Integrating heterogeneous assay data for ML-based ADME prediction

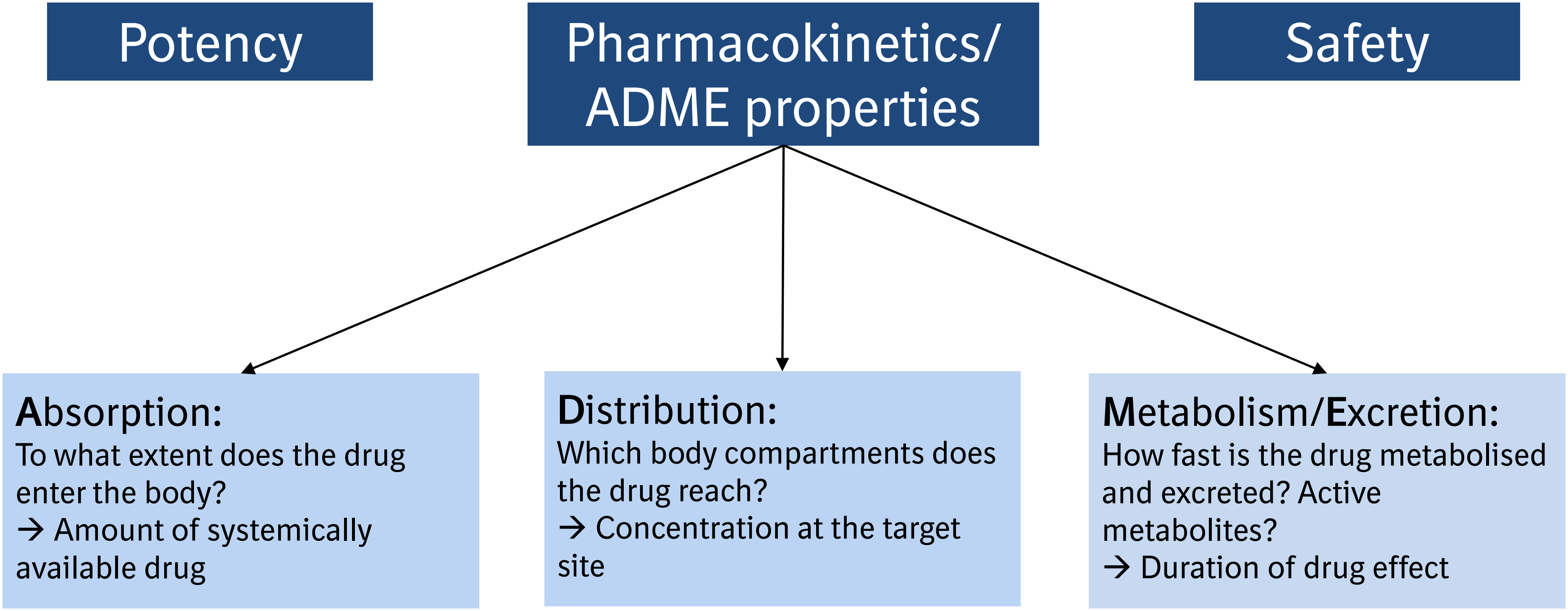
Moritz Walter

9th Joint Sheffield Conference on Chemoinformatics

20/06/2023

Boehringer Ingelheim

Drug discovery as a multi-parameter optimisation problem



Potency

Pharmacokinetics/ ADME properties

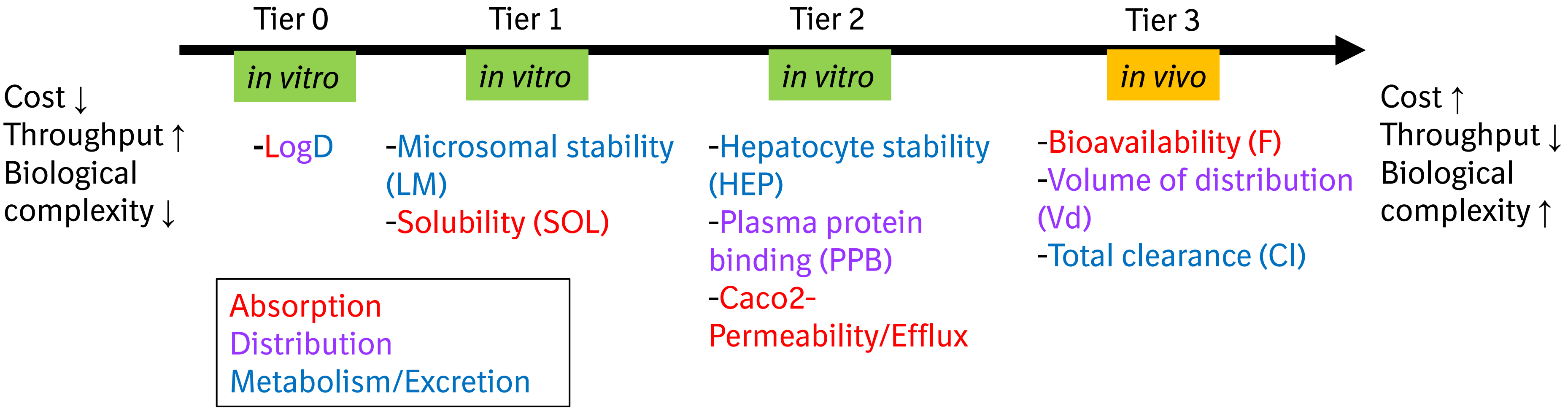
Safety

Absorption:
To what extent does the drug enter the body?
→ Amount of systemically available drug

Distribution:
Which body compartments does the drug reach?
→ Concentration at the target site

Metabolism/Excretion:
How fast is the drug metabolised and excreted? Active metabolites?
→ Duration of drug effect

Measurement of PK/ADME properties



Cost ↓
Throughput ↑
Biological complexity ↓

Tier 0 — *in vitro*

-LogD

Tier 1 — *in vitro*

-Microsomal stability (LM)
-Solubility (SOL)

Tier 2 — *in vitro*

-Hepatocyte stability (HEP)
-Plasma protein binding (PPB)
-Caco2-Permeability/Efflux

Tier 3 — *in vivo*

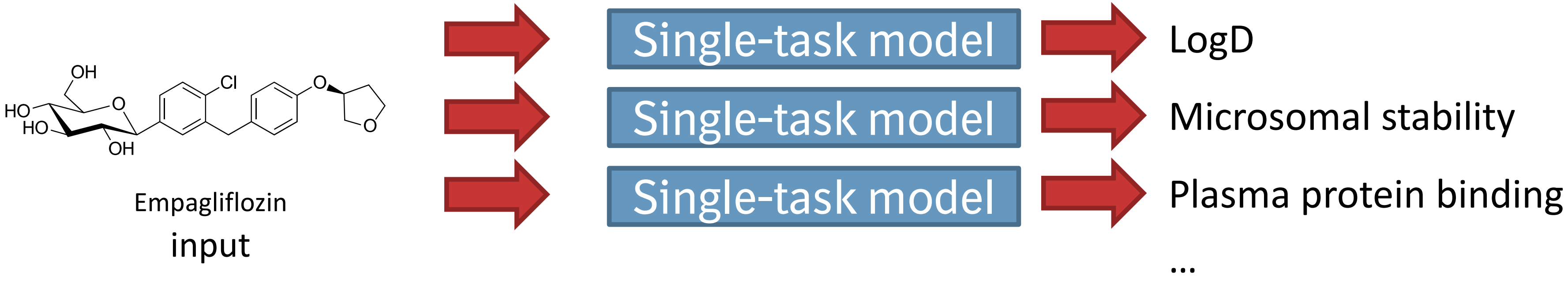
-Bioavailability (F)
-Volume of distribution (Vd)
-Total clearance (Cl)

Cost ↑
Throughput ↓
Biological complexity ↑

Absorption
Distribution
Metabolism/Excretion

- Cheap/high-throughput assays required to test large numbers of compounds
- Only measure promising candidates in complex assays
- Can we use ML predictions to prioritise compounds for synthesis/to replace experiments?

Multi-task modelling



Empagliflozin

input

Single-task model

Single-task model

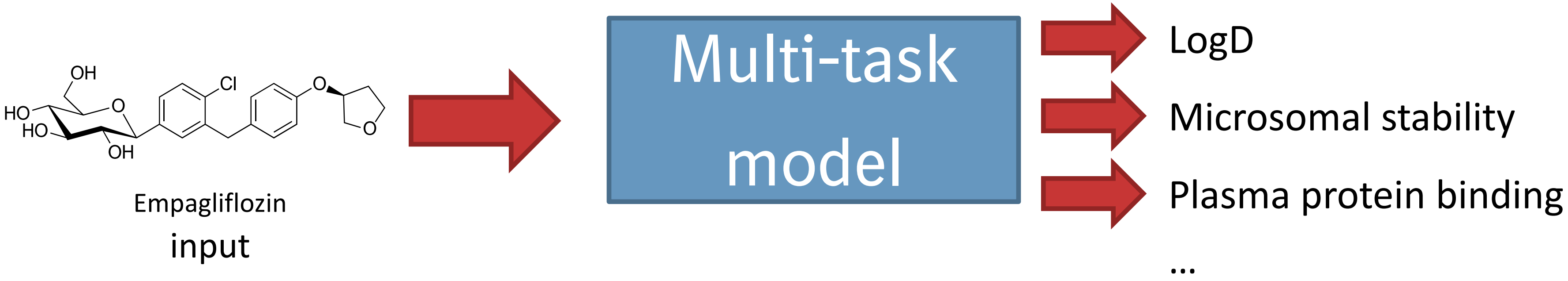
Single-task model

LogD

Microsomal stability

Plasma protein binding

…



Empagliflozin

input

Multi-task model

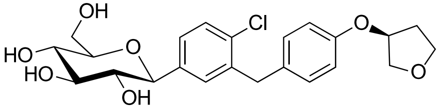
LogD

Microsomal stability

Plasma protein binding

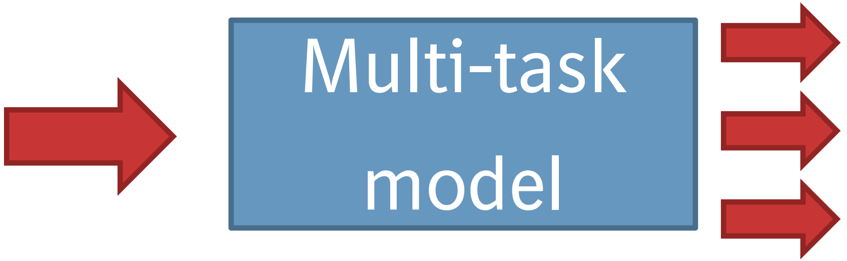
…

Multi-task modelling



Empagliflozin

input



Multi-task model

LogD

Microsomal stability

Plasma protein binding

…

- Motivation: make best use of available data
  – Assays are related
  – Data-poor assays might benefit from signal in data-rich assays
- Implementation: Chemprop[1] (graph-convolutional neural network)
  – Input features: chemical graph with basic information about atoms and bonds
  – Ensembling (we used n=5)

[1] Yang et al 2019, https://doi.org/10.1021/acs.jcim.9b00237

Study design

Goals:

- MT models superior to ST approaches (Chemprop and Random Forest) for data at hand?
- When predicting higher tier assays for a compound: additional benefit of including available experimental data of lower tiers in training?

Data:

- 28 assays arranged in 4 tiers
- Data preparation: curation, filtering, transformations

Model evaluation:

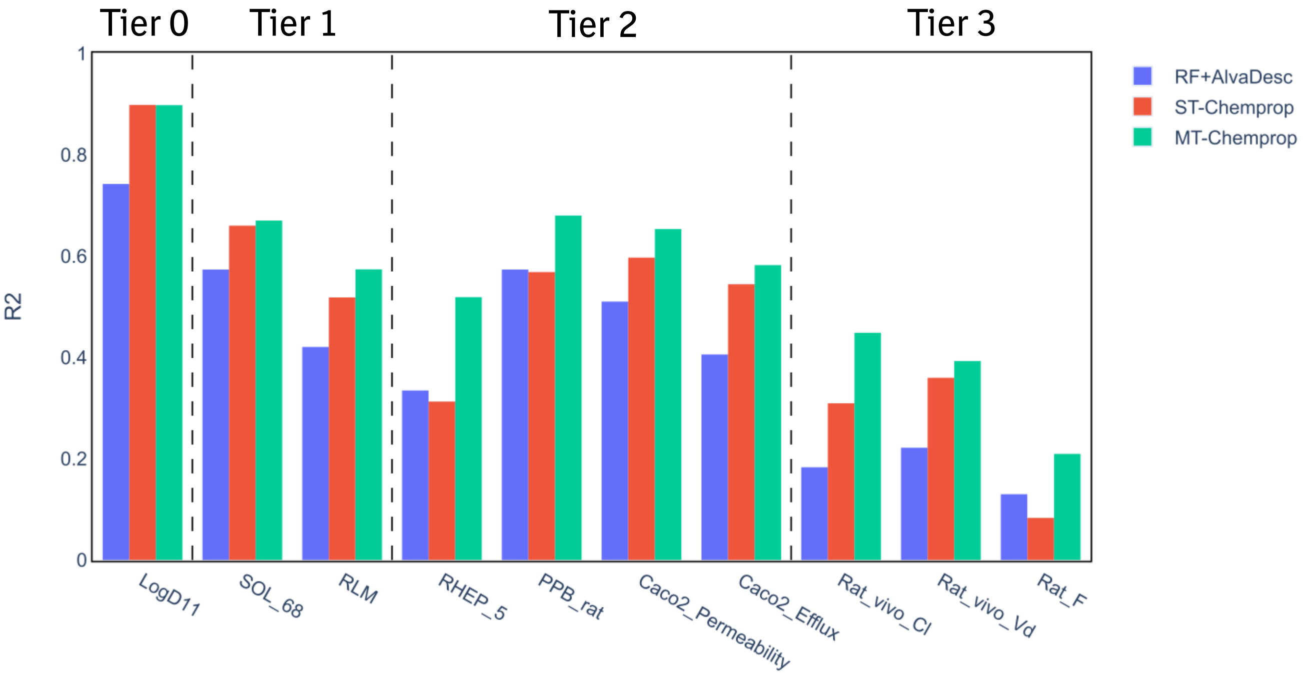
- Temporal data splits (train on up to 31/12/2020, evaluate on 2021)
- R2 (Coefficient of determination) as primary metric

Assay datasets

Tier 0 (n=2)	Tier 1 (n=6)	Tier 2 (n=14)	Tier 3 (n=6)
Training set sizes ~120k Assays • logD at pH 2 and 11	Training set sizes 50k – 125k Assays • SOL (different pH) • LM stability (different species)	Training set sizes 1k – 17k Assays • HEP stability (different species/serum concentrations) • PPB (different species) • Caco2-Permeability/Efflux	Training set sizes 2k – 10k Assays • In vivo PK (Cl, Vd, F) in rat and mouse

Model scores are reported for subset of assays that reflect overall trends

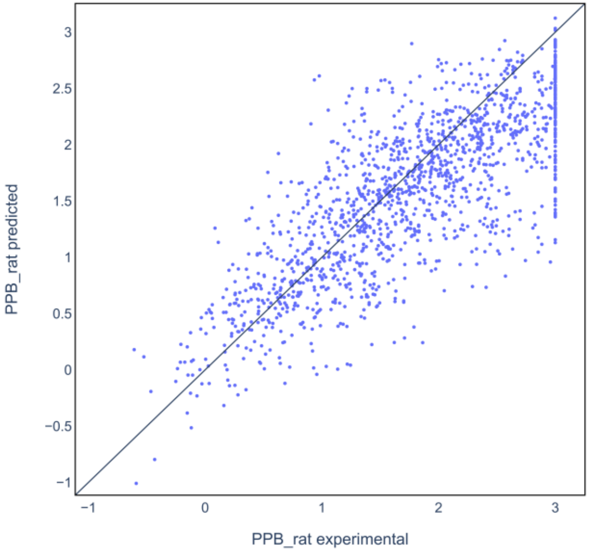
Model evaluation



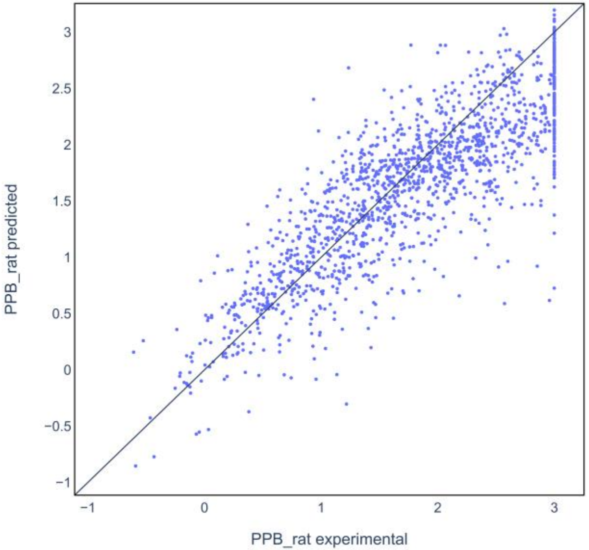
→ MT-Chemprop outperforms ST approaches

PPB_rat: predicted vs experimental

ST-Chemprop (R2 = 0.569)



MT-Chemprop (R2 = 0.680)



PPB [%]	Logit transformed*
50	0
90	0.95
99	2.00
99.9	3.00

$$^{*}y = \log_{10}\frac{x}{1-x}$$

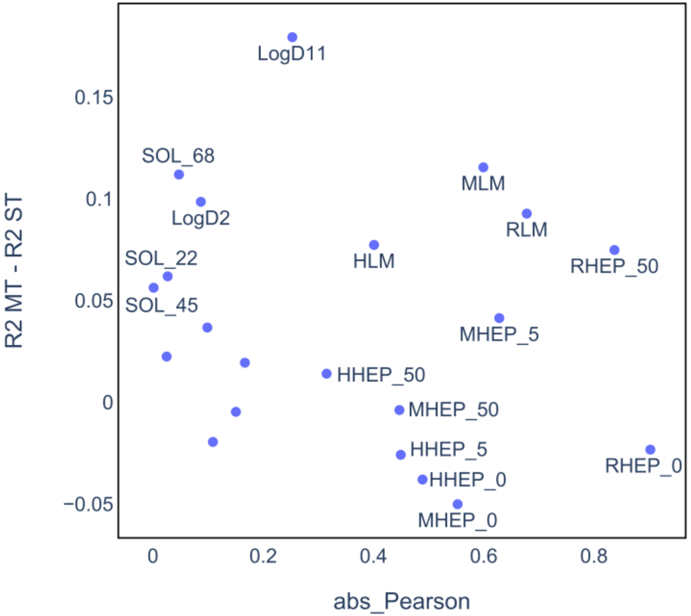
Understanding the success of MT models

- Examplary target assay: RHEP_5
- Which auxiliary assays are the most useful?
  - Train and evaluate pairwise MT-Chemprop models
  - Can we discover factors that determine the success?

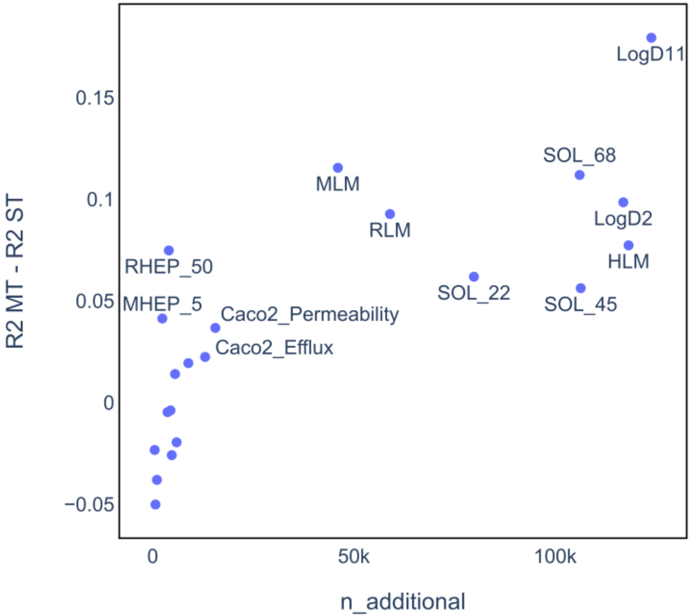
R2 ST-Chemprop	R2 MT-Chemprop
0.313	0.519

Understanding the success of MT models (RHEP_5 example)

x-axis: absolute Pearson correlation coefficient of overlapping training compounds



x-axis: training compounds in auxiliary assay not included in target assay



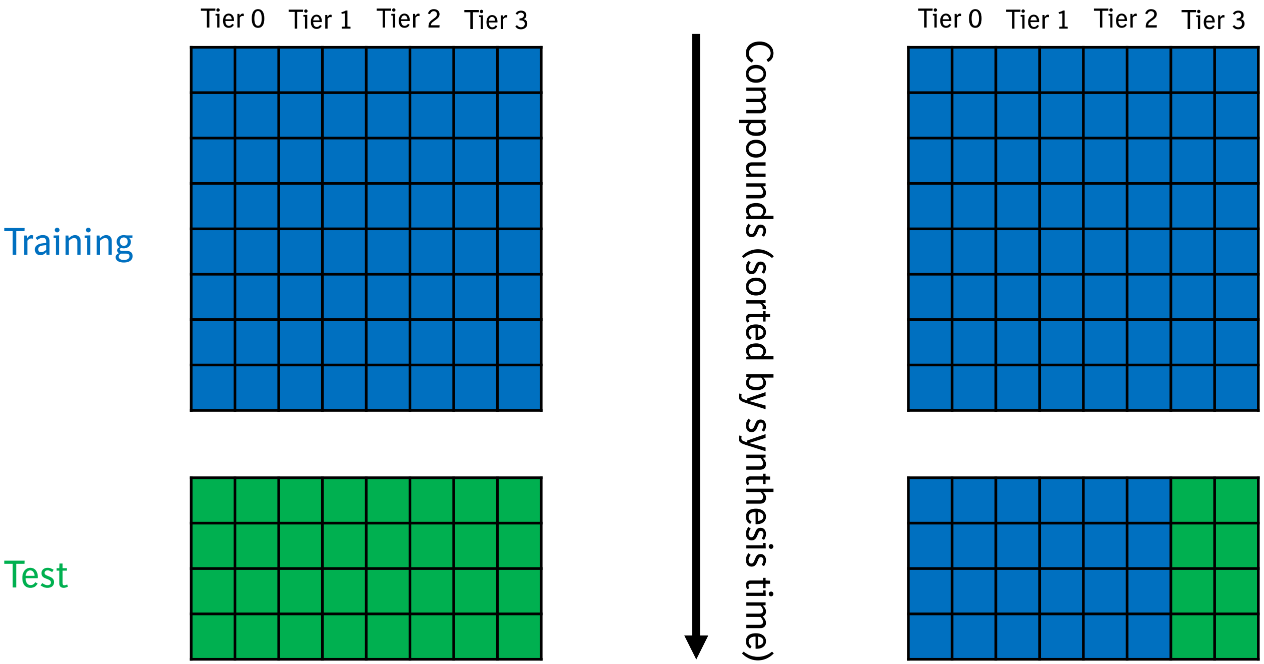
→ Size of auxiliary dataset more relevant than correlation for success of MT model

Boehringer Ingelheim

MT model at design stage vs testing stage

MT model
-> design stage of compound

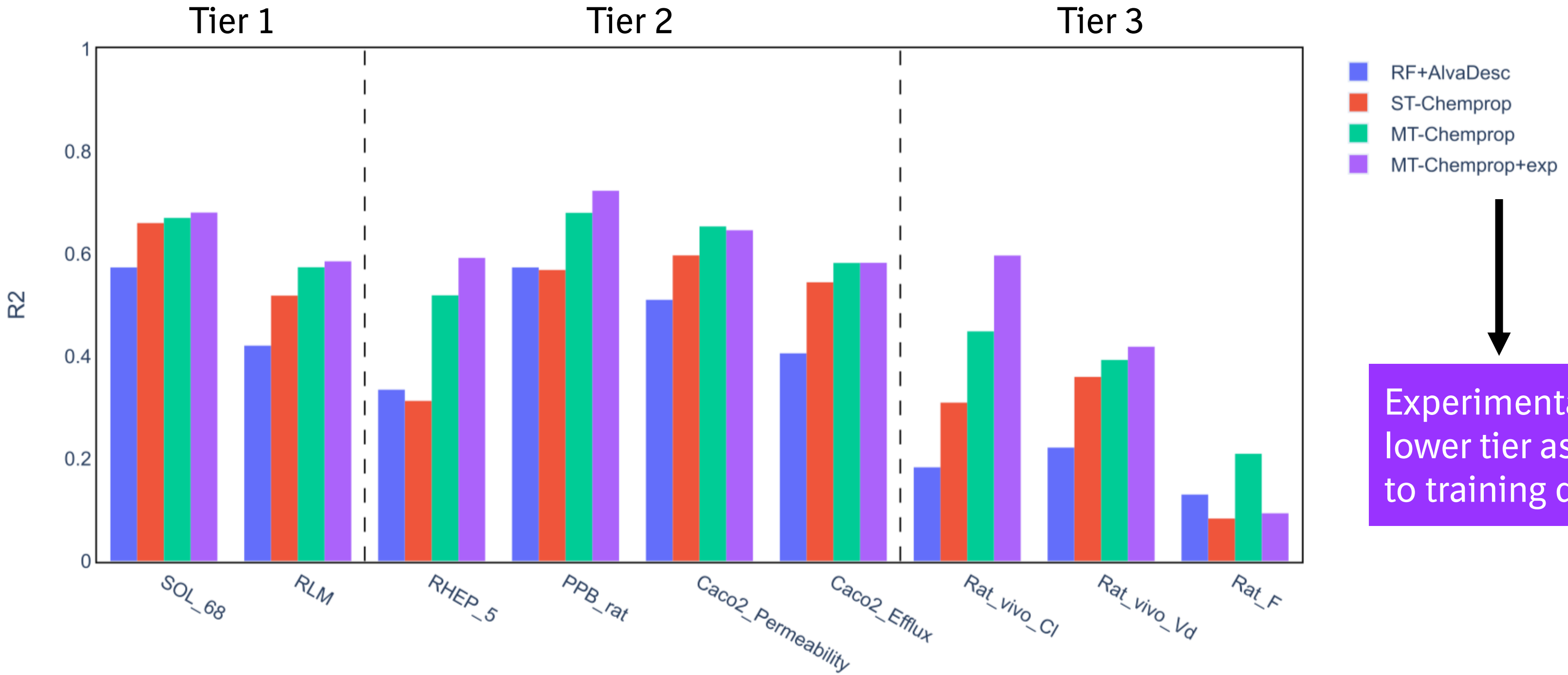
MT model+exp
-> testing stage of compound



“Test compounds“ now both in training and test set

e.g. data for Tier 0 to Tier 2 available for training when predicting Tier 3

Model evaluation



Experimental data of lower tier assays added to training data

→ MT-Chemprop further improved with experimental data of lower tiers available

Conclusion

- MT-Chemprop clearly outperforms ST models on the studied ADME/PK datasets at design stage
- Data-rich assays seem to be the most useful auxiliary assays in the MT-model (despite low correlation to target assay)
- Further improvements possible at testing stage when experimental data (earlier assays) of test compounds is added to the training data

Acknowledgements

CompChem at BI Biberach

Dr. Lina Humbeck
Dr. Miha Skalic