Report for Dr. Rosenfeld

**HI/HA project**

Andy Revell, Monday December 27th, 2021

A picture containing indoor, dark

Description automatically generated

# Table of Contents

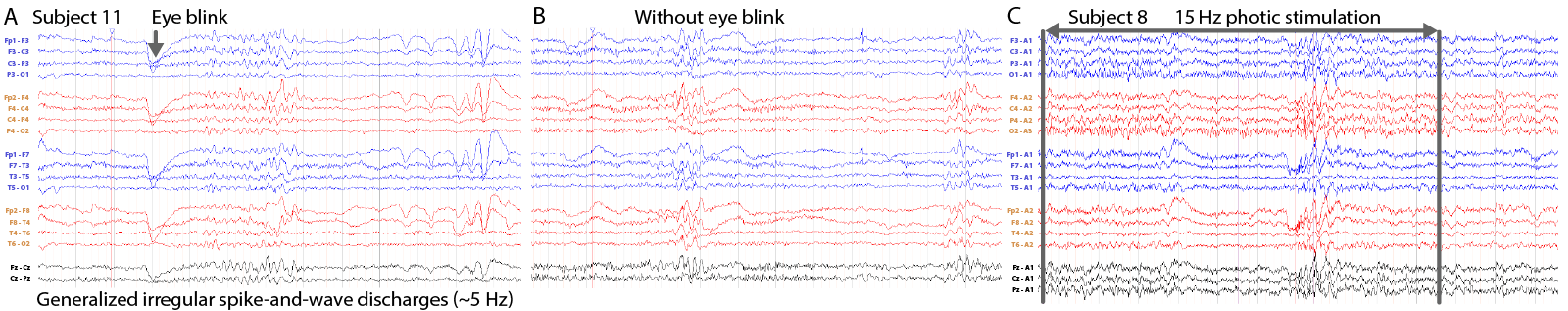
1. Introduction and rationale
2. Methods
3. Results

# Introduction

* **Scientific question**:
  + Are there differences in GluCEST measurements between HI/HA subjects and a healthy reference population?
* **Hypothesis**:
  + GluCEST **measurements** between HI/HA subjects and a healthy reference population are different.
  + We can quantify the differences in GluCEST **measurements** by:
    - Quantifying the differences in means
    - Quantifying the differences in variation (standard deviation)
  + Differences in means
    - Due to the prevalence of seizures in HI/HA syndrome and given the evidence for increased GluCEST values in the affected hippocampus of TLE patients [Davis 2015], it is hypothesized that GluCEST values may be increased in the hippocampus of HI/HA syndrome.
  + Differences in variation
    - However, the etiology of epilepsy in HI/HA patients is different than the TLE subject studied previously, thus we also study the variation of GluCEST values to (1) better understand the neuropathophysiological mechanisms underlying the unique phenotype of HI/HA syndrome and (2) determine the applicability of GluCEST as a biomarker for the [blank] of HI/HA syndrome
      * I don’t know enough about HI/HA syndrome to know if GluCEST can be used as a useful biomarker, but the paper could have this added spin. “Biomarker” has been a big buzzword for the last few years, and it could increase the paper’s citation rate if spun that way. Idk if the variable phenotype of GluCEST patients is a problem clinically (i.e., is the diagnosis hard to make, or does the phenotype from one patient to the next look very different to a clinician?), but maybe GluCEST can be used to better explain that variation. It would be interesting to see GluCEST in HI/HA patients without epilepsy.
* **Scientific contribution** of this study:
  + As stated above,
    - Better understand the pathophysiology of HI/HA
    - Determine the applicability of GluCEST as a biomarker

## Methods

* I got all data from Alfredo
  + T2 scans, b0b1 corrected CEST, and hippocampus segmentation
  + **Controls** (“healthy reference group”: C004-C015.
    - Threw out C012 (no CEST acquired) and C014 (CEST slice fell completely outside of the segmented hippocampus). This brings total to N=10
  + **HI/HA subjects**: CHOP\_01 to CHOP\_12
    - CHOP\_01 and CHOP\_02 threw out (did not even attempt to analyze because the paper draft said that motion artifact was present, and alfredo did not analyze either. Did not confirm myself.)
    - CHOP\_05: there was one patient labeled CHOP\_04\_05 and I considered them as “CHOP\_05”. The original CHOP\_05 folder did not have CEST images and there was already a CHOP\_04 folder with a complete dataset. So I assumed CHOP\_04\_05 was the actual CHOP\_05.
    - CHOP\_09: there was no existing folder.
    - CHOP\_12: folder was given to me, but it was empty (no data inside)
    - This brought the total N=8
      * CHOP\_03, CHOP\_04, CHOP\_04\_05 (“CHOP\_05), CHOP\_06, CHOP\_07, CHOP\_08, CHOP\_10, CHOP\_11
* **CEST measurements**
  + I took the b0b1 corrected CEST images and performed three preprocessing steps:
    - The T2 + hippocampus segmented images are not in same plane/dimensions as the 2D GluCEST, so we must re-slice them using c3D
      * c3D reslice-identity command with interpolation as NearestNeighbor
      * <https://sourceforge.net/p/c3d/git/ci/master/tree/doc/c3d.md#-reslice-itk-resample-image-using-affine-transform>
    - Once the GluCEST image and hippocampus segmentation image were in the same imaging space, I used the hippocampus image as a mask to analyze the CEST values within only the hippocampus (for analyses looking at just the hippocampus)
  + Excluded CEST values below 0 and above 20. These are artifact values, but the cutoff is slightly arbitrary. There are relatively few pixel values above or below this threshold (i.e., they are outliers), but they may alter means and medians.
    - They were converted to NaNs, and when computing means and medians, NaNs were excluded (using python Numpy package commands np.nanmean and np.nanmedian).
  + Measurements collected:
    - ﻿**cest\_total\_mean**: The mean CEST values of the entire CEST slice (not just within the hippocampus).
    - **cest\_total\_median**: The median of the entire CEST slice
    - **cest\_total\_std**: The standard deviation of the entire CEST slice
    - **cest\_total\_pixels**: The total number of pixels captured
    - **cest\_hipp\_mean**: The mean of the CEST values within just the hippocampus. The values from both the right and left hippocampi are pooled
    - **cest\_hipp\_median**: The median of the CEST values within just the hippocampus
    - **cest\_hipp\_std**: The standard deviation of the CEST values within just the hippocampus
    - **cest\_hipp\_pixels**: The number of pixels captured within just the hippocampus
    - **cest\_left\_mean**: Same as cest\_hipp\_mean, but within just the left hippocampus.
    - **cest\_left\_median**: Same as cest\_hipp\_median, but within just the left hippocampus.
    - **cest\_left\_std**: Same as cest\_hipp\_std, but within just the left hippocampus.
    - **cest\_left\_pixels**: Same as cest\_hipp\_pixels, but within just the left hippocampus.
    - **cest\_right\_mean**: Same as cest\_hipp\_mean, but within just the right hippocampus.
    - **cest\_right\_median**: Same as cest\_hipp\_median, but within just the right hippocampus.
    - **cest\_right\_std**: Same as cest\_hipp\_std, but within just the right hippocampus.
    - **cest\_right\_pixels**: Same as cest\_hipp\_pixels, but within just the right hippocampus.
    - **hipp\_total\_volume**: The total hippocampus volume. This is calculated from the hippocampus segmentation image (the 3D volume). This helps us see if the gross hippocampus measurements are different between the groups.
    - **hipp\_left\_volume**: The total left hippocampus volume
    - **hipp\_right\_volume:** The total right hippocampus volume
    - **﻿peak\_cest:** The higher value between **cest\_left\_mean**  and **cest\_right\_mean**
    - **peak\_cest\_zscore:** The mean and standard deviation was calculated for a normal population from the healthy reference subjects. The **peak\_cest\_zscore**  for each subject was calculated from this mean and standard deviation [ (**peak\_cest**  - mean of healthy)/standard deviatiobn of healthy]
  + See spreadsheets “CEST\_measurements.csv” for the measurement outputs
* **Reproducibility:**
  + See GitHub for access to code and anlysis
    - [**https://github.com/andyrevell/HI\_HA\_GluCEST\_2022**](https://github.com/andyrevell/HI_HA_GluCEST_2022)
    - There are 3 main folders:
      * Analysis, Data, and Plots.
      * The data folder is empty for obvious reasons. Left a note saying to contact the corresponding author for any questions.
      * Plots folder contains all the raw plot outputs from Python
      * Analysis folder



**Fig. 1. Representative EEG tracings. | (A)** EEG data from subject 11 demonstrating generalized irregular spike-and-wave discharges at ~5 Hz associated with eye blink (arrow) and **(B)** without eye blink. **(C)** EEG data from subject 8 displays photosensitive generalized and irregular spike-and-wave discharges. Double-headed, horizontal arrow denotes duration of 15 Hz photic stimulation. Normal background EEG activity is observed in both subjects.

* + - * + Analysis.py is the main analysis pipeline
        + Helper.py contains ancillary functions to help with the analysis
        + Constants.py contains any constant values used in the analysis (like colors for the different groups, or the 0-20 CEST threshold)
        + Quick\_measurements.py was from Alfredo to get a start on doing quick CEST measurements
  + Python and package versions
    - Python 3.8.12
    - numpy (1.21.2), pandas (1.2.4), seaborn (0.11.2), nibabel (3.2.1), scipy (1.7.1), statsmodel (0.13.1), matplotlib (3.3.4)
* **Statistics:**
  + Mann Whitney U test was performed when comparing GluCEST measurements (above) between the HI/HA and the healthy reference subjects. Bonferroni correction was applied to correct for multiple comparisons (20 tests performed with a new alpha cutoff as 0.05/20 = 0.0025).
  + Because the GluCEST slice acquisition was manually positioned along the hippocampus, we wanted to see if the size of the hippocampus captured (i.e., the number of pixels, or **cest\_hipp\_pixels**) was related to the mean or variation of CEST values (the **peak\_cest** and **cest\_hipp\_std**, respectively).
    - We fit a linear regression using OLS (ordinary least squares, from the python package statsmodel). Th number of pixels (**cest\_hipp\_pixels**) was used as a coefficient to predict the dependent variable (either **peak\_cest** or **cest\_hipp\_std)**. The p-value of the coefficient was notd from the output of statsmodel.
  + See “CEST\_statistics\_table.csv” for statistics outputs

## Results

* See figures and captions. Hopefully the captions provide enough detail for you to follow and can be cut down as you see fit.
* **Figure 1:** EEG
* **Figure 2:** Overview of GluCEST acquisition and analysis
* **Figure 3:** HI/HA subjects show higher GluCEST variability within the hippocampus compared to a normal, healthy reference population that underwent the same MRI acquisition protocol.

Application

Description automatically generated

**Figure 2. Overview of GluCEST acquisition and analysis. | (A)** GluCEST is measured on an axial slice at the level of the hippocampus using the same MRI acquisition protocol in eight HI/HA subjects and ten healthy reference subjects. Measurements are limited to a single slice (i.e., a single layer of voxels, or equivalently, pixels in 2D). The hippocampus was segmented using the ASHS pipeline in ITK-SNAP. A 3D render of the hippocampus of an example healthy reference subject (C006) is shown along with the 2D GluCEST slice of the same subject (note, the 3D render is used for visualization purposes only). **(B)** The CEST values within the hippocampus are overlaid on the subject’s T2 image from panel A. **(C)** The distribution of CEST values (top, histogram; bottom, ECDF plot) in each hippocampus is shown corresponding to the same subject of panels A and B. The solid vertical line represents the mean of all pixel values in the hippocampus, and the dashed vertical lines represent ± 1 st. dev. **(D)** Heatmap of CEST values are shown for all HI/HA subjects (bottom) and two representative healthy reference subjects (top, C010 and C011). CEST values for both the entire axial slice and within only the hippocampus is shown. Similar to the representative plots of panel C, the distribution of hippocampus CEST values is also shown. Note that CEST values were rounded to the nearest integer at time of acquisition for subject 7, however, the subject was not excluded because the mean and variation could still be reasonably approximated given the data. Subjects 1 and 2 had motion artifacts, subject 9 had an intolerance of the MRI scan, and subject 12 did not acquire all MRI scans to run the full analysis pipeline. These subjects were excluded from analyses. **st. dev.**, standard deviation; **ASHS**, Automatic Segmentation of Hippocampal Subfields; **CEST**, chemical exchange saturation transfer; **ECDF**, empirical cumulative density function.

* **Supplementary Table 1:** Statistics of CEST measurements

Graphical user interface, application

Description automatically generated

**Fig. 3. HI/HA subjects show higher GluCEST variability within the hippocampus compared to a normal, healthy reference population that underwent the same MRI acquisition protocol. | (A)** The variation of CEST values is higher in the HI/HA subjects compared to the healthy reference subjects (p < 5 x 10-5, Mann Whitney U test, N =10 healthy reference, N = 8 HI/HA subjects). In the box and whisker plots, horizontal lines represent medians, box ends represent the 25th and 75th percentiles, and whiskers extend to the minimum and maximums. To the right of each box and whisker includes the mean of each group and 95% CI error bars. **(B)** The difference in CEST variation between the two groups is not due to differences in the amount of the hippocampus captured in the 2D MRI slice of each group. The number of pixels, or equivalently, the number of voxels of the single layer 2D slice, is equal between the two groups (left plot, p > 0.05, Mann Whitney U test). The size of the hippocampus, or the number of pixels, captured is also not associated with the standard deviation of each subject (p > 0.05, middle plot) after fitting a linear regression using OLS with the number of pixels captured for each patient used as a coefficient. Gray bands represent 95% CIs. The appropriateness of using OLS is shown with a plot of the residuals (right plot, dashed line at zero). **(C)** The peak CEST values are similar between each group (p > 0.05, Mann Whitney U test) showing that, while CEST values are more variable in HI/HA subjects, the means are approximately equal to the reference group. Peak CEST is the highest mean CEST value between the left or right hippocampus of each subject. Three subjects, however, are outliers (≥ 2 st. dev.) with respect to their peak CEST values using the healthy reference subjects as the reference population distribution. The middle plot shows that Peak CEST is not related to the size of the hippocampus captured (using the same methods as panel B), and the right plot shows the appropriateness of using OLS. **(D)** Summary of findings. HI/HA subjects have higher variation of hippocampus CEST values (“cold and hot spots”) with approximately equals means compared to a healthy reference population. However, three outlier subjects have higher means (subject 4 = 3.5 st. dev.; subject 5 = 5.3 st. dev.; subject 7 = 3.2 st. dev. above the mean), indicating this study may be underpowered to detect a true difference in means or accurately model the distribution of CEST values in HI/HA syndrome (e.g., a bimodal distribution with a subset of the population having higher means). **St. Dev**., standard deviation; **CI**, confidence interval; **HI/HA**, Hyperinsulinism Hyperammonemia, **CEST**, chemical exchange saturation transfer; **OLS**, ordinary least squares.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| measurement | control\_mean | control\_sd | HIHA\_mean | HIHA\_sd | pvalue |
| cest\_total\_mean | 7.25583219 | 0.35828317 | 7.07967493 | 0.19335154 | 0.35993418 |
| cest\_total\_median | 7.50856731 | 0.37577889 | 7.12383561 | 0.34762865 | 0.04342063 |
| cest\_total\_std | 2.61314345 | 0.11174773 | 2.93905122 | 0.21648078 | 0.0030623 |
| cest\_total\_pixels | 19679 | 1028.7749 | 18834.5 | 1140.75556 | 0.17281411 |
| cest\_hipp\_mean | 8.87407713 | 0.62346467 | 8.7958783 | 0.51273152 | 0.82855706 |
| cest\_hipp\_median | 9.02142853 | 0.56691383 | 8.70180319 | 0.44668508 | 0.23698524 |
| cest\_hipp\_std | 1.50218553 | 0.20779518 | 2.62600886 | 0.62575928 | **4.57E-05\*\*\*** |
| cest\_hipp\_pixels | 983 | 222.577627 | 950.375 | 269.732246 | 0.69646693 |
| cest\_left\_mean | 8.99574696 | 0.60431182 | 8.43991173 | 1.4827281 | 0.20307144 |
| cest\_left\_median | 9.17327132 | 0.55856734 | 8.58986947 | 1.34673376 | 0.23698524 |
| cest\_left\_std | 1.42703356 | 0.22078223 | 2.19665693 | 0.24252311 | **4.57E-05\*\*\*** |
| cest\_left\_pixels | 509.8 | 97.650192 | 483.875 | 127.151915 | 0.63343846 |
| cest\_right\_mean | 8.7345822 | 0.83169373 | 9.13900953 | 1.67215404 | 0.96540061 |
| cest\_right\_median | 8.8710829 | 0.76926546 | 9.22343626 | 1.68011689 | 1 |
| cest\_right\_std | 1.4822575 | 0.23030453 | 2.21335703 | 0.11067366 | **4.57E-05\*\*\*** |
| cest\_right\_pixels | 473.2 | 134.655709 | 466.5 | 158.134436 | 0.96540061 |
| hipp\_total\_volume | 2.95762435 | 0.40463055 | 3.00265969 | 1.06679306 | 0.63343846 |
| hipp\_left\_volume | 1.5198758 | 0.21730149 | 1.51776686 | 0.51552927 | 0.82855706 |
| hipp\_right\_volume | 1.43774855 | 0.19380423 | 1.48489283 | 0.5616924 | 0.76182641 |
| peak\_cest | 9.19863922 | 0.60175793 | 9.87831086 | 1.40356961 | 0.69646693 |

Supplementary Table 1. Statistics of CEST measurements | All CEST measurements calculated with the means and standard deviations of the healthy reference subjects (the “controls”) and HI/HA subjects. P-values were calculated using the Mann Whitney U test, and \*\*\* indicates statistically significant tests after the new alpha threshold of 0.0025 from Bonferroni correction of 20 tests (0.05/20). Note no statistically significant differences in the 3D volume of the hippocampus. **STD** or **SD**, standard deviation.