

Supplemental Digital Content 1

SAMPLE SIZE

The need for sample size calculation is for comparison of AUROCs. Before the study, we hypothesized the base model ROC to be 0.60, and the improved model ROC to be 0.88. We set type I error = 0.05, type II error = 0.10. We hypothesized the rank correlation coefficient in the positive and negative groups were 0.5. For various outcomes, the ratio of positive and negative sample size could range from 5 to 33 (corresponding to positive case percentages of 3% to 10%), the number of positive cases required ranges from 13 to 14; and the number of negative case required ranges from 70 to 429 (calculated by MedCalc). If we set type I error = 0.01, the required positive cases ranges from 18 to 20, and negative cases ranges from 100 to 594.

TEST of GOODNESS-OF-FIT

In evaluating different hemoglobin measures' prediction capability for transfusion, we selected multivariate logistic regression. Instead of using other more sophisticated algorithms such as support vector machine, discriminant analysis, we used a simple model to study the utilities of those measures. There were many multivariate logistic regression models built for different outcomes and different variable combinations, adequateness of each model was investigated by Hosmer-Lemeshow goodness of fit test (analyzed by R ResourceSelection package; 10 groups).

Table S1. H-L goodness of fit test for model Base HR + SpHb

Predictor(s): HR + SpHb	X-Squared	p-value
pRBC1-3hr	12.6756	0.1235
pRBC1-6hr	14.2147	0.0763
pRBC1-12hr	10.8932	0.2078

Table S2. H-L goodness of fit test for model Base HR + lab Hb

Predictor(s): HR + lab Hb	X-Squared	p-value
pRBC1-3hr	4.2134	0.8374
pRBC1-6hr	8.5617	0.3806
pRBC1-12hr	13.292	0.1022

Table S3. H-L goodness of fit test for model Base HR + other lab

Predictor(s): HR + other lab	X-Squared	p-value
pRBC1-3hr	5.7438	0.6759
pRBC1-6hr	5.6534	0.6860

pRBC1-12hr	9.8833	0.2733
------------	--------	--------