What is nowcasting and forecasting?

What is it’s application to the care of acutely and critically ill physiologically monitored patients?

- Nowcasting—using current data to more sensitively and specifically detect the current state
- Forecasting—using current data to sensitively and specifically predict the future state

Example--Weather
Why is nowcasting and forecasting so important in clinical care for acutely and critically ill patients?

- To prevent failure to rescue, but to allow interventions to be applied as early in the FTR cascade as possible.

Benefits:
- To keep a bad situation from getting worse
- To keep a good situation from getting bad
- To improve the quality and safety of care
- To improve the economics of care delivery
Early Detection of Disease Model

Example—Using vital sign data to detect the unstable state more sensitively and specifically with EWS

- Early Warning Scores—assess patterns of changes in charted vital signs that represent departure for “normal”, and assign a single index score.
- Trigger index scores are associated with likelihood of mortality in ward patients
- Can be triggered if one VS deviates markedly, or multiple VS deviate in a lesser manner but cumulatively

Some soft spots in EWS

Based on intermittently obtained VS
Number of variables which contribute to development of the index score are limited
Are dependent upon the accuracy of the charted VS
If calculation is not automated, can be associated with users calculation errors
If alert is not automated, can be associated with users bypassing the call protocol
Example—Using continuously monitored vital sign data to detect the unstable state more sensitively and specifically

- In a static field of single point-in-time multisignal data, health and disease can be separated in stochastic fashion using **Artificial Neural Network** approach to create a probabilistic equation and develop a fused parameter of the *present* state.

**Integrated Monitoring Systems for Early Warning of Departure from Normality**

First utilized in the airline industry as a mechanism for early warning of signs and symptoms of impending engine failure (Boeing)

Also used in automotive industry (Rolls Royce)
**Explanation the Integrated Monitoring Index value**

- Continuous input from 4 vital sign signs undergoes data fusion algorithmic processes to develop a single Visensia Index Score (VSI) score.
- Uses neural networking to recognize changes from normality (defined by a training set).
- Alerts for a single parameter deviating by + 3 SD from “normal” value, or 2-3 parameters moving away from normality by a smaller amount.
- Filtered for noise; requirement for temporal persistence (i.e. 4 out of previous 5 minutes)
- May alarm when all individually measured parameters are still in their “normal” range if the indexed value discerns departure from normality.

**QI Project PURPOSE**

Evaluate the capability of an IMS to improve nurses’ ability to both detect cardiorespiratory instability according to Medical Emergency Team (MET) call criteria in patients on a SpO2 and ECG monitored SDU and shorten duration of instability.

---

**METHODS**

IMS (Visensia™, OBS Medical, Carmel IN, USA), receives continuous input from bedside monitors of HR, RR, BP, and SpO2. Visensia Index (VSI) neural networked value displayed at bedside and in central alarm monitors.

**Phase 1 (P1).** Visensia installed; monitor screen not visible to staff; VSI and VS trend data background recorded. Staff reeducated on UPMC MET triggers and call mechanism.

**Phase 2 (P2).** Monitored information visible, staff educated but no specific direction for use in clinical care. Sensitivity and specificity analysis; clinical algorithm development.

**Phase 3 (P3).** Monitored information visible, audible alerts at hall and central stations; clinical algorithm for response to VSI alerts used.

Detection of VS parameter changes meeting MET trigger values defined instability.

Data comparisons for P1 to P3 used descriptive, Chi-square and students t-test analyses.

**Algorithm--Clinical Response to IMS Alert in P3**

**UPMC Clinical Decision Rules for Visensia™ Red Alerts**

- **VSI Alarm Condition**
  - Go to next step: Is the patient visibly abnormal clinical deterioration?
  - **Yes**
  - **No**
  - Evaluate the 4 monitored VS parameters (HR, RR, BP, SpO2), are sensors attached and signal capture secure?
  - **Yes**
  - **No**
  - Continue nursing evaluation of pat relative to the abnormal vital sign parameter.
  - **Implement Nursing Intervention**
  - Call MD, CRNP if P3 is not received.

**Call MET**

**End**
Metric for Instability

Detection of VS parameter changes meeting UPMC MET trigger values defined instability

- **RR**: <8 or >36
- **SpO2**: < 85% for > 5 min
- **HR**: <40 or >140
- **BP**: <80 OR >200 systolic or 110 diastolic

VS trend plots achieving above criteria blindly judged by senior physician intensivist (Prof. M Pinsky) into instability levels

- CRI hit
- CRI min
- CRI full
- MET actual

Example of a chart judged as CRI hit but reason for VSI elevation is artifactual. Patient has precipitous \( \downarrow \) SpO2 at 0200 but no compensatory change in other vitals, rapid return of SpO2 to prior baseline; judged as transient loss SpO2 signal capture

Example a chart judged as CRI min. Patient has transient elevation of BP and drop in RR across MET call threshold at 04:00 with VSI alert but then reverts to baseline
RESULTS
Phase 1 and 3 patients were similar

<table>
<thead>
<tr>
<th></th>
<th>Phase 1</th>
<th>Phase 3</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td># Admissions</td>
<td>326</td>
<td>308</td>
<td></td>
</tr>
<tr>
<td># Weeks</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Monitoring Hours</td>
<td>18,258</td>
<td>18,314</td>
<td></td>
</tr>
<tr>
<td>Age (mean yrs ± SD)</td>
<td>57.7 ± 19.7</td>
<td>57.4 ± 20.2</td>
<td>ns</td>
</tr>
<tr>
<td>Male Gender (n, %)</td>
<td>191 (58.6%)</td>
<td>180 (58.4%)</td>
<td>ns</td>
</tr>
<tr>
<td>Caucasians (n, %)</td>
<td>240 (73.6%)</td>
<td>221 (71.8%)</td>
<td>ns</td>
</tr>
<tr>
<td>DRG Weights (mean ± SD)</td>
<td>2.75 ± 3.65</td>
<td>2.86 ± 3.60</td>
<td>ns</td>
</tr>
</tbody>
</table>

Incidence of Instability

-37.8%
-30.1%
-68.4%

**p<0.01
CONCLUSIONS

• Using an IMS with continuous derivation of a single instability index value within a diagnosis and management system improved the nurses ability to detect present clinical instability as compared to conventional four channel monitoring in the SDU environment.

• Improved detection of present instability resulted in:
  – a decrease in overall duration of patient instability
  – a decrease in the number of patients who progressed to serious instability
  – an increase in the number of patients with serious instability who had a MET called

But, we noted something curious......
Patterns of instability seem to emerge even far in advance of IMS CRI detection.

Are impending signs of future instability available far in advance of overt instability? Is this “nurse intuition” or is it possible to recognize and quantify these subtle changes far in advance to forecast future instability?

Next steps

- If using the IMS enables nurses to detect present instability more sensitively and specifically............................ wouldn’t it be even more awesome to be able to predict future instability in advance of its overt manifestation based on very subtle changes in VS patterns?
- Move from NOWCASTING to FORECASTING

Machine learning

- **Definition:** Machine learning is a subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence. Machine learning explores the construction and study of algorithms that can learn from and make predictions on data. Such algorithms operate by building a model from example inputs in order to make data-driven predictions or decisions, rather than following strictly static program instructions.
- In a dynamic field of continuously changing but inter-related variables, Machine Learning data-driven classification techniques include: Principal Component Analysis, Support Vector Machines, K Nearest Neighbors, Random Forests, Naïve Bayesian Classifier
Pattern Recognition—Faces
Recognition and interpretation improves as move from low frequency univariate data—to high frequency, multivariate, multisource

Examples of uses of machine learning for pattern detection
“signatures”
- Fingerprints
- DNA sequencing
- Written information
  - Zip Code

Machine Learning Applied to Forecasting Instability
- Supervised Learning Workflow
  - What is the minimal data set needed to forecast instability: Monitoring parsimony
    - Number of independent monitoring variables
    - Lead time
    - Sampling frequency
  - What additional information will improve specificity
    - Derived monitoring variables
    - Other data (age, gender, labs, meds)
NIH National Institute of Nursing Research

RO1NR013912

“Predicting Patient Instability Noninvasively for Nursing Care (PPINNC)”
10/13 to 10/15

This proposal seeks to further apply machine learning and complexity modeling-based algorithms to predict cardiorespiratory instability (CRI) prior to overt instability with sufficient lead-time and accuracy to support a nursing decision for preemptive therapy.

Example of a chart judged as CRI hit (artifactual deviation of VS)

But first….how to deal with the problem of artifact

Can you train the machine learning algorithm to react only to real VS deviations in forecasting?

Example of an SpO2 event epoch annotated by experts as artifact

List of numeric features informative to differentiate between real alerts and artifact

<table>
<thead>
<tr>
<th>Feature</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpO2_median</td>
<td></td>
</tr>
<tr>
<td>SpO2_sd</td>
<td></td>
</tr>
<tr>
<td>SpO2_cv</td>
<td></td>
</tr>
<tr>
<td>SpO2_range</td>
<td></td>
</tr>
<tr>
<td>SpO2_range_ratio</td>
<td></td>
</tr>
<tr>
<td>SpO2_max</td>
<td></td>
</tr>
<tr>
<td>SpO2_range</td>
<td></td>
</tr>
<tr>
<td>SpO2_quar_resvar</td>
<td></td>
</tr>
<tr>
<td>SpO2_min</td>
<td></td>
</tr>
<tr>
<td>SpO2_mean</td>
<td></td>
</tr>
<tr>
<td>SpO2_quad_rsq</td>
<td></td>
</tr>
</tbody>
</table>

Drop off Data Sparseness Oscillation
Example of an SpO2 event epoch annotated by experts as real
Signal integrity good; gradual down slope and upslope, accompanying
tachycardia and tachypnea

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpO2_median</td>
<td></td>
</tr>
<tr>
<td>SpO2_sd</td>
<td></td>
</tr>
<tr>
<td>SpO2_cv</td>
<td></td>
</tr>
<tr>
<td>SpO2_range_ratio</td>
<td></td>
</tr>
<tr>
<td>SpO2_max</td>
<td></td>
</tr>
<tr>
<td>SpO2_range</td>
<td></td>
</tr>
<tr>
<td>SpO2_quar_resvar</td>
<td></td>
</tr>
<tr>
<td>SpO2_min</td>
<td></td>
</tr>
<tr>
<td>SpO2_mean</td>
<td></td>
</tr>
<tr>
<td>SpO2_quad_rsq</td>
<td></td>
</tr>
</tbody>
</table>

**Event Annotation by Experts**

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Events</td>
<td>631</td>
</tr>
<tr>
<td>Real alerts</td>
<td>418</td>
</tr>
<tr>
<td>Artifact</td>
<td>158</td>
</tr>
<tr>
<td>Unable to Classify</td>
<td>55</td>
</tr>
<tr>
<td>Annotated as Real Events Total</td>
<td>418</td>
</tr>
<tr>
<td>Annotated as Real Events by subtype</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>60 (14%)</td>
</tr>
<tr>
<td>RR</td>
<td>138 (32%)</td>
</tr>
<tr>
<td>SpO2</td>
<td>181 (44%)</td>
</tr>
<tr>
<td>BP</td>
<td>45 (11%)</td>
</tr>
</tbody>
</table>

Mean Area Under the Receiver Operating Characteristic Curve (AUC) scores from 10-fold cross-validation (CV) experiments on the array of algorithms for three categories of VS events.

Training and validation (10-fold CV) on the Block 1 (dashed lines) data is displayed, along with the mean AUC scores when testing the algorithms on the Block 2 test set data (solid lines).
Using the 631 labeled events, RF was trained to tell apart real alerts from artifact, then cross-validated to mitigate overfitting.

Next the resulting model was applied to the 795 unseen events, which were then reviewed by the experts for external validation of RF predictions.

**Performance of models when applied to a real time stream of data starting from the time that the VS parameter crosses threshold**

- Also needed: differentiate real alerts vs artifact as the VS abnormality evolves in real time (i.e. not using the full epoch as in previous experiment).
- This approach has implications for providing directive clinical decision support advice to clinicians or differential alarming systems in real time and decreasing alarm fatigue.
- In this experiment the RF classifier is applied to the full set of features but as event evolves in real time.
- We constructed moving windows of 3 mins width (180 sec), for every 20 seconds from the start of the alert and up to 180 seconds.
Now that the problem of artifact is taken care of, can move to models trained to forecast real instability.
AUC Score by lead time modeled with multiple true instability events per patient during their SDU admission

- RF models with binary label (positive/negative);
- Take all CRI into account
- Plot shows the temporal change of model's discriminative power (AUC score) when approaching true CRI event
- This is leave one patient out cross validation result

AUC Score by lead time modeled with multiple true instability events per patient during their SDU admission

- AUC=90%

AUC Score by lead time modeled with only the first true instability event per patient during their SDU admission

- RF models with binary label (positive/negative);
- Only first CRI were modeled
- Plot shows the temporal change of model’s discriminative power (AUC score) when approaching true CRI event
- This is leave one patient out cross validation result

AUC Score by lead time modeled with only the first true instability event per patient during their SDU admission

- AUC=71%

AUC Score by lead time modeled with only the first true instability event per patient during their SDU admission

- AUC=78%
AUC Score by lead time modeled with only the first true instability event per patient during their SDU admission

AUC=82%

AUC Score by lead time modeled with first alert per patient during their SDU admission

AUC=87%

Performance from multi-label model
Prediction of the future instability type (driver)

Leave one patient out result

- RF models with multi-labels: BP, HR, RR, SPO2 and negative;
- Not only predict the probability of real event, also predict the type of CRI events
- Compute most probable class
- Plot shows the temporal change of recall when approaching true event

Performance from multi-label model
Prediction of the future instability type (driver)

Leave one patient out result

- RF models with multi-labels: BP, HR, RR, SPO2 and negative;
- Not only predict the probability of real event, also predict the type of CRI events
- Compute most probable class
- Plot shows the temporal change of recall when approaching true event

<table>
<thead>
<tr>
<th>Total Class</th>
<th>BP</th>
<th>HR</th>
<th>RR</th>
<th>SPO2</th>
<th>neg</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>0</td>
<td>39</td>
<td>62</td>
<td>72</td>
<td>46</td>
<td>189</td>
</tr>
<tr>
<td>Precision</td>
<td>0%</td>
<td>63%</td>
<td>56%</td>
<td>27%</td>
<td>94%</td>
<td>81%</td>
</tr>
<tr>
<td>Overall Acc</td>
<td>0%</td>
<td>63%</td>
<td>56%</td>
<td>27%</td>
<td>94%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Overall Accuracy: 81%
We built machine learning models to predict the onset of future CRI events at various lead time.

The closer to the event, the better performance.

However, the temporal trend varies by type of CRI alert, suggesting heterogeneous risk process toward CRI events.

ML has a potential to identify informative VS abnormalities in advance of overt CRI. These findings have promise for future utilization.

So how can we move from mathematical models which can forecast instability in real time—to clinical action at the beside?

We envision a graphical user interface (GUI) which gives the probability of an event occurring, the time to the event, and the type (physiologic driver) of the event.
What is the future for such forecasting modeling in invasively monitored patients in the ICU?

- Results of an animal (pig) model study
- Controlled bleeding
- HR, RR, invasive monitoring (arterial, CVP, PA)

Identify Onset of Bleeding Earlier

- Train a multivariate regressive model (Random Forest)
- Leave one out

Density of Data Affects Bleeding Detection
Multivariate models for various groups of measurements

Density of Data Affects Bleeding Detection
Multivariate models for various groups of measurements

Implications

- NOWCASTING can directly impact clinical practice by better detecting and recognizing present instability, thereby decreasing preventable mortality.
- FORECASTING can target patients who are most likely to become unstable in the future. Predicting instability at the earliest possible time to intervene carries with it the greatest potential to minimize morbidity. The technology and computational modeling methods proposed have the potential to drastically shift nursing surveillance paradigms and change instability care from reactive to preemptive.