

Supplemental Digital Content

Supplemental Table 1. Measures of Predictive Performance between Black and White cases for the threshold of optimal sensitivity and specificity (VASQIP, RAI 26, 3.5%)

Supplemental Table 2. Measures of Predictive Performance between Black and White cases for the threshold of detectable frailty (NSQIP, RAI 30, 2.3% predicted 30-day mortality)

Supplemental Figure 1. Comparison of ROC curves between Black and White cases in the RAI_{Base} (NSQIP)



Supplemental Table 1. Measures of Predictive Performance between Black and White cases for the threshold of optimal sensitivity and specificity (VASQIP, RAI 26, 3.5%)



		Bias Assessment				Comparative Model Performance		
		RAI _{Base}	RAI _{Balanced}	RAI _{Age-adj}	RAI _{Fair}	Δ from RAI _{Base}		
		Statistic	Statistic	Statistic	Statistic	Δ RAI _{Balanced}	Δ RAI _{Age-adj}	Δ RAI _{Fair}
Accuracy	Black	0.829	0.823	0.803	0.765	-0.6%	-2.6%	-6.4%
	White	0.794	0.786	0.801	0.648	-0.8%	0.7%	-14.6%
	Bias	3.5%	3.7%	0.2%	11.7%	0.2%	-3.3%	8.2%
Balanced Accuracy	Black	0.783	0.782	0.783	0.693	-0.1%	0.0%	-9.0%
	White	0.762	0.76	0.761	0.68	-0.2%	-0.1%	-8.2%
	Bias	2.1%	2.2%	2.2%	1.3%	0.1%	0.1%	-0.8%
True Positive Rate (Sensitivity)	Black	0.733	0.737	0.762	0.616	0.4%	2.9%	-11.7%
	White	0.727	0.733	0.718	0.714	0.6%	-0.9%	-1.3%
	Bias	0.6%	0.4%	4.4%	-9.8%	-0.2%	3.8%	-10.4%
True Negative Rate (Specificity)	Black	0.833	0.827	0.804	0.771	-0.6%	-2.9%	-6.2%
	White	0.797	0.788	0.804	0.645	-0.9%	0.7%	-15.2%
	Bias	3.6%	3.9%	0.0%	12.6%	0.3%	-3.6%	9.0%
False Positive Rate	Black	0.167	0.173	0.196	0.229	0.6%	2.9%	6.2%
	White	0.203	0.212	0.196	0.355	0.9%	-0.7%	15.2%
	Bias	3.6%	3.9%	0.0%	12.6%	0.3%	-3.6%	9.0%
False Negative Rate	Black	0.267	0.263	0.238	0.384	-0.4%	-2.9%	11.7%
	White	0.273	0.267	0.282	0.286	-0.6%	0.9%	1.3%
	Bias	0.6%	0.4%	4.4%	-9.8%	-0.2%	3.8%	-10.4%
F1 score	Black	0.242	0.237	0.223	0.163	-0.5%	-1.9%	-7.9%
	White	0.206	0.2	0.209	0.129	-0.6%	0.3%	-7.7%
	Bias	3.6%	3.7%	1.4%	3.4%	0.1%	-2.2%	-0.2%
MCC	Black	0.274	0.27	0.26	0.171	-0.4%	-1.4%	-10.3%
	White	0.236	0.232	0.239	0.14	-0.4%	0.3%	-9.6%

	Bias	3.8%	3.8%	2.1%	3.1%	0.0%	-1.7%	-0.7%
Brier score	Black	0.0309	0.0308	0.0309	0.0339	0.0001	0.0000	0.0030
	White	0.0306	0.0306	0.0306	0.0335	0.0000	0.0000	0.0029
	Bias	-0.0003	-0.0002	-0.0003	-0.0004	0.0001	0.0000	-0.0001
Harrel's C	Black	0.8546	0.833	0.8334	0.7614	-0.0216	-0.0212	-0.0932
	White	0.8335	0.8545	0.8546	0.7867	0.021	0.0211	-0.0468
	Bias	0.0211	-0.0215	-0.0212	-0.0253	0.0426	0.0423	0.0464
Equal Opportunity^a		1.01	1.01	1.06	0.86	0.0%	5.0%	-15.0%
Predictive Equality^b		1.05	1.05	1	1.2	0.0%	-5.0%	15.0%

^aRatio of Sensitivity between Black and White cases – ratio higher than 1 suggests better sensitivity in Black patients than White, ratio less than 1 suggests better sensitivity in White patients than Black

^bRatio of Specificity between Black and White cases – ratio higher than 1 suggest better specificity in Black patients than White, ratio less than 1 suggests better specificity in White patients than Black

For comparisons of racial bias,  indicates metrics with better performance among Black patients than White, while  indicates metrics with better performance among White patients than Black.

For comparisons of various RAI versions,  indicates improvements in predictive performance, while  indicates reductions in predictive performance.

Abbreviations: RAI, Risk Analysis Index; MCC, Matthew's correlation coefficient



Supplemental Table 2. Measures of Predictive Performance between Black and White cases for the threshold of detectable frailty (NSQIP, RAI 30, 2.3% predicted 30-day mortality)



		RAI_{Base}
		Statistic
Accuracy	Black	0.921
	White	0.906
	Bias	1.5%
Balanced Accuracy	Black	0.73
	White	0.733
	Bias	-0.3%
True Positive Rate (Sensitivity)	Black	0.534
	White	0.557
	Bias	-2.3%
True Negative Rate (Specificity)	Black	0.925
	White	0.91
	Bias	1.5%
False Positive Rate	Black	0.075
	White	0.09
	Bias	1.5%
False Negative Rate	Black	0.466
	White	0.443
	Bias	-2.3%
F1 score	Black	0.119
	White	0.108
	Bias	1.1%
MCC	Black	0.169
	White	0.16
	Bias	0.9%

Brier score	Black	0.0095
	White	0.0097
	Bias	0.0002
Harrel's C	Black	0.8544
	White	0.8476
	Bias	0.0068
Equal Opportunity^a		0.96
Predictive Equality^b		1.02

^aRatio of Sensitivity between Black and White cases – ratio higher than 1 suggests better sensitivity in Black patients than White, ratio less than 1 suggests better sensitivity in White patients than Black

^bRatio of Specificity between Black and White cases – ratio higher than 1 suggest better specificity in Black patients than White, ratio less than 1 suggests better specificity in White patients than Black

For comparisons of racial bias,  indicates metrics with better performance among Black patients than White, while  indicates metrics with better performance among White patients than Black.

For comparisons of various RAI versions,  indicates improvements in predictive performance, while  indicates reductions in predictive performance.

Abbreviations: RAI, Risk Analysis Index; MCC, Matthew's correlation coefficient

Supplemental Figure 1. Comparison of ROC curves between Black and White cases in the RAI_{Base} (NSQIP)

