

薬物血中濃度を予測する ANN-PK モデルのデータ拡張に関する検討

A Study on Data Augmentation of Artificial Neural Networks-Pharmacokinetic for Predicting Drug Concentration

○川村駿太¹, 関弘翔², 宮野咲紀³, 森田直人⁴, 辻泰弘⁵, 細野裕行²

*Shunta Kawamura¹, Hiroto Seki², Saki Miyano⁴, Naoto Morita⁴, Yasuhiro Tsuji⁵, Hiroyuki Hosono²

Abstract: This study examines data augmentation patterns for input and supervised attributes in order to improve the ANN-PK model for predicting drug blood concentration.

1. まえがき

経時的な薬物血中濃度の推移を表す薬物動態 (PK) 解析は、投与量・投与方法の設計や薬効・副作用の予測において重要である。先行研究^[1]では、人工知能を患者個人々の薬物の効果予測に応用できないかを考え、薬学的数理モデルである PK モデル (コンパートメントモデル) と Artificial Neural Networks (ANN) を組み合わせた ANN-PK モデルを開発した。その結果、従来の母集団薬物動態 (PopPK) モデル^[2]よりも血中濃度の予測精度が向上することが報告されている。

ANN-PK モデルでは、取得した患者データを ANN に入力し、体内から薬物を代謝する能力の指標であるクリアランス (CL) 予測データを出力する。出力した CL 予測データ、投与量データ、投与時間データを PK モデルに入力し算出する血中濃度予測データを、患者から採取された実測の薬物血中濃度と比較し、誤差逆伝播学習で最適化する。一般に ANN の学習には大量のデータが必要とされるが、薬物血中濃度を含む臨床データは 100 程度のデータであることが多く、モデルの最適化が不十分となる可能性が考えられる。

本研究では、学習データ不足を補う方法としてデータ拡張に着目し、ANN-PK の学習に有効であるかを検証する。

2. 解析結果

免疫抑制剤シクロスポリン (36 名, 89 データ) を用いて実験を行う。まず入力属性である年齢, 体重, および臨床検査値の ALT, AST, 血清クレアチニンに対して標準偏差の 2 割の範囲で確率的に加算・減算させるデータ拡張を検討した。さらに、それらのデータを用いて、PopPK モデルの式^[2]に当てはめることで、教師属性である血中濃度のデータ拡張を検討した。同じハイパーパラメータで ANN-PK を leave-one-out 法により学習し、データ拡張の有無を含むいくつかのパターンで結果を比較した。評価には二乗平均平方根誤差

(Root Mean Square Error, RMSE) を用いた。Table 1 に様々なデータ拡張パターンにおけるそれぞれの RMSE を示す。Table1 より、10 割で入力属性のみデータ拡張した場合が最も高精度であるとわかる。また、Fig. 1 に示す縦軸を血中濃度の実測値, 横軸を予測値とした散布図相関からも、10 割で入力属性のみを拡張することで予測が実測とより一致する様子が確認できる。教師属性の拡張時における精度低下は、PopPK モデルの精度 RMSE41.1[ng/mL]に起因すると考えられる。

Table 1. RMSE with various data augmentation patterns

Patterns of data augmentation	RMSE[ng/mL]
Without data augmentation	32.28
Only input attributes in 10%	33.53
Both input and supervised attributes in 10%	40.80
Only input attributes in 50%	34.65
Both input and supervised attributes in 50%	41.00
Only input attributes in 100%	30.27
Both input and supervised attributes in 100%	41.02

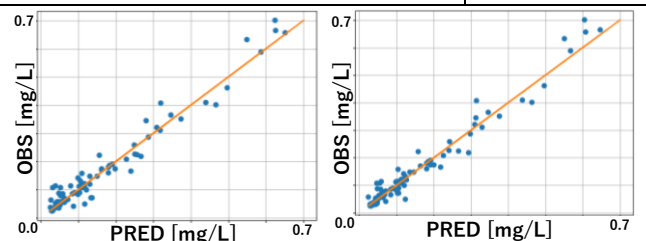


Figure 1. Scatter diagram correlation: the left figure shows without data augmentation, while the right figure shows data augmentation only for input attributes at 10%

3. まとめ

本研究では、薬物血中濃度を予測する ANN-PK モデルに対して、データ拡張の有効性を検証した。結果より、入力属性のデータ拡張が精度向上に寄与することが明らかとなった。

参考文献

[1] C. Ogami, Y. Tsuji, H. Seki, CPT: Pharmacometrics & Systems Pharmacology 2019;10(7):760-768
 [2] Y. Tsuji, N. Iwanaga, A. Mizoguchi, et al., Biol Pharm Bull. 2015;38(9):1265-71

1: 日大理工・学部・情報 2: 日大理工・教員・情報 3: 日大理工・院(後)・情報 4: 日大理工・院(前)・情報 5: 日大薬・教員・薬学