TOWARDS THE DEVELOPMENT OF SUBJECT-INDEPENDENT INVERSE METABOLIC MODELS

Seyedhooman Sajjadi, Anurag Das, Ricardo Gutierrez-Osuna, Theodora Chaspari, Projna Paromita, Laura E. Ruebush, Nicolaas E. Deutz, Bobak J. Mortazavi

COMPUTER SCIENCE & ENGINEERING
TEXAS A&M UNIVERSITY

PATHS-UP
Precise Advanced Technologies and Health Systems for Underserved Populations

Introduction

Motivation

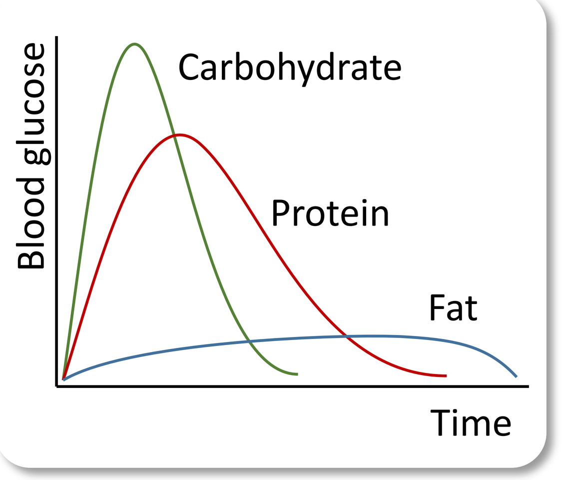
- 30M Americans suffer from diabetes and 84M are pre-diabetic
- In order to prevent pre-diabetics from developing type 2 diabetes, it is important to minimize excess glucose levels

Significance

- Maintaining proper glucose levels requires proper management of diet and exercise
- While exercise tracking exists, current diet monitoring solutions remain impractical or creates user burden, often with abundant manual logging necessary

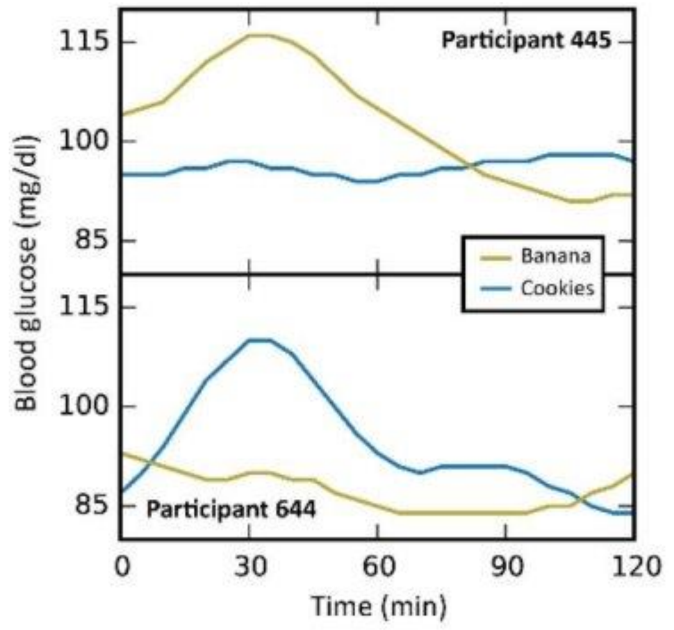
Continuous Glucose Monitoring: An Opportunity

- Continuous glucose monitors (CGM) can measure the post prandial glucose response (PPGR) to any food eaten
- PPGR is known to be impacted by the macronutrient composition of meals (carbohydrates, proteins, and fats)
- **This suggests the shape of the PPGR can be used to estimate macronutrients in a meal.**
- **We call these Inverse Metabolic Models (IMMs)**



Key Challenge

- A landmark study from Zeevi et. al (Cell, 2015) tracked glucose response of meals in 800 participants and identified significant differences in responses to the same meals



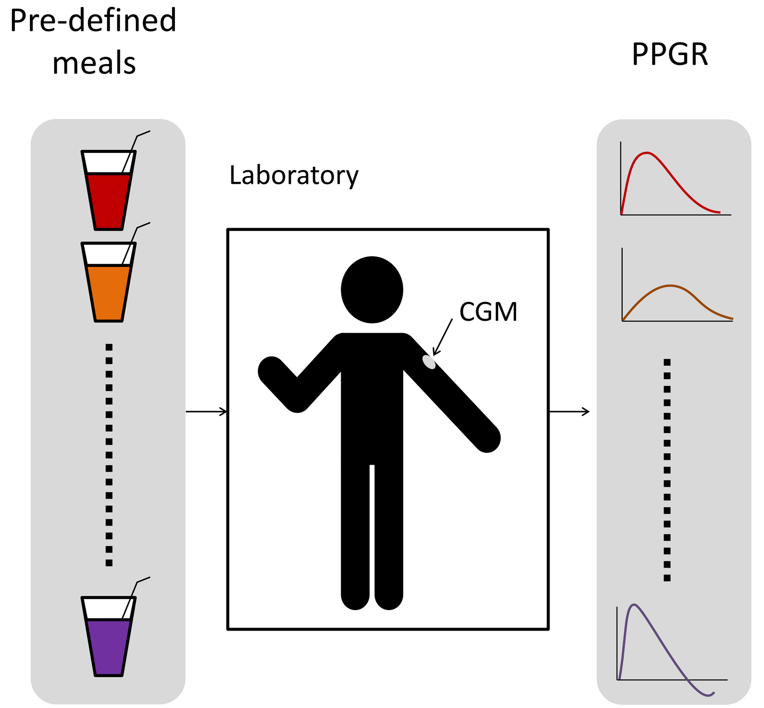
(Zeevi et al, Cell 2015)

Method

Analytic Design

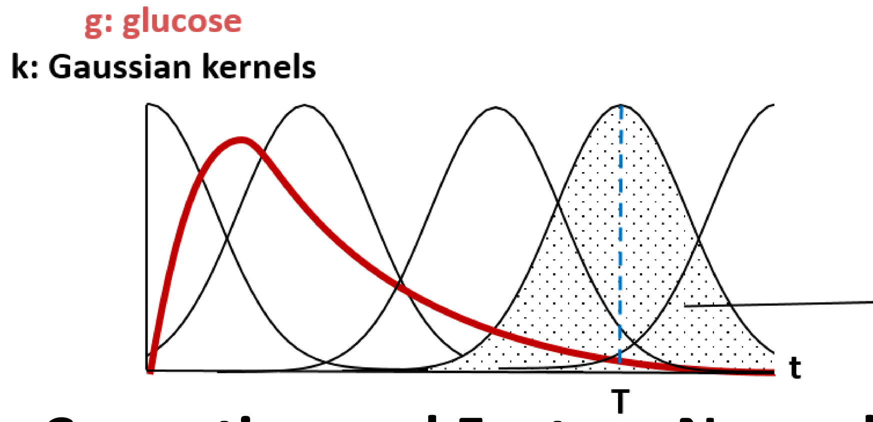
- We provided participants with nine meals of known macronutrient composition and were asked to wear a CGM to capture data
- 15 healthy older adults (60-85 years), BMI of 25-35 (IRB #2017-0886)

Meal	CHO (g)	Protein (g)	Fat (g)	Total E (kcal)
C1P1F1	52.25	15	13	386
C2P2F2	94.75	30	26	733
C3P3F3	179.75	60	52	1427
C1P2F2	52.25	30	26	563
C3P2F2	179.75	30	26	1073
C2P3F2	94.75	60	26	853
C2P1F2	94.75	15	26	673
C2P2F3	94.75	30	52	967
C2P2F1	94.75	30	13	616



Feature Extraction

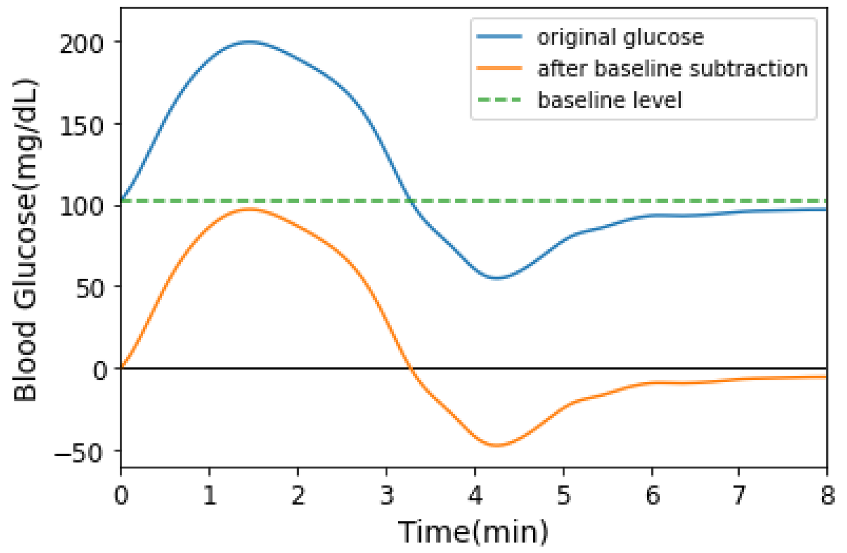
- Gaussian kernels extract 8 features representing area under the PPGR curve at various points in time

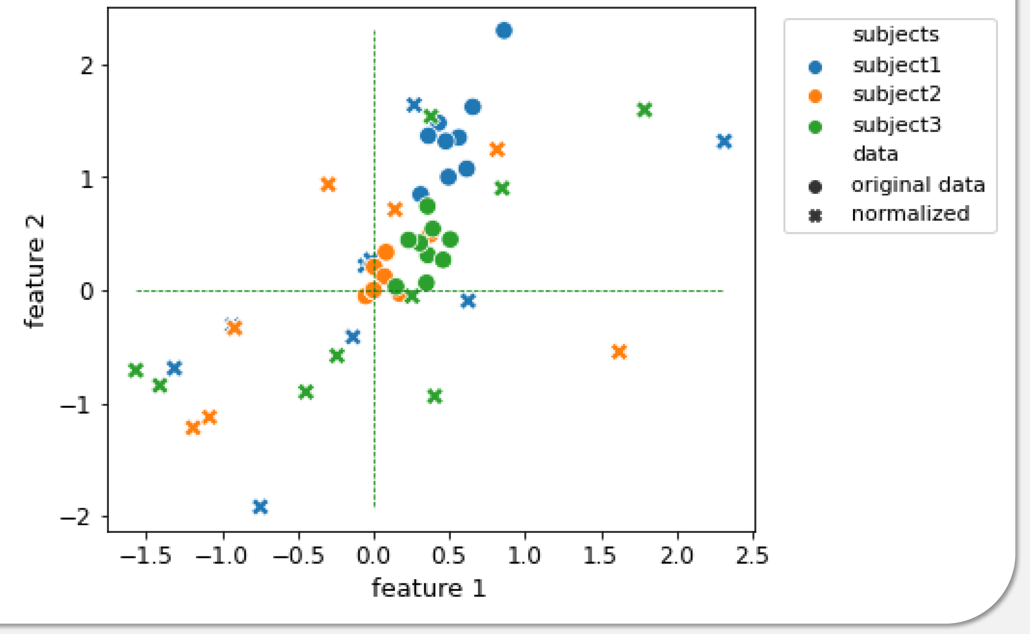


$$f = \int_0^T g(t) e^{\frac{1}{2\sigma_k^2}(t-T_k)} dt$$

Baseline Correction and Feature Normalization

- Center all PPGRs around their initial fasting level
- Normalize features subject-wise to reduce heterogeneity: z-score and min-max normalization





Experiments and Results

- Using a leave-one-subject-out cross-validation, three XGBoost decision tree regression models were trained to estimate quantity of Carbohydrates, Proteins, and Fats.
- We calculated correlation and root mean square relative error (to be able to compare quantity errors with different composition concentrations)
- First, we evaluate the impact of baseline correction:

Results of baseline correction using XGBoost

Model	Correlation C	Correlation P	Correlation F	Mean RMSRE (std) C	Mean RMSRE (std) P	Mean RMSRE (std) F
None	0.55	0.42	0.39	0.41(0.17)	0.51(0.15)	0.51(0.16)
Subtraction	**0.61**	**0.48**	**0.48**	**0.35(0.20)**	**0.50(0.13)**	**0.49(0.15)**
Division	0.59	0.49	0.47	0.34(0.21)	0.49(0.12)	0.51(0.13)

Correlation significance: $p < 0.001$

- This improved correlations to statistical significance, but errors remain large
- Then we implement normalization:

Model	Correlation C	Correlation P	Correlation F	Mean RMSRE (std) C	Mean RMSRE (std) P	Mean RMSRE (std) F
None	0.61	0.48	0.48	0.35(0.20)	0.50(0.13)	0.49(0.15)
min-max	0.77	**0.48**	0.64	0.28(0.16)	**0.47(0.17)**	0.41(0.14)
z-score	**0.83**	0.43	**0.65**	**0.22(0.10)**	0.50(0.12)	**0.40(0.14)**

- Using normalization techniques, we find accurate subject-independent IMMs for C and F. Improvement of P requires additional biomarker data not currently captured by CGMs.

Conclusion

- We evaluated the impact of pre-processing techniques to account for subject-to-subject variability
- We improved accuracy of IMMs
- These results remain stable with the use of only two meals for normalization

Acknowledgements:
This work was supported, in part, by funding from the National Science Foundation (Award #2014475) and Engineering Research Center for Precise Advanced Technologies and Health Systems for Underserved Populations (PATHS-UP) (Award #1648451).