Hybrid^2 Neural ODE Causal Modeling

and an Application to Glycemic Response


Bob Junyi Zou

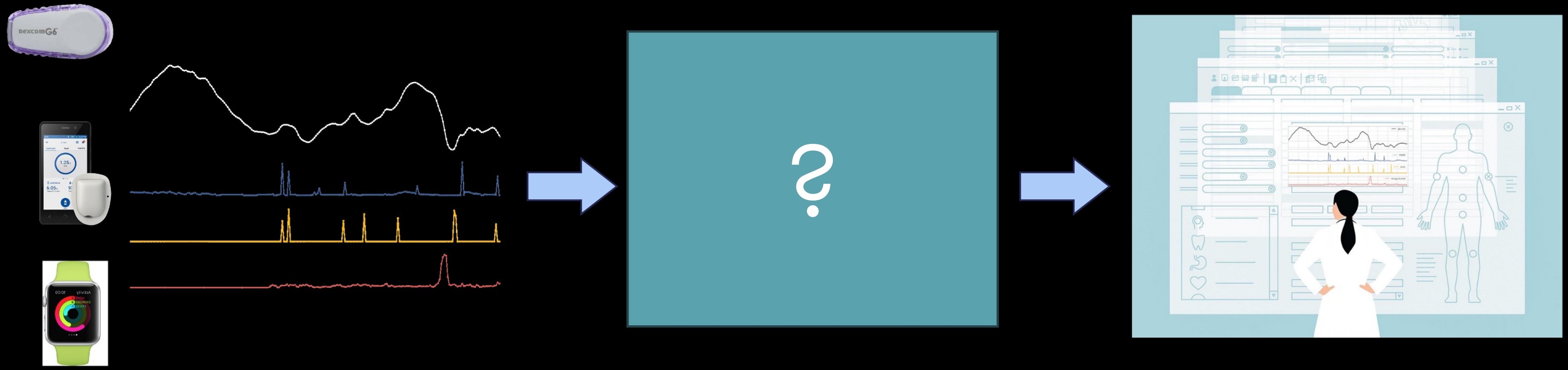

Matthew Levine


Dessi Zaharieva


Ramesh Johari

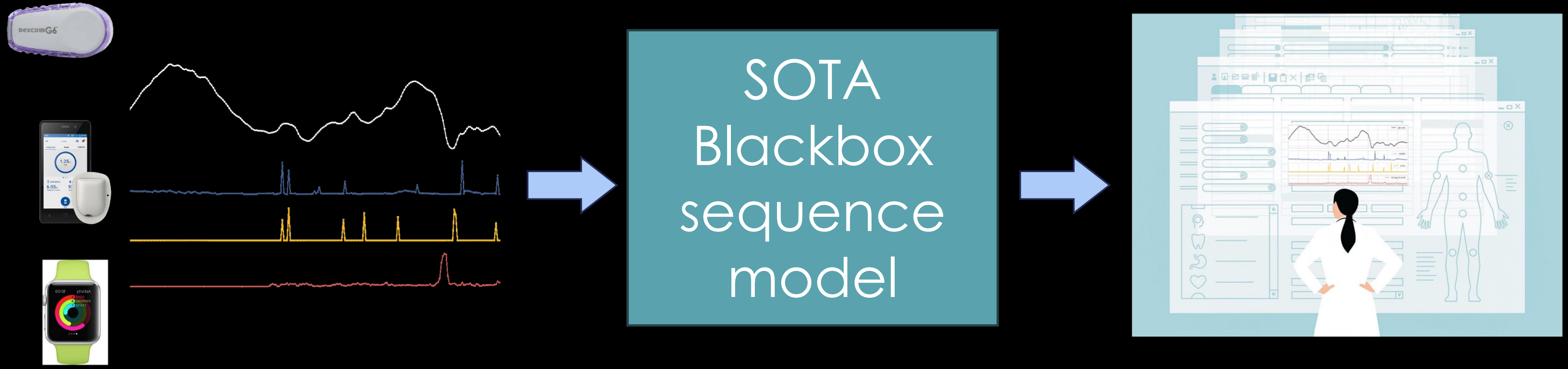

Emily Fox

Health data streams are growing in availability and fidelity…



How can we extract useful knowledge?

Health data streams are growing in availability and fidelity…



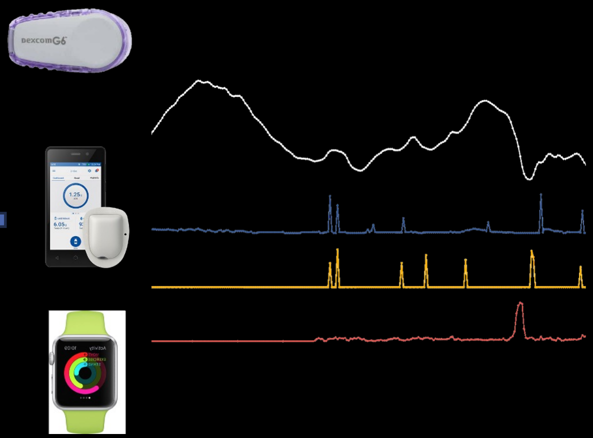
Blackbox models make reasonable predictions on average, but …

Good blackbox prediction

does NOT guarantee

good counterfactual inference

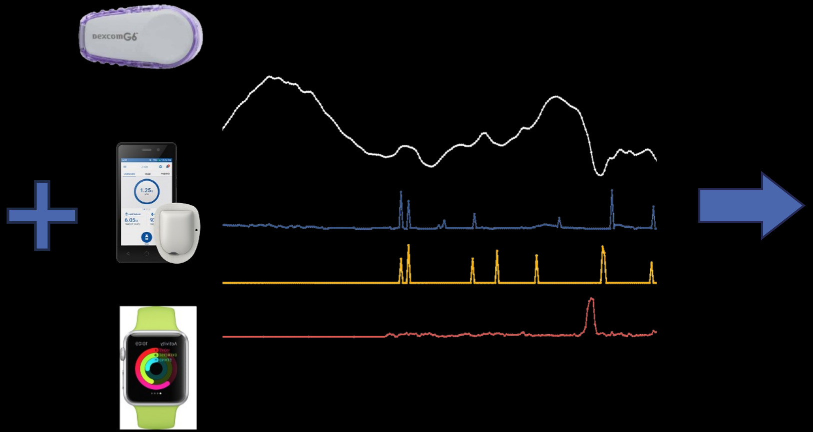
Clinical application:
Post-exercise glycemic response in Type 1 Diabetes

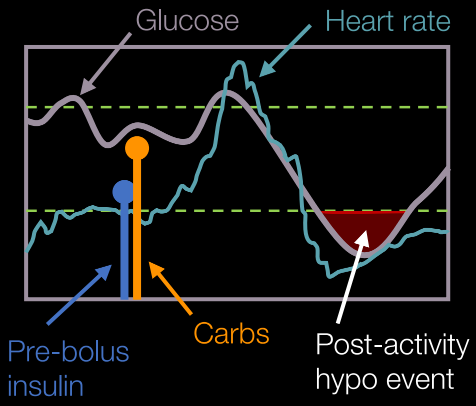


Type 1 Diabetes Exercise Initiative (T1DEXI), *https://doi.org/10.25934/PR00008428*

Clinical application:

Post-exercise glycemic response in Type 1 Diabetes



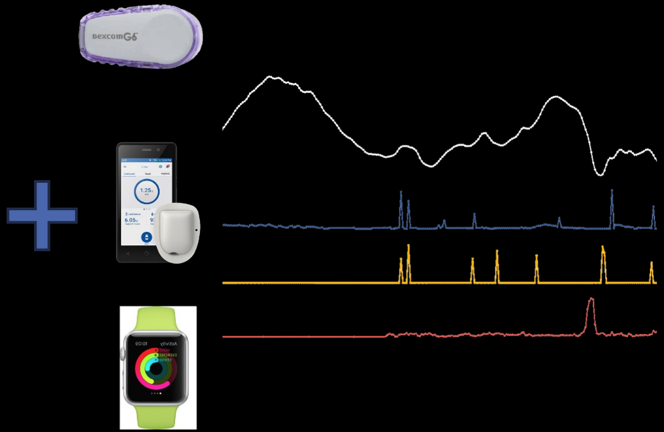


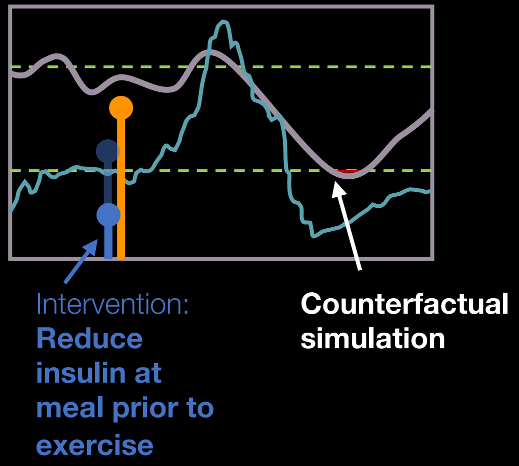
Can we learn dynamic models that:

- Make accurate forecasts?

Type 1 Diabetes Exercise Initiative (T1DEXI), *https://doi.org/10.25934/PR00008428*

Clinical application:
Post-exercise glycemic response in Type 1 Diabetes

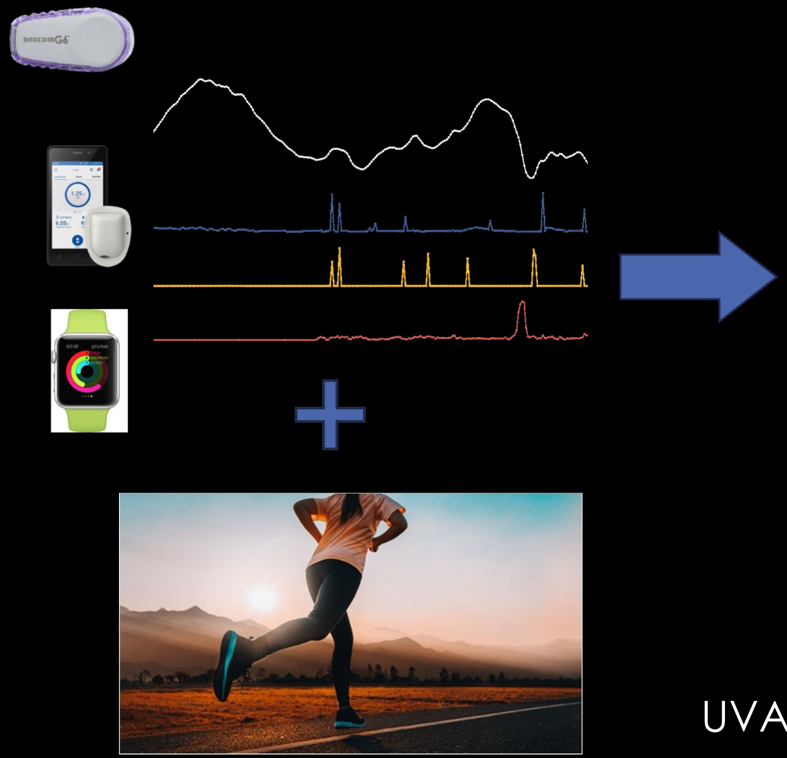




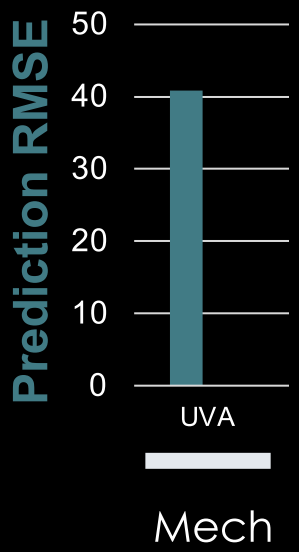
Can we learn dynamic models that:
- Make accurate forecasts?
- Can be used for counterfactual simulations?

Type 1 Diabetes Exercise Initiative (T1DEXI), *https://doi.org/10.25934/PR00008428*

Prediction accuracy: Blackbox models outperform mechanistic model



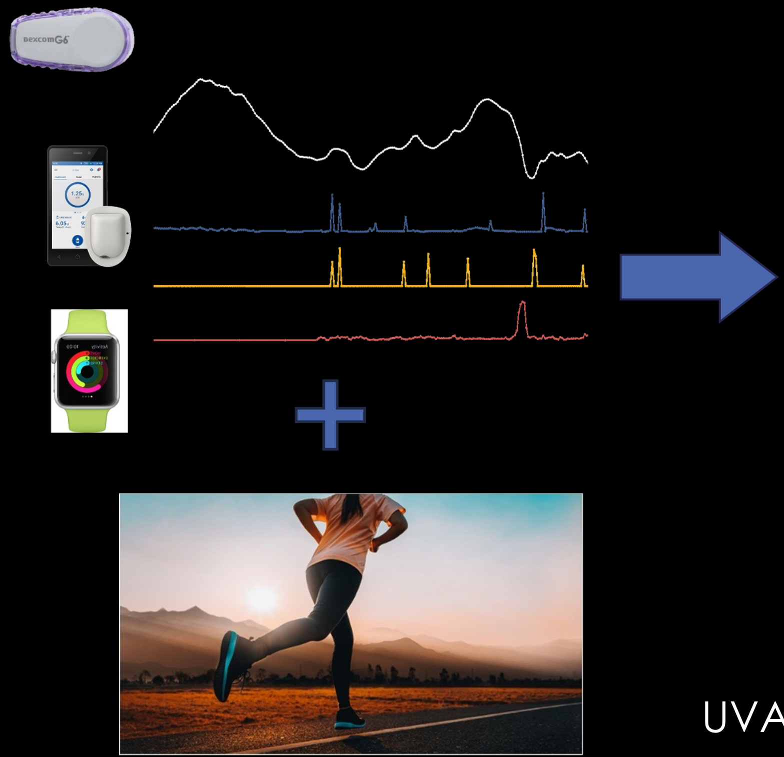
30-min post-exercise forecast RMSE



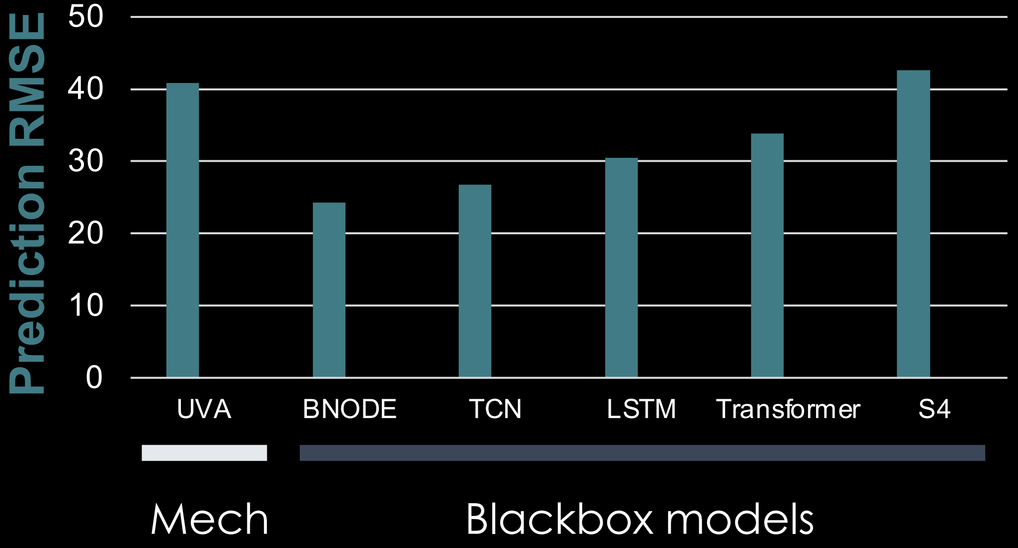
UVA/Padova model (Dalla Man et al. 2014)

Type 1 Diabetes Exercise Initiative (T1DEXI), *https://doi.org/10.25934/PR00008428*

Prediction accuracy:
Blackbox models outperform mechanistic model



30-min post-exercise forecast RMSE

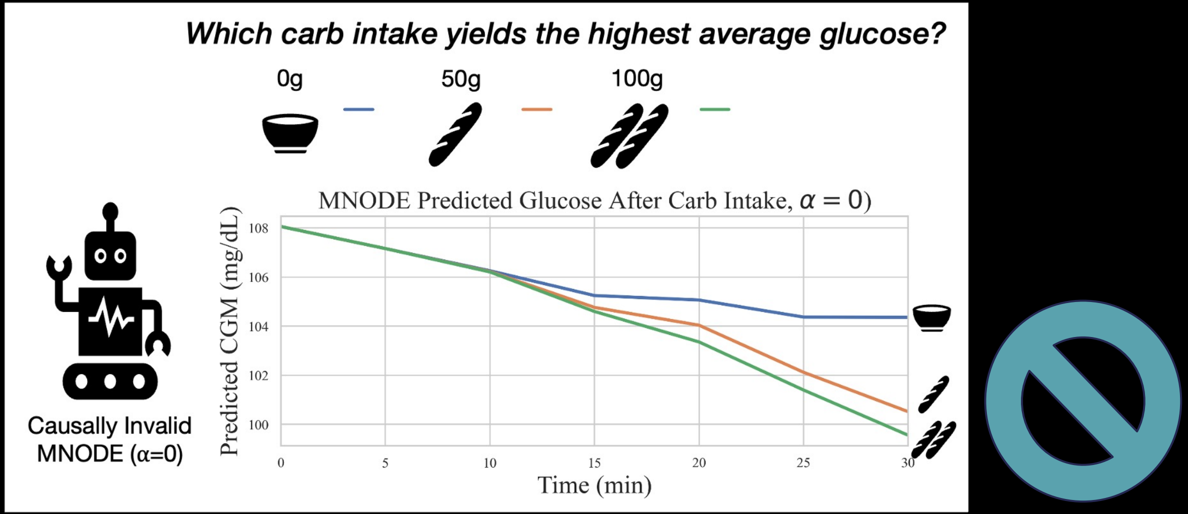


UVA/Padova model (Dalla Man et al. 2014)

Type 1 Diabetes Exercise Initiative (T1DEXI), *https://doi.org/10.25934/PR00008428*

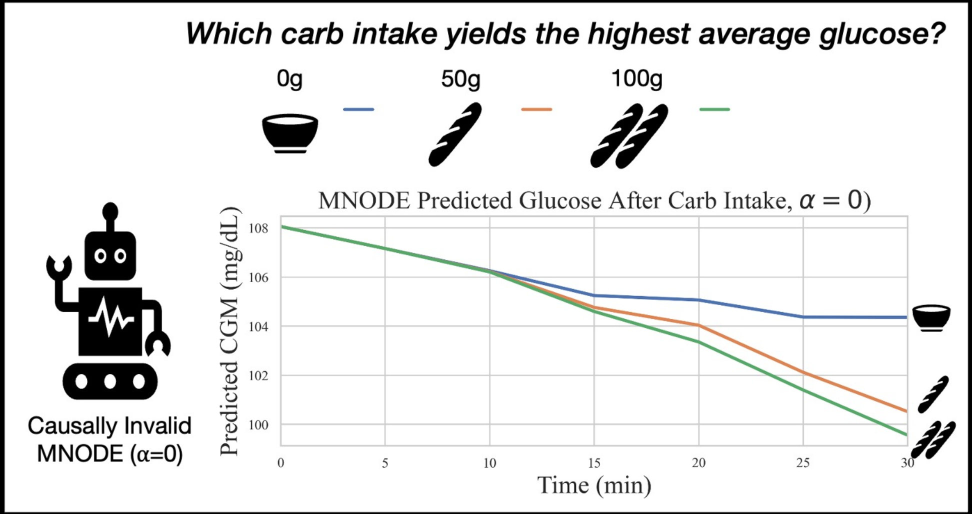
Can we use the trained blackbox models for counterfactual simulations?

Can we use the trained blackbox models for counterfactual simulations? NO.



Often we know things like "*[more] carbohydrates raise blood glucose [more]*"**...and yet predict the opposite**

Can we use the trained blackbox models for counterfactual simulations? NO.



Often we know things like *"[more] carbohydrates raise blood glucose [more]"*...**and yet predict the opposite**

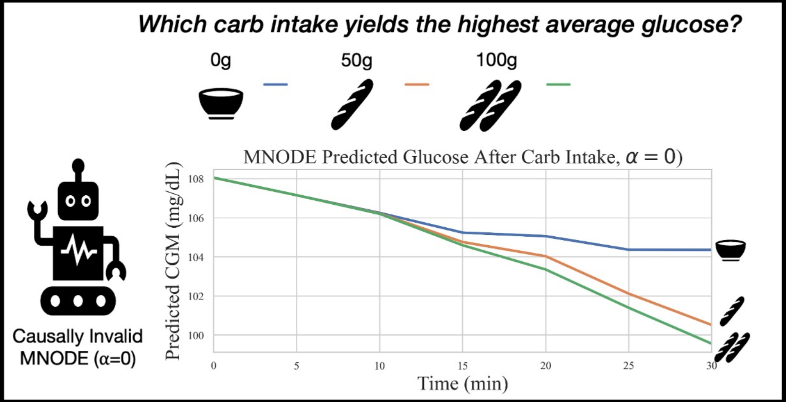
How often do we get the ranking wrong?

Causal error rate

Consider **intervention set**:

$$X^{(i)},\ i = 1, \ldots, K$$

Causal error rate



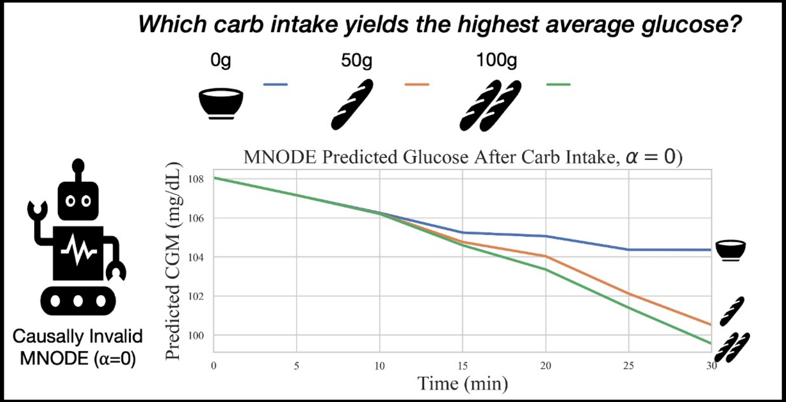
Consider **intervention set**:

$$X^{(i)},\ i = 1, \ldots, K$$

Under fitted model, each yields predictions

$$\hat{Y}^{(i)},\ i = 1, \ldots, K$$

Causal error rate



Consider **intervention set**:

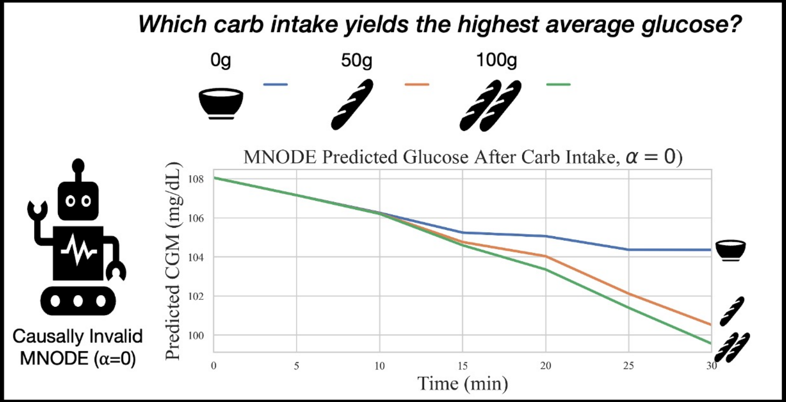
$$X^{(i)}, \ i = 1, \ldots, K$$

Under fitted model, each yields predictions

$$\hat{Y}^{(i)}, \ i = 1, \ldots, K$$

Intervention with **largest estimated effect**: $\hat{\mathcal{I}} = \arg\max_i \ \text{score}\big(\hat{Y}^{(i)}\big)$

Causal error rate



Consider **intervention set**:

$$X^{(i)},\ i = 1, \ldots, K$$

Under fitted model, each yields predictions

$$\hat{Y}^{(i)},\ i = 1, \ldots, K$$

Intervention with **largest estimated effect**: $\hat{\mathcal{I}} = \arg\max_i\ \mathrm{score}(\hat{Y}^{(i)})$

Define error rate:

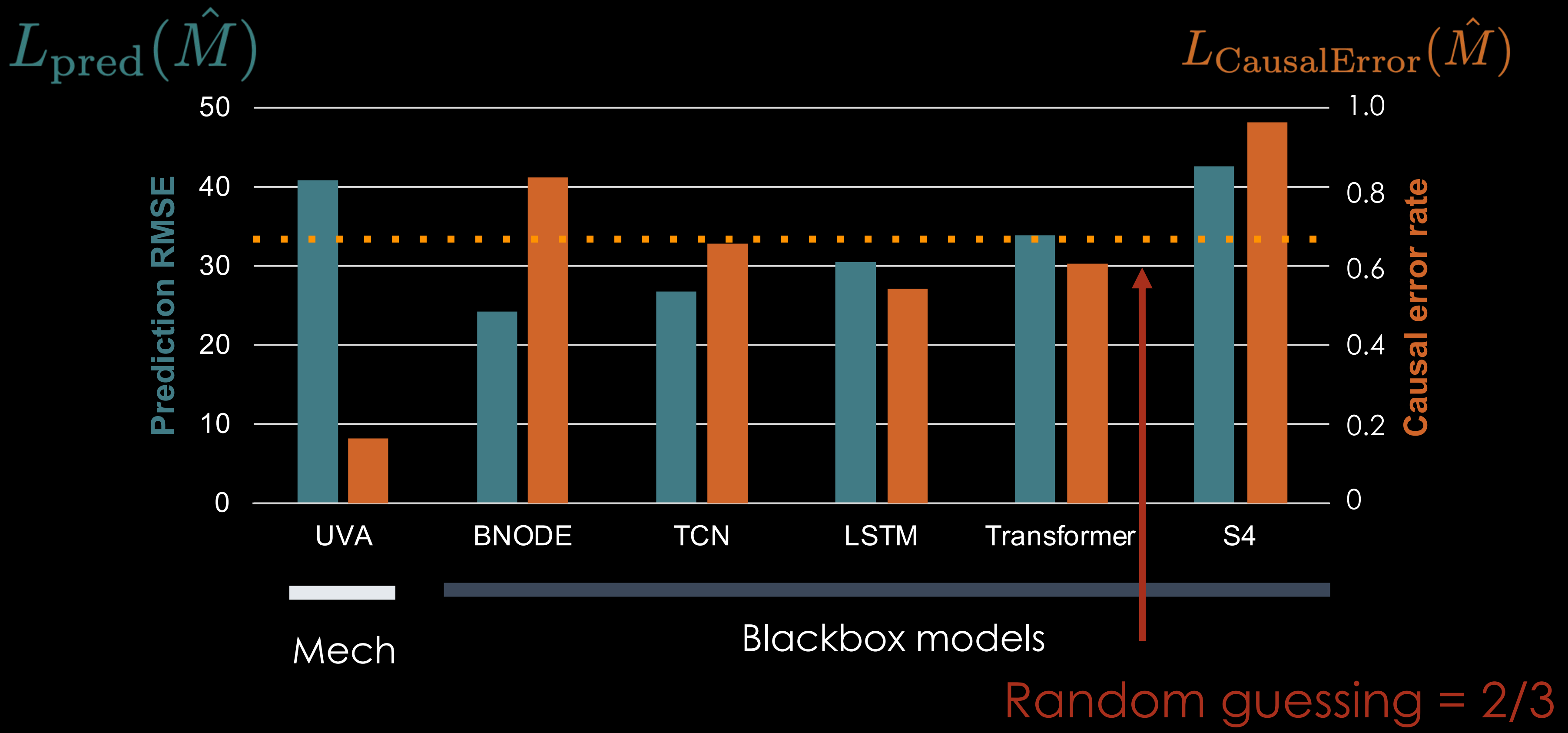
Fitted model

$$L_{\mathrm{CausalError}}(\hat{M}) = \frac{1}{N}\sum_{n=1}^{N} L_{0/1}(\hat{\mathcal{I}}_n, \mathcal{I}_n^*)$$

One-hot encoded domain knowledge of **true max score**

Intervention set index

Blackbox models:
better prediction accuracy + **worse** causal performance

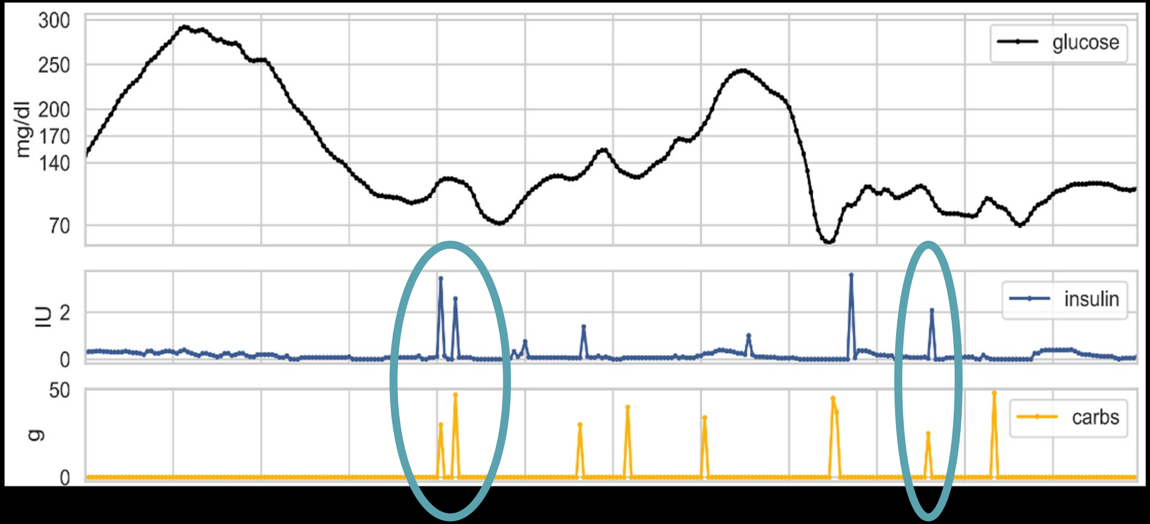


What is going wrong with blackbox models???

Training on real-world observational data:

- Partial observations of a complex system

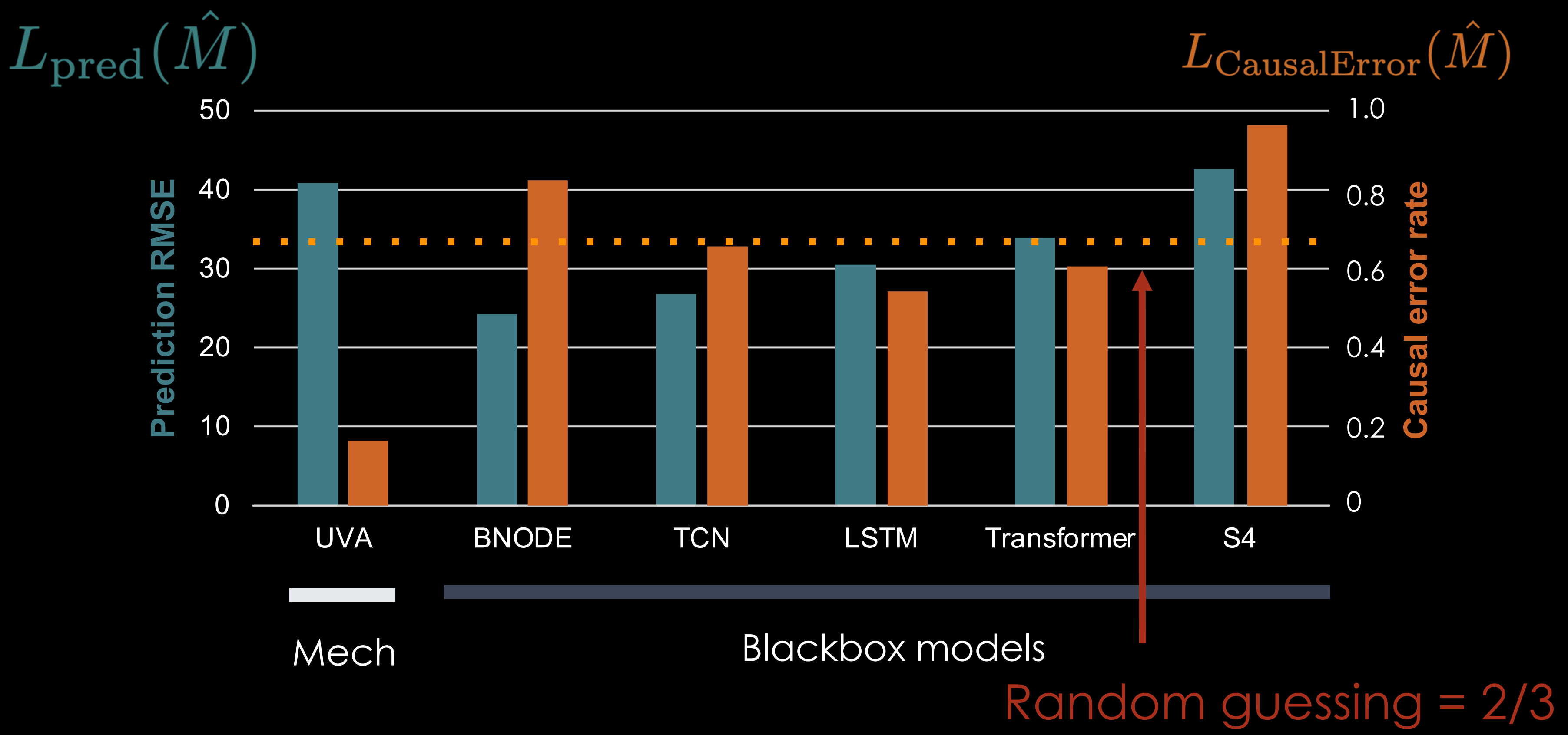
What is going wrong with blackbox models???



Training on real-world observational data:

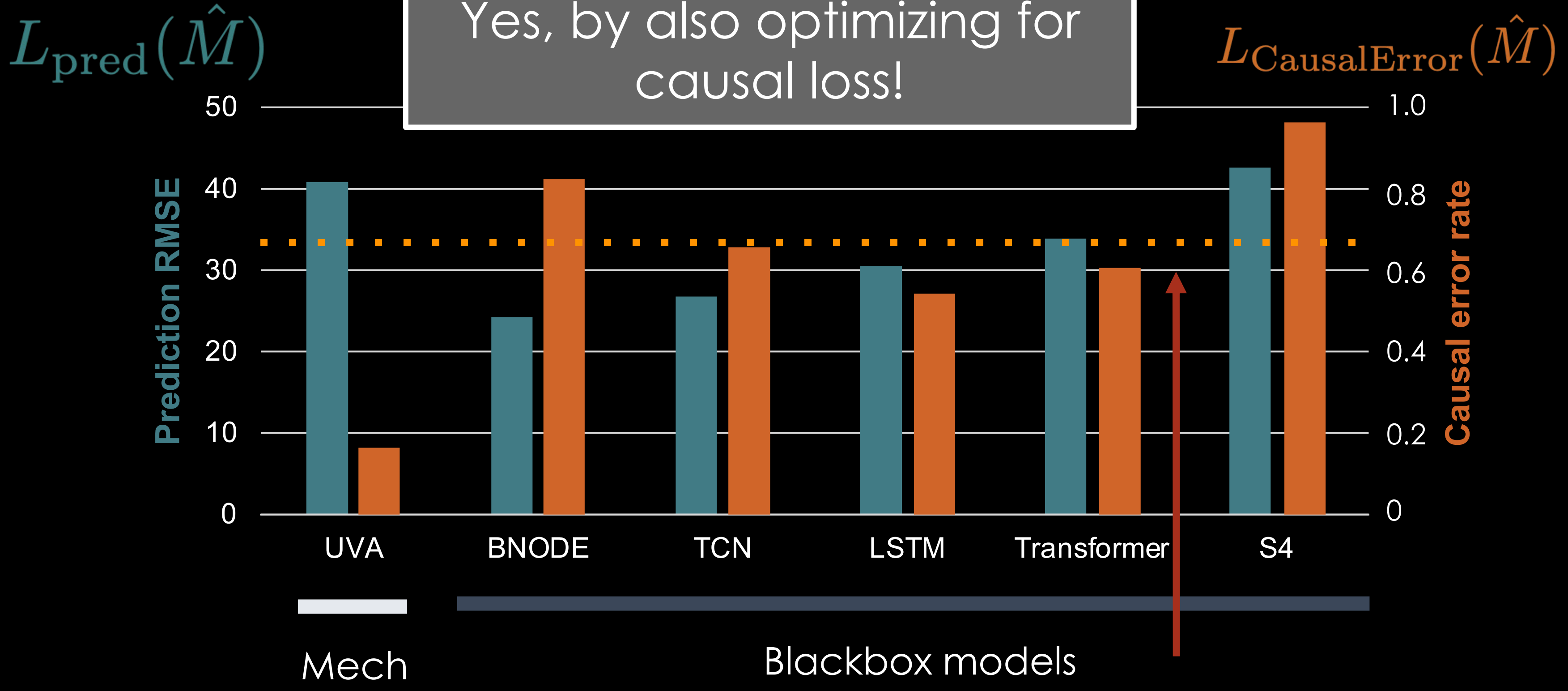
- Partial observations of a complex system
- Biased data collection

Can we rescue **causally invalid** blackbox models?



Can we rescue **causally invalid** blackbox models?

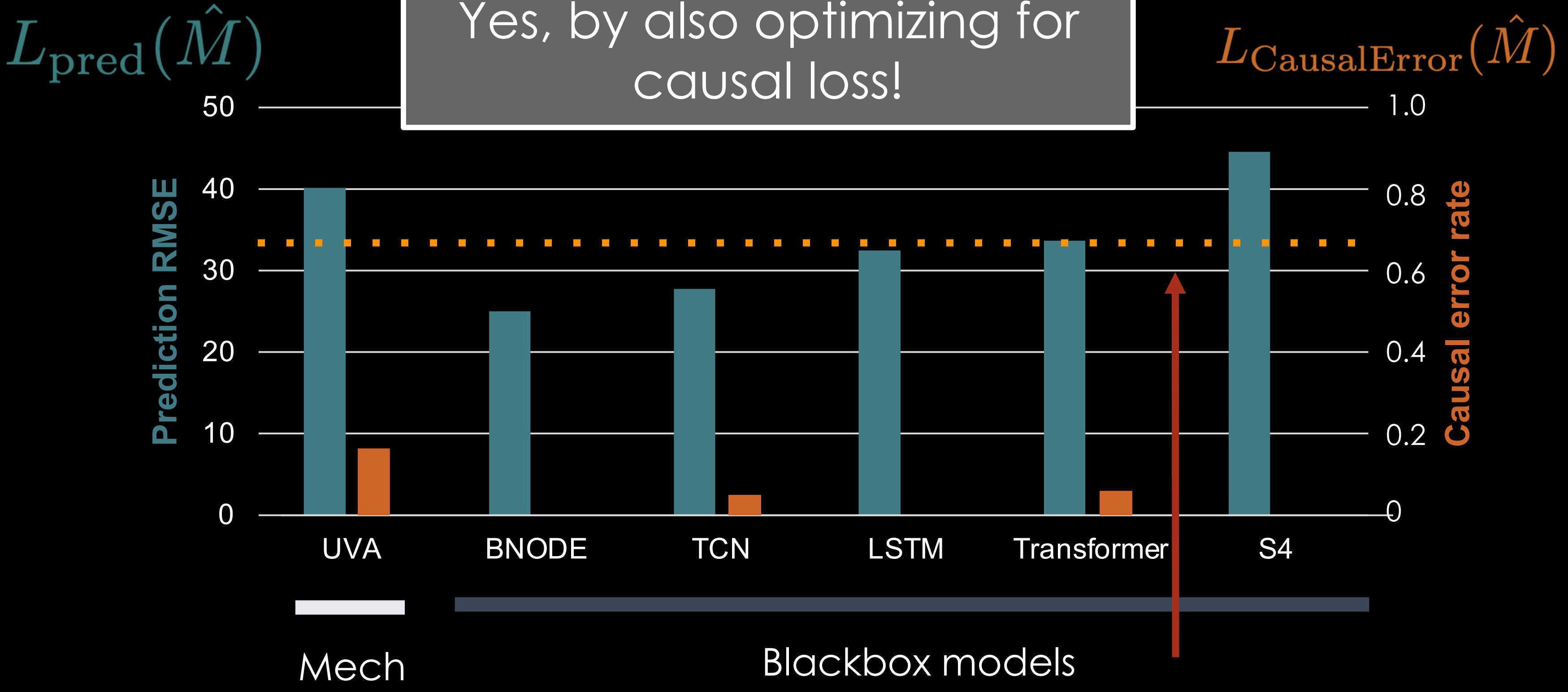
Yes, by also optimizing for causal loss!



Random guessing = 2/3

Can we rescue **causally invalid** blackbox models?

Yes, by also optimizing for causal loss!



Random guessing = 2/3

Hybrid loss

$$L_{\text{hybrid}}(\hat{M}) = (1 - \alpha)L_{\text{pred}}(\hat{M}) + \alpha L_{\text{CausalError}}(\hat{M})$$

Tuning
parameter

Hybrid loss

$$\sum_{t \in \mathcal{T}_{\text{pred}}} \left\| y(t) - \hat{y}(t) \right\|_2^2$$

$$L_{\text{hybrid}}(\hat{M}) = (1 - \alpha) L_{\text{pred}}(\hat{M}) + \alpha L_{\text{CausalError}}(\hat{M})$$

Tuning
parameter

Hybrid loss

$$\sum_{t \in \mathcal{T}_{\text{pred}}} \left\| y(t) - \hat{y}(t) \right\|_2^2$$

$$L_{\text{hybrid}}(\hat{M}) = (1 - \alpha) L_{\text{pred}}(\hat{M}) + \alpha L_{\text{CausalError}}(\hat{M})$$

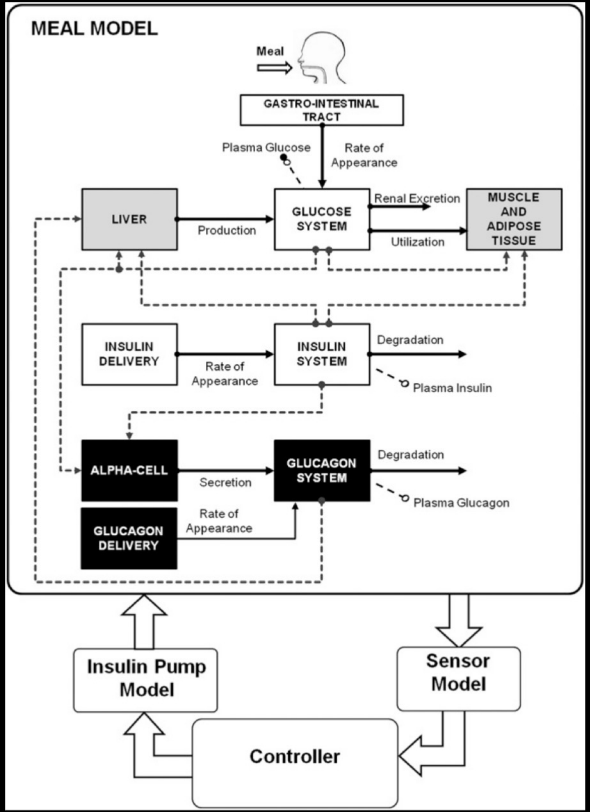
Tuning parameter

$$CrossEntropy(\text{softmax}(\text{scores}), \mathcal{I}^*)$$

Need to pass gradients

Might hybrid architectures that blend **mechanistic** and **data-driven** components provide **adequate causal performance**?

Expert-designed mechanistic models



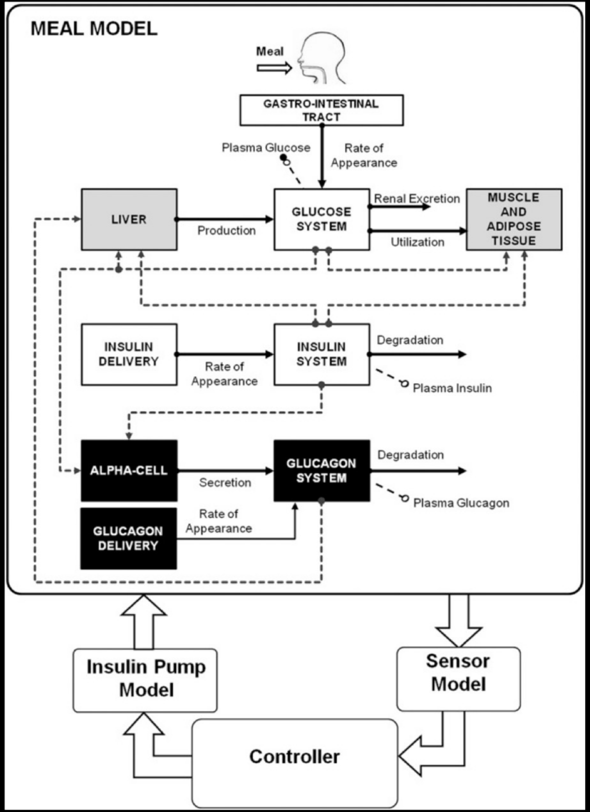
$$\frac{ds}{dt} = f_{\mathrm{mech}}\left(\mathbf{s}, \mathbf{u};\ \theta_{\mathrm{mech}}\right)$$

Physiologic state

External inputs (nutrition, etc.)

UVA model (Dalla Man et al. 2014)

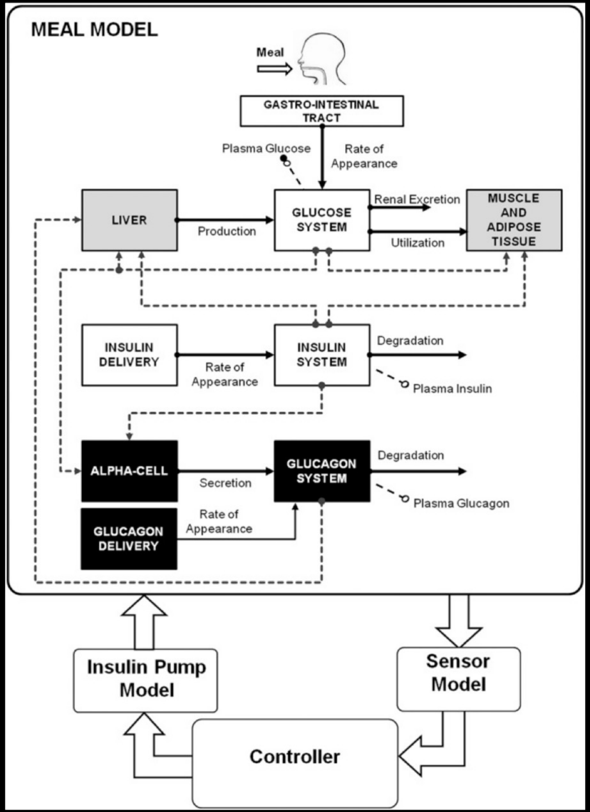
A zoo of hybrid model architectures: **Neural Closure**



$$\frac{d\mathbf{s}}{dt} = f_{\text{mech}}\Big(\mathbf{s};\ \theta_{\text{mech}}\Big)$$

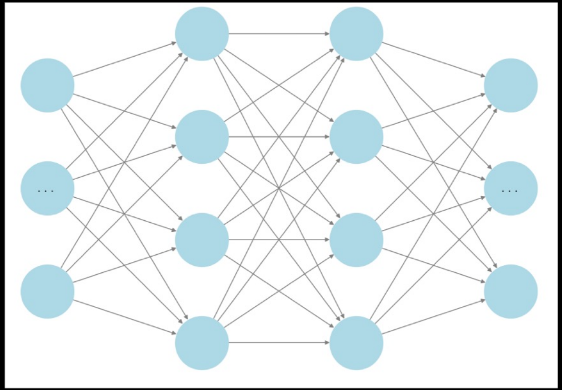
UVA model (Dalla Man et al. 2014)

A zoo of hybrid model architectures: **Neural Closure**



UVA model (Dalla Man et al. 2014)

+

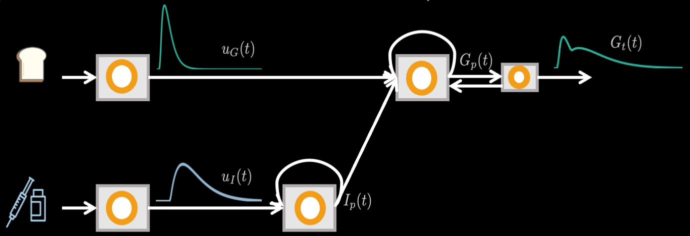


$$\frac{d\mathbf{s}}{dt} = f_{\mathrm{mech}}\Big(\mathbf{s};\ \theta_{\mathrm{mech}}\Big) + \ f_{\mathrm{NN}}\Big(\mathbf{s};\ \theta_{\mathrm{NN}}\Big)$$

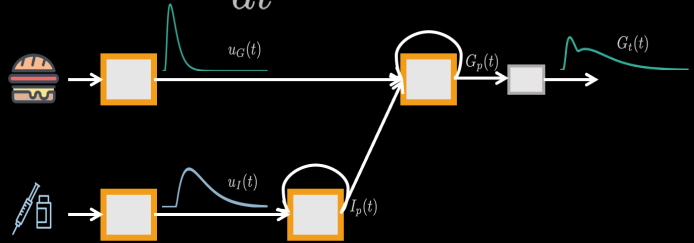
A zoo of hybrid model architectures

$$\frac{ds}{dt} = m(s(t), x(t);\ \beta) + f_\theta(x(t))$$

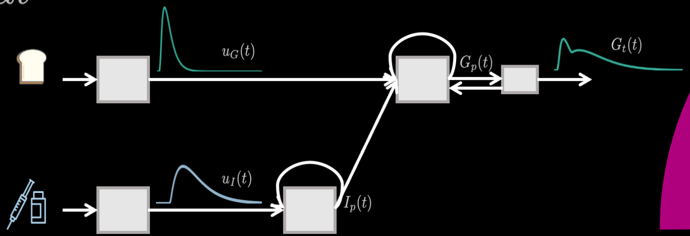
$$\frac{ds}{dt} = m(s(t), x(t);\ \beta,\ f_\theta(x(t)))$$



$$\frac{ds}{dt} = f_\theta(s(t), x(t);\ A_x, A_s)$$



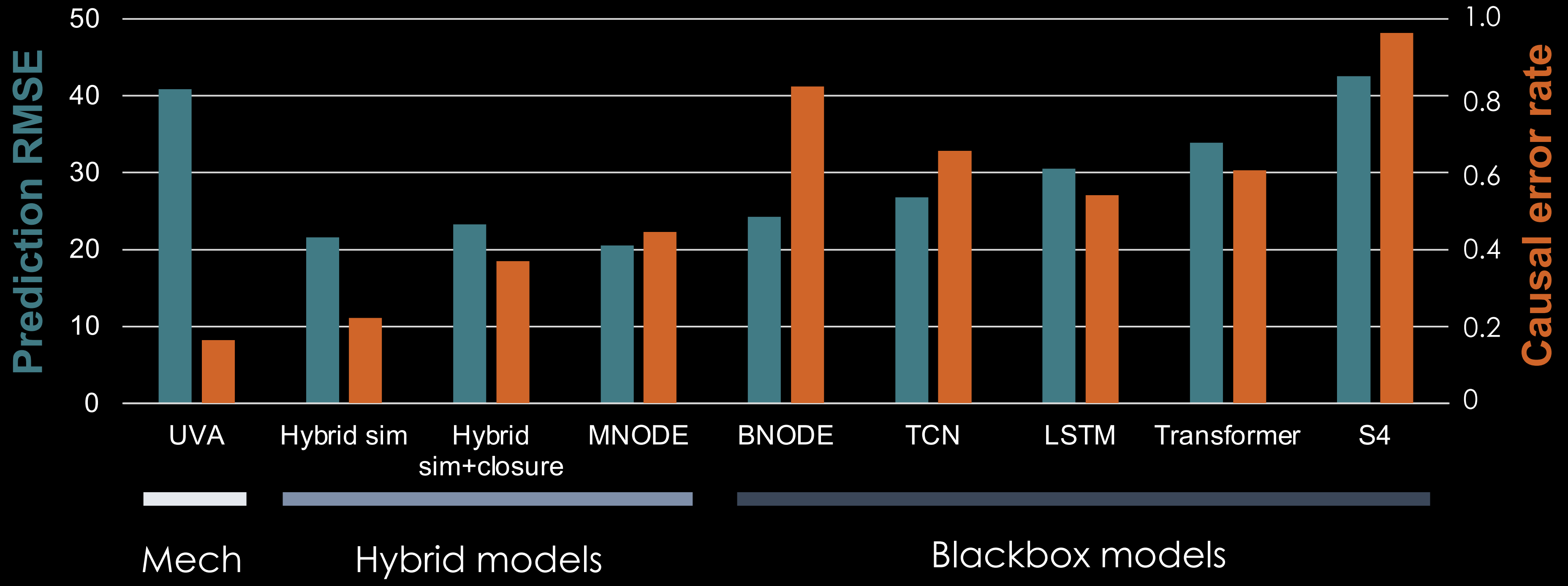
$$\frac{ds}{dt} = m(s(t), x(t);\ \beta)$$



$$\frac{ds}{dt} = f_\theta(s(t), x(t))$$

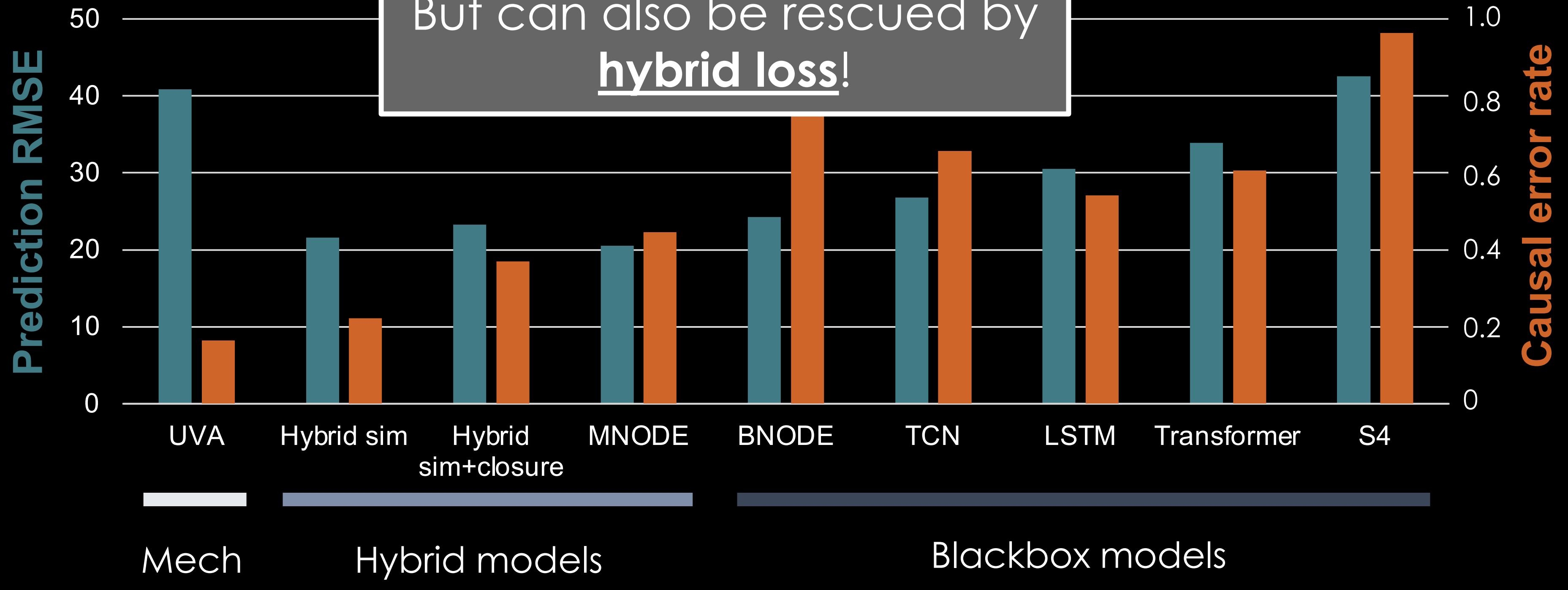
Expert mechanistic model **Black-box model**

Hybrid models can **also** lose causal validity



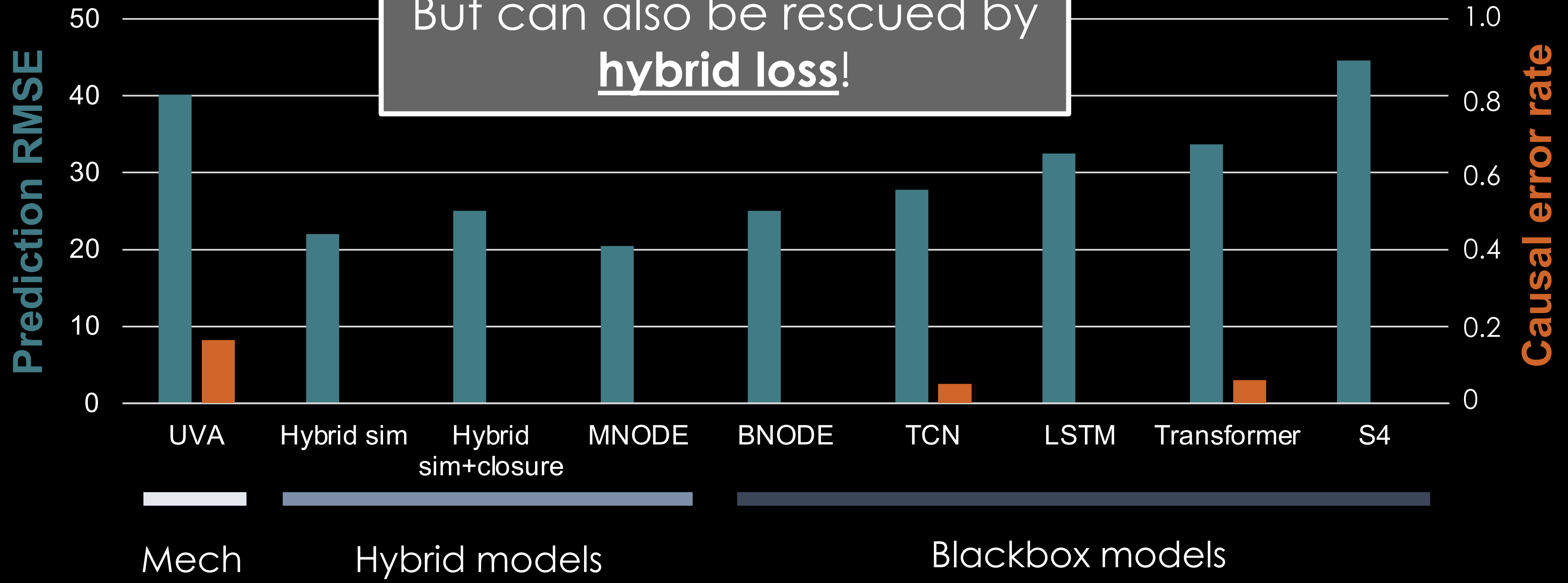
Hybrid models can **also** lose causal validity

But can also be rescued by **hybrid loss**!



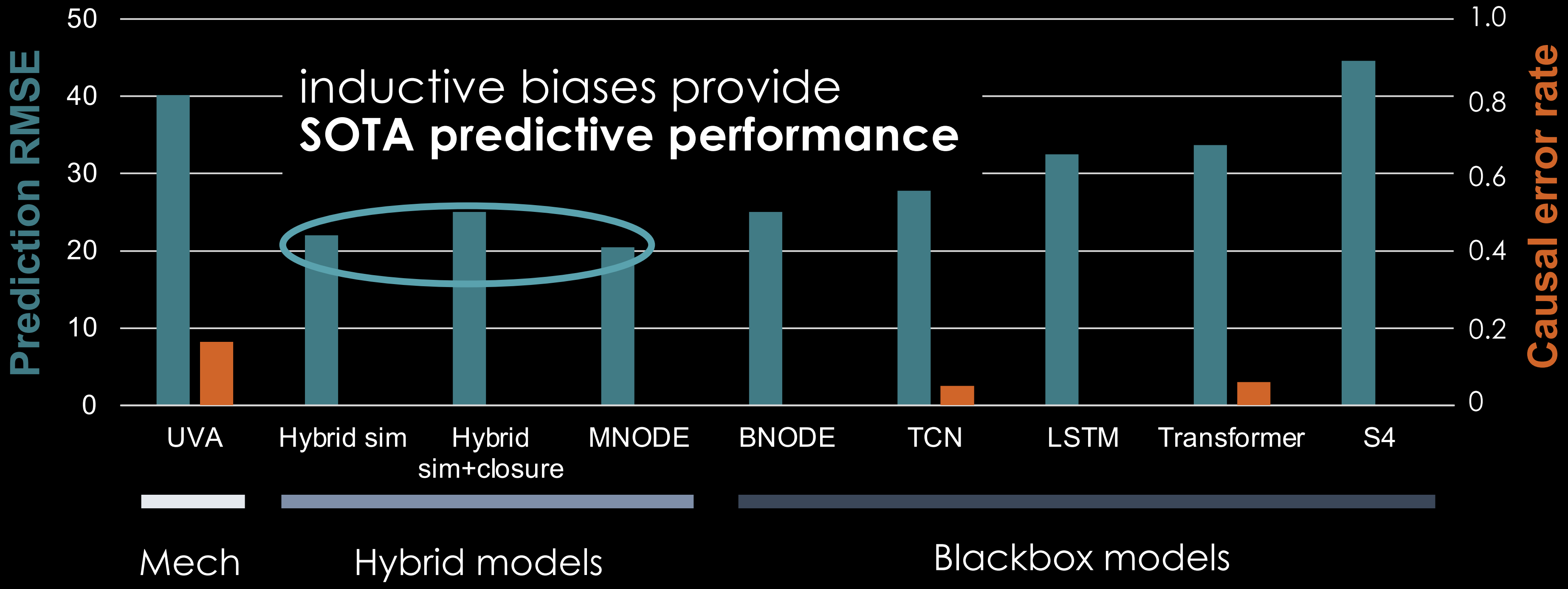
Hybrid models can **also** lose causal validity

But can also be rescued by **<u>hybrid loss</u>**!

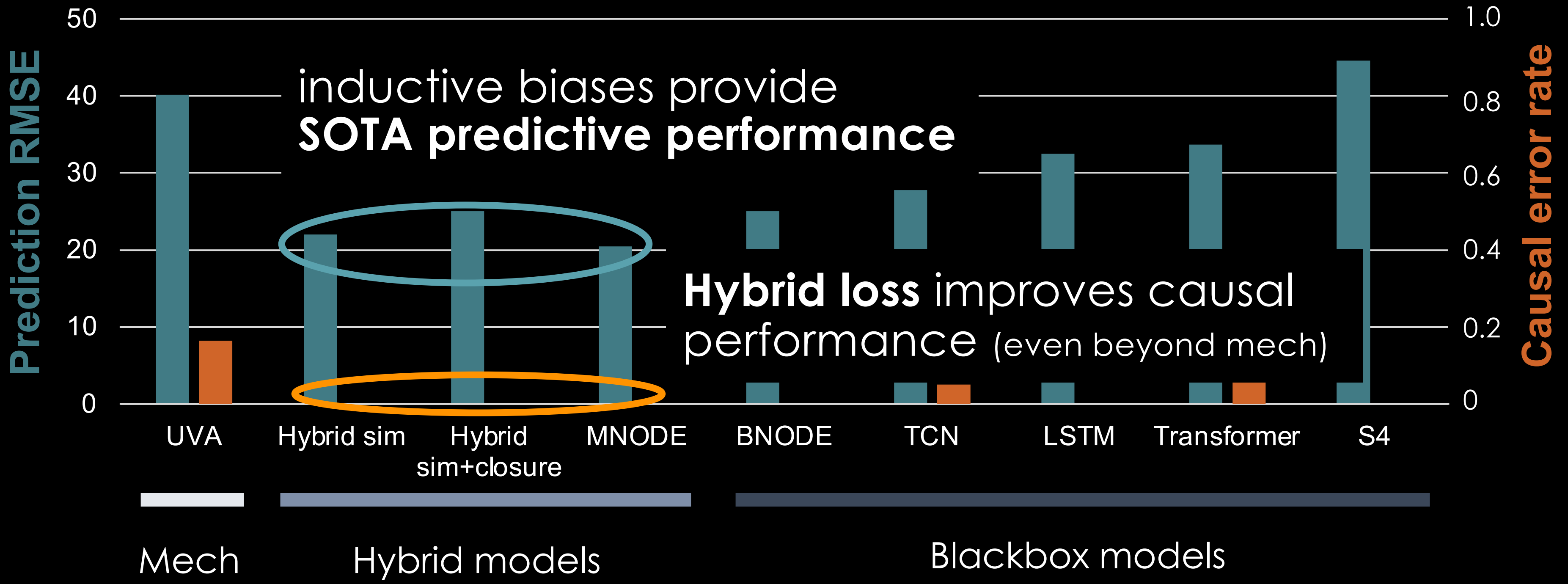


$$L_{\text{hybrid}}(\hat{M}) = (1-\alpha)L_{\text{pred}}(\hat{M}) + \alpha L_{\text{CausalError}}(\hat{M}) \quad \alpha = 0.01$$

Hybrid^2 models: **Best of both worlds**



Hybrid^2 models: **Best of both worlds**



Conclusion

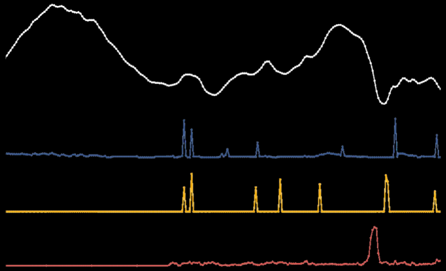
In many open prediction problems, we have **SOME** data and **SOME** knowledge.

The next generation of models will **hybridize** data-driven techniques with mechanistic knowledge.

Conclusion

- We propose, through a **<u>hybrid loss</u>**, a novel way to include inductive bias through known treatment effect rankings
  - **<u>Generality:</u>** It can be applied across methodologies and applications

$$L_{\text{hybrid}}(\hat{M}) = (1-\alpha)L_{\text{pred}}(\hat{M}) + \alpha L_{\text{causal}}(\hat{M})$$

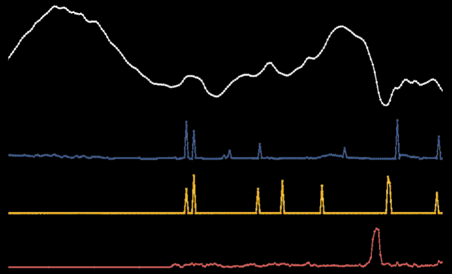


Conclusion

- We propose, through a **<u>hybrid loss</u>**, a novel way to include inductive bias through known treatment effect rankings
    - **<u>Generality:</u>** It can be applied across methodologies and applications

$$L_{\text{hybrid}}(\hat{M}) = (1-\alpha)L_{\text{pred}}(\hat{M}) + \alpha L_{\text{causal}}(\hat{M})$$

- We demonstrate utility of hybrid losses and hybrid models through **<u>hybrid² modeling</u>** in the challenging real-world scenario of modeling post-exercise glycemic response



Acknowledgements

- Helmsley Charitable Trust
- Stanford Institute for Human-Centered AI
- Eric and Wendy Schmidt Center at Broad Institute
- AFOSR, ONR, NSF, NIH
- Vivli
- Alex Wang

Paper link: https://arxiv.org/abs/2402.17233

Bob Zou

Matthew Levine

Dessi Zaharieva

Ramesh Johari

Emily Fox