

HHS Public Access

Author manuscript Resuscitation. Author manuscript; available in PMC 2018 June 01.

Published in final edited form as: Resuscitation. 2017 June ; 115: 5–10. doi:10.1016/j.resuscitation.2017.02.022.

Predicting In-Hospital Mortality for Initial Survivors of Acute Respiratory Compromise (ARC) Events; Development and Validation of the ARC Score

Ari Moskowitz, MD1, **Lars W. Andersen, MD MPH PhD**2,3, **Mathias Karlsson, BS**2,3, **Anne V. Grossestreuer, PhD**2,4, **Maureen Chase, MD MPH**2, **Michael N. Cocchi, MD**2,5, **Katherine Berg, MD**1, and **Michael W. Donnino, MD**1,2 **for the American Heart Association's Get With the Guidelines-Resuscitation Investigators**

¹Beth Israel Deaconess Medical Center, Department of Medicine, Division of Pulmonary, Critical Care, and Sleep Medicine, Boston MA

²Beth Israel Deaconess Medical Center, Department of Emergency Medicine, Boston MA

³Research Center for Emergency Medicine, Aarhus University Hospital, Aarhus, Denmark

⁴University of Pennsylvania, Department of Emergency Medicine, Philadelphia PA

⁵Beth Israel Deaconess Medical Center, Department of Anesthesia Critical Care and Pain Medicine, Division of Critical Care

Abstract

Aim—Acute respiratory compromise (ARC) is a common and highly morbid event in hospitalized patients. To date, however, few investigators have explored predictors of outcome in initial survivors of ARC events. In the present study, we leveraged the American Heart Association's Get With The Guidelines[®]-Resuscitation (GWTG-R) ARC data registry to develop a prognostic score for initial survivors of ARC events.

Methods—Using GWTG-R ARC data, we identified 13,193 index ARC events. These events were divided into a derivation cohort (9,807 patients) and a validation cohort (3,386 patients). A score for predicting in-hospital mortality was developed using multivariable modeling with generalized estimating equations.

Results—The two cohorts were well balanced in terms of baseline demographics, illness-types, pre-event conditions, event characteristics, and overall mortality. After model optimization, nine variables associated with the outcome of interest were included. Age, hypotension preceding the event, and intubation during the event were the greatest predictors of in-hospital mortality. The final score demonstrated good discrimination in both the derivation and validation cohorts. The

Ari Moskowitz, MD: Division of Pulmonary, Critical Care, and Sleep Medicine, 330 Brookline Avenue, Boston, MA 02215, Phone: 617-667-5864, Fax: 617-667-4839, amoskowi@BIDMC.harvard.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

score was also very well calibrated in both cohorts. Observed average mortality was <10% in the lowest score category of both cohorts and >70% in the highest category, illustrating a wide range of mortality separated effectively by the scoring system.

Conclusions—In the present study, we developed and internally validated a prognostic score for initial survivors of in-hospital ARC events. This tool will be useful for clinical prognostication, selecting cohorts for interventional studies, and for quality improvement initiatives seeking to riskadjust for hospital-to-hospital comparisons.

Keywords

Respiratory Failure; Severity of Illness Index; American Heart Association; Hospital Rapid Response Team

Introduction

Acute respiratory compromise (ARC) is a common trigger for hospital-wide rapid response team activation and carries a significant morbidity/mortality burden^{1,2}. Patients suffering from ARC frequently require the initiation of mechanical ventilation and admission to an intensive care unit. Not only are these patients at risk from delayed recognition of clinical decompensation and complications in the peri-intubation period (e.g. hypotension, cardiac arrest),³ but also often go on to develop multi-organ dysfunction⁴. Despite their strong association with poor outcome, ARC events in hospitalized patients are not a well-studied entity. To date, the natural history of hospital ARC events and their relationship to inhospital mortality has not been well described.

Prognostic scores play an important role in critical care medicine where they are used to guide selection of advanced therapies, aid families with difficult decision making, and select patients for enrollment in clinical trials. Using the nationwide Get with the Guidelines Resuscitation® (GWTG-R) cardiopulmonary arrest registry, investigators have recently developed the Cardiac Arrest Survival Post-Resuscitation In-hospital (CASPRI) score^{5,6} and the Good Outcome Following Attempted Resuscitation (GO-FAR) score⁷ to assist with predicting outcomes following in-hospital cardiac arrest. In addition to their prognostic utility, these scores also provide a mechanism for hospital-to-hospital adjusted outcome comparisons to aid in the optimization of quality within health care systems.⁸

While a number of clinical models exist for predicting mortality following cardiac arrest, no models exist to assist with prognostication for patients who have suffered and survived an ARC event. The relatively common occurrence of ARC, coupled with its high morbidity and cost to the healthcare system^{4,9}, justifies further investigation into the predictors of poor outcome following such events. In the present study, we aim to use the GWTG-R ARC database to develop a useful prognostic score for initial survivors of an in-hospital ARC event.

Methods

The GWTG-R database is a large, prospective, quality-improvement registry of in-hospital cardiac arrests, medical emergency team activations, and ARC events sponsored by the American Heart Association. Hospitals participating in the registry submit clinical information regarding the medical history, hospital care, and outcomes for consecutive patients using an online, interactive case report form and Patient Management Tool (Quintiles, Cambridge, Massachusetts). The design, data collection, and quality control of the GWTG-R repository have been described previously^{10,11,12}. In the GWTG-R database, information on ARC events has been captured since 2004. As of the time of this analysis, over 300 unique hospital sites participate in data collection for the GWTG-R ARC data registry. In our study, we included all adult ARC events occurring between June 2005 and June 2015. Authors AM and LA had access to all of the data in the study and take responsibility for the integrity of the data.

Within the GWTG-R data registry, ARC events are defined by absent, agonal or inadequate respiration that requires emergency assisted ventilation, including non-invasive (e.g., mouthto-mouth, mouth-to-barrier device, bag-valve-mask, continuous positive airway pressure or bilevel positive airway pressure) or invasive (e.g., endotracheal or tracheostomy tube, laryngeal mask airway) positive pressure ventilation. Elective intubations and events beginning outside the hospital are not included.¹²

Cohort Selection

From the GWTG-R ARC database we initially identified 34,104 ARC events in adults (age ≥ 18) occurring between June 2005 and June 2015. Of these, we included only index cases, leading to the exclusion of 1,644 non-index events. We also excluded ARC events suffered by visitors/staff $(n=41)$, those that occurred in the emergency department $(n=3,487)$, those in which an invasive airway or non-invasive positive pressure ventilation was already in place $(n=5,676)$, and those that resulted in cardiac arrest $(n=1,755)$. An additional 467 events were excluded due to missing data on in-hospital mortality. Finally, 7,841 patients had missing data on one or more candidate variables. A comparison between the group with missing data and the final study cohort can be found in Table S1 in the Supplementary Materials. The final study cohort included 13,193 unique ARC events. See Figure 1 for complete details on cohort selection.

Candidate Variables and Primary Outcome

Candidate variables were selected on the basis of their potential as predictors of the study outcome. All variables collected in the GWTG-R ARC data registry were screened for inclusion by three clinicians and any discrepancies resolved through consensus. Included candidate variables are listed in Table 1. For definitions of those variables, please refer to Table S2 included in the Supplementary Materials. The outcome of interest was in-hospital mortality.

Statistical Analysis

The entire cohort was divided at random into a derivation (3/4 of the patients) and a validation (1/4 of the patients) cohort. Descriptive statistics were used to summarize each cohort. For the derivation cohort, the *a priori* chosen candidate variables (see Table 1) were entered into a multivariable logistic regression model with in-hospital mortality as the binary outcome of interest. To account for correlations between patients within the same hospital, we performed hierarchic modeling using generalized estimation equations (GEE) with an exchangeable variance-covariance structure. To create a parsimonious model, we next performed backwards selection to eliminate variables not associated with in-hospital mortality. We removed variables one by one (based on the highest p-value) until all variables were associated with the outcome at a p-value < 0.01. We next used the Quasi-likelihood under the Independence model Criterion (QIC) statistic¹³ and removed remaining variables that did not improve the GEE model fit based on the QIC. For categorical variables with more than two categories, categories with similar outcomes (i.e. similar beta coefficients) were collapsed to simplify the model.

We assessed model discrimination in both the full non-parsimonious model and in the final parsimonious GEE model with the c-statistic 14 with 95% confidence intervals estimated by bootstrapping 1000 data sets using unrestricted random sampling. Model calibration was assessed with the Hosmer-Lemeshow test and observed vs. expected plots.

From the final parsimonious GEE model, we used the beta-coefficients to create a clinically useful prediction score. The discrimination and calibration of this score was then assessed first in the derivation cohort and then in the validation cohort. Given the long inclusion period used in this study (2005–2015), and the potential for changes in patient demographics and ARC management strategies over that time period, we additionally performed a post-hoc subgroup analysis in the validation cohort including only patients enrolled in GWTG-R between January 2012 and June 2015.

SAS, version 9.4, (SAS Institute, Cary, NC, USA) was used for all analyses. Quintiles is the data collection coordination center for the American Heart Association/American Stroke Association Get With The Guidelines® programs.

Results

Baseline and event characteristics for the 9,807 patients in the derivation cohort and 3,386 patients in the validation cohort were similar as demonstrated in Table 1. Overall, a large majority of ARC events ended with the patient requiring some form of assisted ventilation (non-invasive positive pressure ventilation or endotracheal intubation). Only 9% of patients in each cohort had return of spontaneous ventilation (no further need for assisted ventilation) at the conclusion of the ARC event. Ultimately, 3863 (39%) patients in the derivation cohort and 1305 (39%) patients in the validation cohort died before hospital discharge.

A total of 21 candidate variables were identified. The c-statistic from the full, nonparsimonious model was 0.68 (95%CI: 0.66, 0.69). After elimination of 12 variables whose exclusion did not worsen the QIC, the c-statistic was unchanged (c-statistic 0.67, 95%CI:

0.66, 0.68). Calibration was assessed by the generation of an observed vs. predicted plot (Figure 2) and the Hosmer-Lemeshow goodness-of-fit test $(p=0.05)$. By visual inspection, the calibration appeared excellent.

Beta-coefficients from the final model were used to create a prediction score (Table 2). Particularly strong, independent predictors of in-hospital mortality were increasing age (especially age >80), hypotension within 4-hours of the event, and need for intubation during the event. The c-statistic for the score was 0.67 and the Hosmer-Lemeshow goodness-of-fit test demonstrated good fit (p=0.64). In the derivation cohort, scores ranged from 0–51 in an approximately normal distribution (Figure 3) with a mean of 21 (\pm 7). The odds ratio for in-hospital mortality per unit increase in the score was 1.10 (95%CI: 1.09, 1.11, p < 0.001). The score also performed well in the validation cohort with an AUC of 0.67 (95%CI: 0.65, 0.69) and the Hosmer-Lemeshow test indicating a good fit ($p = 0.17$). When the validation cohort was limited to patients enrolled in GWTG-R between January 2012 and June 2015, results were similar with an AUC of 0.65 (95%CI: 0.62, 0.69) and the Hosmer-Lemeshow goodness-of-fit test again demonstrating good fit (p=0.93).

Score results in the derivation cohort and validation cohort were divided into 7 bins and then plotted against observed in-hospital mortality as displayed in Figure 4. Given that few patients scored below 5 or above 40, these patients were included in the lowest (0–9) and highest (>34) scoring bins accordingly. As noted in the figure, we found a stepwise increase for in-hospital mortality with increasing score in both the derivation and validation cohorts. The observed average mortality ranged from <10% in the lowest score category of both cohorts to >70% in the highest category illustrating the wide range of mortality separated effectively by our scoring system.

Discussion

In the present study we developed and internally validated a well-calibrated prediction score for initial survivors of an ARC event. To our knowledge, this is the first such prognostic tool for this patient population. Given the large size and nationwide reach of the data repository used for the derivation and validation cohorts, we believe this score is generalizable to initial survivors of ARC events throughout the United States.

The existing evidence base surrounding ARC in hospitalized patients is sparse and suffers from the lack of a standard definition. Previous studies have relied heavily on diagnostic coding within the Nationwide Inpatient Sample to identify patients with acute respiratory failure (ARF). In one study, ARF was defined as a diagnostic code for acute respiratory distress/failure (ICD-9 518.5, 518.81, 518.82) in addition to a procedure code for continuous mechanical ventilation (ICD-9 96.7).⁴ That study identified \sim 330,000 annual cases of ARF in the United States. Stefan et. al., relying on a broader array of diagnostic codes (ICD-9 518.81, 518.82, 518.84, 518.4, 799.1, 786.09), estimated an annual incidence of two-million cases per year.⁹ According to that study, 20.6% of patients admitted with ARF in 2009 did not survive to hospital discharge and overall costs were estimated at \$54.3 billion.⁹ Using data extracted from the GWTG-R ARC database and the narrower definition of ARC

described in the Methods section above, our group has previously estimated an annual incidence of \sim 45,000 ARC events with an estimated mortality of \sim 40%.¹⁵

Few studies have attempted to identify factors associated with increased risk of mortality following ARC. In one study by Behrendt et al., using the Nationwide Inpatient Sample, the predictors of mortality in patients admitted with ARF were age $\,$ 80, multi-organ failure, the human immunodeficiency virus, chronic liver disease, and cancer⁴. In another study of ARF occurring in ICU patients, each additional organ failure measured in the Sequential Organ Failure Assessment (SOFA) score correlated with a stepwise increase risk of death.¹⁶ In our model, age (especially age >80 as found in the Behrendt study), hypotension in the 4 hours preceding the ARC event and need for intubation during the event were the greatest predictors of in-hospital mortality. Interestingly, whether or not an emergency response team was activated and whether or not the event occurred during daytime, weekday hours did not improve the score. Also of note, pre-event pneumonia or congestive heart failure did not independently predict in-hospital mortality.

In order to be used for prognostication by clinicians and families, it is critical that any predictive model be well calibrated. Model discrimination, while important for diagnostic testing, is not particularly useful for predicting outcomes on a patient-to-patient basis. In addition, there is generally a trade-off between calibration and discrimination.17 The predictive score presented here demonstrated good calibration. Although the c-statistic of 0.67 for our prediction score represents far better discrimination than chance alone would predict, it is below that obtained by prediction scores developed for survivors of cardiac arrest using the GWTG database.^{6,7} While we believe that the discrimination achieved by our model does not detract from its validity or utility, there is room for model improvement; possibly with inclusion of variables not captured in the GWTG-R ARC database. In particular, the present score only accounts for pre-event variables and does not include early post-event factors (e.g. organ failure). As previously noted, these may be important independent predictors of mortality.4,16

Beyond its potential clinical applications, the score presented here may be of particular use in research and quality improvement. Investigators studying critically-ill patients are increasingly recognizing the tremendous heterogeneity in baseline risk of death that exists in this population. This can lead to heterogeneity of treatment effect in randomized controlled trials.18 Pre-planned analysis stratified by baseline risk-score may help ameliorate this problem. In addition to its applications in investigation, the score can be used to improve quality within health care systems by allowing for risk-adjusted comparisons between hospitals.¹¹

To put this into context, an 82-year old hypotensive patient who experiences ARC on a medical floor without telemetry and is intubated during the event would have an ARC score of 30. This information may be highly useful to the clinician who can now provide prognostic information to the patient's family—e.g. of all patients with an ARC score of 30, approximately 65% expire in the hospital. In addition, hospitals seeking to compare outcomes amongst patients suffering ARC can look to see how their mortality rate for all patients with ARC scores of 30–34 compares to the mortality rate predicted by the ARC

score. If higher than ~65%, quality improvement initiatives may be warranted. Finally, a researcher seeking to study a novel mechanical ventilation mode for initial survivors of ARC events can identify a cohort of high-risk patients using the ARC score (e.g. ARC $>$ 30) and

The results presented here should be interpreted within the context of a number of limitations. Foremost, our prediction score was developed from data submitted to the GWTG-R ARC data repository and a number of missing data points exist. In addition, given the nature of the dataset, we were limited in those candidate variables we could select from. The possibility remains that other variables not collected in the GWTG-R data repository have more predictive power than those used here. Likewise, information regarding the cause of the ARC events and reason for death (e.g. comorbid withdrawal, progressive hypoxemia) was not available. Finally, the results presented here are derived from hospitals that submit data to the GWTG-R. Although this is a large group of over 300 hospitals, we cannot exclude the possibility that the present score would be less well calibrated at a non-GWTG participating hospital. Future studies will be needed to assess external validity.

determine whether the novel technique results in better than expected outcomes.

The ability to anticipate which hospitalized inpatients will go on to develop ARC is an important corollary to this project. Future studies in this area should focus on exploring the epidemiology of ARC events with an eye towards identifying factors that may prompt prevention or early recognition of respiratory decompensation.

Conclusions

In sum, we present a scoring system for predicting in-hospital mortality in survivors of an acute respiratory compromise event. This scoring system is simple and designed to be useful in a busy clinical setting. Through application of this score, physicians can stratify patients by risk cohort and thereby provide useful prognostic information to patients' families. In addition, the score will be of use to investigators seeking more refined patient cohorts and for quality improvement teams seeking to compare local ARC outcomes with a nationwide sample.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources:

Dr. Donnino is supported by the American Heart Association (14GRNT2001002) and the National Institute of Health (1K24HL127101). Dr. Moskowitz is supported by the National Institute of Health (2T32HL007374-37). Dr. Cocchi is supported by the American Heart Association (15SDG22420010).

The authors wish to thank Francesca Montillo, M.M., for her editorial review of the manuscript.

References

- 1. Parr MJ, Hadfield JH, Flabouris A, Bishop G, Hillman K. The Medical Emergency Team: 12 month analysis of reasons for activation, immediate outcome and not-for-resuscitation orders. Resuscitation. 2001; 50(1):39–44. [PubMed: 11719127]
- 2. Jones DA, DeVita MA, Bellomo R. Rapid-response teams. N Engl J Med. 201365:139–46.
- 3. Jaber S, Amraoui J, Lefrant JY, et al. Clinical practice and risk factors for immediate complications of endotracheal intubation in the intensive care unit: a prospective, multiple-center study. Crit Care Med. 2006; 34:2355–61. [PubMed: 16850003]
- 4. Behrendt CE. Acute respiratory failure in the United States: incidence and 31-day survival. Chest. 2000; 118:1100–5. [PubMed: 11035684]
- 5. Girotra S, Nallamothu BK, Chan PS. Using risk prediction tools in survivors of in-hospital cardiac arrest. Curr Cardiol Rep. 2014; 16(3):457. [PubMed: 24464303]
- 6. Chan PS, Spertus JA, Krumholz HM, Berg RA, Li Y, Sasson C, Nallamothu BK, Get With the Guidelines-Resuscitation Registry Investigators. A validated prediction tool for initial survivors of in-hospital cardiac arrest. Arch Intern Med. 2012; 172(12):947–53. [PubMed: 22641228]
- 7. Ebell MH, Jang W, Shen Y, Geocadin RG, Get With the Guidelines–Resuscitation Investigators. Development and validation of the Good Outcome Following Attempted Resuscitation (GO-FAR) score to predict neurologically intact survival after in-hospital cardiopulmonary resuscitation. JAMA Intern Med. 2013; 173(20):1872–8. [PubMed: 24018585]
- 8. Chan PS, Berg RA, Spertus JA, Schwamm LH, Bhatt DL, Fonarow GC, Heidenreich PA, Nallamothu BK, Tang F, Merchant RM, AHA GWTG-Resuscitation Investigators. Riskstandardizing survival for in-hospital cardiac arrest to facilitate hospital comparisons. J Am Coll Cardiol. 2013; 62(7):601–9. [PubMed: 23770167]
- 9. Stefan MS, Shieh MS, Pekow PS, et al. Epidemiology and outcomes of acute respiratory failure in the United States, 2001 to 2009: a national survey. J Hosp Med. 2013; 8:76–82. [PubMed: 23335231]
- 10. Wang HE, Abella BS, Callaway CW, American Heart Association National Registry of Cardiopulmonary Resuscitation I. Risk of cardiopulmonary arrest after acute respiratory compromise in hospitalized patients. Resuscitation. 2008; 79:234–40. [PubMed: 18692287]
- 11. Peberdy MA, Kaye W, Ornato JP, et al. Cardiopulmonary resuscitation of adults in the hospital: a report of 14720 cardiac arrests from the National Registry of Cardiopulmonary Resuscitation. Resuscitation. 2003; 58:297–308. [PubMed: 12969608]
- 12. Get With The Guidelines–Resuscitation web page. 2015. Accessed 14 September 2015, at [http://](http://www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelines-Resuscitation/Get-With-The-Guidelines-Resuscitation_UCM_314496_SubHomePage.jsp) [www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelines-Resuscitation/Get-With-](http://www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelines-Resuscitation/Get-With-The-Guidelines-Resuscitation_UCM_314496_SubHomePage.jsp)[The-Guidelines-Resuscitation_UCM_314496_SubHomePage.jsp](http://www.heart.org/HEARTORG/HealthcareResearch/GetWithTheGuidelines-Resuscitation/Get-With-The-Guidelines-Resuscitation_UCM_314496_SubHomePage.jsp)
- 13. Pan W. Akaike's information criterion in generalized estimating equations. Biometrics. 2001; 57:120–5. [PubMed: 11252586]
- 14. Hermansen SW. Evaluating Predictive Models: Computing and Interpreting the c Statistic SAS Global Forum. Data Mining and Predictive Modeling. :2008.
- 15. Andersen LW, Berg KM, Chase M, Cocchi MN, Massaro J, Donnino MW, American Heart Association's Get With The Guidelines(®)-Resuscitation Investigators. Acute respiratory compromise on inpatient wards in the United States: Incidence, outcomes, and factors associated with in-hospital mortality. Resuscitation. 2016; 30:105, 123–129.
- 16. Flaatten H, Gjerde S, Guttormsen AB, et al. Outcome after acute respiratory failure is more dependent on dysfunction in other vital organs than on the severity of the respiratory failure. Crit Care. 2003; 7:R72. [PubMed: 12930559]
- 17. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007; 115(7):928–35. [PubMed: 17309939]
- 18. Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC. Implications of Heterogeneity of Treatment Effect for Reporting and Analysis of Randomized Trials in Critical Care. Am J Respir Crit Care Med. 2015; 192(9):1045–51. [PubMed: 26177009]

Figure #2. Model Calibration

Figure #3. Score Distribution: Derivation Cohort

Figure #4. Observed Mortality by Score

Author Manuscript

Author Manuscript

Table 1

Patient and event characteristics

Rhythm at start of event

* Other refers to the following locations: post-anesthesia care unit, ambulatory/outpatient area, cardiac catheterization laboratory, delivery suite, diagnostic/intervention area, operating room, rehab, skilled nursing or mental health unit/facility, and same-day surgical area

Table 2

Final Model from the derivation cohort

* Other refers to the following locations: post-anesthesia care unit, ambulatory/outpatient area, cardiac catheterization laboratory, delivery suite, diagnostic/intervention area, operating room, rehab, skilled nursing or mental health unit/facility, and same-day surgical area