Genentech
A Member of the Roche Group

Neural-PK/PD as Pharmacology-Informed Deep Learning Architecture

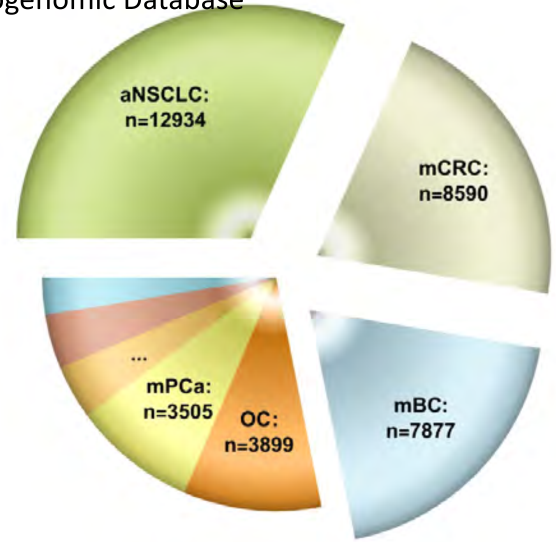
James Lu, Genentech

IQ Workshop on *Machine Intelligence for Quantitative Modeling in Drug Discovery & Development Applications*
15-16 September 2022

Motivation: the Data Challenge in the Digital Age

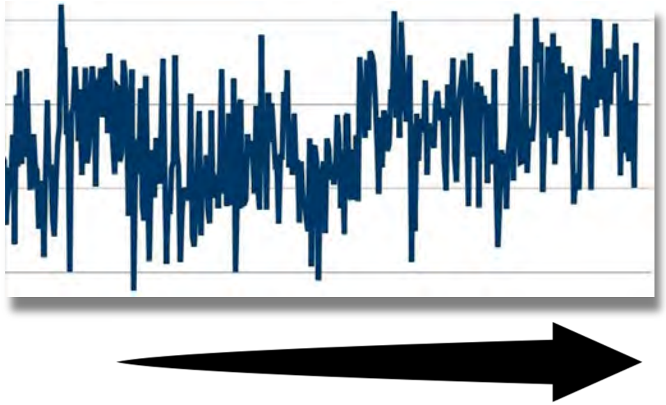
- Current trends in technology & digitization generate growth across:
  - the number of patients
  - the number of longitudinal measurements
  - the dimensionality of data

eg, Flatiron Clinicogenomic Database



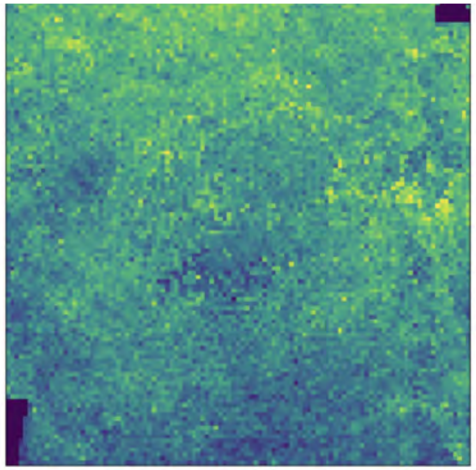
> 40K patients

eg, wearable device



> 1M measurements

eg, gene expression



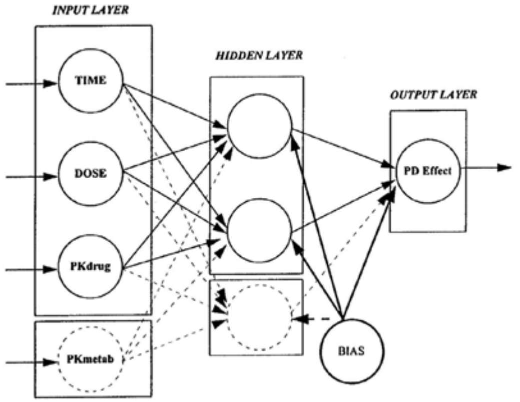
> 10K genes

Cancer Genome Atlas Network, Nature (2012)

Liu *et al*, Nature Med (2022)

A Brief History of Neural Networks for PK & PK/PD Modeling

Feed-Forward Network

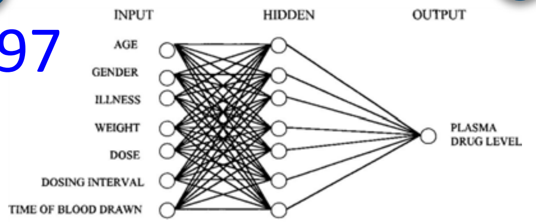


Artificial Neural Networks As a Novel Approach to Integrated Pharmacokinetic–Pharmacodynamic Analysis

JOGARAO V. S. GOBBURU*x AND EMILE P. CHEN‡

Received October 11, 1995, from the *Department of Pharmaceutical Sciences, North Dakota State University, Fargo, ND 58105, and ‡Department of Drug Metabolism and Pharmacokinetics, Hoffmann-La Roche, Nutley, NJ 07110. Accepted for publication February 19, 1996.

1997



Application of Neural Networks to Population Pharmacokinetic Data Analysis

HSIAO-HUI CHOW†x, KRISTIN M. TOLLE‡, DENISE J. ROE§, VICTOR ELSBERRY†, AND HSINCHUN CHEN‡

Received September 24, 1996, from the †Department of Pharmacy Practice and Science, ‡Department of Management Information Systems, and §Arizona Cancer Center, University of Arizona, Tucson, AZ 85721. Accepted for publication March 24, 1997.

'05

Mapping the Dose–Effect Relationship of Orbofiban from Sparse Data with an Artificial Neural Network

DONALD E. MAGER,[1] JASON D. SHIREY,[1] DERMOT COX,[2] DESMOND J. FITZGERALD,[3] DARRELL R. ABERNETHY[1]
[1]Laboratory of Clinical Investigation, National Institute on Aging, Gerontology Research Center, Baltimore, Maryland
[2]Department of Clinical Pharmacology, Royal College of Surgeons in Ireland, Dublin, Ireland
[3]University College Dublin, Dublin, Ireland

Population pharmacokinetic and pharmacodynamic models of remifentanil in healthy volunteers using artificial neural network analysis

S. H. Kang, M. R. Poynton, K. M. Kim, H. Lee, D. H. Kim, S. H. Lee, K. S. Bae, O. Linares, S. E. Kern & G. J. Noh

Deep Learning revolution

Recurrent Neural Network

International Journal of Clinical Pharmacology and Therapeutics, Vol. 59 – No. 2/2021 (138-146)

Long short-term memory recurrent neural network for pharmacokinetic-pharmacodynamic modeling

Xiangyu Liu[1,2], Chao Liu[2], Ruihao Huang[3], Hao Zhu[2], Qi Liu[2], Sunanda Mitra[1], and Yaning Wang[2]
[1]Electrical and Computer Engineering, Texas Tech University, Lubbock, TX, [2]Office of Clinical Pharmacology, Office of Translational Science, Center for Drug Evaluation Research, US Food and Drug Administration, White Oak, MD, and [3]Department of Mathematical Sciences, Michigan Technological University, Houghton, MI, USA

'20

ARTICLE

Introduction of an artificial neural network–based method for concentration-time predictions

Dominic Stefan Bräm[1,2] | Neil Parrott[1] | Lucy Hutchinson[1] | Bernhard Steiert[1]

Neural-ODE Chen *et al*, NeurIPS (2018)

ARTICLES — nature machine intelligence

Deep learning prediction of patient response time course from early data via neural-pharmacokinetic/pharmacodynamic modelling

James Lu, Brendan Bender, Jin Y. Jin and Yuanfang Guan

iScience — CellPress

Article

Neural-ODE for pharmacokinetics modeling and its advantage to alternative machine learning models in predicting new dosing regimens

James Lu, Kaiwen Deng, Xinyuan Zhang, Gengbo Liu, and Yuanfang Guan

ARTICLE

Deep compartment models: A deep learning approach for the reliable prediction of time-series data in pharmacokinetic modeling

Alexander Janssen[1] | Frank W. G. Leebeek[2] | Marjon H. Cnossen[3] | Ron A. A. Mathôt[1] | for the OPTI-CLOT study group and SYMPHONY consortium

'21

ARTICLE

An artificial neural network–pharmacokinetic model and its interpretation using Shapley additive explanations

Chika Ogami[1,2] | Yasuhiro Tsuji[2] | Hiroto Seki[3] | Hideaki Kawano[4] | Hideto To[1] | Yoshiaki Matsumoto[5] | Hiroyuki Hosono[3]

...

PK and PK/PD Modeling from the Dynamical Systems Perspective

- Two distinct types of Dynamical Systems
  - **Autonomous**:

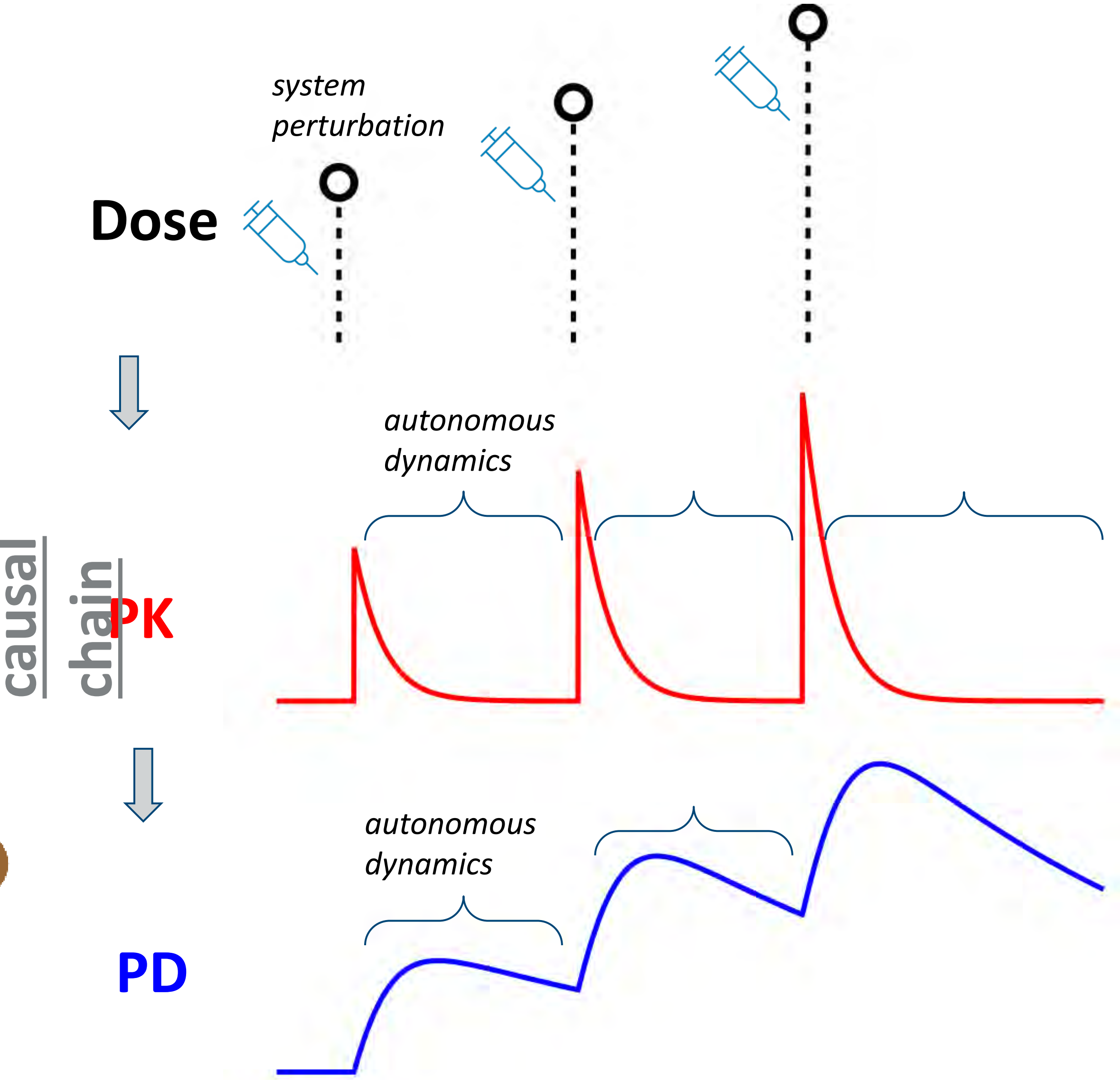
    $$y'(t) = f(y(t), p)$$

  - **Non-autonomous**:

    $$y'(t) = f(y(t), t, p)$$

- PK and PK/PD models are piecewise autonomous systems with dosing introducing time-dependent perturbations:

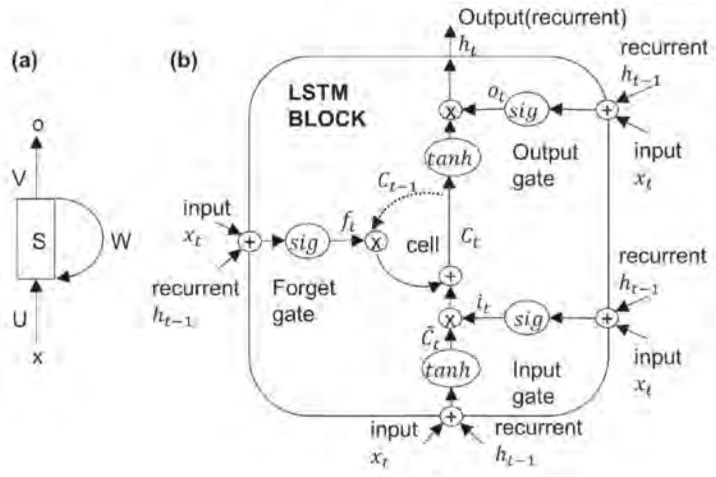
$$y'(t) = \sum_{i=1}^{n} \mathrm{dose}(i)\, \delta(t - T_i) + f(y(t), p)$$



Distinctions with alternative Deep Learning Models for Time Series Data

Standard DL architectures for Time Series data (such as LSTM) do not explicitly encode causality relationships between **Dose**, **PK** and **PD**

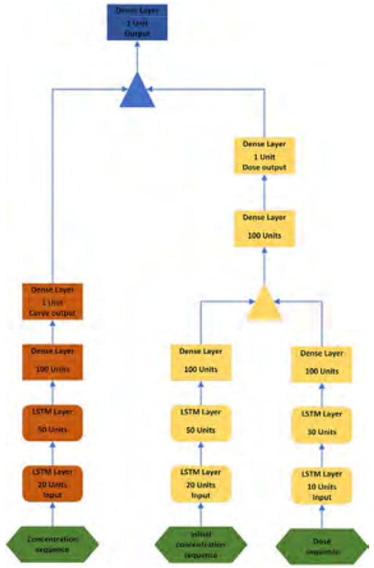
- All input data (x) enters the neural net at an equal footing
- Potentially challenging to perform counterfactual ("what-if") simulations



Liu *et al*, IJCPT (2021)

Necessitated the development of architectures that separates out:

- Concentration sequence data
- Dose sequence data



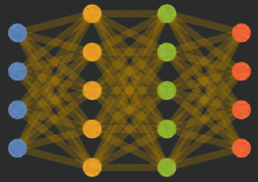
Bram *et al*, CPT:PSP (2022)

Genentech
A Member of the Roche Group

The Merging of Deep Learning with Dynamical Systems

Deep Learning

- Excels in approximating high dimensional/nonlinear functions
- Learn to improve model as data increases

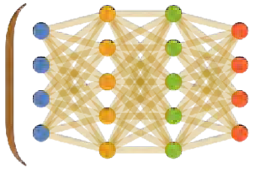


Dynamical Systems

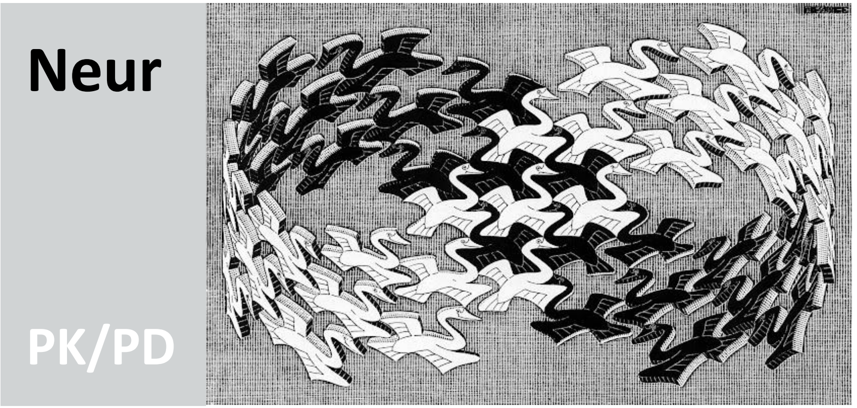
- Enables assumption of causal dynamical relationships
- Enables dosing implementation as perturbation in the system state:

$$\frac{dy(t)}{dt} = \sum_{i=1}^{n} \mathrm{dose}(i)\,\delta(t - T_i) + f(y(t), p)$$

$$\frac{dy(t)}{dt} = \sum_{i=1}^{n} \mathrm{dose}(i)\,\delta(t - T_i) + \Bigg($$



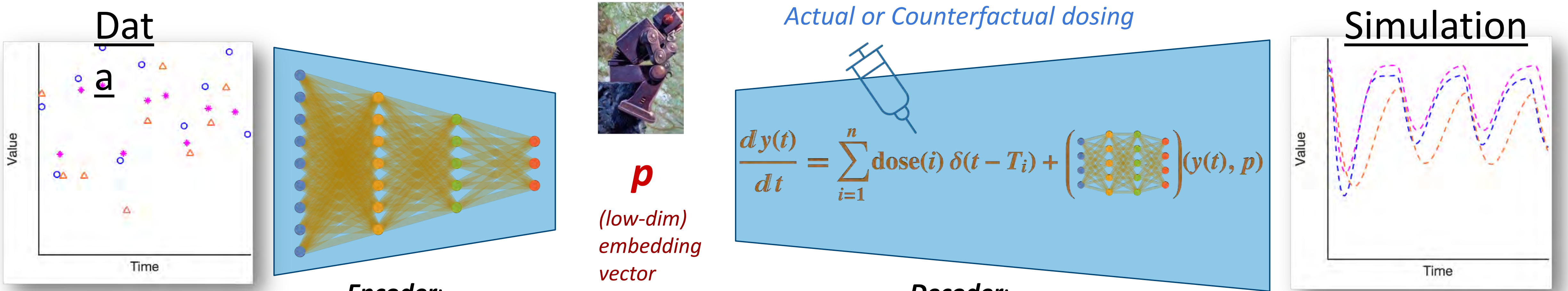
$$\Bigg)(y(t), p)$$



A pharmacology-informed dynamical system that learns from data and enables counterfactual dosing simulations

The Architecture of Neural-PK/PD enhances Interpretability

Pharmacology-Informed Encoder-Decoder Architecture



Encoder:
data abstraction
at patient level

Decoder:
predictions using a <u>single</u>
system of ODEs

Benefits:

- Explicit control on the parameter dimension for characterizing *inter-individual variability*
- Explicit control on the state dimension for reproducing observed dynamics
- Explicit dosing port enables counterfactual simulations under different dosing regimens

Formulations of Population-PK/PD vs Neural-PK/PD

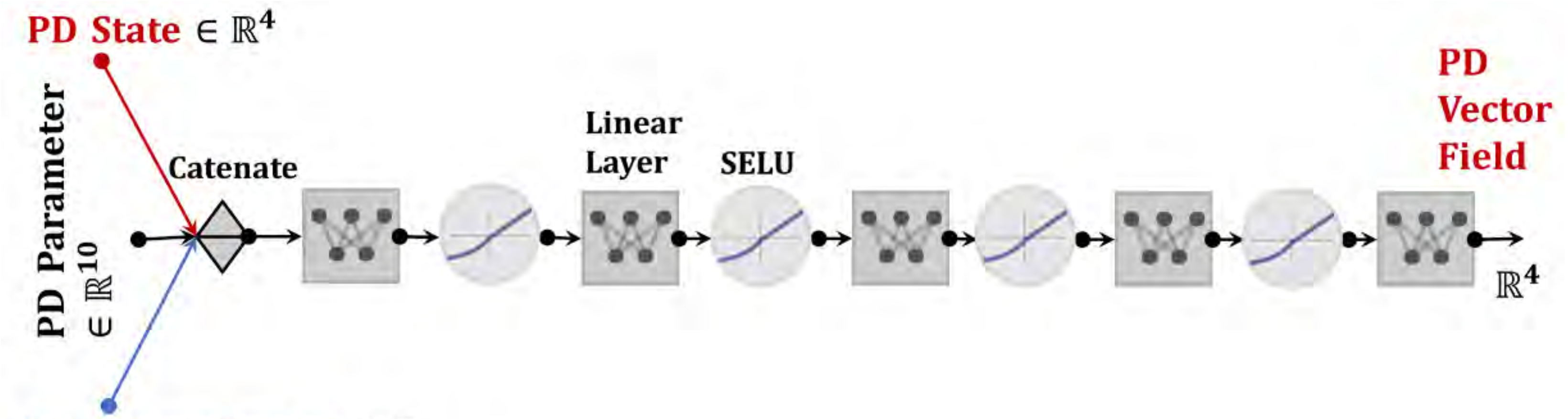
Human Generated Model

Equations based on human experience & data fit

DADT(1)=K21*A(2)-K12*A(1)-K10*A(1)
DADT(2)=K12*A(1)-K21*A(2)
DADT(3)=KEO*CP-KEO*A(3)
DADT(4)=-KTR*A(4)+KTR*A(4)*(1-EFFPLT)*(BT/A(8))GAM**
DADT(5)=-KTR*A(5)+KTR*A(4)
...

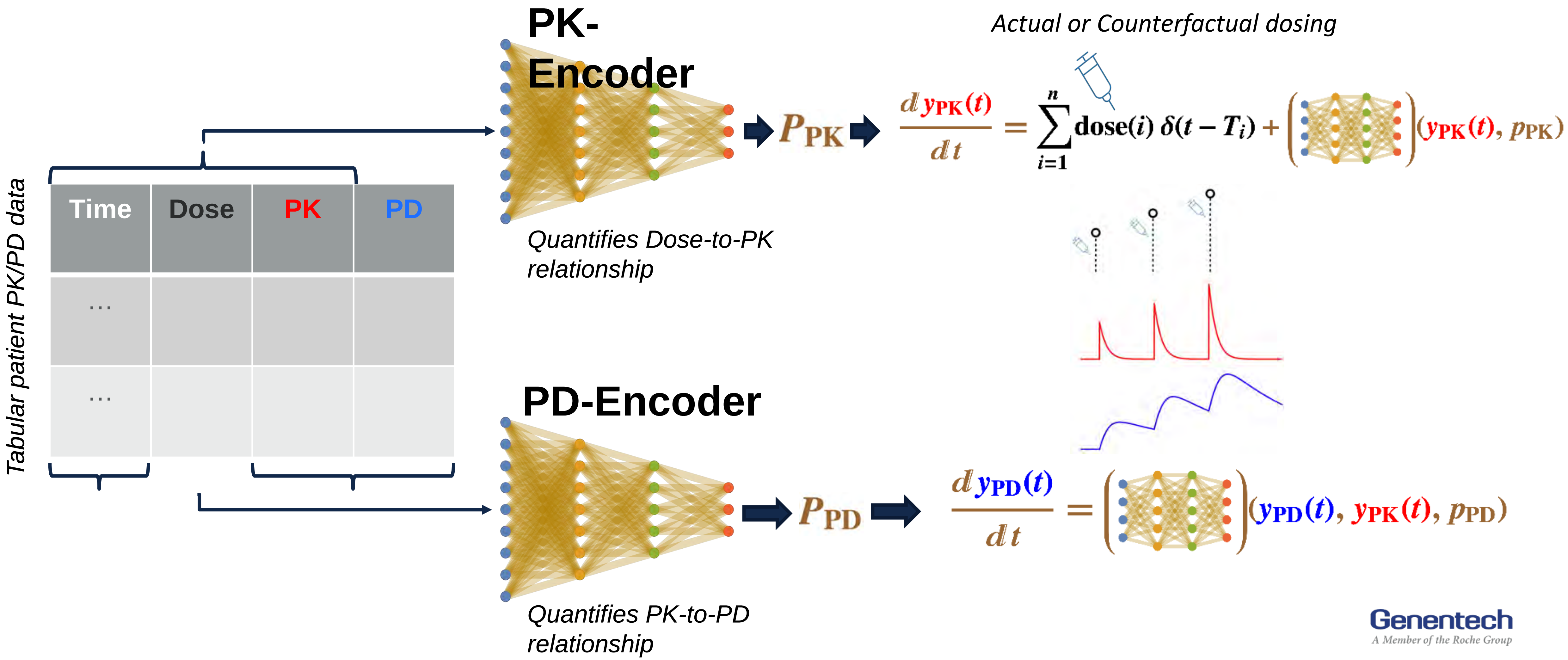
Neural Network Generated Model

Neural network learned from PK & PD data



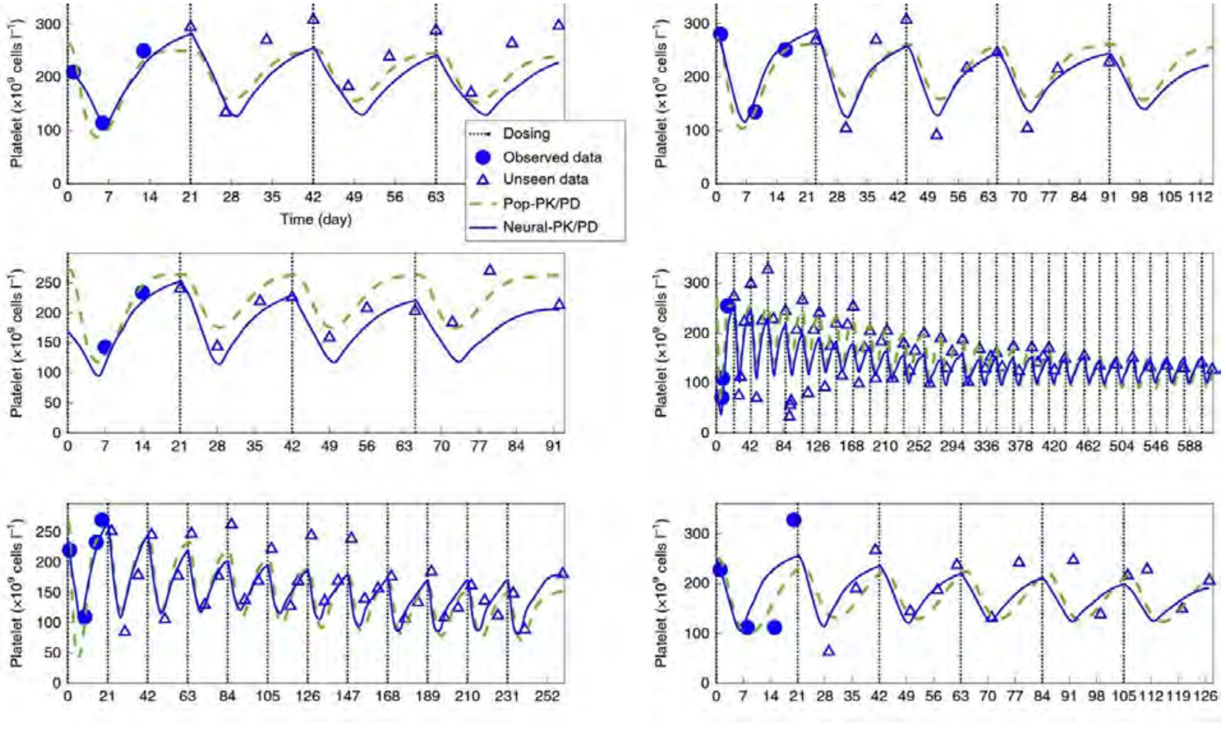
Lu *et al*, Nature Mach Intell (2021)

Encoding Causality Assumption into Neural-PK/PD Architecture



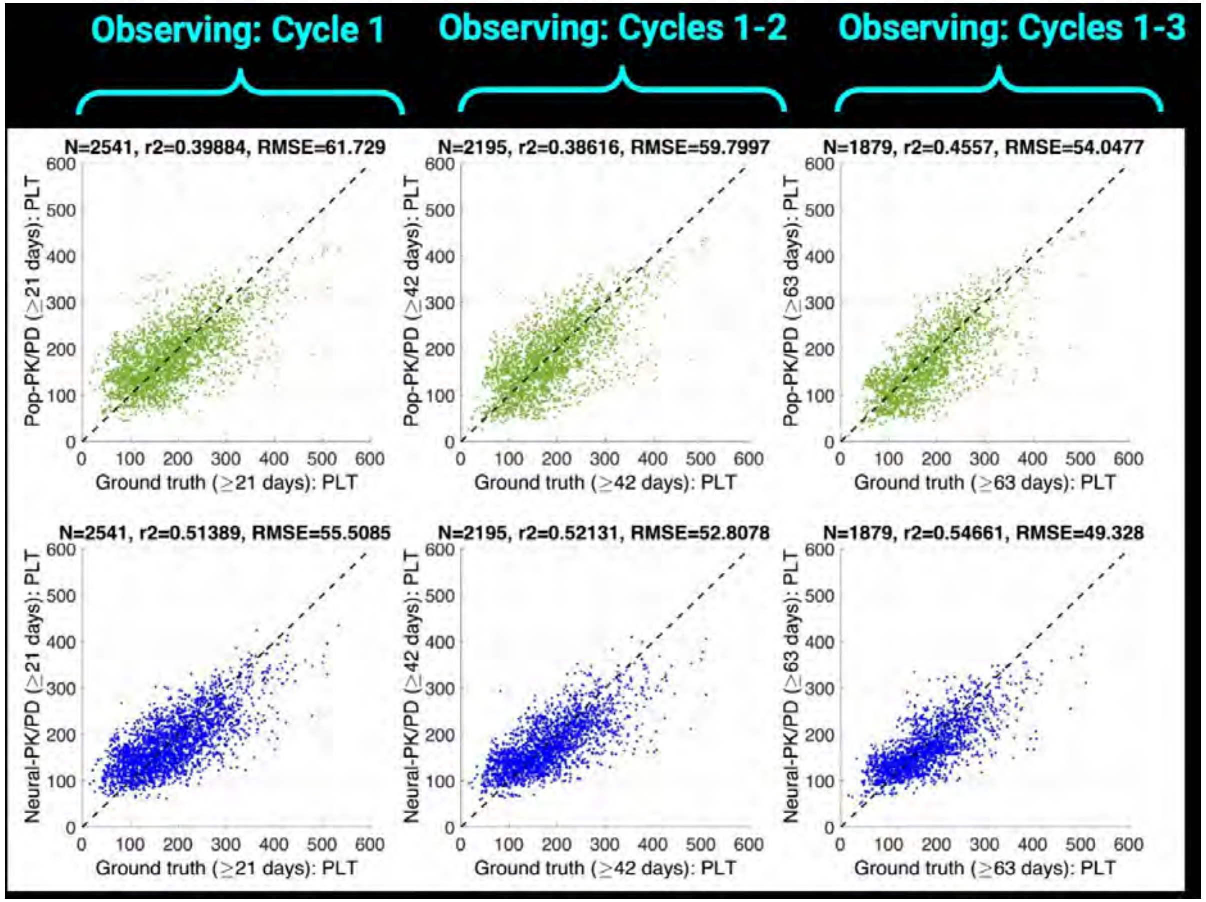
Neural-PK/PD Enables Accurate Predictions in Time

Learned model demonstrates qualitatively similar dynamics but more precise predictions as compared to the state-of-the-art pop-PK/PD model for platelet dynamics



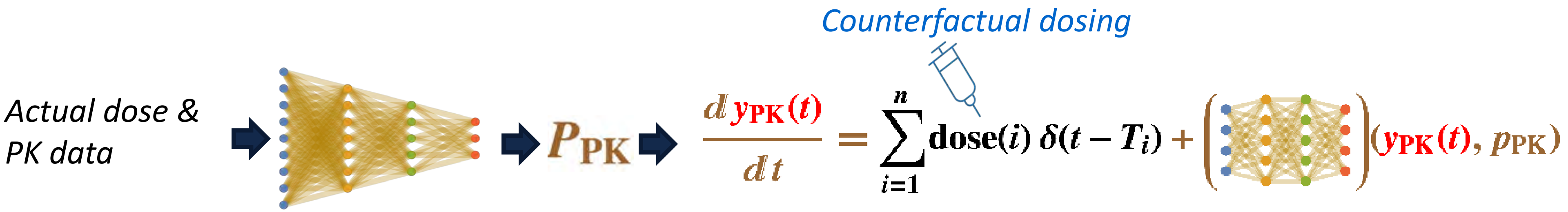
Lu *et al*, Nature Mach Intell (2021)

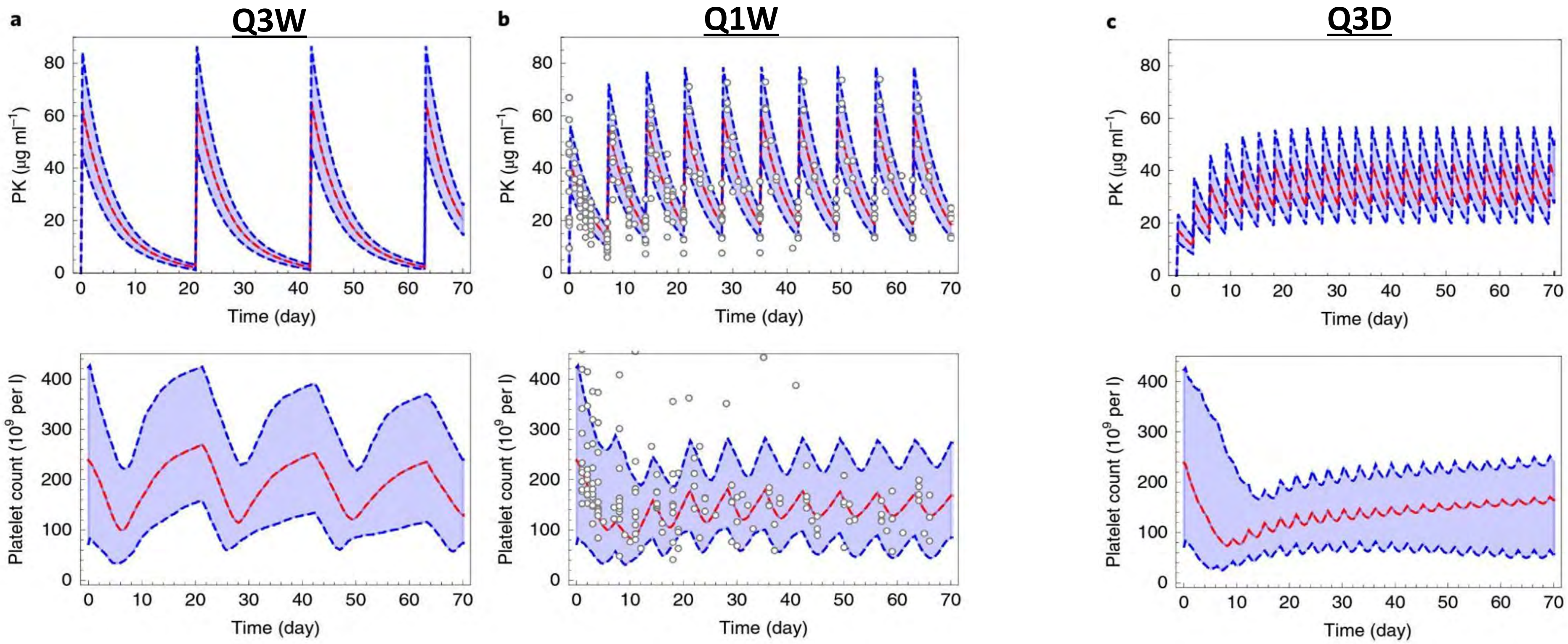
Model demonstrates superior predictive performance



Neural-PK/PD Enables Predictions for Alternate Dosing Regimens

Simulations of counterfactual dosing regimens shown to be consistent with data





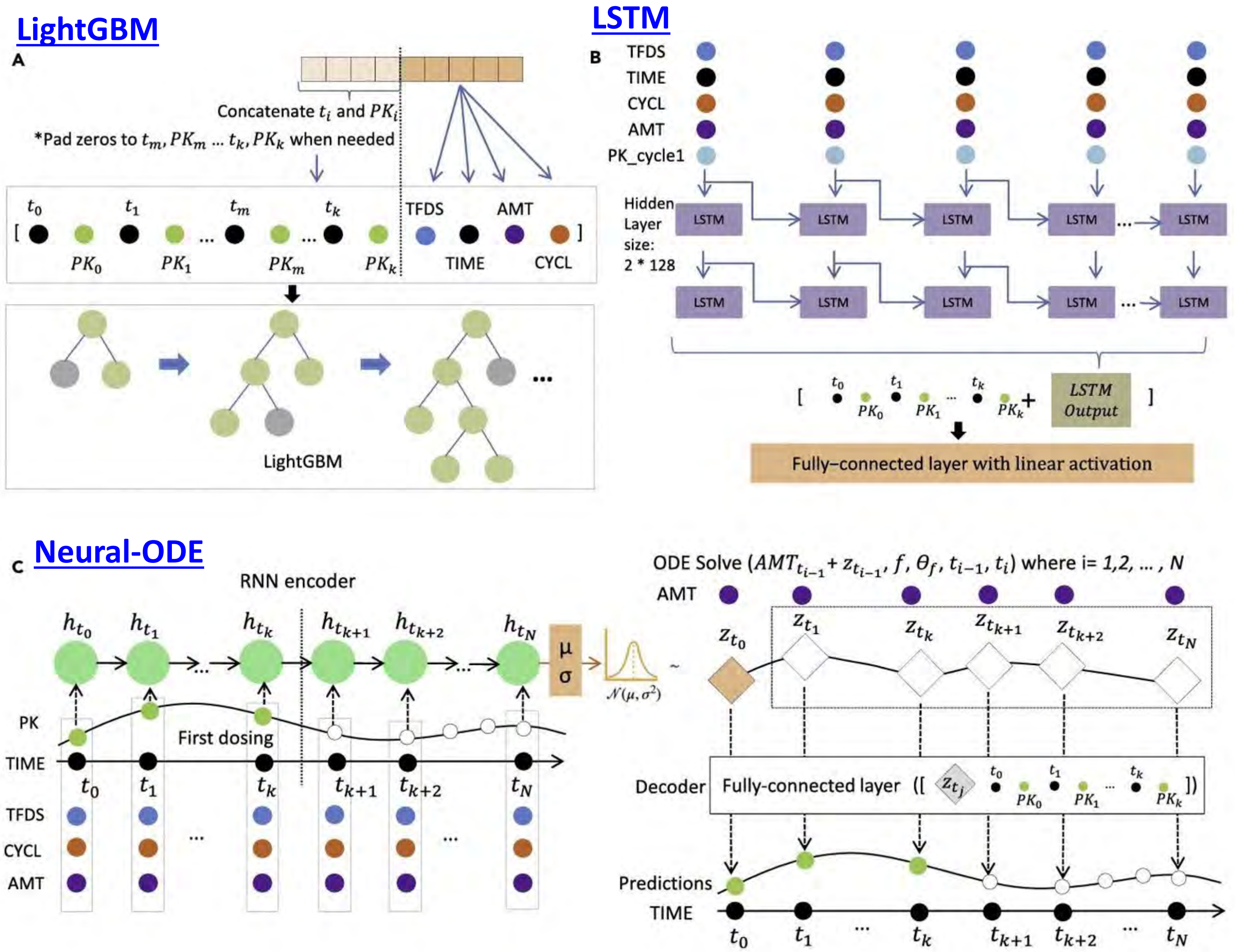
Lu *et al*, Nature Mach Intell (2021)

Genentech
A Member of the Roche Group

The Importance of the Architecture Choice in Generalizability

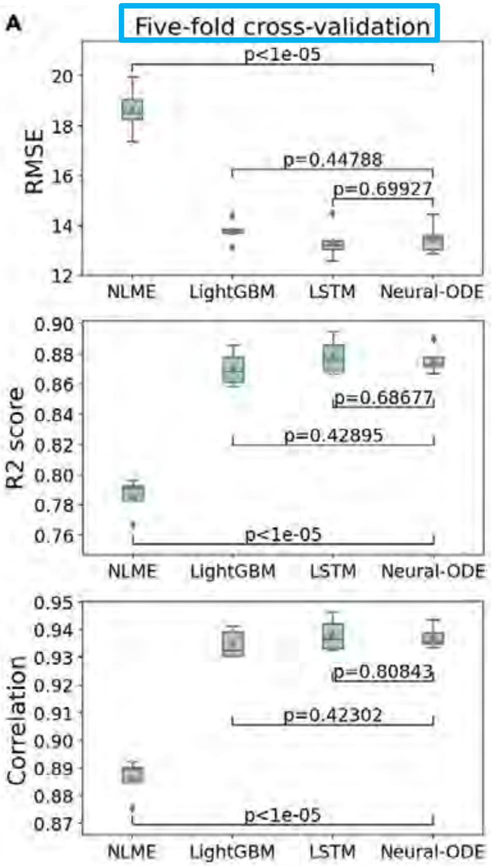
- Comparison of Neural-ODE based architecture vs alternate models:


iScience
CelPress OPEN ACCESS
Article
Neural-ODE for pharmacokinetics modeling and its advantage to alternative machine learning models in predicting new dosing regimens
James Lu, Kaiwen Deng, Xinyuan Zhang, Gengbo Liu, and Yuanfang Guan

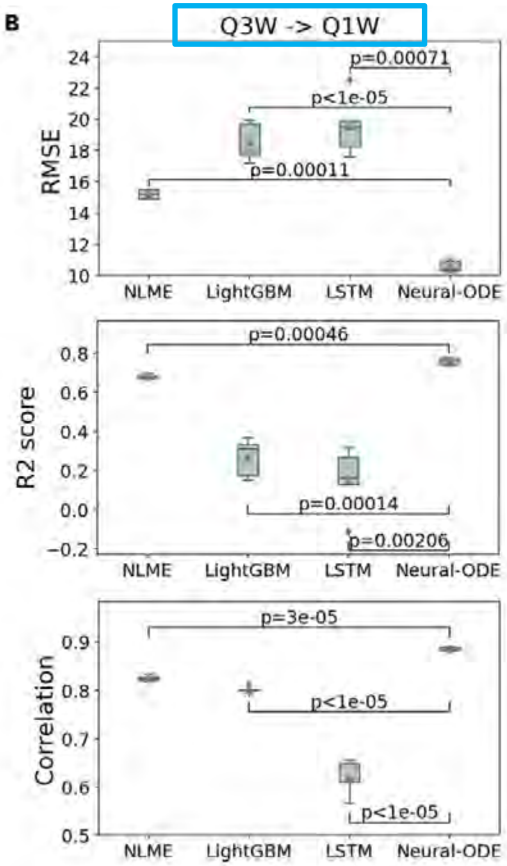


The Importance of the Architecture Choice in Generalizability

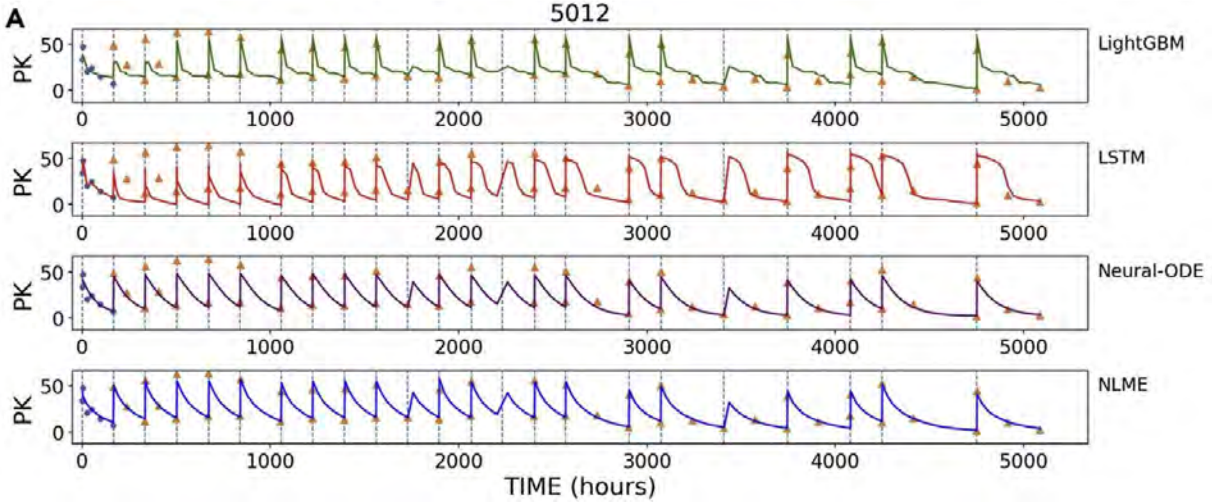
ML/DL models exhibit similar performance when train and test sets are the same distribution



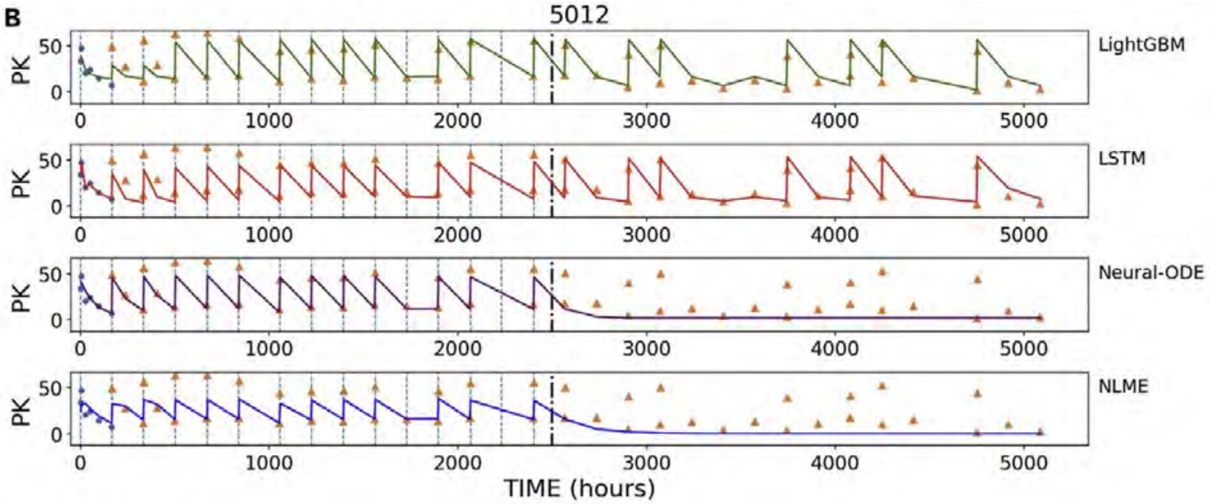
Models such as LightGBM and LSTM do not generalize well across dosing regimens



Comparison of model predictions for a patient



Comparison of model predictions with dosing stopped @ t=2500 hour



Lu *et al*, iScience (2021)

Conclusion

- The rise of complex, high volume PK/PD data in the Digital Age necessitates the development of AI methodologies
- Neural-PK/PD is a pharmacology-informed DL model that encapsulates *dose-concentration-effect* relationship within an encoder-decoder architecture:
  - leverages dynamical systems concepts
  - encoder bottleneck may improve better generalization from limited data
  - decoder enables counterfactual simulations

Acknowledgement

Genentech

- Gengbo Liu, Brendan Bender, Jin Jin, Amita Joshi

University of Michigan/Ann Arbor Algorithms

- Kaiwen Deng, Xinyuan Zhang, Yuanfang Guan

Genentech

A Member of the Roche Group