Non-invasive Total Serum Bilirubin Estimation in Preterm Infants with Modified Mixed-Effect Random Forest

Meng Chen[1], Alain Beuchée[1], Fabrice Tudoret[1], Alfredo I. Hernández[1]
[1] Univ Rennes, CHU Rennes, LTSI-INSERM U-1099, F-35000 Rennes, France

Context: Hyperbilirubinemia affects 60% to 80% of newborns and can lead to irreversible neurotoxicity, especially for those born very premature. Total serum bilirubin (TSB) is measured by repetitive blood sampling. The proposal of a non-invasive TSB estimator in this context might help minimize blood exploitations. This work compares machine learning models for TSB estimation using heart rate variability (HRV) analysis.

Methods: Four-hour raw ECG signals centered on each TSB measurement were acquired. Data processing included RR series estimation and HRV feature extraction. A classical random forest (RF) regressor was used in the form of $y = f(X)$, where y is the TSB to estimate and X the features comprising 18 HRV parameters and 2 age indicators. A novel modified mixed-effect random forest (mMERF) was proposed here, formulated as $y = f(X) + b1_i(Z1_i) + b2_i[g(Z2_i)]$, where $f(X)$ is an RF, $b1_i$ and $b2_i$ are the linear coefficients of two random effects for patient i: $Z1$ (post-menstrual age) and $Z2$ (post-natal age), and $g(\cdot)$ is an exponential decay function characterizing TSB dynamics recently proposed by our team. We hypothesize that this mMERF model, explicitly introducing physiological insights, should provide better performance. Bland-Altman plots were used for comparing observed and estimated TSB.

Results: From 317 newborns, 1652 bilirubin-associated samples are derived. Compared to the baseline RF estimator, mMERF realized a higher Pearson's correlation of 0.84 (RF: 0.53) and improved the 95% limits of agreement by 51% (training) and 34% (testing), while attenuating the undesired proportional bias (linear regression's slope from 0.50 to 0.13 in the training data and from 0.78 to 0.28 in the testing data, see Fig.1).

Conclusion: The proposed mMERF model, incorporating personalized and knowledge-based random effects, achieved higher agreement and lower proportional bias in estimating TSB levels. Although this model requires patient-specific history for initialization, it shows clinical potential in intensive care units where longitudinal data are commonly seen.

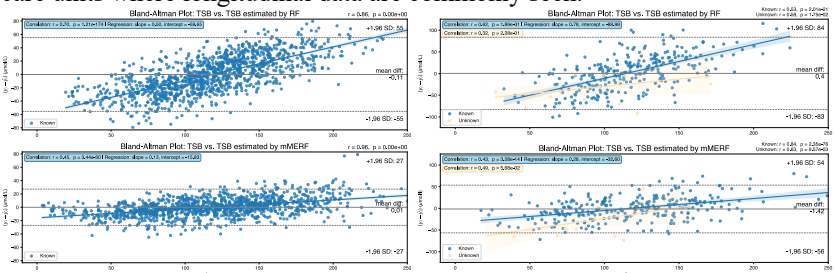

Fig. 1: Bland-Altman plots comparing RF (upper) and mMERF (lower) models via the difference between observed (y_i) and estimated ($\hat{y}_i$) TSB levels. Left panels: training data. Right panels: testing data.