Can Big Data Analytics Support Personalized Sleep Medicine?

Allan I. Pack, M.B.Ch.B., Ph.D.
John Miclot Professor of Medicine
Division of Sleep Medicine/Department of Medicine
University of Pennsylvania Perelman School of Medicine
Philadelphia, Pennsylvania
USA

Outline
Will Use OSA As an Example

• What is personalized medicine?
• Personalizing approaches to obstructive sleep apnea
• What are big data analytics?
• How can big data advance personalized medicine?
Personalized Medicine

“Personalized medicine is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it”

Paraphrased from Dan Ariely, Duke University (via Dan Rader)

Basis of Uniqueness

- Genetics
- Microbiome
- Epigenetics
- Environment

Unique patient

Personalized Phenotypes

CLINICAL CHARACTERISTICS
- ± Sleepiness
- ± Co-morbidities
- Other symptoms

MOLECULAR PROFILES
- Proteomics
- Metabolomics
- Transcriptome

IN-DEPTH PHENOTYPING MEASURES
- Physiologic measures
- Imaging
Fundamental Concept

• ALL PATIENTS WITH APPARENTLY SAME DISORDER ARE NOT IDENTICAL

• Use multiple approaches to evaluate differences
  – Physiological differences
  – Clinical differences
  – OMIC differences (all OMICS)
  – Genetic/epigenetic differences

• Use unbiased, discovery approaches

A Variant of Precision Medicine

The Four Ps
  • Predictive
  • Preventive
  • Personalize
  • Participatory – Great technology for OSA

Concept proposed by Leroy Hood

Personalized Sleep Apnea

• Three main concepts
  – Different physiological risk factors
  – Different physiology in PSG
  – Different symptomatic subtypes

Using PSG to Identify Physiological Risk Factors for OSA

• Pharyngeal collapsibility
• Overall loop gain
• Arousal threshold
• Upper airway muscle responsiveness

?CAN INFORMATION BE OBTAINED FROM ROUTINE PSG
Is This Physiological Phenotyping Ready for Primetime?

• It is based on a simple model of ventilatory control
• Studies to date in small samples
• No reliability assessments, e.g.,
  – Does loop gain change within a night (sleep state)?
  – Does assessment of loop gain change from night to night?

MUCH WORK TO BE DONE

What is Clustering Analysis?

• Classify individuals using input variables such that:
  • Within a cluster, individuals are as similar as possible
  • Between clusters, individuals are as dissimilar as possible

CLUSTERING IS FREQUENTLY USED TO IDENTIFY SUBGROUPS
What About Different Physiological Subtypes Based on Standard Variables Obtained by Scoring PSG

<table>
<thead>
<tr>
<th>Cluster (n)</th>
<th>Cluster label</th>
<th>Median AHI* (events/hour)</th>
<th>Conventional OSA severity*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (533)</td>
<td>Mild</td>
<td>4</td>
<td>None/mild</td>
</tr>
<tr>
<td>B (119)</td>
<td>PLMS</td>
<td>10</td>
<td>Moderate</td>
</tr>
<tr>
<td>C (186)</td>
<td>NREM and poor sleep</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>D (168)</td>
<td>REM and hypopnea</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>E (75)</td>
<td>Hypopnea and hypoxia</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>F (42)</td>
<td>Arousal and poor sleep</td>
<td>68</td>
<td>Severe</td>
</tr>
<tr>
<td>G (124)</td>
<td>Combined severe</td>
<td>84</td>
<td></td>
</tr>
</tbody>
</table>

CPAP reduces CV events

7 groups/only 2 where CPAP ↓ events

What About Symptomatic Subtypes?

Lichuan Ye
Grace Pien
Brendan Keenan
Thorarinn Gislason
Clinical Subtypes of Obstructive Sleep Apnea

OSA = $\sum OSA_1 + \sum OSA_2 + \sum OSA_3$

Cluster 1: OSA + Insomnia
Cluster 2: Asymptomatic
Cluster 3: OSA + Excessive Sleepiness

Epworth Sleepiness Score
- Cluster 1: 9.5 ± 0.7
- Cluster 2: 7.9 ± 0.6
- Cluster 3: 15.7 ± 0.6

Predict Using Clinical Features: Demographics of Clusters

Cluster differences not driven by BMI or AHI

<table>
<thead>
<tr>
<th></th>
<th>Total Cohort</th>
<th>Cluster 1: Disturbed Sleep</th>
<th>Cluster 2: Minimally Symptomatic</th>
<th>Cluster 3: Excessively Sleepy</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects, n (%)</td>
<td>822 (100)</td>
<td>269 (32.7)</td>
<td>203 (24.7)</td>
<td>350 (42.6)</td>
<td>–</td>
</tr>
<tr>
<td>Male, %</td>
<td>81.0%</td>
<td>78.4%</td>
<td>83.7%</td>
<td>81.4%</td>
<td>0.336</td>
</tr>
<tr>
<td>Age, years</td>
<td>54.5 ± 10.6</td>
<td>54.1 ± 11.0</td>
<td>56.6 ± 10.3</td>
<td>53.6 ± 10.3</td>
<td>0.005</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>33.5 ± 5.7</td>
<td>33.3 ± 5.6</td>
<td>33.0 ± 5.6</td>
<td>34.0 ± 5.8</td>
<td>0.120</td>
</tr>
<tr>
<td>AHI, events/hr</td>
<td>44.9 ± 20.7</td>
<td>43.8 ± 20.4</td>
<td>43.1 ± 18.9</td>
<td>46.7 ± 21.7</td>
<td>0.181</td>
</tr>
</tbody>
</table>
Do Different Subtypes have Different Consequences?

**FOCUS ON CARDIOVASCULAR EVENTS**

Diego Mazzotti

---

**Sleep Heart Health Study**

- **Multi-center prospective community-based cohort** study of participants >40 years, designed to assess CV consequences of OSA
- **Baseline** (1995-1998) and **follow-up** (2001-2003) assessments, including **questionnaires** and in-home PSG
- **Standardized surveillance of CV outcomes** until end of follow-up (2008-2011)
- **Data available for 5,804 participants** through the NSRR

Quan et al., 1997 Sleep; Redline et al., 1998 Sleep; Dean et al., 2016 Sleep; Zhang et al., 2018 JAMA
OSA Symptom Subtypes in the Sleep Heart Health Study

K=4

OSA-Related Cardiovascular Risk Comes from Only the Excessively Sleepy Group.
No Increased Risk in Other Subgroups

Adjusted Cox Proportional Hazards Model confirms Excessively Sleepy have significantly worse outcomes vs. controls

Covariates in adjusted analyses: age, sex, body mass index, type 2 diabetes, hypertension, HDL, total cholesterol, triglycerides, alcohol use and smoking
Does This Result in Part Explain Negative Results in SAVE Study?

SAVE excluded very sleepy patients (Epworth Sleepiness Score >15)


• Very helpful review of potential big data approaches
• Applications
  – Identifying patients with likely undiagnosed OSA in EHR (age, gender, BMI, comorbidities)
  – Tracking outcomes – MI, stroke, diabetes, etc.
  – Coupling outcomes with CPAP compliance data - pragmatic trials
  – Obtaining financial information – use to develop bundled payment model
Utilizing Big Data Approaches in OSA

What do we need to do:
• Develop and validate an algorithm to identify OSA in EHR

Brendan Keenan

Algorithm Performance
OSA>Diagnosis Codes on 2 or More Occasions
(Keenan BT, et al, JCSM, 2019, in press)

ALL SITES EXCEPT ONE HAD PPV AND NPV >90%
Utilizing Big Data Approaches in OSA

What do we need to do:
• Develop and validate an algorithm to identify OSA in EHR
• Develop algorithms to identify unrecognized OSA in EHR (age, gender, BMI, associated comorbidities)
• Get CPAP data remotely (Bluetooth) and into EHR
• Get questionnaire data collected routinely
  – What subtype (minimize questions)
• Compare outcomes of CPAP users and non-users (control for covariates; propensity score matching)

Conclusions
• Studies of OSA are identifying subtypes – have implications for treatment and outcomes
• Have validated algorithm to identify OSA cases in EHR
• Big data approaches have great potential
• We are in strong position with ability to obtain remotely data from CPAP machines