Preterm Infants Length of Stay Prediction Achieving Better Accuracy with Greater Usability

Tzviel Frostig[1], Yoav Benjamini[1], Orli Kehat[2], Ahuva Weiss-Meilik[2], Dror Mandel[1,2],

Ben Peleg[1,3], Zipora Strauss[1,3], and Alexis Mitelpunkt[1,2]

[1] Tel Aviv University [2] Tel Aviv Medical Center [3] Sheba Medical Center

Objective Length of stay (LOS) in the neonatal intensive care unit (NICU) has an important implication on hospital occupancy, healthcare costs, management, and parent anxiety. The goal was to obtain a generalizable, and robust model to predict LOS of preterm infants in the NICU.

Data Data collected Tel-Aviv medical center NICU between the years 2011-2018 containing 5,966 newborns with all available data in their electronic health records.
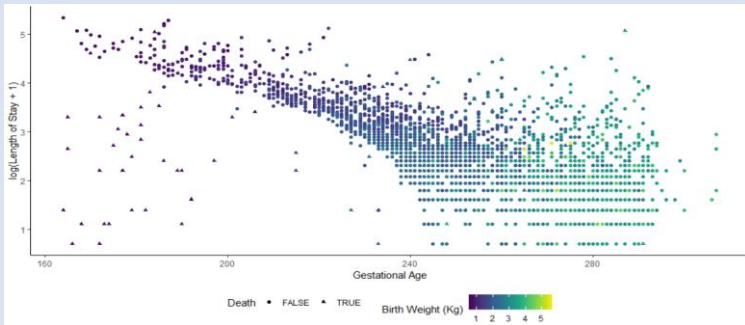
Log(LOS) as a function of GA (Gestational Age, x-axis) and birth-weight (color).

Methods We compared the prediction accuracy of the standard linear regression (LM) and quantile regression (QR) models (with variable selection) for each age-accumulated data (and term vs preterm group). Three variable selection methods were employed: LASSO, AIC step forward selection and step forward selection method based on FDR.

Focusing the prediction The prediction is focused on the first day (earliest prediction) and on preterm infants (tend to stay longer)

	Preterm				Term			
Day	N	LOS	MAE	R^2	N	LOS	MAE	R^2
1	1879	18.6	5.64	0.82	1343	5.9	4.63	0.14
1	1841	18.9	5.47	0.83	1003	7.8	4.96	0.14
7	1713	20.3	5.49	0.82	575	11.6	5.81	0.1

Results of Random Forest for various time durations. Term infants' LOS is harder to predict

Selected model A simple Quantile Regression model is found using AIC forward step selection. It achieves results comparable or better with more sophisticated models. It achieves significantly better results than currently used models.

The selected model is based on APGAR, gestational age, birth-weight and sex.

	MAE	R^2
Selected model	6.26	0.745
Naive regression	8.19	0.697
Israel Reimbursement	12.4	0.279
Discharge at term	9.26	0.62

Comparing various models used for LOS prediction.

Validation The model was externally validated using 6236 infants from Sheba Medical Center.

	N	LOS	MAE	R^2
Training	1879	18.6	6	0.81
Survivors test	1152	18.2	6.3	0.75
Validation	6236	22.2	6.0	0.76

Results for various data-sets used

App A simple to use web-app was created to allow practitioners and parents access to the model, tzviel.shinyapps.io/calcuLOS.