Validation of a 'Usual Care' Model for Vasopressor Initiation in a Cohort of Emergency Department Patients with Sepsis

Baturay Aydemir, Varesh Prasad, James C. Lynch, Brett Biebelberg, Iain Kehoe, Andrew T. Reisner, Thomas Heldt, *Senior Member, IEEE*

Abstract— **Usual care regarding vasopressor (VP) initiation is ill-defined. We aimed to further validate a quantitative model for usual care in the Emergency Department (ED) regarding the timing of VP initiation in sepsis. We retrospectively studied a cohort of adult critically-ill ED patients who also received antibiotics in the ED. We applied a multivariable model previously developed from another patient cohort which distinguishes between time points at which patients were or were not subsequently started on a continuous VP infusion. The model has six independently significant predictors (respiratory rate, Glasgow Coma Scale score, systolic blood pressure, SpO_2, administered intravenous fluids, and elapsed time). The outcome was initiation of VP infusion, either within the ED or within 6 hours after leaving the ED. We applied the model to all time points, beginning when all model input parameters were first available for a given patient, and ending when either VP were first started, or the patient left the ED. Out of 55,963 adult ED patients during the two-year study interval, we identified 1,629 who met our inclusion criteria. The area under the receiver operating characteristic curve was 0.81 for all patients, and 0.72 for the subset with at least one hypotensive blood pressure measurement. At a model threshold with sensitivity and specificity 0.74 and 0.74, respectively, the median advance detection time was 170.5 minutes (IQR 53 – 363).**

Clinical Relevance— **This study establishes that the standard of care for vasopressor initiation in sepsis patients can be modelled with only a few parameters.**

I. Introduction

Sepsis and septic shock are acutely life-threatening conditions that are thought to contribute to one to two of every three hospital deaths in the United States. Recognition of sepsis is complicated by the insidious onset of the condition, and recognition of septic shock technically depends on management decisions (i.e. initiation of vasopressors), as septic shock is defined as "sepsis with persisting hypotension requiring vasopressors to maintain MAP ≥ 65 mmHg" along with hyperlactatemia, according to the 3rd International Consensus Definition (Sepsis-3) (1). The progression of sepsis to septic shock might therefore be better understood if one were to be able to characterize, in a formalized manner, the initiation of vasopressors in septic patients as part of the "usual care" for such patients.

The relevant physiology and associated clinical practice in sepsis patients have been poorly described in the past. In prior reports, usual care has been characterized as administration of "liberal fluids" before moving on to vasopressors when hypotension is refractory (2). However, that characterization did not rely on a formal analysis of a body of data. Additionally, practice guidelines do not define usual care as they tend to be intentionally formulated in a manner to allow for substantial latitude in care practice (3), (4).

Recently, our team found that the decision made by clinicians to begin or forego vasopressors in hypotensive sepsis patients can be well modeled by a small number of clinical factors, including the degree of hypotension, fluid volume administered, and basic vital signs information that are established metrics of disease severity (5). In this report, we sought to further validate our predictive model by applying it to a new and more recent cohort of patients and thereby demonstrating that the concept of usual care can be formalized in septic patients with hypotension.

II. Methods

A. Original Model Development

The model was developed from a cohort of Emergency Department (ED) patients at a large, urban academic hospital over a 2-year period from April 1, 2014, to March 31, 2016. The study was approved by the local Institutional Review Board with requirement for informed consent waived. From 185,949 encounters, 705 encounters were selected who met criteria adapted from the current Centers for Medicare and Medicaid Services Severe Sepsis/Septic Shock Early Management Bundle (SEP-1) definition for septic shock (6). Patients had to meet all of the following criteria: 1) a final hospital ICD-10 diagnosis code for sepsis; 2) either confirmed source of infection or high suspicion for infection documented in the admission note; and 3) development of persistent hypotension (systolic blood pressure < 90 mmHg on at least two measurements), or elevated lactate ≥ 4.0 mmol/L, or use of vasopressor (VP) medication in the ED. Patients were excluded if they received VP prior to ED arrival, had a non-infectious etiology or were made comfort measures only (CMO) in the ED. A final cohort of 589 encounters were retained after applying the exclusion criteria.

For each patient in the cohort; labs, vital signs, patient locations, demographics and hospital outcome were downloaded from the EMR. Two trained chart reviewers independently reviewed clinical documentation and any

This work was supported in part by Nihon Kohden Corporation.

B.A. is with Harvard Medical School, Boston, MA 02115 USA. (e-mail: b.aydemir94@gmail.com).

V.P. was with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA. He is now with Verily Life Sciences (e-mail: varesh.prasad@gmail.com).

A.T.R. is with Massachusetts General Hospital and Harvard Medical School, Boston, MA 02115 USA (e-mail: areisner@mgh.harvard.edu).

J.C.L, B.B., I.K. and T.H. are with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: jclynch@mit.edu, bbiebelberg@gmail.com, kehoe.i@husky.neu.edu, thomas@mit.edu).

disagreements were resolved by majority vote in a review session that included a third (physician) reviewer.

The statistical model was developed with the variables at the final decision time for vasopressor initiation. The final decision time was defined as the last documented time with a systolic blood pressure (SBP) < 90 mmHg prior to VP initiation for patients who received VP for > 24 hours. For patients who did not receive VP, the final decision time was the last documented time of SBP < 90 mmHg. First, all variables were included in a L1-regularized logistic regression (LR) model, using five-fold cross validation, to select the regularization parameter to maximize the area under the receiving operating characteristic curve (AUC). Next, the variables with non-zero regression coefficients were entered into a stepwise forward selection process to select those with P-values less than 0.05 for a final multivariate LR model.

B. Application of the Model to a New Cohort of Patients

We applied this model to a separate cohort of 55,963 adult ED patient encounters between April 1, 2016 – April 5, 2018 who met at least one of the following criteria: Shock Precautions on Triage (SPoT) vitals (SBP < 100mmHg or Heart Rate (HR) > SBP) (7), quick sequential organ failure assessment score (qSOFA) ≥ 2, (8) Systemic Inflammatory Response Syndrome (SIRS) ≥ 2 (9), ICU admission. This data was derived from the same tertiary care center, though represented a patient cohort not used for model development. From this cohort, 2,681 patient encounters were selected in which the patient received intravenous antibiotics in the ED and was subsequently admitted to the intensive care unit (ICU). We used the combination of ED antibiotics and ICU admission as an adjunct criterion to select the majority of encounters for which a clinical VP initiation decision needed to be made for a patient with life-threatening sepsis. For each of these patients ED triage time, SBP, respiratory rate (RR), peripheral capillary oxygen saturation (SpO_2), Glasgow Coma Scale (GCS) score and fluid boluses during hypotension were retrieved from the electronic medical record (EMR). For each of these parameters, at least 20 cases were randomly selected and compared to the patient's chart in the EMR to confirm agreement. The model score was calculated according to

$$Score = 1.059 - (0.458 \times SBP\_cont) + (0.087 \times RR\_exp) - (0.093 \times SpO_2\_min) - (0.910 \times GCS\_min) + (0.347 \times Fluids\_hypotension) - (0.909 \times Time\_since\_triage)$$

SBP_cont is the SBP in mmHg at the time of the model score calculation. *RR_exp* is the exponentially weighted RR in min^{-1}, with each older respiratory rate carrying half as much weight as the more recent one. For *SBP_cont* and *RR_exp*, the most recent values are carried over to the subsequent time points (sample-and-hold) until a new SBP or RR measurement is available. *GCS_min* is the minimum GCS score up until the time of the model score calculation. *Fluids_hypotension* is the sum total fluid amount in liters at the time of the model score calculation involving lactated ringers and 0.9% sodium chloride boluses. A fluid administration was counted if it was administered while the most recent SBP measurement was less than 90 mmHg. *Time_since_triage* is the time elapsed from the ED triage time in hours.

Patients were classified as receiving vasopressors if they received vasopressors in the ED or within 6 hours of the time the patient was discharged from the ED. Vasopressor medications included are dopamine, epinephrine, norepinephrine, and dobutamine.

For a given patient, the maximum model score prior to the administration of vasopressors was used to generate the receiver operating characteristic (ROC) curve. For patients who did not receive vasopressors within 6 hours of leaving the ED, the maximum model score throughout the ED encounter was used. For the study cohort, a threshold score was selected to maximize the difference between true positive rate (TPR) and false positive rate (FPR). Advance detection time was calculated as time difference between VP administration and the first time at which the model score passed the threshold for a given patient encounter.

In an attempt to evaluate model performance in a cohort enriched with sepsis cases, a separate ROC curve was calculated by excluding the encounters without at least one SBP measurement less than 100 mmHg.

C. Statistical Methods

For study cohort group comparisons, two-tailed t-tests were used for continuous variables, and chi-squared tests were used for categorical variables. The ROCs were generated with MATLAB and their AUC calculated with the trapezoid rule.

III. Results

A. Study cohort

Of the 55,963 adult ED patient encounters, 2,681 encounters met the initial inclusion criteria of intravenous antibiotics in the ED and admission to the ICU. After excluding the encounters for which a model score could not be calculated (most commonly due to absence of recorded GCS), and those who received vasopressors prior to the point at which the model score can be calculated, the final study cohort included 1,629 encounters. Study cohort characteristics are presented in Table I. Compared to patients who received vasopressors (VP group), patients who did not receive vasopressors (no-VP group) were older and had a lower percentage of non-white patients. The VP group also had a higher heart rate, lower SBP and lower SpO_2 at ED triage. The VP group also had an average total fluid amount more than twice that of the no-VP group. Triage temperature and respiratory rate did not differ significantly between the two groups.

B. Receiver Operating Characteristic Curve

The ROC curve for the model is shown in Fig. 1. The model achieved an AUC of 0.81 when applied to the entire cohort of ED patients admitted to the ICU after receiving antibiotics in the ED. In an attempt to enrich the population further with sepsis cases, we repeated the analysis by excluding the encounters that did not have an SBP measurement less than 100 mmHg during their stay. In this population, the model achieved an AUC value of 0.72.

TABLE I. STUDY COHORT CHARACTERISTICS. VALUES PRESENTED EITHER AS FRACTIONS OR AS MEANS (STANDARD DEVIATION).

Variable (units)	*Vasopressor encounters (N = 469)*	*No-vasopressor encounters (N = 1,160)*	*P-value*
Age (years)	61.2 (17.5)	66.8 (15.1)	< 0.001
Female %	45	40	0.069
Non-White %	22	15	0.003
Triage temperature (°F)	97.9 (5.49)	98.0 (1.44)	0.387
Triage heart rate (min^{-1})	101.0 (24.6)	95.4 (22.0)	< 0.001
Triage systolic blood pressure (mmHg)	114.1 (30.8)	131.0 (29.0)	<0.0dddd sdsddsad 01
Triage respiratory rate (min^{-1})	20.8 (5.19)	20.5 (5.02)	0.211
Triage SpO_2 %	94.8 (5.64)	95.8 (4.43)	< 0.001
Triage GCS score	13.5 (3.15)	13.d8 (2.8)	0.040
Total fluid boluses (mL)	1739 (1383)	864 (1046)	< 0.001

C. *Distribution of Advance Detection times*

A threshold model score was selected to maximize the difference between TPR and FPR. At this threshold, the model sensitivity and specificity were 0.74 and 0.74, respectively, in the complete study cohort of 1,629 encounters. For the true positive cases, a distribution of advance detection times is presented in Fig 2. The median advance detection time was 170.5 minutes. The 25th and 75th percentiles for advance detection time were 53 and 363 minutes, respectively.

IV. DISCUSSION

This analysis confirms our prior results on an entirely new and more recent validation cohort and demonstrates that the results of the 'usual care' model for vasopressor initiation hold despite significantly increased awareness, and therefore increased clinical vigilance, for sepsis and septic shock. Previously, we obtained an AUC of 0.93 (95% CI 0.90–0.96) on the training cohort and 0.78 (95% CI 0.58–0.99) on the held-out validation cohort across all (hypotensive and non-hypotensive) encounters (5). On encounters with at least one hypotensive reading, we previously obtained an AUC of 0.80 (95% CI 0.78–0.83) in the training cohort and 0.77 (95% CI 0.68–0.86) in the validation cohort (5). The AUCs reported here on a completely independent validation cohort are 0.81 for performance evaluation on all encounters and 0.72 for performance evaluation on hypotensive encounters are in line with the performance on the held-out validation set in our previous analysis.

Taken together, these results provide strong evidence that usual care regarding the initiation of vasopressors at a tertiary care medical center can be reasonably modeled. With regard to fluid administration, the results state, simply put, that the more abnormal the vital signs, the less fluid the patient received before initiation of vasopressors. Specifically, lower SBP, higher RR, lower SpO_2 and lower GCS correspond to a higher likelihood of receiving vasopressors, independent of the amount of fluid a patient receives. This characterization is different from the "liberal fluids" paradigm that has previously characterized usual care (2), and it is notable that the liberal-fluids characterization did not rely on statistical modeling. Additionally, more time elapsed from triage independently predicted a lower likelihood of receiving vasopressor. This is most likely because the patients who did not receive vasopressors for a long time had some observation available to the treating physician that were not captured by the variables of the model.

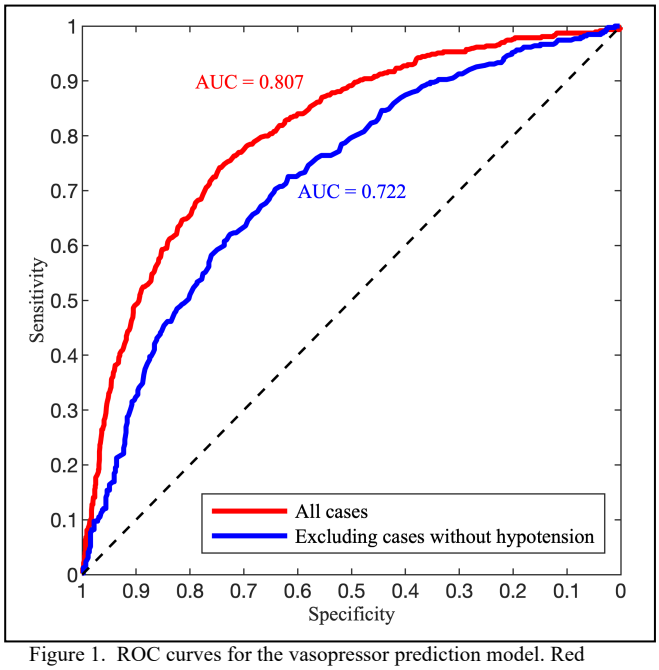

Figure 1. ROC curves for the vasopressor prediction model. Red curve: model performance on all cases; blue curve: model performance on cases who had at least one systolic blood pressure measurement below 100 mmHg.

One reason why the current results are notable is because they represent a completely different, and more recent, patient cohort. The fact that the predictive model yielded comparable performance to the prior validation data suggests that "usual care" has not substantially changed in the time since the predictive model was first developed. Additionally, despite a change in EMR systems, patient characteristics were similarly represented in the older and the more recent patient cohort.

The second reason why the current results are notable is because they were sourced entirely from a data query from a large EMR. The original development of the VP usual-care model arose from a dataset that involved substantial human chart review. The current analysis was based on an electronic query of an EMR with no human-in-the-loop to filter out unreliable data. In theory, it was possible that the model performance would degrade simply because the underlying EMR data contained too many errors (e.g., incorrectly documented vital-signs or incorrectly documented vasopressor infusion times). Of course, it is possible that there were errors and/or biases in the EMR dataset that we analyzed in this report. To rule this out, we are currently performing chart reviews in a convenience cohort to validate these data elements. Yet it is encouraging that, the performance of the model, when applied automatically to a

large EMR data query, was comparable to the prior performance. This is preliminary evidence that large-data analysis on the topic of VP initiation may be valid.

One limitation of our current analysis is that the study cohort was derived from patients who received intravenous antibiotics in the ED and were subsequently admitted to the ICU. While the majority of these patients were cases of sepsis, there were a number of cases with non-infectious etiology (*e.g.*, trauma). We attempted to overcome this limitation by applying our model to a subset of this study population with demonstrated hypotension, assuming that this would enrich the population with true sepsis cases. In this analysis, the model AUC remained high, indicating that our model can still perform well in a more challenging cohort enriched with sicker patients. Additionally, good performance in the original cohort involving a number of non-sepsis cases argues that our model may be generalizable to VP initiation decisions beyond just sepsis. In the future, our team plans to repeat the analysis in a "purer" cohort of sepsis cases based on diagnosis codes, as well as in other critically ill patient populations such as trauma cases. Additionally, a multicenter validation of the predictive model will shed light on the question whether "usual care" for VP initiation in one hospital captures equally well to how patients are treated in a different hospital.

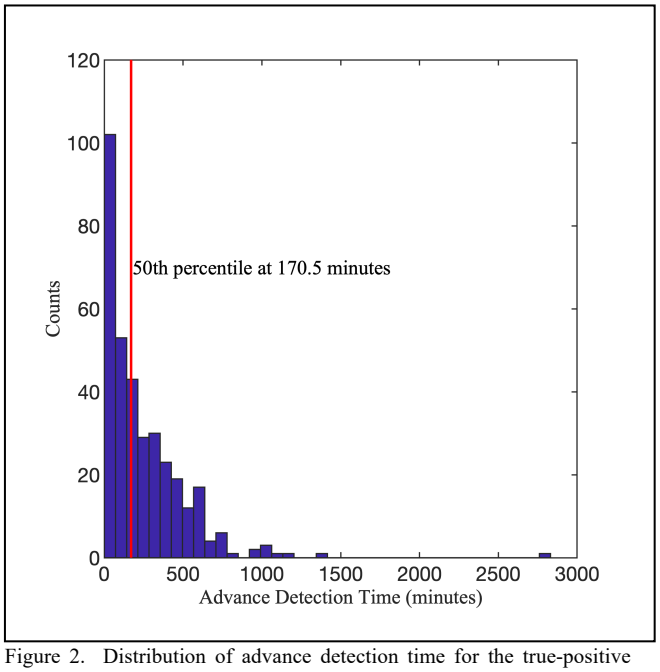

Figure 2. Distribution of advance detection time for the true-positive cases. A threshold score was selected for the study cohort that maximized the difference between TPR and FPR. The 25th and 75th percentiles for advance detection time were 53 and 363 minutes, respectively.

Assuming that this model for usual care is further validated, not only for the EMR data from our medical center (as discussed above) but across multiple medical centers, then this model may be useful for designing and interpreting prospective trials: an objective tool for describing usual care and evaluating when management was atypical. After further validation, our team plans to apply the model to large datasets to assess whether there are differences in outcomes when patients receive VP in accordance to the model, or if VP are imitated earlier than the model predicts, or if VP were initiated late (or never) according to the model.

V. Conclusions

Our work demonstrates that the performance of a previously developed model for usual care in hypotensive sepsis patients holds up in a new and more recent cohort of patients. Our work therefore adds evidence that a small number of clinically relevant variables describes the decision made by clinicians to begin or forego vasopressors in hypotensive sepsis. Expansion of such modeling for multicenter validation is therefore warranted.

References

[1] M. Shankar-Hari, G. S. Phillips, M. L. Levy, C. W. Seymour, V. X. Liu, C. S. Deutschman, et al., "Developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for sepsis and septic shock (Sepsis-3)," *JAMA* vol. 315, issue 8, pp. 775–787, 2016.

[2] W. H. Self, M. W. Semler, R. Bellomo, S. M. Brown, B. P. deBoisblanc, M. C. Exline et al., "Liberal versus restrictive intravenous fluid therapy for early septic shock: Rationale for a randomized trial," *Annals of Emergency Medicine,* vol. 72, issue 4, pp. 457–466, 2018.

[3] M. M. Levy, L. E. Evans, and A. Rhodes, "The surviving sepsis campaign bundle: 2018 update," *Intensive Care Medicine,* vol. 44, issue 6, pp. 925–928, 2018.

[4] A. Rhodes, L. E. Evans, W. Alhazzani, M. M. Levy, M. Antonelli, R. Ferrer et al., "Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016," *Intensive Care Medicine,* vol. 43, issue 3, pp. 304–377, 2017.

[5] V. Prasad, A.T. Reisner, J.C. Lynch, M.R. Filbin, T. Heldt, "Modeling of usual care: vasopressor initiation for sepsis in hypotension," (in review).

[6] Centers for Medicare and Medicaid Services. (2018, Jul. 1). Hospital inpatient specifications manuals, sepsis bundle project (SEP) [Online]. Available: https://qualitynet.org/inpatient/specifications-manuals#tab5

[7] A. T. Reisner, R. Capp, J. E. Thorsen, J. T. Rubin, T. Mooncai et al., "Development and Validation of the SPoT Sepsis (Shock Prevention on Triage) Clinical Decision Rule: 84," *Acad. Emerg. Med.*, vol. 22, issue. S1, pp. S43, 2015.

[8] C. W. Seymour, V. X. Liu, T. J. Iwashyna, F. M. Brunkhorst, T. D. Rea et. al., "Assessment of clinical criteria for sepsis for the third international consensus definitions for sepsis and septic shock (sepsis-3)," *JAMA - J. Am. Med. Assoc.*, vol. 315, no. 8, pp. 762–774, Feb. 2016, doi: 10.1001/jama.2016.0288

[9] R. C. Bone, R. A. Balk, F. B. Cerra, R. P. Dellinger, A. M. Fein *et al.,* "American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis," *Crit. Care Med.*, vol. 20, no. 6, pp. 864–874, 1992