.org/10.1093/CERCOR/BHT425>.
- Fair, D.A., Miranda-Dominguez, O., Snyder, A.Z., Perrone, A., Earl, E.A., Van, A.N., Koller, J.M., Feczko, E., Tisdall, M.D., Van Der Kouwe, A., Klein, R.L., Mirro, A.E., Hampton, J.M., Adeyemo, B., Laumann, T.O., Grattton, C., Greene, D.J., Schlaggar, B.L., Hagler, D.J.N., Dosenbach, N.U.F., 2020. Correction of respiratory artifacts in MRI head motion estimates. *NeuroImage* 208, 116400. <https://doi.org/10.1016/j.neuroimage.2019.116400>.
- Feczko, E., Conan, G., Marek, S., Tervo-Clemmens, B., Cordova, M., Doyle, O., Earl, E., Perrone, A., Sturgeon, D., Klein, R., Harman, G., Kilamovich, D., Hermsillo, R.,

- Miranda-Dominguez, O., Adebimpe, A., Bertolero, M., Cieslak, M., Covitz, S., Hendrickson, T., Fair, D.A., 2021. Adolescent brain cognitive development (ABCD) community MRI collection and utilities. *BioRxiv* 20, 451638. <https://doi.org/10.1101/2021.07.09.451638>.
- Fenske, S.J., Liu, J., Chen, H., Diniz, M.A., Stephens, R.L., Cornea, E., Gilmore, J.H., Gao, W., 2023. Sex differences in resting state functional connectivity across the first two years of life. *Dev. Cogn. Neurosci.* 60. <https://doi.org/10.1016/j.DCN.2023.101235>.
- Filippi, M., Valsasina, P., Misci, P., Falini, A., Comi, G., Rocca, M.A., 2013. The organization of intrinsic brain activity differs between genders: a resting-state fMRI study in a large cohort of young healthy subjects. *Hum. Brain Mapp.* 34 (6), 1330–1343. <https://doi.org/10.1002/HBM.21514>.
- Fine, C., 2014. His brain, her brain? *Science* 346 (6212), 915–916. https://doi.org/10.1126/SCIENCE.1262061/ASSET/68947B62-18EA-4B9E-AEEE-44236EDB49B2/ASSETS/GRAPHIC/346_915_F1.JPEG.
- Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33 (3), 341–355. [https://doi.org/10.1016/S0896-6273\(02\)00569-X](https://doi.org/10.1016/S0896-6273(02)00569-X).
- Fisher, R.A., 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* 10 (4), 507. <https://doi.org/10.2307/2331838>.
- Flach, P.A., Kull, M., 2015. Precision-Recall-Gain curves: PR analysis done right. *Adv. Neural Inf. Process. Syst.* 28. (<http://www.cs>).
- Flint, C., Förster, K., Koser, S.A., Konrad, C., Zwitserlood, P., Berger, K., Hermesdorf, M., Kircher, T., Nenadic, I., Krug, A., Baune, B.T., Dohm, K., Redlich, R., Opel, N., Arolt, V., Hahn, T., Jiang, X., Dannlowski, U., Grotegerd, D., 2020. Biological sex classification with structural MRI data shows increased misclassification in transgender women. *Neuropsychopharmacology* 45 (10), 1758–1765. <https://doi.org/10.1038/s41386-020-0666-3>.
- Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R.Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., Zabs, D., 2018. Recruiting the ABCD sample: design considerations and procedures. *Dev. Cogn. Neurosci.* 32, 16–22. <https://doi.org/10.1016/j.DCN.2018.04.004>.
- Giedd, J.N., Raznahan, A., Mills, K.L., Lenroot, R.K., 2012. Review: magnetic resonance imaging of Male/female differences in human adolescent brain anatomy. *Biol. Sex. Differ.* 3 (1), 1–9. <https://doi.org/10.1186/2042-6410-3-19/FIGURES/1>.
- Gilmore, J.H., Lin, W., Prastawa, M.W., Looney, C.B., Vetsa, Y.S.K., Knickmeyer, R.C., Evans, D.D., Smith, J.K., Hamer, R.M., Lieberman, J.A., Gerig, G., 2007. Regional gray matter growth, sexual dimorphism, and cerebral asymmetry in the neonatal brain. *J. Neurosci.* 27 (6), 1255–1260. <https://doi.org/10.1523/JNEUROSCI.3339-06.2007>.
- Goldstein, J.M., Seidman, L.J., Horton, N.J., Makris, N., Kennedy, D.N., Caviness, V.S., Faraone, S.V., Tsuang, M.T., 2001. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb. Cortex* 11 (6), 490–497. <https://doi.org/10.1093/CERCOR/11.6.490>.
- Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2016. Generation and evaluation of a cortical area parcellation from Resting-State correlations. *Cereb. Cortex* 26 (1), 288–303. <https://doi.org/10.1093/CERCOR/BHU239>.
- Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Grattton, C., Sun, H., Hampton, J.M., Coalson, R.S., Nguyen, A.L., McDermott, K.B., Shimony, J.S., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Nelson, S.M., Dosenbach, N.U.F., 2017. Precision functional mapping of individual human brains. *Neuron* 95 (4), 791–807.e7. <https://doi.org/10.1016/j.neuron.2017.07.011>.
- Gordon, E.M., Chauvin, R.J., Van, A.N., Rajesh, A., Nielsen, A., Newbold, D.J., Lynch, C. J., Seider, N.A., Krimmel, S.R., Scheidter, K.M., Monk, J., Miller, R.L., Metoki, A., Montez, D.F., Zheng, A., Elbau, I., Madison, T., Nishino, T., Myers, M.J., Dosenbach, N.U.F., 2023. A somato-cognitive action network alternates with effector regions in motor cortex. *Nature* 617 (7960), 351–359. <https://doi.org/10.1038/s41586-023-05964-2>.
- Gotts, S.J., Saad, Z.S., Jo, H.J., Wallace, G.L., Cox, R.W., Martin, A., 2013. The perils of global signal regression for group comparisons: a case study of autism spectrum disorders. *Front. Hum. Neurosci.* 7 (JUL), 54731. <https://doi.org/10.3389/FNHUM.2013.00356/BIBTEX>.
- Gu, Q., Zhu, L., Cai, Z., 2009. Evaluation measures of the classification performance of imbalanced data sets. *Commun. Comput. Inf. Sci.* 51, 461–471. https://doi.org/10.1007/978-3-642-04962-0_53/COVER.
- Gur, R.C., Turetsky, B.I., Matsui, M., Yan, M., Bilker, W., Hughtett, P., Gur, R.E., 1999. Sex differences in brain gray and white matter in healthy young adults: correlations with cognitive performance. *J. Neurosci.* 19 (10), 4065–4072. <https://doi.org/10.1523/JNEUROSCI.19-10-04065.1999>.
- Gur, R.E., Gur, R.C., 2016. Sex differences in brain and behavior in adolescence: findings from the Philadelphia neurodevelopmental cohort. *Neurosci. Biobehav. Rev.* 70, 159–170. <https://doi.org/10.1016/j.neubiorev.2016.07.035>.
- Hallquist, M.N., Hwang, K., & Luna, B. (2013). The nuisance of nuisance regression: Spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. <https://doi.org/10.1016/j.neuroimage.2013.05.116>.
- Hand, D., Christen, P., 2018. A note on using the F-measure for evaluating record linkage algorithms. *Stat. Comput.* 28 (3), 539–547. <https://doi.org/10.1007/S11222-017-9746-6/FIGURES/3>.
- Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., Bießmann, F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 87, 96–110. <https://doi.org/10.1016/j.neuroimage.2013.10.067>.
- Herting, M.M., Sowell, E.R., 2017. Puberty and structural brain development in humans. In: *Frontiers in Neuroendocrinology*, 44. Academic Press Inc, pp. 122–137. <https://doi.org/10.1016/j.yfrne.2016.12.003>.
- Hidalgo-Lopez, E., Mueller, K., Harris, T.A., Aichhorn, M., Sacher, J., Pletzer, B., 2020. Human menstrual cycle variation in subcortical functional brain connectivity: a multimodal analysis approach. *Brain Struct. Funct.* 225 (2), 591–605. <https://doi.org/10.1007/S00429-019-02019-Z/FIGURES/4>.
- Hill, A.C., Laird, A.R., Robinson, J.L., 2014. Gender differences in working memory networks: a BrainMap meta-analysis. *Biol. Psychol.* 102, 18–29. <https://doi.org/10.1016/j.biopsycho.2014.06.008>.
- Hjelmervik, H., Hausmann, M., Osnes, B., Westerhausen, R., Specht, K., 2014. Resting states are resting traits – an fMRI study of sex differences and menstrual cycle effects in resting state cognitive control networks. *PLOS ONE* 9 (7), e103492. <https://doi.org/10.1371/JOURNAL.PONE.0103492>.
- Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Hakonarson, H., Gur, R.E., Gur, R.C., Verma, R., 2014. Sex differences in the structural connectome of the human brain. *Proc. Natl. Acad. Sci. USA* 111 (2), 823–828. <https://doi.org/10.1073/pnas.1316909110>.
- Jäncke, L., 2018. Sex/gender differences in cognition, neurophysiology, and neuroanatomy. *F1000Research* 805. <https://doi.org/10.12688/F1000RESEARCH.13917.1>.
- Joel, D., 2020. Beyond sex differences and a Male–female continuum: mosaic brains in a multidimensional space. *Handb. Clin. Neurol.* 175, 13–24. <https://doi.org/10.1016/B978-0-444-64123-6.00002-3>.
- Johnson, L.L., Bradley, S.J., Birkenfeld-Adams, A.S., Radzins Kuksis, M.A., Maing, D.M., Mitchell, J.N., Zuck, K.J., 2004. A parent-report gender identity questionnaire for children. *Arch. Sex. Behav.* 33 (2), 105–116. <https://doi.org/10.1023/B:ASEB.0000014325.68094.F3/METRICS>.
- Kaczurkin, A.N., Raznahan, A., Satterthwaite, T.D., 2018. Sex differences in the developing brain: insights from multimodal neuroimaging. *Neuropsychopharmacology* 44 (1), 71–85. <https://doi.org/10.1038/s41386-018-0111-z>.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., Ryan, N., 1997. Schedule for affective disorders and schizophrenia for School-Age Children–Present and lifetime version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* 36 (7), 980–988. <https://doi.org/10.1097/00004583-199707000-00021>.
- Kim, J.S., Nafziger, A.N., 2000. Is it sex or is it gender? *Clin. Pharmacol. Ther.* 68 (1), 1–3. <https://doi.org/10.1067/MCP.2000.108356>.
- Kodiwera, C., Alexander, A.L., Harezlak, J., McAllister, T.W., Wu, Y.C., 2016. Age effects and sex differences in human brain white matter of young to middle-aged adults: a DTI, NODDI, and q-space study. *NeuroImage* 128, 180–192. <https://doi.org/10.1016/j.neuroimage.2015.12.033>.
- Kong, R., Li, J., Orban, C., Sabuncu, M.R., Liu, H., Schaefer, A., Sun, N., Zuo, X.N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.T., 2019. Spatial topography of Individual-Specific cortical networks predicts human cognition, personality, and emotion. *Cereb. Cortex* 29 (6), 2533–2551. <https://doi.org/10.1093/CERCOR/BHY123>.
- Kranz, G.S., Hahn, A., Kaufmann, U., Küblböck, M., Hummer, A., Ganger, S., Seiger, R., Winkler, D., Swaab, D.F., Windischberger, C., Kasper, S., Lanzenberger, R., 2014. White matter microstructure in transsexuals and controls investigated by diffusion tensor imaging. *J. Neurosci.* 34 (46), 15466–15475. <https://doi.org/10.1523/JNEUROSCI.2488-14.2014>.
- Kurth, F., Gaser, C., Luders, E., 2020. Development of sex differences in the human brain. *Cognitive Neuroscience*. <https://doi.org/10.1080/17588928.2020.1800617>.
- Kurth, F., Gaser, C., Sánchez, F.J., Luders, E., 2022. Brain sex in transgender women is shifted towards gender identity. *J. Clin. Med.* 11 (6), 1582. <https://doi.org/10.3390/JCM11061582>.
- Lenroot, R.K., Giedd, J.N., 2010. Sex differences in the adolescent brain. *Brain Cogn.* 72 (1), 46–55. <https://doi.org/10.1016/j.bandc.2009.10.008>.
- Li, H., Satterthwaite, T.D., Fan, Y., 2017. Large-scale sparse functional networks from resting state fMRI. *NeuroImage* 156, 1–13. <https://doi.org/10.1016/j.neuroimage.2017.05.004>.
- Lotze, M., Domin, M., Gerlach, F.H., Gaser, C., Lueders, E., Schmidt, C.O., Neumann, N., 2019. Novel findings from 2,838 adult brains on sex differences in gray matter brain volume. *Sci. Rep.* 9 (1), 1–7. <https://doi.org/10.1038/s41598-018-38239-2>.
- Luders, E., Kurth, F., 2020. Structural differences between Male and female brains. *Handb. Clin. Neurol.* 175, 3–11. <https://doi.org/10.1016/B978-0-444-64123-6.00001-1>.
- Luders, E., Toga, A.W., 2010. Sex differences in brain anatomy. In: *Progress in Brain Research*, 186. Elsevier B.V, pp. 2–12. <https://doi.org/10.1016/B978-0-444-53630-3.00001-4>.
- Luders, E., Narr, K.L., Thompson, P.M., Woods, R.P., Rex, D.E., Jancke, L., Steinmetz, H., Toga, A.W., 2005. Mapping cortical gray matter in the young adult brain: effects of gender. *NeuroImage* 26, 493–501. <https://doi.org/10.1016/j.neuroimage.2005.02.010>.
- Luders, E., Narr, K.L., Thompson, P.M., Rex, D.E., Woods, R.P., DeLuca, H., Jancke, L., Toga, A.W., 2006. Gender effects on cortical thickness and the influence of scaling. *Hum. Brain Mapp.* 27 (4), 314–324. <https://doi.org/10.1002/HBM.20187>.
- Luders, E., Sánchez, F.J., Gaser, C., Toga, A.W., Narr, K.L., Hamilton, L.S., Vilain, E., 2009. Regional gray matter variation in Male-to-female transsexualism. *NeuroImage* 46, 904–907. <https://doi.org/10.1016/j.neuroimage.2009.03.048>.
- Maccallum, R.C., Zhang, S., Preacher, K.J., Rucker, D.D., 2002. On the practice of dichotomization of quantitative variables. *Psychol. Methods* 7 (1), 19–40. <https://doi.org/10.1037/1082-989X.7.1.19>.

- Mathew, J., Masson, G.S., Danion, F.R., 2020. Sex differences in visuomotor tracking. *Sci. Rep.* 10 (1), 1–12. <https://doi.org/10.1038/s41598-020-68069-0>.
- McNemar, Q., 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12 (2), 153–157.
- Mourão-Miranda, J., Bokde, A.L.W., Born, C., Hampel, H., & Stetter, M. (2005). Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. <https://doi.org/10.1016/j.neuroimage.2005.06.070>.
- Mueller, S.C., Landré, L., Wierckx, K., T'Sjoen, G., 2017. A structural magnetic resonance imaging study in transgender persons on Cross-Sex hormone therapy. *Neuroendocrinology* 105 (2), 123–130. <https://doi.org/10.1159/000448787>.
- Mueller, S.C., De Cuypere, G., T'Sjoen, G., 2017. Transgender research in the 21st century: a selective critical review from a neurocognitive perspective. *Am. J. Psychiatry* 174 (12), 1155–1162. <https://doi.org/10.1176/appi.ajp.2017.17060626>.
- Mueller, S.C., Guillamon, A., Zubiarré-Elorza, L., Junque, C., Gomez-Gil, E., Uribe, C., Khorshad, B.S., Khazai, B., Talaei, A., Habel, U., Votinov, M., Derril, B., Lanzenberger, R., Seiger, R., Kranz, G.S., Kreukels, B.P.C., Kettner, P.T.C., Burke, S. M., Lambalk, N.B., Luders, E., 2021. The neuroanatomy of transgender identity: Mega-Analytic findings from the ENIGMA transgender persons working group. *J. Sex. Med.* 18 (6), 1122–1129. <https://doi.org/10.1016/j.jsxm.2021.03.079>.
- Murphy, K., Birn, R.M., Bandettini, P.A., 2013. Resting-state fMRI acquisitions and cleanup. *NeuroImage* 80, 349–359. <https://doi.org/10.1016/j.neuroimage.2013.04.001>.
- Murray, S.O., Schallmo, M.P., Kolodny, T., Millin, R., Kale, A., Thomas, P., Ramsdayer, T.H., Troche, S.J., Bernier, R.A., Tadin, D., 2018. Sex differences in visual motion processing. *Curr. Biol.* 28 (17), 2794–2799.e3. <https://doi.org/10.1016/j.cub.2018.06.014>.
- National Academies of Sciences, E., Division of Behavioral, M., Social Sciences, Education, Health and Medicine Division, Committee on Population, Board on Health Sciences Policy, Committee on the Use of Race, E., A. as P. D. in G. R., 2023. Using population descriptors in genetics and genomics research: a new framework for an evolving field. National Academies Press (US), pp. 1–217. <https://doi.org/10.17226/26902>.
- Nopoulos, P., Flaum, M., O'Leary, D., Andreasen, N.C., 2000. Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Res. Neuroimaging* 98 (1), 1–13. [https://doi.org/10.1016/S0925-4927\(99\)00044-X](https://doi.org/10.1016/S0925-4927(99)00044-X).
- Paus, T., Keshavan, M., Giedd, J.N., 2008. Why do many psychiatric disorders emerge during adolescence? *2008 9:12 Nat. Rev. Neurosci.* 9 (12), 947–957. <https://doi.org/10.1038/nrn2513>.
- Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Courapeau, D., Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., Dubourg, V., Passos, A., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830. (<http://scikit-learn.sourceforge.net>).
- Pembury Smith, M.Q.R., Ruxton, G.D., 2020. Effective use of the McNemar test. *Behav. Ecol. Sociobiol.* 74 (11), 1–9. <https://doi.org/10.1007/s00265-020-02916-Y/TABLES/5>.
- Petersen, N., Kilpatrick, L.A., Goharad, A., Cahill, L., 2014. Oral contraceptive pill use and menstrual cycle phase are associated with altered resting state functional connectivity. *NeuroImage* 90, 24–32. <https://doi.org/10.1016/j.neuroimage.2013.12.016>.
- Poldrack, R.A., Huckins, G., Varoquaux, G., 2020. Establishment of best practices for evidence for prediction a review. *JAMA Psychiatry* 77 (5), 534. <https://doi.org/10.1001/JAMAPSYCHIATRY.2019.3671>.
- Potter, A., Dube, S., Allgaier, N., Loso, H., Ivanova, M., Barrios, L.C., Bookheimer, S., Chaarani, B., Dumas, J., Feldstein-Ewing, S., Freedman, E.G., Garavan, H., Hoffman, E., McGlade, E., Robin, L., Johns, M.M., 2021. Early adolescent gender diversity and mental health in the adolescent brain cognitive development study. *J. Child Psychol. Psychiatry* 62 (2), 171–179. <https://doi.org/10.1111/JCPP.13248>.
- Potter, A.S., Dube, S.L., Barrios, L.C., Bookheimer, S., Espinoza, A., Ewing, S.W.F., Freedman, E.G., Hoffman, E.A., Ivanova, M., Jefferys, H., McGlade, E.C., Tapert, S.F., Johns, M.M., 2022. Measurement of gender and sexuality in the adolescent brain cognitive development (ABCD) study. *Dev. Cogn. Neurosci.* 53, 101057. <https://doi.org/10.1016/j.dcn.2022.101057>.
- Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* 59, 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>.
- Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* 84, 320–341. <https://doi.org/10.1016/j.neuroimage.2013.08.048>.
- Price, M.A., Hollinsaid, N.L., Bokhour, E.J., Johnston, C., Skov, H.E., Kaufman, G.W., Sheridan, M.K., Oleszki, C., 2023. Transgender and gender diverse Youth's experiences of Gender-Related adversity. *Child Adolesc. Soc. Work J.* 40 (3), 361–380. <https://doi.org/10.1007/s10560-021-00785-6/TABLES/4>.
- Rametti, G., Carrillo, B., Gómez-Gil, E., Junque, C., Segovia, S., Gomez, Á., Guillamon, A., 2011. White matter microstructure in female to Male transsexuals before cross-sex hormonal treatment. A diffusion tensor imaging study. *J. Psychiatr. Res.* 45, 199–204. <https://doi.org/10.1016/j.jpsychires.2010.05.006>.
- Rametti, G., Carrillo, B., Gómez-Gil, E., Junque, C., Zubiarré-Elorza, L., Segovia, S., Gomez, Á., Guillamon, A., 2011. The microstructure of White matter in Male to female transsexuals before cross-sex hormonal treatment. A DTI study. *J. Psychiatr. Res.* 45 (7), 949–954. <https://doi.org/10.1016/j.jpsychires.2010.11.007>.
- Ricard, J.A., Parker, T.C., Dhamala, E., Kwasa, J., Allsop, A., Holmes, A.J., 2022. Confronting racially exclusionary practices in the acquisition and analyses of neuroimaging data. *Nat. Neurosci.* 26 (1), 4–11. <https://doi.org/10.1038/s41593-022-01218-y>.
- Ritchie, S.J., Cox, S.R., Shen, X., Lombardo, M.V., Reus, L.M., Alloza, C., Harris, M.A., Alderson, H.L., Hunter, S., Neilson, E., Liewald, D.C.M., Auyeung, B., Whalley, H.C., Lawrie, S.M., Gale, C.R., Bastin, M.E., McIntosh, A.M., Deary, I.J., 2018. Sex differences in the adult human brain: evidence from 5216 UK biobank participants. *Cereb. Cortex* 28 (8), 2959–2975. <https://doi.org/10.1093/CERCOR/BHY109>.
- Ruigrok, A.N.V., Salimi-Khorshidi, G., Lai, M.C., Baron-Cohen, S., Lombardo, M.V., Tait, R.J., Suckling, J., 2014. A meta-analysis of sex differences in human brain structure. *Neurosci. Biobehav. Rev.* 39, 34–50. <https://doi.org/10.1016/j.neubiorev.2013.12.004>.
- Rutter, M., Caspi, A., Moffitt, T.E., 2003. Using sex differences in psychopathology to study causal mechanisms: unifying issues and research strategies. *J. Child Psychol. Psychiatry* 44 (8), 1092–1115. <https://doi.org/10.1111/1469-7610.00194>.
- Saad, Z.S., Gotts, S.J., Murphy, K., Chen, G., Jo, H.J., Martin, A., Cox, R.W., 2012. Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain Connect.* 2 (1), 25–32. <https://doi.org/10.1089/BRAIN.2012.0080/ASSET/IMAGES/LARGE/FIGURE4.JPEG>.
- Sacher, J., Neumann, J., Okon-Singer, H., Gotowiec, S., Villringer, A., 2013. Sexual dimorphism in the human brain: evidence from neuroimaging. *Magn. Reson. Imaging* 31, 366–375. <https://doi.org/10.1016/j.mri.2012.06.007>.
- Satterthwaite, T.D., Shinohara, R.T., Wolf, D.H., Hopson, R.D., Elliott, M.A., Vandekar, S. N., Ruparel, K., Calkins, M.E., Roalf, D.R., Gennatas, E.D., Jackson, C., Erus, G., Prabhakaran, K., Davatzikos, C., Detre, J.A., Hakonarson, H., Gur, R.C., Gur, R.E., 2014. Impact of puberty on the evolution of cerebral perfusion during adolescence. *Proc. Natl. Acad. Sci. USA* 111 (23), 8643–8648. <https://doi.org/10.1073/PNAS.1400178111/-/DCSUPPLEMENTAL/PNAS.2014001781.PDF>.
- Schaefer, A., Kong, R., Gordon, E.N., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.T., 2018. Local-Global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* 28, 3095–3114. <https://doi.org/10.1093/cercor/bhx179>.
- Schraiber, J.G., Edge, M.D., 2024. Heritability within groups is uninformative about differences among groups: cases from behavioral, evolutionary, and statistical genetics. *Proc. Natl. Acad. Sci. USA* 121 (12), e2319496121. https://doi.org/10.1073/PNAS.2319496121/SUPPL_FILE/PNAS.2319496121.SAPP.PDF.
- Seitzman, B.A., Gratton, C., Marek, S., Raut, R.V., Dosenbach, N.U.F., Schlaggar, B.L., Petersen, S.E., Greene, D.J., 2020. A set of functionally-defined brain regions with improved representation of the subcortex and cerebellum. *NeuroImage* 206, 116290. <https://doi.org/10.1016/j.neuroimage.2019.116290>.
- Sepehrband, F., Lynch, K.M., Cabeen, R.P., Gonzalez-Zacarias, C., Zhao, L., Arcy, M.D., Kesselman, C., Herting, M.M., Dinov, I.D., Toga, A.W., Clark, K.A., 2018. Neuroanatomical morphometric characterization of sex differences in youth using statistical learning. *NeuroImage* 172, 217–227. <https://doi.org/10.1016/j.neuroimage.2018.01.065>.
- Shanmugan, S., Seidlitz, J., Cui, Z., Adebimpe, A., Bassett, D.S., Bertolero, M.A., Davatzikos, C., Fair, D.A., Gur, R.E., Gur, R.C., Larsen, B., Li, H., Pines, A., Raznahan, A., Roalf, D.R., Shinohara, R.T., Vogel, J., Wolf, D.H., Fan, Y., Satterthwaite, T.D., 2022. Sex differences in the functional topography of association networks in youth. *Proc. Natl. Acad. Sci. USA* 119 (33), e2110416119. https://doi.org/10.1073/PNAS.2110416119/SUPPL_FILE/PNAS.2110416119.SAPP.PDF.
- Shaqiri, A., Roishnivil, M., Grzeczowski, L., Chkonia, E., Pilz, K., Mohr, C., Brand, A., Kunchulia, M., Herzog, M.H., 2018. Sex-related differences in vision are heterogeneous. *Sci. Rep.* 8 (1). <https://doi.org/10.1038/s41598-018-25298-8>.
- Siegel, J.S., Mitra, A., Laumann, T.O., Seitzman, B.A., Raichle, M., Corbetta, M., Snyder, A.Z., 2017. Data quality influences observed links between functional connectivity and. *Behav. Cereb. Cortex* 27 (9), 4492–4502. <https://doi.org/10.1093/CERCOR/BHW253>.
- Sokolova, M., Japkowicz, N., Szpakowicz, S., 2006. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. AAAI Workshop Tech. Rep. 24–29. https://doi.org/10.1007/11941439_114/COVER.
- Stoica, T., Knight, L.K., Naaz, F., Patton, S.C., Deput, B.E., 2021. Gender differences in functional connectivity during emotion regulation. *Neuropsychologia* 156, 107829. <https://doi.org/10.1016/j.neuropsychologia.2021.107829>.
- Stumme, J., Jockwitz, C., Hoffstaedter, F., Amunts, K., Caspers, S., 2020. Functional network reorganization in older adults: Graph-theoretical analyses of age, cognition and sex. *NeuroImage* 214, 116756. <https://doi.org/10.1016/j.neuroimage.2020.116756>.
- Tian, L., Wang, J., Yan, C., He, Y., 2011. Hemisphere-and gender-related differences in small-world brain networks: a resting-state functional MRI study. *NeuroImage* 54, 191–202. <https://doi.org/10.1016/j.neuroimage.2010.07.066>.
- Tian, Y., Zalesky, A., 2021. Machine learning prediction of cognition from functional connectivity: are feature weights reliable? *NeuroImage* 245, 118648. <https://doi.org/10.1016/j.neuroimage.2021.118648>.
- Tomasi, D., Volkow, N.D., 2012. Gender differences in brain functional connectivity density. *Hum. Brain Mapp.* 33 (4), 849–860. <https://doi.org/10.1002/HBM.21252>.
- Tunc, B., Solmaz, B., Parker, D., Satterthwaite, T.D., Elliott, M.A., Calkins, M.E., Ruparel, K., Gur, R.E., Gur, R.C., Verma, R., 2016. Establishing a link between sex-related differences in the structural connectome and behaviour. *Philos. Trans. R. Soc. B Biol. Sci.* 371 (1688). <https://doi.org/10.1098/RSTB.2015.0111>.
- U.S. Census Bureau, 2023. American community survey: 2023 data release. U.S. Department of Commerce. (<https://www.census.gov/Programs-Surveys/Acs/>).
- U.S. Department of Health and Human Services, 2017. Annual update of the HHS poverty guidelines. Federal Register, pp. 8831–8833. (<https://aspe.hhs.gov/sites/default/files/documents/c8b85d8360d5a70beef10da5fbb690d/HHS-Poverty-Guidelines-Fed-Register-2017.pdf>).

- Vanston, J.E., Strother, L., 2017. Sex differences in the human visual system. *J. Neurosci. Res.* 95 (1–2), 617–625. <https://doi.org/10.1002/JNR.23895>.
- Vijayakumar, N., Op de Macks, Z., Shirtcliff, E.A., Pfeifer, J.H., 2018. Puberty and the human brain: insights into adolescent development. In: *Neuroscience and Biobehavioral Reviews*, 92. Elsevier Ltd, pp. 417–436. <https://doi.org/10.1016/j.neubiorev.2018.06.004>.
- Weis, S., Hodgetts, S., Hausmann, M., 2019. Sex differences and menstrual cycle effects in cognitive and sensory resting state networks. *Brain Cogn.* 131, 66–73. <https://doi.org/10.1016/j.bandc.2017.09.003>.
- Weissman-Fogel, I., Moayed, M., Taylor, K.S., Pope, G., Davis, K.D., 2010. Cognitive and default-mode resting state networks: do Male and female brains “rest” differently? *Hum. Brain Mapp.* 31 (11), 1713–1726. <https://doi.org/10.1002/HBM.20968>.
- Xi, W., Beer, M.A., 2018. Local epigenomic state cannot discriminate interacting and non-interacting enhancer–promoter pairs with high accuracy. *PLOS Comput. Biol.* 14 (12), e1006625. <https://doi.org/10.1371/JOURNAL.PCBI.1006625>.
- Yeo, T.B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fisch, B., Liu, H., Buckner, R. L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106 (3), 1125. <https://doi.org/10.1152/JN.00338.2011>.
- Yokota, Y., Kawamura, Y., Kameya, Y., 2005. Callosal shapes at the midsagittal plane: MRI differences of normal males, normal females, and GID. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 7, pp. 3055–3058. <https://doi.org/10.1109/IEMBS.2005.1617119>.
- Zhang, C., Dougherty, C.C., Baum, S.A., White, T., Michael, A.M., 2018. Functional connectivity predicts gender: evidence for gender differences in resting brain connectivity. *Hum. Brain Mapp.* 39 (4), 1765–1776. <https://doi.org/10.1002/HBM.23950>.
- Zhou, J.N., Hofman, M.A., Gooren, L.J.G., Swaab, D.F., 1995. A sex difference in the human brain and its relation to transsexuality. *Nature* 378 (6552), 68–70. <https://doi.org/10.1038/378068a0>.
- Zhu, D.C., Tarumi, T., Khan, M.A., Zhang, R., 2015. Vascular coupling in resting-state fMRI: evidence from multiple modalities. *J. Cereb. Blood Flow. Metab.* 35. <https://doi.org/10.1038/jcbfm.2015.166>.
- Zuo, X.N., Kelly, C., Di Martino, A., Mennes, M., Margulies, D.S., Bangaru, S., Grzadzinski, R., Evans, A.C., Zang, Y.F., Castellanos, F.X., Milham, M.P., 2010. Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy. *J. Neurosci.* 30 (45), 15034–15043. <https://doi.org/10.1523/JNEUROSCI.2612-10.2010>.