

# An Examination of the Impact of Cerebromicrovascular Disease on Elderly People with Diabetes

---

Syed Ashemi & Robert Harder & Aiden Van Greuning

# Background Information

- **Cerebromicrovascular disease (CMVD):** Result of significant damage to the brain's small blood vessels, often caused by diabetes.
- **ECG measurements:** Used to indirectly detect diabetes by identifying related cardiovascular abnormalities.
- **Dataset used:** CVES database with 120 subjects (60 controls and 60 diabetics), aged 55-75. ECG data extracted for 13 random controls and 15 diabetics at 1000 Hz sampling over 24 hours.
- **Best results in open literature [1]:**
  - Diabetic subjects showed lower HRV and prolonged QT intervals.
  - 48% of diabetic subjects with CMVD had abnormal ECG results, while 19% of diabetic subjects without CMVD had abnormal ECG results

# Preprocessing Methods

- **Notch Filtering:** Remove 60Hz power line interference noise for control and diabetic groups
- **Pan-Tompkins Algorithm (PT):** Original ECG signal (control & diabetic) → Bandpass Filter → Derivative Filter → Squaring operator → Moving Average Filter
- To remove high/low frequency noises, muscle contraction interference, power-line noise, baseline wander, smoothen ECG signals by averaging a set of neighboring data points
- Done to make it easier to extract ECG features (**PQRST**)

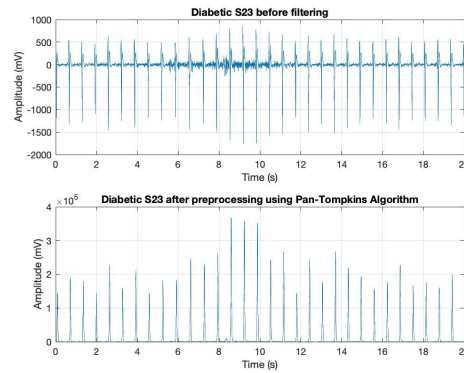
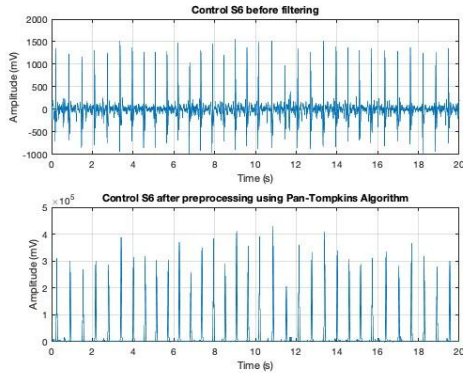
# Feature Analysis

- **PQRST Points:** used ECG toolbox to locate each point based on a predetermined set of criteria
- **PQRST Intervals:** based on the individual points found using ECG toolbox, intervals between points were calculated
- Mean, standard deviation of R-R, S-T, Q-T intervals
  - Using mean, standard deviation equations
- Heart Rate
  - Implementing RR interval into heart rate equation
- Average QRS length
  - Averaging lengths of QRS complexes in a given signal

# Machine Learning

- **Logistic Regression:** supervised technique based on logistic function for two group classification
  - Greater accuracy compared to Linear Discriminant and Decision Tree Models, fits unique S-shaped predictor line to map class labels, can precisely identify and map feature weight and importance
- **5 fold cross validation:** evaluation of model where data is split into 5 ( $k = 5$ ) folds and model trained on  $k - 1$  folds with one fold left to test a model, repeated  $k$  times, bias lessened by averaging accuracy
- Chosen over 10-fold cross validation due to smaller data pool
- **Confusion matrix:** defines overall performance of model by evaluating number of true negatives, false positives, false negatives and true positives in 4 by 4 table
- 0 represents control group and 1 represents diabetic group

# Results



Confusion Matrix:

```

7    6
5    10
    
```

Sensitivity: 66.67%

Specificity: 53.85%

Classification Accuracy: 60.71%

AUC (Area Under Curve): 0.60

Signal	NumberofBeats	RRInterval	SDofRRInterval	AvgQRSLength	HR	MeanSTInterval	SDofSTInterval	MeanQTInterval	SDofQTInterval
ECGS6	102	589.57	40.508	35.912	101.77	114.07	30.766	242.63	32.45
Signal	NumberofBeats	RRInterval	SDofRRInterval	AvgQRSLength	HR	MeanSTInterval	SDofSTInterval	MeanQTInterval	SDofQTInterval
ECGS23	95	633.53	14.84	33.493	94.707	162.52	49.694	336.01	54.311

Overall Testing Accuracy (Testing Percentage): 54.00%

Overall Validation Accuracy (Validation Percentage): 62.25%

# Reasoning of Results

- **ECG values** for both groups aren't as accurate
- **Diabetic patients** should have a higher HR, QRS complex, ST interval, QT interval
- **Control patients** should have a higher RR interval
- Due small sample size and inaccurate calculation, **5 fold cross validation** values aren't as accurate
- **Confusion matrix** is semi accurate (FP & FN too high)
- **Best results** in open literature [2]:
  - Test accuracy (80.86%)
  - Validation accuracy ( $\approx$  75-85%)
  - Sensitivity (46.34%)
  - Specificity (86.66%)
  - Classification accuracy ( $\approx$  66.5%)
  - AUC (0.7932)

# References

- [1] Q. Qiu et al., “Heart rate variability is associated with cerebral small vessel disease in patients with diabetes,” *Frontiers in neurology*, <https://pmc.ncbi.nlm.nih.gov/articles/PMC9685533/> (accessed Nov. 24, 2024).
- [2] Xie, Z., Nikolayeva, O., Luo, J., Li, D., “Building Risk Prediction Models for Type 2 Diabetes Using Machine Learning Techniques”, U.S. Centers for Disease Control and Prevention, [https://www.cdc.gov/pcd/issues/2019/19\\_0109.htm#](https://www.cdc.gov/pcd/issues/2019/19_0109.htm#) (accessed Nov. 24, 2024).