Development of Thalamocortical Structural Connectivity in Typically Developing and Psychosis Spectrum Youth

Supplement

SUPPLEMENTARY METHODS

Participants

Participants were part of the PNC (Study Accession phs000607.v3.p2), a publicly available dataset consisting of approximately 9500 individuals (age range: 8-22 years) (1). All subjects provided medical history, clinical, and cognitive data. A subset of participants completed neuroimaging data collection. Of the 1601 participants with neuroimaging data, the diffusionweighted imaging series was collected on 1396. From this sample, 252 were excluded for a serious medical condition ($n = 53$), insufficient clinical data to reach a diagnosis ($n = 63$), or low data quality (n = 136). This resulted in a final sample of 1144 participants. We compared demographic, clinical, and cognitive characteristics between individuals included in the study and those excluded for data that did not meet our quality standards (see Data quality, below). Compared to individuals included in the study, those excluded for data quality were younger (13.1 vs. 15.3, $p < .001$), had fewer years of education (5.9 vs. 8.0, $p < .001$), and had lower global cognition (CNB) scores ($-.23$ vs $.10$, p $< .001$). Those excluded were 36 typicallydeveloping youth, 33 psychosis-spectrum youth, and 67 youth with other psychopathology. Individual were not more likely to be excluded based on clinical diagnosis (p=.39). Complete demographics are presented in Supplementary Table S1.

Clinical assessment. Psychopathology self-report and other-report (i.e. parent) ratings and medical histories were completed using GOASSESS, a computerized structured clinical interview.

Neuroimaging Data Acquisition, Preprocessing, and Probabilistic Tractography

High resolution T1-weighted structural scans were acquired using the MPRAGE sequence on a Siemens Tim Trio 3T scanner with a 32-channel head coil with 0.93 x 0.93 x 1 mm voxels (160 slices, TR/TE = 1810/3.5, FOV = 180 x 240 x 160, matrix = 192 x 256 x 160; flip angle = 9°). Diffusion-weighted images were obtained using a twice-refocused spin-echo (TRSE) single-shot EPI sequence (TR = 8100 ms, TE = 82 ms, FOV = 240 x 240 mm, 70 slices, 1.875 x 1.875 x 2 mm resolution, gap = 0, total volumes = 71, GRAPPA factor = 3, bandwidth = 2170 Hz/pixel, PE direction = AP). The complete 64-direction diffusion-weighted imaging set was acquired as two independent 32-direction sequences (b value = 1000 s/mm²), with a total of 7 interspersed nondiffusion-weighted images (b value = 0 s/mm^2). The two sequences were concatenated prior to preprocessing and the b=0 images were averaged.

Image processing was performed on the Vanderbilt University Institute of Imaging Science Center for Computational Imaging XNAT platform (2) and in MATLAB (version 2018a). Processing pipelines were containerized using Singularity (3) and built at SingularityHub (https://singularity-hub.org), and are available through github

(https://github.com/baxpr/freesurfer-singularity/releases/tag/v1.0.0;

https://github.com/baxpr/dwipre-PNC/releases/tag/v1.0.0; https://github.com/baxpr/bedpostsingularity/releases/tag/v1.0.0; https://github.com/baxpr/thaltrack-

whole/releases/tag/v3.0.3). Each subject's structural T1-weighted image was segmented using FreeSurfer developmental version 6 (4). Cortical ROIs for each participant were defined based on the Desikan-Killiany (DKT) atlas (5) (see Supplementary Table S2 for list of cortical regions). Six bilateral cortical regions-of-interest (ROIs) and the thalamus were used as targets and the seed, respectively, for probabilistic tractography analysis of diffusion-weighted scans. The six cortical ROIs included the: prefrontal cortex (PFC), motor cortex/supplementary motor area, somatosensory cortex, posterior parietal cortex, temporal cortex, and occipital cortex (see Supplementary Figure S1 for an example segmentation). In addition, each subject's T1-weighted anatomical image was segmented into grey matter, white matter, and CSF tissue classes and

DARTEL normalized (6) to Montreal Neurological Institute (MNI) space using the Computational Anatomy Toolbox 12 (CAT12, version 12.5) for Statistical Parametric Mapping 12 (SPM12).

FMRIB's Diffusion Toolbox (FDT) for FSL v5.0.6 software package

(http://www.fmrib.ox.ac.uk/fsl/) was used to preprocess the diffusion data and perform probabilistic tractography using the same general approach described by Behrens et al., (7). First, eddy current distortions were corrected using the Eddy tool (8). Fractional anisotropy (FA) and mean diffusivity (MD) images were calculated by fitting a diffusion tensor model at each voxel using DTIFIT. The brain was extracted using the brain extraction tool (BET) (9) and the diffusion parameters were estimated using bedpostx. This method allows for modeling of crossing fibres within each voxel of the brain.

Following preprocessing, the probabilistic fibre tracking tool (probtrackx2) was used to quantify anatomical connectivity between the thalamus and each of the six ipsilateral cortical targets. From each thalamic voxel, 5000 samples were sent through the probability distributions on principle fibre direction. With a curvature threshold of 0.2, 2000 maximum number of steps, step length 0.5 mm, and subsidiary fibre volume threshold of 0.01. Modified Euler streaming and distance correction was used to account for the fact that connectivity distribution drops with distance from the seed mask. The left and right hemispheres were analyzed separately and the contralateral hemisphere was included as an exclusion mask to eliminate streamlines reaching the contralateral hemisphere. For each region, the cortical target was included as a stop mask to ensure that streamlines stopped when reaching the target, and the remaining cortical ROIs from the ipsilateral hemisphere were included as exclusion masks to ensure that only fibres reaching the cortical target directly were included.

From probtrackx2, seed-to-target voxel-wise images were generated, one for each cortical target, in which the value of each voxel within the seed mask (i.e. thalamus) represents the number of samples seeded from that voxel reaching the relevant target mask. The connectivity of each cortical region with the thalamus was calculated by dividing the number of samples

3

reaching that region, summed across all voxels in the thalamus, by the total number of samples within the thalamus reaching all cortical regions. This is a measure of total tractography-defined connectivity from the thalamus to a specific cortical area after controlling for overall connectivity of the thalamus. These values, expressed as "percent total connectivity", served as the dependent variables for the primary analysis. Dividing by total thalamus seeds effectively controls for the potential influence of volume differences across groups.

Each participants' probabilistic tractography map, created as a result of the thalamus-to-cortical targets analysis, were carefully checked for coverage of thalamocortical white matter pathways, normalized using the waytotal value, thresholded for error (1.5%), warped to MNI space using the normalization parameters derived from CAT12, and averaged into group maps representing the spatial overlap of each thalamocortical tract. Group maps were carefully visually inspected and a threshold of >75% participant overlap was selected to capture individual spatial variation in tracts while minimizing the inclusion of voxels with low probability of lying in the tract of interest. Thresholded group maps were transformed to each participant's diffusion space and mean diffusion values (FA, MD) were extracted. Group masks of each thalamocortical tract are presented in Supplementary Figure S2.

Data quality

Prior to preprocessing, all diffusion and T1 scans were carefully visually inspected, blind to demographic and clinical data, for acquisition issues (i.e., partial brain coverage or acquisition artifacts), and 2 participants were excluded. Next, the motion metric of mean relative displacement (rDisp) and the contrast-to-noise ratio (CNR) were calculated. Participants were excluded for excessive motion, defined as a mean rDisp value \geq the 95th percentile (rDisp \geq .73), or for a CNR value \leq the 5th percentile (CNR \leq .46). 80 participants were excluded for excessive motion/low CNR. Motion (rDisp) was included in all analyses as a covariate of no interest. Motion and CNR metrics for included participants are displayed in Supplementary Table S3. Finally, preprocessed diffusion and T1 scans were visually inspected for processing failures (i.e., skull-stripping, coregistration, FreeSurfer, or processing failure), and 54 participants were excluded.

4

SUPPLEMENTARY RESULTS

Thalamocortical Percent Total Connectivity

Percent total connectivity distributions for the thalamus to each cortical ROI are presented in Supplementary Figure S3A. Streamline counts distributions for the thalamus to each cortical ROI are presented in Supplementary Figure S3B.

Age effects. Age effects of percent total connectivity for each cortical ROI with the thalamus, broken down by hemisphere, are presented in Supplementary Figure S4. Full statistical results are presented in Supplementary Table S4. Linear effects of age were more pronounced in the right hemisphere and largely consistent across hemispheres, with the exception of increased thalamic-motor cortex connectivity, which showed no relationship with age in the left hemisphere.

We performed planned secondary analyses on streamline counts to determine if age effects were consistent across analyses. Age effects of streamline counts for each cortical ROI with the thalamus are presented in Supplementary Figure S5. Full statistical results for streamline counts are presented in Supplementary Table S5. Regression models were repeated as in the primary analysis, with streamline count as the dependent variable, and with an additional covariate of total streamline count for any connection between the thalamus and cortex. Linear effects of age were consistent across analyses. Streamline counts increased linearly with age for the motor and somatosensory cortex, and decreased linearly with age for the temporal and occipital cortex. No other connections showed a linear effect of age for thalamocortical streamline counts. There were also no quadratic effects of age for streamline counts.

Sex effects. Percent total connectivity of each cortical ROI with the thalamus, broken down by hemisphere and sex, are presented in Supplementary Table S6. Females had greater thalamicoccipital cortex percent total connectivity compared to males. There was a trend for greater thalamic-somatosensory cortex percent total connectivity in males than females (*p* = .01) that

did not reach significance at a Bonferroni-corrected *p* = .008. Sex effects were more pronounced in the right hemisphere. Secondary analyses of streamline counts revealed stronger effects compared to the percent total connectivity analysis, with greater thalamicoccipital cortex streamline counts in females and greater thalamic-somatosensory streamline counts in males, and an additional effect of greater thalamic-motor cortex streamline counts in males. Secondary analyses of streamline counts by sex are presented in Supplementary Table S7.

Group differences. Primary analyses for effects of group, presented in Supplementary Table S8, showed no between-group differences in thalamic percent total connectivity for any cortical ROI. Results were consistent across hemispheres. Secondary analyses of streamline counts, presented in Supplementary Table S9, also showed no between-group differences. Mean volumes of the thalamus and cortical ROIs used as seed/target masks for probabilistic tractography are presented in Supplementary Table S10. Cortical and thalamic ROI volumes were larger in the typically developing group compared to the psychosis spectrum group. Volume differences were controlled in the primary analyses by dividing total connectivity by all thalamic connections to any ROI.

Interaction effects. There was a linear interaction between age and sex, with older females showing more thalamic-occipital percent total connectivity than males. The interaction between linear age and sex for total thalamic-occipital connectivity is presented in Supplementary Figure S6.

Cognitive correlates. There were no correlations between cognitive function, as measured by CNB composite scores, and total thalamic connectivity for any cortical ROI. Cognitive correlations are presented in Supplementary Table S11.

6

Thalamocortical White Matter Microstructure

White matter microstructure (FA) distributions for the thalamus to each cortical ROI are presented in Supplementary Figure S3C.

Age effects. Age effects of FA and MD for each bilateral cortical ROI are presented in Figure 3 and Supplementary Figure S7, respectively. Age effects of FA for each cortical ROI with the thalamus, broken down by hemisphere, are presented in Supplementary Figure S8. Complete statistical results for FA and MD analyses are presented in Supplementary Table S12 and S13, respectively. White matter microstructure, measured as mean FA within thalamocortical tracts, increased linearly with age for all thalamocortical tracts. For each tract, older participants had higher FA values than younger participants. White matter microstructure in tracts linking the thalamus with the motor, somatosensory, posterior parietal, and occipital cortex also showed quadratic associations with age. Higher FA values were observed in late adolescence, while early childhood/early adulthood was associated with lower FA values. Quadratic effects are presented in Supplementary Figure S9. Linear and quadratic effects of age were consistent across hemispheres and when excluding gray matter voxels from tract masks. Follow-up analysis showed that MD values decreased linearly with age across regions. For each region, older participants had lower MD values than younger participants. Follow-up analyses also showed quadratic associations between MD and age across regions (trend effect in thalamicsomatosensory tract). For each pathway, lower MD values were observed in late adolescence, while early childhood/early adulthood was associated with higher MD values. Quadratic effects are presented in Supplementary Figure S10.

Sex effects. White matter microstructure of each cortical ROI with the thalamus, broken down by hemisphere and sex, are presented in Supplementary Tables S14 and S15. Males had higher FA values compared to females in tracts linking the thalamus with the prefrontal, motor, somatosensory, posterior parietal, and occipital cortex. Higher FA values in males were consistent across hemispheres although more pronounced in the left hemisphere. Follow-up analyses showed that males also had lower mean MD in most white matter tracts, with the

exception of the thalamic-occipital tract. Sex effects were consistent when excluding gray matter voxels from tract masks.

Group differences. Thalamocortical white matter microstructure analyses for effects of group, presented in Supplementary Tables S16 and S17, showed that typically-developing youth had higher FA values compared to psychosis-spectrum youth in all tracts linking the thalamus and the cortex. Results were consistent across hemispheres and when excluding gray matter voxels from tract masks. Follow-up analysis showed that typically-developing youth also tended to have lower MD values compared to psychosis-spectrum youth across tracts, with significantly lower values in the thalamic-somatosensory and thalamic-posterior parietal tract.

Sensitivity analysis for whole-brain FA. We performed a sensitivity analysis to determine whether age, sex, and group FA effects in thalamocortical tracts were associated with overall whole brain differences in FA. Whole brain FA values were calculated for each participant and entered into each analysis as a covariate of no interest. Statistical results are presented in Supplementary Table S18. Age effects were consistent with the primary analysis across all brain regions except the temporal cortex, which did not show a significant effect of age when controlling for whole brain FA (*p* = .03). Sex and group effects were consistent with those found in the primary analysis.

Sensitivity analyses for race, parental education, and CNR. Because race, parental education, and CNR differed by diagnostic group, we performed sensitivity analyses to determine whether between-group differences in thalamocortical FA were associated with between-group differences in these demographic and quality characteristics. For each analysis, a matched sample of participants was randomly selected in SAS (PROC SURVEYSELECT) from each diagnostic group (race-matched sample = 96 per diagnostic group, parental education-matched sample = 95 per diagnostic group, CNR-matched sample = 96 per diagnostic group). Statistical results are presented in Supplementary Table S19. Effect sizes for the sensitivity analyses were largely consistent with effect sizes found in the primary analysis of the full sample (ES range:

.011 - .013), with similar or larger between-group effects in all sensitivity analyses (ES range: .010 - .037) with the exception of the somatosensory cortex in the parental education-matched analysis, which was lower (ES = .006) than in the primary analysis (ES = .013).

Interaction effects. There was an interaction between linear age and sex in tracts connecting the thalamus with the prefrontal, motor, somatosensory, and posterior parietal cortex. Older males had higher FA values in thalamic tracts than older females. Interactions are presented in Supplementary Figure S11. Full statistical results are presented in Supplementary Table S20. There were no significant interactions between age and group, or sex and group (all $ps \geq .10$). For reference, FA values are plotted for age by group in Supplementary Figure S12.

Global cognition, z-score $0.10 \pm .5$ $-0.23 \pm .6$ 46.26 $1,1279$ $< .001$ ^{*} WRAT, standard score 100.6 ± 15.1 102.7 ± 16.4 1.98 $1,1277$.16

Supplementary Table S1. Sample characteristics for individuals included vs. excluded from

Note: Asterisk denotes significant *p*-value. Standard deviation (SD); White (W); African-American (AA); Other (O).

Supplementary Table S2. Cortical structures included in each cortical region of interest used as a target in probabilistic tractography

Note: R=Right; L=Left

Supplementary Table S3. Quality metrics.

Note: Asterisk denotes significant *p*-value. Standard deviation (SD); contrast-to-noise ratio (CNR).

Supplementary Table S4. Linear and quadratic effects of age for thalamocortical percent total connectivity

Note: Asterisk denotes significant *p*-value. Sex and rDisp are included as effects of no interest.

Supplementary Table S5. Linear effects of age for thalamocortical streamline counts

Note: Asterisk denotes significant *p*-value. Sex, rDisp, and total streamline count are included as effects of no interest.

Supplementary Table S6. Thalamocortical percent total connectivity values by sex

Note: Asterisk denotes significant *p*-value. Age, quadratic age, and rDisp are included as effects of no interest.

Supplementary Table S7. Thalamocortical streamline counts by sex

Note: Asterisk denotes significant *p*-value. Age, quadratic age, rDisp, and total streamline count are included as effects of no interest.

Supplementary Table S8. Thalamocortical percent total connectivity values by group

Note: Asterisk denotes significant *p*-value. Age, quadratic age, sex, and rDisp are included as effects of no interest.

Supplementary Table S9. Thalamocortical streamline counts by group

Note: Asterisk denotes significant *p*-value. Age, quadratic age, sex, rDisp, and total streamline count are included as effects of no interest.

Supplementary Table S10. Seed/target Volumes (mm3)

Note: Asterisk denotes significant *p*-value. No covariates were included in between-group volume comparisons.

Supplementary Table S11. Association between cognition and thalamocortical percent total connectivity

Note: Asterisk denotes significant *p*-value. For each region tested, all other regions, along with age, sex, group, and rDisp are included as effects of no interest.

Supplementary Table S12. Linear and quadratic effects of age for thalamocortical FA

Note: Asterisk denotes significant *p*-value. Sex and rDisp are included as effects of no interest.

Supplementary Table S13. Linear and quadratic effects of age for thalamocortical MD

Note: Asterisk denotes significant *p*-value. Sex and rDisp are included as effects of no interest.

Supplementary Table S14. Thalamocortical FA values by sex

Note: Asterisk denotes significant *p*-value. Age, quadratic age, and rDisp are included as effects of no interest.

Supplementary Table S15. Thalamocortical MD values by sex

Note: Asterisk denotes significant *p*-value. Age, quadratic age, and rDisp are included as effects of no interest.

Supplementary Table S16. Thalamocortical FA values by group

Note: Asterisk denotes significant *p*-value. Age, quadratic age, sex, and rDisp are included as effects of no interest.

Supplementary Table S17. Thalamocortical MD values by group

Note: Asterisk denotes significant *p*-value. Age, quadratic age, sex, and rDisp are included as effects of no interest.

Supplementary Table S18. Sensitivity analysis for age, sex, and group effects of FA controlling for whole brain FA

Note: Asterisk denotes significant *p*-value. rDisp and whole brain FA are included as effects of no interest for all analyses. Additional effects of no interest by model are: age model: sex; sex model: age and quadratic age; group model: age, quadratic age, and sex.

Supplementary Table S19. Sensitivity analysis for group effects of FA controlling for race, parental education, and CNR

Supplementary Table S20. Interaction effects in thalamocortical white matter microstructure

Note: Asterisk denotes significant *p*-value. Sex, group, and rDisp are included as effects of no interest.

Supplementary Figure S1. Cortical segmentation

Supplementary Figure S1. Six bilateral cortical regions were included as thalamic targets for the probabilistic tractography analysis.

Supplementary Figure S2. White matter tracts between the thalamus and cortical ROIs

Supplementary Figure S2. Group masks for each thalamocortical white matter tract are overlaid on a T1 template.

Supplementary Figure S3. Connectivity distribution by cortical region

Supplementary Figure S3. Distribution of thalamocortical percent total connectivity values (A), streamline counts (B), and white matter FA (C) by cortical target.

Supplementary Figure S4. Linear effects of age were less pronounced in the left hemisphere (A) than the right hemisphere (B).

Supplementary Figure S5. Effects of age, sex, and group for thalamocortical streamline counts

Supplementary Figure S5. Streamline count analyses showed similar age and group effects as percent total connectivity analyses. Sex effects were stronger, with one additional cortical region (motor cortex) showing greater streamline counts in males. 95% confidence intervals are shown for sex and group effects.

Supplementary Figure S6. Interaction of linear age and sex for thalamic-occipital percent total connectivity

Supplementary Figure S6. There was a linear interaction between age and sex, with older females showing more thalamic-occipital percent total connectivity than males.

Supplementary Figure S7. Effects of linear age, sex, and group for MD values

Supplementary Figure S7. (A) MD values decreased linearly with age across regions. (B) Males had lower mean MD in most white matter tracts, with the exception of the thalamic-occipital tract. (C) Typically-developing youth and youth with other psychopathologies had lower MD values compared to psychosis-spectrum youth in the thalamic-somatosensory and thalamicposterior parietal tract. 95% confidence intervals are shown for sex and group effects.

Supplementary Figure S8. Linear age effects of thalamocortical FA values, by hemisphere

Supplementary Figure S8. Linear effects of age were consistent across left (A) and right (B) hemispheres for FA values.

Supplementary Figure S9. Quadratic age effects of thalamocortical FA values

Supplementary Figure S9. FA values in the thalamic-posterior parietal and thalamic-occipital tracts followed and inverted-U shape, with higher FA values observed during adolescence and lower FA values in childhood and early adulthood.

Supplementary Figure S10. Quadratic age effects of thalamocortical MD values

Supplementary Figure S10. MD values followed a U-shaped curve, with lower MD values observed during adolescence and higher MD values in childhood and early adulthood.

Supplementary Figure S11. Interactions of linear age and sex for thalamocortical FA values

Supplementary Figure S11. Older males had higher FA values than older females in tracts connecting the thalamus with the prefrontal motor, somatosensory, and posterior parietal cortex.

Supplementary Figure S12. Linear age by group for thalamocortical FA values

Supplementary Figure S12. Age effects are plotted by (A) typically-developing and psychosisspectrum youth and (B) typically-developing youth and youth with other psychopathologies.

SUPPLEMENTAL REFERENCES

- 1. Calkins ME, Moore TM, Merikangas KR, Burstein M, Satterthwaite TD, Bilker WB, *et al.* (2014): The psychosis spectrum in a young U.S. community sample: Findings from the Philadelphia Neurodevelopmental Cohort. *World Psychiatry*. 13: 296–305.
- 2. Harrigan RL, Yvernault BC, Boyd BD, Damon SM, Gibney KD, Conrad BN, *et al.* (2016): Vanderbilt University Institute of Imaging Science Center for Computational Imaging XNAT: A multimodal data archive and processing environment. *Neuroimage*. 124: 1097–1101.
- 3. Sochat V V., Prybol CJ, Kurtzer GM (2017): Enhancing reproducibility in scientific computing: Metrics and registry for Singularity containers. *PLoS One*. 12: 1–24.
- 4. Dale AM, Fischl B, Sereno MI (1999): Cortical Surface-Based Analysis. *Neuroimage*. 9: 179– 194.
- 5. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, *et al.* (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 31: 968–980.
- 6. Ashburner J (2007): A fast diffeomorphic image registration algorithm. *Neuroimage*. 38: 95– 113.
- 7. Behrens TEJJ, Johansen-Berg H, Woolrich MW, Smith SM, Wheeler-Kingshott CAMM, Boulby PA, *et al.* (2003): Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat Neurosci*. 6: 750–757.
- 8. Andersson JLR, Sotiropoulos SN (2016): An integrated approach to correction for offresonance effects and subject movement in diffusion MR imaging. *Neuroimage*. 125: 1063–1078.
- 9. Smith SM (2002): Fast robust automated brain extraction. *Hum Brain Mapp*. 17: 143–155.