National Cardiac Arrest Audit Public Report 2022–23
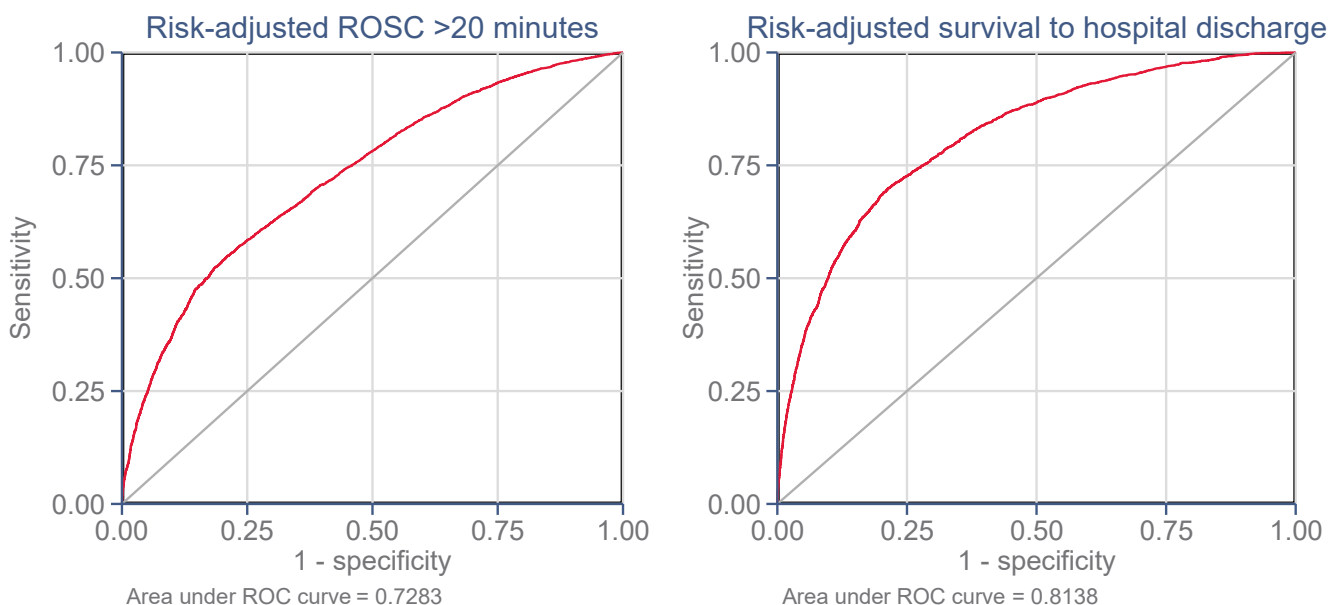
Statistical appendix
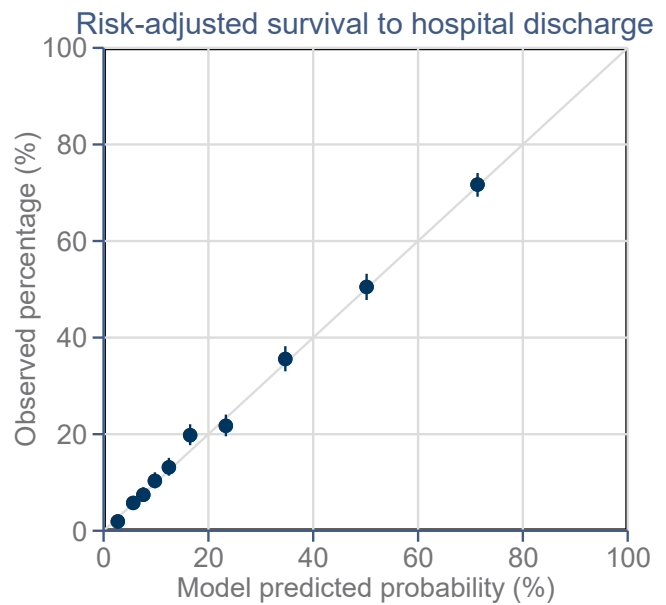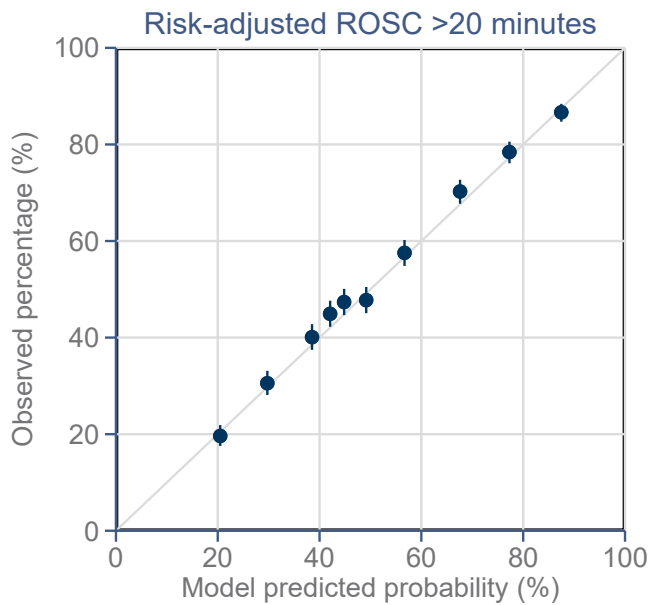
Model fit

The NCAA risk models were developed and published in 2014[1] and were subsequently recalibrated in 2018 and 2023. The $NCAA_{2023}$ risk models are the latest recalibrations using data from 21,646 in-hospital cardiac arrests in 190 hospitals participating in NCAA between 1 April 2021 and 31 December 2022. For the purpose of the NCAA Public Report 2022-23, the fit of the model was assessed based on 14,019 in-hospital cardiac arrests in 184 hospitals using the following methods:

- Discrimination was assessed using the c index,[2] equivalent to the area under the receiver operating characteristic (ROC) curve[3];
- Calibration was assessed graphically by dividing the dataset into 10 equal-sized groups based on quantiles of predicted risk (note that the Hosmer-Lemeshow test for perfect calibration was not used, as in a sample of this size, statistically significant departures from perfect calibration would be expected even with a well-fitting model[4]);
- Overall goodness of fit was assessed with Brier's score, representing the mean squared error between outcomes and predictions.[5]

The c index (area under the ROC curve) was 0.728 (95% confidence interval 0.720 to 0.737) for ROSC > 20 minutes and 0.814 (95% confidence interval 0.805 to 0.822) for survival to hospital discharge. Calibration was qualitatively good (see below). Brier's score was 0.209 and 0.135. The ROC curve and calibration plot for each indicator are shown below:

Risk-adjusted ROSC >20 minutes

Area under ROC curve = 0.7283

Risk-adjusted survival to hospital discharge

Area under ROC curve = 0.8138

Risk-adjusted ROSC >20 minutes — Observed percentage (%) vs Model predicted probability (%)

Risk-adjusted survival to hospital discharge — Observed percentage (%) vs Model predicted probability (%)

Expected value and predicted range

In this report, the observed value for each risk-adjusted outcome is compared against an expected value. The expected value is calculated as the mean predicted probability of ROSC > 20 minutes or survival to hospital discharge from the NCAA$_{2023}$ model for all eligible team visits in that hospital.

To compare the observed value with the expected value, we calculate predicted ranges based on the expected value and the number of eligible team visits using Binomial predictive distributions, equivalent to a slice through a funnel plot specific to the individual hospital.[6] Over-dispersion was adjusted for by estimating a multiplicative over-dispersion factor for each quality indicator based on 10% Winsorisation.[6,7] The estimated over-dispersion factors were: Risk-adjusted ROSC > 20 minutes, 1.24; Risk-adjusted survival to hospital discharge, 1.23. There was no significant over-dispersion in either indicator when restricted to arrests on the ward.

The observed value is plotted against bars corresponding to the 95% and 99.8% predicted ranges, an approach identified as being understandable to the public and reinforcing appropriate interpretation of results.[8]

References

1. Harrison DA, Patel K, Nixon E, Soar J, Smith GB, Gwinnutt C, Nolan JP, Rowan KM on behalf of the National Cardiac Arrest Audit. Development and validation of risk models to predict outcomes following in-hospital cardiac arrest attended by a hospital-based resuscitation team. *Resuscitation* 2014; 85:993-1000.
2. Harrell FE, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. *JAMA* 1982; 247:2543-6.
3. Hanley JA, McNeil BJ. The meaning and use of the area under the receiver operating characteristics (ROC) curve. *Radiology* 1982; 143:29-36.
4. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. *Crit Care Med* 2007; 35:2052-6.
5. Brier GW. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 1950; 75:1-3.
6. Spiegelhalter D. Funnel plots for comparing institutional performance. *Stat Med* 2005; 24:1185-202.
7. Spiegelhalter D. Handling over-dispersion of performance indicators. *Qual Saf Health Care* 2005; 14:347-51.
8. Rakow T, Blackshaw E, Pagel C, Spiegelhalter DS. Comparing what to what, on what scale? The impact of item comparisons and reference points in communicating risk and uncertainty. *J Behav Decis Mak* 2018; 31:547-61.