SHOCK, Vol. 31, No. 6, pp. 574–580, 2009

CAN WE IMPROVE THE CLINICAL UTILITY OF RESPIRATORY RATE AS A MONITORED VITAL SIGN?

Liangyou Chen,* Andrew T. Reisner,*† Andrei Gribok,* Thomas M. McKenna,* and Jaques Reifman*

**Bioinformatics Cell, Telemedicine and Advanced Technology Research Center (TATRC), US Army Medical Research and Materiel Command, Fort Detrick, Maryland; and †Massachusetts General Hospital Department of Emergency Medicine, Boston, Massachusetts*

Received 22 Sept 2008; first review completed 10 Oct 2008; accepted in final form 13 Oct 2008

ABSTRACT—Respiratory rate (RR) is a basic vital sign, measured and monitored throughout a wide spectrum of health care settings, although RR is historically difficult to measure in a reliable fashion. We explore an automated method that computes RR only during intervals of clean, regular, and consistent respiration and investigate its diagnostic use in a retrospective analysis of prehospital trauma casualties. At least 5 s of basic vital signs, including heart rate, RR, and systolic, diastolic, and mean arterial blood pressures, were continuously collected from 326 spontaneously breathing trauma casualties during helicopter transport to a level I trauma center. "Reliable" RR data were identified retrospectively using automated algorithms. The diagnostic performances of reliable versus standard RR were evaluated by calculation of the receiver operating characteristic curves using the maximum-likelihood method and comparison of the summary areas under the receiver operating characteristic curves (AUCs). Respiratory rate shows significant data-reliability differences. For identifying prehospital casualties who subsequently receive a respiratory intervention (hospital intubation or tube thoracotomy), standard RR yields an AUC of 0.59 (95% confidence interval, 0.48 – 0.69), whereas reliable RR yields an AUC of 0.67 (0.57 – 0.77), $P < 0.05$. For identifying casualties subsequently diagnosed with a major hemorrhagic injury and requiring blood transfusion, standard RR yields an AUC of 0.60 (0.49 – 0.70), whereas reliable RR yields 0.77 (0.67 – 0.85), $P < 0.001$. Reliable RR, as determined by an automated algorithm, is a useful parameter for the diagnosis of respiratory pathology and major hemorrhage in a trauma population. It may be a useful input to a wide variety of clinical scores and automated decision-support algorithms.

KEYWORDS—Data quality, respiratory rate, physiological monitoring, decision-support, trauma, vital signs

INTRODUCTION

Background

The respiratory rate (RR) is a classic vital sign, measured and monitored throughout a wide spectrum of health care settings. However, clinical measurements are frequently inaccurate because of poor technique and natural ambiguities inherent in this measurement (1–7). Therefore, a superior method of computing RR might be useful, with potential application to a wide range of medical conditions, spanning respiratory, infectious, neurological, and metabolic pathologies. Given the fundamental importance of RR, an improved measurement method might enhance disease diagnosis, prognosis (i.e., clinical scores), triage, and monitoring (i.e., vigilant detection of unexpected deterioration).

Importance

In this investigation, we explored an automated RR measurement method in a population of prehospital trauma casualties. We investigated the method's capability to identify patients with major respiratory and circulatory pathologies, which is a primary focus of early trauma care (the "ABCs"). Trauma is the leading cause of death for Americans aged 1 to 44 years (8). Superior physiological information related to respiratory and circulatory pathologies would be useful for triage (i.e., prioritization of casualties based on injury severity and determination of whether to send the casualty to a specialized trauma center or a local medical facility), resource mobilization (e.g., activation of trauma teams and operating rooms at a receiving trauma center), and therapeutic decision making.

Goals of this investigation

The new method is an automated algorithm that identifies consistent, rhythmic, and clean respiratory patterns, and computes RR exclusively from those intervals. We performed a retrospective analysis of archived Propaq (Protocol Systems; Beaverton, Ore) monitor data, where the RR is measured by impedance pneumography derived from a standard electrocardiogram (ECG). We compared the diagnostic value of standard Propaq RR measurements versus "reliable" RR data (determined by the automated method), as discriminators of which casualties required major respiratory inventions and major hemorrhage. We hypothesized that the automated method would be diagnostically superior to the conventional method, because there is less ambiguity in the measurement of RR during these intervals. If true, this finding would suggest that we could improve the automated monitoring of a

Address reprint requests to Jaques Reifman, Bioinformatics Cell, TATRC, US Army Medical Research and Materiel Command, ATTN: MCMR-TT, Building 363 Miller Drive, Fort Detrick, MD 21702. E-mail: jaques.reifman@us.army.mil.

This work was partially supported by the Combat Casualty Care Directorate of the US Army Medical Research and Materiel Command, Fort Detrick, MD.

This work was performed at Bioinformatics Cell, TATRC, US Army Medical Research and Materiel Command, Fort Detrick, MD.

DOI: 10.1097/SHK.0b013e318193e885

key physiological parameter. The general principle—focusing on clear, regular, consistent breathing intervals—might be applicable to a range of respiration monitoring modalities, for example, capnometry, nasal thermometry, extensometry, and so on.

MATERIALS AND METHODS

Study design and settings

This is a retrospective study based on physiological time-series data collected from 898 trauma-injured patients during transport by medical helicopter from the scene of injury to the level I unit at the Memorial Hermann Hospital in Houston, Tex (9). Additional attribute data were collected retrospectively via chart review (Table 1). The time-series variables were measured by Propaq 206EL vital-sign monitors (Protocol Systems) during transport, downloaded to an attached personal digital assistant, and ultimately stored in our database (10). The physiological data included the ECG (measured at 182 Hz) and the corresponding monitor-computed heart rate (HR); a respiratory waveform (an impedance pneumogram, derived using the ECG leads and recorded at 23 Hz) and the corresponding monitor-computed RR; and noninvasive measurements of systolic blood pressure (SBP), mean arterial blood pressure (MAP), and diastolic blood pressure (DBP) (using a standard oscillometric device) collected intermittently at multiminute intervals. The patient attribute data included demographics, injury descriptions, prehospital interventions, and hospital treatments. There are 100 attribute parameters for each patient, and these data have undergone prior analysis (11–14). Data were collected and analyzed with approval of the local human-subjects institutional review board.

Data processing

Recent reports described a computer algorithm that evaluated the reliability of RR measurements made by a Propaq transport monitor (which uses impedance pneumography via the ECG leads) (15). The investigational algorithm identifies rhythmic and clean waveform segments both by evaluating the waveform itself and by computing an independent RR that is compared with the Propaq's RR output. Ultimately, this algorithm rates the RR with an integer, from least ("0") to most ("3") reliable. In this investigation, we treat any RR rated by the algorithm 2 or greater as "reliable." The functionality of the RR reliability algorithm is demonstrated with examples from our database, shown in Figure 1.

Selection of participants

From the overall database, we included patients with at least 5 s of consecutive reliable RR data and excluded patients with prehospital intubation. We also excluded patients who did not have at least 5 s of reliable HR, SBP, and DBP data (13, 16). Reliable RR for each casualty was calculated as the average of all segments of at least 5 s of consecutive reliable RR data recorded during patient transport. Standard RR was the average of all nonzero RR Propaq data recorded during patient transport, without regard to their data quality (which might have included a mixture of reliable and unreliable data points and could be computed without any data reliability algorithm).

It is likely that RR patterns could be altered or confounded by injuries or treatments that mechanically affect respiration, independent of any possible hemorrhage. Therefore, we compared reliable RR versus standard RR in the subset of patients that had injuries to the thorax identified by a search of abbreviated injury-scale codes in our database. It is also likely that RR patterns could be altered or confounded by altered mental status, but there were few casualties with a reduced Glasgow (<14) who were spontaneously breathing for any meaningful subset analysis (four with a major hospital respiratory intervention, five with major hemorrhage).

Outcome measures

In terms of outcomes, major respiratory interventions were defined as patients who received emergency department intubation or subsequent tube thoracotomy. Patients who did not receive either of these interventions comprised the control group for the major respiratory intervention population. Patients with major hemorrhage were defined as those who received a blood transfusion in the hospital and also had documented injuries that were consistent with hemorrhage, as determined by chart review. These specific injuries were one or more of the following: (*a*) laceration of solid organs, (*b*) thoracic or

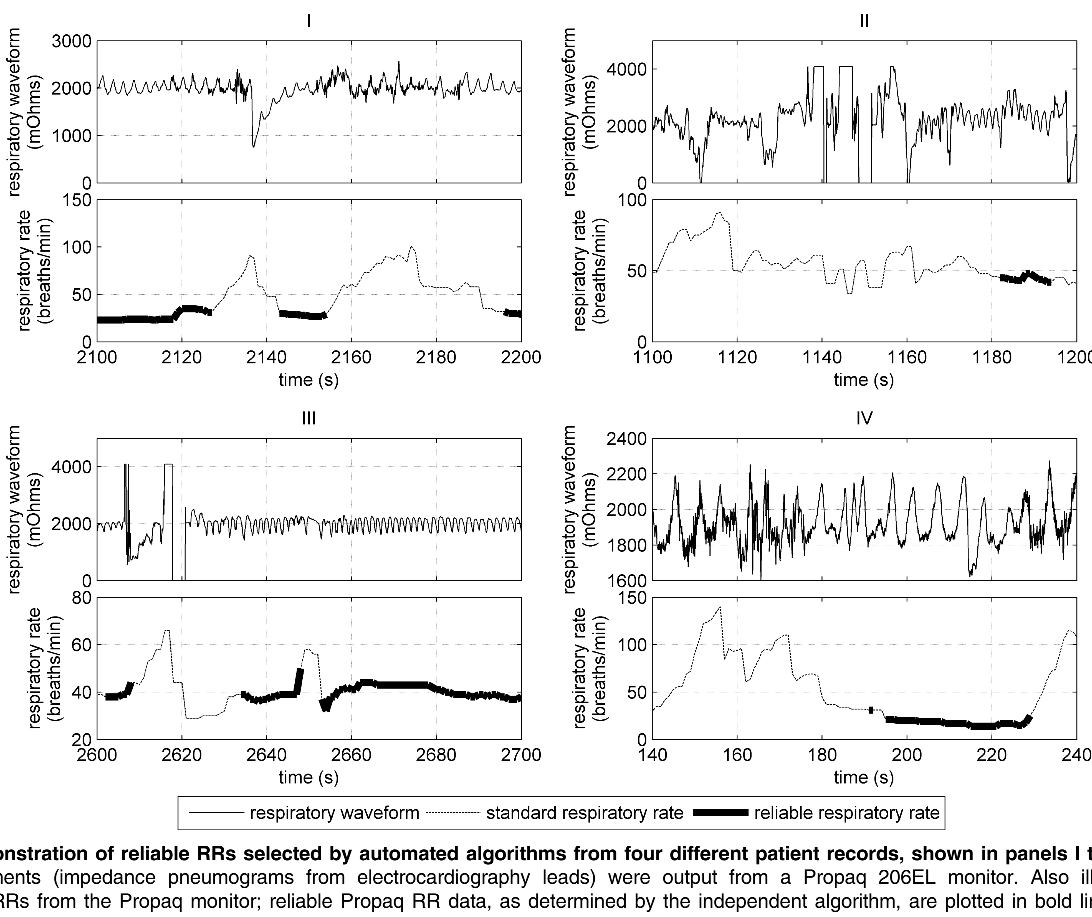

FIG. 1. **Demonstration of reliable RRs selected by automated algorithms from four different patient records, shown in panels I to IV**. Respiratory waveform segments (impedance pneumograms from electrocardiography leads) were output from a Propaq 206EL monitor. Also illustrated are the corresponding RRs from the Propaq monitor; reliable Propaq RR data, as determined by the independent algorithm, are plotted in bold lines, whereas RR measurements of questionable reliability are plotted in thin lines.

abdominal hematomas, (*c*) explicit vascular injury and operative repair, or (*d*) limb amputation. Because we are assessing the performance of vital signs, we could not use vital-sign criteria in our actual definition of major hemorrhage. Using documented injuries and therapies, "perfect" retrospective classification is impossible, but these objective outcome definitions are clinically reasonable and provided a very fair point of objective comparison for standard versus reliable RRs. The remaining patients comprised the control cases for the major hemorrhage population.

Methods of measurement

The diagnostic performances of reliable and standard vital signs were evaluated by constructing receiver operating characteristic (ROC) curves and calculating the areas under the curve (AUCs) for each ROC curve. We used the ROCKIT freeware (University of Chicago) (17) for these analyses, which automatically partitions each variable into at most 20 intervals for the ROC curve construction (18). ROCKIT assumes a binormal ROC model, that is, data for each of the decision outcomes (respiratory intervention or hemorrhage versus their respective controls) are considered to be normally distributed. Under this assumption, each ROC curve is transformed into a straight line on the normal-deviate axes (18), whose ordinate intercept "a" and slope "b" are estimated by the maximum-likelihood method. The AUC is computed based on its mathematical relationship with a and b (18, 19). The ROC curves estimated from this method are smoother than empirically evaluated ROC curves and can better represent the relationship between vital-sign variables and the decision outcomes (18, 19). We performed univariate ROC analyses on each of the vital signs and report the estimated AUC and corresponding 95% confidence interval. Statistical tests of significance between ROC curves were performed within ROCKIT, which uses the z-score test to compare the difference between the areas under two ROC curves (20). All statistical differences were based on paired tests (standard versus reliable RR), where we report the two-tailed P values. A significance level of 0.05 is used in this study. Because all the statistical tests address the same underlying hypothesis, that is, that reliable RR is more clinically useful than standard RR, we did not make any explicit corrections for multiple comparisons.

Benchmark versus other vital signs

To explore how reliable RR might enable novel diagnostic applications, we benchmarked the ROC AUC of other vital signs for the prehospital identification of patients with major hemorrhage. For each patient, HR, SBP, MAP, and DBP were calculated as the average of reliable data recorded during patient transport, using previously reported reliability measures for those vital signs (13, 16). (The HR reliability algorithm, which evaluates the ECG waveform and considers if there is agreement between several different methods of computing HR, was previously compared versus blinded human experts for several hundred ECG waveform excerpt (16). When the HR algorithm identified reliable data, in 97% of the cases, blinded human experts concurred that the waveform was clean and, in 100% of those cases, concurred with the monitor's reported HR. The blood pressure reliability algorithm compares the HR measured by an oscillometric noninvasive blood-pressure cuff versus the ECG HR and also checks that the relationships between SBP, MAP, and DBP are physiologic (13). Reliable SBP, as determined by this algorithm, has been found to be statistically superior to unreliable SBP, as a predictor of major hemorrhage.) In addition, we computed three simple multivariate metrics, based on the combination of vital signs, to explore the interaction of potentially independent variables. We computed the AUC for (*a*) the shock index, defined as the ratio of HR and SBP; (*b*) arterial pulse pressure (PP), defined as the difference between SBP and DBP; and (*c*) the breath index, defined as the ratio of RR and PP. We used shock index as a basis of comparison for the multivariate metrics because it is arguably the best known multivariate discriminator for major hemorrhage (21, 22). The PP was selected because, by report, it has value in the diagnosis of hemorrhagic hypovolemia (9, 21). The breath index is a novel metric, not previously reported, which scales the RR relative to the PP (analogous to the shock index, which scales HR to SBP). These results provide general context for interpreting the reliable RR versus standard RR results. Because our sample size did not permit multiple comparisons between all vital signs and vital-sign combinations, no formal hypothesis testing was conducted.

RESULTS

Characteristics of study subjects

Table 1 summarizes the attributes of the study population. In general, patients with reliable vital signs distribute similarly as the total population in terms of sex, age, and injury type. The study population shows a lower mortality rate after exclusion of prehospital intubated patients (99 patients).

Main results

Table 2 compares reliable versus standard RR for the prediction of in-hospital respiratory intervention and the identification of major hemorrhage. Overall, reliable RR is statistically superior to standard RR for both outcomes.

Sensitivity analyses

Reliable RR trends toward superiority in one subgroup (Table 2, identifying the need for in-hospital respiratory interventions in patients with thoracic injuries), whereas it is clearly statistically superior in the other three subgroups.

In Table 3, we illustrate the ROC AUCs of other basic vital signs and vital-sign combinations, for the prehospital identification of major hemorrhage. The RR, HR, and SBP have similar AUCs. Note that there is a trend toward higher AUC when vital signs were used in combination, and incorporating both the PP and the RR yielded the highest AUC.

Finally, as shown in Figure 2, we compared the distribution of standard and reliable RR in patients of each outcome versus their respective controls. Reliable RR demonstrates fewer extreme cases (e.g., >60 breaths/min) and shows better separation than standard RR for the discrimination of both outcomes.

DISCUSSION

In this investigation, the underlying hypothesis was that *irregular* patterns in the pneumogram waveform yield *unreliable* (i.e., nondiagnostic) measurements of RRs, whereas regular, consistent, and clean waveforms produce *reliable* (i.e., diagnostic) RRs. In practice, this approach should exclude waveforms corrupted by measurement artifact, but also

Table 1. **Demographics and population selection**

Characteristics	Overall database	Patients with reliable vital signs*	Study population†
Population size, n	898	425	326
Male, n (%)	660‡ (73)	317 (75)	247 (76)
Female, n (%)	234‡ (26)	108 (25)	79 (24)
Mean (SD) age, yrs	37 (16)	38 (15)	38 (16)
Blunt injury, n (%)	778 (87)	371 (87)	284 (87)
Mortality, n (%)	94 (10)	38 (9)	8 (2)
Intubated, n (%)	201 (22)	99 (23)	0 (0)
Major respiratory intervention,§ n (%)	102 (11)	51 (12)	33 (10)
Major hemorrhage,‖ n (%)	94 (10)	50 (12)	33 (10)

*Patients with at least 5 s of consecutive reliable RR, HR, and SBP, DBP, and MAP vital signs.
†Patients with reliable vital signs and were spontaneously breathing (i.e., not intubated during transport).
‡Four patients had no assigned sex in the total population.
§Received major respiratory live-saving interventions, including in-hospital intubation and chest tube.
‖Received blood transfusion in the emergency room and also had documented injuries that were consistent with major hemorrhage.

Table 2. **Comparison of standard versus reliable RRs for the prediction of major in-hospital respiratory intervention (Resp., A) and the identification of major hemorrhage (Heme., B)**

A. Respiratory intervention (Resp.)					
	Control (cases)	Resp. (cases)	Standard RR* (AUC†)	Reliable RR‡ (AUC)	P§
Population	293	33‖	0.59 (CI, 0.48 – 0.69)	0.67 (CI, 0.57 – 0.77)	0.03
Thoracic injury	68	25	0.52 (CI, 0.38 – 0.66)	0.63 (CI, 0.51 – 0.75)	n.s.
Nonthoracic	225	8	0.56 (CI, 0.37 – 0.74)	0.73 (CI, 0.49 – 0.89)	0.04
B. Major hemorrhage (Heme.)					
	Control (cases)	Heme. (cases)	Standard RR (AUC)	Reliable RR (AUC)	P
Population	293	33‖	0.60 (CI, 0.49 – 0.70)	0.77 (CI, 0.67 – 0.85)	0.0003
Thoracic injury	77	16	0.56 (CI, 0.39 – 0.71)	0.76 (CI, 0.61 – 0.87)	0.0020
Nonthoracic	216	17	0.60 (CI, 0.45 – 0.73)	0.79 (CI, 0.66 – 0.89)	0.0002

*Variables are computed as the mean of all nonzero RRs.
†Area under the ROC curve.
‡Variables are computed as the mean of all good-quality RRs.
§P value of two-tailed paired comparison of AUCs between standard and reliable RR.
‖16 cases had both major in-hospital respiratory intervention and major hemorrhage.
CI indicates confidence interval; n.s., not significant.

might exclude physiological breathing patterns that were truly irregular. Our major finding was that RR computed from the smooth, regular, and rhythmic breathing (as assessed by our computer algorithm), which we termed *reliable RR*, was significantly more diagnostic than RR from other noisy or arrhythmic intervals, for diagnosing both respiratory and circulatory pathologies. We conclude that these special breathing segments are, on average, physiologically more informative and provide superior clinical information.

This finding is particularly notable given historical issues related to RR (6). It is a difficult vital sign to measure, because partial breaths (which are common) must be either counted or discounted, and breathing is often irregular (e.g., when speaking or swallowing). Also, breathing patterns are volatile, altered by emotions, conscious control, or even the patient's awareness that RR is being measured. There is no widely accepted electronic tool for measuring RR, and caregivers' measurements are often unreliable because of poor technique. For these reasons, Lovett et al. (6). referred to RR as the "vexatious vital." Standard bedside monitors electronically measure RR by impedance pneumography derived from continuous electrocardiography. However, even in a highly controlled intensive care unit setting in which many patients are sedated or paralyzed, the RR reported by bedside monitors is frequently inaccurate (7). This may be because of the inherent ambiguities in measuring RR, or perhaps because the ECG is vulnerable to serious artifact caused by patient movement, imperfect ECG lead attachment, and muscle activity (1, 23).

We speculate that this automated reliable RR method may offer diagnostic and prognostic value in the evaluation of numerous pathologies (e.g., respiratory disorders, metabolic disorders, etc.). For example, RR is an input to the Pneumonia Patient Outcomes Research Team (PORT) score (24) for the prognosis of patients with pneumonia and to several prehospital trauma severity indices, including the Trauma Score (25) and the Prehospital Index (26). In a medical helicopter (the setting of this investigation), caregivers cannot even hear breath sounds (4), so an improved method of automatically monitoring RR would be all the more valuable. Indeed, because accurate RR measurements may be broadly useful, it has been suggested that it is imperative to develop improved electronic methods of measuring RR (6). The method used in this investigation is advantageous in that no additional hardware (besides a standard patient monitor, in this case a Propaq monitor) is required, and its diagnostic capability seems quite promising. A number of alternative options to monitor RR are available, including capnometry, pneumotachography, nasal thermometry, and extensometry (6). The general strategy of seeking regular, rhythmic, and clean breathing intervals might be clinically valid for these sensor modalities as well.

RR and in-hospital respiratory interventions

Reliable RR was statistically superior to standard RR for this outcome, which validates our hypothesis. It was interesting

Table 3. **Performance evaluation of reliable vital signs for the diagnosis of major hemorrhage in trauma casualties**

Reliable vital sign*	AUC†
Breath index (RR/PP‡)	0.85 (CI, 0.77 – 0.91)
PP	0.78 (CI, 0.69 – 0.86)
Shock index (HR/SBP)	0.78 (CI, 0.67 – 0.86)
RR	0.77 (CI, 0.67 – 0.85)
HR	0.74 (CI, 0.63 – 0.83)
SBP	0.71 (CI, 0.61 – 0.80)
MAP	0.60 (CI, 0.49 – 0.71)
DBP	0.55 (CI, 0.43 – 0.67)

Vital-sign data were computed as the means of all good-quality data collected during patient transport. The quality of the vital-sign data was determined by computer algorithms.
*Variables were computed as the mean of all good-quality data during transport.
†Area under the ROC curve.
‡PP = SBP − DBP.

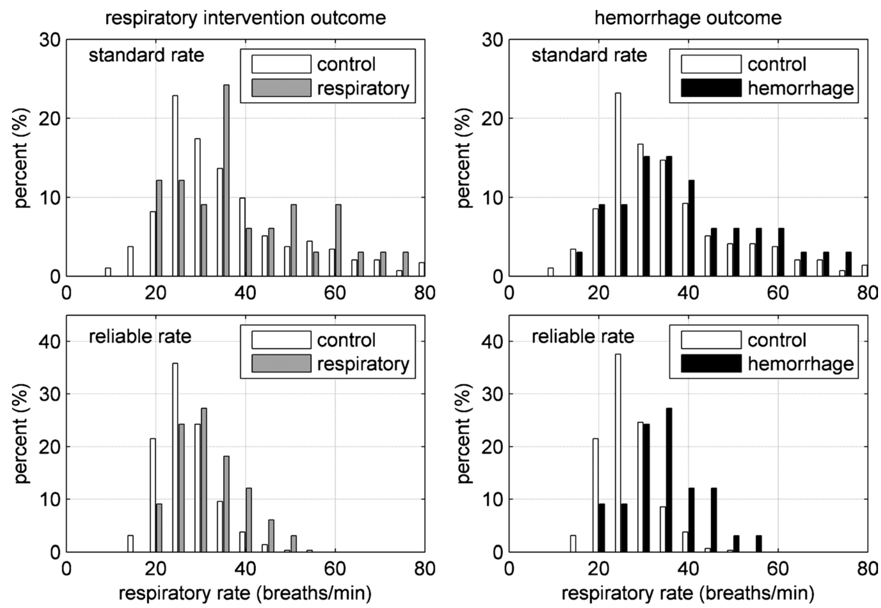

Fig. 2. **Distribution of standard (upper panels) and reliable (lower panels) RRs in high acuity (i.e., with major in-hospital respiratory intervention [left panels] and with major hemorrhage [right panels]) versus control patients.**

to note that the ROC AUCs were mediocre, however. When we excluded patients with chest trauma, reliable RR yielded a trend toward better AUCs. When we examined the subpopulation with documented thoracic injuries, the AUCs were the lowest in our study. It may be that chest injury has an inconsistent effect on RR. We speculate that the respiratory impairment of major chest injury would drive tachypnea, whereas the pain associated with chest injury might retard tachypnea. Although reliable RR seems superior to standard RR for identifying this outcome, its clinical utility may be modest.

RR and major hemorrhage

One interesting finding of this report is the notable value of reliable RR in the diagnosis of major hemorrhage with prehospital physiological data. This finding is one example of how a superior method of measuring RR might have wide clinical utility, especially when combined with additional clinical data. Specifically, in our univariate analysis, reliable RR was diagnostically quite similar to HR and SBP in the diagnosis of major hemorrhage, in terms of ROC AUCs. (By contrast, diagnosing hemorrhage using standard RR, which included noisy or uneven breathing intervals, was barely better than flipping a coin—AUC = 0.60.) Moreover, when scaled by PP, reliable RR provided the highest AUC, as shown in Table 3, although these exploratory findings were not subjected to formal statistical testing and may represent, to some extent, random variability. Future investigation is warranted.

It is not surprising that reliable RR was useful in diagnosing hemorrhage, when one considers decades of prior physiological laboratory research. Based on a feline model, it is known that hemorrhage may induce tachypnea, via a reflex mediated by the carotid body chemoreceptors. A reduction in blood pressure or increase in peripheral vasoconstriction leads to a pronounced reduction of blood flow and oxygen delivery to the chemoreceptors, producing "stagnant hypoxia" within the chemoreceptors (i.e., local tissue hypoxia caused by reduced perfusion). The chemoreceptors then stimulate the medullary respiratory center and trigger tachypnea. The carotid body chemoreceptor serves as a "bellwether" of impaired global circulation because it is exceptionally sensitive to any reduction in perfusion, in terms of developing local tissue hypoxia (27–31). The chemoreceptors may further be excited by metabolic acidosis (specifically, lactic acidosis caused by circulatory shock and perhaps an increase in global metabolic activity mediated by epinephrine) (32). Another respiratory reflex, originating in the arterial baroreceptors, stimulates respiration when blood pressures fall, although this seems to be of secondary importance (27).

Limitations

Like any vital sign, RR must be interpreted in context, and clinical judgment or additional clinical information inevitably enhances its clinical utility. In the case of using RR to detect major hemorrhage, it is possible that pain or fear, common in a trauma population, could be sources of "false positives," because epinephrine and norepinephrine alone can stimulate tachypnea (33). (However, catecholamines also raise blood pressure, which preserves carotid body perfusion and suppresses chemorecepter discharge despite vasoconstriction (34). In the laboratory, supplemental oxygen also suppresses the tachypneic response caused by isolated catecholamines) (33).

Similarly, it is possible that casualties with thoracic injuries would not mount as consistent a tachypneic response to hemorrhage, nor would patients with altered mental status. In our subpopulation analyses, examining those patients with documented thoracic trauma, we found a slight trend toward reduced ROC AUCs, which might mean that the association between chest injury and tachypnea is less consistent (note also that the confidence intervals for those AUCs are wider, reflecting the reduced sample sizes of the subpopulations). We were not able to quantitatively explore the effects of altered sensorium on RR because there were few spontaneously breathing casualties with reduced Glasgow Coma Scale (yielding meaningless AUC confidence intervals that were greater than ±0.25). In any case, when applied to our overall study population, without any special consideration for confounding factors, we found that reliable RR was statistically superior to standard RR for the diagnosis of respiratory and circulatory pathological findings.

A major limitation to our method is that it is not "on-demand." Rather, the method requires passive observation until the patient spontaneously evidences five or more consecutive seconds of clean, regular respiration. Given an average 26 minutes of prehospital data for each subject, we found reliable RR data in only 57% of these cases. This paucity of reliable RR data was primarily because of noise artifacts in the pneumogram/ECG recorded during helicopter transport: when we examined the ECG waveforms from this database of prehospital Propaq records, we found that most of 7-s ECG segments contained enough noise artifact to obscure one or more QRS complexes (16). It is possible that in-hospital data would have fewer artifacts and be more usable. In the future, because the algorithm determines RR reliability automatically, a monitor could certainly indicate whenever the measured RR was unreliable. If notified that the RR is unreliable, the caregiver may be able to rectify the situation, for example, keeping the patient still and silent for several seconds, replacing loose ECG leads, and so on. Overall, this automated method may be most applicable for applications that are not time-sensitive (e.g., PORT scoring for pneumonia patients) rather than applications that are time-sensitive (e.g., monitoring for respiratory arrest, initial triage evaluations, etc.). A hybrid solution may be valuable, in which reliable RR is used whenever it is available, but if not, standard RR is used instead.

The final limitation is the retrospective nature of this analysis. Moreover, there is a technical barrier to implementing these algorithms so that they function in real time as part of prospective investigation and clinical dissemination. We have started to develop a platform that can run, in real time, our RR reliability algorithm, as well as a wide range of additional advanced algorithms to interpret continual physiological data. Also, governmental, corporate, and medical groups are actively planning "plug-and-play" monitoring system architectures (35), which will facilitate the dissemination of novel physiological algorithms (5, 36–41). Ultimately, reliable RR might be indicative of circulatory, respiratory, infectious, neurological, and metabolic pathologies and hence valuable to a wide range of medical applications, including triage, diagnosis, prognosis (i.e., clinical scores), and automated alarms.

ACKNOWLEDGMENTS

The authors thank the staff at the University of Texas Health Science Center and COL John Holcomb and Dr Jose Salinas of the US Army Institute of Surgical Research for providing access to the Vital Signs (Trauma) database.

REFERENCES

1. Kaiser W, Findeis M: Artifact processing during exercise testing. *J Electrocardiol* (suppl 32):212–219, 1999.
2. Tsien CL, Fackler JC: Poor prognosis for existing monitors in the intensive care unit. *Crit Care Med* 25(4):614–619, 1997.
3. Edmonds ZV, Mower WR, Lovato LM, Lomeli R: The reliability of vital sign measurements. *Ann Emerg Med* 39(3):233–237, 2002.
4. Garner DC: Noise in medical helicopters. *JAMA* 266(4):515–516, 1991.
5. Yamakoshi K: Unconstrained physiological monitoring in daily living for health care. *Front Med Biol Eng* 10(3):239–259, 2000.
6. Lovett PB, Buchwald JM, Sturmann 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.
7. 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.
8. Hoyert DL, Mathews TJ, Menacker F, Strobino DM, Guyer B: Annual summary of vital statistics: 2004. *Pediatrics* 117(1):168–183, 2006.
9. Cooke WH, Salinas J, Convertino VA, et al.: Heart rate variability and its association with mortality in prehospital trauma patients. *J Trauma* 60(2): 363–370, 2006; discussion 370.
10. 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.
11. Chen L, Reisner AT, McKenna TM, Gribok A, Reifman J: Diagnosis of hemorrhage in a prehospital trauma population using linear and nonlinear multiparameter analysis of vital signs. *Conf Proc IEEE Eng Med Biol Soc* 1:3748–3751, 2007.
12. 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; discussion 828–829.
13. 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.
14. Chen L, McKenna TM, Reisner AT, Gribok A, Reifman J: Decision tool for the early diagnosis of trauma patient hypovolemia. *J Biomed Inform* 41: 469–478, 2008.
15. Chen L, McKenna TM, Reisner AT, Reifman J: Algorithms to qualify respiratory data collected during the transport of trauma patients. *Physiol Meas* 27(9):797–816, 2006.
16. 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.
17. Laboratories KR: ROCKIT. Available at: http://xray.bsd.uchicago.edu/krl/KRL_ROC/ software_index.htm. Accessed 2007.
18. Metz CE, Herman BA, Shen JH: Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Stat Med* 17(9):1033–1053, 1998.
19. Obuchowski NA: ROC analysis. *AJR Am J Roentgenol* 184(2):364–372, 2005.
20. Metz CE: Some practical issues of experimental design and data analysis in radiological ROC studies. *Invest Radiol* 24(3):234–245, 1989.
21. Birkhahn RH, Gaeta TJ, Terry D, Bove JJ, Tloczkowski J: Shock index in diagnosing early acute hypovolemia. *Am J Emerg Med* 23(3):323–326, 2005.
22. Nakasone Y, Ikeda O, Yamashita Y, Kudoh K, Shigematsu Y, Harada K: Shock index correlates with extravasation on angiographs of gastrointestinal hemorrhage: a logistics regression analysis. *Cardiovasc Intervent Radiol* 30(5): 861–865, 2007.
23. Portet F, Hernandez AI, Carrault G: Evaluation of real-time QRS detection algorithms in variable contexts. *Med Biol Eng Comput* 43(3):379–385, 2005.
24. Fine MJ, Auble TE, Yealy DM, et al.: A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 336(4): 243–250, 1997.
25. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME: A revision of the Trauma Score. *J Trauma* 29(5):623–629, 1989.
26. Koehler JJ, Baer LJ, Malafa SA, Meindertsma MS, Navitskas NR, Huizenga JE: Prehospital Index: a scoring system for field triage of trauma victims. *Ann Emerg Med* 15(2):178–182, 1986.
27. D'Silva JL, Gill D, Mendel D: The effects of acute haemorrhage on respiration in the cat. *J Physiol* 187(2):369–377, 1966.
28. Daly M, Burgh D, Lambertsen CJ, Schweitzer A: Observations on the volume of blood flow and oxygen utilization of the carotid body in the cat. *J Physiol* 125(1):67–89, 1954.
29. Landgren S, Neil E: Chemoreceptor impulse activity following haemorrhage. *Acta Physiol Scand* 23(2–3):158–167, 1951.
30. Winder CV: Combination of hypoxic and hypercapnic stimulation at the carotid body. *Am J Physiol* 136:200–206, 1942.
31. Kenney RA, Neil E: The contribution of aortic chemoceptor mechanisms to

the maintenance of arterial blood pressure of cats and dogs after haemorrhage. *J Physiol* 112(1–2):223–238, 1951.
32. Hertzman AB, Gesell R: The regulation of respiration. *Am J Physiol* 81: 563–578, 1927.
33. Joels N, White H: The contribution of the arterial chemoreceptors to the stimulation of respiration by adrenaline and noradrenaline in the cat. *J Physiol* 197(1):1–23, 1968.
34. Lee KD, Mayou RA, Torrance RW: The effect of blood pressure upon chemoreceptor discharge to hypoxia, and the modification of this effect by the sympathetic-adrenal system. *Q J Exp Physiol Cogn Med Sci* 49:171–183, 1964.
35. Goldman JM, Schrenker RA, Jackson JL, Whitehead SF: Plug-and-play in the operating room of the future. *Biomed Instrum Technol* 39(3):194–199, 2005.
36. Field MJ, Grigsby J: Telemedicine and remote patient monitoring. *JAMA* 288(4):423–425, 2002.
37. Grossman P: The LifeShirt: a multi-function ambulatory system monitoring health, disease, and medical intervention in the real world. *Stud Health Technol Inform* 108:133–141, 2004.
38. Hoyt RW, Reifman J, Coster TS, Buller MJ: Combat medical informatics: present and future. *Proc AMIA Symp* 335–339, 2002.
39. Pino E, Ohno-Machado L, Wiechmann E, Curtis D: Real-time ECG algorithms for ambulatory patient monitoring. *AMIA Annu Symp Proc 2005* pp 604–608, 2005.
40. Seo J, Choi J, Choi B, Jeong DU, Park K: The development of a nonintrusive home-based physiologic signal measurement system. *Telemed J E Health* 11(4):487–495, 2005.
41. Wendelken SM, McGrath SP, Blike GT: A medical assessment algorithm for automated remote triage. In: *25th Annual International Conference of the IEEE and Engineering in Medicine and Biology Society 2003*. Cancun, Mexico, 2003: 3630–3633.