

臨床予測モデルを構築する研究のサンプルサイズ計算： 健康効用値へのマッピングへの応用

これまで多くの臨床予測モデル (clinical prediction model) が構築され、臨床現場で診断、治療選択、患者管理などに用いられてきた。臨床予測モデルをデータから構築する際には、そのデータが十分なサンプルサイズを有していることが望ましい。回帰モデルで臨床予測モデルを構築する際に十分なサンプルサイズを持つか判断する方法に、予測変数 1 つ当たりの対象者数やイベント数を参照する方法がある。しかし、この方法は基準が単純化され過ぎており、予測変数 1 つ当たりの対象者数やイベント数が一定以上あっても良い予測性能が得られない場合がある。

近年、Riley らの研究グループは、回帰モデルを用いて臨床予測モデルを構築する研究のサンプルサイズ計算法を提案した。このサンプルサイズ計算法は、大きな過適合 (overfitting) が生じず、重要なモデルパラメータを精度良く推定できるサンプルサイズを計算するというものである。本抄読会では、主に連続アウトカムを線型回帰モデルで予測する研究のサンプルサイズ計算法を紹介する。Riley らが提案する方法は、(1) 包括縮小係数 (global shrinkage factor) が 0.9 以上、(2) 見かけの決定係数と調整済み決定係数の差が 0.05 以下、(3) 残差の標準偏差の 95%信頼上下限の multiplicative margin of error (MMOE) が 1.1 以下、(4) 切片パラメータの 95%信頼上下限の MMOE が一定 (たとえば 1.1) 以下という 4 基準を満たすことが期待される最小のサンプルサイズを計算するというものである。

Riley らのサンプルサイズ計算法を、発表者がこれまで取り組んできた健康関連 quality of life (QOL) 尺度から健康効用値を予測するモデルの開発に適用した。先行研究をレビューし、サンプルサイズ計算に必要な情報を抽出するとともに、先行研究が十分なサンプルサイズを備えているか評価したところ、一部の研究は最小サンプルサイズを下回っていた。

文献

1. Riley RD, Snell KIE, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: Part I - continuous outcomes. *Stat Med* 2019;38:1262–1275.
2. Riley RD, Snell KIE, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: Part II - binary and time-to-event outcomes. *Stat Med* 2019;38:1276–1296.
3. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020;368:m441.
4. Wailoo AJ, Hernandez-Alava M, Manca A, et al. Mapping to estimate health-state utility from non-preference-based outcome measures: an ISPOR good practices for outcomes research task force report. *Value Health* 2017;20:18–27.
5. Hagiwara Y. Sample size consideration for developing a mapping algorithm using linear regression. Submitted.