The development of ANNs for the prediction of influential individuals and outlying individuals and their application during the model building process

Osama Qutishat[1], Simon J. Carter[1], Rikard Nordgren[1], Alzahra Hamdan[1], Shijun Wang[1], Tianwu Yang[1], Xiaomei Chen[1], Simon Buatois[2], João A. Abrantes[2], Andrew C. Hooker[1] and Mats O. Karlsson[1]

[1.] Department of Pharmacy, Uppsala University, Sweden [2.] Roche Pharma Research and Early Development, Roche Innovation Center Basel, Basel, Switzerland

Introduction

- Outlying individuals (OutIDs) add to OFV but lie outside model prediction distribution (assessed via simeval in PsN [1])
- Influential individuals (InfIDs) can drive population parameter estimates (assessed via Case Deletion Diagnostics (CDD) in PsN [1])
- Both PsN methods are time & computer intensive
- Artificial Neural Networks (ANNs)
  - ✚ useful for complex inter-relationships
  - – need lots of (rich/balanced) data
- ANN = Quicker prediction than CDD and simeval
  - Use multiple layers (connected via nodes)
- **Aims: 1)** Develop ANNs to predict InfIDs and OutIDs **2)** Deploy into Pharmpy [2] **3)** test on 8 drugs during automatic model development (AMD) models at 4 stages

Methods

Database creation

- 13 predictors (see fig.1) from 27 NLME models
- True positives (TP): OutID: $Res_{iOFV} > 3$
  InfID: dOFV > 3.84

Pre-processing

- Normalise predictors from -1 to 1
- Split data (90:10 train:test) using sklearn
- 10-fold cross-validation as sparse data
  - TP distribution must be same across splits
    - Done with modified stratified KFold split

ANN

- Input (predictors), hidden (interactions) & output layers (prediction), learning rate = 7E-5
- Developed in tensorflow (2.6) in Python (3.9.5)
- Permutation importance:
  - Gives importance of predictor to output
- Final ANNs based on: **sensitivity** (TP rate), **specificity** (true negative (TN) rate) and **precision** (true / total positives).

Deployment

- Final ANN models converted into tflite
- Tflite size (~1 MB) << tensorflow (~ 500 MB)
- Included in Pharmpy [2] (0.70.0 or later) for prediction of InfIDs and OutIDs

Testing

- 8 NLME model (2 oral, 6 i.v.) stages:
  - Starting model (1 CMT, 1st order absorption and elimination), end of structural [3], IIV and RUV AMD workflows

Results

Table 1: Final OutID ANN results

OutIDs	TP	TN
Predicted Positive	79	16
Predicted negative	21	1691

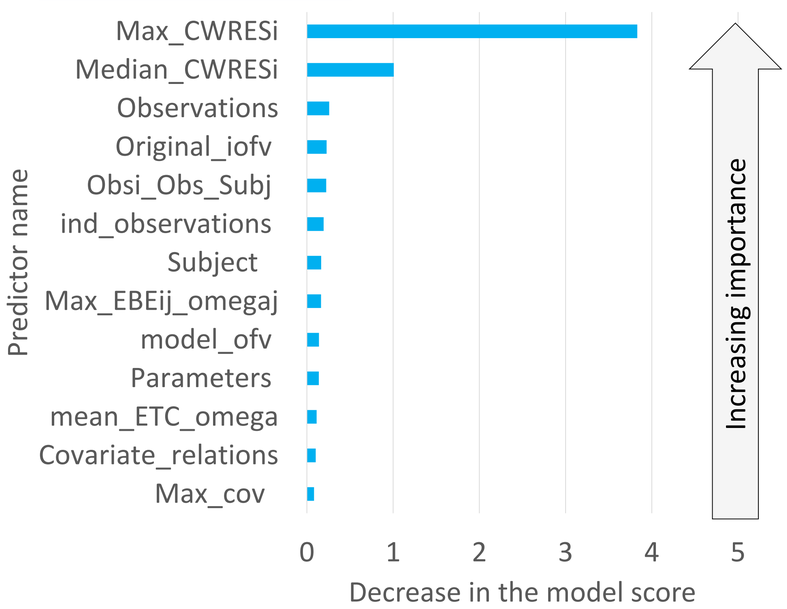

Figure 1: Permutation importance of the OutID ANN predictors*

Table 2: Final InfID ANN results

InfIDs	TP	TN
Predicted Positive	15	9
Predicted negative	11	2538

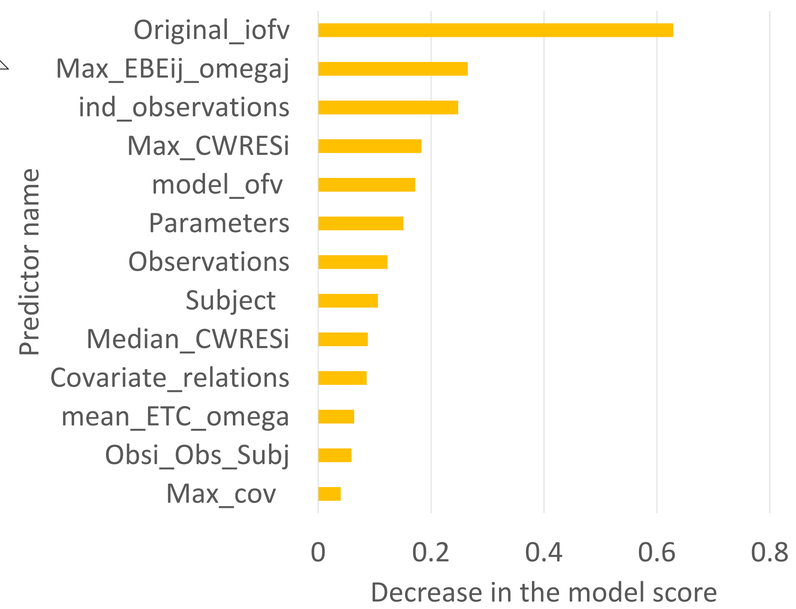

Figure 2: Permutation importance of the InfID ANN predictors*

- True values: 6% OutID (100/1791)
  1% InfID (26/2573)
- Final ANN models:
  - OutID: 4 hidden (128, 96, 64, 32 nodes)
  - InfID: 4 hidden (128, 96, 64, 24 nodes)
- OutID: sensitivity: 79%, specificity: 99.1%, precision: 83.2%
- InfID: sensitivity: 57.8%, specificity: 99.6%, precision: 62.5%
- CWRESi and iOFV important for prediction of OutIDs and InfIDs (fig. 1 & 2)

Table 3: Prediction of number of OutIDs and InfIDs in Pharmpy for starting and final AMD stage models

Drug	Starting		Structural		IIV		RUV	
Daunorubicin	1	0	1	0	3	0	1	0
Desmopressin	2	0	3	0	1	0	1	0
Factor VIII	2	0	1	0	0	0	0	0
Gentamicin	16	0	12	0	17	0	16	0
Lopinavir	0	6	0	6	1	5	0	7
Pefloxacin	6	0	5	0	4	0	0	0
Tobramycin	8	0	9	0	6	0	6	0
Warfarin	3	0	5	0	0	0	0	0

Conclusions

- ANNs successfully developed for both OutIDs and InfIDs
- Lower precision and sensitivity for InfIDs maybe due to too few TP than for OutIDs
- ANNs deployed successfully into Pharmpy via tflite.
  - NB: ANN trained on published models, not on AMD
- Predicted OutIDs from Pharmpy ↓ or ↔ during AMD model development (Table 3)
- Predicted InfIDs from Pharmpy only seen with lopinavir (Table 3)
  - Re-tested using CDD = false positives
- More work required to improve use of ANNs for OutID and InfID
  - Other predictors from models and datasets, more “true” values (TP and TN)

*Max and Median_CWRESi = maximum and median CWRES for individual *i*, Max_EBEij_omega = maximum EBE *i* / variance for parameter, mean_ETC_omega = mean of standard errors of the EBE / variance for parameter. Counts: Observations = total observations in dataset, ind_observations = observations per subject, Parameters in model, Subjects, Covariate_relations

[1] Carter SJ *et al.* (in draft 2022): Assessment of Influential Individuals and Outliers in Pharmacometric Models Through the Use of Case Deletion Diagnostics (CDD) and Simulation Evaluations (simeval) Tools.
[2] Nordgren R, et al. (2021) Pharmpy and assemblerr-Two novel tools to simplify the model building process in NONMEM (https://www.page-meeting.org/default.asp?abstract=9656).
[3] Hamdan A *et al. (2022)* Automatic development of pharmacokinetic structural models. (https://www.page-meeting.org/default.asp?abstract=10020)

Acknowledgements

This work was supported by F. Hoffmann-La Roche Ltd., Basel, Switzerland and Bayer Pharmaceuticals, Berlin, Germany

In collaboration with: