SUPPLEMENTARY MATERIAL

MATERIALS AND METHODS

Subject selection

Recruitment of HCW and AR-PPD groups was balanced across the study. The HCW group included women with an EPDS ≤5 and no current or past psychiatric diagnosis or family history of psychiatric illness, as ascertained by clinical and research interviews [1]. The EPDS was used to assess peripartum depressive and anxiety symptoms [2,3] and a cut-off score of ≥10 was chosen to identify women with current depressive and anxiety symptomatology. As the EPDS is not sufficiently accurate in predicting risk of postpartum depressive symptoms alone [4], the AR-PPD group included women who either had an EPDS score ≥10 (indicating current depressive and/or anxiety symptomatology) or, regardless of antepartum EPDS score at study entry, a personal history of PPD or non-puerperal depression as determined by the Structured Clinical Interview for DSM-IV TR Disorders (SCID-IV), Patient Edition [1]. Since antepartum anxiety and depression symptoms are associated with, or may represent the early presentation of PPD symptomatology, women who met criteria for an anxiety disorder or depressive disorder not otherwise specified were included in the AR-PPD group. Women who met SCID-IV criteria for a major depressive episode (MDE) at study entry were excluded as the main aims of the prospective imaging study were to examine peripartum blood NAS and postpartum intrinsic RSFC and cortical GABA concentrations in those women who developed peripartum adjustment and minor/major depressive disorders compared to women who remained euthymic.

Participants were excluded for multiple gestation pregnancy, lifetime history of manic episode or any psychotic disorder, elevated suicidal risk, and alcohol, nicotine or substance abuse/dependence in the 6 months prior to study entry or use during the study, contraindication to MRI, positive urine pregnancy test at time of MRI. Participants were medication-free except

prenatal vitamins, as needed over the counter antacids, antihistamines and stool softeners were allowed. All participants provided written informed consent and each received monetary compensation for participation. Study data was managed using Research Electronic Data Capture (REDCap)[5].

Study procedures

For this manuscript, the use of "PPD" refers to women who developed a new onset adjustment disorder with depressed mood or depressive disorder (minor or major) during pregnancy or the postpartum period under study. The SCID-IV was completed at visits 1 and 4 or 5 (at time of postpartum MRI), the EPDS was completed at visits 1-5 and telephone EPDS were attempted weekly during the postpartum to monitor the development of PPD symptoms as evidenced by rising total EPDS score. Additional research assessments done at all 5 visits included Structured Interview Guide for Hamilton Depression Rating Scale ($HAM\text{-}D_{17}$) [6], Structured Interview Guide for Hamilton Anxiety Scale (HAM-A)[7,8], Spielberger State-Trait Anxiety Inventory (STAI)[9], Pittsburgh Sleep Quality Index (PSQI) [10], Sheehan Disability Scale (SDS) [11] and urine benzodiazepine test. Assessments done at visit 1 included the SCID-IV and past medical history/demographics. Additional assessments done at the time of postpartum MRI (visit 4 or 5) included a labor and delivery questionnaire, a menses and breastfeeding recording form and a urine pregnancy test. For women AR-PPD, the MRI was scheduled based on when the weekly telephone EPDS total score started to rise and availability for the participant to come to the research center. Diagnosis of adjustment disorder with depressed mood, minor depressive disorder or major depressive disorder with peripartum onset was confirmed by SCID-IV at time of MRI. HCW were scheduled for postpartum MRI with blood draw to match the days since delivery when women with postpartum adjustment and depressive disorders were scanned and blood was drawn so that women would be matched for postpartum time which could affect plasma NAS, RSFC patterns or cortical GABA+/Cr concentrations. Women who

developed adjustment disorder with depressed mood were eligible for imaging as the study was designed to examine a range of depressive symptom severity in alignment with a Research Domain Criteria approach to psychiatric neuroscience[12]. The use of “PPD” for the sole purpose of this manuscript refers to all women who developed a new onset adjustment disorder with depressed mood or depressive disorder (minor or major) during the peripartum period under study. Only AR-PPD women who developed a new onset adjustment disorder with depressed mood or depressive disorder (minor or major) during pregnancy or the postpartum period under study were examined with fMRI and MRS. AR-PPD women do did not develop one of the above diagnoses were not eligible for fMRI and MRS.

Blood samples for neuroactive steroid analyses were obtained in the morning and collected into tubes containing EDTA. Plasma was stored at -80°C until analysis was completed by collaborators blind to the participant group assignment.

fMRI data acquisition and analysis

Image acquisition.

The imaging protocol for fMRI data acquisition was constant for all participants throughout the study and performed on the same research scanner (Aim 1). T1-weighted anatomical MRI images (MPRAGE sequence, 256 × 252 voxels, TR: 6.76 ms, TE: 3.1 msec, FOV: 244 mm × 256 mm × 204 mm, 170 slices) were collected for diagnostic and localization purposes. Additionally, T2-weighted TSE scans were collected (560 × 560 voxels, TR: 3000 ms, TE: 80 ms, FOV: 250 mm × 202 mm × 119 mm, 30 slices) to serve as intermediate registration targets. Resting-state scan images were obtained using an EPI sequence (84 × 81 voxels, TR: 2500 ms, TE: 30 msec, FOV: 256 mm × 256 mm × 150 mm, flip angle: 75°, slice thickness: 3 mm, 50 slices) lasting 406 seconds. All participants underwent the resting-state MRI scan with open eyes and were instructed to attend to a static image while thinking of nothing in

particular[13]. The static image contained a white plus-sign superimposed on the middle of a black background and was projected onto a screen visible through a mirror mounted on the head coil. Resting-state images were preprocessed in FSL (realignment, slice timing correction, spatial smoothing (FWHM=5mm), skull stripping) http://www.fmrib.ox.ac.uk/fsl). Spatial normalization to the standard MNI template was performed using each participant's T2 and T1 anatomical images as intermediate targets. No participants were excluded due to poor quality fMRI data.

^{1}H MRS data acquisition, processing and quantification

The imaging protocol for MRS data acquisition was constant for all participants throughout the study and was used on the same research scanner (Aim 2, 3, 4 + Validation Aim). A three-plane, low-resolution, high-speed scout imaging series was obtained, followed by a series of standardized high-resolution axial, coronal and sagittal T_1- and T_2-weighted scans to enable optimal placement of the ^{1}H MRS voxels of interest. Voxel placement was agreed upon by both an expert spectroscopist (CMM) and the principal investigator (KMD). For each participant, the pgACC voxel of interest was centered midsagittally, anterior to the genu of the corpus callosum (Aim 2) and the OCC voxel of interest (Validation Aim) was centered on the midline and rotated in the sagittal slice to align along the cerebellar tentorium and placed as posterior as possible without including the sagittal sinus or skull[14]. A board certified neuroradiologist reviewed the structural scans to rule out pathology: no abnormalities were identified.

GABA concentration uncertainties that exceeded a Cramer-Rao Lower Bound (CRLB) of 10% were to be considered poor quality and excluded from further analyses however CRLB was <10% for all participants so none were excluded from analysis. An outlier analysis was then completed for GABA+/Cr concentrations. One participant in the PPD group had a pgACC GABA+/Cr concentration +/- 3 standard deviations of the PPD group mean and was not

included in data analyses. The use of a GABA+/Cr ratio instead of absolute concentration is in alignment with past publications[15] and exhibits the best reproducibility among other methods[16].

Assessment of MRS voxel tissue heterogeneity

Structural MRI scans were analyzed using Statistical Parameter Mapping (SPM8-http://www.fil.ion.ucl.ac.uk/spn/software/spm8/) and white (WM) and gray matter (GM) and cerebrospinal fluid (CSF) in each MRS voxel were estimated using Matlab-based code provided by Drs. Nia Goulden and Paul Mullins of Bangor University and available at (https://www.bangor.ac.uk/psychology/biu/Wiki.php.en). This code generates a WM, GM and CSF image, a mask of the voxel location and the WM, GM and CSF percentages within the MRS voxel. To correct for partial volume effects and relaxation, we used the formula published in Gasparovic C et al. 2006 [17] (Aims 2, 3, 4 + Validation Aim).

Urinary benzodiazepine detection

To determine the presence of undisclosed benzodiazepine use, a urine sample was obtained at the time of each blood draw. Urinary benzodiazepine testing was performed since use could interfere with the interpretation of NAS, RSFC and MRS examinations. The urinary benzodiazepine drug test (Innovacon, Inc., San Diego, CA) is a lateral flow chromatographic immunoassay for the qualitative detection of oxazepam (major metabolite) with a cut-off concentration of 300ng/mL. Common benzodiazepines (e.g. alprazolam, clonazepam, diazepam, etc.) are detected with the assay (for rigor, all Aims).

Statistical Analysis

Magnetic resonance spectroscopy

As part of a post-hoc analysis, we examined correlations between OCC and pgACC GABA+/Cr across all participants and within groups and between groups.

Peripartum plasma neuroactive steroid concentrations and analyses of relationships with postpartum mood, RSFC and GABA MRS

We chose to examine allopregnanolone and its isomer pregnanolone based on our previous results in women at-risk for PPD[18]. We started by examining descriptive statistics and performing an exploratory data analysis on allopregnanolone and pregnanolone to identify outliers in the data[19]. We looked at antepartum (visits 1 and 2), delivery (visit 3), and postpartum (visits 4 and 5) time points separately as the range of NAS values were statistically and physiologically different at these visits. We observed outliers graphically with scatter plots of allopregnanolone and pregnanolone over time, identified values that were greater than +/- 3 standard deviations from the mean, and calculated Cook's D statistics in models run separately for peripartum time (antepartum, delivery, postpartum) and adjusted for the gestational age or postpartum time that the measure was taken. Cook's D was calculated with a cut-off of 4 divided by the number of observations in our data, to identify influential values, as has been suggested in previous literature[20,21]. Values that met all three conditions were excluded from our analysis.

Next, we centered visit time on delivery, with timing prior to delivery (visits 1 and 2) coded negatively and timing after delivery (visits 3, 4, and 5) coded positively. Using an autoregressive covariance structure, we used generalized estimating equation (GEE) methods to control for the repeated participants correlation at the five visit time points, similar to the methodology in our previous work[18]. All models were adjusted for the time (centered on delivery), neuroactive steroid concentration and the clinical data measured. *P* values are reported from a *z* test that a single regression coefficient was equal to 0 as well as from the

overall Type 3 test of any difference among levels of a factor (for HCW vs. PPD models), adjusting for other variables in the model. All results are reported with the conventional critical significance level of p=0.05. Analyses were conducted in SAS Version 9.4 (SAS Institute, Inc., Cary, North Carolina) (Aim 4).

We correlated average antepartum NAS and average postpartum NAS values against the functional connectivity maps for the DMPFC seed. Other predictors were group mean, age and total postpartum days at time of MRI. We did not attempt to create GEE models for these tests since connectivity data across the five study visits were lacking. We tested for correlations between the averaged allopregnanolone or pregnanolone concentration and pgACC and OCC GABA+/Cr concentrations.

RESULTS

Demographic and clinical data

Most women reported having completed or partially-completed college (67%) and were currently employed (80%). HCW were more likely to be married compared to women with PPD (75% vs. 48%; p=0.04). More than half of the participants were nulliparous at study entry (57%) compared to primiparous (35%) and multiparous (8%) women. Delivery mode for most births was vaginal (76%) with just over half of labors induced (55%). The majority of women reported full or partial breastfeeding (82%). Participants delivered at 39.3 (±1.3) weeks gestational age, on average, with mean infant birthweights of 3368.2 kg (±543.1). The postpartum MRI scan took place at 34.7 (±16.2) days after delivery, on average (all Aims).

Resting-state functional connectivity analyses

Of the 53 AR-PPD participants enrolled, 25 developed adjustment disorder with depressed mood or minor/major depressive disorder of which 23 completed the RSFC scan (Figure 1).

Two PPD RSFC scans were not completed due to women suffering from a panic attack after MRS data acquisition requiring scan cessation. Of the 35 HCW participants enrolled, 28 completed the RSFC scan with 7 participants either lost-to-follow up, withdrawn due to medical reasons or declining MRI.

^{1}H-MRS GABA+/Cr concentrations in the pgACC and OCC

OCC and pgACC GABA+/Cr concentrations were not correlated with across groups (r=+0.236, p= 0.099). OCC and pgACC GABA+/Cr concentrations were not correlated in either HCW (r=+0.338, *p*=0.078) or in women with PPD (r= -0.046, p=0.838) and the strength of the correlation did not differ by group (p=0.510). All post-hoc tests were uncorrected.

SUPPLEMENTAL TABLE 1. Peripartum Psychometric Scale Total Scores (mean ±SD)

	n	HAM-D_{17}	HAM-A	EPDS	STAI-S	SDS	PSQI
Antepartum Visit 1							
PPD	23	14.13 ± 6.11	18.13 ± 7.75	12.83 ± 3.96	44.22 ± 10.05	11.65 ± 6.06	24.35 ± 5.42
HCW	28	3.50 ± 2.71	6.07 ± 4.22	2.18 ± 2.51	24.82 ± 8.06	1.43 ± 3.37	11.36 ± 5.42
Antepartum Visit 2							
PPD	23	14.65 ± 7.00	16.65 ± 8.27	12.09 ± 4.70	44.83 ± 10.01	10.74 ± 6.60	23.22 ± 5.98
HCW	25	3.72 ± 3.27	5.68 ± 4.61	2.32 ± 2.25	24.80 ± 6.71	1.48 ± 3.85	11.56 ± 5.86
Postpartum Visit 3							
PPD	23	14.04 ± 4.69	15.13 ± 7.56	11.43 ± 4.38	41.52 ± 10.62	8.00 ± 5.58	23.43 ± 8.13
HCW	28	4.61 ± 3.58	5.68 ± 5.88	1.82 ± 2.83	26.04 ± 8.70	1.04 ± 3.56	11.36 ± 6.72
Postpartum Visit 4							
PPD	23	15.43 ± 6.18	17.52 ± 7.53	13.26 ± 3.63	47.74 ± 10.41	12.09 ± 6.05	23.35 ± 7.51
HCW	28	3.82 ± 2.92	4.04 ± 3.69	1.93 ± 2.58	25.71 ± 7.68	1.18 ± 2.87	10.18 ± 5.27
Postpartum Visit 5							
PPD	22	15.18 ± 7.97	15.55 ± 7.80	13.14 ± 5.43	46.00 ± 9.84	9.36 ± 4.86	19.41 ± 6.11
HCW	28	2.21 ± 2.11	2.61 ± 2.78	1.29 ± 1.61	23.14 ± 4.94	0.68 ± 1.52	8.71 ± 5.08

Data was collected across peripartum time where Visit 1 occurred during 22-38 weeks gestational age, Visit 2 occurred during 29-39 weeks gestational age, Visit 3 occurred within 3 days postpartum, Visit 4 during 2-7 weeks postpartum and Visit 5 during 4-11 weeks postpartum. Abbreviations: SD: Standard deviation; PPD: Peripartum depression; HCW: Healthy comparison women; HAM-D_{17}: Hamilton Depression Scale; HAM-A: Hamilton Anxiety Scale; EPDS: Edinburgh Postnatal Depression Scale; STAI-S: Spielberger State-Trait Anxiety Inventory; SDS: Sheehan Disability Scale; PSQI: Pittsburgh Sleep Quality Index

SUPPLEMENTAL TABLE 2: Regions where resting-state functional connectivity with the dorsomedial prefrontal cortex (DMPFC) seed region differs significantly by group.

Group Differences in DMPFC connectivity

Controls > Patients

Structure	Peak locations	Peak Coordinates	Voxels	Max. Z
Medial parietal	Precuneus/Post. Cingulate/Postcentral G.	(-14, -40, 44)	459	3.74
Lateral Parietal	Supramarginal G./Angular G.	(-62, -46, 24)	229	4.17
	Supramarginal G./Angular G.	(-52, -48, 40)		4.17
	Supramarginal G.	(-54, -36, 38)		3.42
	Supramarginal G./Angular G./Sup.Parietal L	(-42, -50, 46)		3.31

Patients > Controls: None

All coordinates in the Montreal Neurological Institute (MNI) atlas space. Peaks whose coordinates have greater than 50% probability of lying in white matter are not listed

Abbreviations: Post. = posterior; G. = gyrus; Sup. = superior; L= lobule

SUPPLEMENTAL TABLE 3: Brain regions where dorsomedial prefrontal cortex (DMPFC) functional connectivity is correlated with pregenual anterior cingulate cortex (pgACC) or occipital cortex (OCC) ^{1}H-MRS GABA+/Cr concentrations

Pregenual GABA correlations with DMPFC connectivity

Positive correlation

Structure	Peak locations	Peak Coordinates	Voxels	Max. Z
Sub-cortical	Insula	(42, 6, -12)	536	4.93
	Temporal Pole / Insula	(46, 12, -10)		4.5
	Insula / Planum temporale	(42, -14, -2)		4.42
	Temporal Pole/Planum temporale	(44, 6, -18)		4.05

Negative correlation: None

Occipital GABA correlations with DMPFC connectivity

Positive correlation

Structure	Peak locations	Peak Coordinates	Voxels	Max. Z
Lateral parietal	Lateral Occip.G./Angular G.	(-40, -60, 42)	831	4.39
	Superior Parietal L.	(-30, -52, 54)		4.35
	Angular G./Supramarginal G.	(-42, -56, 44)		4.14
Temporal pole	Orbitofrontal Ctx	(-20, 10, -18)	744	4.29
	Left Amygdala/Hippocampus	(-16, -12, -14)		4.23
	Left Amygdala	(-18, -8, -16)		4.13
	Left Amygdala	(-30, -4, -20)		4.09
Temporal pole	Temporal Pole	(32, 12, -32)	263	4.07
	Right Amygdala/Parahipp. G.	(30, 4, -22)		3.81
	Right Amygdala/Parahipp. G.	(22, 4, -22)		3.66
	Right Amygdala	(32, 0, -18)		3.52
	Temporal Pole	(36, 22, -32)		3.5
	Insula/Orbitofrontal Ctx/Temporal Pole	(34, 6, -16)		3.49

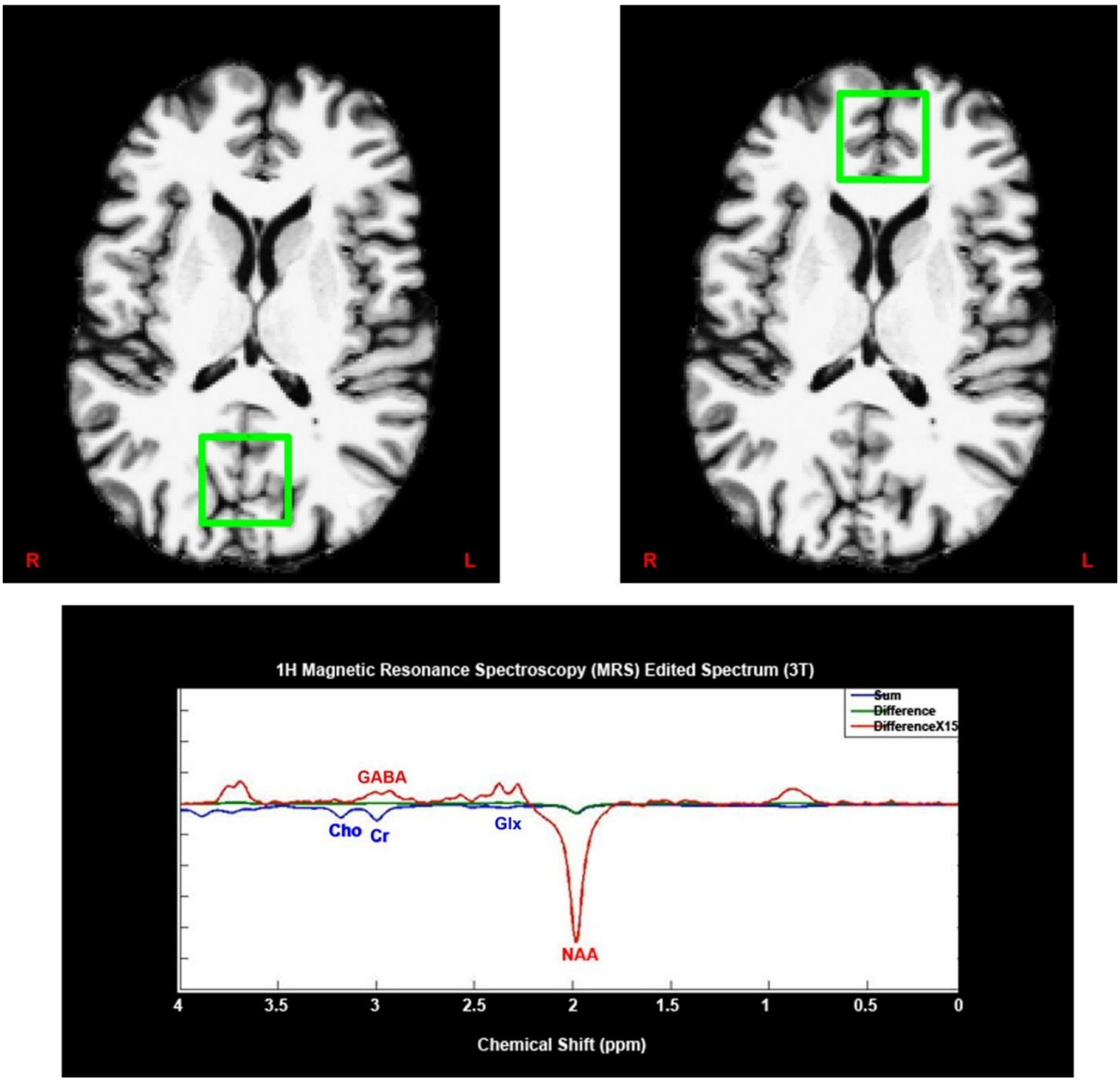
Negative correlation

Medial parietal	Supplementary Motor A/Anterior Cingulate	(10, -2, 48)	614	4.45

All coordinates in the Montreal Neurological Institute (MNI) atlas space. Peaks whose coordinates have greater than 50% probability of lying in white matter are not listed.

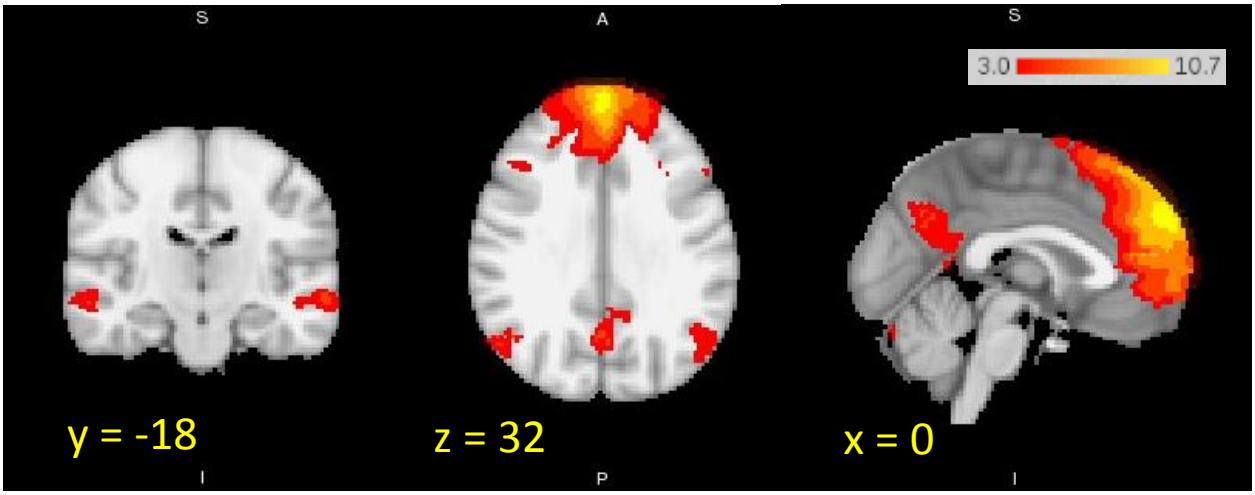
Abbreviations: Occip= occipital; G. = gyrus; L. = lobe; Ctx. = cortex; Parahipp. = parahippocampal; A. = area

SUPPLEMENTAL FIGURE 1: Proton magnetic resonance spectroscopy (^{1}H-MRS) voxel localization and ^{1}H spectra.



Upper part illustrates voxel localization for the occipital cortex (left) and pregenual anterior cingulate cortex (right) superimposed onto a slice from the anatomical scan of a healthy comparison woman. The lower part is a sample representation of an edited ^{1}H spectrum for a healthy comparison woman. The blue line indicates the unedited ^{1}H spectrum, the green line indicates the edited ^{1}H spectrum and the red line indicates an amplified version of the green line.

SUPPLEMENTAL FIGURE 2: Mean correlation, across all participants, of the dorsomedial prefrontal cortex (DMPFC) seed. Significant functional coupling is displayed by voxels ranging from red to yellow for positive correlations.



SUPPLEMENTAL MATERIAL REFERENCES

1. First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders - Patient Edition. New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 2001.
2. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry.* 1987;150:782-786.
3. Lydsdottir LB, Howard LM, Olafsdottir H, Thome M, Tyrfingsson P, Sigurdsson JF. The mental health characteristics of pregnant women with depressive symptoms identified by the Edinburgh Postnatal Depression Scale. *J Clin Psychiatry.* 2014;75(4):393-398.
4. Meijer JL, Beijers C, van Pampus MG, Verbeek T, Stolk RP, Milgrom J, et al. Predictive accuracy of Edinburgh postnatal depression scale assessment during pregnancy for the risk of developing postpartum depressive symptoms: a prospective cohort study. *BJOG : an international journal of obstetrics and gynaecology.* 2014;121(13):1604-1610.
5. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics.* 2009;42(2):377-381.
6. Williams JB. A structured interview guide for the Hamilton Depression Rating Scale. *Arch Gen Psychiatry.* 1988;45(8):742-747.
7. Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol.* 1959;32(1):50-55.
8. Shear MK, Vander Bilt J, Rucci P, Endicott J, Lydiard B, Otto MW, et al. Reliability and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A). *Depress Anxiety.* 2001;13(4):166-178.

9. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. State Trait Anxiety Inventory. 1970; www.mindgarden.com. Accessed February 25, 2011.
10. Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 1989;28(2):193-213.
11. Sheehan DV. *The Anxiety Disease.* New York: Charles Scribner & Sons; 1983.
12. Cuthbert BN. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry.* 2014;13(1):28-35.
13. Patriat R, Molloy EK, Meier TB, Kirk GR, Nair VA, Meyerand ME, et al. The effect of resting condition on resting-state fMRI reliability and consistency: a comparison between resting with eyes open, closed, and fixated. *NeuroImage.* 2013;78:463-473.
14. Epperson CN, O'Malley S, Czarkowski KA, Gueorguieva R, Jatlow P, Sanacora G, et al. Sex, GABA, and nicotine: the impact of smoking on cortical GABA levels across the menstrual cycle as measured with proton magnetic resonance spectroscopy. *Biol Psychiatry.* 2005;57(1):44-48.
15. Brennan BP, Admon R, Perriello C, LaFlamme EM, Athey AJ, Pizzagalli DA, et al. Acute change in anterior cingulate cortex GABA, but not glutamine/glutamate, mediates antidepressant response to citalopram. *Psychiatry Res.* 2017;269:9-16.
16. Bogner W, Gruber S, Doelken M, Stadlbauer A, Ganslandt O, Boettcher U, et al. In vivo quantification of intracerebral GABA by single-voxel (1)H-MRS-How reproducible are the results? *European journal of radiology.* 2010;73(3):526-531.
17. Gasparovic C, Song T, Devier D, Bockholt HJ, Caprihan A, Mullins PG, et al. Use of tissue water as a concentration reference for proton spectroscopic imaging. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine.* 2006;55(6):1219-1226.

18. Deligiannidis KM, Kroll-Desrosiers AR, Mo S, Nguyen HP, Svenson A, Jaitly N, et al. Peripartum neuroactive steroid and gamma-aminobutyric acid profiles in women at-risk for postpartum depression. *Psychoneuroendocrinology.* 2016;70:98-107.

19. Seo S. *A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets*. http://d-scholarship.pitt.edu/7948/1/Seo.pdf: Graduate School of Public Health, University of Pittsburgh; 2006.

20. Cook RD. Detection of Influential Observations in Linear Regression. *Technometrics.* 1977;19(1):15-18.

21. Bollen K, RW J. *Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. Modern Methods of Data Analysis.* Newbury Park, CA: Sage; 1990.