Supplementary Online Content

Javitt DC, Carter CS, Krystal JH, et al. Utility of imaging-based biomarkers for glutamate-targeted drug development in psychotic disorders: a randomized clinical trial. *JAMA Psychiatry*. Published online November 22, 2017. doi:10.1001/jamapsychiatry.2017.3572

eMethods.

eTable. Correlations between biomarkers and ketamine-induced symptoms

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

eMethods.

Overall: The study was implemented at 3 sites on scanners as follows: Columbia University/New York State Psychiatric Institute (3T GE MR750); University of California, Davis (3T Siemens TIM TRIO) and Yale University (3T Siemens TIM TRIO). All sites used a 32-channel head coil.

Prior to commencing the study, cross-site consistency was assessed using Agar phantoms. Similar stability was observed across all sites with ~1% fluctuation in the signal over time. Similar phase map patterns were observed across sites suggesting similar shimming capabilities.

Subjects: Subjects were medically healthy men and women aged 18-55 years without current or past Axis I or II psychiatric or substance history. Exclusion criteria included a history of recreational use of ketamine, PCP or other NMDAR modulator; adverse reaction to therapeutic ketamine; history of first-degree relative with schizophrenia; clinically significant history of violence or suicidality; history of significant medical illness, including high blood pressure (>140/90); significant neurological illness; significant head injury; pregnancy; taking current psychotropic medication; contraindication to an MRI or color blindness. All subjects provided informed consent prior to participation.

Demographics are shown in eMethods Table 1

Study sequence: The sequence of study procedures is shown in eMethods Figure 1. Following randomization, subjects were studied in two separate imaging sessions. On Day 1, MRS scans were obtained both prior to (30 min) and following (60 min) administration of the ketamine bolus and initiation of the ketamine infusion. On Day 14, task-based (RISE) and pharmacoBOLD were acquired.

PharmacoBOLD methods: T2*-weighted EPI scans were collected on day 14 for 10 min prior to and during each ketamine infusion. Each fMRI acquisition consisted of 298 volumes (TR = 2 sec, TE = 30 msec, flip angle = 77^0, with 32 slices at a resolution of 3.4 x 3.4 x 4.0 mm). Gradient echo field maps were also acquired proximal to the fMRI acquisition (TR = 500 msec, TE = 4.92 and 2.46 msec, flip angle = 55^0, with 32 slices at a resolution of 3.4 x 3.4 x 4.0 mm). Each functional scan was rigid-body motion corrected, slice time corrected, spatially smoothed (FWHM = 5 mm), and high-pass filtered (1200 sec) using FSL (FMRIB Software Library). Linear regression was performed using 27 motion parameters including 3 translations, 3 rotations, lag+1 derivatives, lag-1 derivatives, quadratic term, the mean relative motion, and the mean absolute motion. The model residual was then used for quantifying all ketamine-evoked responses

Linear registration to the subject's high-resolution Spoiled Gradient (SPGR) image (6 df) and to the MNI152 standard space template (12 df) was performed to generate spatial transforms. A region of interest (ROI) analysis was performed using the mid-cingulate cortex[1]. The ROI mask was created by first placing individual seeds at (MNI: 4,2,42) and

(MNI: -4,2,42) and then spherically dilating the seeds by 8-mm, which caused them to merge into a single mask. The mask was then transformed to the individual subject's functional space and multiplied by a gray matter mask to minimize the contribution of non-gray voxels. For each subject, the BOLD time series was extracted from the ROI and fit to a three-parameter gamma model:

$$y(t) \ = \ a \left(\frac{t}{t_{max}} \right)^{b \cdot t_{max}} e^{(t_{max}-t) \cdot b}$$

To determine whether head motion could account for group differences in the ketamine-evoked BOLD response, we computed the absolute and relative root mean squared difference. Root mean squared difference calculations were implemented in MCFLIRT:

Motion parameters were calculated using root mean squared difference as implemented in MCFLIRT:

$$\text{rms} = \text{sqrt}(0.2 \cdot R^2 \cdot ((\cos(\theta_x)-1)^2 + \sin(\theta_x)^2 + (\cos(\theta_y)-1)^2 + \sin(\theta_x)^2 + (\cos(\theta_z)-1)^2 + \sin(\theta_z)^2) + \text{trans}_x^2 + \text{trans}_y^2 + \text{trans}_z^2)$$

where R = radius of spherical ROI = 80mm; θ_x, θ_y, θ_z are the three rotation angles; and trans_x, trans_y, trans_z are the three translations.

Absolute motion was calculated relative to the reference volume used in the motion correction. Relative motion was calculated relative to the previous volume.

Potential effects of motion were assessed by correlation relative to the magnitude of the BOLD response. There was no significant effect of relative motion (rho = 0.02, p = 0.90) or absolute motion (rho = -0.15, p = 0.40) on BOLD magnitude within the ketamine group. There was also no significant effect of relative motion (rho = -0.17, p = 0.22) or absolute motion (rho = -0.02, p = 0.90) on the BOLD response across all subjects (ketamine and control group).

We computed framewise displacement as described in Power et al[2]. No significant differences were found (t-test, p = 0.073) between the ketamine (0.13 ± 0.07) and control (0.17 ± 0.14) groups. We computed the temporal signal to noise ratio (TSNR) as described by Triantafyllou et al[3]. No significant differences in TSNR were found (t-test, p = 0.97) between the ketamine (148.2 ± 27.1) and control (147.8 ± 45.4) means.

^{1}H-MRS methods: A series of high-resolution scans, consisting of standardized axial and coronal T1-, T2-, and spin density-weighted scans that were oblique, parallel and orthogonal to the anterior-posterior commissure plane were used for prescribing the ^{1}H-MRS voxel. This was followed by a SPGR T1-weighted volumetric scan (TR/TE = 30/8 ms, flip angle 45⁰, field of view 24 cm, 256 x 256 matrix, 124 coronal slices and a slice thickness of 1.0 mm). The ^{1}H-MRS voxel was placed in the medial prefrontal cortex

(mPFC) anterior to the genu of the corpus callosum, oriented along the AC-PC line and centered on the inter-hemispheric fissure, with dimensions 2.5 cm x 3.0 cm x 2.5 cm (volume, 18.8 cm^3). The ^{1}H-MRS acquisition used a standard Point Resolved Spectroscopy (PRESS) pulse sequence (TR/TE = 2000/80 ms) with 336 excitations and CHESS water suppression[4]. Four acquisitions (15-min each) were conducted over the 1-hr infusion period and compared with the pre-infusion baseline.

For data analysis, eight-channel phased-array coil data were combined into a single regular time-domain free-induction decay signal, using the unsuppressed voxel tissue water signal to derive the relative phased-array coil sensitivities. The *a priori* primary analysis focused on changes in the glutamate+glutamine to creatine ratio in the mPFC, denoted as Glx. For analysis, the unsuppressed water resonance was used for eddy current correction, and the spectra were broadened as necessary for all spectra from an individual session to have the same line width[5] and then fitted with LCModel 6.3[6]. To assess potential changes over time, the spectra acquired before the administration of ketamine/placebo were subtracted from those acquired after the start of the injection, and those difference spectra were fitted with LCModel.

Several measures were taken to insure spectral quality. Studies in which the water linewidth was greater than 15 Hz were not included. For the difference spectra the following criteria were used: 1. Any spectra for which the water was more than 2 Hz greater than the baseline scans was not used. 2 After differencing and residual baseline correction spectra were rejected if they did not have a sufficiently flat baseline throughout after LC model fitting. Difference spectra were rejected if the residual creatine and NAA peaks were more than 1 % of the unsubtracted peak intensity. At this level any subtraction error at the glutamate resonance would be below noise fluctuations. Spectra were rejected if the water suppression failed sufficiently that LCModel was unable to achieve a fit that resembled the data, either in the unsubtracted spectra or the difference spectra.

Voxel location and a representative spectrum are shown in eMethods Figure 2. Mean Glx CRLB, FWHM and SNR values pre/post infusion are shown in eMethods Table 2.

RISE task methods: The RISE task was implemented in accordance with Ragland et al.[7]. Scan parameters were identical to the pharmacoBOLD scans and differed only in length of acquisition: 237 volumes for the encoding scan, 252 volumes for each item recognition scan, and 171 volumes for the associative recognition scan. Briefly, subjects viewed a series of visual depictions of objects (3s each) that were presented in pairs on a presentation screen within the MRI scanner. fMRI data was collected during one encoding phase and two retrieval phases. During encoding, subjects viewed pairs of objects and alternated between performing *item-specific encoding* (e.g. Is either object living?) and *relational encoding* (e.g. Can one object fit inside the other?). During *item recognitio*n, subjects viewed pairs of items that were either previously presented or

previously unstudied (50% each) and indicated whether they were new or old. During *associative recognition*, subjects indicate whether or not the two items had previously been presented together.

Pre-designated primary outcome measures included behavioral performance for relational items and DLPFC BOLD activation during the task. BOLD data were analyzed using multiple linear regression implemented in FMRI Expert Analysis Tool [8]. Functional images were motion corrected and subjects with frame-to-frame movement > 3 standard deviations from the cohort mean were excluded. Images were spatially smoothed with a Gaussian kernel (7mm FWHM, isotropic) and high-pass filtered (< 200 s). Functional images underwent B0 un-warping using the gradient-echo field map estimates for each subject to correct for magnetic field inhomogeneity.

Activity associated with each event type was modeled by convolving a vector of expected neural activity with the 3 basis-function set of FMRIB's Linear Optimal Basis Sets (FLOBS). Only estimates from the first (canonical) basis function were passed up to higher (group-level) analysis. For the encoding task, item-specific and relational encoding were modeled as separate event types, and non-response trials were also modeled as their own event. Item recognition had seven event types: hits and misses separately for "old" targets that underwent item-specific and relational encoding, correct rejections and false alarms for "new" foil items, and non-response trials. Associative recognition had three event types: hits (i.e., correctly identifying a "changed" or "unchanged" target pair), misses (i.e., incorrectly identifying a "changed" target pair as "unchanged" or vice versa), and non-response trials.

Group-level contrasts were made for encoding (relational minus item-specific), item recognition (hits minus misses for item-specific and hits minus misses for relational encoding), and associative recognition (=hits minus misses). Research site was added as a covariate in the group-level GLM designs. [9]Anatomical ROIs for the PFC were identified for the relational minus item-specific encoding contrast with structural masks from the WFU_PickAtlas[9] used to restrict the mask to subgroups of activated voxels within left and right DLPFC [Brodmann areas (BA) 9, 46, and 9/46].

RISE task data integrity evaluation: To evaluate subject motion, framewise displacement metrics were calculated using FSL. For the pre-infusion scans, the amount of relative frame-to-frame motion was higher for the second Item Recognition run (motion=0.168) than during Encoding run (motion=.077); $t=2.7$, $p<.05$. Similarly, subject motion during the second Item Recognition run was greater than during the Associative Recognition run (motion=.095); $t=2.4$, $p<.05$, and was also greater than the first Item Recognition run (motion=.111); $p=3.05$, $p<.01$

For the post-infusion scans, subject motion was greater during Encoding (motion=.147) than during the first Item Recognition run (motion=.104); $t=2.8$, $p<.01$. There were no other condition differences in motion, and it should be noted that all the motion was within 3 SDs of the cohort mean.

For RISE, 20 participants were excluded. None were excluded due to poor image quality, but a number were missing because of administration errors (e.g., giving the wrong version of the task), because patients were not generating valid behavioral responses or because of incomplete or missing data.

eMethods Table 1: Study Demographics[1]

|	Ketamine (n=39)	Placebo (n=20)	Statistics
Age	31.1±9.6	32.2±10.2	n.s.
Male (%)*	54%	80%	n.s.

1. For subjects who completed at least one scan (Day 1). Demographics from all randomized subjects were essentially similar. n.s.= no significant difference on Fisher exact test.

eMethods Table 2: H-MRS analysis parameters from LC model

		mean	SD	mean	SD	Mean	SD
		CRLB Glx (%)		FWHM (Hz)		SNR	
Overall	Pre	10.8	3.9	9.9	2.3	23.3	11.6
	Post	10.9	4.4	9.8	2.2	24.3	12.4
By site:							
Columbia	Pre	13.9	2.7	11.0	2.0	13.1	1.2
	Post	14.7	2.5	10.9	2.0	13.1	1.4
UCDavis	Pre	7.3	1.1	8.1	1.4	34.3	4.9
	Post	6.7	1.0	8.1	1.4	35.8	5.2
Yale	Pre	7.8	1.8	10.1	2.0	32.2	11.2
	Post	7.7	2.4	9.8	2.0	33.7	11.2

eMethods Figure 1: Sequence of procedures, day 1 and 14 scans
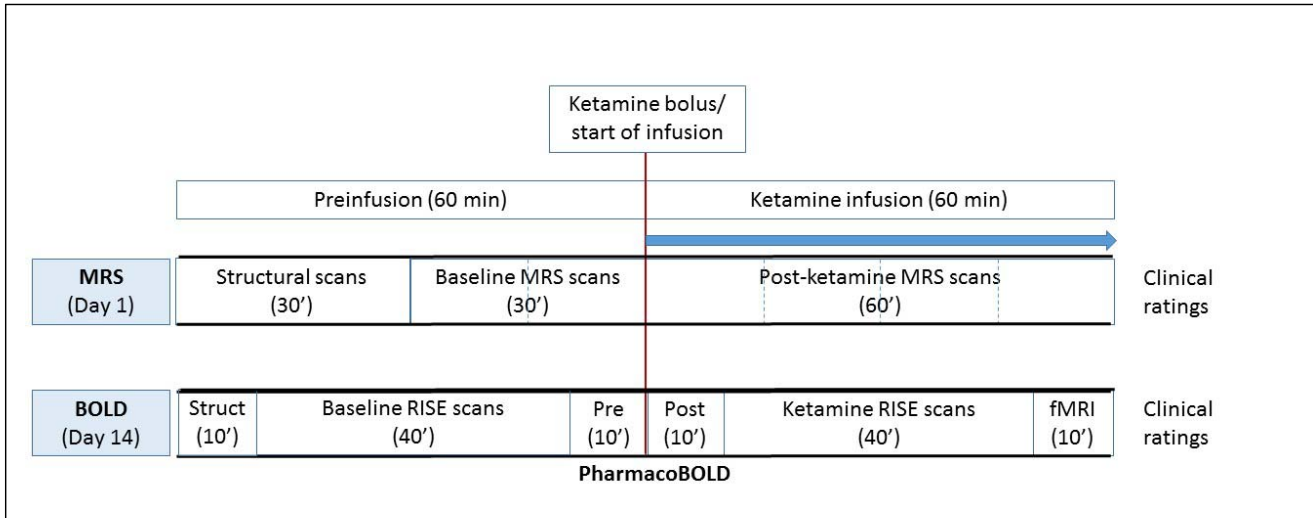
Schematic diagram of the MRS (day 1) and RISE task / pharmacoBOLD imaging day sequences (day 14). Struct indicates Structural Scans

eMethods Figure 2: Voxel location and representative spectrum for 1H-MRS data
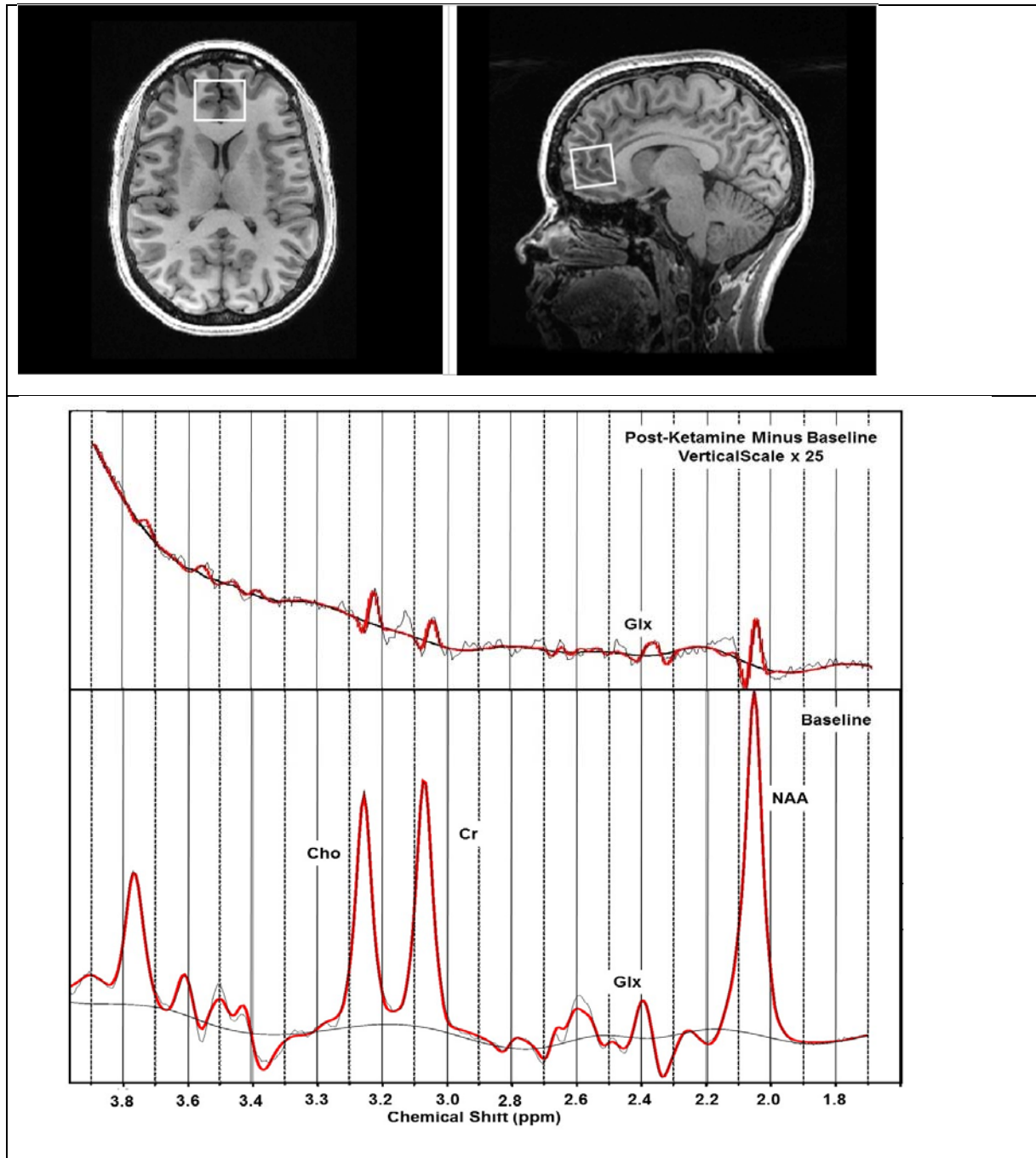
Figure legend: *Top*: Voxel location for ^{1}H-MRS shown in transverse and sagittal section. *Bottom*: Representative data from a single subject. A spectrum acquired before administration of ketamine/placebo (lower panel) was subtracted from the spectra after the injection, yielding differences (upper panel). Because the differences were small, the difference is displayed with a vertical scale 25 times greater than the baseline spectrum. LCModel Version 6.3 was used to fit the baseline and difference spectra (red lines). The subtraction errors in Cr, Cho, and NAA are 0.5% relative to baseline, in contrast with the 2% difference in Glx.

eReferences

1.   De Simoni S, Schwarz AJ, O'Daly OG, et al. Test-retest reliability of the BOLD pharmacological MRI response to ketamine in healthy volunteers. *NeuroImage.* 2013;64:75-90.
2.   Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage.* 2012;59(3):2142-2154.
3.   Triantafyllou C, Hoge RD, Krueger G, et al. Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. *NeuroImage.* 2005;26(1):243-250.
4.   Haase A, Frahm J, Hanicke W, Matthaei D. 1H NMR chemical shift selective (CHESS) imaging. *Phys Med Biol.* 1985;30(4):341-344.
5.   Kreis R. Issues of spectral quality in clinical 1H-magnetic resonance spectroscopy and a gallery of artifacts. *NMR in biomedicine.* 2004;17(6):361-381.
6.   Provencher SW. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magn Reson Med.* 1993;30(6):672-679.
7.   Ragland JD, Ranganath C, Harms MP, et al. Functional and Neuroanatomic Specificity of Episodic Memory Dysfunction in Schizophrenia: A Functional Magnetic Resonance Imaging Study of the Relational and Item-Specific Encoding Task. *JAMA psychiatry.* 2015;72(9):909-916.
8.   Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. Fsl. *Neuroimage.* 2012;62(2):782-790.
9.   Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage.* 2003;19(3):1233-1239.

eTable. Correlations between biomarkers and ketamine-induced symptoms

Variable		pharmacoBOLD r (p)	Glx r (p)	Task fMRI r (p)
All Subjects		N = 53	N = 44	N = 39
BPRS	Activation	0.14 (0.3)	0.056 (0.7)	0.28 (0.10)
	Hostility	0.048 (0.7)	0.021 (0.9)	0.077 (0.65)
	Negative	0.31 (0.03)	-0.031 (0.8)	0.36 (0.027)
	Positive	0.56 (<.0001)	0.19 (0.23)	0.048 (0.8)
	Total	0.43 (0.0014)	0.21 (0.17)	0.30 (.075)
CADSS		0.56 (<.0001)	0.12 (0.43)	-0.12 (0.48)
POMS		0.31 (0.032)	0.031 (0.85)	-0.069 (0.69)
PSI		0.37 (0.007)	0.16 (0.3)	0.21 (0.21)
Ketamine Group		N = 34	N = 29	N = 27
BPRS	Activation	-0.15 (0.40)	-0.17 (0.40)	0.38 (0.06)
	Hostility	0.02 (0.9)	-0.07 (0.7)	0.039 (0.85)
	Negative	0.33 (0.07)	-0.15 (0.46)	**0.43 (0.034)**
	Positive	0.14 (0.45)	-0.26 (0.19)	0.27 (0.20)
	Total	0.085 (0.64)	-0.20 (0.32)	**0.46 (0.02)**
CADSS		0.078 (0.67)	-0.15 (0.46)	0.14 (0.50)
POMS		0.075 (0.68)	-0.19 (0.34)	0.19 (0.37)
PSI		0.089 (0.63)	-0.16 (0.43)	0.29 (0.16)
1Left DLPFC activation, relational encoding				

Bold indicates significant values (at *P* < .05).