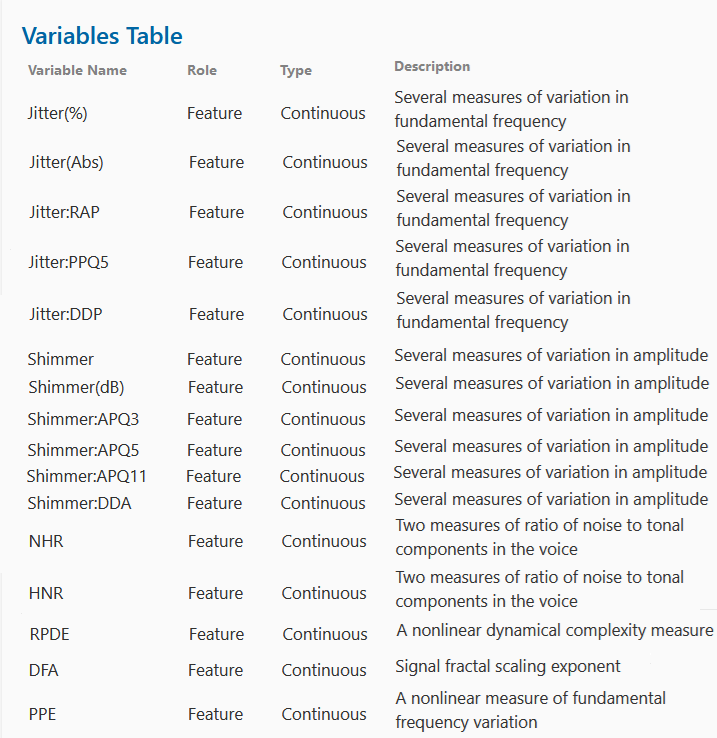
DATA

The data used for our analysis came from STUDY. It consists of 22 columns. Two are identifiers to distinguish patient and specific trial. Two of the data columns are demographical data, age and sex. Two columns are motor and total UPDRS which represent clinician measurement of symptoms and are considered source of truth. The remaining 16 columns encode jitter, shimmer, and 5 other numeric values which describe characteristics of voice thought to be related to Parkinson’s Disease symptom severity.



Custom built equipment was used to collect data from ### distinct patients over ## trials per patient. There are no missing data points.