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

P4 Medicine
PREDICT PREVENT PERSONALIZE PARTICIPATE

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

Concept proposed by Leroy Hood

See: Pack AI, Ann Am Thorac Soc 13:1456, 2016
Lim DC, et al, Respirology 22:849, 2017

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

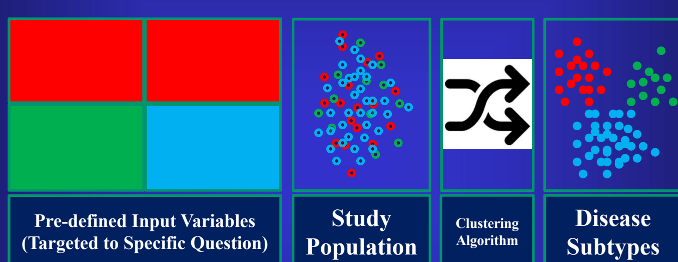
?CAN INFORMATION BE OBTAINED FROM ROUTINE PSG

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
(Zinchuk AV, et al, Thorax Sept 21, 2017)

Description of and labels for the polysomnographic clusters based on distinguishing features

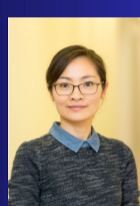
Cluster (n)	Cluster label	Median AHI* (events/hour)	Conventional OSA severity*
A (533)	Mild	4	None/mild
B (119)	PLMS	10	
C (186)	NREM and poor sleep	19	Moderate
D (168)	REM and hypoxia	19	
E (75)	Hypopnoea and hypoxia	44	Severe
F (42)	Arousal and poor sleep	68	
G (124)	Combined severe	84	

→ CPAP reduces CV events

→ CPAP reduces CV events

7 groups/only 2 where CPAP ↓ events

What About Symptomatic Subtypes?


Lichuan Ye


Grace Pien


Brendan Keenan

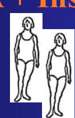

Thorarinn Gislason

Clinical Subtypes of Obstructive Sleep Apnea

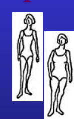
(Ye L, et al, Eur Respir J 44:1600, 2014)

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

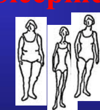
Cluster 1: OSA + Insomnia



Cluster 2: Asymptomatic



Cluster 3: OSA + Excessive Sleepiness



Epworth Sleepiness Score

9.5±0.7 7.9±0.6 15.7±0.6

Predict Using Clinical Features: Demographics of Clusters

(Ye L, et al, Eur Respir J 44:1600, 2014)

Cluster differences not driven by BMI or AHI

	Total Cohort	**Cluster 1:** Disturbed Sleep	**Cluster 2:** Minimally Symptomatic	**Cluster 3:** Excessively Sleepy	**p**
Subjects, n (%)	822 (100)	269 (32.7)	203 (24.7)	350 (42.6)	–
Male, %	81.0%	78.4%	83.7%	81.4%	0.336
Age, years	54.5 ± 10.6	54.1 ± 11.0	56.6 ± 10.3	53.6 ± 10.3	0.005
BMI, kg/m^2	33.5 ± 5.7	33.3 ± 5.6	33.0 ± 5.6	34.0 ± 5.8	0.120
AHI, events/hr	44.9 ± 20.7	43.8 ± 20.4	43.1 ± 18.9	46.7 ± 21.7	0.181

Do Different Subtypes have Different Consequences?

FOCUS ON CARDIOVASCULAR EVENTS

Diego Mazzotti

Sleep Heart Health Study

SLEEP HEART HEALTH

- **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

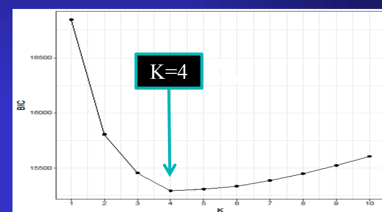
https://sleepdata.org

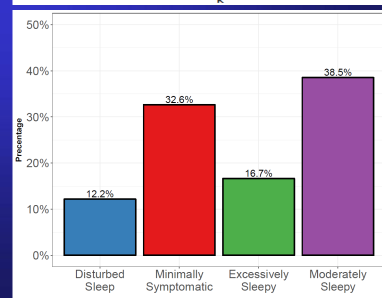
Quan et al., 1997 *Sleep*; Redline et al., 1998 *Sleep*; Dean et al., 2016 *Sleep*; Zhang et al., 2018 *JAMIA*

National Sleep Research Resource

OSA Symptom Subtypes in the Sleep Heart Health Study

(Mazzotti D, et al, AJRCCM 200:493-506, 2019)

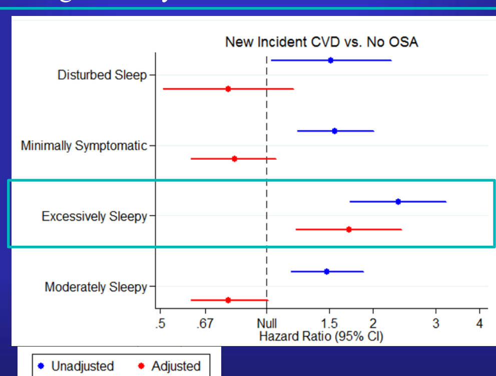




OSA-Related Cardiovascular Risk Comes from Only the Excessively Sleepy Group. No Increased Risk in Other Subgroups

(Mazzotti D, et al, AJRCCM 200:493-506, 2019)

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)

(see Javaheri S, et al, Chest 156:431, 2019)

Big Data in Sleep Apnea: Opportunities and Challenges

(Pepin JL, et al, Respirology, 2019)

- 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

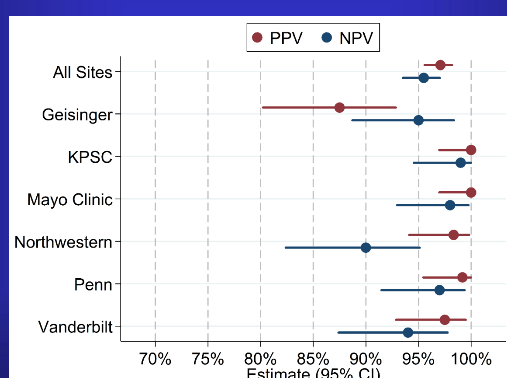
Utilizing Big Data Approaches in OSA

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