Study Protocol & Statistical Analysis Plan

Title: Feasibility of Gamma Transcranial Alternating Current Stimulation to Reduce Beta-amyloid Load and Improve Memory

NCT#: NCT04646499

Last Updated: 11/02/2021

Participants.

This study was preregistered at ClinicalTrials.gov with identifier NCT04646499. The study protocol was approved by the University of California San Francisco (UCSF) Institutional Review Board and all participants signed informed consent documents prior to participating in the study. Participants received $20 per hour for participation and a $50 bonus for completion of the study. To be eligible for this study, all participants were 60-80 years of age, fluent in English, had at least 12 years of education, and had normal or corrected-to-normal vision. In addition, participants had to have aMCI (see below for aMCI inclusion criteria) and be able to complete cognitive tasks and cooperate with all study procedures. The exclusion criteria were as follows: neurological or psychiatric disorders other than aMCI, receiving investigational medications or participated in a clinical trial with medications within the previous 30 days, family history of epilepsy, implanted electronic devices (e.g. pacemaker), prior serious head trauma, pregnant, or IQ under 80. In addition, participants were not on cholinesterase inhibitors, memantine, psychotropics, anti-depressants, or anti-anxiety medications. Finally, participants did not have color vision deficiency, a history of substance abuse, glaucoma, macular degeneration, amblyopia, or strabismus.

Fourteen older adults with aMCI were enrolled and 13 completed the entire intervention in a single-arm design (age mean: 73.62 years (SD: 5.44), Montreal Cognitive Assessment mean (MoCA): 23.39 (SD: 3.66), Education years mean: 17.69 (SD: 2.78), 7 Male, 11 right-handed). Prior to enrolling in the study, all participants completed the Montreal Cognitive Assessment (MoCA). In the current study, participants were considered aMCI by scoring between 27 and 16 on the MoCA and had a self-complaint of memory difficulty. In addition, participants needed an age-matched Z-score of at least -1 on immediate memory or delayed memory (as measured by the California Verbal Learning Test, short version) *and* at least -1 Z-score on verbal fluency (D words), semantic fluency (animals), processing speed (digit symbol and number trails tasks), or task switching (number letter trails task). All participants also scored as having intact activities of daily living by scoring as mostly independent on the IADL. Finally, all participants had the presence of both AB and tau in the pre-tACS phlebotomy biomarker analyses, indicating that MCI status is likely due to progression towards AD as opposed to other potential causes of cognitive decline. One participant withdrew voluntarily prior to any cognitive assessment or tACS.

Study Timeline.

The study required 10 visits to UCSF over 5 weeks. During visit 1, participants completed baseline assessments of the instrumental activities of daily living (IADL) survey, cognitive tests assessing baseline episodic memory performance, and an MRI scan (Week 1). On the following week, participants received gamma tACS while engaged in the Stimulation Tasks (see details below) on five consecutive weekdays (Week 2, visits 2-6, Monday-Friday). On one day during each of the three following weeks, participants also completed the Stimulation Tasks while receiving gamma tACS (Weeks 3-5, visits 7-9). These sessions occurred on the same day of the week for each participant (e.g. consecutive Tuesdays), with a small number occurring ± one day. Finally, during the 5th week (visit 10), participants completed the same procedure as the baseline assessment (visit 1), which included the IADL survey, tests of episodic memory, and an MRI scan. All study activities took place at UCSF, except for one participant who completed the eight tACS sessions in their home. For this participant, a researcher was present with the required tACS equipment during the at-home sessions, and the participant completed the MRI and outcome assessments at UCSF.

Surveys.
Participants completed the Instrumental Activities of Daily Living (IADL) during the initial pre-tACS and post-tACS follow-up Cognitive Task sessions (Weeks 1 and 5, respectively) . Additionally, following each tACS session, participants filled out a tolerability survey of the following 11 measures on a scale of 0 (not noticeable) to 10 (not tolerable): headache, neck pain, scalp pain, tingling, itching, burning sensation, increased alertness, increased sleepiness, trouble concentrating, acute mood change, and presence of phosphenes. Prior to the initial tACS session (Week 1) participants completed surveys that measure their anticipation, anxiety, and expected benefits on both memory and cognition following tACS on a scale of 1 - 10. Following the final tACS session (Week 5), participants were probed on the same categories for their perceived benefits of the tACS intervention.

Magnetic Resonance Imaging.
Participants completed two MRI sessions with the same protocol prior to and following the tACS intervention (i.e., during the pre- and post-tACS phases, respectively). All data was collected by a Siemens 3T MAGNETOM Trio MRI using a 64-channel head coil. First, high-resolution T1-weighted anatomical images were acquired (1 x 1 x 1 mm voxel size, FOV = 160 x 240 x 256 mm, repetition time (TR) = 2300 ms, echo time (TE) = 3 ms, flip angle (FA) = 9°). For MRS, two voxels were placed in the left Pars, and left Hipp (3x3x2 cm each) based on the T1 structural scan. We then collected GABA-edited MEGA-PRESS scans (TR/TE = 2000/68 ms, 128 averages, FA = 90°), with editing pulses at 1.9 edit-ON and 7.2 edit-OFF ppm. The edit-ON and edit-OFF differences yield the peaks affected by the editing pulses.

Next, participants completed approximately six-minutes of eyes-closed resting-state functional MRI (rs-fMRI) using a T2*-weighted echoplanar imaging (EPI) sequence with the following parameters: 560 volumes, TR = 850 ms, TE = 32.8 ms, FA = 45°, in-plane resolution = 2.2 mm^2, 66 total 2.2 mm slices using a multiband acceleration factor of 6. Participants were instructed to close their eyes, remain awake, and be as still as possible. Last, DTI data was collected with 10 non-diffusion weighted images (b = 0 s/mm^2) followed by 96 diffusion weighted images (b = 2500 s/mm^2; TR = 2420 ms, TE= 72.2 ms, and FA = 85°). One participant did not have usable rs-fMRI at their post-tACS session due to imaging artifacts, resulting in usable datasets from 12 participants for these analyses. One participant did not have baseline DTI data due to time constraints in the MRI, resulting in usable datasets from 12 participants for these analyses.

Phlebotomy.
Participants had two phlebotomy sessions immediately after their pre- and post- MRI sessions, on days with baseline behavioral assessments in Week 1 and the final tACS session in Week 5, respectively. Before both sessions, participants were instructed to fast from any food intake prior to arrival for a minimum of 4 hours. Following the session, participants had an opportunity to rest and eat food. On both sessions, participants underwent a draw of 20 mL of blood in two 10 mL vials. These vials were immediately delivered to the sample processing lab located on the floor above the phlebotomy lab where the samples were centrifuged at 200 x g for 15 min at 4 °C then put in storage until analysis. At the end of data collection, the entire sample was gradually brought to room temperature for analysis. All analyses were conducted by a board-certified laboratory technician blinded to the hypotheses of our study. This pre- and post-tACS analyses of the blood draw samples allow for changes in measures of amyloid, tau, and NfL.

Neuromodulation.
Participants were fitted with a neoprene head cap with electrodes located at 8 locations bilaterally (F7, F8, FT7, FT8, T7, T8, P7, and P8; 10-20 EEG system). The tACS was delivered through an 8-channel mobile Starstim device (Neuroelectrics, Spain) with NG Pistim electrodes (contact area: 3.14 cm^2). On each session the tACS current was ramped up slowly over 4 minutes to a total of 1.6 mA (0.4 mA per electrode baseline to peak, with the four contralateral electrodes set to 180° offset), maintained full strength for 52 minutes, then ramp back down to 0 mA over 4 min. We sought to target multiple regions of the cortex known to be related to MCI, AD, and episodic memory processing: the Hipp, IPL, and Pars. Based on our preliminary research in mouse models of AD, which tested whether there was a dose-response curve with greater stimulation strengths and durations[1], we observed that the greatest doses of gamma oscillatory current had the strongest effects. Therefore, we applied the highest dose of stimulation that we believed would be tolerable to adults with MCI, without inducing overly distracting side effects such as phosphenes or physical sensations that are more likely at higher stimulation doses. During each of the eight 60-minute tACS sessions, participants engaged in the Stimulation Tasks (see below for details). Following the end of each tACS session, participants filled out a survey of side effects (see Surveys above).

Cognitive Testing.
During Week 1, prior to any tACS sessions, participants first completed the California Verbal Learning Test (CVLT) to assess aspects of episodic memory, including delayed free-recall, and cued recall/recognition. The CVLT is a task known to positively correlate with hippocampal size, act as an early detector of AD, and track disease progression such as demyelination in multiple sclerosis. Participants completed the standard and alternate versions of the CVLT during the initial cognitive testing session (Week 1) and the final Cognitive Testing session (Week 5), with the order counterbalanced across participants. The CVLT has a 20-minute and a 10-minute break period, with the former separating the short-delay free recall (SDFR) from the long-delay free recall (LDFR), and the latter break separating the LDFR from the force-choice recognition portion.

During the CVLT 20-minute break, participants completed the Paired Associates Task (PAT). The PAT assesses episodic memory via recall and recognition of pairs of face and scene stimuli and, therefore, was selected so that the stimuli would not interfere with the verbal words encoded during the CVLT. We used a modified version of the PAT, where the stimulus presentation timing was increased by 500 ms to 2500 ms; however, the trial count and task demands were the same as the PAT in Voskontas et al.[2]. The task begins with an encoding period consisting of 10 face-scene pairs. After the presentation of each face-scene pair (2500 ms), participants judged whether the person is in an indoor (‘D’ key) or outdoor scene (‘K’ key). The 10 face-scene pairs were presented randomly six times each. Immediately following the PAT encoding period, participants were tested in the recall period where they were presented with one face or scene (2500 ms) followed by three items of the other category (scenes or faces, respectively). Participants then responded (untimed) with which of the three items they believed was paired with the initial stimulus presented during the encoding phase. Each face and scene item pair were used as the probe three times resulting in 60 total recall trials. Following the end of the PAT recall period, participants were presented with a recognition task where they were shown either correctly matched or incorrectly matched face-scene pairs (untimed). Correct face-scene pairs were presented three times in random order each and incorrect pairs were presented 60 times resulting in 90 trials. An

alternate version of the PAT task was used during the final cognitive testing session with unique stimuli (Week 5) and these two versions were counterbalanced between participants.

During the CVLT 10-minute break, participants completed a verbal and semantic fluency task where they had 60 seconds to say as many words as possible that begin with the given letter or fit the category (excluding proper nouns). During the pre- and post-tACS intervention (visits 1 and 10) participants completed the categories F, A, S, and animals. The cognitive testing always occurred at the beginning of the day, so that no fatigue effects lowered participant performance and to ensure that participants could successfully complete both standardized memory tests (CVLT) as well as the computerized memory test (PAT).

Stimulation Tasks.

Once the tACS began, participants completed a series of Stimulation Tasks designed to engage episodic memory. First, they completed the Mnemonic Discrimination Task (MDT^3), which assesses high-fidelity long-term memory (LTM). During each tACS session, participants completed the MDT with unique stimuli consisting of common real-world items. This resulted in eight versions of the MDT, which were counterbalanced in order between participants (forwards or backwards order). Immediately after the tACS began, participants completed the encoding portion of the MDT task. During encoding, participants were cued first with "will the object fit inside a lady's shoe box?" Following the prompt, participants viewed 39 items in a random order and responded to the question for each item with an untimed button press. Next, participants were presented with a second cue "can you carry the object across the room using only one hand?" Participants then responded to this question following a random presentation of the same 39 items. All items were presented for 2500 ms followed by a screen that that listed the button press responses ("D" for yes, "K" for no). A fixation cross was presented on the screen for 1000 ms between trials. The MDT encoding portion of the task is timed to last 8-10 minutes. All versions of the MDT task were created using PsychoPy.

Following the MDT encoding period, participants completed 20 minutes of a commercially available tablet game, Spot the Difference. This game was chosen so that participants are mentally engaged in a visual search task that has no interference with recently encoded items from the MDT task and not a task that would evoke stress due to task difficulty. Participants completed this task at their own pace and advanced to new images without any score being kept. Next, participants completed the semantic fluency task with three letters and a category similar to the cognitive testing with F, A, S, and animals (see above). Here, we created 8 unique sets each consisting of three letter prompts and one semantic category. Thus, the fluency task was not repeated across any of the tACS sessions and was different from the fluency tasks used as our Cognitive Tasks outcome measures. Participants were instructed to name as many items as possible.

During the final 10 minutes, participants completed three MDT test blocks where they viewed 13 items seen during encoding, 13 items that were an alternate (similar) version of the encoded items, and five novel lures. For each item participants responded as to whether they were the same exact item from the encoding task from the beginning of stimulation. Each item was viewed for 2500 ms and followed by a screen where participants would respond whether the item was definitely old ("D" key), maybe old ("F" key), maybe new ("J" key), or definitely new ("K" key). A fixation cross was presented on the screen for 1000 ms between trials. The order of the target, lure, and novel items was random within each of the three testing blocks. In total, the tasks for the participant to complete during stimulation are equal to the length of the tACS session (1 hour).

Data Processing and Statistical Analysis.
Given the relatively small number of participants this study, we used non-parametric tests for all statistical analyses to reduce influence from potential extreme values. Changes in outcome metrics were assessed with non-parametric Wilcoxon sign-ranked tests. We also report rank-biserial correlations (r_{rb}) to represent effect sizes of these changes, similar to Cohen's d. We set a significance threshold of $p < 0.05$ and report non-significant 'trends' at $p < 0.10$. As this is a hypothesis-generating pilot study, we did not correct for multiple comparisons for each test conducted.

Primary Outcomes.
Our primary outcome measures were feasibility and tolerability of eight sessions of gamma tACS. To assess these, we measured dropout rate of participants and the reported side effects collected following each stimulation session. Analysis of the side effects collected following the end of tACS consist of the average across all 8 sessions for the 11 potential categories. These metrics were rated on a scale of 0 (not noticeable) to 10 (not tolerable). As this was a single-arm design, we compared side effect ratings from this study to those from a recently completed a multi-session tACS intervention in healthy older adults of the same age range, in which we administered 1 mA of 6 Hz tACS to the prefrontal cortex (F3-F4, 10-20 EEG system). This allows for an age-matched comparison between prior stimulation protocols and one that is higher in dose (1 mA vs 1.6 mA), frequency (6 Hz and 40 Hz), and duration (~20 minutes and 60 minutes) employed in the current study. As both studies included the same side-effects questionnaire categories, we conducted a non-parametric Mann-Whitney U test to assess differences in the average ratings for each metric between the two studies.

Secondary Behavioral Outcomes.
Secondary outcome measures assessed changes in performance on the memory and fluency tasks (CVLT, PAT, fluency). The CVLT is a standardized cognitive test with different task demands at each stage. For CVLT, we focused on the correct items named during the short-delay free recall (SDFR) and long-delay free recall (LDFR) sections. For the PAT, we measured accuracy for both the recall and recognition portions. For the semantic (animals) and verbal (F, A, S) fluency task, we measured the number of correct non-repeated words.

Secondary Phlebotomy Outcomes.
Blood was collected by phlebotomy in EDTA tubes and centrifuged at 2,000 g for 10 min at 4°C. Plasma was then aliquoted in 500 microliter polypropylene tubes, and stored at -80°C until analyses, with an average needle-to-frozen storage time < 2 hours. Plasma was analyzed according to vendor protocols, using commercially available kits Neuro 4-PLEX E [amyloid $\beta_{1\text{-}42}$ (Aβ42), amyloid $\beta_{1\text{-}40}$ (Aβ40), neurofilament light chain (NfL), and glial fibrillary acidic protein (GFAP)] and P-tau181 for single molecule arrays in an HD-X analyzer (Quanterix, Billerica, MA). Lowest level of quantification and average coefficient of variations were 0.378 pg/mL, 4.1% for Aβ42; 1.02 pg/mL, 3.5% for Aβ40; 0.4 pg/mL, 4.3% for NfL; 2.89 pg/mL, 4.6% for GFAP; and 0.4 pg/mL, 10.1% for P-tau181. Analyzed samples underwent only one thaw cycle prior to use. Samples were run in duplicate, with kits from the same lot. Following analysis, we were able to measure the picograms per milliliter (pg/mL) of each of the three AD biomarkers at both time points, to quantify stimulation-related changes in AD biomarker levels between the two time points.

Exploratory Behavioral Outcomes.
We examined changes in MDT task performance as a non-registered exploratory measure, focusing on the changes in the lure discrimination index (LDI, proportion lure correct rejection—proportion novel false alarm) between the initial tACS session and final tACS session. Changes in performance were assessed with a Wilcoxon sign-rank test. As the MDT task was completed concurrently with the tACS, it is important to note that changes in MDT performance reflect gains during tACS from repeated test performance, rather than transfer to an untrained task, as with our secondary outcomes. In addition to examining changes in this study population, however, we also compared initial and final LDI performance in these MCI participants to a placebo training control group of healthy older adults from a previous study, where no training-related MDT gains were expected. For this analysis, we compared both initial and final performance between these groups using non-parametric Mann-Whitney U tests.

Further, we examined changes in IADL from pre-tACS to post-tACS as an exploratory measure on each of the categories: bathing, dressing, grooming, mouth care, toileting, transferring bed/chair, walking, climbing stairs, eating, shopping, cooking, managing medications, phone use, housework, driving/transport, and managing finances. Specifically, we measured change in the IADL metrics from the pre-tACS to post-tACS session by measuring the average rating across all IADL categories and comparing these in non-parametric Wilcoxon signed-rank tests between the pre-tACS baseline and post-tACS follow-up sessions.

Resting State Functional MRI.
We conducted exploratory analyses on the rs-fMRI data collected during the pre- and post-tACS scans. Standard preprocessing of MRI data for functional connectivity analyses was carried out with AFNI and FreeSurfer. First, participants' T1-weighted anatomical images were skull-stripped and segmented into gray matter, white matter, and cerebrospinal fluid (CSF) using FreeSurfer. Next, EPI data was pre-processed using afni_proc.py in AFNI. First, the de-spike option was applied to interpolate extreme outlier timepoints from the BOLD signal intensity time courses. Next, the EPI data was motion corrected (aligning each EPI volume to the volume with the minimum outlier fraction), co-registered to the T1-weighted anatomical image, warped to MNI space, and resampled to an isotropic resolution of 2 mm^3. Further, we implemented noise reduction by regressing out several 'nuisance' variables: (1) 12 motion regressors (6 realignment parameters and their derivatives), (2) voxel-wise local white matter regressors using AFNI's fast ANTATICOR method, and (3) the top 3 principal components from lateral ventricle voxels. EPI data was also bandpass filtered (0.01-0.10 Hz) and volumes with excessive motion (> 1 mm root mean square motion) and the volume prior were censored. Nuisance regression, bandpass filtering, and censoring were performed in a single step.

To measure functional connectivity, we quantified Pearson's correlations (and applied a Fisher z-transform) between ROI timeseries from Pars, IPL, and Hipp (left and right) during the baseline and follow-up MRI sessions. Pars and Hipp ROIs were automatically output from the MRS scan (see below) and we generated an additional IPL ROI in FSL. We copied each left hemisphere mask to the right hemisphere to generate masks for all six regions. We calculated functional connectivity as the average connectivity between each ROI pair between and within hemispheres (e.g., IPL-Hipp connectivity calculated as the average of right IPL-Hipp, left IPL-Hipp, right IPL – left Hipp, and left IPL – right Hipp). We note that RSFC for one participant had relatively high values at baseline for each ROI pairing (average RSFC: 0.64) compared to the other participants, despite having no visible artifacts or excessive motion. Importantly, due to

technical issues, this participant did not have baseline MRS or DTI measures and therefore did not overly influence correlation results with those metrics.

Magnetic Resonance Spectroscopy.
Proton 1H MRS data was collected in two ROIs (left Pars and left Hipp). MRS data from our third ROI (IPL) was not collected due to time constraints in the scanner. During the second MRI session we duplicated voxel placement based on images from the first MRI session. All GABA and Glx concentrations were analyzed using GANNET 3.1, which is a specialized toolbox for Matlab designed to analyze GABA by calculating the area under the curve. The E/I ratio of Glx to GABA within ROIs was calculated (Glx/GABA+, both CSF-corrected with a water reference file). This ratio measured the excitatory/inhibitory balance within the ROI at both the pre-tACS and post-tACS MRI sessions. Two participants had poor shimming at baseline and were excluded from analyses, resulting in usable datasets from 11 participants for these analyses.

Diffusion Weighted Imaging.
DTI data was preprocessed using FMRIB's Diffusion Toolbox (FDT fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT), a part of Functional MRI of the Brain software (FSL; Analysis Group, FMRIB, Oxford, United Kingdom). We first removed non-cortical tissue with the Brain Extraction Tool within FSL. To correct for head movement and eddy current distortions, we used Eddy Correct. Next, we used DTIfit to calculate fractional anisotropy (FA) maps for each participant. The FA maps were input in to Tract-based Spatial Statistics (TBSS) in order to create a mean FA skeleton by aligning to a common space (FMRIB58_FA; FA > 0.2 threshold). We analyzed individual white matter tracts generated with the Johns Hopkins University white matter atlas included with FSL. As we did not expect significant change in FA in a month, we only report the results of the pre-tACS baseline DTI scan.

We selected four specific white matter tracts that serve our functional ROIs (Pars, IPL, Hipp). First, we selected the Fornix (FX) as it serves as the dominant outflow tract of the hippocampus. Second, we selected the Arcuate Fasciculus (AF) as it connects the frontal and temporal lobes by passing through the parietal lobe, therefore acting as an important tract for the IPL and Pars ROIs. Third, we selected the Superior Longitudinal Fasciculus (SLF) as it directly connects the IPL and Pars ROIs. Fourth, we selected the Uncinate (UN) as it connects the Pars and Hipp ROIs.

Electrical Field Modeling.
Current modeling of the tACS was performed using the Realistic vOlumetric Approach to Simulate Transcranial Electric Stimulation (ROAST) software toolbox for Matlab to map electrical field (EF) changes throughout the cortex. ROAST is an open-source MATLAB-based, automated pipeline that applies SPM segmentation to the head and neck. Following segmentation, typical isotropic electrical conductivities are assigned to the tissues and electrodes, typical boundary conditions are assigned to the surfaces, and simulation of current flow is achieved by solving the Laplace equation ($\nabla \cdot (\sigma \nabla V) = 0$), where V is potential and σ is conductivity. Current modeling was conducted on the skull-striped cortex of each participant's baseline T1 and T2 scan then fit to the MNI-152 standard head.

Exploratory Neural Analyses.
To measure changes in neural metrics, we assessed changes in RSFC (pairwise connectivity between Pars, Hipp, and IPL ROIs), and MRS-based E/I within in voxels of interest (Pars and Hipp). Changes in neural measures were assessed with non-

parametric Wilcoxon signed-rank tests. To link behavioral tACS changes with underlying neural changes, we conducted correlations between the outcome metrics and neural metrics that had significant changes in the prior analyses. All correlation analyses were conducted using non-parametric Spearman's correlations (rho).

References

1. Liu, Q. *et al.* Intracranial alternating current stimulation facilitates neurogenesis in a mouse model of Alzheimer's disease. *Alzheimers. Res. Ther.* **12**, 1–12 (2020).

2. Viskontas, I. V., Knowlton, B. J. & Fried, I. Responses of neurons in the medial temporal lobe during encoding and recognition of face-scene pairs. *Neuropsychologia* (2016) doi:10.1016/j.neuropsychologia.2016.07.014.

3. Wais, P. E., Arioli, M., Anguera-Singla, R. & Gazzaley, A. Virtual reality video game improves high-fidelity memory in older adults. *Sci. Rep.* **11**, 2552 (2021).