Medical Knowledge Engineering Session

Are Standard Vital Signs Useful in Early Trauma Casualty Care? A New Look at an Old Problem	Dr. Jaques Reifman BIC/BHSAI/TATRC/MRMC
Visensia – A System for Early Warning of Clinical Deterioration	Mr. Wayne Nethercutt OBS Medical
A Platform for Testing and Comparing of Real-Time Decision-Support Algorithms in Mobile Environments	Mr. Maxim Khitrov BIC/BHSAI/TATRC/MRMC
A Vital Sign Based Decision Support System for Evaluating Post-Surgical Hospital Patients	Dr. Neal Sanders Vanderbilt University

Are Standard Vital Signs Useful in Early Trauma Casualty Care?

A New Look at an Old Problem

Jaques Reifman, Ph.D.
BIC/TATRC/MRMC

St. Pete Beach, Florida
12 August 2009

opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the U.S. Army or the U.S. Department of Defense

BIC

TATRC

Primary Goal

Develop fieldable decision-support algorithms for triage and diagnosis of trauma casualties

Requirements:

- Handle challenges of real-world data
  - Imperfect measurements and often missing data
  - Confounding effects and imperfect knowledge of outcomes
- Decision support applicable to each and every individual casualty

First Step: Develop decision-support algorithms through retrospective analysis of prehospital standard vital signs, while addressing requirements

- establish a benchmark for basic vital signs -

Strategy to Establish a Benchmark

- Systematic retrospective analysis of vital-sign data
  - Step-by-step, evidence-based analysis
  - Numerous sensitivity analyses
- Explicitly identify and address potential sources of error
  - Potential for bias
- Parsimonious approach
  - All other things being equal, the simplest solutions are the best

Prehospital Trauma Data for Retrospective Analysis

- 783 patients (~25 min of transport time/patient)
- High-frequency vital-sign waveforms (3 waveforms)
  - ECG, SpO2, Respiratory
- Low-frequency vital-sign time series (9 variables)

  Derived variables
  - ECG heart rate
  - SpO2 heart rate
  - SaO2 arterial O2 saturation
  - Respiratory rate

  Measured variables
  - NIBP (systolic, diastolic, MAP)
  - NIBP heart rate
  - End tidal CO2
- Discrete patient attribute data (100 variables)
  - Demographics, injury description, prehospital interventions, etc.

Vital signs used in decision-support algorithms

Propaq

HR
RR
SaO2
SBP
DBP

Decision Support for Hemorrhage Discrimination

Outcomes: Patients were Divided into Two Classes

- Hemorrhage class (hemorrhage, 79 patients)

  Received in-hospital blood (≥1 unit within 24 h upon arrival)

  AND

  Any documented injury associated with major hemorrhage
  - Laceration of solid organs
  - Blood collection in abdomen, thorax, or pelvis
  - Explicit vascular injury and operative repair
  - Limb amputation

- Control class (control, 600 patients)

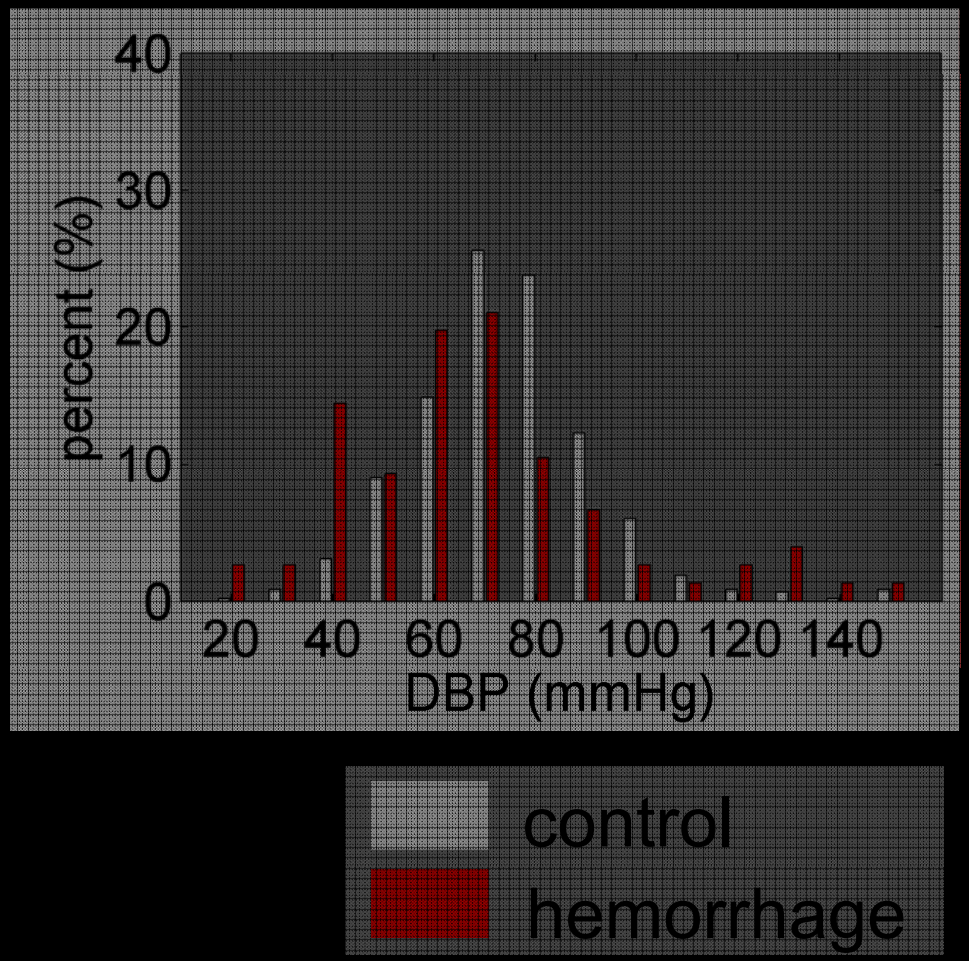
  No in-hospital blood

A Naïve Approach

Use ***first*** vital-sign measurement to identify major hemorrhage cases

Vital Sign (# of Patients)	ROC AUC (95% CI)*
HR (679)	0.58 (0.51-0.66)
RR (637)	0.51 (0.39-0.62)
SaO2 (659)	0.58 (0.46-0.69)
SBP (661)	0.71 (0.64-0.77)
DBP (661)	0.61 (0.53-0.68)

*Area under the receiver operating characteristic curve and 95% confidence interval, computed by the ROCKIT software (U. Chicago)



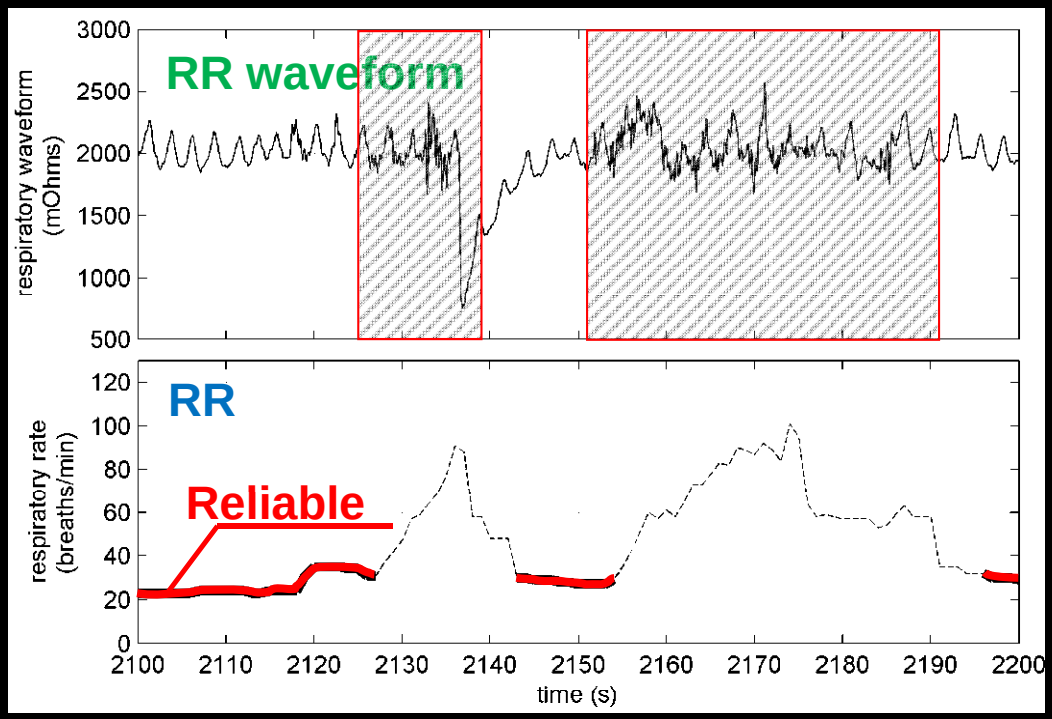
Are vital signs useful for identifying major hemorrhage?

Problem of Data Reliability

Vitals signs, e.g., respiratory rate (RR), often unreliable

- Even in an ICU, monitors usually inaccurate[1]
- Clinicians inaccurate too[2]
- On a helicopter, cannot hear breath sounds[3]

Examples from our database:



RR: 27% reliable

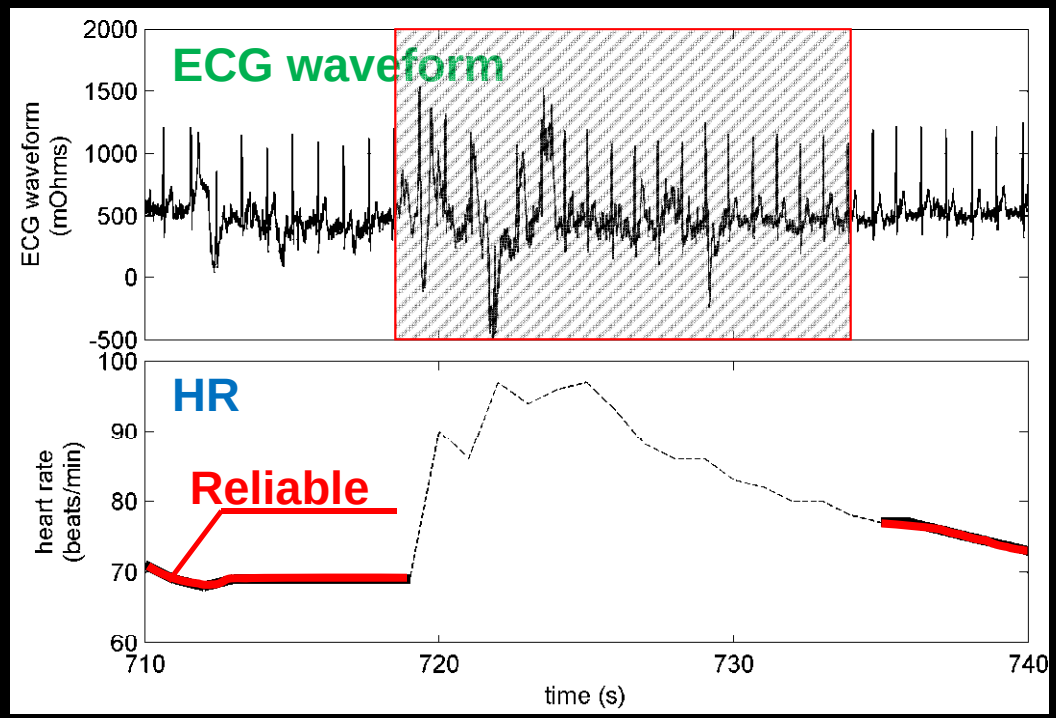
Our automated algorithm matched expert's annotations in 90% of the cases (Chen et al., *Physiol Meas* 2006)

[1]Friesdorf et al., *Int J Clin Mon Comput* (1994)
[2]Lovett et al., *Ann Emerg Med* (2005)
[3]Garner, *JAMA* (1991)

Problem of Data Reliability

Other novel algorithms for other vital signs, reliable vs. unreliable measurements

Example: unreliable ECG waveform & heart rate (HR)

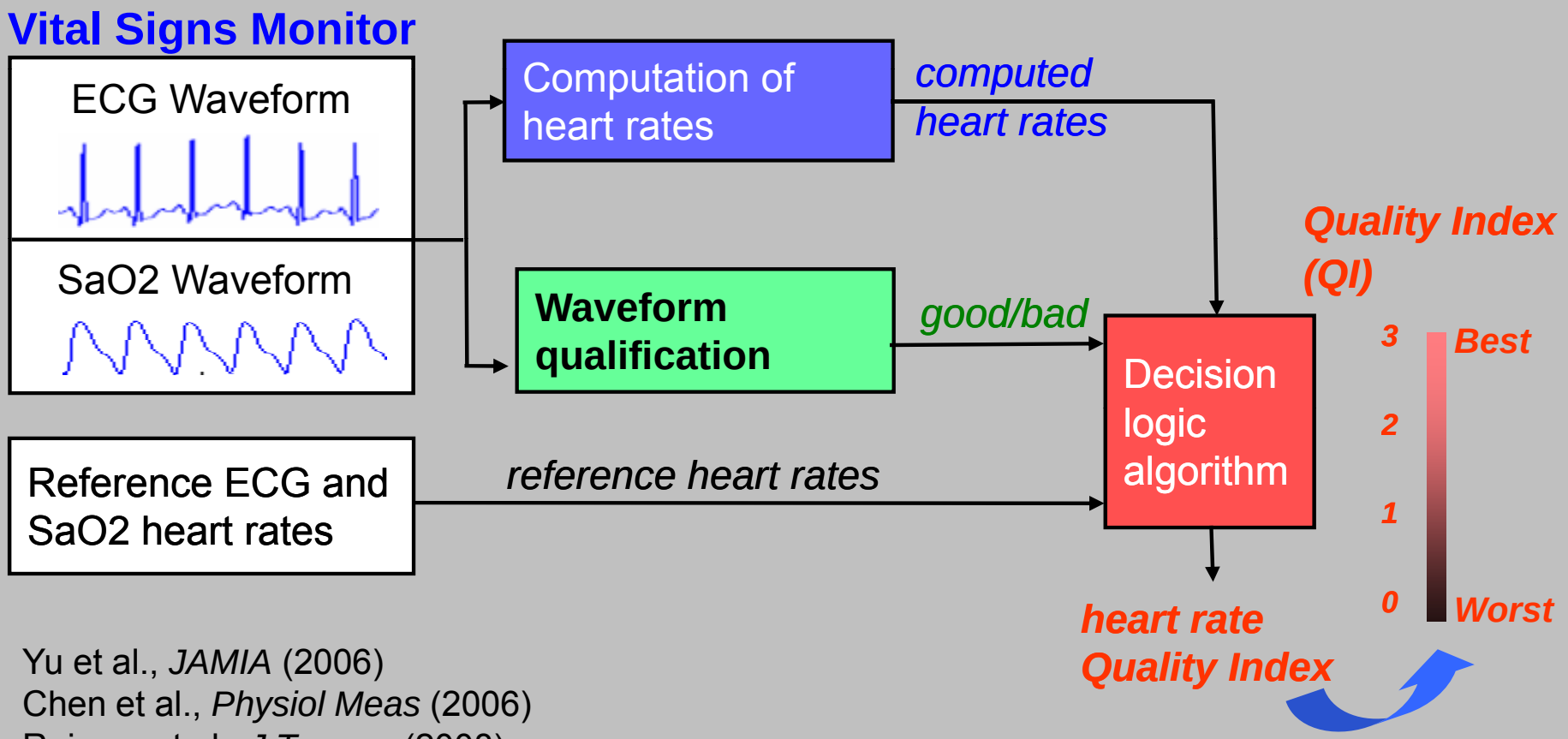


HR: 44% reliable

Our automated algorithm matched expert's annotations >90% of the cases; ~80% agreement between experts (Yu et al., *JAMIA* 2006)

Decision Logic Algorithm

Signal Validation for: ECG, SPO2 (PPG), and RR Waveforms
HR, RR, SaO2, BPs (1-sec data)



Yu et al., *JAMIA* (2006)
Chen et al., *Physiol Meas* (2006)
Reisner et al., *J Trauma* (2008)

"Reliable" vs. "Unreliable" Vital Signs
(as determined by algorithms)

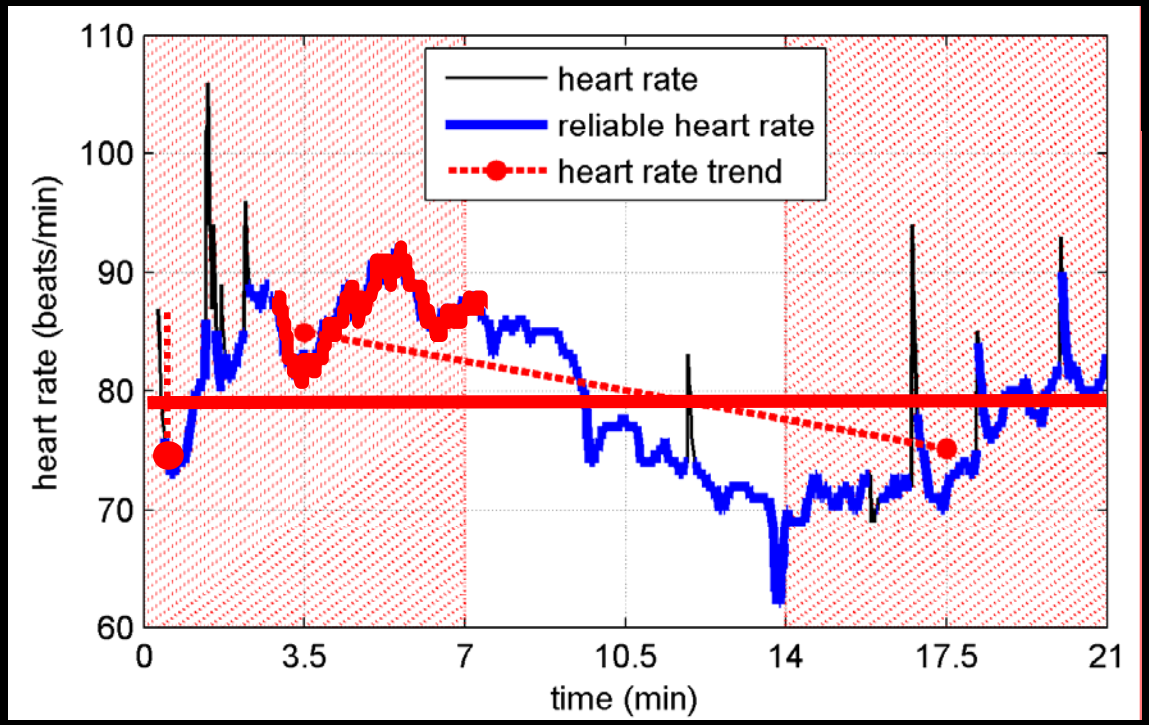
- "Reliable" vital signs superior to "unreliable" vital signs

Vital Sign	ROC AUC (95% CI)*	
	First Data	First Reliable
HR	0.58 (0.51-0.66)	**0.71 (0.64-0.77)**
RR	0.51 (0.39-0.62)	**0.64 (0.56-0.72)**
SaO2	0.58 (0.46-0.69)	0.59 (0.45-0.72)
SBP	0.71 (0.64-0.77)	0.73 (0.67-0.79)
DBP	0.61 (0.53-0.68)	0.63 (0.55-0.70)

*Each AUC was obtained from the largest available population.

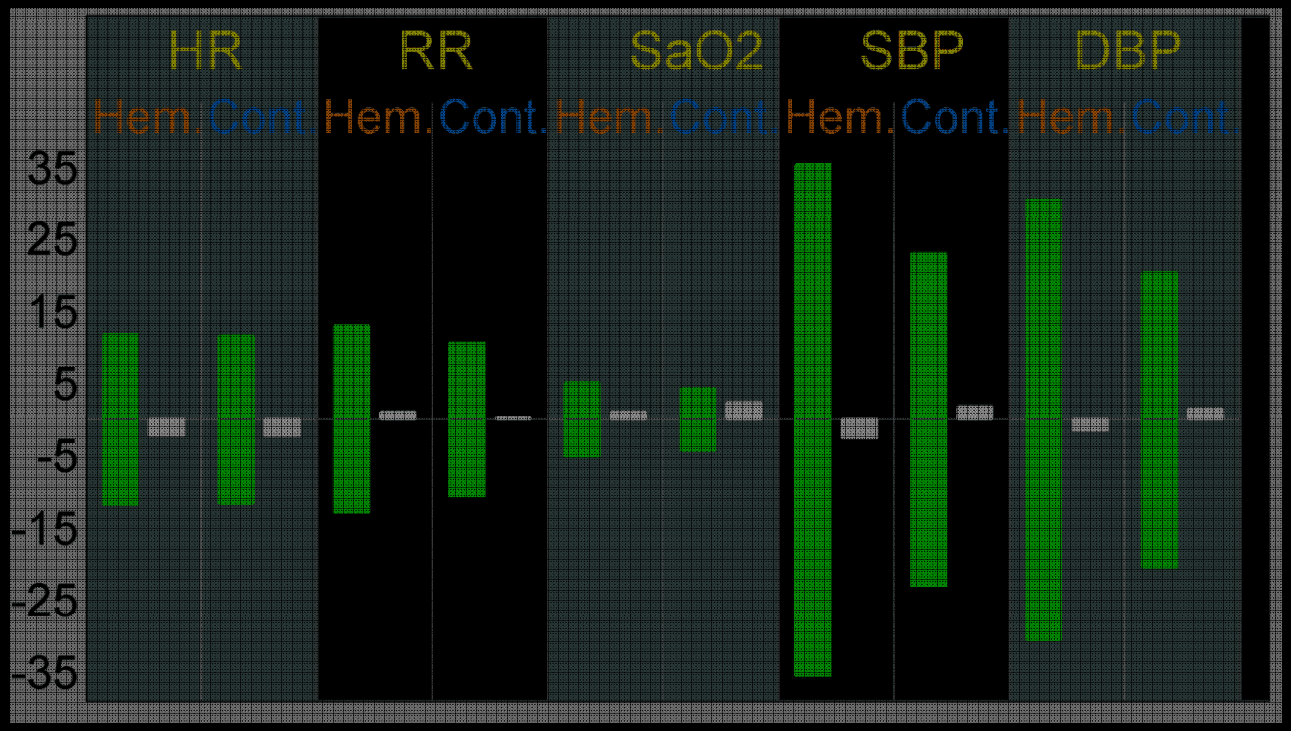
Data Analysis Through Time

- Use the first reliable data?
- Take the average?
- Use trend information?
- Exploit temporal shapes?



Data Analysis Through Time: Trend

- Average variability of vital signs over 21 minutes
- Are trends useful, or is there too much "physiologic noise"?



Chen et al., *Prehospital Emergency Care* (2009)

95% data variability (avg.)
trend (avg.)

Vital Sign	ROC AUC (95% CI)*
ΔHR	0.50 (0.39-0.61)
ΔRR	0.55 (0.41-0.68)
ΔSaO2	0.57 (0.41-0.72)
ΔSBP	0.55 (0.45-0.66)
ΔDBP	0.54 (0.44-0.65)

*Area under the receiver operating characteristic curve and 95% confidence interval.

Prehospital vital-sign trends are *not* discriminatory

Data Analysis Through Time: Avg. Reliable

- Simple but effective:

  A running average to filter-out temporal variability

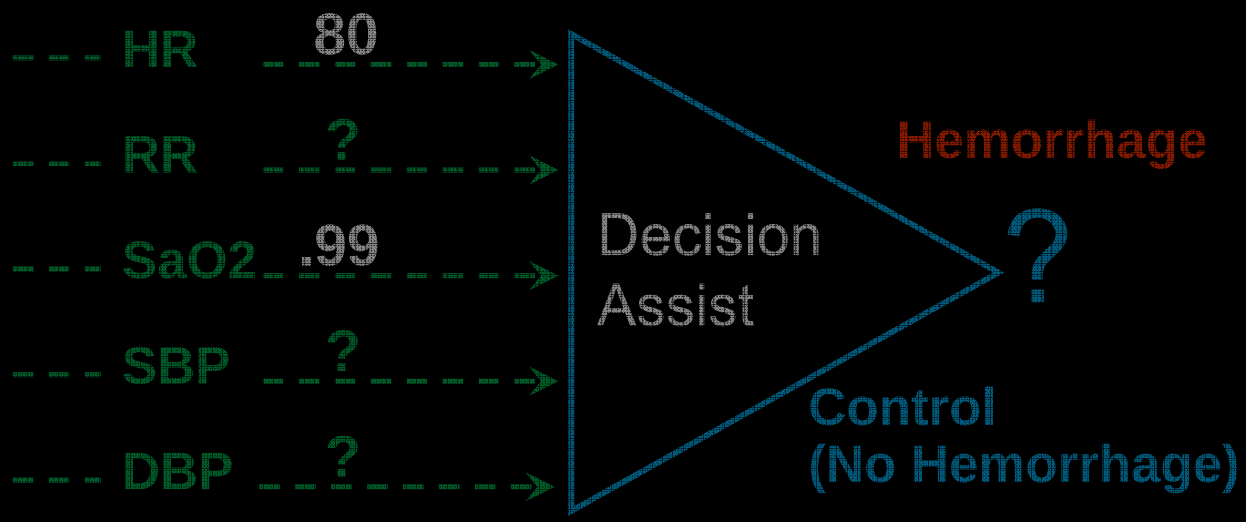
Vital Sign	ROC AUC*		
	First	First Reliable	Average Reliable
HR	0.58 (0.51-0.66)	0.71 (0.64-0.77)	0.74 (0.66-0.81)
RR	0.51 (0.39-0.62)	0.64 (0.56-0.72)	0.68 (0.59-0.76)
SaO2	0.58 (0.46-0.69)	0.59 (0.45-0.72)	0.58 (0.47-0.68)
SBP	0.71 (0.64-0.77)	0.73 (0.67-0.79)	0.78 (0.71-0.84)
DBP	0.61 (0.53-0.68)	0.63 (0.55-0.70)	0.59 (0.50-0.67)

Multivariate Analysis

Can discrimination be improved by simultaneously employing multiple vital signs?

Problem: only **402 (59%)** patients have ≥ 1 sec of reliable vital signs (all 5) **at some time** during transport

Ensemble Classifier



Chen et al., *Journal of Biomedical Informatics* (2008)

Ensemble Classifier only requires *one second* of *any* reliable vital sign during transport [674 patients (99%)]*:

AUC = 0.85 (0.80 – 0.90)

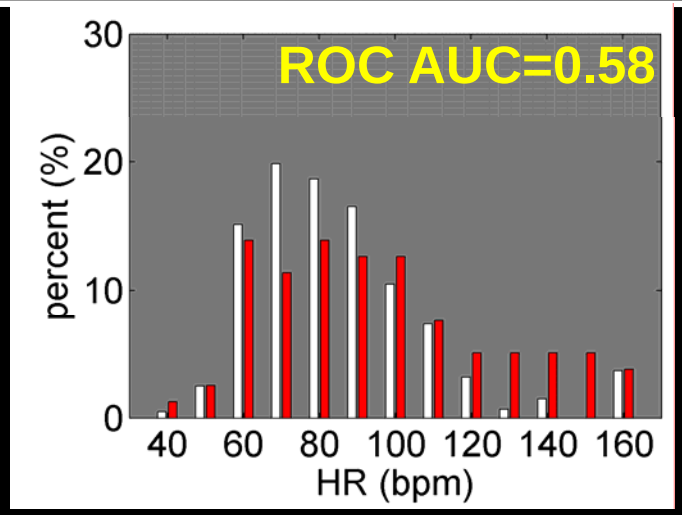
*Patients were randomly divided into non-overlapping training (50%) and testing (50%) sets with balanced classes. The average test AUC and 95% confidence interval of 100 random runs is shown.

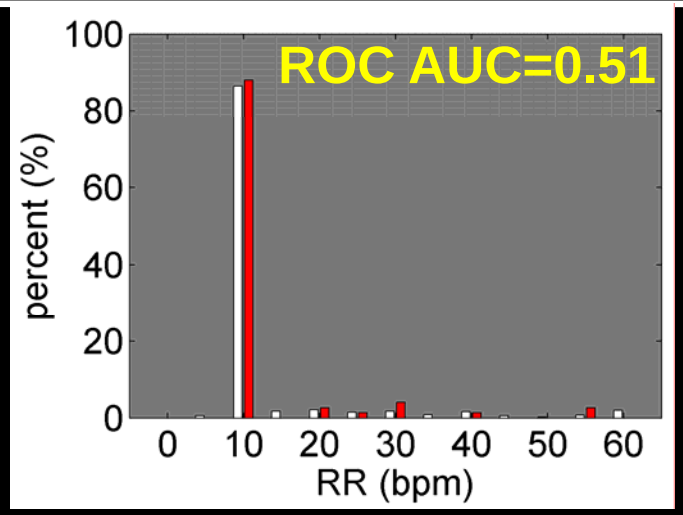
Putting All Together

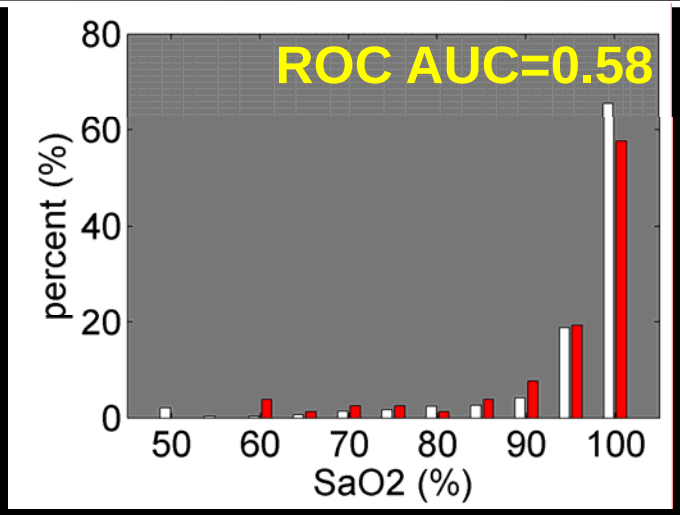
Vital Sign	ROC AUC*		
	First	First Reliable	Average Reliable
HR	0.58 (0.51-0.66)	0.71 (0.64-0.77)	0.74 (0.66-0.81)
RR	0.51 (0.39-0.62)	0.64 (0.56-0.72)	0.68 (0.59-0.76)
SaO2	0.58 (0.46-0.69)	0.59 (0.45-0.72)	0.58 (0.47-0.68)
SBP	0.71 (0.64-0.77)	0.73 (0.67-0.79)	0.78 (0.71-0.84)
DBP	0.61 (0.53-0.68)	0.63 (0.55-0.70)	0.59 (0.50-0.67)
Ensemble	0.71 (0.62-0.82)	0.77 (0.64-0.86)	0.85 (0.80-0.90)

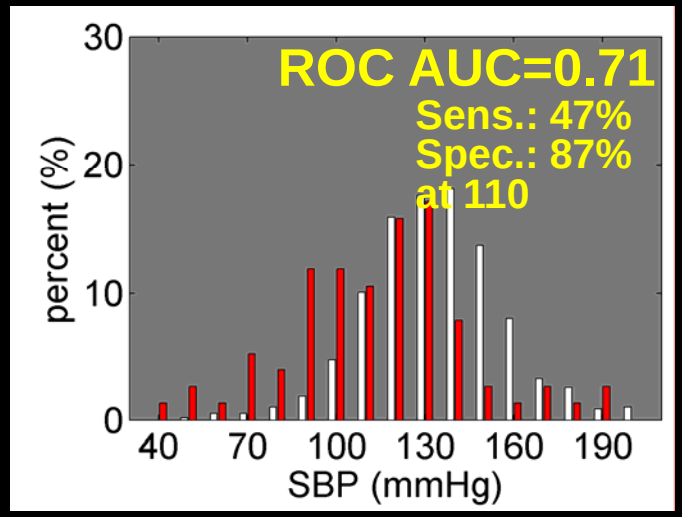
*Each AUC was obtained from the largest available population.

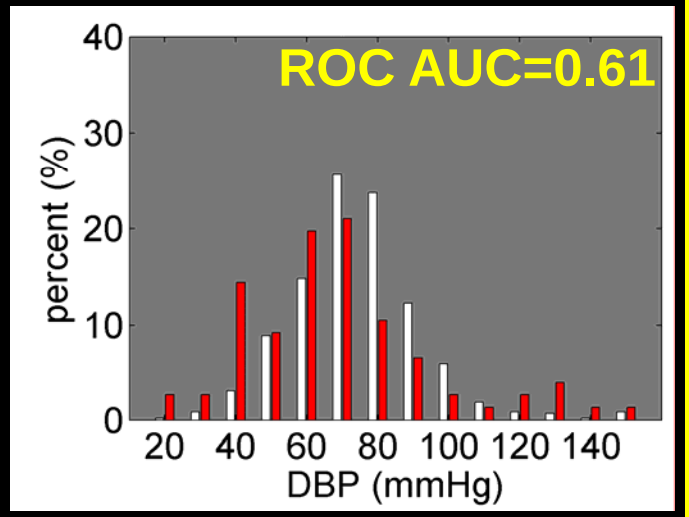
Comparison of First Data vs. Ensemble

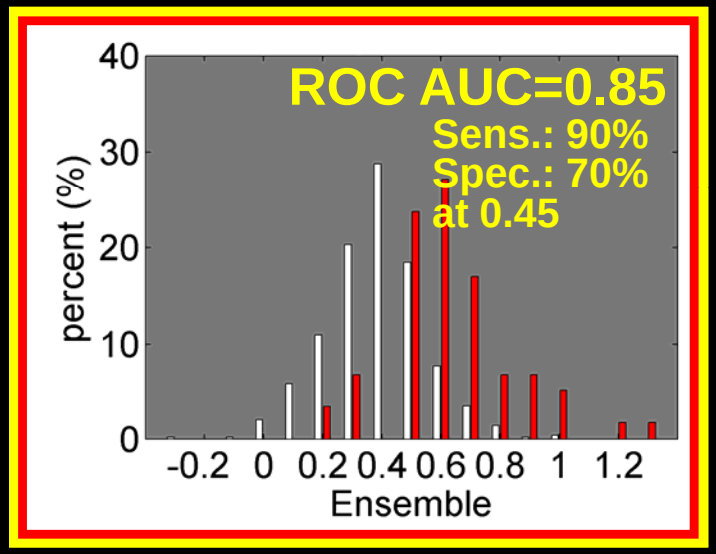












control

hemorrhage

Improved discrimination with multivariate, *Ensemble Classifier*

Summary

Established a benchmark for basic vital signs

- **For Major Hemorrhage decision support**
  - ROC AUC of 0.85 (0.80 – 0.90 with 95% CI)
- **To be considered, other sensor modalities and parameters need to surpass benchmark**

Next Steps

- **Employ decision support to predict clinical outcomes during transport of trauma patients (APPRAISE)**

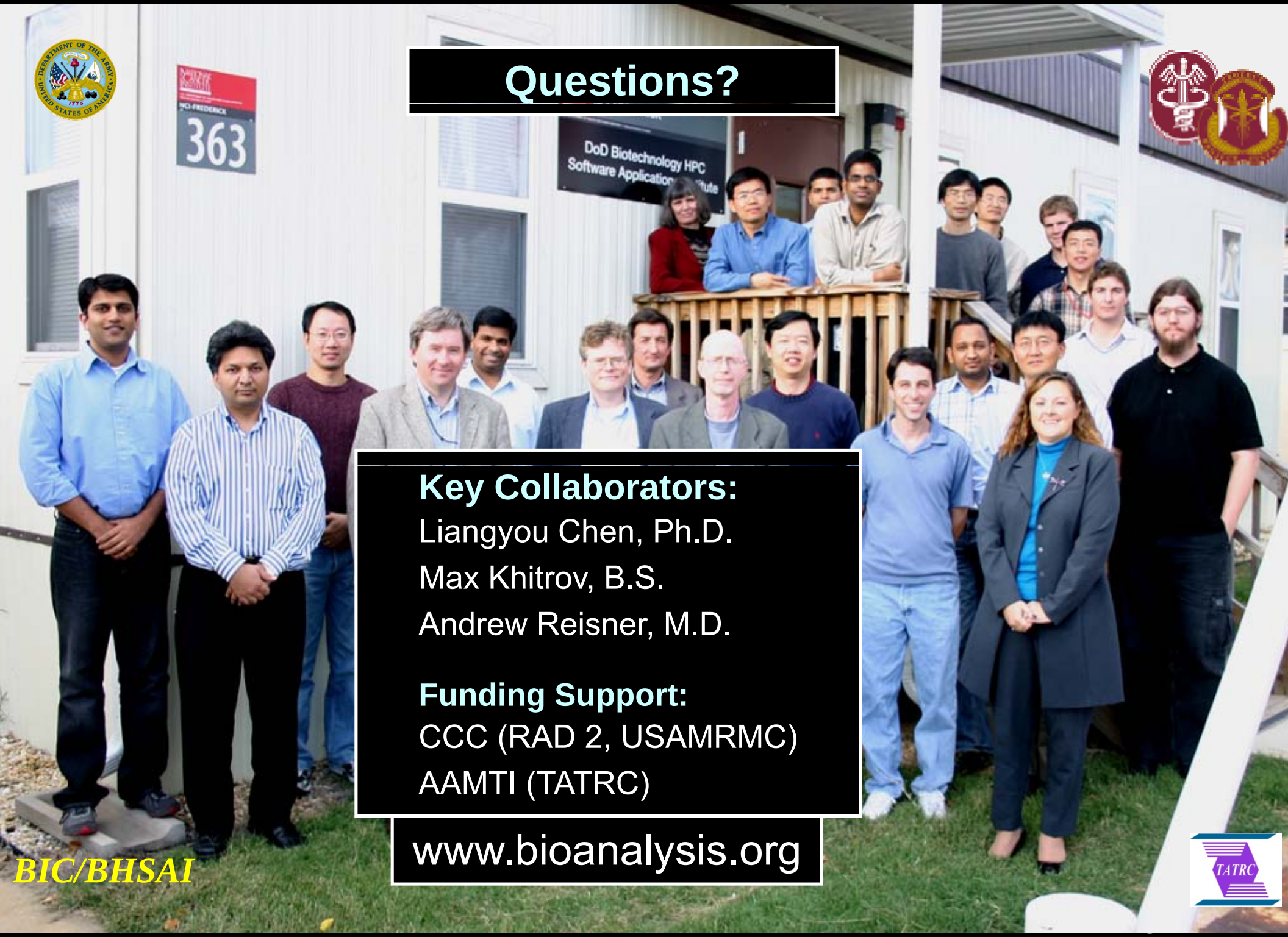
- **Evaluate other sensor modalities against benchmarked results**

Boston Medflight

APPRAISE

Propaq Monitor

Ruggedized PC & Algorithms



Questions?

Key Collaborators:

Liangyou Chen, Ph.D.

Max Khitrov, B.S.

Andrew Reisner, M.D.

Funding Support:

CCC (RAD 2, USAMRMC)

AAMTI (TATRC)

www.bioanalysis.org

BIC/BHSAI