Supplementary Online Content

Peterson BS, Rauh VA, Bansal R, et al. Effects of prenatal exposure to air pollutants (polycyclic aromatic hydrocarbons) on the development of brain white matter, cognition, and behavior in later childhood. *JAMA Psychiatry*. Published online March 25, 2015. doi:10.1001/jamapsychiatry.2015.57.

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Supplemental Methods

Measures of PAH Exposure We represented prenatal airborne PAH exposure by the sum of 8 nonvolatile PAH: benzo[a]anthracene, chrysene/iso-chrysene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[a]pyrene, indeno[1,2,3-c,d]pyrene, dibenzo[a,h]anthracene, and benzo[g,h,i]perylene. These 8 PAH were measured using personal air monitoring of the mothers over a 48-hour period in the third trimester of pregnancy.[1] Vapors and particles ≥2.5 μg in diameter were collected on a pre-cleaned quartz microfiber filter and a pre-cleaned polyurethane foam cartridge backup. The samples were analyzed at Southwest Research Institute (San Antonio, TX) for concentrations of all 8 individual PAH compounds. The flow rate, time, and completeness of documentation of each personal monitoring measure were assessed for accuracy. Monitoring of the mothers of all 40 children in the present study provided samples of acceptable quality.

In addition, we previously validated 48-hour personal air monitoring measures of prenatal exposure to these 8 PAH [2] against measures from concurrent 2-week monitoring of residential air samples in a representative subset of homes (n=101) over 6 weeks in the final trimester of pregnancy.[3] The levels of the pollutants measured in weeks 1-2 correlated significantly with levels measured in weeks 2-6 ($r=0.0.57\text{-}0.76$, $p<0.001$). The indoor air levels of the pollutants over the 6 weeks of monitoring correlated significantly with PAH levels in the single 48-hour prenatal personal air sample ($r=0.44$, $p<0.001$). A separate study of pregnant women has shown a high inter-correlation of concurrent indoor, outdoor, and personal monitoring of PAH (pairwise Spearman's coefficients for 9 PAHs ≥ 0.84, $p < 0.01$).[4] These findings indicate that our single measure of PAH in prenatal personal air is a reasonable indicator of chronic prenatal exposure via inhalation.

Postnatal Urinary PAH Metabolites A suite of 9 PAH metabolites was measured at the CDC in spot urine (collected from the child at age 5) using automated liquid-liquid extraction and gas chromatography/isotope dilution high-resolution mass spectrometry.[5-7] Although a relatively short-lived biomarker (half-life 6-35 hours),[8] in conditions of chronic exposure such as in our study population, the metabolites provide a useful measure of exposure to PAH.[5,6] Specific gravity (SG) measurements, obtained using a handheld refractometer (Urine-Specific-Gravity-Refractometer-PAL-10-S-P14643C0;TAGO USA, Inc., Bellevue, WA), were used to control for urinary dilution of the samples with the formula: fresh weight metabolites for the child*(mean SG-1)/(SG for that child-1).[9] The 9 PAH metabolites were summed and then natural log-transformed for use in statistical analyses. The 9 metabolites included: 1-Hydroxynaphthalene, 2-Hydroxynaphthalene, 2-Hydroxyfluorene, 3-Hydroxyfluorene, 9-Hydroxyfluorene, 1-Hydroxyphenanthrene, 2-Hydroxyphenanthrene, 3-Hydroxyphenanthrene, and 4-Hydroxyphenanthrene.

Cotinine Levels Maternal blood (30-35 ml) was collected in the hospital within 24 hours of delivery, and umbilical cord blood (30-60 ml) was collected at delivery. Samples were transported to the laboratory immediately, where buffy coat, packed red blood cells, and plasma samples were separated and stored at -70°C. A portion of each sample was shipped to the Centers for Disease Control for analysis of plasma cotinine using high-performance liquid chromatography atmospheric-pressure ionization tandem mass spectrometry.[10] The maternal and cord plasma concentrations of cotinine were significantly intercorrelated ($r=0.88$, $p<0.001$), and maternal values were substituted for missing cord cotinine values.

MRI Acquisition Anatomical images were acquired on a GE Signa 3 Tesla whole body scanner (Milwaukee, WI) equipped with a body transmitter coil and an 8-channel head receiver

coil. High-resolution, T1-weighted images of the brain were acquired using a Fast Spoiled Gradient Recall sequence: Inversion Time=500 ms, Repetition Time=4.7 ms, Echo Time=1.3 ms, Field of View=24 cm, image matrix=256x256, acceleration factor=2, number of slices=160, slice thickness=1 mm encoded for sagittal slice reconstruction, providing voxel dimensions of 0.9375x0.9375x1.0mm).

Image Processing All processing was performed on Sun Ultra 10 workstations using ANALYZE 8.0 software Biomedical Imaging Resource, Mayo Foundation, Rochester, MN), together with software developed in-house, while blind to participant characteristics and hemisphere (images were randomly flipped in the transverse plane prior to preprocessing). All morphometric analyses were performed with the MRI dataset resliced to correct for any residual head rotation, tilt, or flexion/extension.

Preprocessing Large-scale variations in image intensity were corrected using a validated algorithm developed at the Montreal Neurological Institute [11]. Extracerebral tissues were removed using an automated tool for extracting the brain [12]. This tool first smooths image intensity using an anisotropic filter, detects three-dimensional edges using Marr-Hildreth edge detector [13], and then selects as the brain the largest connected component with a closed boundary. Connecting dura was removed manually on each slice in sagittal view and checked in the orthogonal views. The brainstem was transected at the pontomedullary junction.

Cortical Gray Matter Segmentation Gray-scale values of "pure" representations of cortical gray matter and white matter were sampled bilaterally at 4 standard locations throughout the brain (frontal, temporal, occipital, parietal) using an 8x8=64 pixel array, sufficiently large to provide statistical stability but small enough to avoid partial volume effects that include other tissue types. These 4 values were then averaged for each tissue type, and a threshold value halfway

between the mean gray matter and mean white matter values was computed for a slice in the imaging volume. We invoked these threshold values on a slice-by-slice basis to provide an initial rough classification of cortical gray and white matter throughout the cerebrum. This classification was then hand edited in the coronal and transverse views to provide the most accurate segmentation possible of the cortical mantle. The intraclass correlation coefficient calculated using a 2-way random effects model[14] as a measure of the reliability of cortical gray matter volumes was 0.98.

Overview of the Analysis of Surface Morphologies

The analysis of cerebral surfaces previously has been rigorously validated using synthetic and real-world datasets.[15] Comparing surfaces within one group or across two or more groups of brains requires determining which points on the surface of each brain correspond anatomically with the points on the surfaces of all the other brains in the comparison. When these points of correspondence have been established, a signed distance (the Euclidean distance, with positive distances for outward deformations and negative distances for inward deformations) between the corresponding points within and between groups can be evaluated statistically. Our method for determining these points of anatomical correspondence across individual brain surfaces uses a two-step procedure for determining the point correspondences between the surface of each participant brain and the surface of a template brain. In the first step, each brain is coregistered using a similarity (rigid body with global scaling) transformation to the template brain such that the cerebral surfaces are moved to a close approximation to the template surface. In the second step, each participant brain is treated as a fluid flowing into the template brain, and therefore constitutes a high-dimensional, nonlinear warping of the entire 3D volume of each participant brain to the entire 3D volume of the template brain. Each participant brain therefore appears

exactly as the template brain appears, including every point of their outer surfaces, thereby establishing precise point-to-point correspondences across the surface of each participant brain with the template brain. The nonlinear warping of each participant brain is then reversed to bring each brain into the initial correspondence established by the similarity coregistration, bringing each label for correspondence matching of cerebral surfaces that had been established previously through the nonlinear warping. The physical distances are measured from each point on the surface of each participant's brain to the corresponding point on the surface of the template brain. Those distances at each point of the template brain, one for each participant, constitute a continuous variable that can either be compared across groups of participants, related to another variable within a group of participants, or simply quantified using traditional descriptive statistics. The parameters of those statistical analyses are then color-coded and displayed at each voxel on the surface of the template brain. Finally, measures of cortical thickness at each point on the surface of each participant's brain can replace distance from the template brain as the continuous variable of interest in those statistical analyses. The following sections provide details of the steps in coregistration, the procedures used in the selection of the template brain, and the statistical modeling procedures used in our population of participants.

Initial Similarity Coregistration

The brains in our dataset were first brought into a common coordinate or template space. Following isolation of the brain from non-brain tissue (see the main text), we used mutual information (MI) of gray scale values to register the brain of each participant to the template brain.[16] The template brain and each of the other brains in our sample were treated as rigid bodies as the optimal translation, rotation, and scaling parameters maximized the mutual information $I(F;R)$ between each brain and the template. Let $F = \{F_1, F_2, \ldots, F_n\}$ denote the

random field from which one of the volumes, called the *participant* volume, was sampled. Similarly, let $R = \{R_1, R_2, \ldots, R_m\}$ be the random field from which the other volume, called the *template* volume, was sampled. Assuming that the random variables $\{F_1, F_2, \ldots, F_n\}$ were independently and identically distributed (i.i.d.), let F be the random variable which represented the voxel intensities in the participant volume. Similarly, assuming that $\{R_1, R_2, \ldots, R_m\}$ were i.i.d, let R be the random variable representing the voxel intensities in the template volume. Then the mutual information $I(F;R)$ between the two random variables F and R with marginal probability density functions $P_F(f)$ and $P_R(r)$ and joint probability density function $P_{F,R}(f,r)$ is defined to be $I(F,R) = \sum_{f,r} P_{F,R}(f,r) \log_2 \left(\frac{P_{F,R}(f,r)}{P_F(f) \times P_R(r)} \right)$. Let α denote the set of transformation parameters -- 3 translations, 3 rotations, and scale -- by which the participant volume was transformed. Then the probability density function of the transformed participant volume, as a function of α, is denoted as $P_{F,\alpha}(f)$. Similarly, the joint density function is denoted by $P_{F,R,\alpha}(f,r)$. The probability density function, $P_{R,\alpha}(r)$, is a function of α because we were evaluating the mutual information of overlapping regions in the two volumes for various values of α. The joint and the marginal probability density functions were estimated using the Parzen window method[17]. Using these notations, the mutual information, $I_\alpha(F,R)$, was evaluated as a function of α as:

$$I_\alpha(F,R) = \sum_{f,r} P_{F,R,\alpha}(f,r) \log_2 \left(\frac{P_{F,R,\alpha}(f,r)}{P_{F,\alpha}(f) \times P_{R,\alpha}(r)} \right). \quad (1)$$

The optimal set of transformation parameters α that maximized $I_\alpha(F,R)$ were then found using the method of stochastic gradient descent.[16]

Nonlinear Warping Using Fluid Flow Dynamics

Following the similarity transformation of each brain to the template brain described above, we used the method of *fluid flow* (FF) to refine this coregistration. FF uses a nonlinear deformation of one brain surface to another by treating the gray scale image of one brain as a fluid that flows into the gray scale image of another brain so as to match its gray scale characteristics precisely.[18] [3,6]. The FF algorithm therefore uses a high dimensional, nonlinear warping of images to identify corresponding points on the participant and reference surfaces. This algorithm modeled the warping of gray scale in the images as the flow of fluid using the partial differential equations (PDEs) of the Navier-Stokes equations. In fluid flow dynamics, nonlinear deformations of large magnitude are permitted because the stress (force per unit area) that restrains local deformations relaxes over time. The PDE governing the deformation is as follows[18]:

$$\mu\nabla^2\vec{v}(\lambda+\mu)\vec{\nabla}(\vec{\nabla}\cdot\vec{v})+\vec{b}(\vec{u})=\vec{0}, \tag{2}$$

where $\nabla^2=\nabla^T\nabla$ is the Laplacian operator, $(\vec{\nabla}\cdot\vec{v})$ is the divergence operator, μ and λ are the viscosity coefficients, and $\vec{v}(\vec{x},t)$ is the velocity of the particle at time t and position $\vec{x}$ in the Eulerian reference frame. In the Eulerian reference frame, a voxel located at $\vec{x}$ at time t, with a displacement vector $\vec{u}(\vec{x},t)$, originated from a point $\vec{x}-\vec{u}(\vec{x},t)$. In Eqn. (2), the term $\vec{b}(\overrightarrow{u(\vec{x},t)})$ denotes the body force acting on a particle located at $\vec{x}$ at time t, which is the driving force for the local nonlinear deformations of an image.

The PDE in Eqn. (2), defined on a domain $\Omega=[0,1]^3$, was a boundary value problem that was solved numerically using the successive over-relaxation (SOR) method with checker board updates.[19] Zero boundary conditions were assumed for $\vec{v}(\vec{x},t)$ along the boundary $\partial\Omega$ of the domain Ω at all t, and therefore the displacements of the voxels at $\partial\Omega$ is zero. Because the

displacement field $\vec{u}(\vec{x},t)$ was given in the Eulerian reference frame, the velocity field $\vec{v}(\vec{x},t)$ and the displacement fields $\vec{u}(\vec{x},t)$ were related as :

$$\vec{v} = \frac{d\vec{u}}{dt} = \frac{\partial \vec{u}}{\partial t} + \sum_i v_i \frac{\partial \vec{u}}{\partial x_i}. \quad (3)$$

We used the *Forward Time Centered Space* (FTCS)[19] method for the numerical estimation of the various partial derivatives in Eqn. (2).

The modeling of fluid dynamics has been used extensively in various settings for the registration of 3D images.[18] We used this algorithm, however, for identifying the corresponding points on the surface of the participant and template brains, rather than for simple registration. Using this high dimensional warping algorithm, each participant brain was warped to the template brain, and points on the surface of the template brain were used to label points on the surface of each participant's brain. Each participant brain was then unwarped and brought back along with the labeled points to its original conformation in the common coordinate space established in the initial rigid body registration. Signed Euclidean distances between the corresponding labeled points of the template and participant brain were then calculated.

Controlling for Morphological Scaling Effects

Larger body size predicts larger overall brain size, and larger overall brain size predicts larger sizes of individual brain regions, a phenomenon termed "morphological scaling". The similarity transformation that we use for coregistration includes a term for overall scale of the brain being registered to the template brain, which brings all brains in the sample to the same volume as the template brain and thereby controls for morphological scaling.

Selection of the Template Brain We used a two-step procedure to select a template brain that was most representative of our cohort. First, we identified a brain that was as representative as

possible of the demographics of the sample in terms of age, weight, and height. We coregistered the brains for all remaining participants to this preliminary template to identify the point correspondences across the surfaces of the brains, and we measured the distances of each point on the surface of each brain from the corresponding point on the surface of the template brain. We then identified the brain in our sample for which all points across the surface were closest (in terms of least squares) to the average of the computed distances, in effect identifying the brain whose surface contour was closest to the average contour in the overall sample. We then used this brain as the final template brain. All brains then underwent a second coregistration to identify point correspondences across brains, but this time to the final template brain that was morphologically most representative of the surface features of the entire sample. Note that we used a single representative brain as a template rather than an averaging of gray-scale images from many brains because a single brain has well-defined tissue interfaces, such as the CSF-gray matter or gray-white matter interfaces. Averaging images for a template blurs tissue boundaries and increases registration errors that are subtle but important when distinguishing subtle effects across individuals and populations. The selection of the individual template brain has little influence on the final statistical findings, as the same statistical analyses performed either on the preliminary template brain or on a randomly selected brain yields the same findings.

Cortical Thickness We subtracted the cortical mantle from the coregistered brain of each participant. We then used a three-dimensional morphological operator to distance-transform this brain without the cortex from the coregistered brain of the same participant containing the cortex.[20,21] This operation calculated cortical thickness in each coregistered brain as the smallest distance of each point on the external cortical surface from the outermost surface of the white matter. As we measured these thicknesses in template space, their values inherently accounted

for generalized scaling effects within the cerebrum because the brain and its local features, such as cortical thickness, already had been scaled during the similarity transformation of that brain to the template.

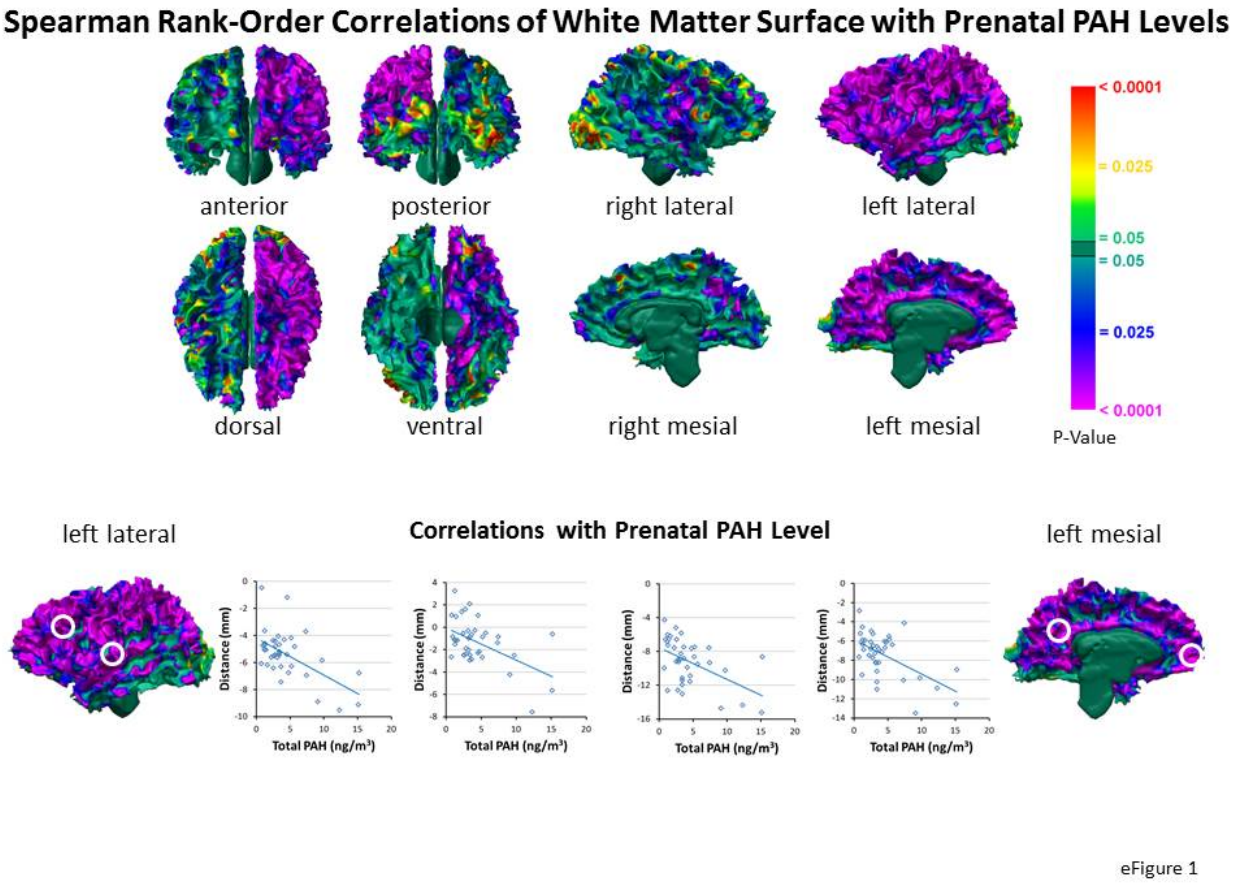
Sulcal Overlay We displayed statistical results on the ICBM (International Consortium for Brain Mapping) high-resolution, single-participant template [22]. The delineation of cortical gyri was performed manually by an expert neuroanatomist, and this 3-D set of labels was mapped onto the cortical surface of our template brain to delineate regional boundaries to aid localization of findings.

Mediator Analyses We used mediator analyses to test our hypothesis that brain surface measures mediated the relationship between prenatal PAH exposure level and processing speed in later childhood. At each point on the surface of the template white matter, we tested the statistical significance that variable M (white matter measures) mediated the association between the independent variable, X (prenatal PAH levels), and the dependent variable, Y (processing speed), using three regression equations: (1) $Y=i_1+cX+e_1$; (2) $Y=i_2+c'X+bM+e_2$; and (3) $M=i_3+aX+e_3$, where i_1, i_2, and i_3 were intercepts, *a, b, c, and c'* were coefficients, and e_1, e_2, and e_3 were error terms for the regressions.[23] Testing the statistical significance of mediation required demonstrating the significance of both the coefficient *c* (using a t-test in the linear regression) and the product *a x b* (also using a t-test, which we approximated as $ab/\sqrt{b^2\sigma_a^2+a^2\sigma_b^2}$, where σ_a^2 and σ_b^2 were the standard deviations of *a* and *b* estimated from the corresponding linear models). We then plotted the p-value for *a x b* at each voxel across the cerebral surface after using False Discovery Rate to correct for the number of statistical comparisons.

eDiscussion. Supplemental Discussion

One plausible alternative explanation for the PAH-associated ADHD symptoms in the present study is that ADHD symptoms could have been transmitted genetically to the children from mothers who themselves have ADHD symptoms, and those mothers could have been more likely than mothers without ADHD to have been exposed to higher levels of air pollutants during pregnancy. If such a gene-environment correlation were indeed responsible for our findings, then presumably the same correlation would hold for any group of women with ADHD, including those in the prior imaging studies of ADHD. At least, it would not be clear why the gene-environment correlation would operate in our sample and not the samples from prior ADHD studies. The dissimilarity of morphological findings associated with ADHD symptoms in our sample from those in prior studies of ADHD therefore suggests that, rather than higher PAH levels being a consequence of ADHD in the mothers we recruited, prenatal exposure to higher PAH levels likely contributed to the ADHD symptoms in our sample. An analogous line of reasoning obtains for the associations of PAH levels with processing speed and externalizing problems.

eFigure 1. Nonparametric Correlations and Scatterplots for Prenatal PAH Exposure



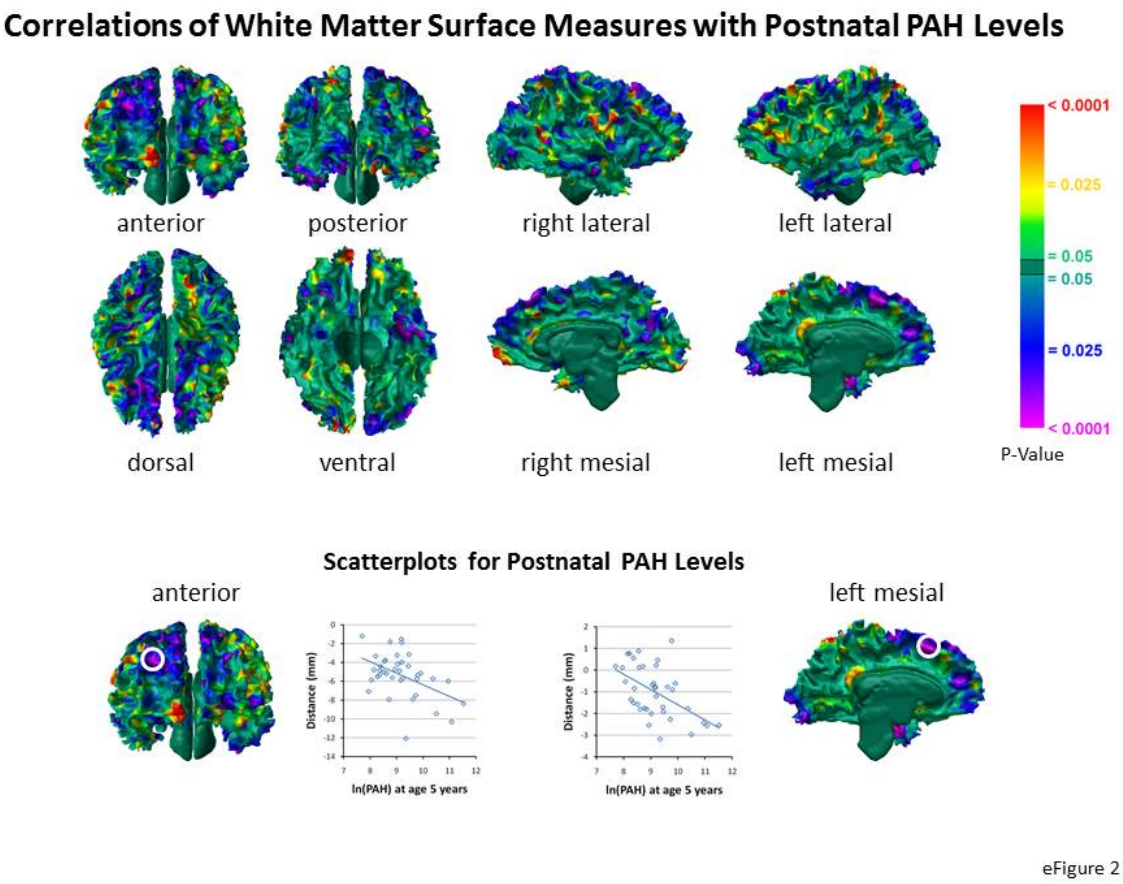
eFigure 1

To ensure that outliers were not unduly influencing the correlations of PAH levels with white matter surface measures, we assessed those associations using the nonparametric Spearman Rank-Order correlation coefficient and visual inspection of the scatterplots for the associations.

Upper panel: Spearman correlation after regressing the white matter surface measures on age and sex of the participants. The findings were unchanged from those using parametric analyses (Fig.1 main text).

Lower panel: Scatterplots show that the association was not driven by the presence of outliers in the dataset. The white circles indicate where in the brain the dataset was sampled to generate the scatters.

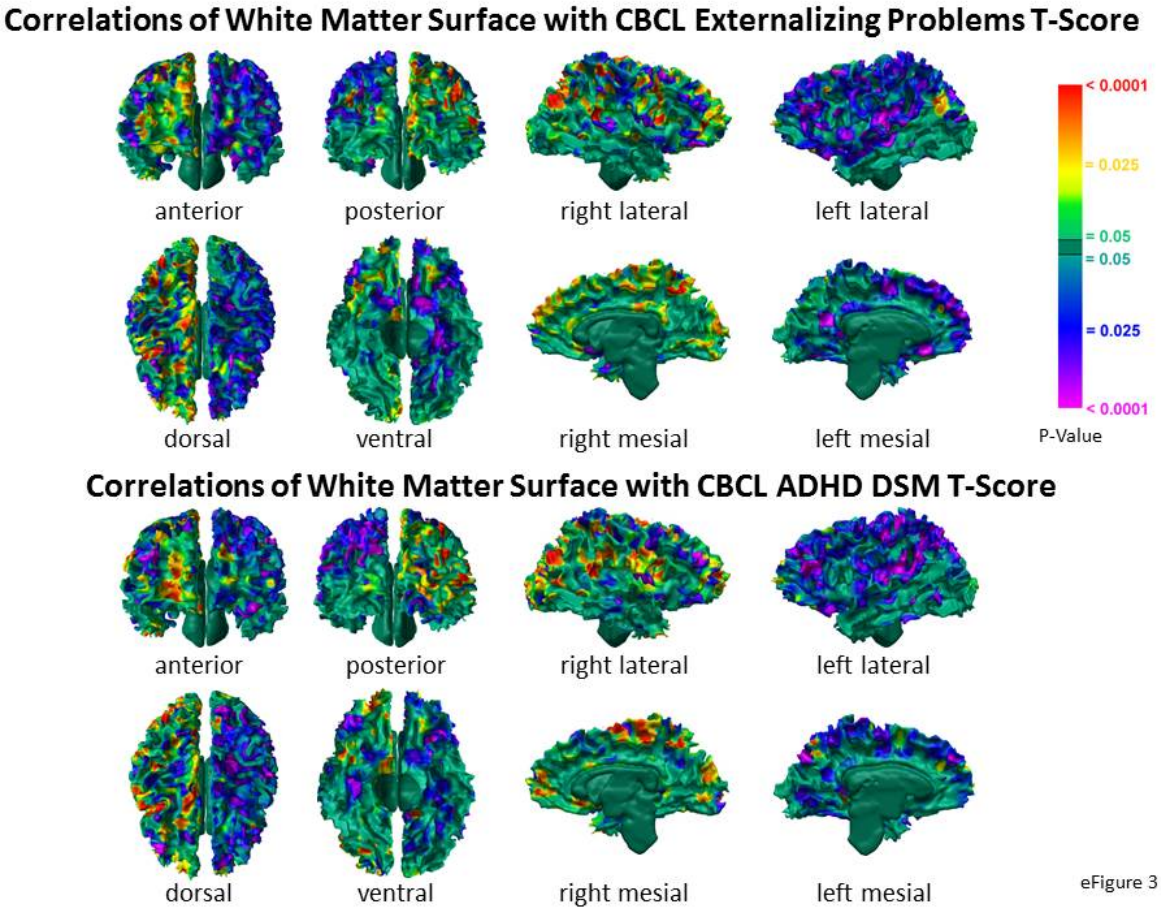
eFigure 2. Correlations of Postnatal PAH Levels With White Matter Surface Measures



Upper panel: Maps for these correlations are as described for Figure 1 of the main text, except the regressions are for postnatal PAH exposure levels measured at age 5. The analyses covaried for age, sex, and prenatal PAH levels. Significant inverse correlations of postnatal PAH levels with white matter measures are most prominent over the surfaces of the dorsolateral prefrontal regions bilaterally, especially over the superior frontal gyri. The values for postnatal PAH metabolite levels have been natural log-transformed.

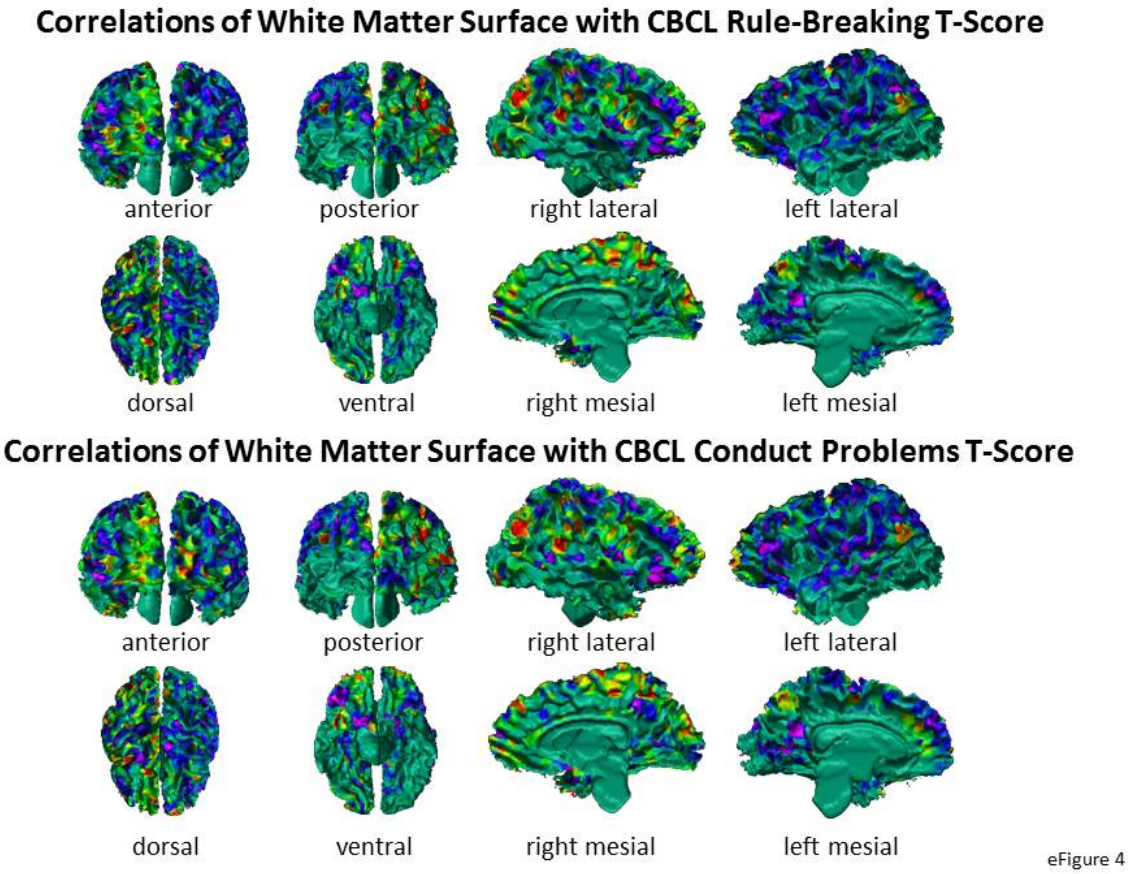
Lower panel: Scatterplots show that the significant correlations derive from the entire range of postnatal PAH values and are not driven by outliers.

eFigure 3. Correlations of White Matter Surface Measures With CBCL Externalizing Problems and ADHD-DSM Symptoms

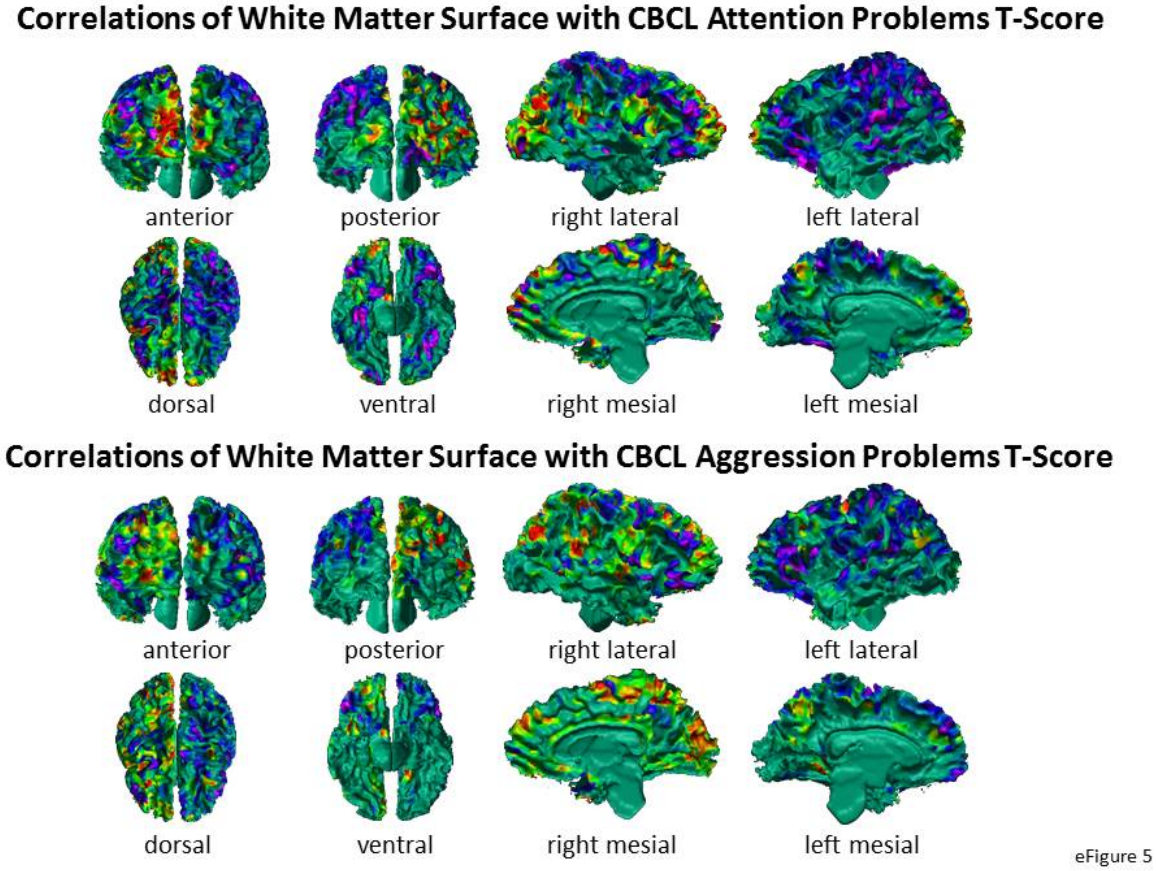


P-values that are FDR-corrected for multiple comparisons are plotted for partial correlations of behavioral measures with distances at each point on the white matter surface from the corresponding point on the white matter surface of the template brain. Warm colors (yellow, orange, and red) represent significant positive correlations and cool colors (blue and purple) representing inverse correlations. A sea-green color indicates correlations that are not statistically significant. Significant findings indicate that reductions in white matter surface measures of the left hemisphere accompany more severe externalizing problems and more severe symptoms of ADHD on the CBCL.

eFigure 4 and eFigure5. Correlations of White Matter Surface Measures With Additional CBCL Scores



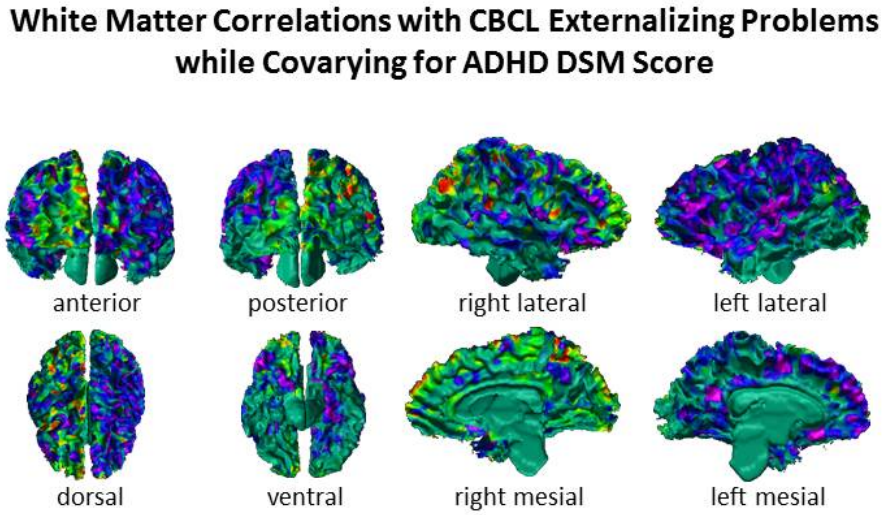
eFigure 4



eFigure 5

The correlations with additional CBCL T-scores include those for Rule-Breaking, Conduct Problems, Attention Problems, and Aggression Problems. P-values that are FDR-corrected for multiple comparisons are plotted for partial correlations of behavioral measures with distances at each point on the white matter surface from the corresponding point on the white matter surface of the template brain. Warm colors (yellow, orange, and red) represent significant positive correlations and cool colors (blue and purple) representing inverse correlations. A sea-green color indicates correlations that are not statistically significant.

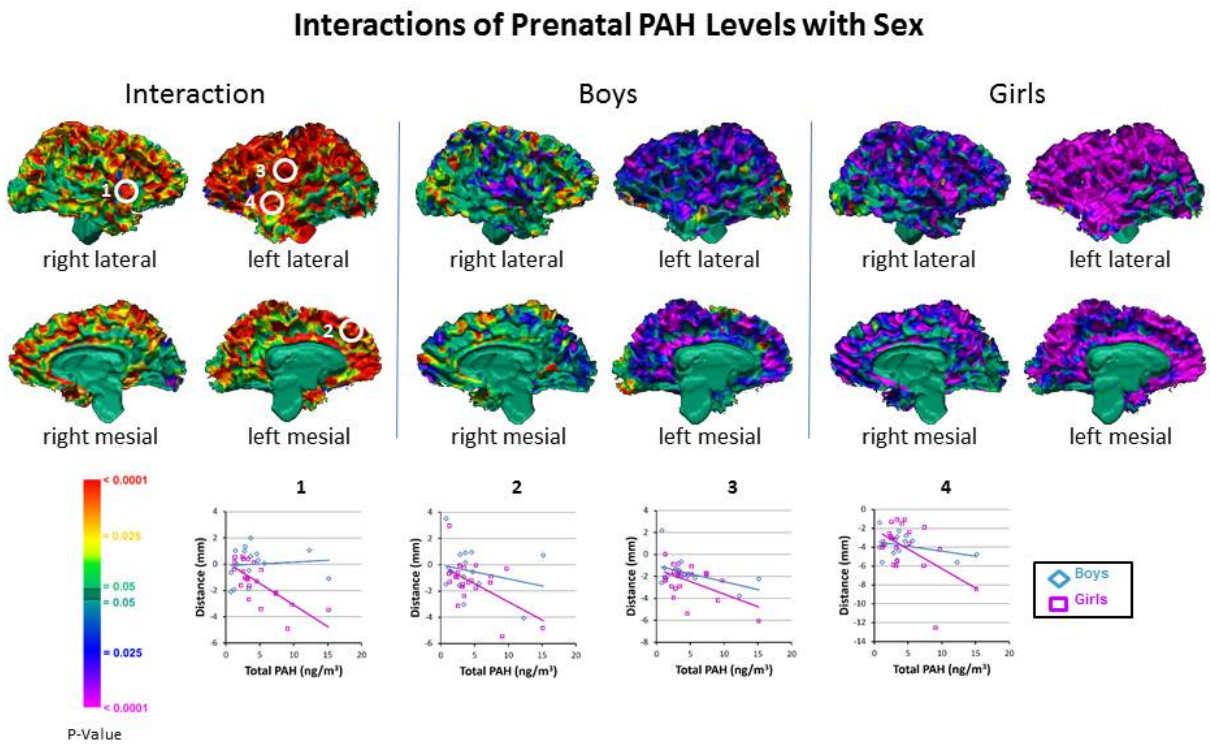
eFigure 6. Correlation of White Matter Surface Measures With CBCL Externalizing Problems While Covarying for ADHD DSM Scores



eFigure 6

This figure is the same as that in the top panel for Figure 5 of the main text, except that it covaries for ADHD DSM T-scores in addition to age and sex. The findings are essentially unchanged from those when not covarying for ADHD DSM scores.

eFigure 7. Interactions of Prenatal PAH Exposure Levels With Sex on White Matter Surface Measures



eFigure 7

The significant inverse associations of prenatal PAH exposure with white matter surface measures were present in both boys (middle panel) and girls (right panel) of the sample, but interaction analyses (left panel) indicate that the associations were stronger in girls. Scatterplots, however, suggested that these sex differences were driven by relatively few participants of each sex, and therefore the findings should be regarded with caution and as preliminary. Representative points sampled on the cerebral surface are numbered in the interaction maps on the left, and the corresponding scatterplots are shown in the bottom row. Color coding is as described in prior figures.

ONLINE-ONLY SUPPLEMENTAL REFERENCES

1. Perera FP, Rauh V, Tsai WY, et al. Effects of transplacental exposure to environmental pollutants on birth outcomes in a multiethnic population. *Environ Health Perspect.* 2003;111(2):201-205.
2. Choi H, Jedrychowski W, Spengler J, et al. International studies of prenatal exposure to polycyclic aromatic hydrocarbons and fetal growth. *Environ Health Perspect.* 2006;114(11):1744-1750.
3. Rundle A, Hoepner L, Hassoun A, et al. Association of childhood obesity with maternal exposure to ambient air polycyclic aromatic hydrocarbons during pregnancy. *Am J Epidemiol.* 2012;175(11):1163-1172.
4. Choi H, Perera F, Pac A, et al. Estimating individual-level exposure to airborne polycyclic aromatic hydrocarbons throughout the gestational period based on personal, indoor, and outdoor monitoring. *Environmental health perspectives.* 2008;116(11):1509-1518.
5. Li Z, Romanoff LC, Trinidad DA, et al. Measurement of urinary monohydroxy polycyclic aromatic hydrocarbons using automated liquid-liquid extraction and gas chromatography/isotope dilution high-resolution mass spectrometry. *Anal Chem.* 2006;78(16):5744-5751.
6. Li Z, Sandau CD, Romanoff LC, et al. Concentration and profile of 22 urinary polycyclic aromatic hydrocarbon metabolites in the US population. *Environ Res.* 2008;107(3):320-331.
7. Miller RL, Garfinkel R, Lendor C, et al. Polycyclic aromatic hydrocarbon metabolite levels and pediatric allergy and asthma in an inner-city cohort. *Pediatr Allergy Immunol.* 2010;21(2 Pt 1):260-267.
8. Jongeneelen FJ, van Leeuwen FE, Oosterink S, et al. Ambient and biological monitoring of cokeoven workers: determinants of the internal dose of polycyclic aromatic hydrocarbons. *Br J Ind Med.* 1990;47(7):454-461.
9. Hauser R, Meeker JD, Park S, Silva MJ, Calafat AM. Temporal variability of urinary phthalate metabolite levels in men of reproductive age. *Environ Health Perspect.* 2004;112(17):1734-1740.
10. Bernert JT, Jr., Turner WE, Pirkle JL, et al. Development and validation of sensitive method for determination of serum cotinine in smokers and nonsmokers by liquid chromatography/atmospheric pressure ionization tandem mass spectrometry. *Clin Chem.* 1997;43(12):2281-2291.
11. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging.* 1998;17(1):87-97.
12. Shattuck DW, Leahy RM. BrainSuite: an automated cortical surface identification tool. *Med Image Anal.* 2002;6(2):129-142.
13. Marr D, Hildreth E. Theory of edge detection. *Proc R Soc Lond B Biol Sci.* 1980;207(1167):187-217.
14. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin.* 1979;86:420-428.
15. Bansal R, Staib LH, Wang Y, Peterson BS. ROC-based assessments of 3D cortical surface-matching algorithms. *Neuroimage.* 2005;24:150-162.

16. Wells WM, 3rd, Viola P, Atsumi H, Nakajima S, Kikinis R. Multi-modal volume registration by maximization of mutual information. *Med Image Anal.* 1996;1(1):35-51.
17. Duda RO, Hart PE, Stork DG. *Pattern classification.* 2nd ed. New York: Wiley; 2001.
18. Christensen GE, Joshi SC, Miller MI. Volumetric transformation of brain anatomy. *IEEE Trans Med Imaging.* 1997;16(6):864-877.
19. Press WH. *Numerical recipes in C++ : the art of scientific computing.* 2nd ed. Cambridge, UK ; New York: Cambridge University Press; 2002.
20. Haralick R, Shapiro L. Chapter 5. *Computer and Robot Vision.* Vol 1: Addison-Wesley Publishing Company; 1992.
21. Rosenfeld A, Pfaltz J. Distance Functions in Digital Pictures. *Pattern Recognition.* Vol 11968:33-61.
22. Mazziotta J, Toga A, Evans A, et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences.* 2001;356(1412):1293-1322.
23. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol.* 2007;58:593-614.

eTable. Correlation Matrix for PAH Levels and Select Behavioral Measures on the CBCL and WISC-IV

	Prenatal PAH	Postnatal PAH	Anxiety-Depression	Attention	DSM-ADHD	DSM-Conduct	Externalizing	Processing Speed	FSIQ
Prenatal PAH	1								
Postnatal PAH	-0.10, 0.54	1							
Anxiety-Depression	0.09, 0.60	0.08, 0.64	1						
Attention	-0.19, 0.24	-0.23, 0.17	0.26, 0.12	1					
DSM-ADHD	-0.04, 0.80	-0.01, 0.94	0.30, 0.08	0.82, 0.0001	1				
DSM-Conduct	-0.04, 0.82	-0.10, 0.57	-0.10, 0.57	0.63, 0.0001	0.70, 0.0001	1			
Externalizing	-0.15, 0.39	-0.02, 0.91	0.44, 0.006	0.57, 0.0001	0.66, 0.0001	0.78, 0.0001	1		
Processing Speed	-0.32, 0.047	-0.14, 0.40	0.003, 0.99	0.03, 0.88	0.08, 0.63	-0.03, 0.86	0.07, 0.69	1	
FSIQ	-0.08, 0.64	0.02, 0.91	0.01, 0.97	-0.22, 0.18	-.16, 0.35	-0.31, 0.06	-0.03, 0.87	0.62, 0.0001	1

Each cell of the table shows first the Pearson correlation coefficient associating each pair of variables and then the associated p-value for that correlation coefficient. Statistically significant correlations are highlighted in red.
FSIQ: full scale IQ