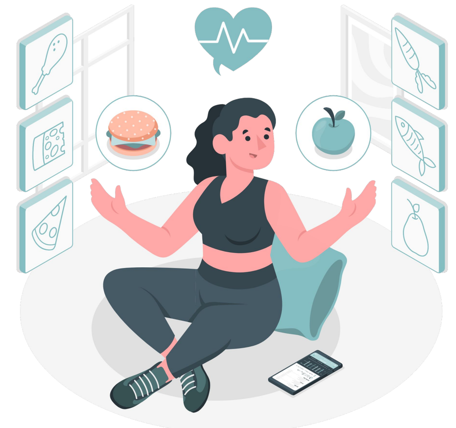
Interpretable Mechanistic Representations for Meal-level Glycemic Control in the Wild



Ke Alexander Wang

Emily B. Fox

	Expert CGM features (TIR, MAGE, ...)	Mechanistic ODE features	Pure neural network features	Hybrid VAE features (ours)
Noninvasive, inexpensive	Yes	No	Yes	Yes
Adaptable to new data	No	No	Yes	Yes
Physiologically Interpretable	No	Yes	No	Yes
Robust to in-the-wild data	Yes	No	Yes	Yes

Our health application

Type-2 diabetes disease subtyping

ML problem

How do we learn a physiologically-grounded embedding space to describe glycemic control?

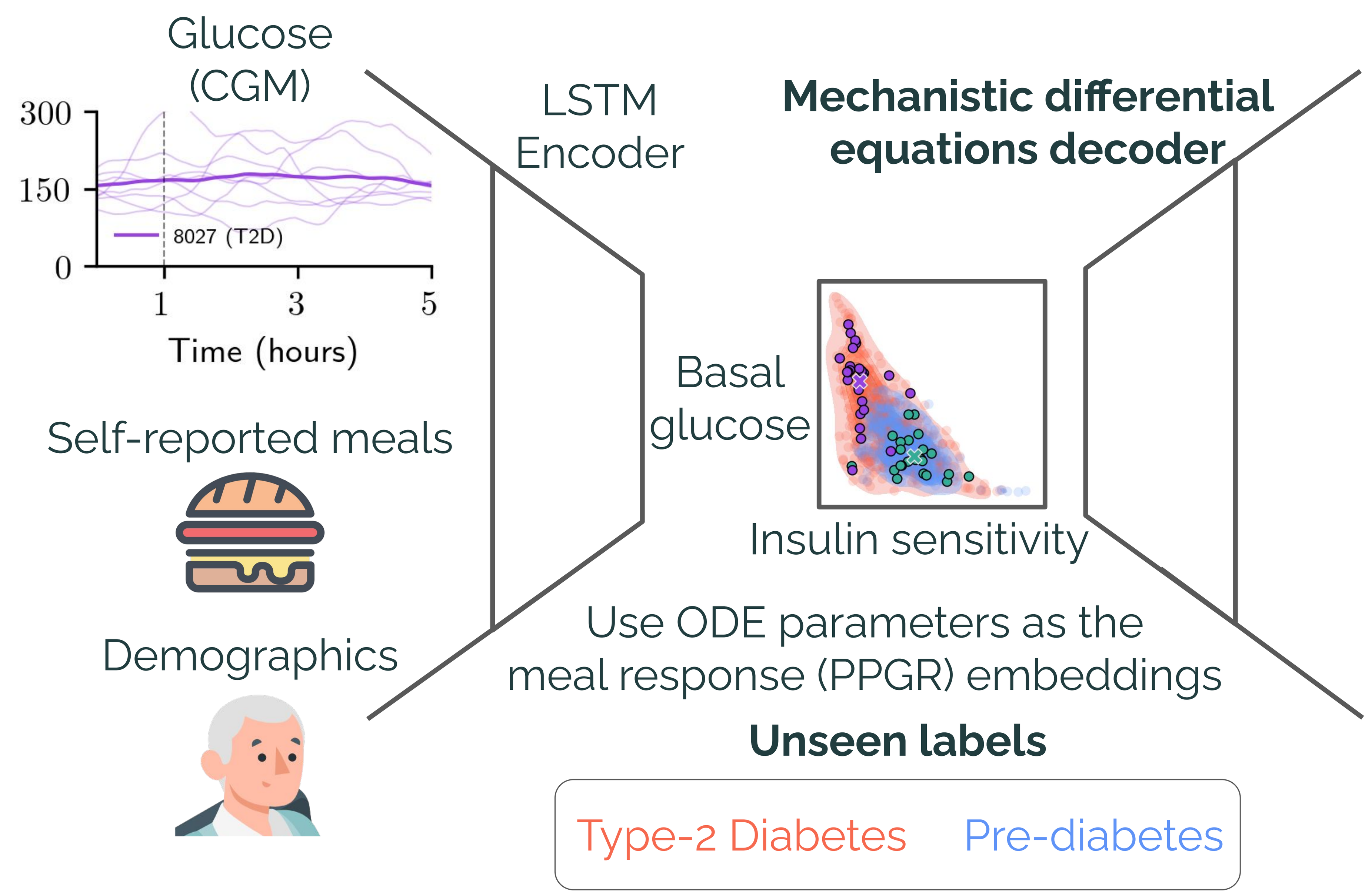
Our constraints (common in healthcare)

- Unlabeled data
- “Small data” regime
- Real-world self-reported data with missingness
- Want to incorporate prior expert knowledge

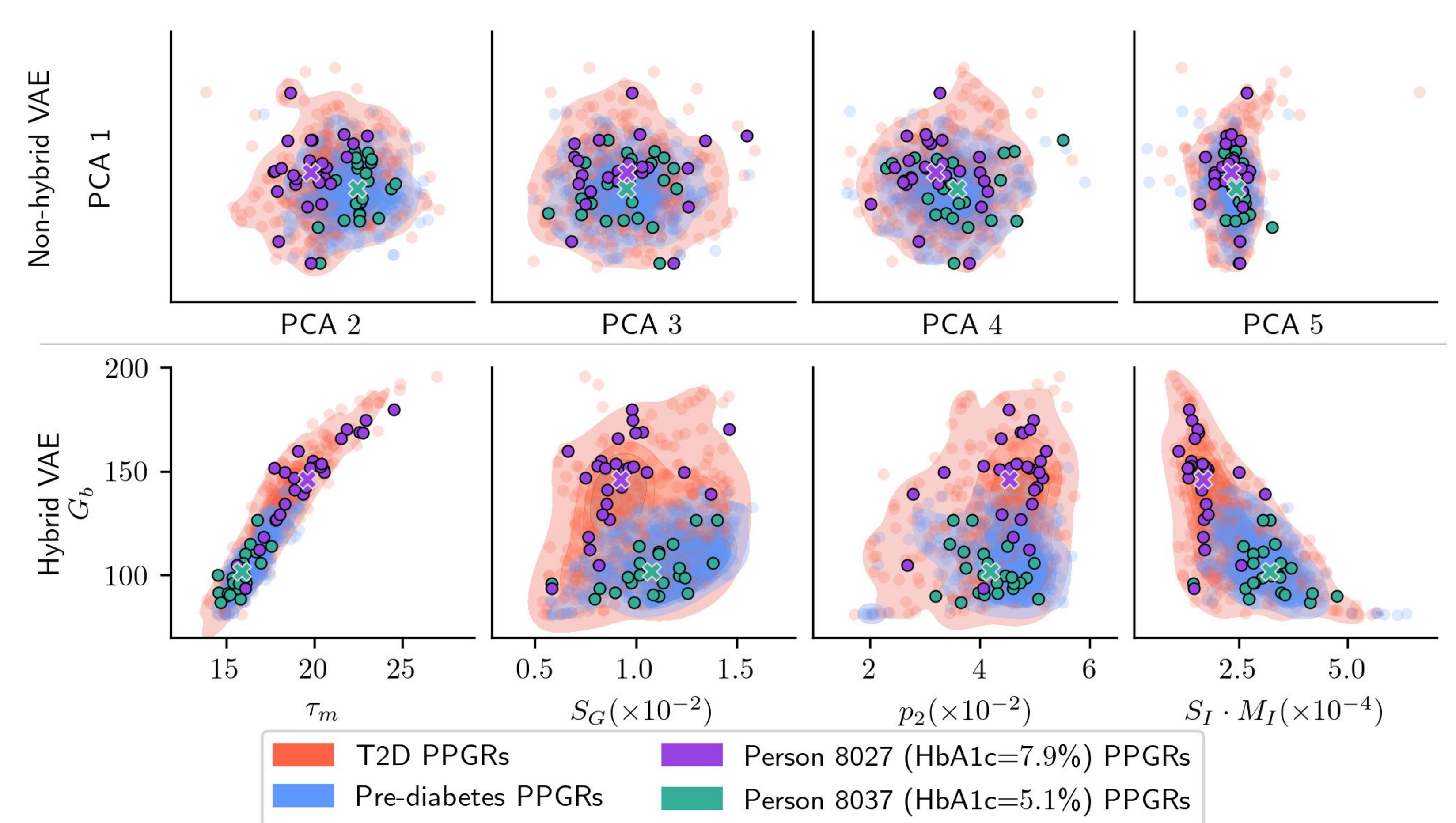
Code on GitHub

Follow me @therealalexwang

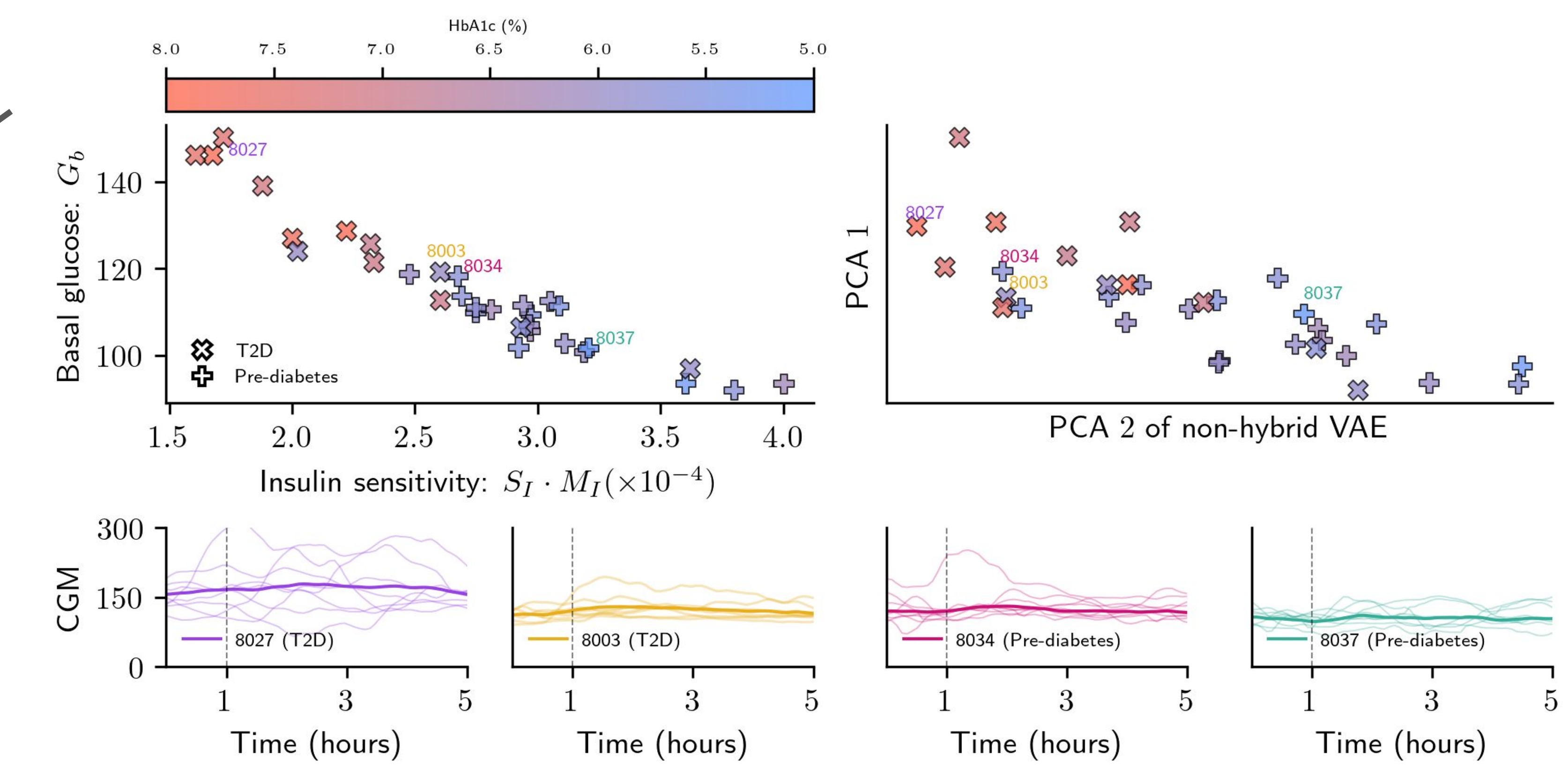
Our model



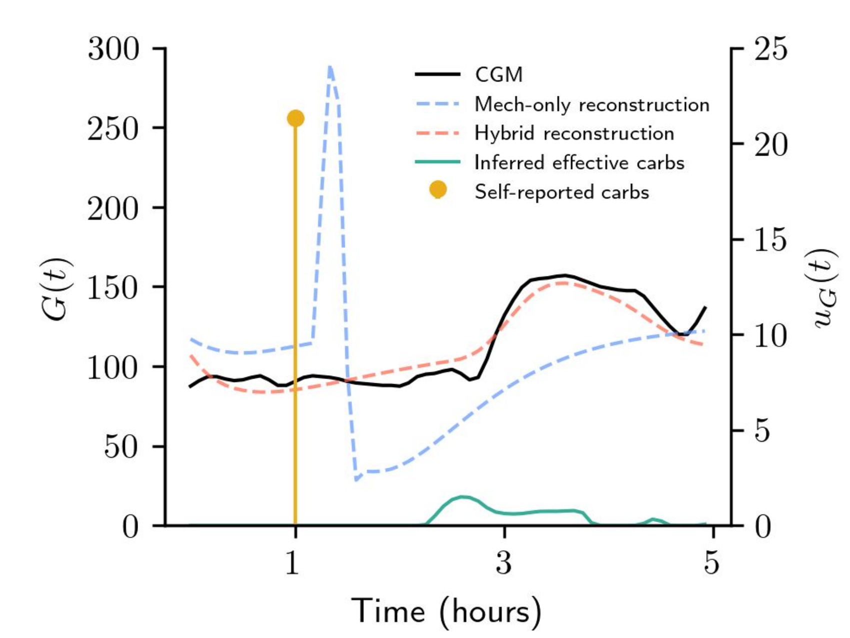
Result: A physiologically interpretable embedding space



Result: Embedding space agrees with clinical standard HbA1C (unseen)



Result: Inferring effective carbs makes model robust to errors in self-reported meal data



Can’t blindly trust self-reported meals!

Trick: infer effective carbs w/ end-to-end training

Result: Quantitatively better clustering than other methods

	NMI	AMI	Hom.	Comp.
Raw CGM ($d = 60$)	0.41	0.39	0.36	0.47
Raw CGM + DTW ($d = 60$)	0.54	0.53	**0.62**	0.51
Expert features ($d = 10$)	0.25	0.25	0.61	0.21
TCL + Average ($d = 32$)	0.24	0.21	0.19	0.33
TCL + Concat ($d = 3 \times 32$)	0.29	0.27	0.24	0.38
Black-box VAE ($d = 32$)	0.21	0.19	0.22	0.21
Mechanistic ODE ($d = 7$)	0.14	0.12	0.14	0.14
Hybrid VAE ($d = 7$)	**0.54**	**0.53**	0.51	**0.58**

More accurate and more interpretable