Pre-registration of Secondary Data Analysis

Based on template available at https://osf.io/zpfnb/; see also Weston et al. (2018)

Study Information

Title: Neuroimaging correlates and predictors of response to repeated-dose intravenous ketamine in PTSD

Authors: Agnes Norbury, Abigail Collins, James Murrough, Adriana Feder

Research Questions and Hypotheses/Estimates:

Background. Although so far primarily explored as a therapy for treatment-resistant depression (e.g. Zarate et al. 2006; Murrough et al. 2013), promising initial data shows that intravenous (IV) administration of the glutamate NMDA receptor antagonist ketamine may also improve symptoms of post-traumatic stress disorder (PTSD) - over and above its effects on experience of depression (Feder et al. 2014). It is hypothesised that ketamine administration results in increased synaptogenesis in the prefrontal cortex and hippocampus, which may alleviate symptoms of affective disorders via reversing characteristic structural deficits in these regions observed following chronic stress (Li et al. 2010, 2011; Duman et al. 2012). Previous studies suggest that improvement in depression scores following IV ketamine are associated with changes in neural connectivity metrics during both emotion perception and under task-free conditions (e.g. Murrough et al. 2015; Abdallah et al. 2017). This study will analyse a rich battery of neuroimaging data collected as part of a wider randomized clinical trial of repeated-dose IV ketamine for PTSD. Given the relatively small sample size, we will harness knowledge derived from previous functional imaging studies to define a set of candidate neuroimaging metrics that may *a priori* be expected to be related to both experience of PTSD symptoms and ketamine-related changes in neural function. We will use a statistical approach appropriate for such datasets (i.e., where the number of measures is greater than number of observations) in order to probe whether any of these candidate measures are robustly related to symptom change data following repeated-dose ketamine treatment.

Research Question 1. If changes in candidate neural processing indices are relevant to PTSD aetiology, then changes in function in these measures would be expected to accompany changes in PTSD symptom scores over the course of the trial. If these changes are specific to ketamine (as opposed to general treatment effects), then these changes should be of greater magnitude in individuals who received ketamine. If these changes are related specifically to PTSD symptoms, then these relationships will persist after accounting for any concomitant changes in depression symptom scores.

> ***Hypothesis 1.1***. Are changes in any of our predefined candidate brain imaging measures reliably related to changes in PTSD symptoms following treatment?
>
> > ***Statistical Test 1.1.1***. To test this, we will run a robust leave-one(subject)-out cross-validated regularized multiple regression model (elastic net regression), where the independent (target) variable is change in PTSD symptom score, and predictor variables are changes in candidate imaging measures, with potential

confounds regressed out (i.e. the residuals of a model where imaging change scores are regressed against potential confounding covariates). Candidate imaging measures consist of an *a priori* set of task-free functional connectivity, anatomical connectivity, and task-based emotional-processing metrics (for a full list see Current Study: Variables). Relevant subject-level covariates will also be included in the model, including participant age, gender identity, and mean drug side effects score (for infusion sessions vs baseline).

***Hypothesis 1.2*.** If changes are specific to drug treatment, then we would expect to observe a significant factor*treatment interaction when drug treatment received is included in the model.

Statistical Test 1.2.1. Re-run model 1.1.1, with a between-subjects factor of drug.

Hypothesis 1.3. If changes are specific to improvements in PTSD, rather than depression symptom levels, then these predictors should remain significant when change in depression score is added to the model.

Statistical Test 1.3.1. Re-run model as in 1.2.1, but with individual changes in depression symptom scores added to the model as a predictor.

Research Question 2. As individuals with a diagnosis of PTSD are heterogeneous in terms of symptom experience, with likely variation in underlying pathological mechanisms, it is reasonable to expect that ketamine therapy may be more effective for some individuals than others. We will therefore examine whether candidate predictors of treatment response (baseline neuroimaging and other measures) are related to change in symptom scores pre *vs* post receiving the repeated-dose treatment regime. Again, we will then examine whether any such predictors are specific to individuals who actually received ketamine, and to PTSD (*cf* depression) symptom improvement.

Hypothesis 2.1. Are there baseline individual differences which reliably predict whether an individual with a diagnosis of PTSD may respond to repeated-dose drug therapy?

Statistical test 2.1.1. We will examine whether any potential moderators of ketamine effectiveness measured at baseline predict changes in PTSD score at follow up, in a similar model to Research Question 1, but using baseline imaging measures rather than change scores as predictors. Other potential response predicting variables available at baseline will also be included in the model, specifically: general executive function test scores, history of alcohol use disorder in first degree relatives, perceived degree of social support, and average intensity of dissociative symptoms experienced at 40 minutes post- first infusion (see Current Study: Variables for full details and supporting literature).

Hypothesis 2.2. Is this predictive relationship greater in individuals who actually received ketamine?

Statistical test 2.2.1. As 2.1.1, but with drug treatment group added to the model.

Hypothesis 2.3. Are these predictors specific to PTSD symptom (*cf* depression symptom) improvement?

Statistical test 2.3.1. As 2.2.1, but with change in depression added to the model as a predictor/covariate.

Data Description

Brief description of data set(s): Imaging data were acquired from a subset of participants of a clinical trial of repeated-dose IV ketamine for PTSD (GCO 15-0265, see https://clinicaltrials.gov/ct2/show/NCT02397889).

Is this data open or publically available? Currently, No (clinical trial is still ongoing)

How can the data be accessed? Provide link if available online: Currently N/A

Date of download or access: Initial access to raw functional imaging data was granted to to the primary analyst (AN) on 09/07/18. Access to relevant clinical data (demographic information and clinical measure scores) and behavioural data (for functional scans) was granted to AN on 10/04/18. At this point, AN *only* was partially unblinded to drug condition by the trial pharmacist (i.e. conditions were labelled as Drug A or Drug B, but not as ketamine or midazolam). Data were collected by research staff under the supervision of the clinical trial Research Coordinator (AC) and the Principal Investigator (AF), who remained fully blinded throughout.

Data Source: Own Lab Collection - Data were connected by one of the analysts' lab (but primary analyst was not involved in data collection)

Codebook: Not Applicable

Sampling and data collection procedures:
Recruitment for the clinical trial was primarily via media advertisements to the general public in New York City (print, online, radio). Patients were additionally recruited from clinical practices at Mount Sinai, including the World Trade Center Health Program and the Outpatient Psychiatry Department, as well as Veterans organizations and other community clinics (e.g. crime victims treatment programs).

Eligibility criteria for the clinical trial are described at clinicaltrials.gov/ct2/show/NCT02397889. Patients with a diagnosis of major depressive disorder were not excluded from the trial in order prevent limiting the generalizability of any findings (almost 50% of patients with PTSD have a comorbid diagnosis of depression). For similar reasons, a THC-positive urine screen on any of the drug administration days did not result in subsequent trial exclusion (it is relatively common for individuals with PTSD in the United States to self-medicate with marijuana).

Upon recruitment to the trial, participants were randomly allocated to a drug condition (ketamine or midazolam) by the trial pharmacist. Participants and all administering research staff were blind to drug conditions. Trial participants visited to complete initial screening tests, then, following a drug washout period (for counter-indicated substances: see clinicaltrials.gov/ct2/show/NCT02397889), received 6 IV doses of drug, at a rate of 3 a week for two weeks. Side-effects measures were collected during each of the drug infusion sessions, and

primary clinical outcome measures (PTSD and depression symptoms) at 24 hours, 1 and 2 weeks post first-infusion.

Participants for the imaging study were recruited from the wider clinical trial cohort on an *ad hoc* basis, based on their consent to take part and additional MRI eligibility criteria. Exclusion criteria for the MRI study were claustrophobia, any trauma or surgery which may have left magnetic material in the body, presence of magnetic implants or pacemakers, and inability to lie still for 1 hour or more.

Imaging data consisted of a battery of MRI measures (a T1-weighted structural scan, diffusion-weighted anatomical scan, and one task-free and two task-based T2*-weighted functional imaging scans) that were collected during a baseline session (at the end of the washout period, prior to administration of any drugs), and again during a post-infusion session (following the 4th drug infusion, approximately 1 week post first-infusion). The imaging data analyst was partially unblinded during MRI data analysis - i.e. had access to drug groupings, but not drug identity labels for each participant. No other researchers or study staff were partially unblinded at this point.

Knowledge of Data

Prior work based on the dataset: Never worked with this dataset

Prior research activity: I have never analysed these data before

Prior knowledge of current dataset: No prior knowledge

Moment of preregistration: Registration prior to any researcher on this team analysing the data

Current Study: Variables

Manipulated variables:
Participants were randomly allocated to one of two drug conditions, receiving multiple IV doses (3 infusions per week for two weeks) of either ketamine (0.5mg/kg) or the active control drug midazolam (0.045mg/kg).

Measured variables:
Demographic information:

- **Age** (years)
- **Gender identity**
- **Education level** (recorded as a categorical variable ranging from 1 [‘grade 6 or less’] to 8 [‘completed graduate/professional school’])
- **Income** (recorded as a categorical variable ranging from 1 [‘0-25K’] to 5 [‘249K+’])

Clinical measures:

- Clinician-Administered PTSD Scale for DSM-5 (**CAPS**-5; Weathers et al. 2013). The CAPS-5 is a structured clinical interview that is both able to diagnose presence/absence of PTSD according to DSM-5 criteria, and provide a quantitative measure of symptom presence and severity (a composite metric of frequency and intensity for each symptom, based on a 5 point scale ranging from 0 [‘Absent’] to 5 [‘Extreme/Incapacitating’]). CAPS-5 total (sum) score will be used for all confirmatory analyses listed here (scored as per Weather et al., 2013; i.e., the sum of severity scores for items 1-20). As part of the wider clinical trial, this was measured at baseline and at 1 and 2 weeks post-first infusion (*NB*, all clinician-rated measures of psychological symptoms were made by blinded raters, who were not present during the drug infusion sessions). For the purposes of this analysis, we will calculate symptom change scores as trial endpoint post-infusion score (2-weeks post first infusion exit visit) minus baseline score (pre-infusion at visit 1).

- Montgomery Asberg Depression Rating Scale (**MADRS**; Montgomery and Åsberg 1979). The MADRS is a 10-item instrument used for the evaluation of depressive symptoms via a structured clinical interview. 10 core symptoms of depression are rated on a scale of 0 to 6, the summing of which yields a total score. As part of the wider clinical trial, this was measured at baseline, 24 hours, 1 and 2 weeks post first-infusion. For the purposes of this analysis, we will calculate symptom change scores as trial endpoint post-infusion score (2-weeks post first infusion exit visit) minus baseline score (pre-infusion at visit 1).

- Current use of prescription medication (**MHQ_prescription**; binary, measured at screening visit and then held constant over course of the trial)

Neuroimaging measures:

Due to our relatively small sample size, we have decided to restrict our analyses to measures drawn from an *a priori* target circuit, defined on the basis of previous functional imaging studies in individuals with an diagnosis of PTSD. The most commonly implicated brain regions for BOLD signal differences in PTSD vs control groups (at rest and during 'emotional processing' tasks) are the ventromedial prefrontal cortex (vmPFC), rostral and dorsal anterior cingulate cortex (r and dACC), anterior insula cortex (a.insula), amygdala (AMG), and anterior hippocampus (aHC) (Pitman et al. 2012; Yehuda et al. 2015; Akiki, Averill, and Abdallah 2017; for commonly implicated interconnections between these brain regions, see list below). Interestingly, signal changes in or between several of these brain regions have been also identified in response to IV ketamine in both healthy individuals and patient groups (e.g. Grimm et al. 2015; Becker et al. 2017; J. W. Murrough et al. 2015; Abdallah et al. 2017).

- A **diffusion-weighted anatomical connectivity (dwMRI) scan**. Will be used to derive anatomical connectivity values between the following regions (our target network, 9 measures total):
    - vmPFC-AMG; vmPFC-a.insula; vmPFC-aHC
    - rACC-AMG; rACC-a.insula; rACC-aHC
    - dACC-AMG; AMG-a.insula; a.insula-aHC

- A T2*-weighted single echo **task-free (resting state) functional connectivity scan** (**rsfMRI**; 12 minutes, eyes open). Functional connectivity metrics will be extracted between the following regions (our target network, 9 measures total):
    - vmPFC-AMG; vmPFC-a.insula; vmPFC-aHC
    - rACC-AMG; rACC-a.insula; rACC-aHC
    - dACC-AMG; AMG-a.insula; a.insula-aHC

- Two T2*-weighted single echo **task-related functional MRI (fMRI) scans**:
  1) Emotional face processing task (Hariri et al. 2002) (11 measures):
    - mean signal for faces>shape matching contrast across the AMG, a.insula, vmPFC, and rACC
    - LDC multivariate similarity metric for the fearful *vs* happy faces contrast in the AMG and rACC (as there may be limited test-retest reliability for mean univariate signal in these ROIs on this task; Nord et al. 2017)
    - functional connectivity for face>shape matching contrast for the specified network connections between these regions (i.e. vmPFC-AMG; vmPFC-a.insula; rACC-AMG; rACC-a.insula; AMG-a.insula)

  2) Emotional interference (face Stroop) task (Offringa et al. 2013) (8 measures):
    - mean signal for the incongruent>congruent contrast in the rACC, dACC, and AMG
    - mean signal for the fear>happy incongruent contrast in the rACC, dACC, and AMG
    - functional connectivity for the incongruent>congruent contrast for rACC-AMG dACC-AMG

- A T1-weighted (**MPRAGE**) **structural image**. This image will be used in analysis of other signal modalities and also to extract gross intracranial volume (ICV).
- We will record mean framewise displacement (FD) across for each task (including rest) as an index of participant movement during the scan.

Other drug-related measures:

- The Patient Rated Inventory of Side Effects (**PRISE**; Rush et al. 2004) gives a measure of total general adverse side-effects across multiple somatic domains, which are sensitive to both IV ketamine and midazolam administration (Murrough et al., 2013; Feder et al., 2014).This scale is a 7 item assessment of the side effects in the following symptom areas; gastrointestinal, heart, skin, nervous system, eyes/ears, genital/urinary, sleep, sexual functioning, and other. Each domain has multiple symptoms and for each domain the patient rates whether these symptoms are tolerable or distressing. For each timepoint, presence of side effects will be summed across all domains, to yield a total score (PRISE total). Mean drug-related side effects will then be calculated as the average post infusion (+240 minutes) total score, minus the average baseline (pre-infusion) score.

- **YMRS-1**: a single item from the Young Mania Rating Scale (YMRS; Young et al. 1978). A clinician-rated index of participant elevated mood, previously shown to be sensitive to mood-related effects of ketamine (vs midazolam) at 40 minutes post-infusion (Murrough et al., 2013, Feder et al., 2014) - but, importantly, not to predict treatment response to ketamine (Luckenbaugh et al. 2014 - see below). Consists of a single item scoring “elevated mood’ from 0 (Absent) to 4 (Euphoric; inappropriate laughter; singing). Mean drug-related effects will be calculated as the average post infusion (+40 minutes) score, minus the average baseline (pre infusion) score.

- **BPRS-positive**: a 4 item measure of positive psychotic-like symptoms from the Brief Psychiatric Rating Scale (BPRS; Overall and Gorham, 1988), a clinician rated measure previously shown to be sensitive to psychotomimetic effects of ketamine (vs midazolam) at 40 minutes post-infusion (Murrough et al., 2013, Feder et al., 2014), but not to predict treatment response to ketamine (Luckenbaugh et al. 2014 - see below). The BPRS positive subscore consists of 4 items rated from 1 (not present) to 7 (extremely severe) that describe the presence of psychotic-like symptoms (suspiciousness, hallucinations, unusual thought content, and conceptual disorganisation). For each time point, total score will be calculated by summing across the 4 items. Mean drug-related effects will be calculated as the average post infusion (+40 minutes) score, minus the average baseline (pre infusion) score.

- The Clinician-Administered Dissociative States Scale (**CADSS**; Bremner et al. 1998). The CADSS is used to measure dissociative effects of drug administration; specifically impairment in body perception, environmental perception, time perception, memory impairment, and feelings of unreality. The scale includes 19 questions and 8 observer ratings scored from 0 (not at all) to 4 (extremely), and has previously been shown to be sensitive to both psychoactive effects induced by IV ketamine, and potential short-term

differences in magnitude of these effects observed under IV ketamine vs IV midazolam (and therefore potential blinding failure in both participants and observing clinicians/researchers; Murrough et al., 2013; Feder et al., 2014). CADSS was measured for the clinical trial on each infusion visit pre drug-administration and +40, +120 and +240 minutes post-infusion (previous studies suggest peak effect, and peak difference to midazolam control, is observed at +40 minutes post-infusion, *ibid*).

    - Interestingly, a meta-analysis of previous studies suggests that +40 min post-infusion change in CADSS dissociation score (but not blood pressure, pulse, or psychotomimetic effects) predicts depression symptom improvement at +24 hours (Luckenbaugh et al., 2014, *N*=108). It is therefore possible that degree of dissociative effects experienced during infusion is related to the treatment effect, rather than just effect on unblinding *per se*. We will therefore not include drug-related CADSS changes as a nuisance regressor in the symptom change data model for RQ1, but will include CADSS change during the first infusion (+40 minutes post-infusion 1 minus pre-infusion 1 score) as a quasi-'baseline' predictor for RQ2.

- Drug urine screen results for scan days (**THC_presence**; binary)

Additional baseline measures:

- **CogState** computerized test battery of executive/WM function: 1. Shopping List Test, 2. Groton Maze Learning Test, 3. Detection Task, 4. Identification Task, 5. One Card Learning Task, 6. One Back Task, 7. Delayed Recall Shopping List and Groton Maze Learning Tasks. As per the CogState analysis guide, performance across tasks on a primary index measure will be z scored across participants (for each task), then averaged (sum individual task scores/N tasks) to yield a composite score for executive function, across participants.

- Self-reported family history of alcohol use disorders in first degree relative (**FPH_ETOH**; has been previously been shown to relate to response to single dose IV ketamine in TRD, see Niciu et al. 2014; *N*=71-108). We will record *N* relatives reported (out of mother/father/sibling/child/grandparent).

- The Medical Outcomes Study Social Support Survey (**MOS-SS**; Sherbourne and Stewart 1991) is a 19-item self-report measure designed to assess perceived levels of functional social support. The MOS-SS has two subscales (emotional and instrumental social support) to identify potential social support deficits. We will use total score at baseline (pre-infusion visit 1).

Scales:

No custom aggregate scales will be used in this analysis. Please see 'Measured variables' for details of how previously published and validated instruments will be scored (and relevant references).

Indices:
Please see 'Measured variables: Other drug-related measures' for details of how information about treatment side-effects will be collated across infusion sessions and measures.

Transformations:
All self-reported predictor variables (demographic and questionnaire measures) will be feature-scaled by z-scoring (across subjects) before inclusion in the model.

Data Inclusion/exclusion:
Clinical measurements (for PTSD and depression symptom score changes) to be used in the analysis for RQ1 and RQ2 will be drawn from baseline (pre-infusion visit 1) and endpoint (exit visit, post infusion 6) study visits only. All participants who completed the full study (including both neuroimaging sessions) will be included in the analysis for RQ1. Additional participants who completed the clinical trial (all infusion and exit visits) and baseline imaging session but declined to or were unable to complete the post-infusion neuroimaging battery will also be included in the baseline prediction analysis (RQ2). Participants who tested positive for THC on drug neuroimaging visits will not be excluded from the analysis (as this is a small pilot study and self-medication with marijuana is fairly common in US PTSD patients), but drug urine screen status will be added to the analysis as a potential confounding variable.

Outliers:
We will exclude any measurements >3SD away from the mean for that measurement.

Weights
Not Applicable

Sample size:
Imaging data were acquired from a subset of participants of a larger clinical trial of repeated-dose IV ketamine for PTSD. Participants from this trial were not required to consent to the imaging sessions, and not all participants met criteria for the imaging phase, therefore participants were recruited into the add-on study on an *ad hoc* basis according to their wishes and eligibility. Further, some participants completed a baseline imaging session and the full drug infusion regime, but declined or were unable to complete a second (post-infusion) scan. The current sample size is therefore currently N=14 complete datasets (Research Question 1) and N=15 baseline-only imaging data with complete clinical outcomes (Research Question 2). As the parent clinical trial is currently nearing the end of recruitment, we will include any further participants who consent to take part in the imaging phase, but do not expect the number of final participants to rise significantly (final projected N<35).

This is a small sample size pilot study, as IV ketamine is still at the proof-of-concept stage as a therapy for PTSD. However, we believe this investigational status also provides a strong clinical motive for the analysis of this data, on which future (larger) studies may be based. A small sample size means that our analyses are likely to be underpowered for small-to-moderate size effects, although we should be adequately powered to detect moderate-to-large (potentially

most clinically relevant) effect sizes (for a simple bivariate relationship at *N*=15, we should be able to detect an *r* of 0.55 or above for alpha=0.05, one tailed, and 80% power - although *nb* regularization shrinks correlation coefficients to more conservative weight estimates). Further, we have attempted to guard against primarily fitting noise or reporting of spurious results via detailed pre-registration of our analysis plan, selection of literature-derived imaging measures previously related to our condition and treatment of interest, use of cross-validation and regularization terms in regression (prediction) models, and conducting the analysis blind to drug condition labels. Where possible (e.g. for drug side effects scales), scores were averaged across time points, in order to reduce measurement error.

Missing data:

- Face Stroop task data was unfortunately not collected for 3 participants (at *N*=15)
- CogState (baseline cognitive function) data was not collected for one individual due to a technical issue (at *N*=15)

Missing data will be dealt with using multiple imputation (this method replaces missing values with plausible simulated data which leave the mean and variance of the distribution unchanged).

Current Study: Analyses

Statistical models:

Research Question 1
Hypothesis 1.1
Model 1.1:

- Elastic net regularized multiple regression model with minimal mean square error as determined by leave-one(subject)-out cross-validation

Predictors:

- *Neuroimaging measures:* change in candidate measures between post-infusion and pre-infusion (baseline) visits for 9 rsfMRI, 9 dwMRI, 11 task-based fMRI (Hariri task), and 8 task-based fMRI (face Stroop task) measures. Each measure to have relevant confounds regressed out of individual session data prior to calculation of change scores (for rsfMRI, mean FD and *N* censored frames; for dwMRI, ICV; for task-based fMRI, mean FD)
- *Subject level covariates:* age, gender identity, mean drug side effects measures (YMRS-1, BPRS-positive, and PRISE scores), education level, prescription medication status, urine drug screen status

Target variable:

- Change in CAPS-5 score (exit minus baseline visit)

Hypothesis 1.2
Model 1.2:

- Model 1.1 with the additional between-subjects predictor variable of drug (drug A *vs* drug B)

Hypothesis 1.3
Model 1.3:

- Model 1.2 with the additional predictor variable (subject-level covariate) of change in MADRS score (exit minus baseline visit)

Research Question 2
Hypothesis 2.1
Model 2.1:

- Elastic net regularized multiple regression model with minimal mean square error as determined by leave-one(subject)-out cross-validation

Predictors:

- *Neuroimaging measures:* candidate measures from the pre-infusion (baseline) visit for 9 rsfMRI, 9 dwMRI, 11 task-based fMRI (Hariri task), and 8 task-based fMRI (face Stroop task) measures. Each measure to have relevant confounds regressed out of individual session data prior to entering into the model (for rsfMRI, mean FD and *N* censored frames; for dwMRI, ICV; for task-based fMRI, mean FD)

- *Other candidate predictor variables:* CogState total score (baseline visit), self-reported perceived social support (baseline visit), self-reported income (baseline visit), family history of alcohol use disorders (baseline visit), CADSS response on first infusion visit (+40mins post infusion minus pre-infusion score), baseline PTSD severity (CAPS-5 score at baseline visit)
- *Subject level covariates:* age, gender identity, mean drug side effects measures (YMRS-1, BPRS-positive, and PRISE scores), education level, prescription medication status, urine drug screen status

Target variable:

- Change in CAPS-5 score (exit minus baseline visit)

Hypothesis 2.2

Model 2.2:

- Model 2.1 with the additional between-subjects predictor variable of drug (drug A *vs* drug B)

Hypothesis 2.3

Model 2.3:

- Model 2.2 with the additional predictor variable (subject-level covariate) of change in MADRS score (exit minus baseline visit)

Follow-up analyses:

- For robustness, we will check for any significant differences in movement parameters (mean FD), and relevant task fMRI behavioural performance parameters (i.e. reaction times and response accuracy) between drug groups for the post-infusion session data. We will also check if any gross BOLD signal changes between baseline and post-infusion scan sessions in a control ROI differ between drug conditions.

- PTSD is not a unitary construct - PTSD symptoms on the CAPS-5 can be parsed into several subdimensions of PTSD symptomatology: specifically, intrusive (re-experiencing) symptoms, avoidance, negative alterations in cognition and mood, and arousal/reactivity symptoms (hypervigilance) (Weathers et al., 2013). Further, different dimensions of PTSD symptoms have previously been linked to different indices of neural function (e.g. Abdallah et al., 2017), and may be *a priori* expected to respond differentially to different forms of treatment. Planned follow-up analyses will therefore be carried out to determine the relationship between our various neuroimaging measures and individual CAPS-5 subscale scores, in the analyses listed for Research Questions 1 and 2.

Inference criteria:

For each hypothesis test we will fit a single multiple regression model to our data, where our criteria for reliable predictive value will be inclusion in the model that minimises mean square error (as determined by leave-one-subject out cross-validation), across all searched parameter values.

Sensitivity Analyses:
Please see the statistical model descriptions above for an explanation of how we will use both cross-validation and regularization within our multiple regression model, to maximise sensitivity, and minimise likelihood of overfitting our dataset.

Statistical Analysis Backup Plan:

- As we may be underpowered to detect anything smaller than moderate to large effects under our primary analysis plan, a lack of findings under this model will not allow us to draw any strong conclusions about the likelihood of the null hypothesis (i.e., that none of our candidate neuroimaging measures are related to symptom change scores). If our planned regularized regression models fail to detect any relationship between candidate markers and PTSD symptom change scores, we will therefore re-run them as normal multiple regressions, as a less robust test that may still furnish preliminary support for future studies.

Exploratory analysis:

- Under evidence that relationships between neuroimaging measures and symptom changes scores appears non-linear, non-linear regression analyses may be run as an exploratory test, for example using Gaussian Processes regression.

- If no candidate imaging markers show reliable relationships symptom change scores, we may also perform whole-brain analyses across various imaging modalities (with appropriate FDR multiple comparisons correction), both regression of symptom changes scores, and secondarily to examine differences between participants who received drug A vs drug B.

References

Abdallah, Chadi G., Lynnette A. Averill, Katherine A. Collins, Paul Geha, Jaclyn Schwartz, Christopher Averill, Kaitlin E. DeWilde, et al. 2017. “Ketamine Treatment and Global Brain Connectivity in Major Depression.” *Neuropsychopharmacology* 42 (6): 1210–19. https://doi.org/10.1038/npp.2016.186.

Akiki, Teddy J., Christopher L. Averill, and Chadi G. Abdallah. 2017. “A Network-Based Neurobiological Model of PTSD: Evidence From Structural and Functional Neuroimaging Studies.” *Current Psychiatry Reports* 19 (11): 81. https://doi.org/10.1007/s11920-017-0840-4.

Becker, Benjamin, Maria Steffens, Zhiying Zhao, Keith M. Kendrick, Claudia Neumann, Bernd Weber, Johannes Schultz, Mitul A. Mehta, Ulrich Ettinger, and Rene Hurlemann. 2017. “General and Emotion-Specific Neural Effects of Ketamine during Emotional Memory Formation.” *NeuroImage* 150 (April): 308–17. https://doi.org/10.1016/j.neuroimage.2017.02.049.

Bremner, J. Douglas, John H. Krystal, Frank W. Putnam, Steven M. Southwick, Charles Marmar, Dennis S. Charney, and Carolyn M. Mazure. 1998. “Measurement of Dissociative States with the Clinician-Administered Dissociative States Scale (CADSS).” *Journal of Traumatic Stress* 11 (1): 125–36. https://doi.org/10.1023/A:1024465317902.

Duman, Ronald S., Nanxin Li, Rong-Jian Liu, Vanja Duric, and George Aghajanian. 2012. “Signaling Pathways Underlying the Rapid Antidepressant Actions of Ketamine.” *Neuropharmacology*, Anxiety and Depression, 62 (1): 35–41. https://doi.org/10.1016/j.neuropharm.2011.08.044.

Feder, Adriana, Michael K. Parides, James W. Murrough, Andrew M. Perez, Julia E. Morgan, Shireen Saxena, Katherine Kirkwood, et al. 2014. “Efficacy of Intravenous Ketamine for Treatment of Chronic Posttraumatic Stress Disorder: A Randomized Clinical Trial.” *JAMA Psychiatry* 71 (6): 681–88. https://doi.org/10.1001/jamapsychiatry.2014.62.

Grimm, Oliver, Natalia Gass, Wolfgang Weber-Fahr, Alexander Sartorius, Esther Schenker, Michael Spedding, Celine Risterucci, et al. 2015. “Acute Ketamine Challenge Increases Resting State Prefrontal-Hippocampal Connectivity in Both Humans and Rats.” *Psychopharmacology* 232 (21): 4231–41. https://doi.org/10.1007/s00213-015-4022-y.

Hariri, Ahmad R., Alessandro Tessitore, Venkata S. Mattay, Francesco Fera, and Daniel R. Weinberger. 2002. “The Amygdala Response to Emotional Stimuli: A Comparison of Faces and Scenes.” *NeuroImage* 17 (1): 317–23. https://doi.org/10.1006/nimg.2002.1179.

Li, Nanxin, Boyoung Lee, Rong-Jian Liu, Mounira Banasr, Jason M. Dwyer, Masaaki Iwata, Xiao-Yuan Li, George Aghajanian, and Ronald S. Duman. 2010. “MTOR-Dependent Synapse Formation Underlies the Rapid Antidepressant Effects of NMDA Antagonists.” *Science* 329 (5994): 959–64. https://doi.org/10.1126/science.1190287.

Li, Nanxin, Rong-Jian Liu, Jason M. Dwyer, Mounira Banasr, Boyoung Lee, Hyeon Son, Xiao-Yuan Li, George Aghajanian, and Ronald S. Duman. 2011. “Glutamate N-Methyl-D-Aspartate Receptor Antagonists Rapidly Reverse Behavioral and Synaptic Deficits Caused by Chronic Stress Exposure.” *Biological Psychiatry*, Serotonin and Depression - Revisited, 69 (8): 754–61. https://doi.org/10.1016/j.biopsych.2010.12.015.

Luckenbaugh, David A., Mark J. Niciu, Dawn F. Ionescu, Neal M. Nolan, Erica M. Richards, Nancy E. Brutsche, Sara Guevara, and Carlos A. Zarate. 2014. “Do the Dissociative Side Effects of Ketamine Mediate Its Antidepressant Effects?” *Journal of Affective Disorders* 159 (April): 56–61. https://doi.org/10.1016/j.jad.2014.02.017.

Montgomery, Stuart A., and Marie Åsberg. 1979. “A New Depression Scale Designed to Be Sensitive to Change.” *The British Journal of Psychiatry* 134 (4): 382–89. https://doi.org/10.1192/bjp.134.4.382.

Murrough, J. W., K. A. Collins, J. Fields, K. E. DeWilde, M. L. Phillips, S. J. Mathew, E. Wong, C. Y. Tang, D. S. Charney, and D. V. Iosifescu. 2015. “Regulation of Neural Responses to Emotion Perception by Ketamine in Individuals with Treatment-Resistant Major Depressive Disorder.” *Translational Psychiatry* 5 (2): e509. https://doi.org/10.1038/tp.2015.10.

Murrough, James W., Andrew M. Perez, Sarah Pillemer, Jessica Stern, Michael K. Parides, Marije aan het Rot, Katherine A. Collins, Sanjay J. Mathew, Dennis S. Charney, and Dan V. Iosifescu. 2013. “Rapid and Longer-Term Antidepressant Effects of Repeated Ketamine Infusions in Treatment-Resistant Major Depression.” *Biological Psychiatry*, Depression: Risk, Rhythms, and Response, 74 (4): 250–56. https://doi.org/10.1016/j.biopsych.2012.06.022.

Niciu, Mark J., David A. Luckenbaugh, Dawn F. Ionescu, Sara Guevara, Rodrigo Machado-Vieira, Erica M. Richards, Nancy E. Brutsche, Neal M. Nolan, and Carlos A. Zarate. 2014. “Clinical Predictors of Ketamine Response in Treatment-Resistant Major Depression.” *The Journal of Clinical Psychiatry* 75 (5): e417–23. https://doi.org/10.4088/JCP.13m08698.

Nord, C. L., A. Gray, C. J. Charpentier, O. J. Robinson, and J. P. Roiser. 2017. “Unreliability of Putative FMRI Biomarkers during Emotional Face Processing.” *NeuroImage* 156 (August): 119–27. https://doi.org/10.1016/j.neuroimage.2017.05.024.

Offringa, Reid, Kathryn Handwerger Brohawn, Lindsay K. Staples, Stacey J. Dubois, Katherine C. Hughes, Danielle L. Pfaff, Michael B. VanElzakker, F. Caroline Davis, and Lisa M. Shin. 2013. “Diminished Rostral Anterior Cingulate Cortex Activation during Trauma-Unrelated Emotional Interference in PTSD.” *Biology of Mood & Anxiety Disorders* 3 (1): 10. https://doi.org/10.1186/2045-5380-3-10.

Overall, John E., and Donald R. Gorham. 1988. “The Brief Psychiatric Rating Scale (BPRS): Recent Developments in Ascertainment and Scaling.” *Psychopharmacology Bulletin* 24 (1): 97–99.

Pitman, Roger K., Ann M. Rasmusson, Karestan C. Koenen, Lisa M. Shin, Scott P. Orr, Mark W. Gilbertson, Mohammed R. Milad, and Israel Liberzon. 2012. “Biological Studies of Post-Traumatic Stress Disorder.” *Nature Reviews Neuroscience* 13 (11): 769–87. https://doi.org/10.1038/nrn3339.

Rush, A. John, Maurizio Fava, Stephen R Wisniewski, Philip W Lavori, Madhukar H Trivedi, Harold A Sackeim, Michael E Thase, et al. 2004. “Sequenced Treatment Alternatives to Relieve Depression (STAR*D): Rationale and Design.” *Controlled Clinical Trials* 25 (1): 119–42. https://doi.org/10.1016/S0197-2456(03)00112-0.

Sherbourne, Cathy Donald, and Anita L. Stewart. 1991. “The MOS Social Support Survey.” *Social Science & Medicine* 32 (6): 705–14. https://doi.org/10.1016/0277-9536(91)90150-B.

Weathers, F.W., D.D. Blake, P.P. Schnurr, D.G. Kaloupek, B.P. Marx, and T.M. Keane. 2013. “Clinician-Administered PTSD Scale for DSM-5 (CAPS-5).” General Information. 2013. https://www.ptsd.va.gov/professional/assessment/adult-int/caps.asp.

Weston, Sara J., Stuart J. Ritchie, Julia M. Rohrer, and Andrew K. Przybylski. 2018. “Recommendations for Increasing the Transparency of Analysis of Pre-Existing Datasets.” *PsyArXiv*, July. https://doi.org/10.31234/osf.io/zmt3q.

Yehuda, Rachel, Charles W. Hoge, Alexander C. McFarlane, Eric Vermetten, Ruth A. Lanius, Caroline M. Nievergelt, Stevan E. Hobfoll, Karestan C. Koenen, Thomas C. Neylan, and Steven E. Hyman. 2015. “Post-Traumatic Stress Disorder.” *Nature Reviews Disease Primers* 1 (October): 15057. https://doi.org/10.1038/nrdp.2015.57.

Young, R. C., J. T. Biggs, V. E. Ziegler, and D. A. Meyer. 1978. “A Rating Scale for Mania: Reliability, Validity and Sensitivity.” *The British Journal of Psychiatry* 133 (5): 429–35. https://doi.org/10.1192/bjp.133.5.429.

Zarate, Carlos A., Jaskaran B. Singh, Paul J. Carlson, Nancy E. Brutsche, Rezvan Ameli, David A. Luckenbaugh, Dennis S. Charney, and Husseini K. Manji. 2006. “A Randomized Trial of an N-Methyl-D-Aspartate Antagonist in Treatment-Resistant Major Depression.” *Archives of General Psychiatry* 63 (8): 856–64. https://doi.org/10.1001/archpsyc.63.8.856.