Mediating Effects of Neural Targets on Depression, Weight and Anxiety Outcomes of an Integrated Collaborative Care Intervention: The ENGAGE-2 Mechanistic Pilot Randomized Controlled Trial

Supplement

Table S1. Reasons for Ineligibility

 a Includes weight over 350 pounds, traumatic brain injury, tumor or any other known structural abnormality in brain, bullet, shrapnel, or other projectile above the shoulder, not being able to lie down in an fMRI scanner for about an hour due to claustrophobia, personal history of epilepsy, convulsions, or seizures, have piercings that cannot be removed.

b Includes stage 4 or greater renal disease, liver failure, cancer (other than non-melanoma skin cancer) that is/was active or treated with radiation or chemotherapy within the past year.

Supplementary Methods

1. Supplementary Functional Neuroimaging Methods

Viewing of Facial Expressions Task

A standardized set of 3D evoked facial expression stimuli were presented in pseudorandom order, with 5 repeated blocks of 8 stimuli per block for sad, fear, anger, and happy relative to neutral blocks (1). Threat stimuli is the combination of fear and anger stimuli relative to neutral blocks. During the conscious viewing condition, each face was presented for 500 ms, with an interstimulus interval of 750 ms. We created a context for participants to continuously view the faces by instructing them that they would be asked post-scan questions about these faces. To elicit the negative affect circuit in response to non-conscious threat stimuli, we presented the same fear and anger stimuli in a backward-masking design to prevent awareness. In this nonconscious condition, face stimuli were presented for 10ms followed immediately by a neutral face mask stimulus for 150 ms, and with a stimulus onset asynchrony of 1250 ms to match that of the conscious condition (2).

Imaging Sequences

BOLD contrast functional images were acquired with echo-planar T2*-weighted imaging using a GE MR750 3T scanner (GE Healthcare, Milwaukee, Wisconsin) with a NOVA 32-channel head coil. Head motion was restricted with foam pads.

Each whole brain volume consisted of 45 interleaved 3mm thick axial/oblique slices (74 x 74 matrix; TR=2000ms; TE=27.5ms; voxel size=3x3x3mm; FOV=222mm; flip angle=77°). One hundred fifty-four volumes were acquired over 5 minutes and 8 seconds for both tasks. A highresolution T1-weighted structural scan was acquired using GE's BRAVO sequence at the end of the imaging session for use in normalization of the fMRI data into standard space with the following parameters: TR=0.008, TE=0.003; voxel size=1x1x1mm; number of slices=176; FOV=256x256; flip angle=11 $^{\circ}$.

Image Pre-processing

Pre-processing and data analysis were performed using Statistical Parametric Mapping (SPM) software implemented in Matlab (SPM8; Wellcome Department of Cognitive Neurology) and the FSL (3) in a manner similar to that of our prior publications (2, 4). Briefly, motion correction was performed by realigning and unwarping the fMRI images to the first image of each task run after removal of the three dummy scans acquired at the start of the scanning session. Images were normalized to the stereotactic space of the Montreal Neurological Institute template (5). T1 weighted data were normalized to standard space using the FMRIB nonlinear registration tool, and the functional echo-planar image data were co-registered to the T1 data using the FMRIB linear registration tool. Prior to computing brain activation values, physiological noise was estimated using the time series from an eroded mask within the ventricles and white matter and was removed from the motion-corrected fMRI time series. Functional data were then smoothed using an 8 mm Gaussian kernel and high-pass filtered using a cutoff period of 128 seconds.

Following realignment and unwarping, quality control diagnostics were completed on the time series data for each run. Quality control diagnostics included removing scans with incidental findings, scanner artefacts and signal dropout. Participants' data were included if no more than 25% (38/151) of time points were censored for frame-wise displacement or variance spikes.

This resulted in total of $n = 82$ and $n = 59$ for the baseline and 2-month imaging sessions respectively.

Defining regions of interest

Our regions of interest for the negative affect circuit engaged by threat and sad were defined in our protocol (6-8) and pre-planned analytic plan was established in a prior systematic procedure validated with the same facial emotion task as used in the present ENGAGE-2 trial (9). Primary target regions of interest were the subgenual ACC (sgACC) and amygdala (bilaterally) for threat and the pregenual ACC (pgACC), amygdala (bilaterally), and anterior insula (bilaterally) for sad. Functional connectivity between ROIs and a global circuit dysfunction score for negative affect circuit engaged by threat in the non-conscious viewing condition and sad in the conscious viewing condition were also computed as the secondary neural targets. Other secondary neural targets included ROIs for the cognitive control circuit using the go-no go task, the default mode circuit, the negative affect circuit engaged by threat faces in the conscious viewing condition, and the positive affect circuit engaged by happy faces in the conscious viewing condition. These regions were defined *a priori* and not derived using a discovery analysis with the present ENGAGE-2 sample. Our *a priori* focus on these regions and pre-planned analytic strategy to test hypotheses, as outlined in the ENGAGE-2 protocol (8) was informed by our synthesis of the imaging findings for depression (6, 10) and prior trials in which imaging was included at the pretrial baseline to predict outcomes for both behavioural and pharmacological interventions. We have demonstrated that masks to define these *a priori* regions are reliably generated using the meta-analytic platform Neurosynth (11).

Specifically, the meta-analytic platform Neurosynth (11) with the search term "threat" was used to define the negative affective network. Analysis of Functional Neuroimaging's (AFNI's) 3dExtrema function was then used to identify peaks corresponding to our a priori regions of interest. Because some terms yielded maps with excessively large spatial extent, we imposed a restriction that each peak have a minimum z-score of 6 and each region extend no farther than 10mm from the peak. For the amygdala, Neurosynth maps were intersected with anatomically defined boundaries from the Automated Anatomical Labeling atlas (12). Finally, all ROIs were intersected with each individual's grey matter mask. Thus, each ROI was specific to the gray matter anatomy of each individual.

Procedures for Quantifying Neural Circuit Mediator Targets

Quantification of activation and connectivity for these ROIs also followed our previously established systematic procedure and incorporated a sample of 50 healthy individuals without depression or obesity (mean age 32.48 years, SD 11.95, 56% female, 54% non-Hispanic White, mean BMI 23.52, SD 3.32, and mean PHQ-9 0.84, SD 1.78) (9), as outlined in the following sections for both the healthy reference sample and the primary sample.

Computing Circuit Function for the Healthy Reference Sample

As was done with activation, region-to-region connectivity for ROIs was first quantified for the healthy reference sample available to this study. Connectivity between ROIs was quantified using psychophysiological interaction (PPI) analysis. For each ROI (used in PPI as a seed region), we calculated the first eigenvariate of that ROI's time course, and deconvolved this based on task events. Finally, we conducted a first-level general linear model consisting of the psychological variable (task contrast of interest), the physiological variable (deconvolved time course of the seed ROI), and the interaction between the psychological and physiological variables (PPI effect of interest). This process was repeated across all voxels and task contrast, yielding estimates of the contrast-dependent connectivity for each seed region. Because PPIbased connectivity estimates can differ slightly based on the ROI seed, we computed region-toregion PPI estimates using each region in the pair as a seed and averaging the results, yielding a single input for each connection.

Resulting activation and connectivity values were mean-centered and scaled to be expressed as standard deviation units. We defined neural circuit function both by individual component values (i.e. regional activation and region-to-region connectivity from each region of interest) as well as by global circuit scores (computed by averaging the constituent activation and connectivity component values for the each circuit). These data served as a healthy reference standard for computing extent of neural circuit dysfunction in the clinical participants.

Computing Circuit Dysfunction for the Present Primary Sample

For each individual participant in the primary sample, we computed activation and connectivity for the ROIs established using our prior systematic procedures. We expressed the extent of dysfunction in these values in terms of standard deviation units referenced to the mean of the healthy reference sample. This process resulted in values for each participant that quantified circuit dysfunction in each region of interest and region-to-region connectivity as well as global circuit dysfunction scores reflecting the average of these values. Through this procedure, global circuit dysfunction scores were interpretable relative to a healthy reference mean of zero. The direction of each regional input to the global circuit dysfunction score was oriented so that greater scores indicated greater dysfunction according to our theoretical framework (9).

In addition, the activation of bilateral regions of interest were significantly and strongly correlated for the negative affect circuit engaged by threat (non-conscious and conscious) and sad (conscious), cognitive control circuit, and positive affect circuit engaged by happy (conscious). These findings suggest a strong level of internal consistency between left and right-sided regions of interest at both baseline and 2-months follow up (see Table SS1).

Table SS1. Cross hemisphere consistency of bilateral regions of interest.

 $CI =$ confidence interval. dlPFC: Dorsal Lateral Prefrontal Cortex. $P = P$ value at an uncorrected threshold of .05.

Comparison with Healthy Reference Sample

At baseline, ENGAGE-2 participants showed elevated activity of dACC in the cognitive control circuit (ES=0.47, 95% CI 0.08 to 0.86) and reduced connectivity of multiple neural targets in the default mode circuit (Medial amPFC to Left AG: ES=-0.43, 95% CI -0.78 to -0.07; Medial amPFC to Right AG: ES=-0.51, 95% CI -0.85 to -0.17; Medial PCC to Medial amPFC: ES=- 1.00, 95% CI -1.37 to -0.64) and negative affect circuit engaged by conscious threat (Medial dACC to Left Amygdala: ES=-0.30, 95% CI -0.60 to -0.01). At 2 months, reduced connectivity of some neural targets in the default mode circuit persisted in the intervention and usual care group (See Table SS2 below).

Table SS2. Comparison between ENGAGE-2 participants and healthy controls at baseline and 2 months, for primary and secondary neural circuit targets.

^a Single letter indicates task activation; paired letters indicate task-related connectivity

b Represents the initial 2-month intervention phase of the I-CARE2 program that implemented a 7-step problem-solving process as its core component

^c Circuit dysfunction score (9)

Abbreviations: AG: Angular Gyrus; amPFC: anterior Medial Prefrontal Cortex; CI = confidence interval; dACC: Dorsal Anterior Cingulate Cortex; dlPFC: Dorsal Lateral Prefrontal Cortex; ES = standardized effect size; Hemi.: hemisphere; L: left; M: medial; PCC: Posterior Cingulate Cortex; pgACC: pregenual anterior cingulate cortex; R: right; sgACC: subgenual anterior cingulate cortex; vMPFC: ventral medial Prefrontal Cortex; vStriatum: ventral Striatum. P= P value at an uncorrected threshold of .05 prior to adjustment for FDR; P_{adj} = P value adjusted for FDR within neural target family (see eAppendix 5).

2. Sample Size Calculation

The focus of this mechanistic pilot trial was on the magnitude and precision (95% CI) of estimates for changes in neural targets in relation to clinical outcomes, with the goal of generating strong hypotheses about specific neural targets as causal effect mediators. Accordingly, the sample size was determined to focus on medium or larger effects given the likely limited clinical relevance of small effects. When considering all combinations of treatmentto-mediator and mediator-to-outcome effects in a suitable effect size metric by Fritz and MacKinnon (13), S=0.14 is small (akin to Cohen's d=0.20), H=0.26 is halfway between small and medium (d=0.35), $M=0.39$ is medium (d=0.50), and L=0.59 is large (d=0.80). Thus, effect size combinations of MM, ML, and LL for the joint mediation were in the medium to large range, which we considered reasonable given our prior findings (14, 15). Accordingly, assuming ≥85% retention over 6 months, a sample size of 105 (70 intervention, 35 control) was chosen as it would bound the 95% CI with a 2-sided standardized half-width of 0.50 (16, 17).

3. Changes in Response to the COVID-19 Pandemic

Recruitment and baseline data collection were not impacted by the pandemic. In-person visits at follow-up were suspended on 3/16/2020 and restarted on 7/10/2020; final data collection ended on 8/31/2020 (Figure S2). After 3/16/2020, delivery of the intervention sessions was changed from in-person to phone or Zoom videoconference.

References:

1. Korgaonkar MS, Grieve SM, Etkin A, Koslow SH, Williams LM (2013): Using standardized fMRI protocols to identify patterns of prefrontal circuit dysregulation that are common and specific to cognitive and emotional tasks in major depressive disorder: first wave results from the iSPOT-D study. *Neuropsychopharmacology*. 38:863-871.

2. Williams LM, Korgaonkar MS, Song YC, Paton R, Eagles S, Goldstein-Piekarski A, et al. (2015): Amygdala Reactivity to Emotional Faces in the Prediction of General and Medication-Specific Responses to Antidepressant Treatment in the Randomized iSPOT-D Trial. *Neuropsychopharmacology*. 40:2398-2408.

3. Tziortzi AC, Searle GE, Tzimopoulou S, Salinas C, Beaver JD, Jenkinson M, et al. (2011): Imaging dopamine receptors in humans with [11C]-(+)-PHNO: dissection of D3 signal and anatomy. *NeuroImage*. 54:264-277.

4. Goldstein-Piekarski AN, Korgaonkar MS, Green E, Suppes T, Schatzberg AF, Hastie T, et al. (2016): Human amygdala engagement moderated by early life stress exposure is a biobehavioral target for predicting recovery on antidepressants. *Proc Natl Acad Sci U S A*. 113:11955-11960.

5. Ashburner J, Friston KJ (2005): Unified segmentation. *NeuroImage*. 26:839-851.

6. Williams LM (2017): Defining biotypes for depression and anxiety based on large-scale circuit dysfunction: a theoretical review of the evidence and future directions for clinical translation. *Depress Anxiety*. 34:9-24.

7. Williams LM, Pines A, Goldstein-Piekarski AN, Rosas LG, Kullar M, Sacchet MD, et al. (2018): The ENGAGE study: Integrating neuroimaging, virtual reality and smartphone sensing to understand self-regulation for managing depression and obesity in a precision medicine model. *Behav Res Ther*. 101:58-70.

8. Lv N, Ajilore OA, Ronneberg CR, Venditti EM, Snowden MB, Lavori PW, et al. (2020): The ENGAGE-2 study: Engaging self-regulation targets to understand the mechanisms of behavior change and improve mood and weight outcomes in a randomized controlled trial (Phase 2). *Contemp Clin Trials*. 95:106072.

9. Goldstein-Piekarski AN, Ball TM, Samara Z, Staveland BR, Keller AS, Fleming SL, et al. (2022): Mapping Neural Circuit Biotypes to Symptoms and Behavioral Dimensions of Depression and Anxiety. *Biol Psychiatry*. 91:561-571.

10. Williams LM (2016): Precision psychiatry: a neural circuit taxonomy for depression and anxiety. *Lancet Psychiatry*. 3:472-480.

11. Yarkoni T, Poldrack R, Nichols T, Van Essen D, Wager T (2011): NeuroSynth: a new platform for large-scale automated synthesis of human functional neuroimaging data. *Frontiers in Neuroinformatics Conference Abstract: 4th INCF Congress of Neuroinformatics*, pp doi: 10.3389/conf.fninf.2011.3308.00058.

12. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*. 15:273-289.

13. Fritz MS, Mackinnon DP (2007): Required sample size to detect the mediated effect. *Psychol Sci*. 18:233-239.

14. Goldstein-Piekarski AN, Wielgosz J, Xiao L, Stetz P, Correa CG, Chang SE, et al. (2021): Early changes in neural circuit function engaged by negative emotion and modified by behavioural intervention are associated with depression and problem-solving outcomes: A report from the ENGAGE randomized controlled trial. *EBioMedicine*. 67:103387.

15. Lv N, Lefferts WK, Xiao L, Goldstein-Piekarski AN, Wielgosz J, Lavori PW, et al. (2021): Problem-solving therapy-induced amygdala engagement mediates lifestyle behavior change in obesity with comorbid depression: a randomized proof-of-mechanism trial. *The American journal of clinical nutrition*.

16. Julious SA, Patterson SD (2004): Sample sizes for estimation in clinical research. *Pharmaceut Statist*. 3:213-215.

17. Julious SA (2004): Sample sizes for clinical trials with normal data. *Stat Med*. 23:1921- 1986.

Figure S1. Study Timeline

^a Student's t test was conducted to compare changes in clinical outcomes before and after 3/16 for overall sample, intervention, and usual care group separately. Linear mixed models were conducted to compare treatment effects before and after 3/16. The fixed effects of each model included baseline value of the outcome, randomization covariates, group (intervention or control), time point (2 or 6 months), group-by-time interaction, COVID indicator (an indicator of whether a participant's outcome was assessed before or after the 3/16 lockdown), and group-by-time-by COVID indicator interaction. The random effects accounted for repeated measures with an unstructured covariance matrix.

Figure S2. Conceptual Framework

^a Please refer to Table S3 below for the specification of primary and secondary neural targets.

Table S3. Families of tests used to determine 'm' value and control false discovery rate

^a Single letter indicates task activation; paired letters indicate task-related connectivity

b Circuit dysfunction score (9)

Abbreviations: AG: Angular Gyrus; amPFC: anterior Medial Prefrontal Cortex; dACC: Dorsal Anterior Cingulate Cortex; dlPFC: Dorsal Lateral Prefrontal Cortex; Hemi.: hemisphere; L: left; M: medial; PCC: Posterior Cingulate Cortex; pgACC: pregenual anterior cingulate cortex; R: right; sgACC: subgenual anterior cingulate cortex; vMPFC: ventral medial Prefrontal Cortex; vStriatum: ventral Striatum.

^a Depression treatment response is defined as ≥50% decrease in SCL-20 scores from baseline.

^b Depression remission is defined as SCL-20 scores<0.5.

^c Anxiety treatment response is defined as ≥50% decrease in GAD-7 scores from baseline.

^d Anxiety remission is defined as GAD-7 scores<5.

Figure S4. Intervention Effects on Outcomes at 2 Months, Overall and by Subgroup.

Table S4. Association of changes in neural targets at 2 months and changes in outcomes at 2 months

Abbreviations: AG: Angular Gyrus; amPFC: anterior Medial Prefrontal Cortex; BMI, body mass index; dACC: Dorsal Anterior Cingulate Cortex; dlPFC: Dorsal Lateral Prefrontal Cortex; GAD-7, 7-item Generalized Anxiety Disorder Scale; Hemi.: hemisphere; L: left; M: medial; PCC: Posterior Cingulate Cortex; pgACC: pregenual anterior cingulate cortex; R: right; SCL-20, Depression Symptom Checklist-20; sgACC: subgenual anterior cingulate cortex; vMPFC: ventral medial Prefrontal Cortex; vStriatum: ventral Striatum.

a Ordinary least square regression model including baseline of outcome, indicator of the outcome data collected before or after COVID-19 shut down at study site (3/16/2020), biotype, treatment, and interaction of biotype*treatment.

b Single letter indicates task activation; paired letters indicate task-related connectivity.

^c Global circuit dysfunction score, composite of primary and secondary neural targets.

Note: Bolded results indicate that 95% CIs do not include null.

Table S5. Comparison of baseline characteristics between ENGAGE and ENGAGE-2 sample

^a Single letter indicates task activation; paired letters indicate task-related connectivity.

^b Global circuit dysfunction score, composite of primary and secondary neural targets.