Final Project Evaluation and Summary

December 07, 2022

1 Evaluation

1.1 Three Test Sites

As stated in the subsection *Training and Evaluation Split*, three hospitals were randomly chosen as the sites for evaluation: 5, 16, and 17. These three sites were left out during the exploratory data analysis, model training, and model stability assessment.

1.2 Baseline Performance

We first applied the decision rule derived by Leonard et al. (2021) to the three test sites to evaluate the baseline performance, to which the relative performance of our models would later be evaluated. Using the 6-variable decision rule (altered mental status, focal neurologic deficit, complaint of neck pain, substantial injury to the torso, high-risk motor vehicle crash, and diving), the sensitivity and specificity of identifying cervical spine injury by the presence of at least one of these six factors were 85% (95% CI 78% to 92%) and 44% (95% CI 39% and 49%), respectively.

Table 1: Decision Rule from Leonard et al. (2011)

	Estimate	Lower CI Bound	Upper CI Bound
Sensitivity	0.85	0.78	0.92
Specificity	0.44	0.39	0.49

1.3 Forward Variable Selection Performance

We then applied the decision rule derived by using forward variable selection on the decision rule by Leonard et al. (2011) to the three test sites. This decision rule consisted of four additional variables: conditions predisposing to cervical spine injury, high-risk hit by car, axial load to top of the head, and injury by clothes-lining. (Note that at each forward step, the forward selection process adds the one variable that gives the single best improvement to the model. Hence, it is possible for the final model to contain variables that are significant when added at each respective step but are no longer as significant in the presence of subsequently added variables (Leonard et al. 2021).) The sensitivity and specificity of identifying cervical spine injury by the presence of at least one of these ten factors was 88% (95% CI 82% to 95%) and 30% (95% CI 26% and 35%), respectively. Compared to the baseline, there was a slight increase in the sensitivity (3% increase) and decrease in the specificity (14% decrease).

Table 2: Forward Variable Selection

	Estimate	Lower CI Bound	Upper CI Bound
Sensitivity	0.88	0.82	0.95
Specificity	0.30	0.26	0.35

2 Single Decision Tree Performance

We tested the single decision tree model (using the aforementioned ten variables) to the three test sites. The sensitivity and specificity was 91% (95% CI 85% to 96%) and 47% (95% CI 42% and 52%), respectively. Compared to the baseline, there were increases in both the sensitivity (6% increase) and in the specificity (3% increase).

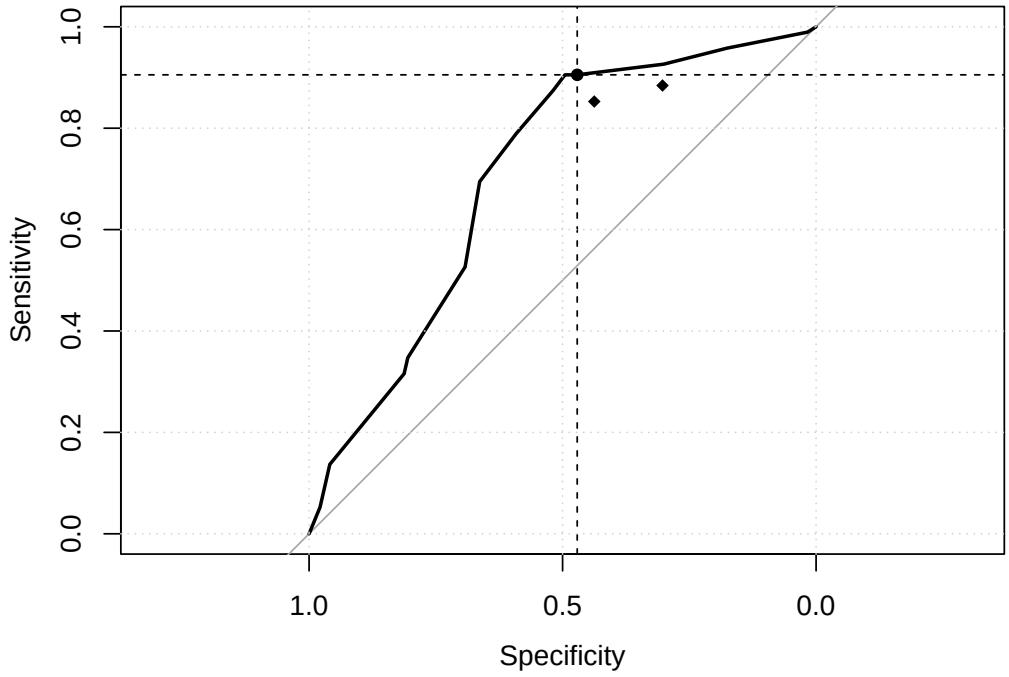


Table 3: Single Decision Tree with 10 Variables

	Estimate	Lower CI Bound	Upper CI Bound
Sensitivity	0.91	0.85	0.96
Specificity	0.47	0.42	0.52

3 Gradient-Boosted Tree Performance

We tested the gradient-boosted decision tree model to the three test sites. The sensitivity and specificity was 92% (95% CI 86% to 97%) and 41% (95% CI 36% and 46%), respectively. Compared to the single-tree model, there was a slight increase in the sensitivity (1% increase) and decrease in the specificity (6% decrease). Compared to the baseline, there was an increase in the sensitivity (7% increase) and a decrease in the specificity (3% decrease).

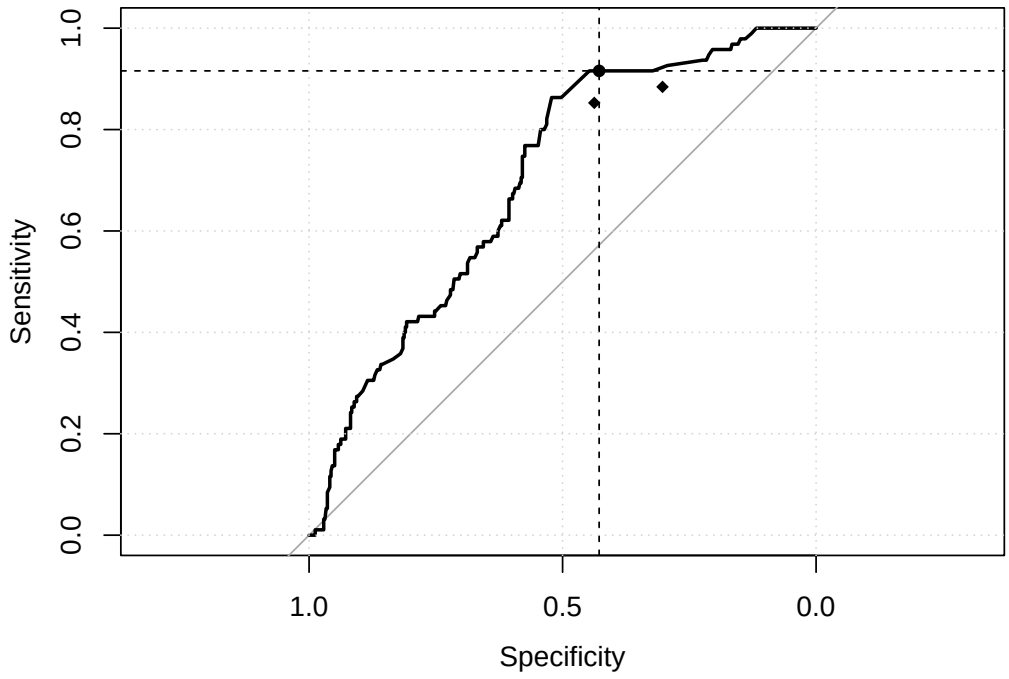


Table 4: Gradient-Boosted Tree

	Estimate	Lower CI Bound	Upper CI Bound
Sensitivity	0.92	0.86	0.97
Specificity	0.43	0.38	0.48

4 Logistic Regression

We tested the logistic regression model (based on all variables from feature selection) to the three test sites. The sensitivity and specificity was 88% (95% CI 82% to 95%) and 43% (95% CI 38% and 48%), respectively. Compared to the baseline, there were increases in both the sensitivity (3% increase) and in the specificity (1% increase).

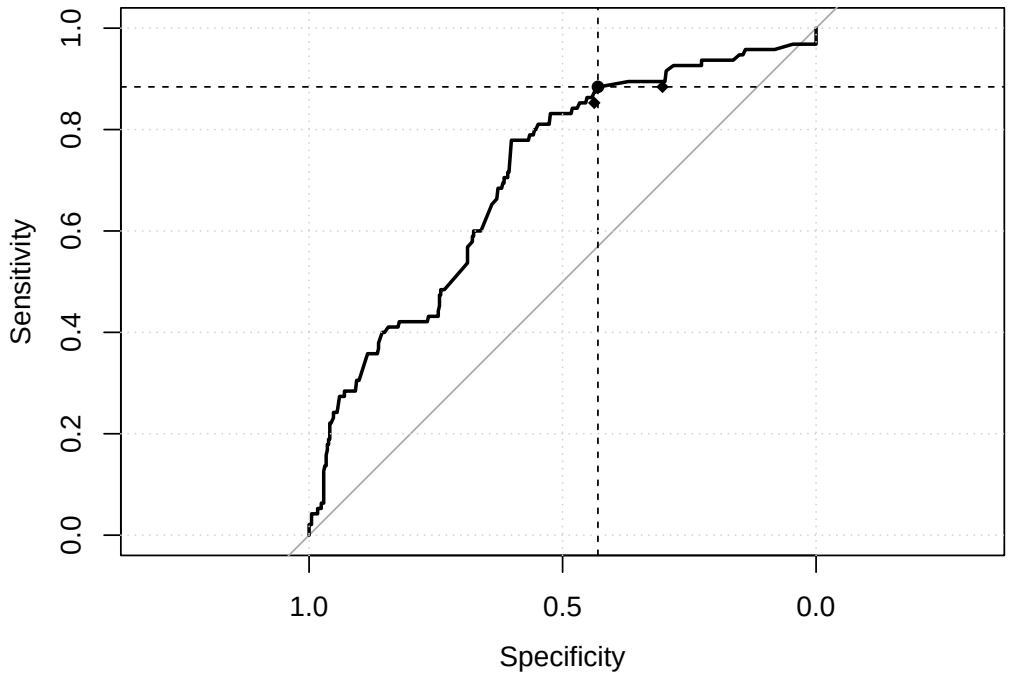
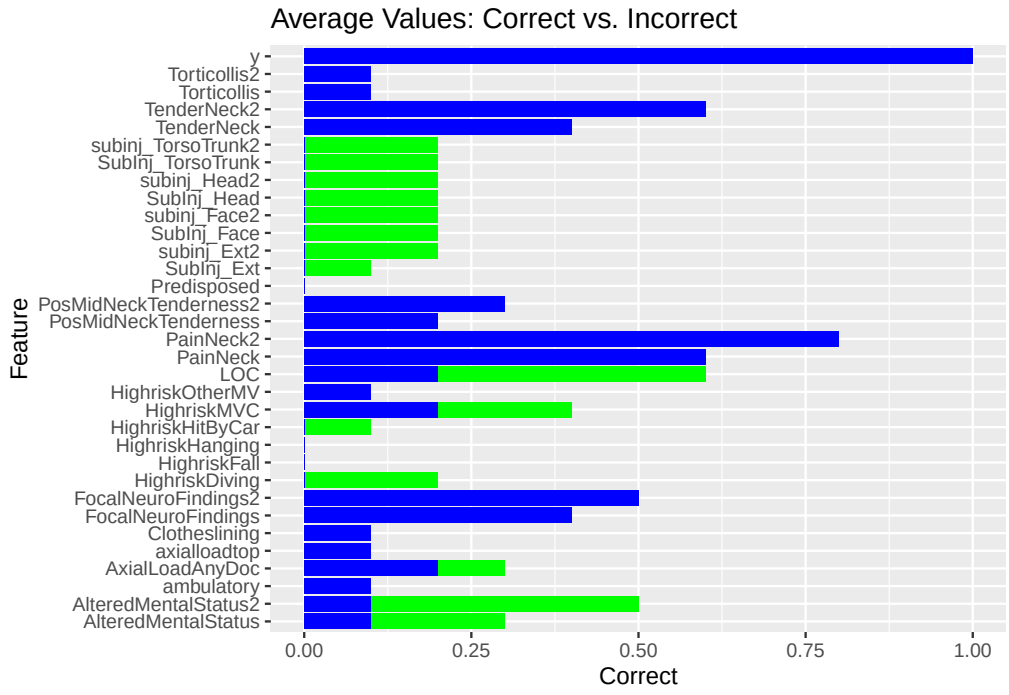


Table 5: Logistic Regression

	Estimate	Lower CI Bound	Upper CI Bound
Sensitivity	0.88	0.82	0.95
Specificity	0.43	0.38	0.48



5 Summary

By investigating primary source data for cervical spine injury (CSI) in children, we derived and validated a clinical decision rule that aims to guide imaging decisions for children who experienced blunt trauma (e.g., altered mental status, focal neurologic deficits, complaint of neck pain, torticollis). Based on our results and stability check, we believe that our decision rule will (1) provide satisfactory performance when applied to future data and (2) perform better than the baseline case in terms of both specificity and sensitivity.