Heart Failure Hospitalization Prediction in Remote Patient Management Systems

Pechenizkiy, M., Vasilyeva, E., Žliobaite, I., Tesanovic, A. & Manev, G.

Project page: http://www.win.tue.nl/~mpechen/projects/rpm/ Supported by KWR MIP project
Focus on Cardio:
• Treatment and management of chronic cardiovascular diseases
Outline

• Remote Patient Management (RPM)
• Patient modeling in RPM
• Heart Failure Hospitalization (HFH) Prediction
  • Problem formulation
  • Available data sources
• Our approach for HFH predicting modeling
  • Construction of positive and negative training instances
• Experimental results
• Conclusions and further work
Health and cost problem

- **Cardio-vascular diseases**
  - Major cause of death and health costs in EU and US
  - Affect 40% of overall population
  - After first event is chronic: heart failure, coronary artery disease,…
  - Shift from patient care in the hospital to care at home
Chronic Heart Failure

- Inability of the heart to pump blood adequately around the body
- Prevalence:
  - 5.7 million people in the US \(^1\)
  - 14 million people in Europe \(^2\)
  - will be doubled by 2020 \(^2\).
- Patients are around 75 year \(^3\) having several co-morbidities
- Prognostics:
  - 40 percent will die within one year \(^2\)
  - 25 ~ 38 percent will survive more than five years \(^2\)

\(^1\) AHA Heart Disease and Stroke Statistics 2009
\(^2\) Shape, study group on heart failure awareness and perception in Europe
\(^3\) ESC Heart Failure Guidelines  2008
Chronic Heart Failure – Management

• **Costs:**
  - $38.4B annually in the US
  - Half of the cost is due to re-admissions
    - 20%-30% patients readmitted within 30 days
    - 56% patients readmitted within a year

• **Guidelines for Management**
  - **Invasive:** devices (implants) and surgery
  - **Non-Invasive:** pharmacological therapy
  - **Non-pharmacological management**
    - Nurse led Self Management programs
    - Telehealth and guidance

ESC Heart Failure Guidelines 2008
Current limitations

- Telehealth programs are not flexible
- Educational material is too generic
- Alerting power is limited
- Generally, it is not clear what the model of RPM technology acceptance is
- Follow presentation by Seppo Puuronen tomorrow
Telehealth and guidance

Patient at home

Measuring Devices

Application Hosting Device

Monitoring & Management Server

Feedback Device

Feedback

Patient RPM data and alerts

Professionals (Cardiologist, HF nurse, GP)

Professional’s Site

Technische Universiteit Eindhoven
University of Technology
Example: Teleheath data

- Education
- Surveys
- Vitals
- RPM data
- Monitoring & Management Server
- Application Hosting Device
- Measuring Devices
- Feedback Device
- Patient at home
Data sources

• Cardiology information systems (IS)
  • Imaging, labs, medical history, medications, some vitals
  • Longitudinal data (multiple-years)
• Hospital IS and electronic patient records
  • Medical history, labs/biomarkers, other diseases, causes, vitals
  • Longitudinal data (multiple-years)
• Telehealth IS
  • Vitals, symptoms, education, medical history, medications
  • Logs of system usage
  • Operational data, one-two years per patient
• Clinical studies
  • Vitals, symptoms, medications, medical history
  • Average one and a half years per patient
## Clinical data – availability and reliability

<table>
<thead>
<tr>
<th>Data classes</th>
<th>Collected via</th>
<th>(Typical) Frequency</th>
</tr>
</thead>
</table>
| Medical history         | **Causes**  
Face to face meeting at a medical professional’s institution  
**Co-mobidities**  
**Prior hospitalizations**  
**Implanted devices**  
  
  
  
  | Once, when diagnosis for chronic condition is made |                                                                                                                                                                           |
| Vital signs             | **Weight**  
-Few months at office  
**Blood pressure**  
-Daily using the system at a patient’s home  
**Pulse**  
  
  
  
  | Daily |                                                                                                                                                                           |
| Bio-markers             | Face to face meeting at a medical professional’s institution  
  
  
  
  | (Few) months |                                                                                                                                                                           |
| Medications             | **Disease related drugs**  
  
  
  
  | Few weeks to few months  
**Non-disease related drugs**  
  
  
  
  |                                                                                                                                                                           |

11
## Clinical data – availability and reliability

<table>
<thead>
<tr>
<th>Data classes</th>
<th>Collected via</th>
<th>(Typical) Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline data</strong></td>
<td></td>
<td><strong>Vitals</strong></td>
</tr>
<tr>
<td></td>
<td>Face to face meeting at a medical professional's institution; before discharge or use of telehealth</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Height</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Other diagnosis</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Lab results</strong></td>
</tr>
<tr>
<td><strong>Questionnaires</strong></td>
<td>Several alternatives:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- A system at a patient’s home, but also can be collected:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Via a telephone contact by a medical professional</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Via face to face meeting during regular checkups at medical professional institution</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Symptoms</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Depression</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Anxiety</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Overall health</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Overall QoL</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Stress</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Sleep patterns</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Fatigue</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Loneliness</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Data collection methods and frequencies can vary depending on the protocol of care and individual patient needs.*
Next generation architecture

Tesanovic et al. CBMS 2009
Next generation architecture
Knowledge discovery process

- Feature extraction
  - Relevant data selection
  - Event-pattern analysis
  - Time-series analysis

- Data to be mined
  - Data Mining techniques
  - Actionable patterns

Input: RPM data
Event analysis

<table>
<thead>
<tr>
<th>Start of a week</th>
<th>Monthly contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hosp. admission</td>
<td>Hosp. discharge</td>
</tr>
<tr>
<td>Weight meas.</td>
<td>BP and HR meas.</td>
</tr>
</tbody>
</table>
Facilitating feature extraction

The graphs illustrate the changes in weight over time, with specific annotations for hospitalization admission and discharge, as well as symptoms worsening, improving, and no change in symptom quality.
HFH Prediction Problem Formulation

- Based on the available data about a patient at moment $t_i$ cast a prediction (alert /no alert) whether the hospitalization for this patient is likely within next 14 days period, $(t_i ; t_{i+14}]$
- Predictions are casted on the daily basis
How do we form positive training instances?

- For negative examples we choose two MC between which no hospitalization happened and select an arbitrary reference point “no hospitalization”

- Some tricks for dealing with missing data, noise, outliers, etc
Experiment design: TEN-HMS database

- 426 patients with cardiovascular diseases
- 143 patients of which were HF patients home tele-monitored during
- 2 years, 92 hospitalizations, 43 patients hospitalized at least once due to HF
- SVN, J48, JRip and other popular approaches (WEKA)
- Train/test random split with classifier/parameter/feature subset selection by CV on the training data
  - Test: 29 patients (9/20)
  - Training: 114 patients (34/80)
- We do not relearn the models every day to keep the setup simple but prequential testing could be used as well
TPR-FPR trade-off

Parameter and algorithm selection on training data

Final estimates with the test set
Our approach performs consistently better

<table>
<thead>
<tr>
<th>Classification model</th>
<th>TPR</th>
<th>FPR</th>
<th>YIndex</th>
<th>HRate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule 1</td>
<td>0.038</td>
<td>0.019</td>
<td>0.019</td>
<td>0.316</td>
</tr>
<tr>
<td>Rule 2</td>
<td>0.123</td>
<td>0.028</td>
<td>0.095</td>
<td>0.211</td>
</tr>
<tr>
<td>Rule 3</td>
<td>0.128</td>
<td>0.012</td>
<td>0.115</td>
<td>0.211</td>
</tr>
<tr>
<td>Rule 4</td>
<td>0.010</td>
<td>0.022</td>
<td>-0.012</td>
<td>0.053</td>
</tr>
<tr>
<td>Rule 5</td>
<td>0.180</td>
<td>0.038</td>
<td>0.143</td>
<td>0.263</td>
</tr>
<tr>
<td>Rule 6</td>
<td>0.189</td>
<td>0.058</td>
<td>0.131</td>
<td>0.316</td>
</tr>
<tr>
<td>Rule 7</td>
<td>0.209</td>
<td>0.055</td>
<td>0.153</td>
<td>0.421</td>
</tr>
<tr>
<td>Rule 8</td>
<td>0.217</td>
<td>0.075</td>
<td>0.142</td>
<td>0.474</td>
</tr>
<tr>
<td>S+D (SVM)</td>
<td>max TPR</td>
<td>0.582</td>
<td>0.215</td>
<td>0.367</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.555</td>
<td>0.124</td>
<td>0.371</td>
</tr>
<tr>
<td>S+D (JRF)</td>
<td>max TPR</td>
<td>0.527</td>
<td>0.162</td>
<td>0.365</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.495</td>
<td>0.124</td>
<td>0.371</td>
</tr>
<tr>
<td>S+D (J48)</td>
<td>max TPR</td>
<td>0.400</td>
<td>0.112</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.345</td>
<td>0.090</td>
<td>0.256</td>
</tr>
<tr>
<td>S+D+H (SVM)</td>
<td>max TPR</td>
<td>0.427</td>
<td>0.196</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.418</td>
<td>0.126</td>
<td>0.292</td>
</tr>
<tr>
<td>S+D+H (JRF)</td>
<td>max TPR</td>
<td>0.627</td>
<td>0.293</td>
<td>0.334</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.464</td>
<td>0.188</td>
<td>0.276</td>
</tr>
<tr>
<td>S+D+H (J48)</td>
<td>max TPR</td>
<td>0.573</td>
<td>0.239</td>
<td>0.334</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.564</td>
<td>0.231</td>
<td>0.332</td>
</tr>
<tr>
<td>S+D+H+FS (SVM)</td>
<td>max TPR</td>
<td>0.432</td>
<td>0.172</td>
<td>0.259</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.409</td>
<td>0.062</td>
<td>0.348</td>
</tr>
<tr>
<td>S+D+H+FS (JRF)</td>
<td>max TPR</td>
<td>0.441</td>
<td>0.173</td>
<td>0.268</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.318</td>
<td>0.082</td>
<td>0.237</td>
</tr>
<tr>
<td>S+D+H+FS (J48)</td>
<td>max TPR</td>
<td>0.432</td>
<td>0.172</td>
<td>0.259</td>
</tr>
<tr>
<td></td>
<td>min FPR</td>
<td>0.341</td>
<td>0.078</td>
<td>0.263</td>
</tr>
</tbody>
</table>
Conclusions and Further work

• HFH prediction is difficult
• Automated classification works better than any individual trigger suggested by experts
• The challenge is in combing data of different types (with different availability and reliability)
• It is important to model availability and reliability explicitly
• Context-awareness in time and space
• Can we improve the accuracy of models if we take educational (usage) data, and feedback patients receive into account
What if…

• … the patient’s condition deteriorates
• … the patient’s habits change
• … the seasons changes
• …

Challenges
• discover the hidden context
• adapt rules (second-order adaptation)
• incorporate new knowledge as it appears
Handling changes in the data

- Expected yet not directly observable
  - Number of co-morbidities
- Unforeseen
  - Changes in hidden contexts

- Handling Concept Drift in Medical Applications
  - 3rd Tutorial on October 14

- Welcome!
Questions?

• Thank you!
References

• Puuronen et al., CBMS 2010

• Pechenizkiy et al., CBMS 2010

• Tesanovic et al., CBMS 2009
  • Aleksandra Tesanovic, Goran Manev, Mykola Pechenizkiy, Ekaterina Vasilyeva: eHealth personalization in the next generation RPM systems. CBMS 2009: 1-8

• preprints available at www.win.tue.nl/~mpechen/publications/