COMPARATIVE ASSESSMENT OF RISK SCORE PERFORMANCE FOR PREDICTING POST RUPTURED AAA REPAIR MORTALITY: AN IN-DEPTH ANALYSIS IN THE NSQIP AND VQI REGISTRIES

Andrew S. Warren BS BA[1], Elina Quiroga MD[1], Thomas FX O'Donnell MD[2], Marc L. Schermerhorn MD[3], Sara L. Zettervall MD MPH[1], Kirsten Dansey MD MPH[1]

[1]University of Washington Division of Vascular Surgery, [2]Division of Cardiac, Thoracic, and Vascular Surgery, New York-Presbyterian Columbia University Medical Center, [3]Division of Vascular and Endovascular Surgery, Beth Israel Deaconess Medical Center

Background

- Ruptured abdominal aortic aneurysm (rAAA) repair remains a high-risk surgical procedure
- Six scoring measures have been created to predict risk (Glasgow Aneurysm Score (GAS), Modified Harborview Risk Score (HRS), Hardman Index, Leiden Score, Medicare Risk Score, VSGNE Score)

Objective

- Critically evaluate and compare the performance of these six risk scores in NSQIP and VQI patients undergoing rAAA repair

Methods

- Patients who underwent rAAA repair in NSQIP and VQI were included
- Each risk score was calculated for all included patients
- Discrimination was assessed using the receiver operator characteristic area under the curves (AUC) and compared within each dataset using the Delong test
- Calibration was assessed using graphical calibration curves and quantified via the integrated calibration index (ICI)

Table 1: Risk Score Components

Risk Score	Components
Glasgow Aneurysm Score*	Age, shock, myocardial disease, cerebrovascular disease, renal disease
Modified Harborview Risk Score*	Age, hypotension, creatinine, INR
Hardman Index*	Age, creatinine, LOC, hemoglobin, electrocardiographic ischemia
Leiden Score*	Age, sex, prior MI, CHF, electrocardiographic ischemia, renal failure, COPD
Medicare Risk Score†	Age, sex, renal failure, CHF, vascular disease, repair type (open vs endovascular)
VSGNE Score†	Age, cardiac arrest, LOC, use of suprarenal clamp

**Exclusively Pre-Operative Variables, †Includes Intra-Operative Variables*

Figure 1: Receiver Operator Characteristic Curves

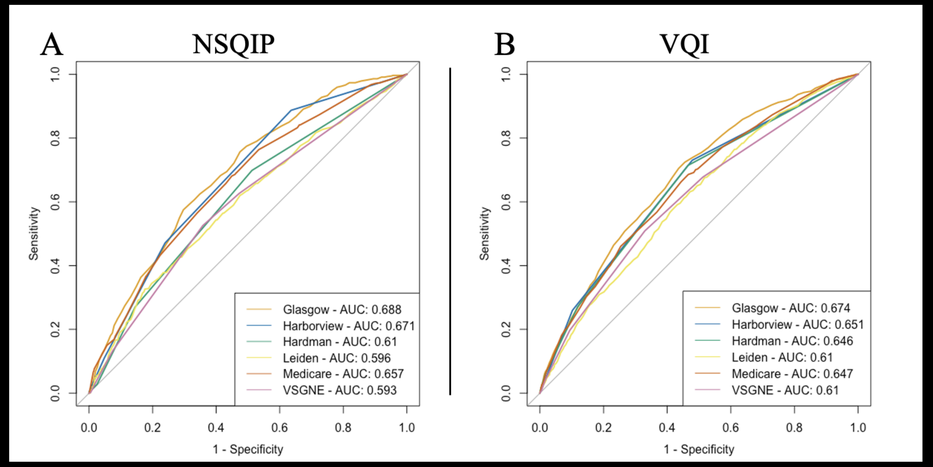
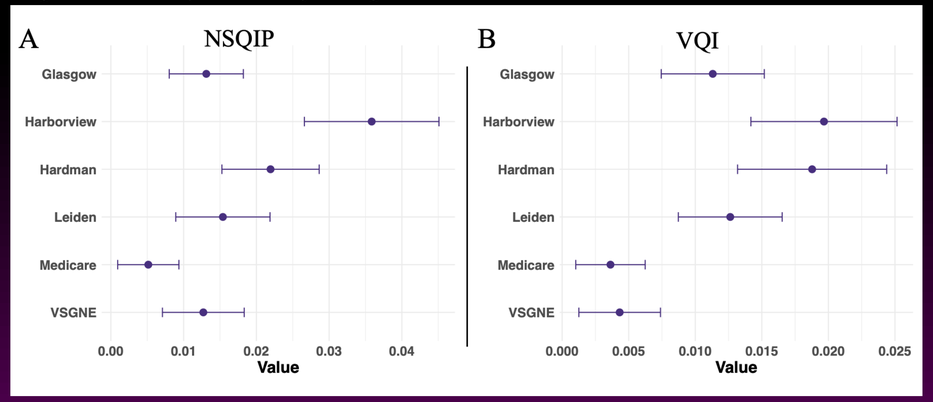


Figure 2: Comparison of Integrated Calibration Index



Results

Discrimination

- In both NSQIP and VQI, the GAS demonstrated a superior ability to discriminate (AUC 0.69 NSQIP, 0.67 VQI) vs. all other risk scores ($p<0.01$)
- The modified HRS (AUC 0.67 NSQIP, 0.65 VQI) and Medicare risk score (AUC 0.66 NSQIP, 0.65 VQI) were comparable ($p=0.37$ NSQIP, $p=0.54$ VQI) to each other, following the GAS in discrimination ability
- The Hardman Index performed similarly to the modified HRS and Medicare risk score in VQI ($p=0.26$ vs modified HRS, $p=0.43$ vs Medicare) but worse in NSQIP ($p<0.01$ vs both)

Calibration

- In NSQIP, all risk scores were well calibrated with predicted and observed mortality differing by <4% for all scores
  - Medicare risk score performed best with an ICI of 0.005
- In VQI, all risk scores again were well calibrated with predicted and observed mortality differing by <2% for all scores
  - Medicare risk score performed best with an ICI of 0.004

Conclusions

- All risk scores display satisfactory ability discerning between high and low-risk patients as marked by their AUC values
- All risk scores are well calibrated to accurately predict risk
- The Glasgow Aneurysm Score and modified Harborview Risk Score appear to be the best pre-operative predictors of post ruptured abdominal aortic aneurysm repair mortality
- The Medicare Risk Score, while also robust, is limited to retrospective use