

Phonatory Analysis on Parkinson's Disease Patients Attending Singing and Discussion Therapy (Parkinsonics) using Signal Processing Techniques

Casey Chen, Laureano Moro-Velazquez, Alex Ozbolt, Ankur Butala,
Alex Pantelyat, Najim Dehak

Motivations

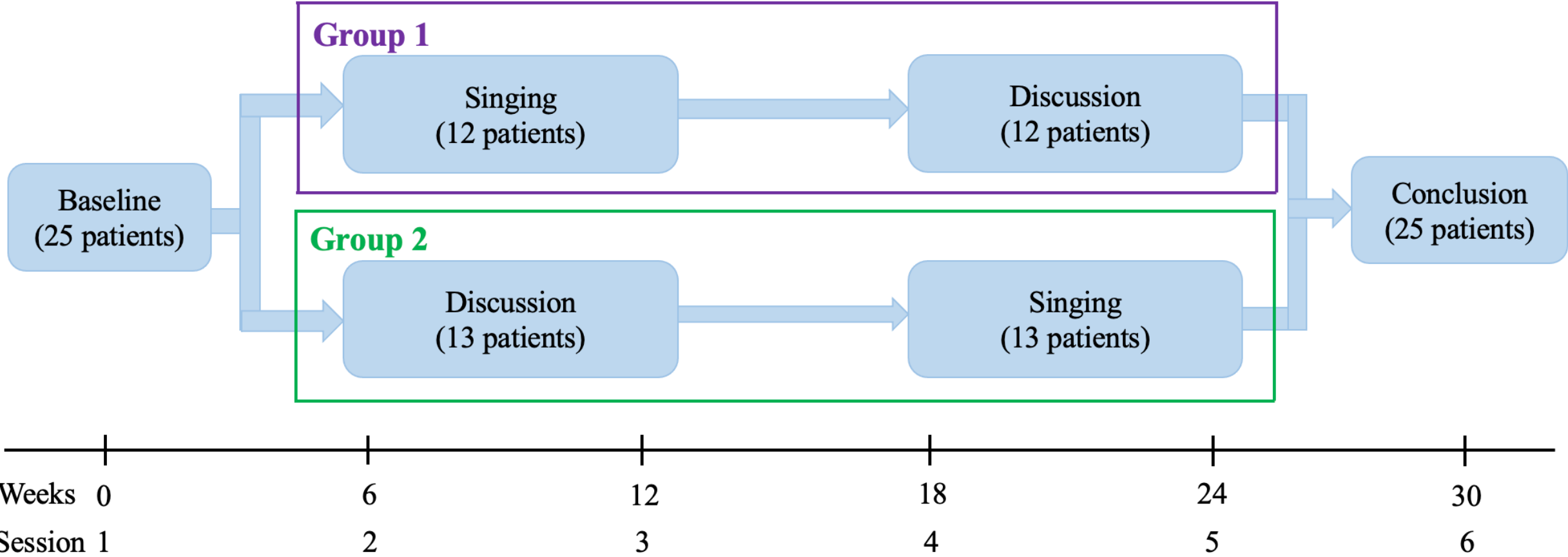
- Parkinson's Disease (PD) is a neurological condition that leads to visual and speech impairments
- In order to relieve some of these symptoms, PD patients may attend different types of therapy

Goals

1. Identify if there are signal processing-based features that can track changes in the voice of PD patients along time
2. Analyze how these features change with therapy and the advance of the disease

Materials

- Parkinsonics data set contains audio recordings of 25 patients who attended speech and singing therapy

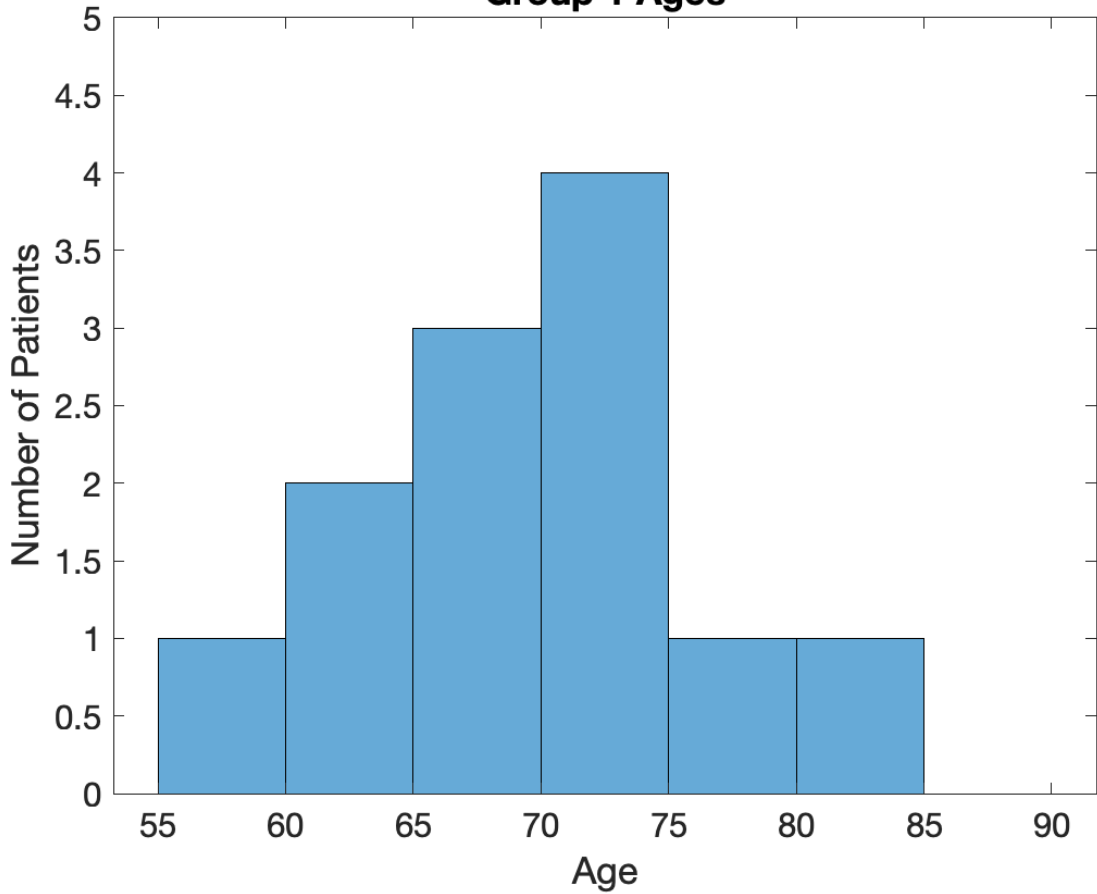


Timeline of the different types of therapy that the patients attended and the times at which the recordings were taken

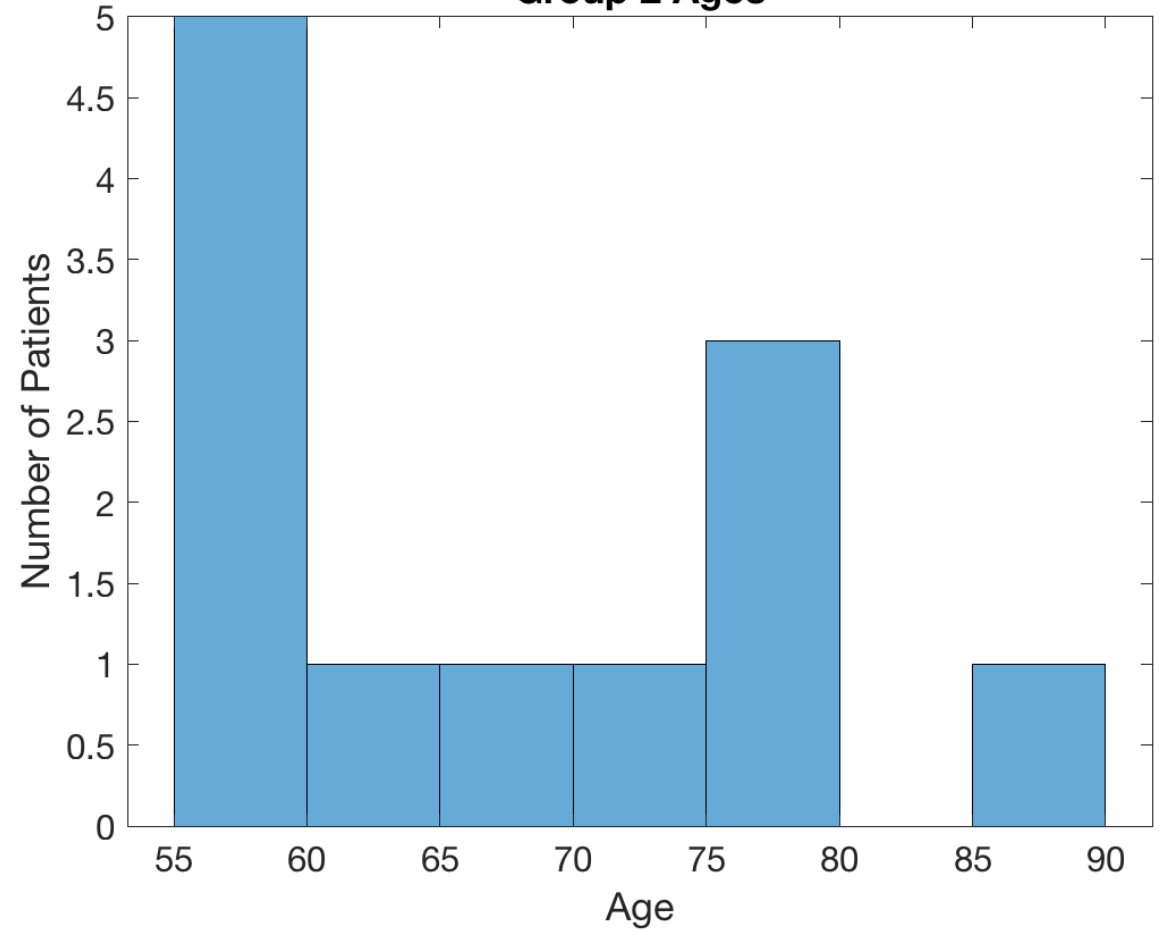
Recording Sessions

- Each recording session involved patients holding the vowel sound /a:/ and /e:/ for as long as possible in one breath
- 3 trials at high intensity
 - A loud
 - E loud
- 3 trials at a normal speaking volume
 - A norm
 - E norm

Group 1 Ages



Group 2 Ages



Comparison can only be done within a group and not between groups because the two age distributions for the groups do not have the same distribution.

Methods of Analysis

- AVCA (Automatic Voice Condition Analysis) library
 - Produced feature vectors of dimension 69
- Kruskal Wallis test and the Wilcoxon Rank Sum test
 - Applied the false discovery rate correction
- Calculated the slopes of the features' trends over the six sessions

Table 1. AVCA library features employed in this study

Feature family	Coefficients
Amplitude and frequency perturbation and fluctuation	Absolute and relative jitter and shimmer, RAP, PPQ5, APQ3, APQ5, FTRI, ATRI, and statistics about HNR, NHR, CHNR, NNE, and GNE
Spectral-Cepstral	LHr
Complexity	D2, LZC, and statistics about LLE, ApEn, SampEn, GSampEn, FuzzyEn, mSampEn, PE, RPDE, and DFA
Modulation Spectra	MSP, and statistics about MSH, MSW, CIL, RALA, and LMR

Results

- Coefficient column: shows the significant features
- Average change column: average difference of patients' vocal feature values between session 6 compared to session 1
- Standard deviation column: the standard deviation of the average change
- Slope column: slope of each patients' linear trend line of their average coefficient values for sessions 1 through 6

Group 1 Results

Coefficient	Average Change	Standard Deviation	Slope
A Norm			
-	-	-	-
A Loud			
CHNR_std	-1.08	0.95	-0.18
CHNR_mean	-1.03	0.92	-0.17
E Norm			
CHNR_std	-0.9	1.39	-0.19
CHNR_mean	-0.82	1.36	-0.17
E Loud			
CHNR_std	-1.13	1.12	-0.24
CHNR_mean	-1.09	1.03	-0.22
GNE_mean	-0.26	1.46	-0.067

All of these features are statistically significant.

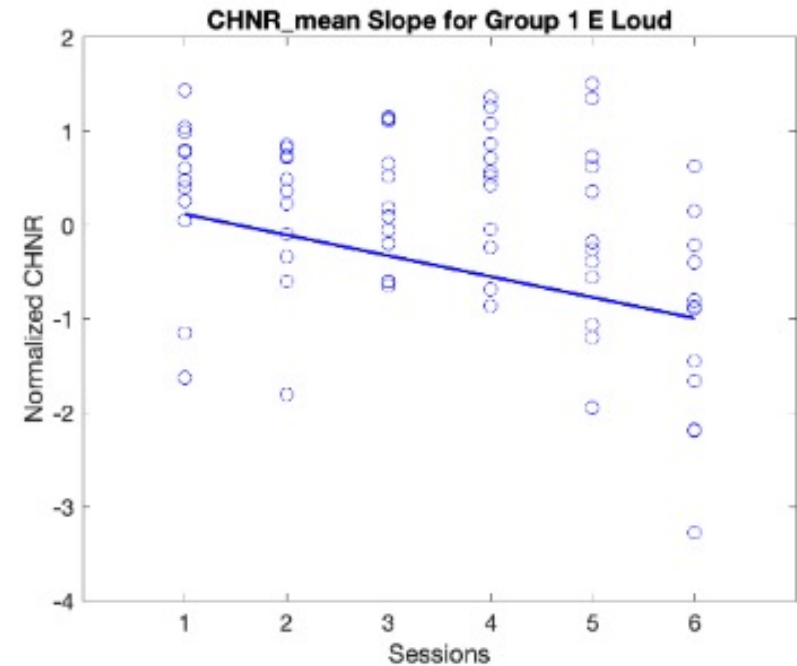


Figure 3. Slope of CHNR mean for E Loud Group 1. This figure depicts how the mean CHNR decreases with time for Group 1.

Group 2 Results

Coefficient	Average Change	Standard Deviation	Slope
A Norm			
LMR_std*	-0.82	1.15	-0.087
A Loud			
-	-	-	-
E Norm			
LMR_std	-1.23	0.99	-0.15
GNE_mean	-0.79	1.26	0.0027
CHNR_std	-0.74	0.92	-0.15
MSHmod_mean	-0.65	1.31	-0.094
E Loud			
CHNR_std	-1.03	1.20	-0.17
CHNR_mean	-0.94	1.07	-0.15
GNE_std	0.82	1.22	0.14
GNE_mean	-0.77	1.09	0.056
LMR_std	-0.81	0.90	-0.071

All of these features are statistically significant.

*Negative correlation of ≤ -0.5 between change in TSM and change in feature values

Discussion

- Significant features for both groups:
 - cepstral-harmonics-to-noise ratio (CHNR) mean and standard deviation
 - glottal-to-noise excitation ratio (GNE) mean
- For group 2, the Modulation Spectrum Homogeneity (MSHmod) mean decreases
- For CHNR, GNE, and MSHmod, the means decreased when comparing session 6 to session 1 which indicates more audible noise in patients' voices and a breathier voice

Conclusions

- Results suggest that therapy did not have a significant impact in most of the vocal features in the AVCA library
- Only features measuring vocal noise were significant when comparing pre- and post-therapy recordings
 - The voice of participants tends to be noisier with time
- Cannot conclude that therapy is completely ineffective
 - No control group
- In future work, additional speech-related features should be employed
 - Use connected speech instead of sustained vowels

Questions?