The Cart Ran Over the Horse: Consumer Sleep Technologies in the Practice of Sleep Medicine

Cathy Goldstein, MD, MS
Associate Professor of Neurology
# Conflict of Interest Disclosures for Speakers

1. I do not have any relationships with any entities producing, marketing, re-selling, or distributing health care goods or services consumed by, or used on, patients, OR

2. I have the following relationships with entities producing, marketing, re-selling, or distributing health care goods or services consumed by, or used on, patients:

<table>
<thead>
<tr>
<th>Type of Potential Conflict</th>
<th>Details of Potential Conflict</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grant/Research Support</td>
<td></td>
</tr>
<tr>
<td>Consultant</td>
<td></td>
</tr>
<tr>
<td>Speakers’ Bureaus</td>
<td></td>
</tr>
<tr>
<td>Financial support</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
</tbody>
</table>

3. The material presented in this lecture has no relationship with any of these potential conflicts, OR

4. This talk presents material that is related to one or more of these potential conflicts, and the following objective references are provided as support for this lecture:

1. 
2. 
3. 
SLEEP 2019 Photography Policy

• Photography IS permitted during this lecture.
• Photography of slides featuring the icon on the left is not permitted.
• Photographs from this lecture are only allowed for personal, social, or non-commercial use.
• Attendees may not use flash photography or otherwise distract the presenters and/or attendees.
We haven’t yet figured out how to utilize consumer sleep technologies (CSTs) for clinical purposes...

But like it or not....

THEY ARE HERE
Even as we sit here...
USE CASE

WEARABLES, NEARABLES, APPS (associated with devices and alone) sleep tracking, “diagnostic”, interventions

CONSUMER FACING, WEARABLE SLEEP ESTIMATION DEVICES WITH MOTION AND HEART RATE SENSORS

NOT… A COMPREHENSIVE REVIEW OF EVERY CONSUMER SLEEP TECHNOLOGY
The sleep measurement paradox...

Sleep deprivation and sleep disorders exert symptoms through...

**CHRONIC** exposure to

**NIGHTLY** disturbances in sleep quality or duration

What is our gold-standard objective sleep measure?

**POLYSOMNOGRAM** great for OSA diagnosis and PAP titration;

**IMPRACTICAL** beyond 1-2 nights
Current longitudinal, objective sleep measurement: Actigraphy

FDA cleared as a class II medical device through the 510 (k) mechanism

Wrist worn accelerometer (solid state piezo-electric or more commonly NOW, Micro-Electro-Mechanical Systems [MEMS] type)

Awake movement > asleep movement

patient wears device → acceleration data digitized into activity counts → device is downloaded → software package uses an equation to estimate sleep

Output validated against PSG in MANY, MANY peer-reviewed publications

CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device) not
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device)

Just measures movement (sometimes light)
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device)

requires IN PERSON initialization and data recovery via a USB port!
No real time data transmission!
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

requires **IN PERSON** initialization and data recovery via a **USB port**! No real time data transmission!
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device)

Weighted sum algorithms used to distinguish sleep and wake
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device)

High sensitivity but low specificity for sleep (>90%, ≈50%); misinterprets non-moving wakefulness as sleep

sleep disruption ↑  accuracy ↓
**An aside on device validation**

**SENSITIVITY**

= number of patients with positive test who have disease (true positives)/all patients with disease

= number of 30S epochs scored as sleep by device that are sleep on PSG/all 30S sleep epochs on PSG

**SPECIFICITY**

= number of patients who have a negative test and are disease free (true negative)/all patients without the disease

= number of 30S epochs scored as wake by device that are wake on PSG/all 30s wake epochs on PSG

‘**ACCURACY**’ is deceiving in sleep validation in normal individuals - most of the overnight period is spent asleep; therefore an algorithm that arbitrarily scores every 30S as sleep (100% sensitive) will appear (deceptively) highly accurate.
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device)

Does not interface with EHR (or anything for that matter)
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device)

Set-up, recording, interpretation is not reimbursed; not practical for clinical workflow
CON’S OF TRADITIONAL ACTIGRAPHY

SINGLE BIOLOGICAL MEASURE Just

OLD DATA ACQUISITION AND STORAGE METHODS

RUDIMENTARY ANALYSIS OF DATA

VALIDATED, BUT NOT GOOD

CLOSED SOFTWARE

RESOURCE INTENSIVE

EXPENSIVE ($800+ a device)
‘It is strongly recommended that the MSLT be preceded by at least one week of actigraphic recording’

‘Sleep log, and whenever possible, actigraphy monitoring ...(DSWPD, ASWPD, ISWRD)

‘Daily sleep logs AND actigraphy, for at least 14 days.’(N24SWD)

‘Sleep log AND actigraphy monitoring (whenever possible and with concurrent light exposure measurement) (SWD)

Knowing the AASM recommendations and the pro's and con's of actigraphy...

In the evaluation of the patient with excessive daytime sleepiness (EDS) thought not secondary to sleep disordered breathing (SDB)....

WHAT REALLY HAPPENS IN CLINIC?
SELF-REPORT SLEEP MEASURES SUBSTITUTED

_for example:

“How many hours do you sleep, on average?”
SELF-REPORT SLEEP MEASURES SUBSTITUTED

for example:

“How many hours do you sleep, on average?”
“How many hours do you sleep at night?”
SELF-REPORT SLEEP MEASURES SUBSTITUED

for example:

“How many hours do you sleep, on average?”

“How many hours do you sleep at night?”

“How many hours do you sleep in a 24 hour period, including naps?”
SELF-REPORT SLEEP MEASURES SUBSTITUTED for example:

“How many hours do you sleep, on average?”
“How many hours do you sleep at night?”
“How many hours do you sleep in a 24 hour period, including naps?”
“What time do you go to bed and when do you wake up?”
SELF-REPORT SLEEP MEASURES SUBSTITUTED for example:

“How many hours do you sleep, on average?”
“How many hours do you sleep at night?”
“How many hours do you sleep in a 24 hour period, including naps?”
“What time do you go to bed and when do you wake up?”
“During the work week? During the weekend?”
**LONGITUDINAL SELF-REPORT**

---

### Two Week Sleep Diary

**Instructions:**
1. Write the date, day of the week, and type of day: Work, School, Day Off, or Vacation.
2. Put the letter "C" in the box when you have coffee, tea, or alcohol. Put "M" when you take any medication. Put "A" when you drink alcohol. Put "E" when you exercise.
3. Put a line (I) to show when you go to bed. Shade in the box that shows when you think you fell asleep.
4. Shade in all the boxes that show when you are asleep at night or when you take a nap during the day.
5. Leave boxes unshaded to show when you wake up at night and when you are awake during the day.

**Sample Entry Below:**
On a Monday when I worked, I logged on my lunch break at 1 PM, had a glass of wine with dinner at 6 PM, and asked watching TV from 7 to 9 PM, went to bed at 10:30 PM; fell asleep around Midnight, woke up and couldn’t go back to sleep at about 4 AM, went back to sleep from 5 to 7 AM, and had coffee and medicine at 7:00 in the morning.

<table>
<thead>
<tr>
<th>Today’s date</th>
<th>Sample</th>
<th>Consensus Sleep Diary Core</th>
<th>ID/Name:__</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/5/14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. What time did you go to bed?</td>
<td>10:15 PM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. What time did you try to go to sleep?</td>
<td>11:30 PM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. How long did it take you to fall asleep?</td>
<td>55 min.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. How many times did you wake up, not counting your final awakening?</td>
<td>3 times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. In total, how long did these awakenings last?</td>
<td>1 hour 10 min.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. What time was your final awakening?</td>
<td>6:35 AM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. What time did you get out of bed for the day?</td>
<td>7:20 AM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. How would you rate the quality of your sleep?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very poor</td>
<td>Very poor</td>
<td>Very poor</td>
</tr>
<tr>
<td></td>
<td>Fair</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>9. Comments (if applicable)</td>
<td>I have a cold</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

LONGITUDINAL SELF-REPORT

Patient burden
Incomplete
More difficult for certain populations

We suggest that clinicians use actigraphy to estimate sleep parameters in adult patients with insomnia disorder. [Conditional]

We suggest that clinicians use actigraphy in the assessment of pediatric patients with insomnia disorder. [Conditional]

We suggest that clinicians use actigraphy in the assessment of adult patients with circadian-rhythm sleep-wake disorder. [Conditional]

We suggest that clinicians use actigraphy in the assessment of pediatric patients with circadian-rhythm sleep-wake disorder. [Conditional]

We suggest that clinicians use actigraphy integrated with home sleep apnea test devices to estimate total sleep time during recording [in the absence of alternative objective measurements of total sleep time] in adult patients suspected of sleep-disordered breathing. [Conditional]

We suggest that clinicians use actigraphy to monitor total sleep time prior to testing with the Multiple Sleep Latency Test in adult and pediatric patients with suspected central disorders of hypersomnolence. [Conditional]

We suggest that clinicians use actigraphy to estimate total sleep time in adult patients with suspected insufficient sleep syndrome. [Conditional]

We recommend that clinicians not use actigraphy in place of electromyography for the diagnosis of periodic limb movement disorder in adult and pediatric patients. [Strong]
OBJECTIVE SLEEP MEASURES PREFERRED OVER SELF-REPORT FOR RESEARCH PURPOSES

Hypothesis generation of mechanisms behind this relationship is difficult because of the uncertainty of measurement; how do we educate people?

U-shaped association with self-report sleep duration and multiple conditions and mortality


Editorial
An open request to epidemiologists: please stop querying self-reported sleep duration
SEEMS LIKE THERE IS ANY EASY SOLUTION TO THE PROBLEM OF LONGITUDINAL, OBJECTIVE SLEEP MEASUREMENT
CONSUMER SLEEP TRACKING TECHNOLOGIES!

Wrist worn sleep estimation devices - Fitbit, Withings, Polar, MisFit, Xiaomi

Rings - Oura, Thim

Dry EEG headbands - Dreem, Muse

‘Nearables’ - Beddit, SleepScore

NOT ONLY DO THESE TRACK, THEY INTERVENE!
How do motion and heart rate based sleep trackers work?

MEMS triaxial accelerometers considered valid sensors for the measurement of physical activity and energy expenditure used in 100’s of peer-reviewed works in the exercise science literature.

Consumer sleep tracking technologies-PROs

Small
Cheap
Ubiquitous
Owned by consumer not health system/investigator
Multisensor (acceleration + heart rate at minimum, temp)
Connectivity via bluetooth
Passive sleep detection mode
Modern computational techniques like machine learning (more flexible than weighted sums!)
Consumer sleep tracking technologies—PROs

Small
Cheap
Ubiquitous
Owned by consumer, not health system/investigator
Multisensor (acceleration + heart rate at minimum)
Connectivity via Bluetooth
Passive sleep detection mode
Modern computational techniques like machine learning (more flexible than weighted sums!)

So what is the problem?
Consumer sleep tracking technologies-CONs

Output from CST devices typically **NOT** validated as an accurate measure of sleep

Hardware, firmware, **AND** software change more rapidly than we can validate them
### Consumer sleep tracking technologies-CONs

<table>
<thead>
<tr>
<th>CLASSIC</th>
<th>FORCE</th>
<th>BLAZE</th>
<th>ALTA HR</th>
<th>CHARGE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ULTRA</td>
<td>CHARGE</td>
<td>ALTA</td>
<td>ICONIC</td>
<td>INSPIRE</td>
</tr>
<tr>
<td>ONE</td>
<td>CHARGE HR</td>
<td>CHARGE 2</td>
<td>VERSA</td>
<td>INSPIRE HR</td>
</tr>
<tr>
<td>FLEX</td>
<td>SURGE</td>
<td>FLEX 2</td>
<td>ACE</td>
<td>VERSA LIGHT</td>
</tr>
</tbody>
</table>

**20 DEVICES IN 10 YEARS FROM ONE MANUFACTURER!**

*This doesn’t even take any algorithm changes into account*
FOR COMPARISON  

<table>
<thead>
<tr>
<th>Year</th>
<th>First author</th>
<th>Device (s)</th>
<th>Performance *</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>Montgomery</td>
<td>Fitbit &quot;original&quot;</td>
<td>Se 0.98, Sp 0.20</td>
</tr>
<tr>
<td></td>
<td>Downs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>Meltzer</td>
<td>Fitbit Ultra</td>
<td>Se 0.87; 0.70, Sp 0.52; 0.70</td>
</tr>
<tr>
<td>2015</td>
<td>Toon</td>
<td>Jawbone UP</td>
<td>Se 0.92, Sp 0.66</td>
</tr>
<tr>
<td>2015</td>
<td>De Zambotti</td>
<td>Jawbone UP</td>
<td>Se 0.96, Sp 0.37</td>
</tr>
<tr>
<td>2016</td>
<td>De Zambotti</td>
<td>Fitbit Charge HR</td>
<td>Se 0.97, Sp 0.42</td>
</tr>
<tr>
<td>2017</td>
<td>Cook</td>
<td>Fitbit Flex</td>
<td>Se 0.98; 0.78, Sp 0.35; 0.80</td>
</tr>
<tr>
<td>2017</td>
<td>Kang</td>
<td>Fitbit Flex</td>
<td>Se 0.97; 0.65, Sp 0.36; 0.82</td>
</tr>
<tr>
<td>2017</td>
<td>Maskevich</td>
<td>Jawbone UP2</td>
<td>Se 0.99, Sp 0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fitbit One</td>
<td>Se 0.99, Sp 0.27</td>
</tr>
<tr>
<td>2017</td>
<td>De Zambotti</td>
<td>OURA ring</td>
<td>Se 0.96, Sp 0.48; N1 + N2 =0.65, N3 0.51, REM 0.61</td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>De Zambotti</td>
<td>Fitbit Charge 2</td>
<td>Se 0.96, Sp 0.61; NREM REM cycles 0.82%</td>
</tr>
<tr>
<td>2018</td>
<td>Cook</td>
<td>Jawbone UP3</td>
<td>Se 0.97, Sp 0.39; N1 + N2 =0.60, N3=0.49, REM=0.30</td>
</tr>
<tr>
<td>2018</td>
<td>Cook</td>
<td>Fitbit Alta HR</td>
<td>Se 0.96, Sp 0.58; N1 + N2 =0.73, N3=0.67, REM=0.74</td>
</tr>
</tbody>
</table>

* Epoch-by-epoch analyses. Se=sensitivity, Sp=specificity; sleep staging performance summarized by agreement

6 YEARS OF PEER REVIEWED VALIDATION OF WEARABLE CONSUMER SLEEP TECHNOLOGIES AGAINST PSG

12 manuscripts

Sensitivity 87-99%, Specificity 20-66% (normal modes)

Sleep staging agreement N1+N2=60-73%, N3 49-67%, REM 30-74%

11 DEVICES → Jawbone OUT

→ 8 DEVICES
Consumer sleep tracking technologies—CONs

20 devices in 10 years from one manufacturer!

This doesn’t even take any algorithm changes into account
Even if we continue with the same validation framework

Manufacturers **DON’T DISCLOSE**

*algorithm methodology AND don’t allow raw data access*—validation studies are **specific to the device and app as a whole, not generalizable**

*characteristics of population* used to train algorithms—devices and algorithms are **population specific**, not generalizable unless tested on various groups of patients
“Given the lack of validation and United States Food and Drug Administration (FDA) clearance, CSTs cannot be utilized for the diagnosis and/or treatment of sleep disorders at this time…The ubiquitous nature of CSTs may further sleep research and practice. However, future validation, access to raw data and algorithms, and FDA oversight are needed.”
<table>
<thead>
<tr>
<th>Year</th>
<th>First Author</th>
<th>Device(s)</th>
<th>Population</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>Meltzer</td>
<td>Fitbit Ultra</td>
<td>Children/adolescents with suspected SDB</td>
<td>↑ PSG-device discrepancy with ↑ OSA severity</td>
</tr>
<tr>
<td>2015</td>
<td>Toon</td>
<td>Jawbone UP</td>
<td>Children/adolescents with suspected SDB</td>
<td>No differences in PSG-device discrepancy with OSA severity, but many outside a priori acceptable ranges</td>
</tr>
<tr>
<td>2015</td>
<td>De Zambotti</td>
<td>Jawbone UP</td>
<td>Adults (F) with insomnia</td>
<td>No differences in PSG-device discrepancy between insomnia patients and normals.</td>
</tr>
<tr>
<td>2017</td>
<td>Kang</td>
<td>Fitbit Flex</td>
<td>Adults with insomnia</td>
<td>TST, SE overestimated by device insomnia patients; frequency of acceptable PSG-device agreement significantly lower in insomnia patients</td>
</tr>
<tr>
<td>2018</td>
<td>De Zambotti</td>
<td>Fitbit Charge 2</td>
<td>Adults with PLMS</td>
<td>No differences in PSG-device discrepancy between PLMS patients and normal</td>
</tr>
<tr>
<td>2018</td>
<td>Cook</td>
<td>Jawbone UP3</td>
<td>Adults with suspected central disorders of hypersomnolence and mild OSA</td>
<td>No differences PSG-device discrepancy difference between narc type II, IH, hypersomnia NOS, OSA</td>
</tr>
<tr>
<td>2018</td>
<td>Cook</td>
<td>Fitbit Alta HR</td>
<td>Adults with suspected disorders of hypersomnolence</td>
<td>No differences PSG-device discrepancy difference between narc type II, IH, hypersomnia NOS</td>
</tr>
</tbody>
</table>
Can we take the **RAW** motion and heart rate data available to developers from the Apple Watch and create our own PSG verified models to estimate sleep?
Ambulatory recording

Final night simultaneous recording with in lab PSG

Well-validated circadian prediction model

"Clock proxy"

“Clock proxy”

Motion during PSG night

PPG heart rate during PSG night

ALGORITHM TRAINING WITH GROUND TRUTH PSG

- Logistic regression
- Machine learning techniques
  - random forest
  - k-nearest neighbors
  - neural net

WAKE
SLEEP
(NREM, REM)

Walch O, Huang Y, Forger D, Goldstein C. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. Submitted.
Checked our work...on a novel data set

Multi-Ethnic Study of Atherosclerosis

- **What**: Multi-center longitudinal investigation of factors associated with cardiovascular disease
- **Who**: 6,834 black, white, Hispanic, and Chinese-American men and women
- **When**: 2000 to present
- **Funding**: National Heart, Lung, and Blood Institute

Motion from actigraphy on PSG night (activity counts)
Heart rate from oximeter on PSG night
“Clock proxy” from ambulatory actigraphy

Best model determined from our data (MLP w/ motion, HR, clock)

WAKE
SLEEP (NREM, REM)

Compare to PSG epoch-by-epoch

https://sleepdata.org/datasets/mesa
Measurement and Analysis of Sleep and Circadian Dimensions

1:45 PM – 2:45 PM | Room 217AB
Co-Chairs: Gregory Belenky, MD and Rebecca Robillard

2:00 PM - 2:15 PM
Sleep Stage Prediction with Raw Acceleration and Photoplethysmography Heart Rate Data Derived from a Consumer Wearable Device
Walch O, Huang Y, Forger D, Goldstein C
If the cart has already come and ran over the horse…

WHERE DO WE GO FROM HERE?
Approach to the future use consumer sleep technologies in the practice of sleep medicine - PRACTICAL

RELATIVELY low risk measure without insurance reimbursement - harness already available sensors?

Digital Health Software Precertification (Pre-Cert) Program

“manufacturers who have demonstrated a robust culture of quality and organizational excellence”

Apple
FitBit
Johnson & Johnson
Pear therapeutics
Phosphorus

Roche
Samsung
Tidepool
Verily
Approach to the future use consumer sleep technologies in the practice of sleep medicine - REPRODUCIBLE AND TRANSPARENT

Can’t just use black box ML in isolation - *a priori* knowledge of the homeostatic and circadian regulation of sleep could benefit classifiers

Open source code? Unlikely from manufacturers

We have provided a github link to our code in our submitted paper
Approach to the future use consumer sleep technologies in the practice of sleep medicine - GENERALIZABLE

Flexibility between tri-axial acceleration and activity counts allows the POTENTIAL for previously acquired cohorts that co-recorded PSG and actigraphy to act as testing sets for algorithms developed with modern devices.

Analyze cohorts with different collection methods with the same algorithms.

Disclose characteristics of training and testing datasets to understand ability to generalize to certain individuals.

Approach to objective, longitudinal sleep tracking—VALIDATED? VERIFIED?

Epoch-by-epoch validation of device output against PSG

Summary PSG vs. Summary wearable metrics (TST, etc)

MISLEADING IN HEALTHY SLEEPERS

Comparison only to actigraphy → INSUFFICIENT, surrogate to surrogate
WHAT would be possible if we had consumer facing wearables using sensors that adhered to certain technical standards, allowed access to raw signal, and made algorithm code open source?

CROWDSOURCE validation to researchers all over world
GOOD
Sleep health ← both objective and self-report combine wearable with mobile app used for both
Monitor other measures of health along with sleep (‘EKG’)
LONG term data collection not possible with health system owned, non-consumer adopted actigraphy
Precise and real time inventions
Increase the reach of sleep medicine in underserved areas

BAD
device and/or associated app delivering unvetted interventions
patient impressions-I don’t feel good but device says sleeping well so don’t seek care (and opposite)
physician impressions-over evaluation when nothing is wrong (CPAP download data) or failure to further evaluate based on device data

UGLY
Data collection ≠ data review litigation if bad outcome (driving drowsy accident due to sleep deprivation, occupational issues)
Privacy issues
Modern wearables use machine learning algorithms, dependent on training data, overrepresentation groups WHO HAVE ACCESS TO DEVICES → inaccurate in underrepresented groups