Team 15: Featurization of long term CGM data to characterize menstrual cycle patterns in female patients with diabetes

Jane Li[1], Nathaniel Tillison[1], Jeyasri Venkatasubramani[1]
[1] Department of Bioengineering, University of California, San Diego
Advisors: Dr. Benjamin Smarr, Saara Kriplani

dexcom

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Shu Chien-Gene Lay Department of Bioengineering

Background

Diabetes Mellitus (DM): Metabolic disease where blood glucose levels are abnormally regulated.

- **589 million adults worldwide**[1]
- Continuous Glucose Monitoring System (CGM) to monitor glucose levels continuously

Estrogens and Glucose: Higher estrogens are associated with decreasing glucose levels and increased insulin sensitivity[3]

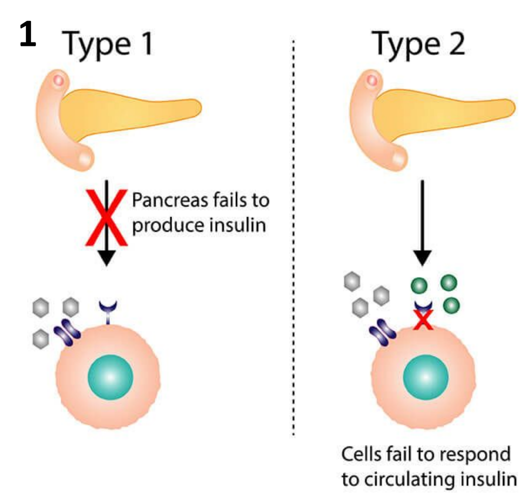


Figure 1: DM mechanism in DM 1 & DM 2[2]

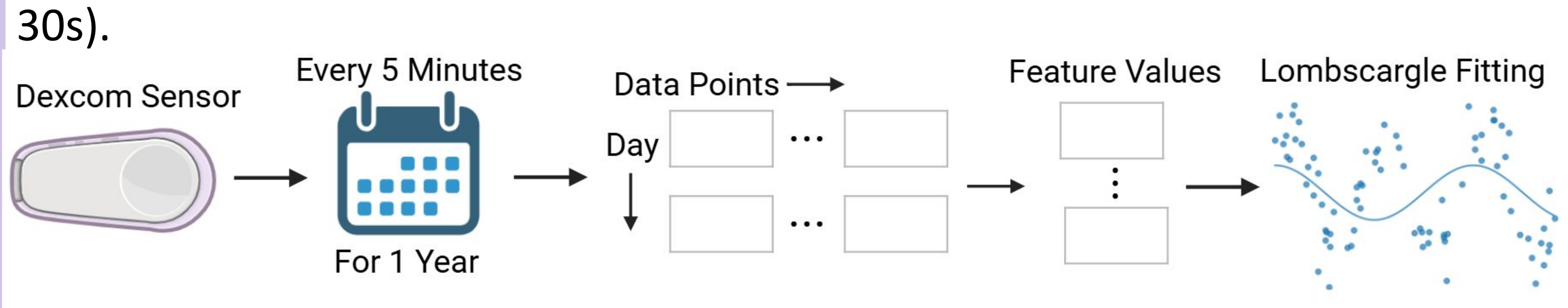
Long term biological rhythm sex differences in DM → contribute to personalized treatments for women based on menstrual cyclicity status

- Multiple studies have shown **phase based glucose changes**: increasing in ovulation and luteal phase, dropping in late-follicular phase[4,5,6]
- 2025's Team 29 discovered **clear monthly cycles in daily CGM statistics** – minutes out of range & 95 percentile – in cyclical women

Investigate prevalence and engineer multiscale model glucose features with **large cohort longitudinal CGM data** from Dexcom

Objectives & Methods

Objective: Identify and analyze features displaying monthly cyclicity in women agnostically determined to be undergoing regular menstruation (women in 30s).



Dexcom sensors measure glucose every 5 minutes from diabetes patients for 1 year. Features are acquired from the resulting data and Lomb-Scargle is run on the data to model the cyclical patterns and strength.

Subproject 1 (Clean Data): Remove data abnormalities that could hinder the outcomes of the lomb scargle pipeline.

Subproject 2 (Develop Multiscale Model Features): Test features that describe nuanced characteristics of rhythms: peak parametrization, circadian fluctuations, and within day cyclicity.

Subproject 3 (Feature Clustering and Stability): Characterize distinct groups of populations and identify features shared amongst individuals.

Cohort Table:

Sex	Age	Diabetes Type	Count
Female	30s	T1D	5000
		T2D	5000
		Total	10000
	60s	T1D	1000
		T2D	999
		Total	1999
	Total		11999
Male	30s	T1D	1000
		T2D	1000
		Total	2000
	60s	T1D	1000
		T2D	1000
		Total	2000
	Total		4000
Total			15999

Results

Subproject 1: Remove Data Abnormalities

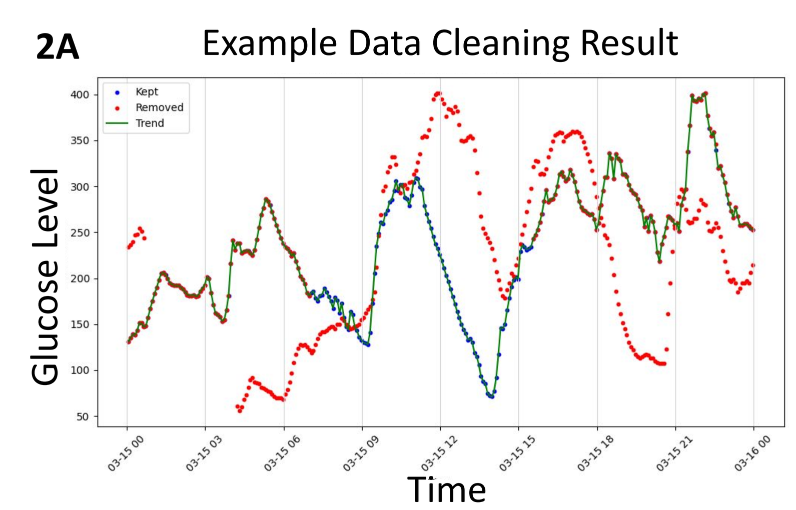


Figure 2A: Sample data cleaning result from one day of one subject.

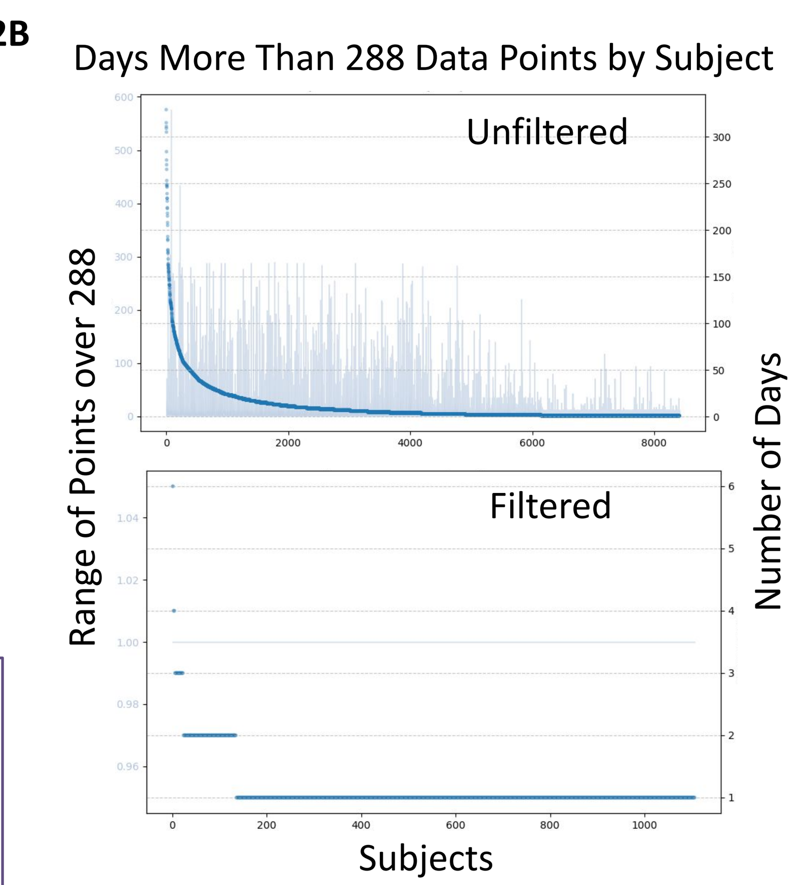


Figure 2B: Number of days and range over 288 data points of all subjects with at least one abnormal day before and after filtering

- 0.14% of data removed
- Visually confirmed correct removal
- After filtering, only 1106 subjects have 1 datapoint over the limit compared to 8415 subjects before (Figure 2B)

Subproject 2: Test Multiscale Model Features

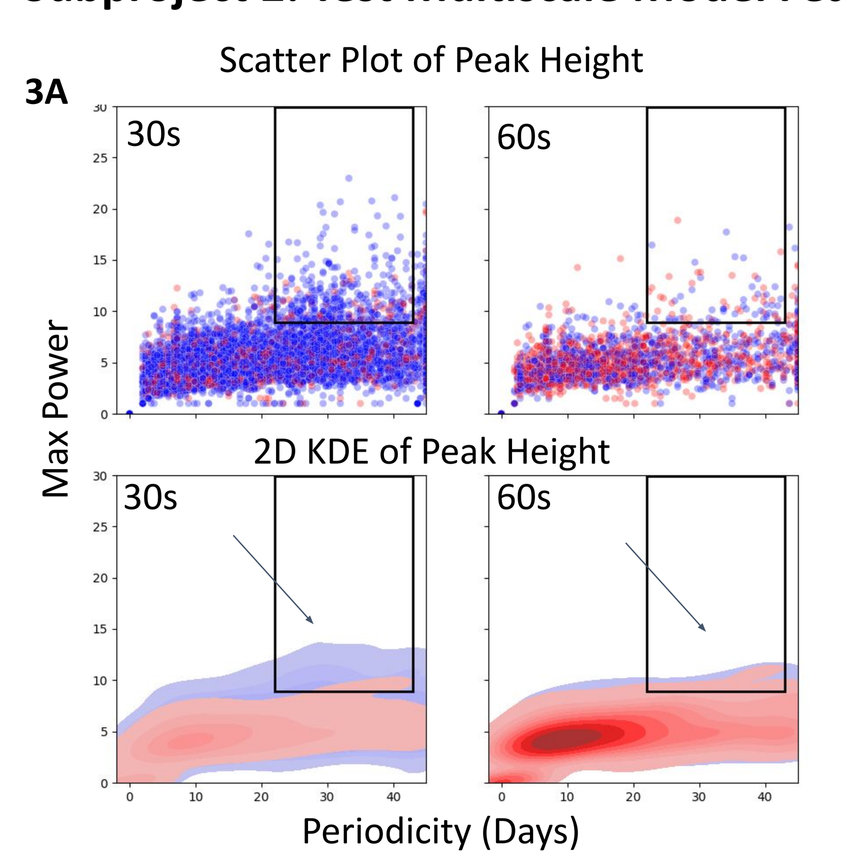


Figure 3A: Example Lomb-Scargle of female vs. male glucose metrics

Table 1: Chi-Square Test P-Values

Comparison	30s Males vs. 30s Females	60s Males vs. 60s Females
Peak Height	1.63E-4	0.70
Peak Mid Height	3.35E-5	0.31
Peak Prominence	0.89	0.53
Range	0.32	0.83
Within Day Cyclicity	1	0.65
Circadian (Hourly Mean)	0.20	0.60
Circadian (Hourly Variance)	0.08	0.35
• Female • Male	Significant for p ≤ 0.05	

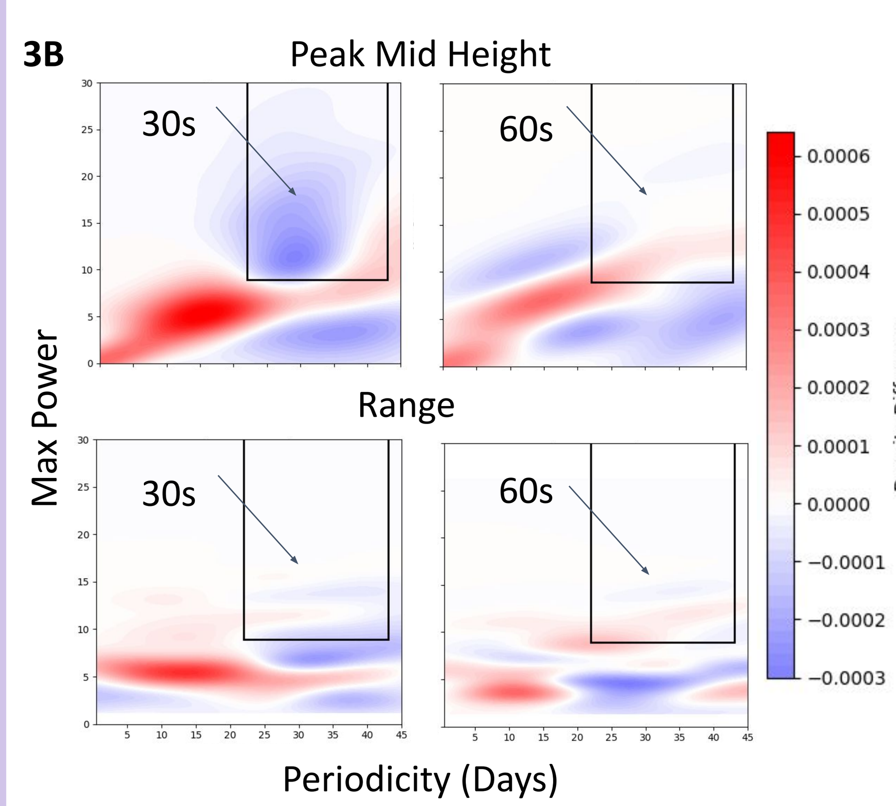


Figure 3B: KDE of density differences (males - females)

- Outlined box = 22-43 day peaks above a 8.94 power threshold (cyclic range)
- Figure 3A shows a **high amount of blue** population within the box; depicting **increased cyclicity for females**.
- The surge of **blue seen in the top left** graph in figure 3B, and **lack thereof in the top right** graph, strengthens the claim that **cyclicity is specific to women in 30s**.

Results Cont.

Subproject 3: Feature Stability and Clustering

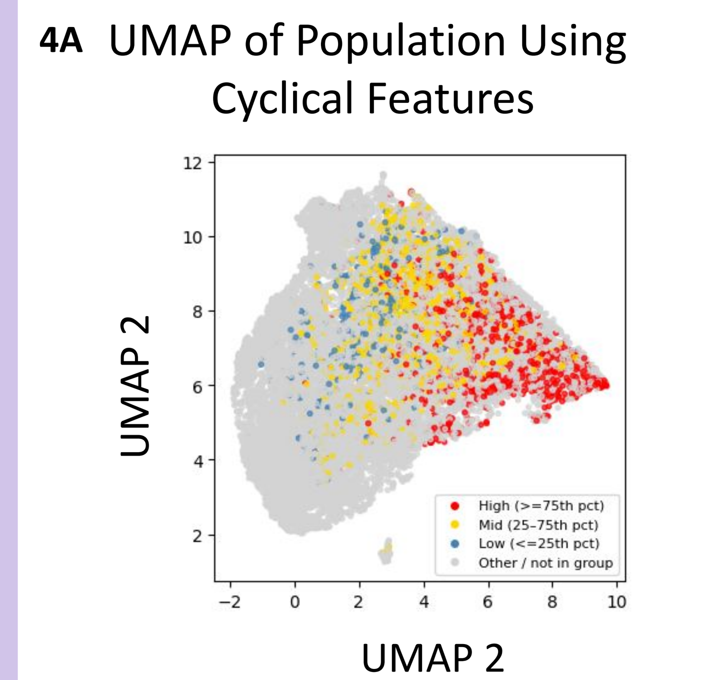


Figure 4B: Dimensional Reduction using UMAP from max power and period of the eight features.

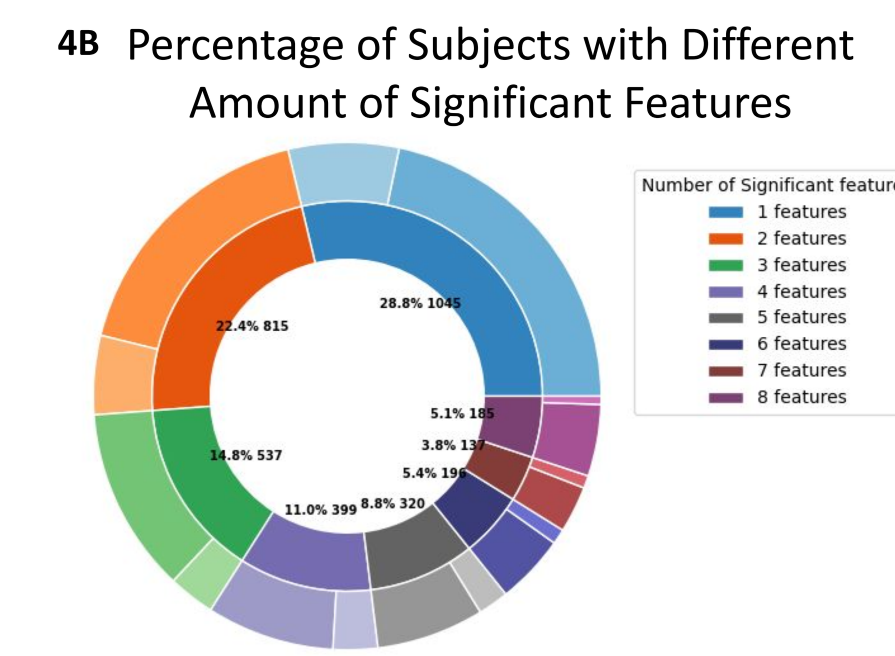


Figure 4A: Distribution of individuals with significance in X amount of features. Lighter color = males, darker color = females

- The UMAP reveals a continuous gradient, with samples arranged according to their level of cyclicity (Fig. 4A).
- The results suggest that there are independent groups of people who show cyclicity in some features but not in others (Fig. 4b).

Conclusion & Future Directions

- Certain features pertaining to **peaks,** which could characterize post meal time spikes, **had the most significance** in cyclicity. Relative height, which captures the height of the peak from the trough to crest, is more subject to change due to external factors. However, absolute height**,** a more simplistic measure, showed significance.
- **Simpler features capture cyclicity best**. It seems that complex features are heterogenous and are more so affected by external factors.
- Without **labels of those external factors**, it is difficult to extract true signal from complex features and validate our findings, which is a **major limitation** of the study.

Future directions:

- Refining dimensionality reduction by filtering for cyclical range to better visualize groups with shared cyclic features.
- Identifying cyclical glucose patterns aligned with menstrual phases using the mcPHASES dataset.
- Long term goal - contribute to **personalized DM treatment plans** based on an individual's cyclicity status.

Acknowledgements & References

We would like to thank our mentors Dr. Smarr, Saara Kriplani and former Smarr Lab member Dr. Jamie Burks. We would also like to thank Dr. Taylor and the BENG 187 instructional team. Finally, thank you to Dexcom for providing the datasets for our project.

References: