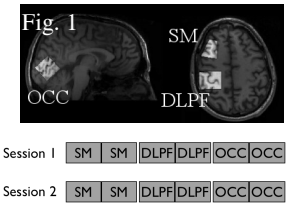
On the repeatability and required sample sizes of GABA-MRS experiments.

C. John Evans[1], Nicolaas Puts[2,3], Richard Edden[2,3], and David McGonigle[1,4]

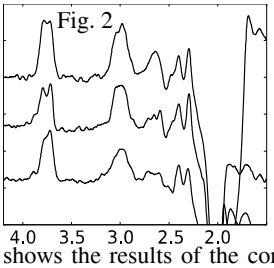
[1]*CUBRIC, School of Psychology, Cardiff University, Cardiff, Wales, United Kingdom,* [2]*Russel H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, MD, United States,* [3]*FM Kirby Center for Functional Brain Imaging, Kennedy Krieger Institute, MD, United States,* [4]*School of Biosciences, Cardiff University, Cardiff, Wales, United Kingdom*

Target Audience: Neuroscientists planning studies to measure GABA using MEGA-PRESS.

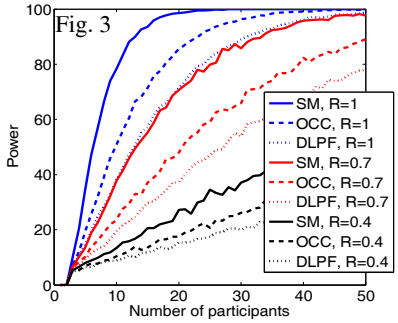
Purpose: Measuring GABA to investigate individual differences in healthy controls requires high measurement repeatability due to smaller effect sizes than in clinical or drug studies. A number of studies have reported the repeatability of GABA MRS using coefficients of variation[1] (CV), however, the utility of CVs is limited without an estimate of either the effect size of interest or the true variation in GABA concentration between participants. The aims of this study are to; (i) evaluate the measurement repeatability of GABA, within and between sessions, in three different regions, (ii) measure the genuine variability in GABA concentration between healthy control participants, and, (iii) use these data to estimate the required sample sizes for experiments investigating correlations between GABA and behaviour in healthy controls.



Methods: 11 healthy male participants (age 21-36) were scanned using a GE 3T system. Participants gave informed consent for this study, which was approved by the local IRB. Each participant was scanned in two sessions (on different days). During each session, two GABA (+MM) acquisitions were performed in 3 counterbalanced voxels - occipital (OCC), sensorimotor (SM) and dorsolateral pre-frontal region (DLPF), (Fig. 1) using MEGA-PRESS[2] (TR/TE=1800/68ms, 332 averages). All spectra were checked for artefacts, resulting in the rejection of DLPF data from 1 participant. A T_1-weighted structural image was acquired for tissue correction. *Coefficients of variation:* Two CVs (as defined by Bland[3]) are reported; CV_{wp}, the *within-participant CV* (from all 4 measurements acquired on each participant), and CV_{ws} the *within-session* CV (treating sessions independently). To determine the robustness of the CV estimates, a bootstrap analysis was performed, sampling with replacement 100,000 times. *Sources of variance:* Total variance in GABA+ values within the dataset, σ^2_{tot}, contains contributions from; (i) variance in GABA+ between participants, σ^2_p; (ii) variance between sessions, σ^2_s (systematic difference between sessions); (iii) within-session error variance σ^2_e; (iv) interaction σ^2_{ps}, (non-systematic difference between sessions). These components of variance were separated using an 2-way ANOVA with interaction, with participant and session as factors, and expressed as percentages of the total variance. *Correlation experiment sensitivity*. Simulations were performed to calculate the required sample size for a study correlating GABA against parameter, Y, (e.g. a metric of behaviour or neuroimaging signal change). Three simulated datasets of 10,000 pairs of $GABA_{sim}$ & Y values were created with correlations of R=0.4,0.7,1.0. The mean and standard deviation of $GABA_{sim}$ were set to the measured *in vivo* values. On 10,000 occasions, N_p samples were taken from this simulated dataset, adding noise to the $GABA_{sim}$ value according to the total noise variance (containing σ^2_s, σ^2_{ps} and σ^2_e). Values of N_p ranged from 2 to 50 to span a typical range of numbers of participants. For each value of N_p and R, the statistical power of the experiment was estimated (the percentage of times that the (true) correlation between $GABA_{sim}$ and Y was successfully detected). *Group experiment sensitivity*. To model a study investigating group differences in GABA concentrations, F-distributions were calculated using the total noise variance (σ^2_s, σ^2_{ps} and σ^2_e), at a variety of simulated effect sizes, ε, and numbers of participants per group, N_{pg}. For each for each N_{pg}, the critical ε required to detect an effect with 80% power was calculated.



Results & Discussion: Typical spectra from a single participant OCC (top), SM, DLPF (bottom) are shown in Fig 2. The within-participant CVs are higher than the within-session CVs (Table 1), due to between-session effects (e.g. repositioning participant), but there is significant overlap between the distributions. Separating the sources of variance (Table 1) shows that the relative contribution of the variance of interest, the *between participant variance,* σ^2_p, is highest for SM (48.4%), but only 13.7% for DLPF. In all voxels, more variance originates from within-session effects ($\sigma^2_e > \sigma^2_{ps}$, $\sigma^2_e > \sigma^2_s$), again suggesting that within-session variability dominates over between-session variability. Fig. 3 shows the results of the correlation experiment sensitivity analysis. For high 'true' correlations (R=1, blue) 80% power can be attained with a moderate number of participants (10-20, depending on voxel), but for a more realistic R=0.7 (red), from 20 to >50 participants are required. For the group experiments, typically >10 participants per group are required to detect an effect size of 15% or greater in OCC, SM. The higher CV of DLPF increases this to ~30 participants for a comparable effect size.



Conclusions: Variation in GABA levels between individuals can contribute up to 50% of the variance in the data, but this varies considerably between regions, with the SM showing the highest between-participant variance. For 80% power in correlation experiments, group sample sizes should be >20 participants.

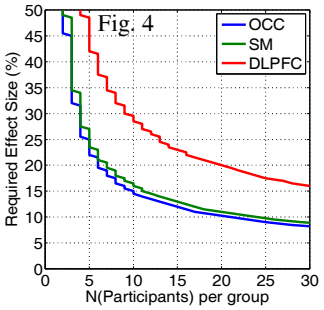


Table 1	[GABA+] Mean ± σ_{tot}	CV_{ws} (%)	IQR (25,75%)	CV_{wp} (%)	IQR (25,75%)	σ^2_p (%)	σ^2_s (%)	σ^2_{ps} (%)	σ^2_e (%)
OCC	1.50 ± 0.15 (9.9%)	7.4	(6.5,8.1)	8.7	(8.3,9.2)	29.5	0.0	28.3	42.2
SM	1.56 ± 0.18 (11.5%)	8.6	(7.5,9.4)	9.4	(8.4,10.2)	48.4	1.1	12.1	38.5
DLPF	1.08 ± 0.20 (18.5%)	16.2	(13.8,18.0)	17.0	(15.1,18.7)	13.7	0.0	15.8	70.5

References: 1. O'Gorman RL, *et al,* In vivo detection of GABA and glutamate with MEGA-PRESS: Reproducibility and gender effects. J. Magn. Reson. Imaging 2011; 33,1262. **2.** Mescher, M., et al, Simultaneous in vivo spectral editing and water suppression. NMR Biomed. 1998;11, 266. **3.** Bland, JM & Altman, DG. Measurement error. Brit. Med. J. 1996 313, 744.