Prediction of Critical Illness During Out-of-Hospital Emergency Care

Christopher W. Seymour, MD, MSc

Jeremy M. Kahn, MD, MSc

Colin R. Cooke, MD, MSc

Timothy R. Watkins, MD

Susan R. Heckbert, MD, PhD

Thomas D. Rea, MD, MPH

Hospitals vary widely in quality of critical care.[1] Consequently, the outcomes of critically ill patients may be improved by concentrating care at more experienced centers.[1-3] By centralizing patients who are at greater risk of mortality in referral hospitals, regionalized care in critical illness may achieve improvements in outcome similar to trauma networks.[4] In 2006, the Institute of Medicine called for a regionalized, coordinated system of emergency care for high-risk patients,[5] one in which patients in most need of high-intensity acute care are distributed to centers with the greatest expertise in caring for the critically ill.

Current out-of-hospital triage of non-injured, critically ill patients uses dispatch criteria,[6] subjective emergency medical services (EMS) assessments,[7,8] coordination by medical command officers,[9] and patient preference.[10] In specific conditions such as coronary artery disease and stroke, out-of-hospital care providers use objective tools to triage and risk-stratify prehospital patients for early treatment and choice of destination.[11-13] However, these subjective and disease-specific assessments alone may not be sufficient for triage in general populations at risk of critical ill-

For editorial comment see p 797.

Context Early identification of nontrauma patients in need of critical care services in the emergency setting may improve triage decisions and facilitate regionalization of critical care.

Objectives To determine the out-of-hospital clinical predictors of critical illness and to characterize the performance of a simple score for out-of-hospital prediction of development of critical illness during hospitalization.

Design and Setting Population-based cohort study of an emergency medical services (EMS) system in greater King County, Washington (excluding metropolitan Seattle), that transports to 16 receiving facilities.

Patients Nontrauma, non–cardiac arrest adult patients transported to a hospital by King County EMS from 2002 through 2006. Eligible records with complete data (N=144 913) were linked to hospital discharge data and randomly split into development (n=87 266 [60%]) and validation (n=57 647 [40%]) cohorts.

Main Outcome Measure Development of critical illness, defined as severe sepsis, delivery of mechanical ventilation, or death during hospitalization.

Results Critical illness occurred during hospitalization in 5% of the development (n=4835) and validation (n=3121) cohorts. Multivariable predictors of critical illness included older age, lower systolic blood pressure, abnormal respiratory rate, lower Glasgow Coma Scale score, lower pulse oximetry, and nursing home residence during out-of-hospital care ($P<.01$ for all). When applying a summary critical illness prediction score to the validation cohort (range, 0-8), the area under the receiver operating characteristic curve was 0.77 (95% confidence interval [CI], 0.76-0.78), with satisfactory calibration slope (1.0). Using a score threshold of 4 or higher, sensitivity was 0.22 (95% CI, 0.20-0.23), specificity was 0.98 (95% CI, 0.98-0.98), positive likelihood ratio was 9.8 (95% CI, 8.9-10.6), and negative likelihood ratio was 0.80 (95% CI, 0.79- 0.82). A threshold of 1 or greater for critical illness improved sensitivity (0.98; 95% CI, 0.97-0.98) but reduced specificity (0.17; 95% CI, 0.17-0.17).

Conclusions In a population-based cohort, the score on a prediction rule using out-of-hospital factors was significantly associated with the development of critical illness during hospitalization. This score requires external validation in an independent population.

JAMA. 2010;304(7):747-754   www.jama.com

ness.[8,14-16] Future development of regionalized systems of acute care will require objective, routinely measured predictors that are associated with important clinical end points in a heterogeneous population. An objective triage tool may also identify patients for early treatment by out-of-hospital care providers.

We sought to develop a tool for prediction of critical illness during out-of-hospital care in noninjured, non–cardiac arrest patients. Using a population-based cohort of EMS records linked to hospital discharge data, we hypothesized that objective, out-of-hospital factors could discriminate between patients who were and

Author Affiliations are listed at the end of this article.
Corresponding Author: Christopher W. Seymour, MD, MSc, Division of Pulmonary and Critical Care Medicine, University of Washington, Harborview Medical Center, PO Box 359762, Seattle, WA 98104 (seymoc @u.washington.edu).
Caring for the Critically Ill Patient Section Editor: Derek C. Angus, MD, MPH, Contributing Editor, *JAMA* (angusdc@upmc.edu).

were not likely to develop critical illness during hospitalization.

METHODS

Study Design, Population, and Setting

We conducted a retrospective cohort study among patients who activated EMS during the 5-year period from 2002 through 2006 in King County, Washington, excluding metropolitan Seattle. King County has a heterogeneous population of 1.7 million persons residing in rural, suburban, and urban areas. Residents are served by a 2-tier EMS system accessed by calling 911. First-tier response is provided by emergency medical technician–fire fighters who provide basic life support (BLS) care. The second tier is provided by paramedics who are trained in advanced life support (ALS) and respond to more severely ill patients based on protocols and assessments by both emergency medical dispatchers and BLS responders. Patients encountered by King County EMS may be transported to 1 of 16 hospitals.

We linked EMS records to the Washington State Comprehensive Hospital Abstract Reporting System (CHARS) database from 2002 through 2007. We excluded patients with traumatic injury and cardiac arrest as determined by EMS documentation. Patients with traumatic injury are already triaged under explicit clinical criteria,[17] and patients with cardiac arrest have a near certain likelihood of requiring intensive care unit (ICU) admission.[18] Among remaining patients, we included EMS encounters that met the following 3 criteria: patient age of at least 18 years; documentation of out-of-hospital vital signs/physical examination; and transport to a receiving facility. The final sample was randomly allocated into development (60%) and validation (40%) cohorts.

Variable Definitions

We defined critical illness as severe sepsis, delivery of mechanical ventilation, or death at any point during hospitalization. We use this definition of critical illness rather than simple admission to an ICU, which can be influenced by emergency department disposition, ICU bed availability, and local practice variation. We used a clinically validated, administrative definition for severe sepsis based on *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* diagnosis codes for a bacterial or fungal infectious process and the presence of acute organ dysfunction.[19] Because of changes in coding for sepsis, we also included *ICD-9-CM* codes 995.91 (systemic inflammatory response syndrome due to infectious process without acute organ dysfunction), 995.92 (severe sepsis), and 785.52 (shock without mention of trauma; septic shock) in our definition. We used the *ICD-9-CM* procedure code (96.7x) to define the need for mechanical ventilation.[20] We defined hospital death using discharge disposition in CHARS. We abstracted out-of-hospital clinical data from the King County EMS database, including dispatch, demographic, physical examination, procedure, and transport data. We evaluated only initial out-of-hospital vital signs, documented by first-arriving EMS personnel.

Record Linkage

In the King County EMS database, each patient has a BLS record, yet additional responders (both BLS and ALS) may create duplicate records for the same patient incident. No unique identifier is present across these duplicate records. To identify the first responder, we used probabilistic matching to link ALS encounters (n = 106 694) and BLS encounters (n = 436 159) that represent the same patient (LinkPlus software, version 2.0, Centers for Disease Control and Prevention Cancer Division, Atlanta, Georgia).[21,22] We successfully matched 90 191 ALS records (85%) to corresponding BLS records and removed unmatched ALS records (15%), which are likely to represent redundant ALS responders. Two King County epidemiologists assessed probabilistic match quality through manual review of 200 ALS-BLS pairs, blinded to probabilistic match out-come, of which 192 (96%) were correctly matched. We then applied exclusion criteria to the final matched data set (n = 436 159), then linked eligible encounters (n = 166 908) using direct identifiers (first/last name, age, sex, receiving hospital, transport/admission date) to hospital discharge data using a hierarchical, deterministic matching algorithm.

Model Development

We developed a multivariable model for critical illness during hospitalization in 4 steps: (1) assessment of candidate variable quality and categorization of continuous predictors; (2) construction of a parsimonious model; (3) development of a point score; and (4) internal validation in a separate cohort of patients.[23,24] When choosing candidate variables for the model, we considered clinical relevance, generalizability (inclusion in the National EMS Information System database),[25] and timing of the exposure during out-of-hospital care, in that order. Continuous candidate predictors were categorized using a priori–determined cut points based on clinical relevance and natural distributions. We considered age (<45, 45-64, or ≥65 years), sex, initial systolic blood pressure (≤90, 91-140, 141-180, or >180 mm Hg), initial heart rate (≤60, 61-99, 100-119, or ≥120/min), initial respiratory rate (<12, 12-23, 24-35, or ≥36/min), initial Glasgow Coma Scale score (15, 12-14, 8-11, or <8), initial pulse oximetry (93%-100%, 88%-92%, 80%-87%, or <80%), and out-of-hospital location (nursing home, adult medical facility, home, public building, or street/highway). For simplicity, we did not consider multiplicative or additive interactions as candidate predictors or out-of-hospital procedures that may occur after initial EMS assessments. In our primary analysis, we used single imputation with normal value substitution for variables presumed to be clinically normal if not measured (eg, Glasgow Coma Scale, pulse oximetry), a method previously used for critically ill patients.[26] We constructed a multivariable logis-

tic regression model from candidate predictors and used backward selection with the Akaike information criterion to populate the model.[27] This measure of model fit penalizes models with a large number of variables and attempts to reduce overfitting.

We assigned a point score to each covariate by rounding regression coefficients in the final model to the nearest integer.[28] We generated predicted probabilities of critical illness for each value of the point score using logistic regression with the Huber-White estimator to generate standard errors for regression coefficients. In the validation cohort, we determined overall performance using the Brier score[29] and the McKelvey and Zavoina R^2 statistic.[30] We assessed discrimination using the area under the receiver operating characteristic curve for the composite outcome, as well as each contributing component (death, mechanical ventilation, and severe sepsis).[31] We assessed calibration using the Hosmer-Lemeshow statistic with $P < .10$ indicating that fit was inadequate.[32] Since small, clinically insignificant differences in predicted and observed outcomes result in a significant Hosmer-Lemeshow statistic in large samples,[33] we also calculated the calibration slope.[34] Further details of model performance measures are available in the eAppendix; available at http://www.jama.com. We calculated the sensitivity, specificity, and positive and negative likelihood ratios for each point score as a threshold and grouped patients into 3 categories by risk of critical illness: low ($<10\%$), intermediate (10%-20%), and high ($>20\%$).

We performed several sensitivity analyses to assess the robustness of our findings. To determine if our handling of missing data introduced bias,[35] we assumed data were missing at random and performed multiple imputation for all missing values using a regression switching approach (multiple imputation by chained equations).[36] We then repeated both model development and validation steps on imputed data. To determine the performance of a more flexible but complex model, we

Figure 1. Participant Accrual
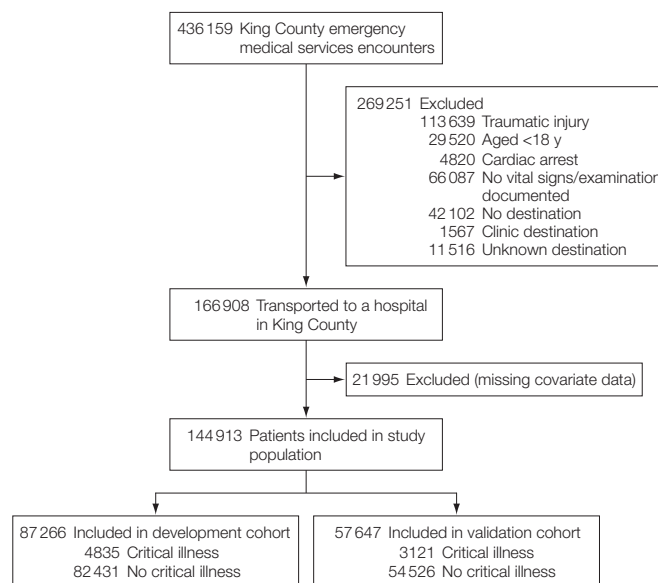
Missing covariate data do not include participants who were missing Glasgow Coma Scale or pulse oximetry data.

reanalyzed model fit after rounding regression coefficients to the nearest half integer. Finally, we repeated our analysis using a different definition of critical illness that included diagnostic and procedure codes for cardiac and respiratory arrest, hypotension, shock, cardiopulmonary resuscitation, and acute respiratory failure.[37] Additional details of sensitivity analyses are provided in the eAppendix. All analyses were performed with Stata software, version 10.0 (Stata Corp, College Station, Texas). All tests of significance used a 2-sided $P \le .05$. This study was approved by the institutional review boards for the Washington State Department of Health, King County Emergency Medical Services, and University of Washington.

RESULTS

Among the 436 159 unique prehospital patients (**FIGURE 1**), the primary reasons for exclusion were traumatic injury (26%), no transport to a hospital (14%), and absence of physical examination by EMS responders (15%). We evaluated 166 908 eligible patients for data quality of candidate predictors (eTable 1) and split complete cases

(n=144 913) into development (n=87 266) and validation (n=57 647) cohorts. Critical illness during hospitalization occurred in approximately 5% of both the development (n=4835) and validation (n=3121) cohorts. Patients experiencing critical illness were older and more likely to be transported from nursing homes, receive ALS care, and present with out-of-hospital respiratory symptoms (**TABLE 1**). In general, patients with critical illness during hospitalization presented with greater alterations in initial out-of-hospital vital signs. Among patients with and without critical illness, EMS response and transport intervals were similar, while total time from patient side to departure was greater for both BLS and ALS personnel responding to patients who developed critical illness.

In the development cohort, we identified 8 objective out-of-hospital predictors of critical illness in our final model (**TABLE 2**). We used point values generated from the rounded regression coefficients to develop a score. The coefficients for sex and nursing home residence rounded to 0 and did not contribute to our final score (eTable 2). We entered the point total for each patient

Table 1. Participant, Incident, and Outcome Data for the Development and Validation Cohorts[a]

	Development		Validation	
	Critical Illness (n = 4835)	No Critical Illness (n = 82 431)	Critical Illness (n = 3121)	No Critical Illness (n = 54 526)
Age, mean (SD), y	69 (17)	59 (21)	69 (17)	60 (21)
Male	2516 (46)	38 479 (40)	1597 (45)	25 409 (40)
Out-of-hospital emergency location				
Routine/home	3319 (67)	55 045 (67)	2062 (66)	36 289 (67)
Public building	90 (2)	5316 (6)	54 (2)	3907 (7)
Street/highway	68 (2)	3520 (4)	61 (2)	2404 (4)
Adult care home	189 (4)	2343 (3)	150 (5)	1515 (3)
Medical facility	107 (2)	3066 (4)	68 (2)	1992 (4)
Nursing home	926 (19)	6048 (7)	654 (21)	3907 (7)
Other	117 (2)	6606 (8)	64 (2)	4557 (8)
Received ALS care	3085 (64)	29 556 (36)	1975 (63)	19 375 (35)
EMS severity code[b]				
Life-threatening	1044 (22)	2118 (3)	611 (20)	1334 (2)
Urgent	2428 (50)	38 564 (47)	1606 (52)	25 495 (47)
Nonurgent	1312 (27)	40 639 (49)	873 (28)	26 960 (49)
Initial out-of-hospital vital signs				
Systolic blood pressure, mean (SD), mm Hg	130 (42)	140 (33)	129 (41)	140 (33)
Diastolic blood pressure, mean (SD), mm Hg	74 (23)	81 (18)	74 (22)	81 (18)
Heart rate, mean (SD), beats/min	99 (27)	90 (23)	98 (27)	91 (23)
Respiratory rate, mean (SD), breaths/min	24 (10)	19 (6)	24 (10)	19 (6)
Glasgow Coma Scale score, median (IQR)	15 (9-15)	15 (15-15)	15 (9-15)	15 (15-15)
Oxygen saturation, mean (SD), %	89 (11)	95 (7)	90 (11)	95 (7)
Diagnostic category[c]				
Respiratory	1362 (28)	8872 (11)	869 (28)	5786 (11)
Neurological	770 (16)	12 342 (15)	467 (15)	8225 (15)
Cardiovascular	580 (12)	15 310 (19)	385 (12)	9894 (18)
Abdominal	336 (7)	8029 (10)	257 (8)	5402 (10)
Alcohol-/drug-related	319 (7)	4806 (6)	211 (7)	3131 (6)
Metabolic/endocrine	194 (4)	2560 (3)	128 (4)	1581 (3)
Psychiatric	47 (1)	3399 (4)	16 (1)	2375 (4)
Fall	26 (1)	821 (1)	26 (1)	532 (1)
Anaphylaxis	11 (<1)	897 (1)	3 (<1)	567 (1)
Obstetric/gynecologic	8 (<1)	1353 (2)	4 (<1)	967 (2)
Medical (other)	870 (18)	16 411 (20)	549 (18)	10 976 (20)
Unknown	264 (6)	5725 (7)	177 (5)	3755 (7)
Other	33 (1)	1518 (2)	16 (1)	1032 (2)
BLS response intervals, median (IQR), min				
Call receipt to unit notification	0.6 (0.2-1.0)	0.6 (0.1-1.0)	0.6 (0.2-1.0)	0.6 (0.1-1.0)
Unit notification to responding	1.6 (1.0-2.0)	1.6 (1.0-2.0)	1.6 (1.0-2.0)	1.6 (1.0-2.0)
Responding to arrival at scene	3.0 (2.0-4.2)	3.0 (2.0-4.3)	3.0 (2.0-4.1)	3.0 (2.0-4.3)
Patient side to departure	33 (24-45)	25 (19-33)	33 (24-45)	25 (19-33)
ALS response intervals, median (IQR), min				
Call receipt to unit notification	2 (1.0-8)	1.8 (1.0-6)	2.0 (1.0-8.8)	1.8 (1.0-6.0)
Unit notification to responding	1.6 (1.0-2.0)	1.6 (1.0-2.0)	1.6 (1.0-2.0)	1.6 (1.0-2.0)
Responding to arrival at scene	6.0 (4.0-8.0)	6.0 (4.0-8.0)	6.0 (4.0-8.0)	6.0 (4.0-8.0)
Patient side to departure	26 (20-34)	21 (16-27)	26 (19-35)	21 (16-27)
Scene–to–hospital destination time, median (IQR), min	9 (6.0-15)	11 (7.0-16)	10 (6.0-14)	11 (7.0-16)
Prehospital interventions				
Intubation	1107 (23)	292 (<1)	669 (21)	211 (<1)
Bag/valve/mask ventilation	833 (17)	557 (1)	534 (17)	347 (1)
Supplemental oxygen	3840 (79)	47 817 (58)	2506 (80)	31 315 (57)
Electrocardiographic monitoring	2838 (59)	27 656 (34)	1858 (60)	18 048 (33)
Peripheral or central intravenous access	2182 (45)	14 968 (18)	1382 (44)	9531 (17)
Mode of transport from scene				
County ambulance	3125 (65)	35 772 (44)	2032 (65)	23 520 (43)
Helicopter/rotary wing	32 (1)	89 (<1)	17 (1)	62 (<1)
Private ambulance	1535 (33)	41 053 (51)	973 (32)	27 185 (52)
Private automobile	31 (1)	3500 (4)	22 (1)	2416 (4)

(continued)

Table 1. Participant, Incident, and Outcome Data for the Development and Validation Cohorts[a] (continued)

	Development		Validation	
	Critical Illness (n = 4835)	No Critical Illness (n = 82 431)	Critical Illness (n = 3121)	No Critical Illness (n = 54 526)
Components of critical illness during hospitalization[d]				
In-hospital mortality	1718 (35)		1108 (36)	
Severe sepsis	2988 (61)		1907 (61)	
Need for mechanical ventilation	2331 (48)		1495 (47)	
Hospital length of stay, median (IQR), d	5 (2-9)	3 (2-5)	5 (2-9)	3 (2-5)

Abbreviations: ALS, advanced life support; BLS, basic life support; EMS, emergency medical services; IQR, interquartile range.
[a] Data are expressed as No. (%) of participants unless otherwise indicated.
[b] The EMS severity code was determined by the on-scene responder.
[c] Diagnostic category was determined by EMS impression.
[d] Primary outcome components are not mutually exclusive.

in a logistic regression model to generate the individual predicted probability of critical illness. For scores ranging from 0 to 8, the mean predicted probabilities of critical illness were 1.2% (95% confidence interval [CI], 1.0%-1.4%), 2.9% (95% CI, 2.7%-3.0%), 6.7% (95% CI, 6.4%-7.1%), 15% (14%-16%), 30% (29%-33%), 52% (47%-56%), 72% (64%-79%), 88% (69%-98%), and 100% (64%-100%), respectively.

In our independent validation sample, the critical illness score demonstrated satisfactory discrimination (area under the receiver operating characteristic curve, 0.77; 95% CI, 0.76-0.78; Brier score, 0.04). The critical illness score had similar performance for each component end point of our primary outcome (hospital mortality, 0.78 [95% CI, 0.77-0.79]; severe sepsis, 0.76 [95% CI, 0.75-0.77]; delivery of mechanical ventilation during hospitalization, 0.81 [95% CI, 0.80-0.82]). Calibration of the model was acceptable at low and intermediate score values but decreased at higher score values (**FIGURE 2**). The Hosmer-Lemeshow goodness-of-fit test demonstrated statistical evidence of inadequate fit ($\chi^2_7=47$; $P<.001$), but the calibration slope (1.0) suggested little overfitting. We calculated sensitivity, specificity, and diagnostic likelihood ratios for each critical illness score grouped by low-, intermediate-, and high-risk categories (**TABLE 3**). Using a cut point of 4 or higher to identify patients who develop critical illness (>20% expected risk), we observed a sensitivity of 0.22 (95% CI, 0.20-0.23), a specificity of 0.98 (95% CI, 0.98-0.98), a positive likeli-

hood ratio of 9.8 (95% CI, 8.9-10.6), and a negative likelihood ratio of 0.8 (95% CI, 0.79-0.82). If this cut point were used to triage patients to a regional referral center for critical care services, we would transport 1887 patients (3.2%) to the regional center, of whom 1211 (64%) would be mistriaged (ie, brought to the regional center but not subsequently develop critical illness). Of 55 760 patients (97%) brought to nonreferral centers, 2445 (4.4%) would be mistriaged (ie, brought to a nonreferral center but subsequently develop critical illness). Using a threshold of 1 or greater to identify critical illness (>2% expected risk) would increase sensitivity to 0.98 (96% CI, 0.97-0.98) but decrease specificity to 0.17 (95% CI, 0.17-0.17). Triage using this cut point would result in the transport of 48 286 patients (84% of the total) to the regional center, of whom 45 231 (94%) would be mistriaged. Of 9361 patients (16%) brought to nonreferral centers, 66 (1%) would be mistriaged.

In our sensitivity analysis, multivariable logistic regression estimates derived after multiple imputation were similar to our primary analysis in the development cohort (eTable 3). Only a point score between 12 and 14 on the Glasgow Coma Scale decreased from 1 to 0 after rounding. Performance of the imputed model in imputed validation data sets was also similar (eTable 4). We observed comparable results when we derived a point score by assigning half-integers to regression coefficients and used an alternative definition of critical illness (eTable 4).[37]

Table 2. Regression Coefficients in the Final Multivariable Logistic Regression Model for Critical Illness During Hospitalization in the Development Cohort (n = 87 266)

Predictor[a]	Regression Coefficient, β (95% CI)	Score[b]
Male	0.22 (0.15 to 0.28)	0
Age, y		
<45	0 [Reference]	0
45-64	0.91 (0.80 to 1.02)	1
≥65	1.32 (1.22 to 1.43)	1
Respiratory rate, breaths/min		
<12	1.35 (1.22 to 1.43)	1
12-23	0 [Reference]	0
24-35	0.79 (0.72 to 0.86)	1
≥36	1.54 (1.43 to 1.64)	2
Systolic blood pressure, mm Hg		
≤90	0.92 (0.82 to 1.0)	1
91-140	0 [Reference]	0
141-180	−0.37 (−0.45 to −0.30)	0
>180	−0.11 (−0.22 to −0.01)	0
Heart rate, beats/min		
≤60	0.09 (−0.03 to 0.21)	0
61-99	0 [Reference]	0
100-119	0.44 (0.36 to 0.52)	0
≥120	0.77 (0.68 to 0.85)	1
Pulse oximetry, %		
≥93	0 [Reference]	0
88-92	0.43 (0.24 to 0.61)	0
80-87	0.83 (0.61 to 1.04)	1
<80	1.08 (0.82 to 1.35)	1
Glasgow Coma Scale score		
15	0 [Reference]	0
12-14	0.51 (0.38 to 0.63)	1
8-11	1.24 (1.10 to 1.39)	1
<8	1.96 (1.81 to 2.10)	2
Nursing home location	0.46 (0.36 to 0.54)	0

Abbreviation: CI, confidence interval.
[a] All vital sign data are initial out-of-hospital measurements documented by first-arriving emergency medical services responder.
[b] Total score ranges from 0 to 8. Scores are derived by rounding the regression coefficient to the nearest integer.

COMMENT

We developed and internally validated an out-of-hospital model that predicts critical illness during hospitalization in a heterogeneous medical population. The critical illness score incorporates a small number of objective out-of-hospital variables, such as systolic blood pressure, heart rate, respiratory rate, Glasgow Coma Scale score, and pulse oximetry. We demonstrate the role that simple physiologic assessment can play in risk stratification in the prehospital period among noninjured patients. The model provides an important foundation for future efforts to identify patients at greatest risk of critical illness using information from the out-of-hospital phase of emergency care.

We developed the current model in direct response to recent calls for centralized, coordinated care in patients with acute illness.[5,38] A system that regionalizes patients with critical illness to centers with greater hospital admission volume has the potential to significantly improve outcomes but is generally not feasible at this time.[39] One major challenge to an optimal regionalized system is identifying which patients will benefit from admission to a regional referral center.[4,40] This challenge is particularly salient in the out-of-hospital setting, where variable utilization of EMS,[41] heterogeneous reasons for dispatch,[42] incomplete information,[43] and subjective assessments[8] may limit accurate identification of critical illness. Disease-specific tools such as the prehospital 12-lead electrocardiogram[44] and Los Angeles or Cincinnati stroke scales[7,45] or trauma triage guidelines[46] facilitate complex management decisions. Yet they are limited to their disease context and may lead to multiple, overlapping systems with less certain benefits to overall system efficiency. Objective tools that apply to a heterogeneous, noninjured population may improve discrimination of critical illness and potentially better match early treatment with patient needs.

Figure 2. Observed vs Expected Probability of Critical Illness Across Scores in the Development and Validation Cohorts
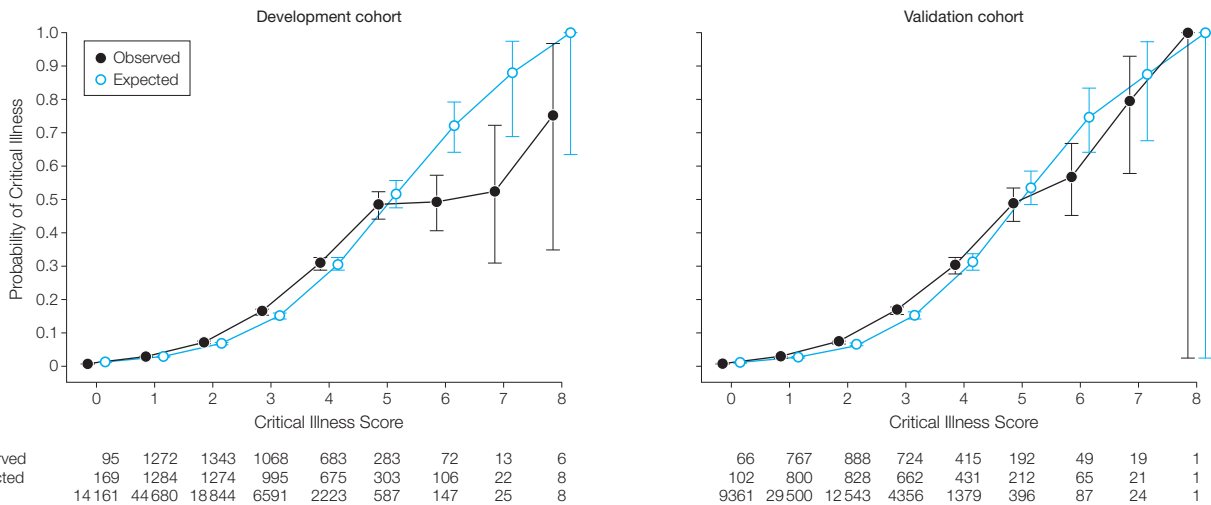
Error bars represent 95% confidence intervals (CIs). Data points are staggered for clarity. If patients with a critical illness score of 6 or higher are collapsed into a single group, the probability of critical illness is 0.76 (95% CI, 0.69-0.82) in the development cohort and 0.78 (95% CI, 0.69-0.85) in the validation cohort.

Table 3. Operating Characteristics for Each Threshold of the Critical Illness Score[a]

	Low Risk		Moderate Risk	High Risk			
	≥1	≥2	≥3	≥4	≥5	≥6	≥7
Patients							
True positive	3055	2288	1400	676	261	69	20
False positive	45 231	16 498	4843	1211	247	43	5
True negative	9295	38 028	49 683	53 315	54 279	54 483	54 521
False negative	66	833	1721	2445	2860	3052	3101
Operating characteristics							
Sensitivity	0.98 (0.97-0.98)	0.73 (0.72-0.75)	0.45 (0.43-0.47)	0.22 (0.20-0.23)	0.08 (0.07-0.09)	0.02 (0.02-0.03)	0.01 (0.0-0.1)
Specificity	0.17 (0.17-0.17)	0.70 (0.69-0.70)	0.91 (0.91-0.91)	0.98 (0.98-0.98)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Positive likelihood ratio	1.18 (1.17-1.19)	2.4 (2.36-2.48)	5.1 (4.8-5.3)	9.8 (8.9-10.6)	18.5 (16-22)	28 (19-41)	70 (26-186)
Negative likelihood ratio	0.12 (0.1-0.16)	0.38 (0.36-0.4)	0.61 (0.59-0.63)	0.80 (0.79-0.82)	0.92 (0.91-0.93)	0.98 (0.97-0.98)	0.99 (0.99-0.99)

[a]Data not shown for cutoffs ≥0 and ≥8.

The current model demonstrated good discriminative capacity. Yet the model tended to overidentify critical illness among those judged at high risk and underidentify critical illness among those judged at low risk. These errors in calibration have important implications for the development of regionalized systems of emergency care. Concerns about overwhelming referral centers are a major barrier to regionalization.[47] Should the current tool be applied, a significant number of patients would be unnecessarily brought to regional referral centers. Likewise, many patients who subsequently develop critical illness would be brought to hospitals with fewer resources to manage critical illness, reducing the clinical effect of a regionalized system. More accurate models are needed before regionalization can be practically implemented. Although the current investigation provides a necessary framework for prehospital triage, future research should be directed to develop accurate, feasible predictive models, perhaps incorporating biomarker measurement (eg, prehospital lactate level).[48] Because complex prediction models are difficult to implement during emergency care,[49] information technology solutions may be needed to integrate real-time physiologic data, biomarkers, and clinical decision making. Additionally, novel collection methods may be required to assess some variables in our model.

The current model requires prospective, external validation. With future linking of national EMS databases to hospital outcomes,[25] we recommend validation in temporally and geographically distinct populations using contemporary data. Since we used objective measures that are relatively easy to assess and included first vital sign measurements, we believe the model will perform well in other EMS system structures (eg, 1-tier systems). Model variables would also be unaffected by the presence of physicians during out-of-hospital care[50] or the longer transport times found in rural settings.[51] In addition to prospective validation, ran-domized trials are needed to evaluate the clinical and economic effect of the triage tool in a real-world setting. Even triage tools that identify critical illness with great accuracy might have unintended, adverse consequences when used in regionalized system. Large, population-based trials that capture a broad range of outcomes for patients triaged to both regional referral and community hospitals are necessary.

We recognize several limitations of our study. We defined critical illness using a composite outcome derived from administrative diagnosis and procedure codes at any time during hospitalization. Although our definition encompasses a broad range of severely ill patients at high risk of death, it may misclassify some patients who either do or do not truly require critical care. For example, some low-risk sepsis patients may be included in our definition while some high-risk patients that do not meet our definition of severe sepsis may not be included. Missing data for Glasgow Coma Scale and pulse oximetry was common, either due to absent documentation or lack of assessment. We used simple imputation with normal value substitution, a method commonly used in ICU severity scores.[26] Because this method may introduce bias, we performed a sensitivity analysis using multiple imputation that showed minimal change in our model or performance in the validation cohort. We also observed statistically significant evidence of inadequate fit using the Hosmer-Lemeshow statistic, and differences between observed and expected probabilities of critical illness at greater score values.[32] Yet the calibration slope analysis confirmed that overfitting was unlikely, and our plot of observed vs expected probabilities demonstrated adequate calibration for low and intermediate scores. We did not evaluate important predictors of critical illness, such as race/ethnicity[52] and individual-level socioeconomic status,[53] because these were not documented in our data source and could not be objectively assessed at the scene. Finally, we catego-rized continuous covariates to facilitate usability in future studies, a step that adds a conservative bias to our performance estimates. All predictive models must trade off accuracy for simplicity, and in this case, we favored a simple model for more practical application in the field.

In summary, we developed a prediction rule using best available out-of-hospital data, and the score on the prediction rule was significantly associated with development of critical illness during hospitalization in noninjured patients. Although improved accuracy and external validation are required, this model provides a foundation for future efforts to identify noninjured patients who may benefit from coordinated systems that regionalize emergency care.

Author Affiliations: Division of Pulmonary and Critical Care Medicine, Harborview Medical Center (Drs Seymour and Watkins), Research Division, Puget Sound Blood Center (Dr Watkins), and Department of Epidemiology (Dr Heckbert) and King County Medic One, Division of General Internal Medicine (Dr Rea), University of Washington, Seattle; Division of Pulmonary, Allergy, and Critical Care, Leonard Davis Institute for Health Economics and Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, University of Pennsylvania Medical Center, Philadelphia (Dr Kahn); and Division of Pulmonary and Critical Care Medicine and Robert Wood Johnson Clinical Scholar Program, University of Michigan, Ann Arbor (Dr Cooke). Dr Kahn is now with the Clinical Research, Investigation, and Systems Modeling of Acute Illness Laboratory, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, and the Department of Health Policy and Management, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania.

REFERENCES

1. Glance LG, Li Y, Osler TM, Dick A, Mukamel DB. Impact of patient volume on the mortality rate of adult intensive care unit patients. *Crit Care Med*. 2006; 34(7):1925-1934.
2. Kahn JM, Goss CH, Heagerty PJ, Kramer AA, O'Brien CR, Rubenfeld GD. Hospital volume and the outcomes of mechanical ventilation. *N Engl J Med*. 2006; 355(1):41-50.
3. Durairaj L, Torner JC, Chrischilles EA, Vaughan Sarrazin MS, Yankey J, Rosenthal GE. Hospital volume-outcome relationships among medical admissions to ICUs. *Chest*. 2005;128(3):1682-1689.
4. Kahn JM, Branas CC, Schwab CW, Asch DA. Regionalization of medical critical care: what can we learn from the trauma experience? *Crit Care Med*. 2008; 36(11):3085-3088.
5. Institute of Medicine Committee on the Future of Emergency Care in the United States Health System. *Hospital Based Emergency Care: At the Breaking Point.* Washington, DC: National Academies Press; 2007.
6. Sporer KA, Johnson NJ, Yeh CC, Youngblood GM. Can emergency medical dispatch codes predict prehospital interventions for common 9-1-1 call types? *Prehosp Emerg Care*. 2008;12(4):470-478.
7. Kidwell CS, Starkman S, Eckstein M, Weems K, Saver JL. Identifying stroke in the field. Prospective validation of the Los Angeles prehospital stroke screen (LAPSS). *Stroke*. 2000;31(1):71-76.
8. Brown LH, Hubble MW, Cone DC, et al. Paramedic determinations of medical necessity: a meta-analysis. *Prehosp Emerg Care*. 2009;13(4): 516-527.
9. Studnek JR, Fernandez AR, Margolis GS, O'Connor RE. Physician medical oversight in emergency medical services: where are we? *Prehosp Emerg Care*. 2009; 13(1):53-58.
10. Yarris LM, Moreno R, Schmidt TA, Adams AL, Brooks HS. Reasons why patients choose an ambulance and willingness to consider alternatives. *Acad Emerg Med*. 2006;13(4):401-405.
11. Brown JP, Mahmud E, Dunford JV, Ben-Yehuda O. Effect of prehospital 12-lead electrocardiogram on activation of the cardiac catheterization laboratory and door-to-balloon time in ST-segment elevation acute myocardial infarction. *Am J Cardiol*. 2008;101 (2):158-161.
12. Sivagangabalan G, Ong AT, Narayan A, et al. Effect of prehospital triage on revascularization times, left ventricular function, and survival in patients with ST-elevation myocardial infarction. *Am J Cardiol*. 2009; 103(7):907-912.
13. Brooks SC, Allan KS, Welsford M, Verbeek PR, Arntz HR, Morrison LJ. Prehospital triage and direct transport of patients with ST-elevation myocardial infarction to primary percutaneous coronary intervention centres: a systematic review and meta-analysis. *CJEM*. 2009;11(5):481-492.
14. Feldman MJ, Verbeek PR, Lyons DG, Chad SJ, Craig AM, Schwartz B. Comparison of the medical priority dispatch system to an out-of-hospital patient acuity score. *Acad Emerg Med*. 2006;13(9):954-960.

15. Levine SD, Colwell CB, Pons PT, Gravitz C, Haukoos JS, McVaney KE. How well do paramedics predict admission to the hospital? a prospective study. *J Emerg Med*. 2006;31(1):1-5.
16. Price TG, Hooker EA, Neubauer J. Prehospital provider prediction of emergency department disposition: implications for selective diversion. *Prehosp Emerg Care*. 2005;9(3):322-325.
17. Mackersie RC. History of trauma field triage development and the American College of Surgeons criteria. *Prehosp Emerg Care*. 2006;10(3):287-294.
18. Langhelle A, Tyvold SS, Lexow K, Hapnes SA, Sunde K, Steen PA. In-hospital factors associated with improved outcome after out-of-hospital cardiac arrest: a comparison between four regions in Norway. *Resuscitation*. 2003;56(3):247-263.
19. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Crit Care Med*. 2001;29(7):1303-1310.
20. Quan H, Parsons GA, Ghali WA. Validity of procedure codes in *International Classification of Diseases, 9th Revision, Clinical Modification* administrative data. *Med Care*. 2004;42(8):801-809.
21. Dean JM, Vernon DD, Cook L, Nechodom P, Reading J, Suruda A. Probabilistic linkage of computerized ambulance and inpatient hospital discharge records: a potential tool for evaluation of emergency medical services. *Ann Emerg Med*. 2001;37(6): 616-626.
22. Fellegi IP, Sunter AB. A theory for record linkage. *J Am Stat Assoc*. 1969;64:1193-1210.
23. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009;338:b605.
24. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. 2009;338:b604.
25. Dawson DE. National Emergency Medical Services Information System (NEMSIS). *Prehosp Emerg Care*. 2006;10(3):314-316.
26. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE—Acute Physiology and Chronic Health Evaluation: a physiologically based classification system. *Crit Care Med*. 1981;9(8):591-597.
27. Akaike H. Information theory and an extension of the maximum likelihood principle. Paper presented at: 2nd International Symposium on Information Theory; September 2-8, 1971; Tsahkadsor, Armenia.
28. Moons KG, Harrell FE, Steyerberg EW. Should scoring rules be based on odds ratios or regression coefficients? *J Clin Epidemiol*. 2002;55(10):1054-1055.
29. Brier GW. Verification of forecasts expressed in terms of probability. *Mon Weather Rev*. 1950; 78:1-3.
30. Mckelvey R, Zavoina W. A statistical model for the analysis of ordinal level dependent variables. *J Math Sociol*. 1975;4:103-120.
31. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36.
32. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982; 115(1):92-106.
33. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. *Crit Care Med*. 2007; 35(9):2052-2056.
34. Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. *Med Decis Making*. 1993;13(1):49-58.
35. Marshall A, Altman DG, Royston P, Holder RL.

Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. *BMC Med Res Methodol*. 2010; 10(1):7.
36. Royston P. Multiple imputation of missing values. *Stata J*. 2004;4(3):227-241.
37. Ehlenbach WJ, Hough CL, Crane PK, et al. Association between acute care and critical illness hospitalization and cognitive function in older adults. *JAMA*. 2010;303(8):763-770.
38. Barnato AE, Kahn JM, Rubenfeld GD, et al. Prioritizing the organization and management of intensive care services in the United States: the PROMIS conference. *Crit Care Med*. 2007;35(4):1003-1011.
39. Kahn JM, Linde-Zwirble WT, Wunsch H, et al. Potential value of regionalized intensive care for mechanically ventilated medical patients. *Am J Respir Crit Care Med*. 2008;177(3):285-291.
40. Carr BG, Conway PH, Meisel ZF, Steiner CA, Clancy C. Defining the emergency care sensitive condition: a health policy research agenda in emergency medicine. *Ann Emerg Med*. 2010;56(1):49-51.
41. Wang HE, Weaver MD, Shapiro NI, Yealy DM. Opportunities for emergency medical services care of sepsis. *Resuscitation*. 2010;81(2):193-197.
42. Neely KW, Eldurkar J, Drake ME. Can current EMS dispatch protocols identify layperson-reported sentinel conditions? *Prehosp Emerg Care*. 2000;4(3): 238-244.
43. Newgard CD. The validity of using multiple imputation for missing out-of-hospital data in a state trauma registry. *Acad Emerg Med*. 2006;13(3): 314-324.
44. Weaver WD, Eisenberg MS, Martin JS, et al. Myocardial Infarction Triage and Intervention Project—phase I: patient characteristics and feasibility of prehospital initiation of thrombolytic therapy. *J Am Coll Cardiol*. 1990;15(5):925-931.
45. Kothari R, Barsan W, Brott T, Broderick J, Ashbrock S. Frequency and accuracy of prehospital diagnosis of acute stroke. *Stroke*. 1995;26(6):937-941.
46. Sasser SM, Hunt RC, Sullivent EE, et al; National Expert Panel on Field Triage, Centers for Disease Control and Prevention (CDC). Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage. *MMWR Recomm Rep*. 2009;58(RR-1):1-35.
47. Kahn JM, Asch RJ, Iwashyna TJ, et al. Physician attitudes toward regionalization of adult critical care: a national survey. *Crit Care Med*. 2009;37(7):2149-2154.
48. Pearse RM. Extending the role of lactate measurement into the prehospital environment. *Crit Care*. 2009;13(1):115.
49. Sasson C, Hegg AJ, Macy M, Park A, Kellermann A, McNally B; CARES Surveillance Group. Prehospital termination of resuscitation in cases of refractory out-of-hospital cardiac arrest. *JAMA*. 2008; 300(12):1432-1438.
50. Huemer G, Pernerstorfer T, Mauritz W. Prehospital emergency medicine services in Europe: structure and equipment. *Eur J Emerg Med*. 1994;1 (2):62-68.
51. Grossman DC, Kim A, Macdonald SC, Klein P, Copass MK, Maier RV. Urban-rural differences in prehospital care of major trauma. *J Trauma*. 1997; 42(4):723-729.
52. Barnato AE, Alexander SL, Linde-Zwirble WT, Angus DC. Racial variation in the incidence, care, and outcomes of severe sepsis: analysis of population, patient, and hospital characteristics. *Am J Respir Crit Care Med*. 2008;177(3):279-284.
53. Rehkopf DH, Berkman LF, Coull B, Krieger N. The non-linear risk of mortality by income level in a healthy population: US National Health and Nutrition Examination Survey mortality follow-up cohort, 1988-2001. *BMC Public Health*. 2008;8:383.