

## 論文の内容の要旨及び論文審査の結果の要旨

本論文は、過剰治療・過小治療を減らし、医療従事者および患者の負担を減らすため、臨床予測モデルの較正指標であるIntegrated Calibration Index (ICI)に着目し、ICIを最適化するための臨床予測モデルの構築方法について取り組んでいる。特に、具体的な研究対象として、3つの課題に取り組んでいる。まず、第2章では2つの規模の異なる臨床データを用い、データセットの規模に応じたICIを最適にするモデル及び変数選択法の検討を行った。解析手法として臨床予測モデルで良く用いられる Logistic regressionモデル (LR)、Linear Probability モデル (LPM)およびA Classification And Regression Tree (CART)の3つを用いた。さらに6つの変数選択法を検討した。結果、臨床データ1ではLPMのLasso法、臨床データ2ではLPMにおけるBackward法(P値基準、AIC基準)、Forward法、Stepwise法がICIを最適にしており、データの規模で最適な変数選択法は異なるがいずれの臨床データにおいてもLPMの有用性を示唆した。

次に、第3章では2つの規模の異なるデータを用い、Event per variable (EPV)に応じたICIを最適にするモデル及び変数選択法の検討を行った。2値アウトカムに対して多く用いられるlogistic回帰を用い、4つの罰則付き回帰を比較した。結果、Adaptive double Lassoが臨床データ1(非常に低いEPV)でcalibrationとdiscriminationの観点から優れていた。臨床データ2(低いEPV)において他の方法に比べsparseでかつ同程度の予測性能を示した。この結果は、Adaptive double Lassoが高次元設定において臨床予測モデルを開発するための強力な変数選択方法であることを示唆した。

さらに、第4章ではICIの定義を変更することで実際の臨床診療における意思決定を反映した新しいcalibration指標を提案した。また提案した指標を最適化するために、決定木とLogisticモデルの組み合わせたモデルを構築し、提案した指標を最適にできるか、シミュレーションを用いて検討した。結果、決定木とLogisticモデルの組み合わせたモデルがいずれの症例数設定(n=500,1000,5000)、強いリスク因子の有無に関わらず、他の手法に比べ提案したcalibration指標を最適にしていた。臨床予測モデルで用いられている多くのcalibration指標は全確率区間のcalibrationの一致を測るものであり、臨床の現状を反映した指標は提案されておらず、これを反映させた提案指標の提案の意義は大きいと考えられる。

これらは、臨床的意義を反映した新たな較正指標の開発及び最適化として、極めて意欲的な試みである。新規性、応用可能性の観点から博士の学位に相応しい内容と判断した。また、筆頭著者有査読論文7報(英文5報、邦文2報)、共著有査読論文45報(英文43報、邦文2報)を公表しており、この点でも早期の学位取得が適切であると判断した。