

## 予測モデルの性能評価指標について

予測モデルとは、既知のデータを用いて新しいデータや別のデータの結果の予測を行うための統計的モデルである。医療において患者個人や集団の疾病の発生、治療後の予後などのイベントを予測することは、治療法の選択などの意思決定や患者のインフォームド・コンセントの取得などにおいて有用であり、予測モデルが用いられている。一般に予測モデルでは患者の臨床的特徴や人口統計的特徴などを用いて作成され、予測を行う。

作成した予測モデルを使用するために必要なことは、予測モデルの性能の評価を行うことである。性能の評価のためにはさまざまな方法や指標が開発されてきた。予測モデルの評価はの主な構成要素は予測力(calibration)と識別力(discrimination)の2つである。予測力とは実際のアウトカムと予測モデルによるアウトカムの予測の近さであり、Hosmer-Lemeshow 統計量やキャリブレーションプロットによって示される。識別力とは予測モデルがイベントを起こす対象と起こさない対象を区別する能力であり、C 統計量や ROC で示される。このような従来用いられてきた指標に加え、比較的最近 reclassification table や net reclassification improvement(NRI)、integrated discrimination improvement(IDI)などの識別力の新たな指標が提案された。また意思決定における予測モデルの有用性についての評価のために decision curve analysis なども提案され使用されている。

そこで本抄読会では予測モデルの性能評価の指標について、従来使用されてきた指標や比較的最近提案された指標を紹介し、それぞれの特徴についてまとめたい。

## 主要文献

Steyerberg EW et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology* 2010; **21**(1): 128 – 138.

Whittemore, AS. Evaluating Health Risk Models. *Stat Med.* 2010; **29**(23): 2438-2452.