Hi everyone, I'm here to talk to you about my poster, "Validating multi-echo fMRI analysis methods across a range of acquisitions".

We'll start with the **Introduction**. From our previous work with multi-echo, we've seen that validation of multi-echo data analysis methods benefits from the presence of a dataset with predictable signal changes across multiple brain regions and acquisition options. And this is important because there has been a growing interest in the use of multi-echo for a wide variety of applications, and a corresponding increase in options for multi-echo data analysis, including new approaches being developed through the tedana software collaboration.
There are some openly available multi-echo datasets, but only a few use tasks, which have expected responses that can be useful for validation, as well as a range of acquisition options to make sure an algorithm is not optimizing for just one type of scan. We are collecting a dataset with a consistent task and several acquisition options to aid in the process of validating and comparing existing and new changes to data preprocessing and analysis methods.

Moving on to the **Methods** section, one of the things we are aiming to do with this dataset is compare data preprocessing methods. So for this analysis, we processed our data in three different ways. The first method was using a *single echo*. I'll talk more about the task design later, but our dataset includes both 3-echo and 5-echo data, so our single-echo preprocessing stream involved using only the 2nd echo out of these data. One thing to note is that the second echo is near the peak T2* weighting for the 3-echo data, but not the 5-echo data, so that's a caveat to consider and it'll help explain some of the results that we show later on in the poster.
We also *optimally combine* the echoes. This is a weighted average of all the echoes and it helps optimize T2* weighting, which has been demonstrated to give a reliable boost in data quality. If you direct your attention to Figure 1, you'll see that the optimally combined images to the very right of the figure help combat the signal dropout that we see in some of the individual echo images to the left.
And finally, we *denoise* the data with the tedana pipeline. This involves independent component analysis, which regresses out components classified as noise. If you direct your attention to figure 2, you'll see that ICA has separated an example dataset into different components, 3 of which are shown in the figure. That makes it possible to remove the components that are non-BOLD, which we classify as noise.
And our last step was to compare brain activation and task response across these three types of preprocessing in multiple subjects using a group analysis.

Next, we'll move on to **Task Design.** We've had 13 healthy volunteers who performed a language localizer task in which they saw or heard 4-second groupings of audio or visual stimuli. The visual stimuli were either written English words or false font words, and the audio stimuli were either spoken English words or vocoded sounds derived from words. The subjects were then instructed to press a button if the same word or nonword appeared twice in a grouping of four. This was a jittered event-related design in which each of the four conditions appeared 34 times over 22 minutes. From this task, we were able to derive two different contrasts; a word-nonword contrast, and a visual-audio contrast. This means that a spatially diffuse number of brain regions will show activation in response to this task, and validation of multi-echo methods will include a range of response magnitudes and signal dropout regions. We also had 3 different acquisition types, including 3-echo and 5-echo data with different voxel sizes. We did not

include our 3-echo data with the smaller voxels in the results because we had fewer volunteers complete this last acquisition type due to time constraints.

Now, we'll move on to **Results**. As I mentioned before, our task gives us both a visual-audio and a word-nonword contrast. Figure 4 shows group maps from our 13 subjects and we can compare the t-statistic of brain activation across our three preprocessing types. We see that in general, single-echo data shows us less robust activation than optimally combined and denoised data, for both contrasts. But one thing to note is that single-echo data is significantly worse in the 5-echo maps, and that's because we looked at the 2nd echo instead of the 3rd echo, which would have been closer to the optimal BOLD weighting.
If you then look at Figure 5, we have dot plots for the two contrasts as well. We took our t-statistics and transformed them to Fisher-z values, which is a rough measure of contrast-to-noise because it's a measure of how big the signal is compared to the unmodeled variance. So in these dot plots we see average Fisher-z values in significant ROIs from the group maps. We can see that effect magnitudes vary across ROIs, and there's some variability in the 3-echo vs. 5-echo data. But one important takeaway is that values tend to be slightly larger after optimal combination and denoising, which is what we expect.

Now, we'll discuss the **Conclusions.** From the group maps and the dot plots, we see that the task generates robust and consistent contrasts with distinct acquisition parameters. In this analysis, we examined the 3 and 5-echo data both with the 3.3mm^3 voxel size, but as I mentioned before, we also have 3-echo data with a smaller voxel size that we can run the same group analysis on in the future. And because we see these robust and consistent contrasts, we can use this dataset to validate and compare different data preprocessing streams, like the ones being developed in the tedana software collaboration.

Finally, we'll address **Future Steps.** First, we want to continue evaluating current analysis pipelines using these data and we'll be increasing the sample size in this process as well. We also want to fully document and share these data, which involves converting them into BIDS format. And lastly, we want to make this data an option in the tedana integration and testing process.

To conclude, we would like to thank the NIH HPC group, as our data were processed using HPC Biowulf resources. We would also like to thank Laura Mesite and Sibylla Leon Guerrero for the PsychoPy script that we adapted for this study. I would like to thank everyone in the Section on Functional Imaging Methods at the NIH for their contributions to this project. And thank you for listening to this presentation!