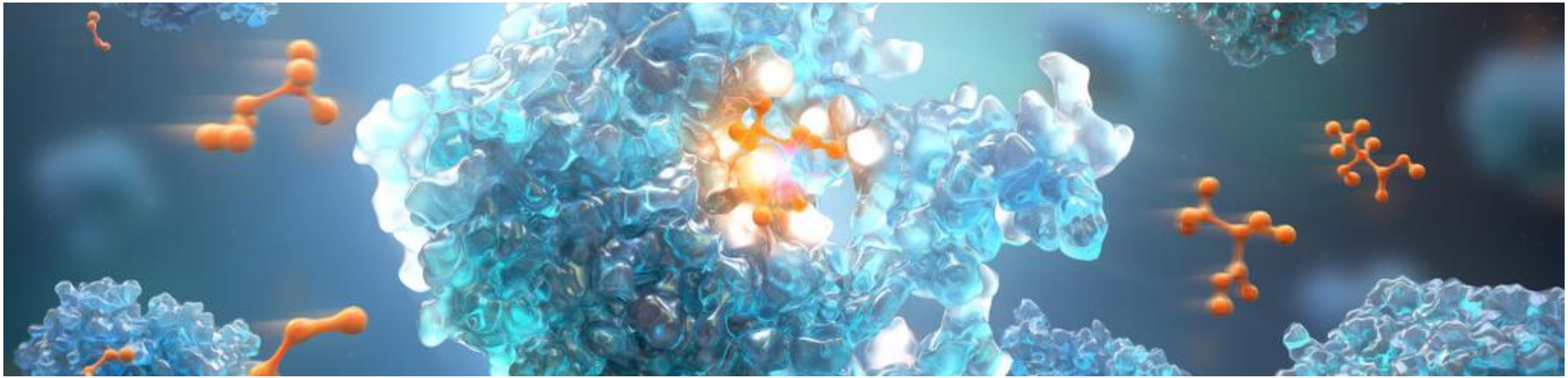
Towards improved safety and efficacy profiles of compounds by predicting *in vivo* pharmacokinetics using machine learning

Dr Olga Obrezanova, Imaging and Data Analytics, Clinical Pharmacology & Safety Sciences, R&D, AstraZeneca, Cambridge, UK

***In Silico* Toxicology Network Meeting 2020**

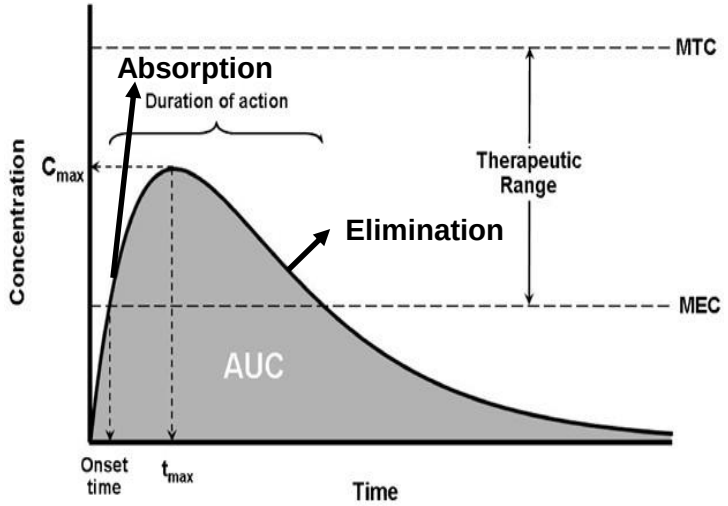
30 September 2020



Importance of pharmacokinetic studies in drug discovery

- Animal pharmacokinetic (PK) data and human and animal *in vitro* systems are utilised
  - to define the rate and route of drug elimination
  - to ensure sufficient compound concentration, for the required duration to achieve efficacy
  - but also low and short enough for an appropriate safety
  - to best inform first time in human dosing
- DMPK studies are critical in maximising the probability of developing successful drugs
- Accurate prediction of PK parameters in animals
  - provides a degree of confidence for extrapolation to human
  - can be used to improve design during drug discovery
  - help to select compounds with better properties

'The dose makes the poison'
The dose also makes for an efficacious and safe drug



MEC - minimum effective concentration
MTC - maximum tolerated concentration

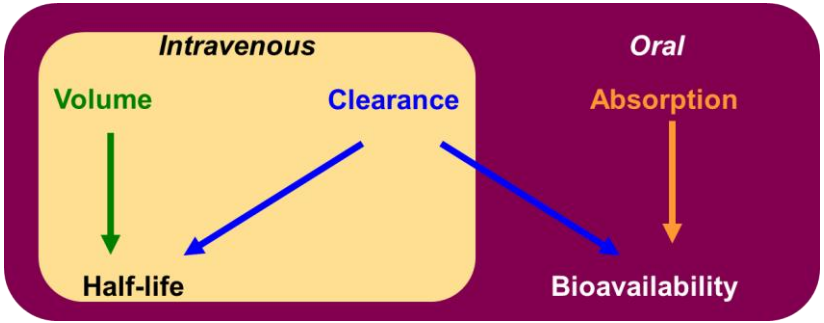
In vivo rat PK model

- **The model** predicts rat *in vivo* PK parameters and concentration-time profiles from chemical structure & measured and/or predicted *in vitro* ADME properties
- **Advantages and purpose of model application**
  - Enable prediction of PK for virtual compounds at the point of design
  - Focus designs to answer specific hypotheses
  - Drive prioritisation of compounds for *in vivo* assays - the right compound to be tested at the right time in the right assay
  - Reduce the number of *in vivo* experiments
- The model is a stepping stone for modelling human PK
- The project started as collaboration with Optibrium and Intellegens
  - Exploration of various machine learning approaches to build the model

optibrium intellegens

Overview of key PK parameters

- **Absorption:** Oral drug delivery projects strive for high oral absorption and bioavailability (F) to achieve optimal in vivo exposure
  - Absorption is influenced by the physicochemical properties, and by solubility and cell permeability and efflux
- **Distribution:** Volume (Vss) is a measure of the extent of distribution and binding of a compound in organs and tissues
- **Metabolism & Elimination:** Projects strive for low clearance (CL) to achieve acceptable duration of target engagement
  - Clearance is one of the most challenging parameters to optimise in drug discovery
  - Hepatic metabolic elimination is the predominant route of clearance for most drugs



- Volume and clearance dictate half-life
- Absorption and clearance drive oral bioavailability

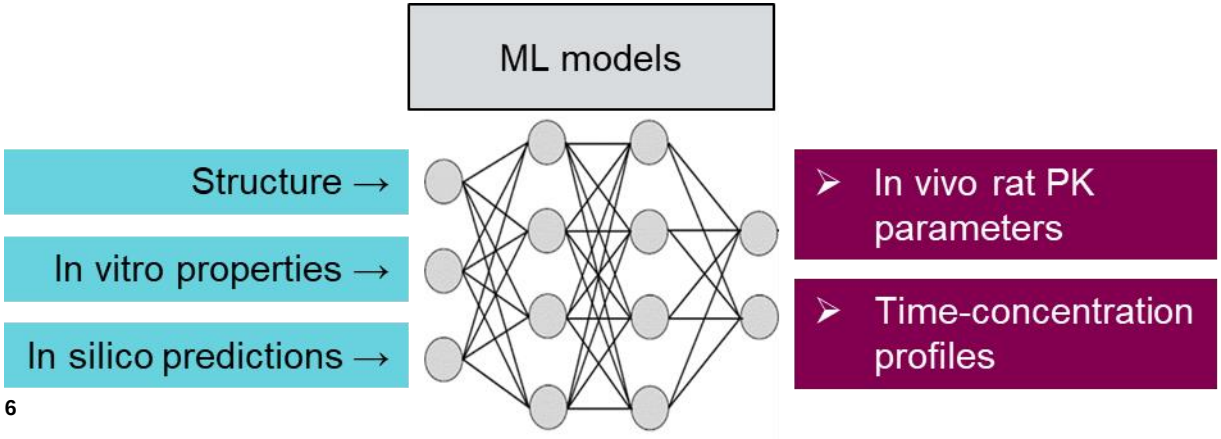
Available data

- Rat pharmacokinetic studies for i.v. and p.o. administration for around 3000 compounds
  - Two replicates for majority of compounds
- Dose dependent time-concentration curves
  - 9-11 points
- 9 PK parameters
  - **Bioavailability**
  - **Clearance** (i.v.)
  - $\mathbf{C_{max}}$ - maximum concentration (i.v. and p.o.)
  - $\mathbf{t_{1/2}}$ (i.v. and p.o.)
  - $\mathbf{V_{ss}}$ - volume of distribution (i.v.)
  - **AUC** (i.v. and p.o.)
- 9 measured *in vitro* ADME properties are used as features
  - Corresponding *in silico* predictions are available

***In vitro* ADME properties**
LogD
Caco2 intrinsic permeability
Human liver microsome intrinsic clearance
Rat hepatocyte intrinsic clearance
Rat Plasma Protein Binding
Human Plasma Protein Binding
Solubility
Fraction unbound in the rat hepatocyte
Caco2 efflux ratio

Modelling approaches

- Chemical descriptors – graph convolutions, molecular properties, ECFP4, signatures
- Addition of 9 measured *in vitro* properties (*in silico* predictions used for missing values)
- Validation on 10% temporal test set
- State-of-the-art machine learning (ML) methods explored
- Prediction of PK parameters and time-concentration curves



ML algorithms:

- Deep Neural Networks (DNN)
  - ChemProp
  - Alchemite™
- Gradient Boosting
- SVM
- Gaussian Processes

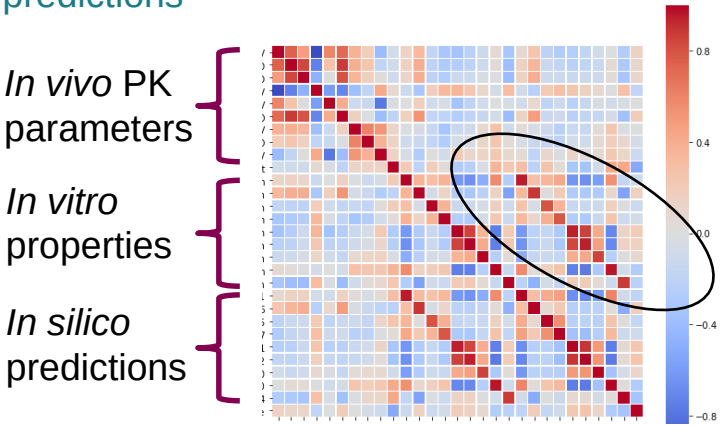
Imputation of missing experimental data

- 10-55% of *in vitro* ADME data is missing depending on the property, on average 25%
- C-Lab models are internal AstraZeneca models predicting ADME properties
  - High correlation between *in vitro* properties and corresponding *in silico* predictions
  - *In vitro* data is training data for C-Lab models

Two different approaches for data imputation

- **AZ approach** - missing *in vitro* properties are replaced with *in silico* C-Lab predictions
- **Alchemite™** imputation approach – embedded in the NN training
  - Both *in vitro* ADME properties and corresponding *in silico* predictions are available to train the model

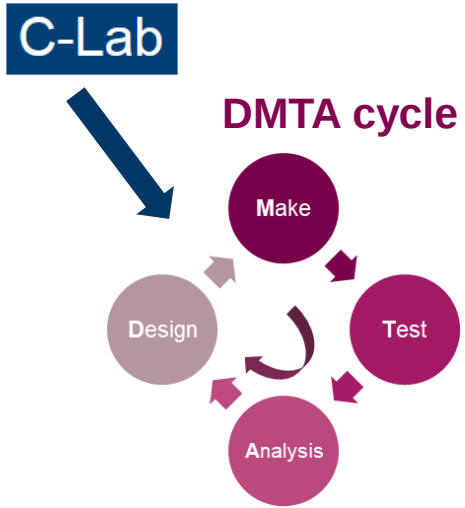
Correlation between *in vivo* PK parameters, *in vitro* ADME properties and *in silico* predictions



ADME & Safety modelling at AstraZeneca

C-Lab models

- **Global** models for physico-chemical, ADME and safety properties based on in-house data (around 40 models)
  - Dataset sizes range from 4000 – 170,000 compounds
- Available through various DMTA tools
  - Can be used during compound design or later in discovery to assist with selecting tool compounds or candidates
  - Approximately 1,200,000 uses per month
- Methods and approaches
  - Temporal-based test sets for real life scenario validation
  - Algorithms: DNN, SVM with conformal prediction framework, RF, kNN
  - Descriptors: 1/2D molecular descriptors, fingerprints and signatures
  - Updated on a regular basis, depending on the amount of new data



C-Lab models influence the decision to-make and increase the success rate of compound reaching development

AstraZeneca computational approaches

ChemProp

- Deep NN with graph convolutions
- NN encoded descriptors
- Used in multi and single task modes
- Yang *et al*. JCIM 2019
- MIT industry consortium
  http://chemprop.csail.mit.edu

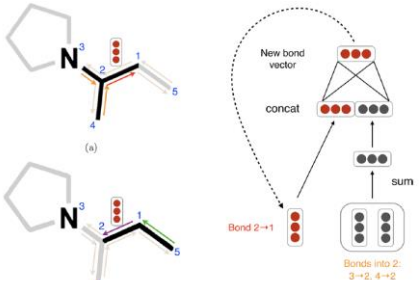
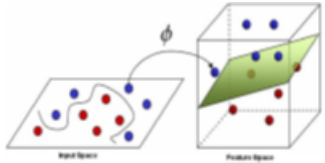


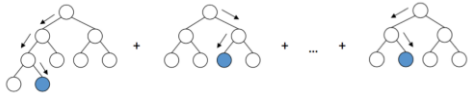
Image from Yang *et al*. JCIM 2019

Support Vector Machine (SVM)

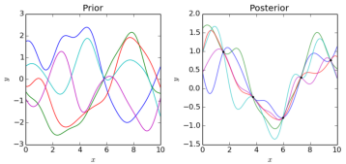
- Signature descriptors



Gradient Boosting regression



Gaussian Processes



2D molecular property descriptors

were used with all algorithms

- As well as *in vitro* ADME properties

MIT MLPDS
Machine Learning for Pharmaceutical Discovery and Synthesis Consortium

Alchemite™ method

- Alchemite™ - deep learning technology, able to handle sparse and noisy experimental data
  - Imputation of missing values
    - The Y values (endpoints) and molecular descriptors are treated as both inputs and outputs of the neural network
    - The neural network iteratively improves that initial estimate for the missing values
  - Used in multi-task mode
  - Whitehead *et al* JCIM 2019
- StarDrop descriptors (330)
  - whole molecule properties
  - counts of substructural fragments (SMARTS patterns)
- In addition, *in vitro* properties + C-Lab predictions

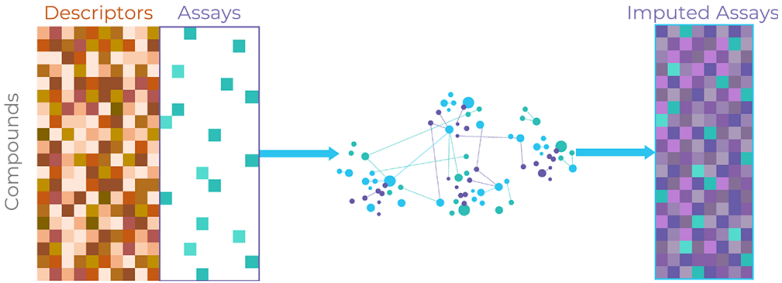


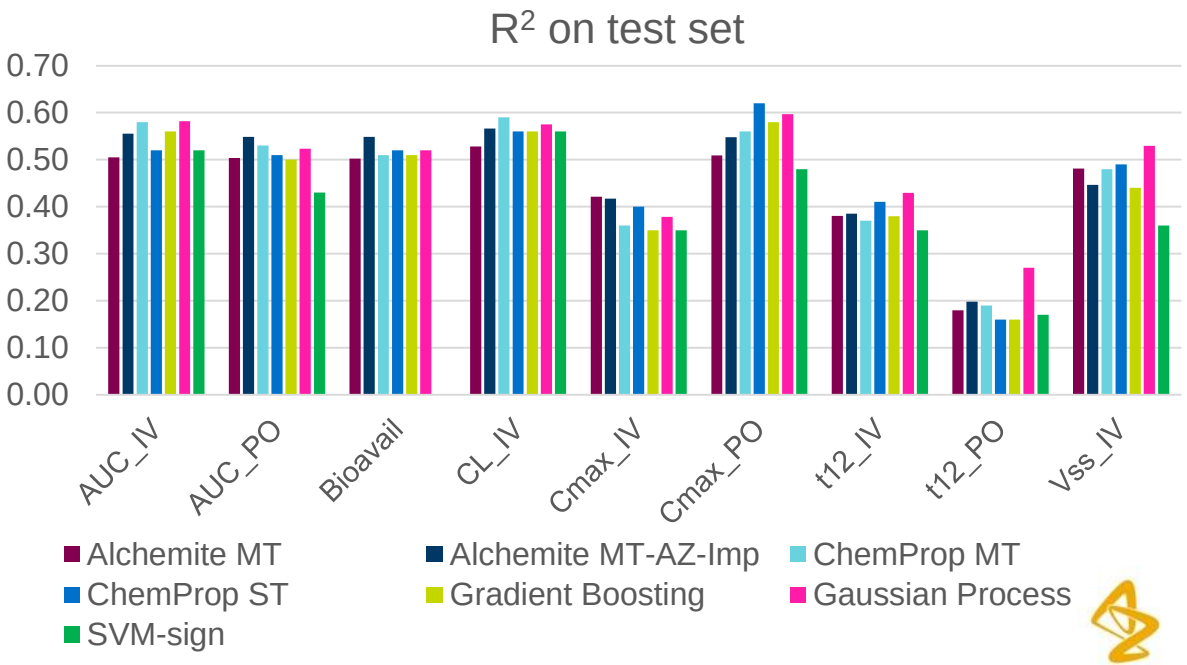
Image from Whitehead *et al*. JCIM 2019

optibrium intellegens

Summary of models performance

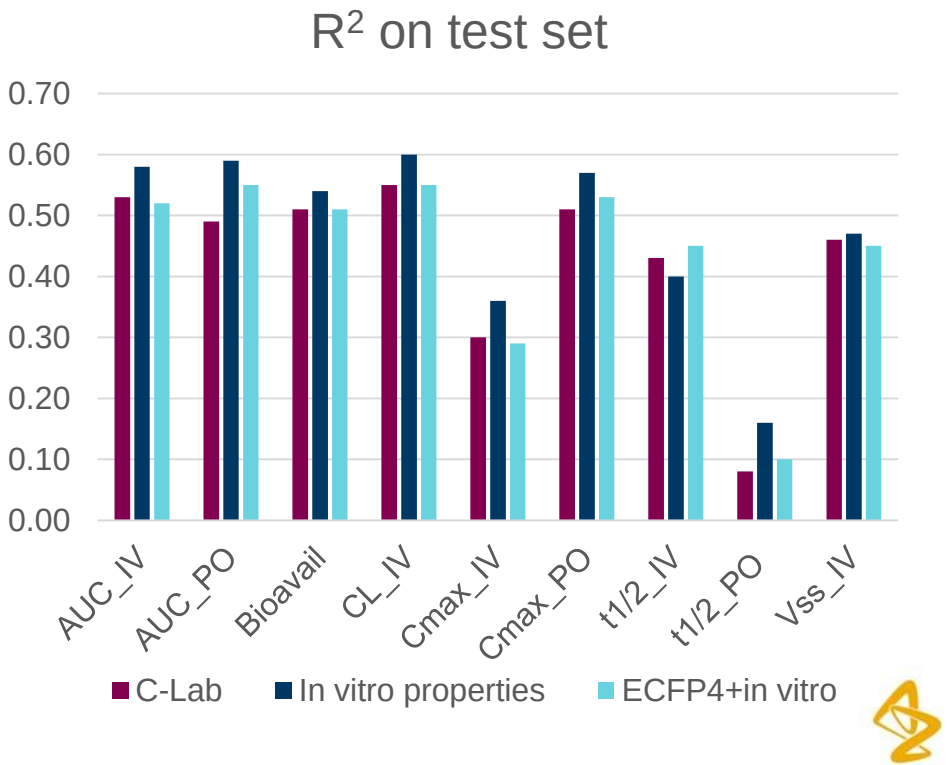
Algorithms show equivalent performance for most parameters

- Good models for majority of PK parameters
- C_{max} iv and half life are difficult to predict
- AZ imputation approach provides better results for most properties in comparison with Alchemite approach
- N_{train} = 2758, N_{test} = 312 compounds
- All PK parameters were log-transformed except half-life (no transformation) and bioavailability (logit)
- Achemite MT-AZ-Imp – AZ way of imputation, missing *in vitro* data replaced with *in silico*



In silico predictions vs *in vitro* properties as features

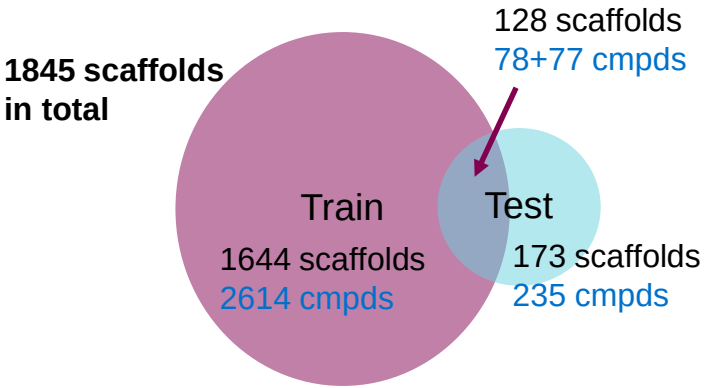
- Gradient Boosting Regression
- 2D molecular property descriptors with addition of
  - C-Lab ADME predictions **OR**
  - *In vitro* ADME properties **OR**
  - Circular fingerprints ECFP4s + *in vitro*
- **Addition of *in vitro* properties improves models in most cases**
- **Addition of ECFP4s does not improve models for most parameters**



Similarity of the test set to training set

Bemis-Murcko scaffolds

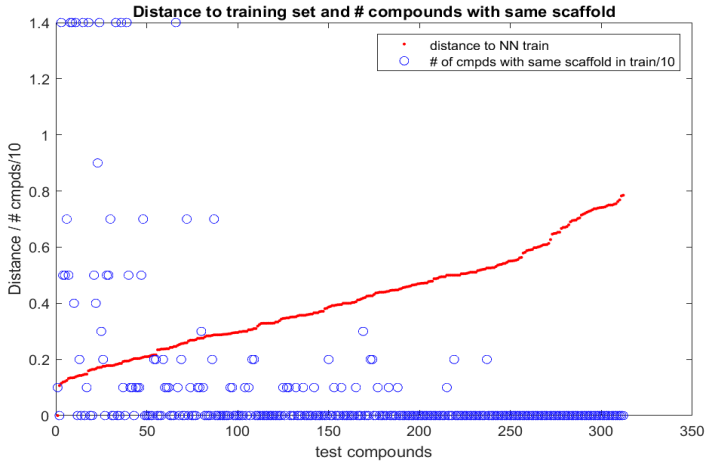
- **75% of the test set have unique scaffolds**



- Scaffold frequencies:
  - # scaffolds with 1 compound = 1389
  - # scaffolds with 2 compounds = 258

Distance to nearest neighbour in training set

- Soergel distance in the ECFP4 space
  - the complement of the Tanimoto coefficient
- **35% of test compounds have distance to training < 0.3**
  - 16% compounds have distance < 0.2



Final model performance

- **ChemProp multi-task model** was selected as the final model (DNN with graph convolutions)
- Model selection factors:
  - Performance
  - Descriptors
  - Confidence estimation
  - Interpretability
  - Ease of implementation
  - Speed of prediction
- The model is being rolled-out in AZ
  - Integration into Augmented Drug Discovery tools
  - Access via chemistry applications

Performance on the test set for most important endpoints

	CL	F%	V_{ss}
R^2	0.57	0.48	0.50
RMSE	0.28	0.72	0.28
% cmpds with error < 2-fold	75%	38%	72%
% cmpds with error < 3-fold	90%	59%	92%
% cmpds with error < 5-fold	97%	76%	98%
Experimental variability	0.18	0.55	0.21

The performance measures are shown for log-transformed values for CL, V_{ss}, for logit-transformed values for %F

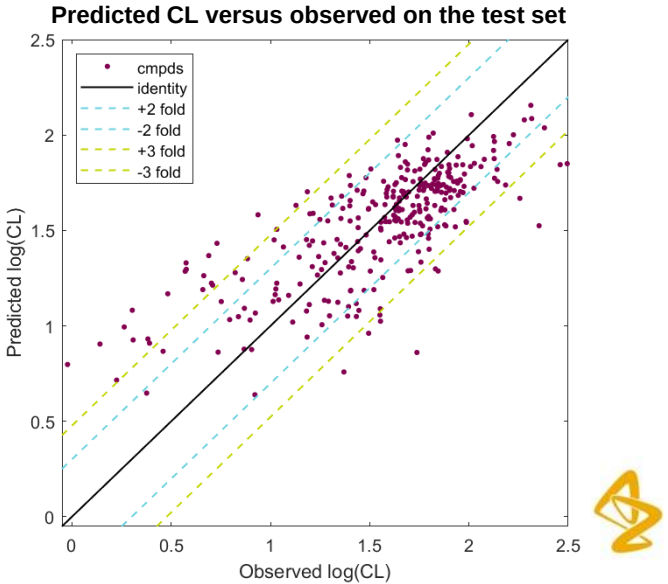
Clearance model performance

- RMSE of CL (clearance) model is close to experimental variability
- 75% of compounds have error within 2-fold (< 0.3 log units)
- In addition, good performance in classifying compounds

Predicted

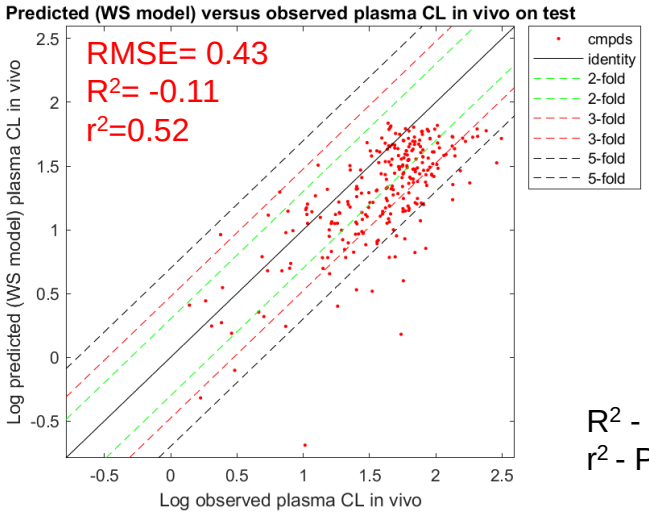
Observed	≤50	>50	Accuracy
≤50 ml/min/kg	153	25	86%
>50 ml/min/kg	48	86	64%
Precision	76%	77%	**77%**

R^2	0.57
RMSE	0.28
% cmpds with error < 2-fold	75%
% cmpds with error < 3-fold	90%
Experimental variability	0.18

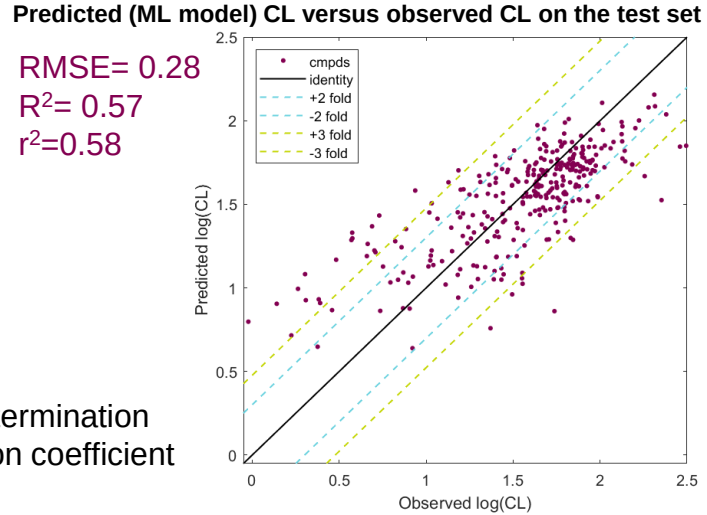


Performance of ML clearance model versus well-stirred model

- Well-stirred model is used to predict *in vivo* PK from *in vitro* measurements and to understand mechanism of clearance
  - Prediction of hepatic metabolic clearance



- ML model predicts total *in vivo* clearance
  - It can predict CL higher than liver blood flow (WSM is restricted by liver blood flow)
- Higher accuracy
- Additional tool alongside WSM



R^2 - Coefficient of determination
r^2 - Pearson correlation coefficient

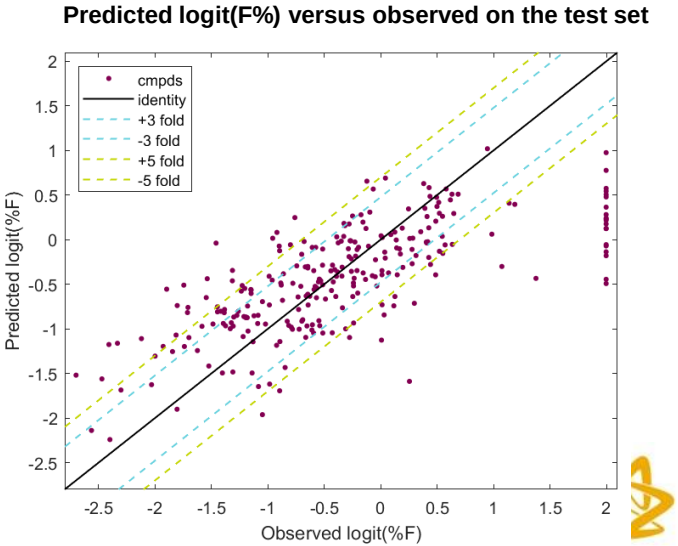
Bioavailability model performance

- Bioavailability model has an acceptable accuracy
- Very good performance in classifying compounds

Predicted

Observed	≤30%	>30%	Accuracy
≤30%	94	30	76%
>30%	18	124	87%
Precision	84%	81%	**82%**

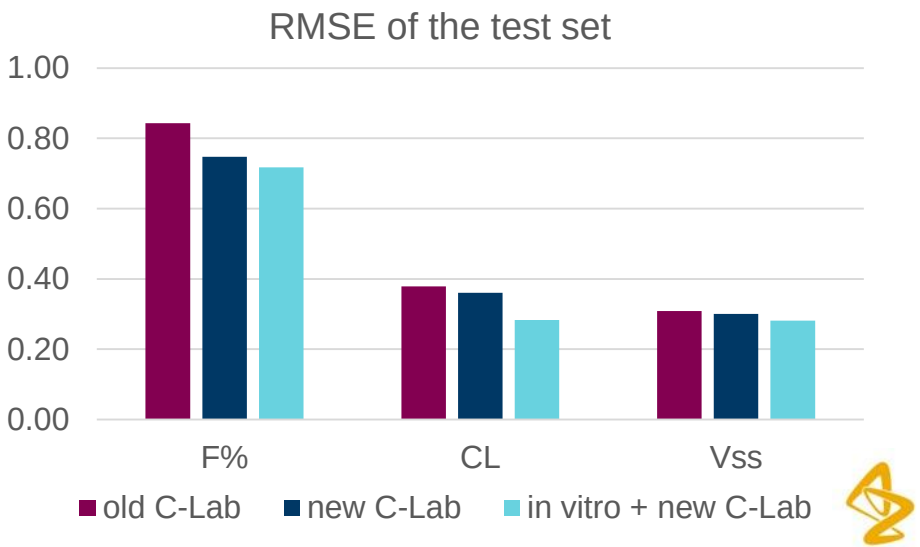
R^2	0.48
RMSE	0.72
% cmpds with error < 2-fold	38%
% cmpds with error < 3-fold	59%
% cmpds with error < 5-fold	76%
Experimental variability	0.55



Validation of rat PK model without *in vitro* features

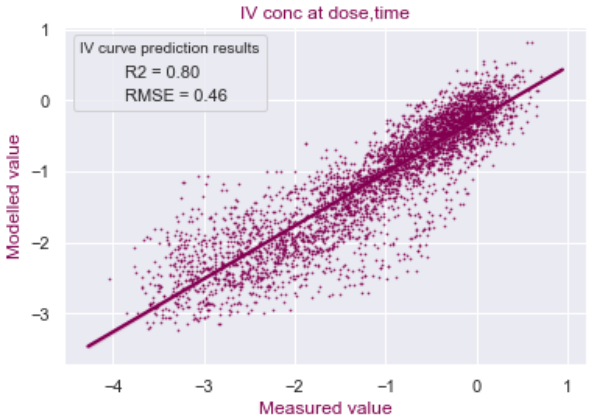
- **Can we predict virtual compounds purely from chemical structure?**
  - Use older C-Lab models for *in vitro* properties
  - New C-Lab models for *in vitro* properties may contain the PK test set compounds in their training data

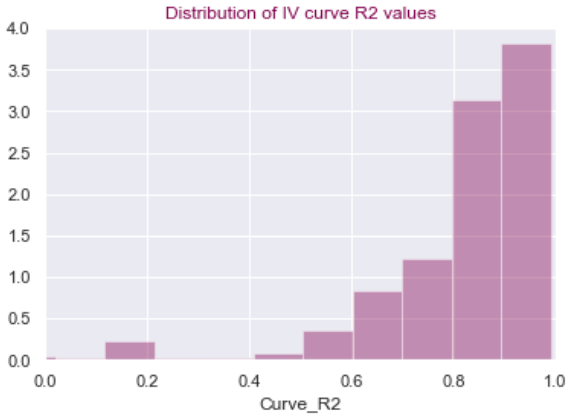
- Accuracy is reduced if not using in vitro properties (new C-Lab bars)
- Accuracy is further reduced if using older C-Lab models (old C-Lab bars)
  - Increase in RMSE by 0.1 log unit for F% and CL
  - Only marginal increase for Vss
- **The models are useful at the point of design**



Modelling PK curves

- **Predict concentration as a function of dose/time**
  - Single task – dose and time become descriptors, Y is concentration
  - Multi task – multiple properties are concentrations at different times
  - Customised CV partition – all points for a compound assigned to one group
- **Example**: Gradient Boosting Regression with 2D molecular property descriptors and *in vitro* ADME properties
- **Good accuracy achieved in IV curves prediction**
  - Overall accuracy $R^2 = 0.8$, RMSE = 0.46 (log units)
  - 63% of IV curves have $R^2 > 0.8$
  - 87% of IV curves have $R^2 > 0.6$

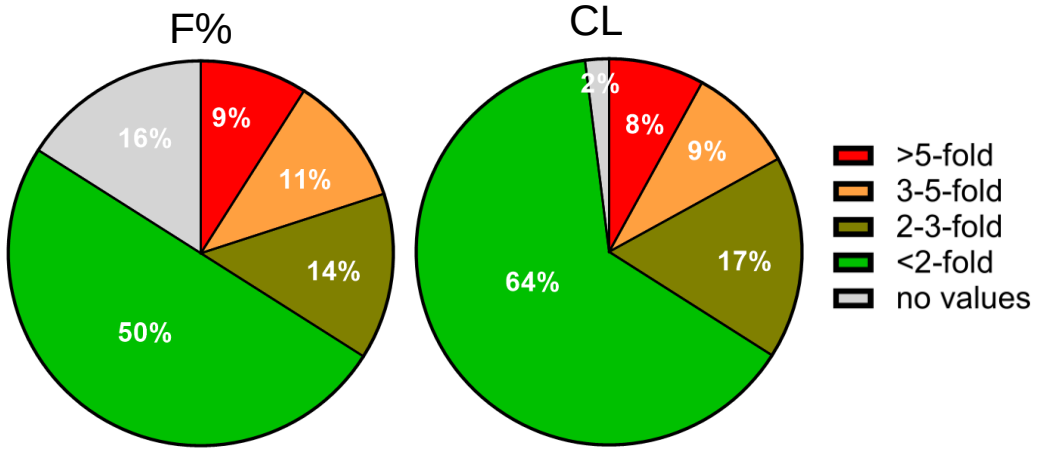




Prospective validation of model performance

- ~300 compounds measured post model building and dataset extraction
- Predicted bioavailability (F%) and clearance (CL) versus measured

Percentage of compounds with 2-3-5-fold error



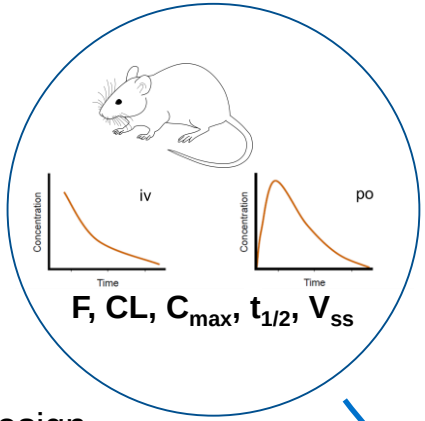
The model performance is consistent between prospective and test set validation

- F% predicted within 2-fold for 50% of compounds (64% within 3-fold)
- CL predicted within 2-fold for 64% of compounds (81% within 3-fold)

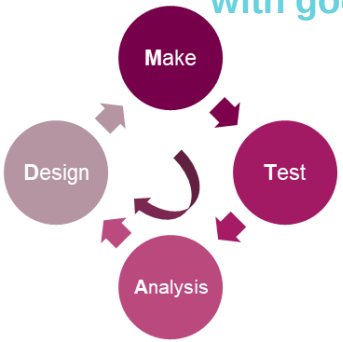
Summary

Successful prediction of *in vivo* rat PK parameters from chemical structure & measured and/or predicted *in vitro* ADME properties

- Design of molecules with a desired multi-objective profile early in drug discovery
- Drive prioritisation of compounds for *in vivo* assays
- Extending knowledge to modelling human PK
- Earlier prediction of DMPK and safety
- Enable human PK prediction at point of design
- Prediction of drug therapeutic index

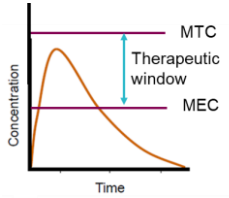


Design compounds with good PK



Enable human PK prediction at point of design

Human PK data



Acknowledgements

AstraZeneca

Data Science & AI, Clinical Pharmacology & Safety Sciences

- Andreas Bender
- Piotr Grabowski
- Nigel Greene
- Anton Martinsson
- Filip Miljkovic
- Olga Obrezanova
- Ioana Oprisiu
- Graham Smith

DMPK Oncology

- Beth Williamson

DMPK Cardiovascular, Renal and Metabolism

- Susanne Winiwarter

Intellegens

- Tom Whitehead

Optibrium

- Matt Segall
- Ben Irwin
- Samar Mahmoud

Confidentiality Notice

This file is private and may contain confidential and proprietary information. If you have received this file in error, please notify us and remove it from your system and note that you must not copy, distribute or take any action in reliance on it. Any unauthorized use or disclosure of the contents of this file is not permitted and may be unlawful. AstraZeneca PLC, 1 Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge, CB2 0AA, UK, T: +44(0)203 749 5000, www.astrazeneca.com