AUTOMATED ANALYSIS OF VITAL SIGNS TO IDENTIFY PATIENTS WITH SUBSTANTIAL BLEEDING BEFORE HOSPITAL ARRIVAL: A FEASIBILITY STUDY

Jianbo Liu,* Maxim Y. Khitrov,* Jonathan D. Gates,[†] Stephen R. Odom,[‡]
Joaquim M. Havens,[†] Marc A. de Moya,[§] Kevin Wilkins,[‖] Suzanne K. Wedel,[‖]
Erin O. Kittell,[¶] Jaques Reifman,* and Andrew T. Reisner[¶]

*Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, Fort Detrick, Maryland; and [†]Department of Surgery, Division of Trauma, Burns and Surgical Critical Care, Brigham and Women's Hospital, Boston, Massachusetts; [‡]Department of Surgery, Division of Acute Care Surgery, Trauma, and Surgical Critical Care, Beth Israel Deaconess Medical Center, Boston, Massachusetts; and [§]Department of Surgery, Division of Trauma, Emergency Surgery and Surgical Critical Care, Massachusetts General Hospital, Boston, Massachusetts; [‖]Boston MedFlight, Bedford, Massachusetts; and [¶]Department of Emergency Medicine, Massachusetts General Hospital, Boston, Massachusetts

ABSTRACT—Trauma outcomes are improved by protocols for substantial bleeding, typically activated after physician evaluation at a hospital. Previous analysis suggested that prehospital vital signs contained patterns indicating the presence or absence of substantial bleeding. In an observational study of adults (aged ≥18 years) transported to level I trauma centers by helicopter, we investigated the diagnostic performance of the Automated Processing of the Physiological Registry for Assessment of Injury Severity (APPRAISE) system, a computational platform for real-time analysis of vital signs, for identification of substantial bleeding in trauma patients with explicitly hemorrhagic injuries. We studied 209 subjects prospectively and 646 retrospectively. In our multivariate analysis, prospective performance was not significantly different from retrospective. The APPRAISE system was 76% sensitive for 24-h packed red blood cells of 9 or more units (95% confidence interval, 59% – 89%) and significantly more sensitive ($P < 0.05$) than any prehospital Shock Index of 1.4 or higher; sensitivity, 59%; initial systolic blood pressure (SBP) less than 110 mmHg, 50%; and any prehospital SBP less than 90 mmHg, 50%. The APPRAISE specificity for 24-h packed red blood cells of 0 units was 87% (88% for any Shock Index ≥1.4, 88% for initial SBP <110 mmHg, and 90% for any prehospital SBP <90 mmHg). Median APPRAISE hemorrhage notification time was 20 min before arrival at the trauma center. In conclusion, APPRAISE identified bleeding before trauma center arrival. *En route*, this capability could allow medics to focus on direct patient care rather than the monitor and, via advance radio notification, could expedite hospital interventions for patients with substantial blood loss.

KEYWORDS—Trauma, hemorrhage, massive transfusion, decision-support systems, prehospital emergency care

INTRODUCTION

Background

Hemorrhage is recognized as the leading treatable cause of death after injury (1). Improved outcomes in trauma patients have been shown when trauma centers apply specific protocols for patients with substantial bleeding (2, 3). These protocols encompass damage-control resuscitation, including aggressive measures to avoid coagulopathy (via permissive hypotension that slows blood loss, adequate restoration of coagulation factors via transfusion, and minimization of hypothermia), which is important because trauma-induced coagulopathy affects between 24% and 56% of critically injured patients (4). For these patients, massive transfusion of packed red blood cells (PRBCs), that is, 10 or more units in 24 h (5), is often necessary. Damage-control resuscitation is paired with damage-control surgery, the operative strategy of prioritizing early surgical control of bleeding, while sparing noncritical surgical repairs that are undertaken only after the patient has sufficiently recovered.

Although management protocols for patients with substantial bleeding are associated with mortality benefits (3, 6), there are no widely accepted criteria for their initiation. Holcomb and Gumbert (2) commented that, in the report by Cotton et al. (3), activation had been subjective after a surgeon's evaluation of the patient. Riskin et al. (6) reported that the Stanford Protocol was activated subjectively "at the discretion of the attending physician." Several clinical scores to predict whether trauma patients will require massive transfusion have been developed, including the McLaughlin score, the Trauma Associated Severe Hemorrhage (TASH) score, and the Assessment of Blood Consumption (ABC) score (5). These scores are based on vital sign data; mechanism of injury or anatomic details (TASH score and ABC score); and abdominal sonography

(ABC score), laboratory testing (McLaughlin score), or both (TASH score).

Our group was interested in whether it would be feasible to identify patients with substantial bleeding before arrival at the hospital, without relying on testing or expertise that is normally hospital based, such as sonography or laboratory testing. Growing evidence shows that assessment of multiple vital signs together may be more effective than univariate approaches for detecting hemorrhagic hypovolemia (7, 8). In addition, there have been encouraging reports of computational techniques (9, 10) to account for the fact that high-acuity trauma patients demonstrate complex temporal fluctuations in their prehospital vital signs (11–13), and to identify unreliable vital signs (14), because spurious measurements are so common (15–18).

In this report, we evaluate the hypothesis that it is feasible to identify patients with substantial 24-h PRBC transfusion requirements by automated analysis of prehospital vital signs. To test this prospectively, a specialized real-time computing platform was developed and deployed into an active prehospital operation (19). If it is feasible to identify patients with substantial bleeding by automated analysis of prehospital vital signs, there might be improved *en route* care as well as in-hospital care. *En route*, caregivers could focus more on patient care rather than split attention with reexamining and reevaluating the vital sign monitor. The automated system could notify the caregivers when the vital signs were statistically consistent with bleeding and display an on-screen checklist of expected responses. The receiving hospital could be provided with advance radio notification, offering a head start for careful preparation of a patient with major hemorrhage, for example, prewarming of the patient's bay (to prevent hypothermia), preparation of fresh frozen plasma for immediate transfusion, and mobilization of surgical assets (for early surgical intervention).

METHODS

Setting and study population

We examined a convenience sample of adult (aged ≥18 years) trauma patients transported by air emergency medical service (EMS) to participating level I trauma centers. With institutional review board approval, we collected a prospective data set from Boston MedFlight (BMF, Bedford, Mass) and

compared the findings with an archival data set originally collected from Memorial Hermann Life Flight (MHLF, Houston, Tex) by Cooke et al. (20) and Holcomb et al. (21). In both data sets, we analyzed all subjects with at least one recorded non-zero systolic blood pressure (SBP). Patients who died before hospital admission (e.g., in the emergency department) were excluded from analysis, because resuscitation was often terminated before large-volume PRBC transfusion could be completed, regardless of whether or not the patient had significant hypovolemia.

Our primary study outcome was 24-h PRBC transfusion volume in patients with hemorrhagic injury, defined as a documented hemorrhagic injury that unequivocally caused some loss of blood (laceration or fracture of a solid organ; documented hematoma within the thorax, peritoneum, retroperitoneum, or pelvis; vascular injury that required operative repair; or limb amputation) and PRBC transfusion within 24 h. Patients who received PRBCs but lacked a documented hemorrhagic injury were excluded from the primary analysis because, in the absence of an explicitly hemorrhagic injury, it was challenging to determine whether the transfusion was clinically indicated. Whether the patient had documented hemorrhagic injury was determined by automated text search, searching for injuries that met the aforementioned criteria (records were also jointly reviewed by two investigators, J.L. and A.T.R., who confirmed that the automated text search had not omitted any applicable hemorrhagic injuries nor included nonhemorrhagic injuries).

The excluded patients who lacked explicitly hemorrhagic injuries were reincluded and analyzed in a sensitivity analysis (see Appendix, Supplemental Digital Content 1, at http://links.lww.com/SHK/A267).

Vital sign data processing

For the prospective cohort, we deployed the APPRAISE (Automated Processing of the Physiological Registry for Assessment of Injury Severity [19]) system onto two active BMF helicopters between February 5, 2010, and December 31, 2012. The APPRAISE system consists of a Propaq 206 patient monitor (Welch-Allyn, Beaverton, Oreg) networked to the GoBook ultracompact ruggedized personal computer (General Dynamics Itronix, Sunrise, Fla) running analytic algorithms developed for this research project (19). As a practical matter, this meant that all vital sign data processing and analyses for BMF were done automatically and in real time.

The following routine vital signs were monitored by the Propaq 206 monitor: heart rate (HR), respiratory rate (RR), oscillometric SBP, and pulse pressure ([PP], the difference between SBP and diastolic blood pressure). The APPRAISE software created an electronic record of the Propaq data, analyzed the vital sign data in real time using algorithms described below and archived the results. The results of the automated analysis were not visible to the flight crew so that the investigational system would not affect clinical decision making (this was a matter of human subject protection for a diagnostic system that had not yet been validated during clinical operation).

The retrospective data originally had been collected onboard MHLF helicopters between August 2001 and April 2004 using a personal digital assistant networked to a Propaq 206 patient monitor to archive the vital sign data (21). Subsequently, those data were uploaded to our data warehousing system (22) and analyzed offline.

We analyzed the prospective and the retrospective Propaq 206 data using the exact same computational methodology. First, the automated algorithms identified and excluded unreliable vital sign measurements (Fig. 1). The reliability
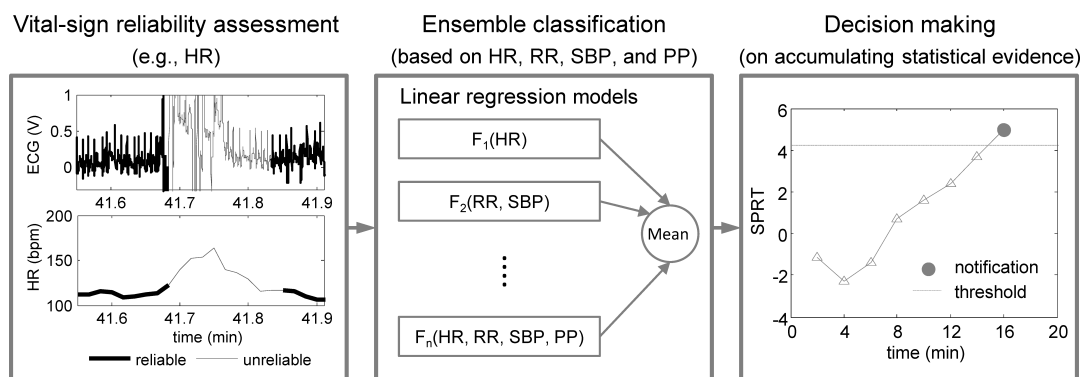
FIG. 1. **Analytic methodology for hemorrhage identification**. In the first step (left panel), algorithms were used to identify, and exclude, unreliable vital signs. In the second step (middle panel), ensemble classification was applied, which consisted of a set of different linear regression models, $F_1, F_2, \cdots, F_n$, that were subsequently averaged together. Ensemble classification is useful when missing data are commonplace: different regression models contain different combinations of the vital signs, and it is possible to omit any of those models that contain a missing input parameter. In the third step (right panel), the mean ensemble classifier was evaluated by the SPRT, a statistical test of whether or not measurements repeated across time are consistent with a control distribution or with a different (e.g., hemorrhagic patient) distribution. bpm—beats per minute; ECG—electrocardiography; V—volt.

algorithms for HR and RR involved analysis of the electrocardiography (ECG) and impedance pneumography waveforms. This allowed us to discriminate between a clean source signal versus an unreliable segment caused by signal artifacts (23, 24). The SBP and PP reliability algorithms assessed signal quality by analyzing the relationship between systolic, diastolic, and mean arterial pressures and by comparing HR measured by ECG versus HR measured by oscillometry (25). These automated algorithms, which have been shown to agree with human experts' opinions (23, 24), can significantly increase the diagnostic value of vital signs by removing spurious measurements (25, 26).

The second step of real-time analysis involved an ensemble classifier, which is a set of multivariate regression models whose numerical outputs were averaged to yield the final output (Fig. 1). We trained the multivariate regression models (i.e., set the weights for the input variables) for a binary outcome as per Chen et al. (27), using the initial 15 min of vital sign data from each MHLF subject. For the model training, the binary outcome was whether patients received 1 or more PRBCs for an unambiguous hemorrhagic injury or not. This model training yielded a classifier that, on the basis of the input vital signs (HR, RR, SBP, and PP), quantified whether the pattern was similar to the population with hemorrhage (output closer to 1) or to the nonhemorrhagic control population (output closer to 0). The ensemble classifier was originally developed for use at a single time point, for example, on 15 min of prehospital data collection, for prediction of 24-h PRBC more than 0, and it was cross-validated using 50%/50% training/testing (27). There were no significant differences ($P > 0.05$) when the receiver operating characteristic area under the curve (ROC AUC) of 10-fold cross-validation was compared with the ROC AUC for 100%/100% training/testing (ΔROC AUC $\pm$ 0.01). Compared with routine multivariate regression, an ensemble classifier can provide two advantages. First, the ensemble can still classify patients even when a complete set of reliable vital signs is unavailable. Second, it can offer performance that is more consistent from one data set to the next (27, 28).

Every 2 min, this analysis was repeated. For the prospective trial, this occurred in real time. For the retrospective analysis, we reapplied the algorithms at every 2-min mark of the patient's electronic record, simulating real-time application. Every time the ensemble classifier was applied (i.e., every 2 min), we analyzed the time-averaged value of all *reliable* HR, RR, SBP, and PP measured since the beginning of the record up to the time of analysis. (For example, at $t = 6$ min, all vital sign data from $t = 0$ to $t = 6$ min were analyzed. At $t = 8$ min, all vital sign data from $t = 0$ to $t = 8$ min were analyzed.) The rationale for analyzing data reaching back to the start of the mission arose from previous analysis suggesting that prehospital vital signs contained enormous variability—likely caused by pain, medications, or other transient stimuli—and that time averaging was an effective method to remove some of the confounding data perturbations and achieve superior diagnostic performance (9).

The third and final step of real-time analysis involved the Wald sequential probability ratio test (SPRT) for determining whether to issue a "hemorrhage notification" on the basis of accumulated evidence from the ensemble classifier outputs (Fig. 1). The SPRT (29) is a useful statistical technique for determining whether repeated measurement samples are consistent with one statistical distribution (e.g., a normal population) versus a second statistical distribution (e.g., an abnormal population). Thresholds for the SPRT were set as per Chen et al. (10), where the SPRT was shown to reduce false alarms at the expense of some alarm latency.

Clinical outcomes

For the BMF data set, a research nurse collected patient attributes and outcome data via retrospective chart review of the receiving hospitals' medical records (i.e., Beth Israel Deaconess Medical Center, the Brigham and Women's Hospital, and the Massachusetts General Hospital. The data were archived electronically using REDCap (30). We obtained injury severity scores from each hospital's trauma registry. For the MHLF data set, a chart review was conducted by the original study authors (21).

Statistical analysis

Data were summarized using mean with standard deviation or median with interquartiles for continuous variables, and frequency with percentage for categorical variables. We computed the proportion of patients who received a hemorrhage notification as a function of the number of units of PRBCs that each patient received during the initial 24 h in the hospital ("24-h PRBC volume"). For comparison, we also computed the proportion of patients with other hemodynamic abnormalities: initial SBP less than 110 mmHg, any prehospital SBP less than 90 mmHg, or any prehospital Shock Index (SI) of 1.4 or higher (where SI = HR/SBP). The threshold for SI was chosen based on the findings in Mutschler et al. (31). We tested for significant differences between those proportions using McNemar's test.

For BMF patients, MHLF patients, and the pooled data set, we developed logistical regression models to quantify the likelihood of a patient receiving a

hemorrhage notification as a function of 24-h PRBC volume. We also tested whether the likelihood of a patient receiving a hemorrhage notification differed between the two populations (BMF and MHLF) controlling for the 24-h PRBC volume. Finally, to investigate other factors that may have influenced whether a patient received a hemorrhage notification or not, we applied multivariate logistical regression to a set of parameters quantifying potential sources of variability: age, mechanism of trauma, prehospital factors (elapsed time since injury, volume of resuscitation, endotracheal intubation, duration of transport), and anatomy of the injuries based on the trauma registry Abbreviated Injury Scale scores. Two-sided values of $P < 0.05$ were considered as statistically significant.

RESULTS

Of the 999 patients with electronic data available (MHLF, 757; BMF, 242), we excluded 22 who lacked a non-zero blood pressure measurement (MHLF, 20; BMF, 2) and 33 who did not survive to admission (MHLF, 27; BMF, 6). Also, there were 89 patients who received 24-h PRBC transfusion while lacking explicitly hemorrhagic injuries (MHLF, 64; BMF, 25); these patients were examined in the sensitivity analysis (see Appendix, Supplemental Digital Content 1, at http://links.lww.com/SHK/A267). Table 1 describes the primary study population.

Correlation of basic vital signs and 24-h PRBC volume

In the MHLF data set, SBP, PP, HR, and RR were significantly ($P < 0.001$) correlated with 24-h PRBC transfusion volume: $\rho = -0.32$, $\rho = -0.36$, $\rho = +0.24$, and $\rho = +0.24$, respectively.

In the BMF data set, SBP ($P < 0.001$) and PP ($P < 0.01$) were significantly correlated with 24-h PRBC transfusion volume: $\rho = -0.30$ and $\rho = -0.23$, respectively. Heart rate showed a nonsignificant trend ($P = 0.051$), with $\rho = +0.14$, whereas RR was not significantly correlated.

Diagnostic test characteristics

Table 2 shows the relationship between the incidence of APPRAISE hemorrhage notification and 24-h PRBC transfusion volume. With increasing 24-h PRBC transfusion volume,

TABLE 1. **Study population characteristics**

	Memorial Hermann Life Flight	Boston MedFlight
Population, n	646	209
Sex, male, n (%)	479 (74)	155 (74)
Age, mean (SD), years	38 (15)	45 (20)
Blunt, n (%)	577 (89)	188 (90)
Penetrating, n (%)	61 (9)	21 (10)
ISS, median (IQR)	16 (9 – 34)	16 (9 – 26)
Interhospital transfer, n (%)	0 (0)	103 (49)
Prehospital airway intubation, n (%)	111 (17)	80 (38)
Prehospital GCS, median (IQR)	15 (13 – 15)	15 (8 – 15)
Prehospital blood transfusion, n (%)	0 (0)	15 (7)
24-h PRBC volume >0 unit, n (%)	75 (12)	31 (15)
24-h PRBC volume ≥3 units, n (%)	57 (9)	18 (9)
24-h PRBC volume ≥10 units, n (%)	22 (3)	8 (4)
Survival to discharge, n (%)	608 (94)	191 (91)

GCS—Glasgow Coma Scale; IQR—interquartile range; ISS—Injury Severity Score; SD—standard deviation.

TABLE 2. **Relationship between prehospital APPRAISE hemorrhage notification versus 24-h PRBC transfusion volume**

	24-h PRBC volume (units)				
	0	1 – 2	3 – 8	≥9	Total
Total patients, n	**749**	**31**	**41**	**34**	**855**
MHLF patients, n	571	18	32	25	646
BMF patients, n	178	13	9	9	209
Hemorrhage notification, n (%)	**96 (13)**	**12 (39)**	**26 (63)**	**26 (76)**	
MHLF, n (%)	79 (14)	9 (50)	22 (69)	19 (76)	
BMF, n (%)	17 (10)	3 (23)	4 (44)	7 (78)	
Any SI ≥1.4, n (%)	**92 (12)**	**8 (26)**	**21 (51)**	**20 (59)**	
MHLF, n (%)	70 (12)	6 (33)	18 (56)	14 (56)	
BMF, n (%)	22 (12)	2 (15)	3 (33)	6 (67)	
Initial SBP <110 mmHg, n (%)	**87 (12)**	**9 (29)**	**22 (54)**	**17 (50)**	
MHLF, n (%)	67 (12)	5 (28)	18 (56)	11 (44)	
BMF, n (%)	20 (11)	4 (31)	4 (44)	6 (67)	
Any SBP <90 mmHg, n (%)	**73 (10)**	**9 (29)**	**24 (59)**	**17 (50)**	
MHLF, n (%)	51 (9)	6 (33)	18 (56)	11 (44)	
BMF, n (%)	22 (12)	3 (23)	6 (67)	6 (67)	

the proportion of APPRAISE-positive subjects exhibited an increasing trend in both the MHLF and BMF studies. In the pooled data set (MHLF and BMF), sensitivities for 24-h PRBC transfusion volume of 9 or more units for APPRAISE notification, SI of 1.4 or higher, initial SBP less than 110 mmHg, and any hypotension (SBP <90 mmHg) were 76% (59% – 89%), 59% (41% – 75%), 50% (32% – 68%), and 50% (32% – 68%), respectively, and we found that the sensitivity of APPRAISE notification was significantly higher than SI of 1.4 or higher ($P = 0.014$), initial SBP less than 110 mmHg ($P = 0.007$), and any hypotension, that is, SBP less than 90 mmHg ($P = 0.007$). The sensitivities of APPRAISE notification for 24-h PRBC transfusion volume of 9 or more units were similar for the MHLF versus BMF data sets: 76% (55% – 91%) and 78% (40% – 97%), respectively.

In the pooled data set (MHLF and BMF), specificities for 24-h PRBC transfusion volume of 0 units (i.e., no blood transfusion at all) for APPRAISE notification, SI of 1.4 or higher, initial SBP less than 110 mmHg, and any hypotension (SBP <90 mmHg) were 87% (85% – 89%), 88% (85% – 90%), 88% (86% – 91%), and 90% (88% – 92%), respectively, and we found that the specificity of APPRAISE was not significantly different from initial SBP less than 110 mmHg or any prehospital SI of 1.4 or higher. Compared with any prehospital SBP less than 90 mmHg, APPRAISE notification showed a significantly lower specificity ($P < 0.05$), although the absolute magnitude of the difference was 3%. The specificities of APPRAISE notification for 24-h PRBC transfusion volume of 0 units were 86% (83% – 89%) for the MHLF data set and 90% (85% – 94%) for the BMF data set. In the pooled data set, negative predictive values (24-h PRBC transfusion volume = 0 units vs. ≥1 unit) for APPRAISE

notification, SI of 1.4 or higher, initial SBP less than 110 mmHg, and any hypotension (SBP <90 mmHg) were similar: 94% (92% – 96%), 92% (90% – 94%), 92% (90% – 94%), and 92% (90% – 94%), respectively.

Incidentally, there were three subjects who received prehospital needle decompression, and all received hemorrhage notifications during transport (24-h PRBC volumes for these subjects were 0, 4, and >20, respectively).

Timelines

Figure 2 illustrates prehospital timelines for all subjects with 24-h PRBC volume of 9 or more units, showing the timing of blood pressure measurements, of APPRAISE hemorrhage notifications, and episodes of hypotension (SBP <90 mmHg). The median notification time after the start time of transport was 6 min (interquartiles 4 – 16) for MHLF and 10 min for BMF (interquartiles 8 – 40). The median notification time before arrival at the hospital was 17 min for MHLF and 52 min for BMF, and the difference was largely caused by shorter transport times for MHLF (the median transport time for subjects with 24-h PRBC volume ≥9 units was 28 min [interquartiles 24 – 36] for MHLF and 65 min [interquartiles 35 – 78] for BMF). Combining the two populations, APPRAISE notification occurred in the first half of the transportation in 73% of the cases.

Nine subjects returned to APPRAISE-negative status after a hemorrhage notification: six MHLF subjects who were actually false positive (i.e., 24-hr PRBC = 0) and three BMF subjects who were true positive (i.e., 24-hr PRBC ≥ 1) and received prehospital PRBC transfusion.

Multivariate logistic regression

Using logistic regression to model the likelihood of APPRAISE hemorrhage notification as a function of 24-h PRBC transfusion volume further demonstrated that the results were similar in both data sets (Fig. 3). Each PRBC unit transfused was associated with a 43% (95% CI, 30 – 57%) increase in the odds of APPRAISE hemorrhage notification for MHLF and a 44% (95% CI, 24 – 67%) increase for BMF. The odds ratio of APPRAISE notification per unit of PRBC transfused was not significantly different between the two data sets (i.e., BMF versus MHLF) when fitting a regression model to the pooled data set ($P = 0.635$). However, there was a nonsignificant trend toward a lower overall likelihood of hemorrhage notification in the BMF data set when compared with the MHLF data set ($P = 0.053$), including a lower likelihood of hemorrhage notifications in patients without bleeding (24-h PRBC = 0) and with substantial bleeding (24-h PRBC ≥9), which is apparent in the offset between the two regression curves (Fig. 3). Note that the specificities and sensitivities extracted from Figure 3 are slightly different from those reported in Table 2 because of the nature of the regression fit.

We investigated the factors associated with whether subjects received an APPRAISE hemorrhage notification; see univariate and multivariate logistic regression analyses in Table 3. In multivariate analysis, four independent factors were significant predictors of whether the patient received an APPRAISE notification: increasing 24-h PRBC volume, increasing severity
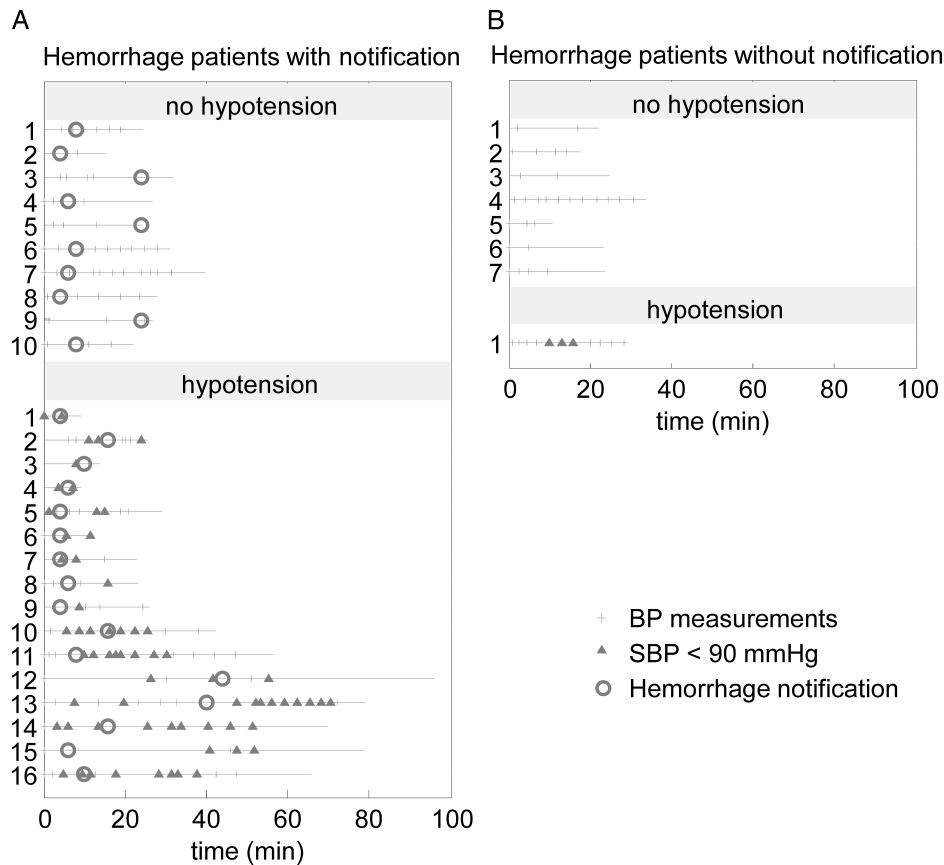
Fig. 2. **Timelines for patients with substantial bleeding (i.e., 24-h PRBC volume ≥9 units) indicating time of hypotensive episodes (SBP <90 mmHg) and hemorrhage notification during prehospital transport.** (A) Bleeding patients who received an APPRAISE hemorrhage notification (i.e., true positives) and (B) bleeding patients who did not receive a notification (i.e., false negatives).

of chest injury, longer flight duration, and younger age. Neither abdominal, nor head, nor extremity injury severity had a significant association with false-negative alarms. Prehospital PRBC transfusion was only found in the BMF cohort, and those patients had a significantly increased risk of APPRAISE hemorrhage notification.

DISCUSSION

This investigation demonstrated that there was a strong association between 24-h PRBC transfusion volume and abnormal prehospital vital signs, and that the majority of patients with large transfusion requirements could be distinguished from other trauma patients using techniques for time series and multivariate analysis. The automated APPRAISE system required neither oversight nor input by the flight crew; it operated wholly autonomously, only requiring that the flight crew use their Propaq transport monitor as per standard procedure. The performance of the APPRAISE algorithms for early identification of patients with 24-h PRBC of 9 or more units was quite similar in actual prospective real-time use (the BMF data set) versus simulated real-time use (the MHLF data set).

Potential benefits of prehospital identification of substantial bleeding

Automated functionality that reliably provides a notification whenever important patterns develop would permit the caregiver to focus much more on the patient (e.g., better pain

control, better management of retching patients who could aspirate, etc.) and not constantly split attention between the patient and the travel monitor. Consistent fully automated detection of hypovolemic vital signs may be most clinically valuable if the EMS caregiver is inexperienced, fatigued, or distracted.

With reliable notification that a bleeding patient is about to arrive, the receiving facility could prepare for hemorrhage-specific management. Today's typical practice involves a trauma team evaluation—postarrival—before deciding whether to activate protocols for substantial bleeding (2). At best, this adds a small delay to care and, in some cases, resultant delays can be substantial. In one report describing the benefits of an institutional protocol for substantial bleeding, interventions such as transfusion of fresh-frozen plasma were not initiated for several hours in many cases (6). By analogy, the common practice of activating the cardiac catheterization team when the prehospital ECG shows ST-elevation myocardial infarction in a patient with chest pain illustrates the potential value of readying the hospital for an exsanguinating patient based on a simple objective prehospital indicator: by initiating in-hospital preparations based on prehospital notification, the time delay to percutaneous coronary intervention can be reduced (32). Of note, cardiologists still conduct expert evaluations before undertaking catheterization, and prehospital notification does not remove clinical authority from hospital caregivers.
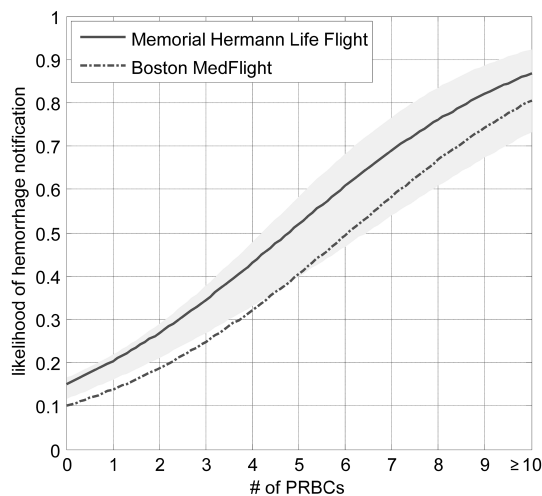
Fig. 3. **Modeling the rate of APPRAISE hemorrhage notification using logistic regression**. The slopes of the BMF and MHLF curves were the same: each PRBC unit transfused was associated with a 43% (95% CI, 30% − 57%) increase in the odds of MHLF hemorrhage notification and also a 44% (95% CI, 24% − 67%) increase for BMF hemorrhage notification. The offset between MHLF and BMF was not statistically significant: further inclusion of a population parameter (1 for MHLF and 0 for BMF) into the regression model for the combined (BMF and MHLF) data set did not yield a statistically significant coefficient for the population parameter. Shaded areas are the 95% confidence intervals for the combined population regression model. The confidence interval becomes wider as a result of a smaller patient population with larger 24-h PRBC volumes. Patients who received 10 or more units of PRBCs were combined into a single category.

It may be most clinically valuable if, rather than a simple alert or notification, the automated system were to display an on-screen list of bulleted action items to remind the EMS caregiver of each and every expected action item for trauma patients with abnormal circulation, for example, check for compressible hemorrhage, check for tension pneumothorax (as noted, the APPRAISE system generated a hemorrhage notification for all three subjects with documented prehospital needle decompression), hold fluids unless SBP was less than 90 mmHg, keep patient warm, and so on. Note that protocol compliance is an underlying challenge throughout health care (33), and checklists are a valuable tool to improve protocol compliance (34–36).

The clinical benefits of this system are speculative because we did not assess clinical impact in the current investigation (an institutional review board–related matter; see Methods). Yet, it seems reasonable to move toward bedside computing for certain tasks, such as statistical analyses that can quantify whether a sequence of vital signs is abnormal, and thereby permit caregivers to focus on quality bedside care.

Physiological interpretation of the findings

At a rudimentary level, this study suggests that patients with massive 24-h blood transfusion requirements demonstrate hypovolemic physiology before hospital arrival. This intuitive finding is consistent with other prediction rules for massive transfusion where hypotension and tachycardia are established predictive factors for massive transfusion (5).

Unlike the other massive transfusion prediction rules, the APPRAISE system only involves vital sign data analyzed during prehospital transport. The APPRAISE system uses well-known statistical techniques, such as time averaging and

the SPRT, for analyzing data that fluctuate through time, and it detects hemorrhage by considering the temporal accumulation of evidence. The system does *not* seek to identify trends through time (e.g., downward drifts in SBP), which may seem counterintuitive, but it has been clearly demonstrated that prehospital vital signs fluctuate substantially frequently without obvious overt directional trends (10, 12–14).

In addition to time series techniques, another common sense principle incorporated in the APPRAISE system was multivariate analysis. Like several prediction rules for massive transfusion (5), the APPRAISE system's algorithms used the independent diagnostic information from more than one vital sign. This is consistent with recent reports that the SI (the ratio of HR to SBP) is a valuable diagnostic tool for identification of hemorrhage (7, 8). The APPRAISE system identifies hypovolemia by a combination of low SBP, low PP, high HR, and high RR. A minority of the massive transfusion patients were not detected by the APPRAISE system; those generally lacked hypotension (Fig. 2B), suggesting that they were not substantially hypovolemic during transport.

There were also APPRAISE hemorrhage notifications in patients who did *not* require massive transfusion. These patients were likely hypovolemic during transport yet without the ongoing blood losses that necessitate massive transfusion (of note, among patients who never needed any PRBCs, those who

TABLE 3. **Factors associated with APPRAISE hemorrhage notification**

Factor	Univariate		Multivariate	
	Odds ratio	95% CI	Odds ratio	95% CI
24-h PRBC transfusion volume (per 1 unit)	1.43‡	(1.32 − 1.55)	1.4‡	(1.29 − 1.52)
Demographics				
Age (per 10 years)	0.91	(0.82 − 1.02)	0.87*	(0.77 − 0.99)
Prehospital course				
Endotracheal intubation (y/n)	2.22‡	(1.53 − 3.20)		
IVF (per 500 mL)	1.51‡	(1.27 − 1.79)		
Time to begin transport (per 10 min)	1.05	(0.99 − 1.12)		
Duration of transport (per 10 min)	1.18†	(1.05 − 1.32)	1.17*	(1.02 − 1.35)
Injury mechanism				
Blunt trauma (y/n)	1.28	(0.70 − 2.33)		
Penetrating trauma (y/n)	0.88	(0.48 − 1.62)		
Injury description				
Head AIS ≥3 (y/n)	1.03	(0.67 − 1.58)		
Abdomen or pelvis AIS ≥ 3 (y/n)	4.41‡	(2.73 − 7.12)		
Extremity, not pelvis AIS ≥3 (y/n)	1.26	(0.80 − 1.98)		
Thorax AIS ≥3 (y/n)	3.73‡	(2.51 − 5.53)	2.58‡	(1.64 − 4.04)

AIS—Abbreviated Injury Scale; CI—confidence interval; IVF—intravenous fluids; (y/n)—binary variables.
Odds ratio significantly different from 1.0: *$P < 0.05$, †$P < 0.01$, ‡$P < 0.001$.

received an APPRAISE hemorrhage notification had significantly higher average injury severity scores). Whether the alert would offer clinical value in this population without massive transfusion requirements is an open question. As discussed above, the APPRAISE system would not obviate the need for clinical assessments by prehospital and receiving facility personnel. Rather, the system is a tool for optimizing vital sign information, offering automated consistent notification when the patterns suggest hypovolemia, and these patterns are strongly associated with subsequent blood transfusion requirements.

Limitations

The study outcome, hemorrhage severity, was quantified by each patient's 24-h PRBC volume. However, the quantity of PRBCs that a patient actually receives is a function of multiple factors, including the speed and effectiveness of surgical hemorrhage control, and some subjective clinical decision making. The generalizability of the findings, that is, the notification incidence versus 24-h PRBC volume, and their applicability to guiding initial resuscitation may have limitations. Yet, the notable consistency (Fig. 3) between the MHLF and BMF results during aeromedical transport to one and three distinct trauma centers, respectively, suggests that such confounding factors can average out across different trauma systems, yielding consistent relationships between prehospital notification incidence and hemorrhage severity.

The prospective BMF arm of this study was sufficient to demonstrate that the real-time system can perform encouragingly well (seven of nine massive transfusion BMF subjects received a real-time prehospital notification). However, the BMF data set was too small to directly compare test characteristics of the APPRAISE notification versus hypotension or SI.

CONCLUSIONS

We conclude that real-time multivariate time series analysis of vital signs is a feasible means of identifying prehospital trauma patients with substantial bleeding, and that prospective investigation of the clinical value of this automated methodology is justified.

REFERENCES

1. Kauvar DS, Lefering R, Wade CE: Impact of hemorrhage on trauma outcome: an overview of epidemiology, clinical presentations, and therapeutic considerations. *J Trauma* 60(6 suppl):S3–S11, 2006.
2. Holcomb JB, Gumbert S: Potential value of protocols in substantially bleeding trauma patients. *Curr Opin Anaesthesiol* 26(2):215–220, 2013.
3. Cotton BA, Reddy N, Hatch QM, LeFebvre E, Wade CE, Kozar RA, Gill BS, Albarado R, McNutt MK, Holcomb JB: Damage control resuscitation is associated with a reduction in resuscitation volumes and improvement in survival in 390 damage control laparotomy patients. *Ann Surg* 254(4):598–605, 2011.
4. Curry N, Davis PW: What's new in resuscitation strategies for the patient with multiple trauma? *Injury* 43(7):1021–1028, 2012.
5. Elmer J, Wilcox SR, Raja AS: Massive transfusion in traumatic shock. *J Emerg Med* 44(4):829–838, 2013.
6. Riskin DJ, Tsai TC, Riskin L, Hernandez-Boussard T, Purtill M, Maggio PM, Spain DA, Brundage SI: Massive transfusion protocols: the role of aggressive resuscitation versus product ratio in mortality reduction. *J Am Coll Surg* 209(2):198–205, 2009.
7. King RW, Plewa MC, Buderer NM, Knotts FB: Shock index as a marker for significant injury in trauma patients. *Acad Emerg Med* 3(11):1041–1045, 1996.
8. Sohn CH, Kim WY, Kim SR, Seo DW, Ryoo SM, Lee YS, Lee JH, Oh BJ, Won HS, Shim JY, et al.: An increase in initial shock index is associated with the requirement for massive transfusion in emergency department patients with primary postpartum hemorrhage. *Shock* 40(2):101–105, 2013.
9. Chen L, Reisner AT, Gribok A, Reifman J: Exploration of prehospital vital sign trends for the prediction of trauma outcomes. *Prehosp Emerg Care* 13(3):286–294, 2009.
10. Chen L, Reisner AT, Chen X, Gribok A, Reifman J: Are standard diagnostic test characteristics sufficient for the assessment of continual patient monitoring? *Med Decis Making* 33(2):225–234, 2013.
11. Morris JA Jr, Auerbach PS, Marshall GA, Bluth RF, Johnson LG, Trunkey DD: The Trauma Score as a triage tool in the prehospital setting. *JAMA* 256(10):1319–1325, 1986.
12. Shapiro NI, Kociszewski C, Harrison T, Chang Y, Wedel SK, Thomas SH: Isolated prehospital hypotension after traumatic injuries: a predictor of mortality? *J Emerg Med* 25(2):175–179, 2003.
13. Lipsky AM, Gausche-Hill M, Henneman PL, Loffredo AJ, Eckhardt PB, Cryer HG, de Virgilio C, Klein SL, Bongard FS, Lewis RJ: Prehospital hypotension is a predictor of the need for an emergent, therapeutic operation in trauma patients with normal systolic blood pressure in the emergency department. *J Trauma* 61(5):1228–1233, 2006.
14. Reisner AT, Chen L, Reifman J: The association between vital signs and major hemorrhagic injury is significantly improved after controlling for sources of measurement variability. *J Crit Care* 27(5):533.e1–533.e10, 2012.
15. Edmonds ZV, Mower WR, Lovato LM, Lomeli R: The reliability of vital sign measurements. *Ann Emerg Med* 39(3):233–237, 2002.
16. Jones DW, Appel LJ, Sheps SG, Roccella EJ, Lenfant C: Measuring blood pressure accurately: new and persistent challenges. *JAMA* 289(8):1027–1030, 2003.
17. Lovett PB, Buchwald JM, Stürmann K, Bijur P: The vexatious vital: neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage. *Ann Emerg Med* 45(1):68–76, 2005.
18. Friesdorf W, Konichezky S, Gross-Alltag F, Fattroth A, Schwilk B: Data quality of bedside monitoring in an intensive care unit. *Int J Clin Monit Comput* 11(2):123–128, 1994.
19. Reisner AT, Khitrov MY, Chen L, Blood A, Wilkins K, Doyle W, Wilcox S, Denison T, Reifman J: Development and validation of a portable platform for deploying decision-support algorithms in prehospital settings. *Appl Clin Inform* 4(3):392–402, 2013.
20. Cooke WH, Salinas J, Convertino VA, Ludwig DA, Hinds D, Duke JH, Moore FA, Holcomb JB: Heart rate variability and its association with mortality in prehospital trauma patients. *J Trauma* 60(2):363–370; discussion 370, 2006.
21. Holcomb JB, Salinas J, McManus JM, Miller CC, Cooke WH, Convertino VA: Manual vital signs reliably predict need for life-saving interventions in trauma patients. *J Trauma* 59(4):821–828, 2005.
22. McKenna TM, Bawa G, Kumar K, Reifman J: The physiology analysis system: an integrated approach for warehousing, management and analysis of time-series physiology data. *Comput Methods Programs Biomed* 86(1):62–72, 2007.
23. Yu C, Liu Z, McKenna T, Reisner AT, Reifman J: A method for automatic identification of reliable heart rates calculated from ECG and PPG waveforms. *J Am Med Inform Assoc* 13(3):309–320, 2006.
24. Chen L, McKenna T, Reisner A, Reifman J: Algorithms to qualify respiratory data collected during the transport of trauma patients. *Physiol Meas* 27(9):797–816, 2006.
25. Reisner AT, Chen L, McKenna TM, Reifman J: Automatically-computed prehospital severity scores are equivalent to scores based on medic documentation. *J Trauma* 65(4):915–923, 2008.
26. Chen L, Reisner AT, Gribok A, McKenna TM, Reifman J: Can we improve the clinical utility of respiratory rate as a monitored vital sign? *Shock* 31(6):574–580, 2009.
27. Chen L, McKenna TM, Reisner AT, Gribok A, Reifman J: Decision tool for the early diagnosis of trauma patient hypovolemia. *J Biomed Inform* 41(3):469–478, 2008.
28. Dietterich TG: Ensemble methods in machine learning. In *Multiple Classifier Systems*. Vol. 1857. Berlin, Germany: Springer, pp 1–15, 2000.
29. Wald A: Sequential tests of statistical hypotheses. *Ann Math Stat* 16(2):117–186, 1945.
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG: Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 42(2):377–381, 2009.
31. Mutschler M, Nienaber U, Munzberg M, Wolfl C, Schoechl H, Paffrath T, Bouillon B, Maegele M, The TraumaRegister, DGU: The Shock Index revisited—a fast guide to transfusion requirement? A retrospective analysis on 21,853 patients derived from the TraumaRegister DGU®. *Crit Care* 17(4):R172, 2013.
32. Bradley EH, Nallamothu BK, Curtis JP, Webster TR, Magid DJ, Granger CB, Moscucci M, Krumholz HM: Summary of evidence regarding hospital strategies

to reduce door-to-balloon times for patients with ST-segment elevation myocardial infarction undergoing primary percutaneous coronary intervention. *Crit Pathw Cardiol* 6(3):91–97, 2007.

33. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, Rubin HR: Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 282(15):1458–1465, 1999.

34. Spector JM, Lashoher A, Agrawal P, Lemer C, Dziekan G, Bahl R, Mathai M, Merialdi M, Berry W, Gawande AA: Designing the WHO Safe Childbirth Checklist program to improve quality of care at childbirth. *Int J Gynaecol Obstet* 122(2):164–168, 2013.

35. Arriaga AF, Bader AM, Wong JM, Lipsitz SR, Berry WR, Ziewacz JE, Hepner DL, Boorman DJ, Pozner CN, Smink DS, et al. Simulation-based trial of surgical-crisis checklists. *N Engl J Med* 368(3):246–253, 2013.

36. Pronovost P: Interventions to decrease catheter-related bloodstream infections in the ICU: the Keystone Intensive Care Unit Project. *Am J Infect Control* 36(10):S171.e1–S171.e5, 2008.