DeepPumas for automatic discovery of individualizable functions governing longitudinal patient outcomes

Niklas Korsbo[1], Mohamed Tarek[1,2], Chris Elrod[1], Antoine Soubret[3], Francesco Brizzi[3], Christopher Rackauckas[1,4], Jogarao Gobburu[1,5], Vijay Ivaturi[1,5]

[1] Pumas-AI, [2] Business School, University of Sydney, [3] Roche, [4] Massachusetts Institute of Technology, [5] University of Maryland

Introduction

The recent advent of Scientific Machine Learning (SciML) is enabling us to combine the wealth of knowledge that modelers, biologists, pharmacologists, and clinicians already possess with the pattern recognition ability of machine learning. Such hybrid models of mechanistic insight, and data-driven function and pattern identification, have proven both data efficient and good at extrapolating longitudinal predictions beyond the time points used to train the model (1,2,3). So far, however, such SciML models has been used to model single or aggregated time course data without handling the between-subject variability that is often central to pharmacometrics analyses.

To address this gap, we here explore mixed effect neural networks (MeNets) for identification of parameterized and *individualizable* functions affecting patient outcomes (4,5). Being mathematical *functions,* these MeNets are usable anywhere in DeepPumas NLME models, including as terms in dynamical systems. However, to isolate the effect of using random effects as part of the input to a neural network, we here explore a very simple MeNet model that uses time-after-first-dose and random effects to directly predict individual longitudinal outcomes.

Methods

We defined five neural network models processing time (t) and a set of 0 to 4 patient-specific random effects (η), respectively. A longitudinal biomarker was modelled as the output of this neural network, with Normal additive observational noise.

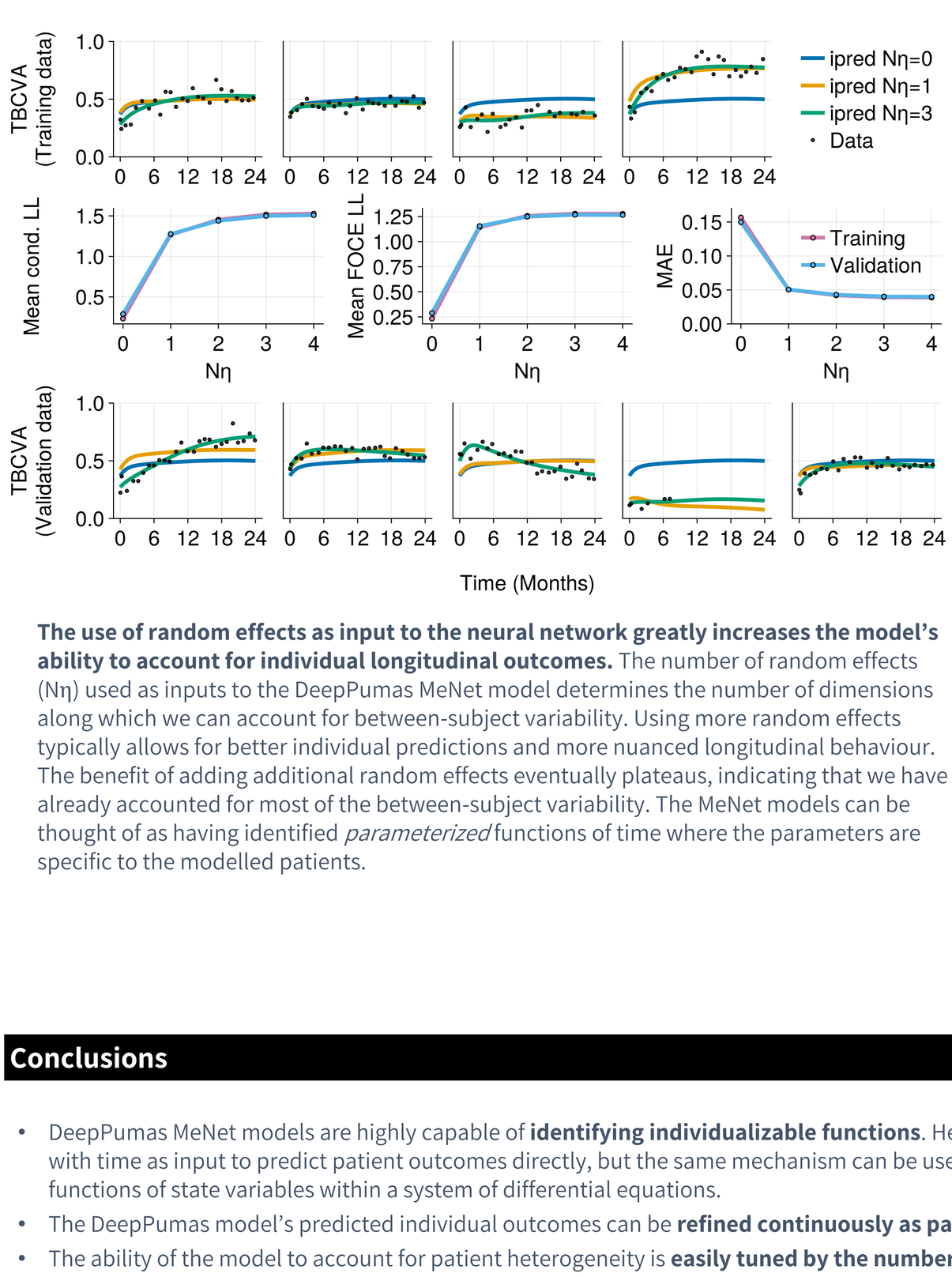
$$\eta_i \sim \text{Normal}\left(0, 0.01\right) \quad \forall i \in [1..N\eta]$$
$$Y\left(t\right) = \text{NN}\left(t, \eta_1, \ldots, \eta_{N\eta}\right)$$
$$TBCVA\left(t\right) \sim \text{Normal}\left(Y(t), \sigma\right)$$

Where NN is a multi-layer perceptron (a simple neural network) with Nη+1 inputs, two hidden layers of 14 nodes each and softplus activation, and an output layer with a single node and no activation function. The weights and biases of NN was fitted together with σ towards a marginal (FOCE approximated) and a conditional loglikelihood. L2 regularization (Normal priors) were added to the weights and biases that link the two hidden layers.

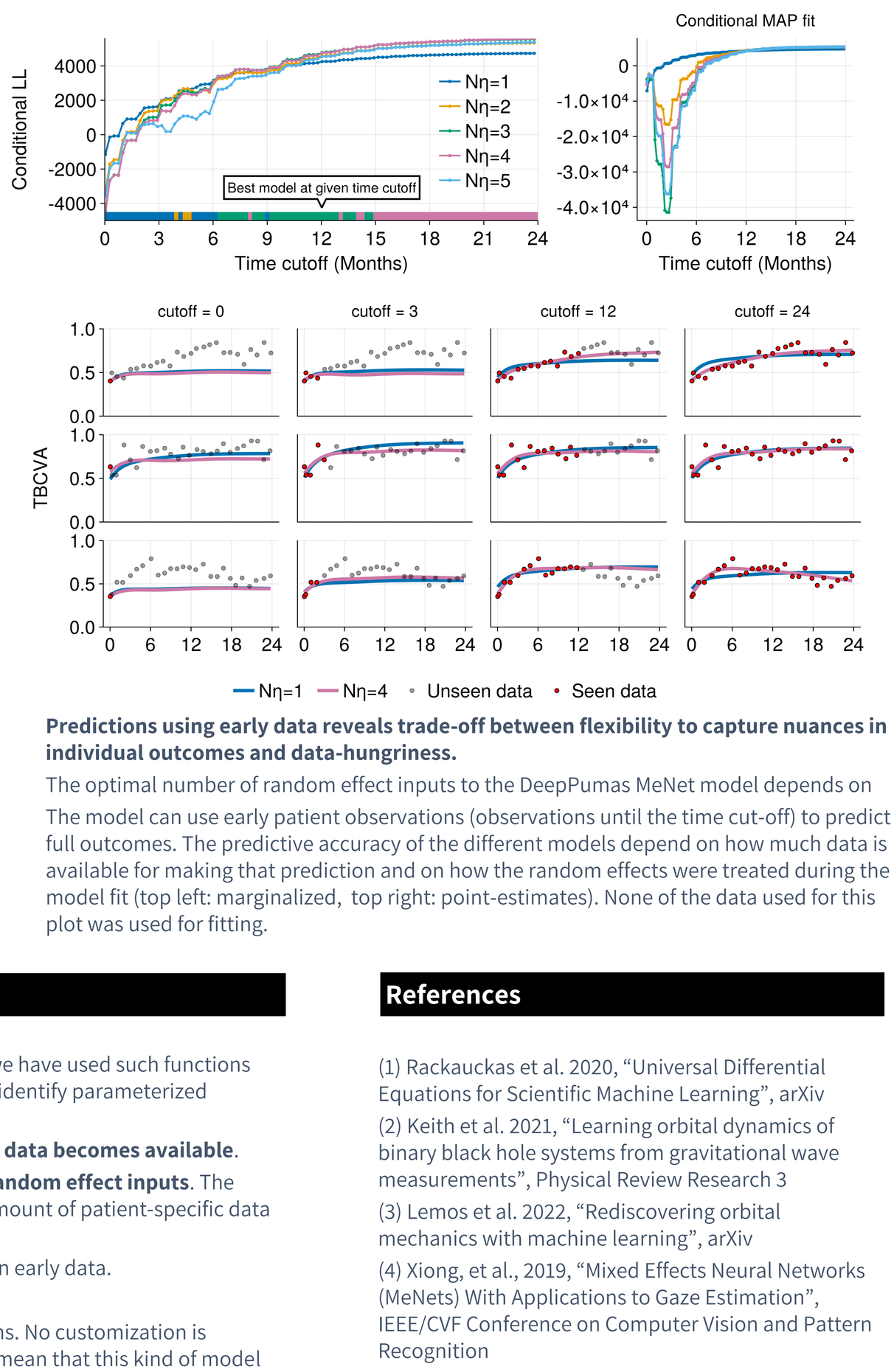
The model was applied to a synthetic data set of transformed best-corrected-visual-acuity (TBCVA). The model was fitted on data from 604 synthetic patients and the analysis was validated on 151 patients that were withheld from the fitting.

Early data predictions were made by using empirical bayes estimates computed from data before a specified time cut-off to then predict the patient's full longitudinal outcomes.

Results



The use of random effects as input to the neural network greatly increases the model's ability to account for individual longitudinal outcomes. The number of random effects (Nη) used as inputs to the DeepPumas MeNet model determines the number of dimensions along which we can account for between-subject variability. Using more random effects typically allows for better individual predictions and more nuanced longitudinal behaviour. The benefit of adding additional random effects eventually plateaus, indicating that we have already accounted for most of the between-subject variability. The MeNet models can be thought of as having identified *parameterized* functions of time where the parameters are specific to the modelled patients.



Predictions using early data reveals trade-off between flexibility to capture nuances in individual outcomes and data-hungriness.

The optimal number of random effect inputs to the DeepPumas MeNet model depends on

The model can use early patient observations (observations until the time cut-off) to predict full outcomes. The predictive accuracy of the different models depend on how much data is available for making that prediction and on how the random effects were treated during the model fit (top left: marginalized, top right: point-estimates). None of the data used for this plot was used for fitting.

Conclusions

- DeepPumas MeNet models are highly capable of **identifying individualizable functions**. Here, we have used such functions with time as input to predict patient outcomes directly, but the same mechanism can be used to identify parameterized functions of state variables within a system of differential equations.
- The DeepPumas model's predicted individual outcomes can be **refined continuously as patient data becomes available**.
- The ability of the model to account for patient heterogeneity is **easily tuned by the number of random effect inputs**. The optimal choice here is dependent on both the amount of data available for training *and* on the amount of patient-specific data you expect to have at the time when the model will be used to inform a decision.
- **Marginalizing** out the random effects during a fit greatly **improves predictive accuracy** based on early data.

The specific model used here is highly flexible and does not rely on any problem-specific assumptions. No customization is needed to apply the same model to data from different diseases altogether. However, this does not mean that this kind of model cannot benefit from using prior scientific knowledge. DeepPumas supports embedding of such mixed-effect neural networks in, for example, dynamical systems. The MeNet is then constrained to capturing only an unknown term within your dynamical system which may otherwise be defined based on all the scientific knowledge you already have about your problem. Furthermore, once individualizable functions has been identified, DeepPumas can use other machine learning tools to identify prognostic factors from covariates to improve individualized early predictions.

References

(1) Rackauckas et al. 2020, "Universal Differential Equations for Scientific Machine Learning", arXiv

(2) Keith et al. 2021, "Learning orbital dynamics of binary black hole systems from gravitational wave measurements", Physical Review Research 3

(3) Lemos et al. 2022, "Rediscovering orbital mechanics with machine learning", arXiv

(4) Xiong, et al., 2019, "Mixed Effects Neural Networks (MeNets) With Applications to Gaze Estimation", IEEE/CVF Conference on Computer Vision and Pattern Recognition

(5) Ngufor, et al. 2019, "Mixed effect machine learning: A framework for predicting longitudinal change in hemoglobin A1c", Journal of Biomedical Informatics