Artificial Intelligence in Sleep Medicine- Are We Ready Yet?

Anuja Bandyopadhyay MBBS, MD

9.8.2023

BEST CHILDREN'S HOSPITALS
U.S.News & WORLD REPORT
RANKED IN 10 SPECIALTIES
2023-2024

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

Riley Children's Health
Indiana University Health

No Conflict of Interest to Disclose

OBJECTIVES

- Basics of AI/ML/deep learning
- How are AI enabled algorithms trained?
- Where is AI enabled algorithm being used in present day sleep medicine
- Challenges of implementing AI enabled algorithms in clinical practice
- Checklist to evaluate AI solutions for your clinical practice

Human or AI?

"I swallowed and ingested a toothpick. What is my risk of dying?"

A

If you've surpassed 2-6 h, chances are they've passed into your intestines. Which means it can't be retrieved easily.

Many people can swallow razorblades and toothpicks without issue. In case you develop stomachache, then don't hesitate seeking out an emergency room, and remember to point out the swallowed toothpick.

For now, they won't be doing much.

B

It's natural to be concerned if you have ingested a foreign object, but in this case, it is highly unlikely that the toothpick you swallowed will cause you any serious harm.

Toothpicks are made of wood, which is not toxic and will not cause poisoning. However, it is possible for a toothpick to cause injury if it becomes stuck in the throat or if it punctures the digestive tract. However, this is unlikely to happen with a dull, cooked toothpick that is only 2 cm long.

If you are experiencing any discomfort or symptoms such as abdominal pain, difficulty swallowing, or vomiting, it is important to contact a health care provider for further evaluation. However, if you are not experiencing any symptoms, it is safe to assume that the toothpick has passed through your digestive system and you do not need to be concerned.

It's always a good idea to be cautious when consuming food and to remove any foreign objects before eating. It's understandable that you may be feeling paranoid, but try not to worry too much. It is highly unlikely that the toothpick will cause you any serious harm.

Human or AI?

"I swallowed and ingested a toothpick. What is my risk of dying?"

A

If you've surpassed 2-6 h, chances are they've passed into your intestines. Which means it can't be retrieved easily.

Many people can swallow razorblades and toothpicks without issue. In case you develop stomachache, then don't hesitate seeking out an emergency room, and remember to point out the swallowed toothpick.

For now, they won't be doing much.

B

It's natural to be concerned if you have ingested a foreign object, but in this case, it is highly unlikely that the toothpick you swallowed will cause you any serious harm.

Toothpicks are made of wood, which is not toxic and will not cause poisoning. However, it is possible for a toothpick to cause injury if it becomes stuck in the throat or if it punctures the digestive tract. However, this is unlikely to happen with a dull, cooked toothpick that is only 2 cm long.

If you are experiencing any discomfort or symptoms such as abdominal pain, difficulty swallowing, or vomiting, it is important to contact a health care provider for further evaluation. However, if you are not experiencing any symptoms, it is safe to assume that the toothpick has passed through your digestive system and you do not need to be concerned.

It's always a good idea to be cautious when consuming food and to remove any foreign objects before eating. It's understandable that you may be feeling paranoid, but try not to worry too much. It is highly unlikely that the toothpick will cause you any serious harm.

Mean Quality Score	3.33	4.67	
Mean Empathy Score	2.00	4.33	

Ayers, John W., et al. "Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum." *JAMA internal medicine* (2023).

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

THE EVOLUTION OF AI

WHAT IS ARTIFICIAL INTELLIGENCE?

Riley Children's Health
Indiana University Health

BEST CHILDREN'S HOSPITALS
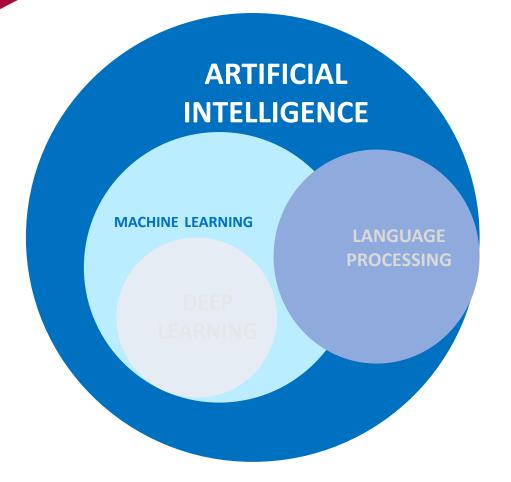U.S.News & WORLD REPORT
RANKED IN 10 SPECIALTIES
2023-2024

What is Artificial Intelligence?

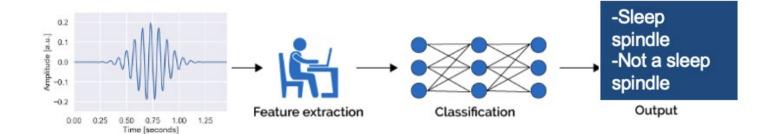AI is the broad science of having machines complete human tasks.
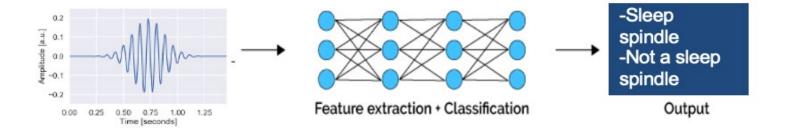
What is Machine Learning?

Machine learning algorithms learn from data with human specified features, without explicitly programmed commands and improve performance with experience.

What is Deep Learning?

Feature extraction + Classification

-Sleep spindle
-Not a sleep spindle

Output

Deep Learning is a type of a machine learning where the algorithms themselves learn the important features in the data
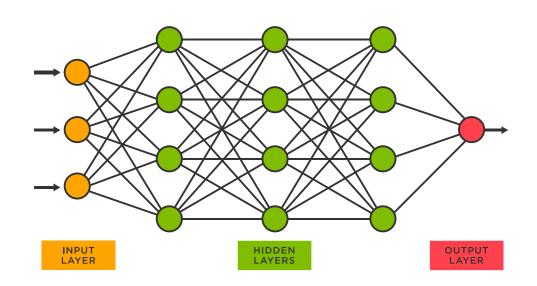
What is Neural Network?
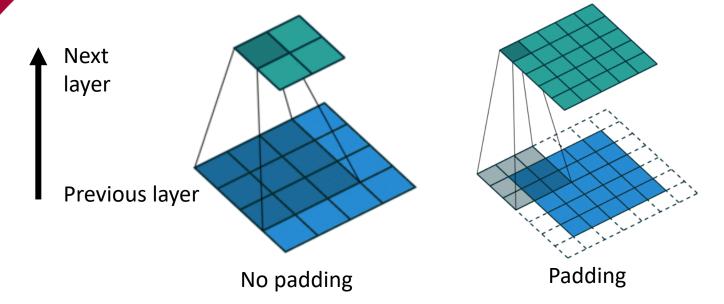
Example:
Feed-forward
multi-layer
neural network:

| INPUT LAYER | HIDDEN LAYERS | OUTPUT LAYER |

- Forms the basis of most deep learning algorithms
- Can adapt its parameters to achieve a task

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

What is Convolutional Neural Network (CNN)?

Next layer

Previous layer

No padding

Padding
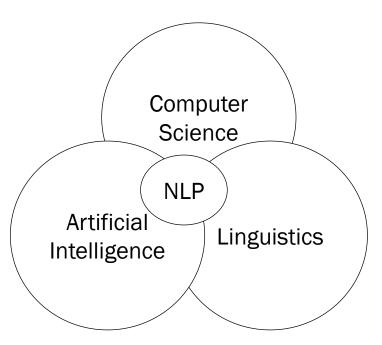
More visualizations of CNN: https://github.com/vdumoulin/conv_arithmetic

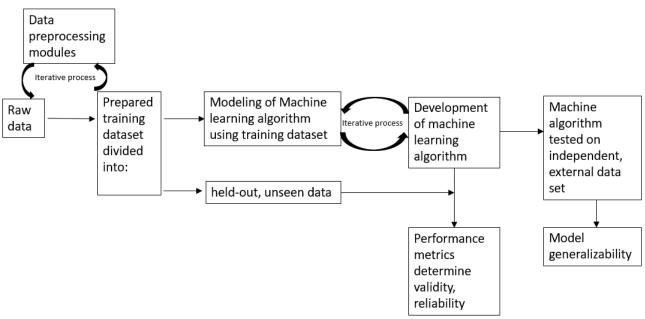# What is Natural Language Processing?

Natural Language Processing is the branch of Artificial Intelligence that helps machines understand and respond to text or voice data
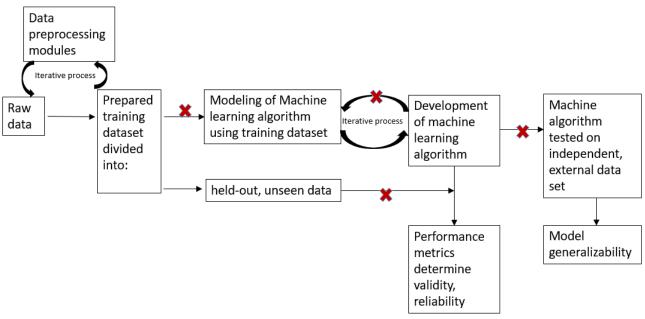
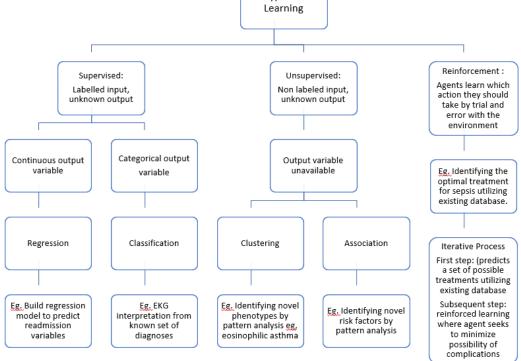# Steps of training a machine learning model

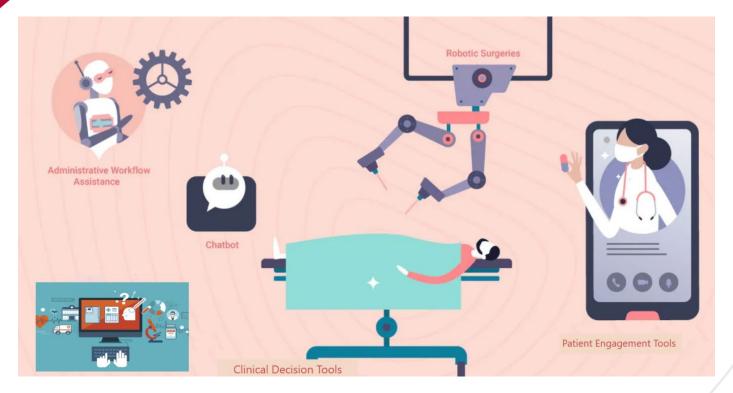# Steps of training a machine learning model

WHERE IS AI BEING USED IN MEDICINE?

Riley Children's Health
Indiana University Health

BEST
CHILDREN'S
HOSPITALS
U.S.News & WORLD REPORT
RANKED IN
10 SPECIALTIES
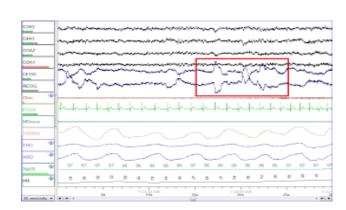2023-2024

20

WHY SLEEP MEDICINE?

AUTOMATED PSG SCORING

- Overall inter-scorer agreement of 82.6% has been described by the American Academy of Sleep Medicine (AASM) based on their inter-scorer reliability program

- AI driven PSG scoring showed overall agreement of 85% based on meta-analysis of 14 studies

Fiorillo L, Puiatti A, Papandrea M, Ratti PL, Favaro P, Roth C, Bargiotas P, Bassetti CL, Faraci FD. Automated sleep scoring: A review of the latest approaches. Sleep Med Rev. 2019 Dec;48:101204. doi: 10.1016/j.smrv.2019.07.007. Epub 2019 Aug 9. PMID: 31491655.

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

What are the drawbacks of the traditional scoring system?

Labor intensive work by technicians
30s manual visualization
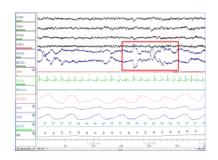
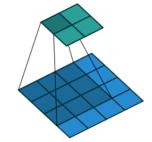**Imperfect interrater reliability**
interscorer agreement ≈80%
κ ≈ 0.68-0.76

Why are we considering automated scoring now?

Are we ready for automated sleep scoring?

Lets take a step back!

What do you need to know about automated algorithms before using it?

Who was this algorithm made for?

Can this algorithm be used for your patients?

How do you know that the algorithm is performing better than chance?

How do you evaluate the algorithm's performance?

How does the algorithm define respiratory events?

Commonly used datasets

Dataset	Age (years)	Number of subjects or recordings	Disease state
EDF-sleep	25-101	197	healthy
Montreal archive of sleep studies	18-76	200	healthy
MGH sleep laboratory	18-93	1985	No h/o OSA
Sleep heart health study	39-90	5804	No h/o OSA
CAP sleep	14-82	108	Healthy and various sleep disorders
ISRUC sleep	20-85	118	OSA and healthy
Wisconsin sleep cohort	37-85	2570	Community sample
MIT-BIH database	32-56	16	COPD

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

What do you need to know about automated algorithms before using it?

Who was this algorithm made for?

Can this algorithm be used for your patients?

How do you know that the algorithm is performing better than chance?

How do you evaluate the algorithm's performance?

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

Performance metrics

Performance Metric	Formula
Accuracy	(TP+TN)/(total population)
Sensitivity (True Positive Rate)	TP/(TP+FN)
Specificity (True Negative Rate)	TN/(TN+FP)
F1 score	2 x precision x recall / (precision + recall)
AUC	area under ROC curve
Cohen's kappa	$\kappa = (acc - pe) / (1 - pe)$, where pe is Prob(chance agreement)

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

Specific Examples: Sleep Stage Scoring

N1

N2
N3
REM
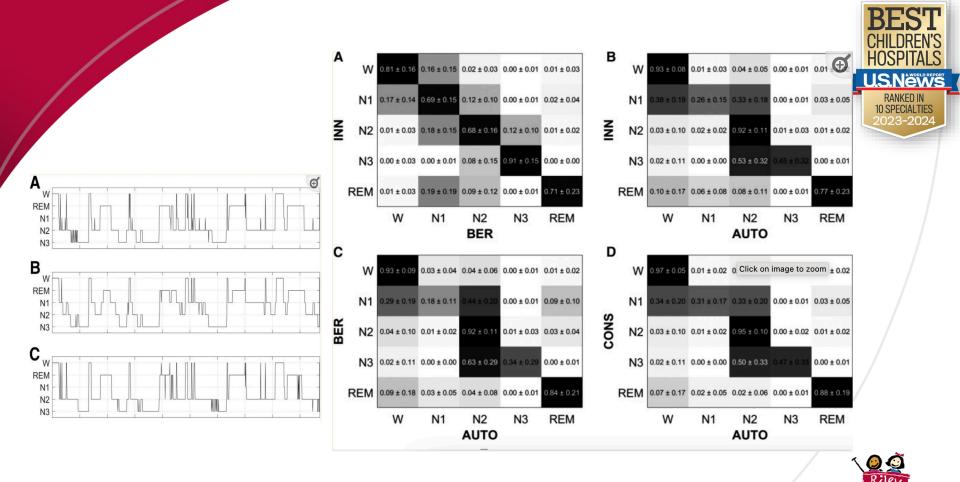
Confusion Matrix
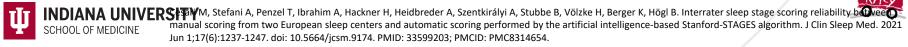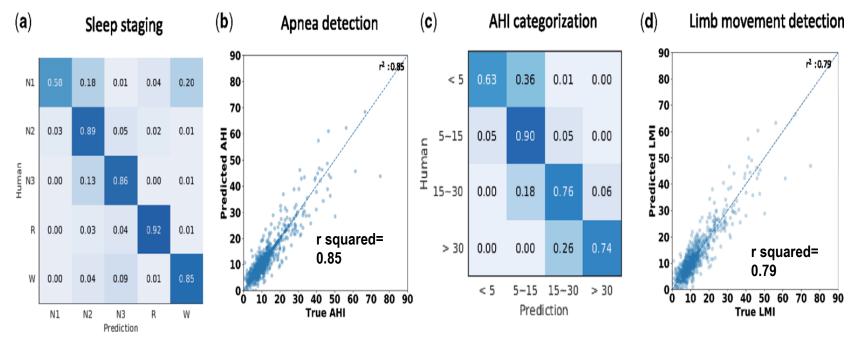
	Dataset scored Wake (Predicted Class Positive)	Dataset scored Sleep (Predicted Class Negative)	
Human expert scored Wake (Actual Class Positive)	True Positive (TP)	False Negative (FN)	Sensitivity (True Positive Rate, Recall): TP/(TP+FN)
Human expert scored Sleep (Actual Class Negative)	False Positive (FP)	True Negative (TN)	Specificity (True Negative Rate): TN/(TN+FP)
	Positive Predictive Value (Precision): TP/(TP+FP)	Negative Predictive Value: TN/(TN+FN)	Accuracy: (TP+TN)/(TP+TN+FP+FN)

Psihodium M, Stefani A, Penzel T, Ibrahim A, Hackner H, Heidbreder A, Szentkirályi A, Stubbe B, Völzke H, Berger K, Högl B. Interrater sleep stage scoring reliability between manual scoring from two European sleep centers and automatic scoring performed by the artificial intelligence-based Stanford-STAGES algorithm. J Clin Sleep Med. 2021 Jun 1;17(6):1237-1247. doi: 10.5664/jcsm.9174. PMID: 33599203; PMCID: PMC8314654.

(a) Sleep staging

(b) Apnea detection — r squared = 0.85

(c) AHI categorization

(d) Limb movement detection — r squared = 0.79

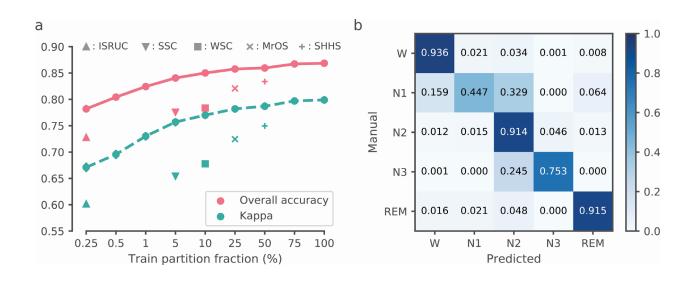**B, D simple correlation** of AHI and LMI (algorithm predicted versus 'true' expert scored)

A, C confusion matrices (probability that model predicts sleep stage/AHI class i, when truth [human scoring] is j;
allows identification of misclassification tendencies

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

Biswal S, Sun H, Goparaju B, Westover MB, Sun J, Bianchi MT. Journal of the American Medical Informatics Association. 2018 Dec;25(12):1643-50.

How can we improve autoscoring algorithm performance?
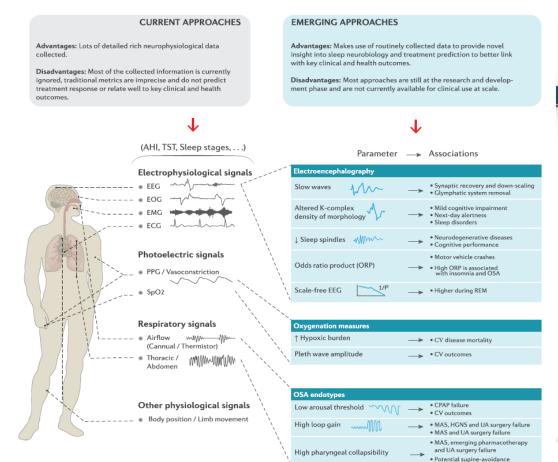
AI Beyond Autoscoring

Can we use AI to get more from what we already record?

Lechat B, Scott H, Naik G, Hansen K, Nguyen DP, Vakulin A, Catcheside P, Eckert DJ. New and Emerging Approaches to Better Define Sleep Disruption and Its Consequences. Front Neurosci. 2021 Oct 7;15:751730. doi: 10.3389/fnins.2021.751730. Erratum in: Front Neurosci. 2021 Nov 18;15:804589. PMID: 34690688; PMCID: PMC8530106.
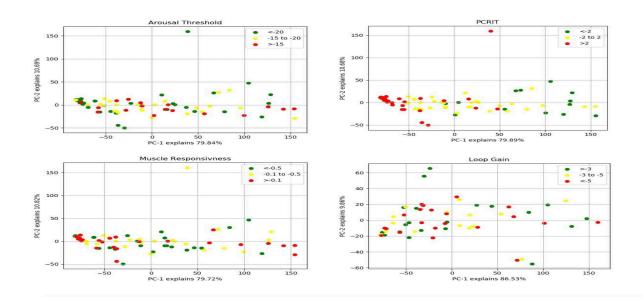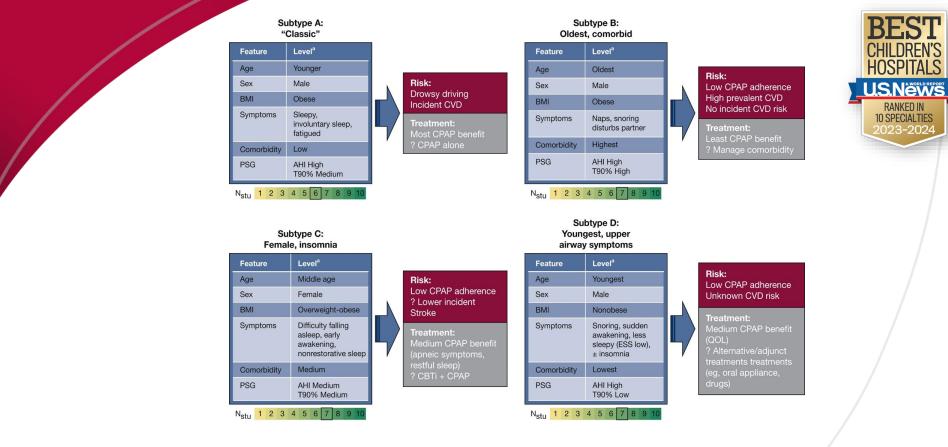
Endotyping OSA severity

Dutta R, Delaney G, Toson B, et al. A Novel Model to Estimate Key Obstructive Sleep Apnea Endotypes from Standard Polysomnography and Clinical Data and Their Contribution to Obstructive Sleep Apnea Severity. *Annals of the American Thoracic Society.* 2021;18(4):656-667.

Cluster analysis to identify OSA phenotype subgroups → outcomes and treatment response
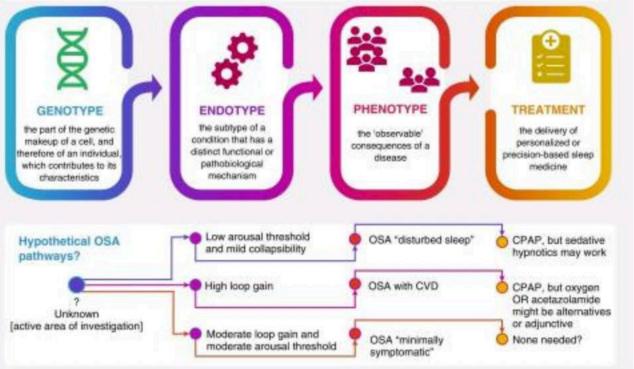
Future of AI driven personalized treatment of OSA

Edwards BA, Redline S, Sands SA, Owens RL. 2019 Sep 15;200(6):691-703.
Javaheri S, Brown LK, Abraham WT, Khayat R. AChest. 2020;157(2):394-402

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

41

Identifying sleep neurophysiological parameters associated with cognitive function in older adults

| 173 objective sleep metrics in 2 cohorts | dimension reduction/cluster analysis and principal spectral component analysis → | Neuropsychological measures |

Macro structure

REM duration

sleep efficiency

Micro structure

sigma/beta ratio
fast spindle activity
slow spindle activity
slow oscillation (SO)
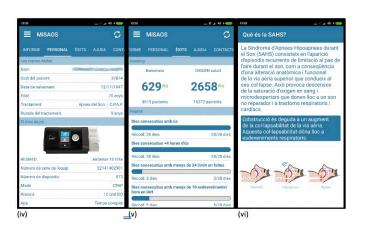duration
Absolute total (slow)
power
Slow, delta relative power

- micro-architecture prediction of cognitive performance <u>INDEPENDENT</u> of macrostructure

- subjective measures with only <u>WEAK</u> associations

Djonlagic I, Mariani S, Fitzpatrick AL, Van Der Klei VMGTH, Johnson DA, Wood AC, Seeman T, Nguyen HT, Prerau MJ, Luchsinger JA, Dzierzewski JM, Rapp SR, Tranah GJ, Yaffe K, Burdick KE, Stone KL, Redline S, Purcell SM. Macro and micro sleep architecture and cognitive performance in older adults. Nat Hum Behav. 2021 Jan;5(1):123-145. doi: 10.1038/s41562-020-00964-y. Epub 2020 Nov 16. Erratum in: Nat Hum Behav. 2020 Dec 16;. PMID: 33199858.
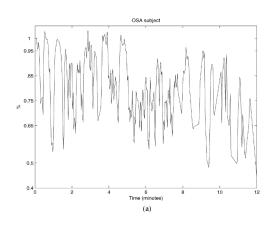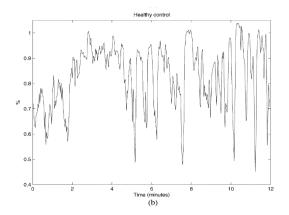
PAP adherence

Turino C, Benítez ID, Rafael-Palou X, Mayoral A, Lopera A, Pascual L, Vaca R, Cortijo A, Moncusí-Moix A, Dalmases M, Vargiu E, Blanco J, Barbé F, de Batlle J. Management and Treatment of Patients With Obstructive Sleep Apnea Using an Intelligent Monitoring System Based on Machine Learning Aiming to Improve Continuous Positive Airway Pressure Treatment Compliance: Randomized Controlled Trial. J Med Internet Res. 2021 Oct 18;23(10):e24072. doi: 10.2196/24072. PMID: 34661550; PMCID: PMC8561405.

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

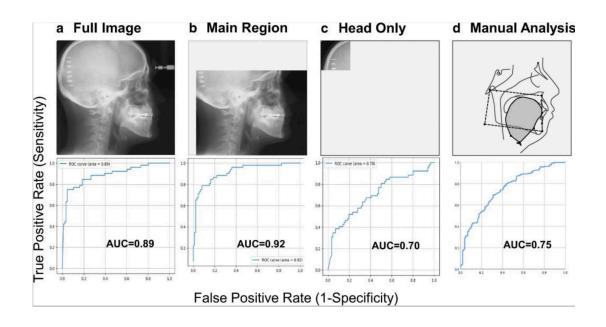# Diagnosis of OSA based on alternative parameters

Liu D, Pang Z, Lloyd SR. A neural network method for detection of obstructive sleep apnea and narcolepsy based on pupil size and EEG. IEEE Trans Neural Netw. 2008 Feb;19(2):308-18. doi: 10.1109/TNN.2007.908634. PMID: 18269961.

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

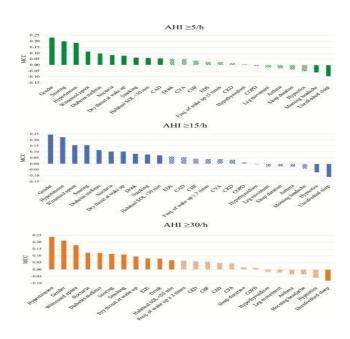# Diagnosis of OSA based on alternative parameters

Tsuiki S, Nagaoka T, Fukuda T, Sakamoto Y, Almeida FR, Nakayama H, Inoue Y, Enno H. Machine learning for image-based detection of patients with obstructive sleep apnea: an exploratory study. Sleep Breath. 2021 Dec;25(4):2297-2305. doi: 10.1007/s11325-021-02301-7. Epub 2021 Feb 8. PMID: 33559004; PMCID: PMC8600677.

45

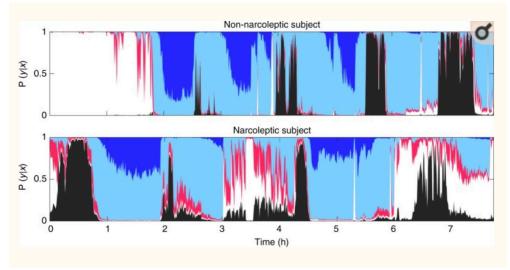# OSA screening using machine learning

Huang WC, Lee PL, Liu YT, Chiang AA, Lai F. Support vector machine prediction of obstructive sleep apnea in a large-scale Chinese clinical sample. *Sleep*. 2020;43(7).

Diagnosis of narcolepsy with single night PSG using hypnodensity features
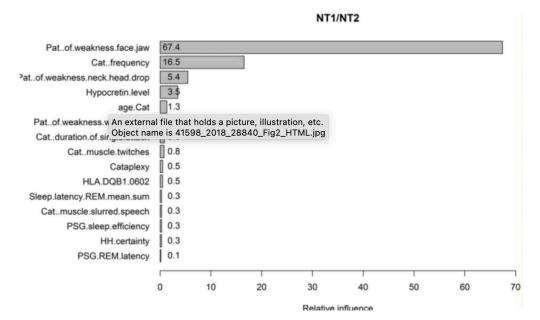

sensitivity of 91% and specificity of 96% compared to PSG-MSLT

Stephansen JB, Olesen AN, Olsen M, Ambati A, Leary EB, Moore HE, Carrillo O, Lin L, Han F, Yan H, Sun YL, Dauvilliers Y, Scholz S, Barateau L, Hogl B, Stefani A, Hong SC, Kim TW, Pizza F, Plazzi G, Vandi S, Antelmi E, Perrin D, Kuna ST, Schweitzer PK, Kushida C, Peppard PE, Sorensen HBD, Jennum P, Mignot E. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. Nat Commun. 2018 Dec 6;9(1):5229. doi: 10.1038/s41467-018-07229-3. PMID: 30523329; PMCID: PMC6283836.

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

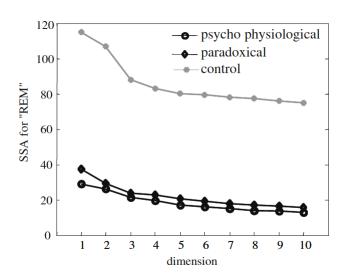# Applying ML to identify clinical features of narcolepsy from Network Database
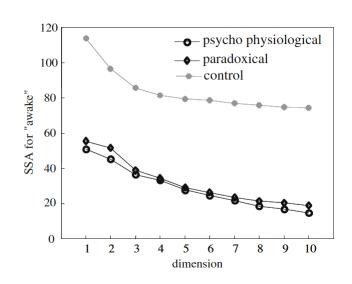
Zhang Z, Mayer G, Dauvilliers Y, Plazzi G, Pizza F, Fronczek R, Santamaria J, Partinen M, Overeem S, Peraita-Adrados R, da Silva AM, Sonka K, Rio-Villegas RD, Heinzer R, Wierzbicka A, Young P, Högl B, Bassetti CL, Manconi M, Feketeova E, Mathis J, Paiva T, Canellas F, Lecendreux M, Baumann CR, Barateau L, Pesenti C, Antelmi E, Gaig C, Iranzo A, Lillo Triguero L, Medrano-Martínez P, Haba-Rubio J, Gorban C, Luca G, Lammers GJ, Khatami R. Exploring the clinical features of narcolepsy type 1 versus narcolepsy type 2 from European Narcolepsy Network database with machine learning. Sci Rep. 2018 Jul 13;8(1):10628. doi: 10.1038/s41598-018-28840-w. PMID: 30006563; PMCID: PMC6045630

Distinguishing insomnia phenotypes using singular spectral analysis (SSA)

Aydın S, Saraoğlu HM, Kara S. Singular spectrum analysis of sleep EEG in insomnia. J Med Syst. 2011 Aug;35(4):457-61. doi: 10.1007/s10916-009-9381-7. Epub 2009 Sep 30. PMID: 20703545.

Delivery of digital behavioral intervention

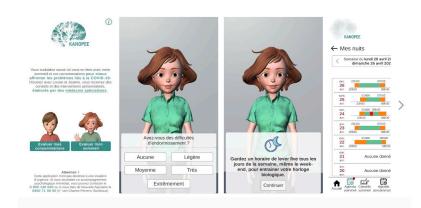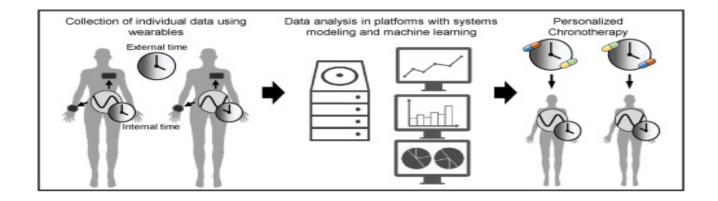# Predicting circadian rhythm for optimal timing of medications

Longitudinal data tracking for sleep health

Kim, Science, 2011; Yetisen. Adv Mat, 2018; Rao. Sensors, 2019

Summary of applications of AI in sleep medicine
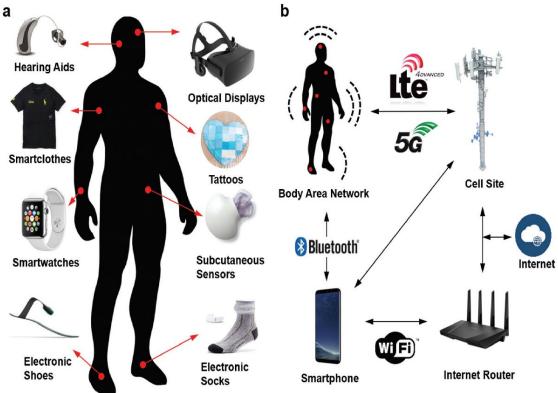
- Automated scoring of PSG (sleep stages, respiratory events, movement)

- Alternative screening methods

- Alternative diagnostic methods

- Unsupervised learning to reveal previously undescribed disease subphenotypes (role in outcomes)

- Identify likely disease endotypes (with relevance for personalized treatment)

- Longitudinal sleep assessments non-traditional sensors (consumer sleep technologies)

- Longitudinal patient generated data to guide precision, dynamic interventions; prediction

CHALLENGES OF USING AI IN SLEEP MEDICINE

Susan Redline[1,*] and Shaun M. Purcell[2]

[1]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA and [2]Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

Heterogeneity of:

physiological endotypes
symptoms
comorbidities
treatment responses
prognoses
genetic risk factors

→ Same diagnostic code

55

Why are we not using this technology?

-Lack of Diversity of Training data set

DESIGN AND DEVELOPMENT

- Heterogeneity in disease
- Difficult to generalize
- Lack of inclusive training data sets
- Disconnect between the problem and the solution
- Not all stake holders are involved early in the process

SELECTION AND USE

- One size may not fit all
- Reliability of Data
- Automation bias
- Assess the specific need and the infrastructure required to support the use

ONGOING SURVEILLANCE

- Usage
- Usability
- Interpretability
- Up-to-date capabilities
- Safety
- Unintended consequences
- Ethical issues/Biases
- Organizational/Social impacts

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

Riley

CHECKLIST TO EVALUATE AI SOLUTIONS FOR YOUR CLINICAL PRACTICE

✓ CLINICAL

✓ OPERATIONAL

✓ FINANCIAL

✓ DATA STORAGE AND SECURITY

✓ LEGAL

✓ QUALITY ASSURANCE

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

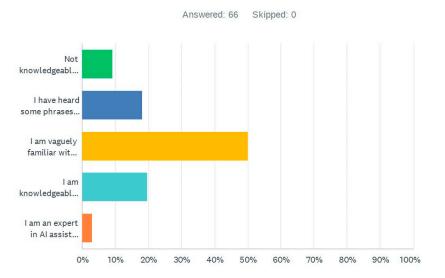# ARE WE READY FOR AI IN SLEEP MEDICINE?

Q4 How would you rate your knowledge about AI assisted algorithms used in clinical practice?

Answered: 66   Skipped: 0
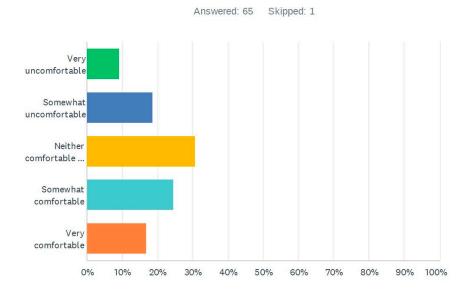
ANSWER CHOICES	RESPONSES	
Not knowledgeable at all	9.09%	6
I have heard some phrases and terms but don't know what those mean	18.18%	12
I am vaguely familiar with some basic concepts	50.00%	33
I am knowledgeable in AI assisted algorithms, know how they work and know how to evaluate them as clinical tools	19.70%	13
I am an expert in AI assisted algorithms	3.03%	2
TOTAL		66

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

61

Q6 How comfortable are you in adapting your clinical practice to integrate AI enabled clinical tools?

Answered: 65   Skipped: 1

ANSWER CHOICES	RESPONSES	
Very uncomfortable	9.23%	6
Somewhat uncomfortable	18.46%	12
Neither comfortable nor uncomfortable	30.77%	20
Somewhat comfortable	24.62%	16
Very comfortable	16.92%	11
TOTAL		65

INDIANA UNIVE
SCHOOL OF MEDICINE

Riley

AI as a helpful tool
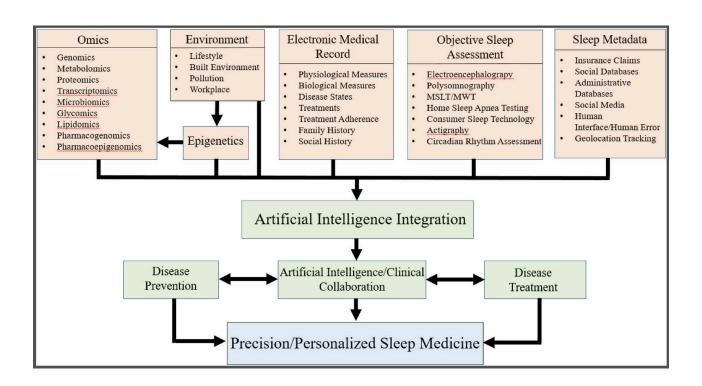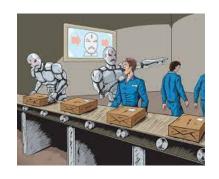
Service Components	AI Category: Assistive	AI Category: Augmentative	AI Category: Autonomous
Primary objective	Detects clinically relevant data	Analyzes and/or quantifies data in a clinically meaningful way	Interprets data and independently generates clinically meaningful conclusions
Provides independent diagnosis and/or management decision	No	No	Yes
Analyzes data	No	Yes	Yes
Requires physician or other QHP interpretation and report	Yes	Yes	No
Examples in CPT code set	Computer-aided detection (CAD) imaging (77048, 77049, 77065-77067, 0042T, 0174T, 0175T)	Continuous glucose monitoring (CGM) (95251), external processing of imaging data sets	Retinal imaging (92229)

Miles to go before we sleep

Thank you!

CHECKLIST TO EVALUATE AI SOLUTIONS FOR YOUR CLINICAL PRACTICE
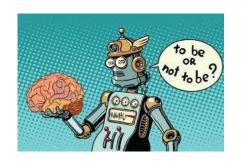
CLINICAL

- Is the AI solution solving a real clinical need?

- Will the AI solution be generalizable to my patients/population?

- Will the AI solution be an asset to key stakeholders? If so, to whom?

- How will the AI solution minimize psychometric biases (Table3)

- Will the AI solution be generalizable and which patients/populations might be missed?

CHECKLIST TO EVALUATE AI SOLUTIONS FOR YOUR CLINICAL PRACTICE

OPERATIONAL

- Is there a clear workflow that integrates the AI solution?
- Is the AI solution easy to use and implement for patients/population?
- Are the data outputs easy enough to manage?
- How will the workflow impact clinical staff and care coordination?
- Will additional training be needed? If so, to whom and how often?
- Are there coding and compliance implications?
- How will we report failures?

CHECKLIST TO EVALUATE AI SOLUTIONS FOR YOUR CLINICAL PRACTICE

FINANCIAL

- What is the purchasing and licensing model? Is it a onetime purchase with support or a subscription model?

- Will technical support and upgrades be included? If so, how often and for how long?

- Will there be additional revenue from the solution?

- Will costs be passed on to consumers, payors?

- What happens financially if the company either fails or is acquired?

DATA STORAGE AND SECURITY

- What safeguards are present to provide security and privacy? Is the database HIPAA compliant?

- Who will have access to the data, and who owns the data?

- What happens to the data if the company either fails or is acquired?

- Are the data properly labeled to be compatible with future applications?

- How much interoperability is present or needs to be created?

DATA STORAGE AND SECURITY

- Does an HL7 interface exist already with your current electronic health record (EHR), or will you help develop the HL7 interface which the manufacture will then license and sell?

- If an HL7 interface does not exist, where will the data be kept? This also means that providers will have to log into a separate program/database.

- Does an application programming interface (API) exist to facilitate data extraction from the company database into a database outside of an EHR?

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

CHECKLIST TO EVALUATE AI SOLUTIONS FOR YOUR CLINICAL PRACTICE

LEGAL

- If a solution is not FDA approved or cleared, are proper legal agreements in place for regulatory and legal resolutions that might arise?

- Do additional patient/population disclosures need to be obtained? If so, how, when, and how often?

- If the AI solution is used for research, do all parties have proper research training and credentialing? Does Institutional Review Board (IRB) need to be notified?

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

CHECKLIST TO EVALUATE AI SOLUTIONS FOR YOUR CLINICAL PRACTICE

LEGAL

- Are intellectual property (IP) rights in place? To what extent and are all parties aware?
- Are there real or perceived conflicts of interest (COI) that need to be disclosed?
- To patients
- To employers
- To regulatory entities
- Are legal processes outlined to deal with software and data if the company either fails or is sold/acquired?

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

QUALITY ASSURANCE

- How is success and failure measured?

- What data will be reviewed, how often, and by whom?

- What are the processes for review of unanticipated issues that might arise during the AI solution evaluation and implementation process?

- Is it possible to have an independent review committee for the data and processes involved?

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

QUALITY ASSURANCE

- Is there a governance structure for the solution, especially when/if scaled to a larger group? If so, are key stakeholders represented?

- What benchmarks exist to evaluate this solution?

- What are the unintended consequences of the solution on other populations or stakeholders?

- At the organizational level

- At the societal level

INDIANA UNIVERSITY
SCHOOL OF MEDICINE