## Online supplement 2 Preprocessing

**Preprocessing of functional MRI data**

Data preprocessing was performed using DPABI V5.1 ([Dpabi.org](http://rfmri.org/dpabi)) and SPM12 ([SPM12.ucl](http://www.fil.ion.ucl.ac.uk/spm/software/spm12/)). The preprocessing steps included 1) discarding the first 10 time points; 2) slice-timing correction; 3) head-motion correction; 4) spatial normalization to the unified segmented anatomical image using forward deformation fields and resampling to 3×3×3 mm3; 5) spatial smoothing with a 3D isotropic Gaussian kernel with a full width at half maximum of 6 mm; 6) linear trend removal; 7) nuisance covariate regression, including head-motion covariates using the Friston 24-parameter model as well as signals from white matter and cerebrospinal fluid; and 8) bandpass filtering of the data with a frequency range of 0.01-0.08 Hz.

**Constructing FC matrices**

All FC measurements were calculated using the Gretna toolbox (J. Wang et al., 2015). Based on previous studies on network construction methods (J. Zhang et al., 2011; Z. Zhang, Telesford, Giusti, Lim, & Bassett, 2016), we applied a functional template called "Dosenbach160" (Dosenbach et al., 2010), which provides 160 functionally isolated regions of interest (ROIs) covering most of the cortical and cerebellar regions of the brain. We used the CONN toolbox ([CONN.org](https://www.nitrc.org/projects/conn/)) to compute the whole-brain networks for each participant. Specifically, BOLD signals were extracted from each ROI, and pairwise correlations between each pair of ROIs were computed to obtain the 160×160 FC matrix for each participant.

**Multisite effect correction**

Site effects on the modular coefficient were removed using the ComBat function available in MATLAB ([ComBatHarmonization.com](https://github.com/Jfortin1/ComBatHarmonization)) to account for site, collection time, and data acquisition parameter variability across each of the data collection sites in the ABIDE data. This approach has been shown to effectively account for scanner-related variance in multisite resting-state fMRI datasets (De Rosa et al., 2023; Maltbie, Yousefi, Zhang, Kashyap, & Keilholz, 2022; Yamashita et al., 2019). With the ComBat function, the diagnosis was treated as the biological variable of interest, and a parametric prior method was used in the empirical Bayes procedure.

As for why this procedure is necessary, some studies have demonstrated that employing a hierarchical Bayesian regression model during normative modeling allows for the consideration of variability and correlations across different levels (e.g., sites, individuals, observations) (Bayer et al., 2021). In this way, site effect could have been corrected and incorporated within the normative model itself. However, the Bayesian model utilized in our study, which originates from Bethlehem et al., only addresses variability at the individual level, thus requires the ComBat pipeline. We chose this model due to its advantages in terms of computational efficiency, automatic parameter learning, and the mitigation of issues related to local optima and overfitting.