Mobile Health

PPG (2)

Cecilia Mascolo

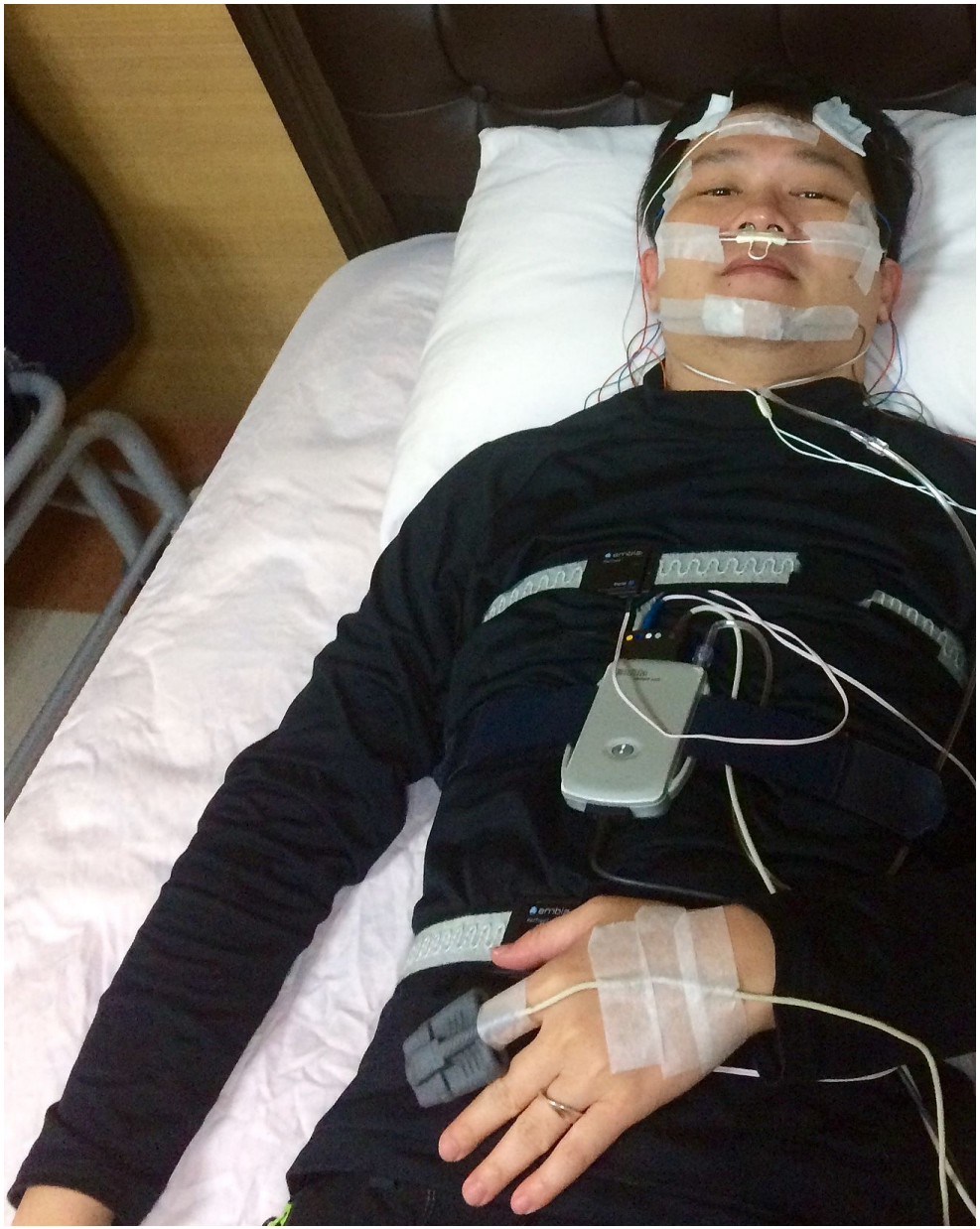
UNIVERSITY OF CAMBRIDGE

Sleep Epochs

- 30-second duration epochs usually manually labelled.
- There are up to six sleep stages:
  - awake;
  - rapid eye movement sleep (REM);
  - non-rapid eye movement (Non-REM);
  - sleep stage NREM 1 (N1);
  - sleep stage NREM 2 (N2);
  - sleep stage NREM 3 (N3).
- Each night has 4-6 cycles (90mins)

PSG: Sleep Monitoring Gold Standard

- Polysomnography (PSG) is a multi-sensor approach
  - electroencephalography (EEG),
  - electromyography (EMG)
  - electrooculography (EOG
- Together facilitate the measurement of brain activity, alongside both muscle and eye movement.
- Measurements of respiratory and cardiac activity are also often included.



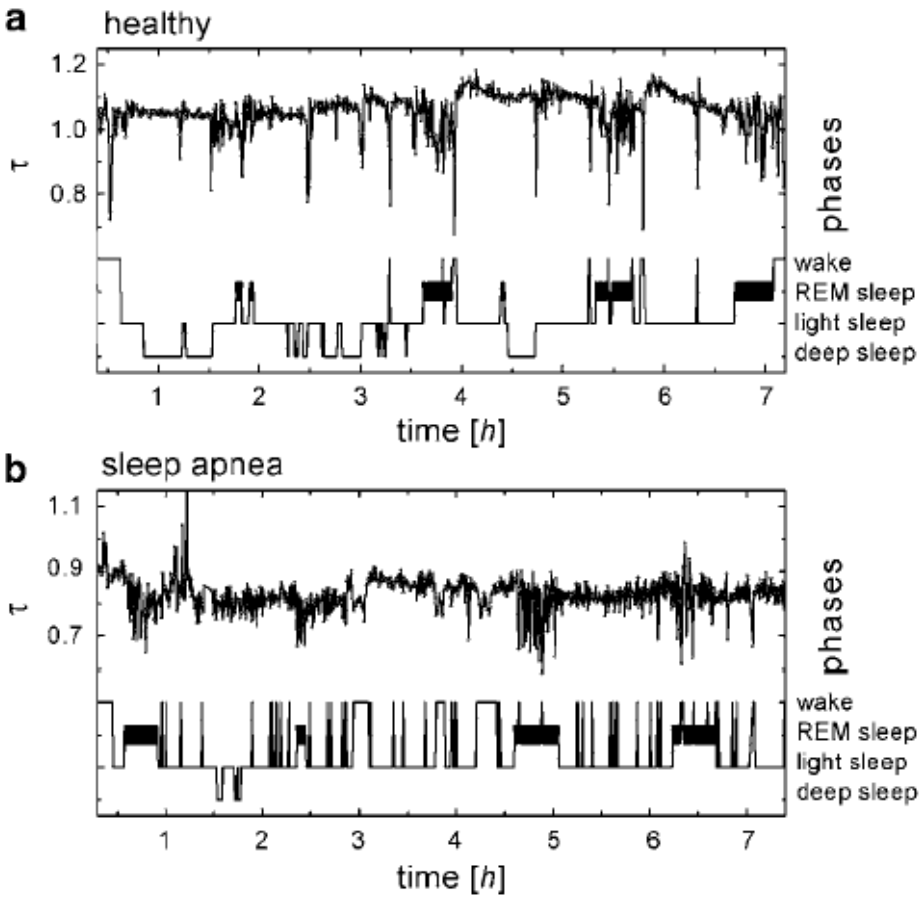
Alternatives…

- Basic measures of sleep eg duration, awake episodes, sleeping pattern, bedtime routine, perception after sleep.
- These could be asked through a questionnaire/sleep diary.
- Or
- Smartphones basic features:
  - Has the phone been used (when).
  - Accelerometer of phone placed on mattress can measure movements.
  - Microphone (for sleep apnea).
- Under mattress (acceleration/pressure).
- Contactless Radio (its own lesson).

Sleep Apnea

- Sleep Apnea occurs when throat muscles relax and block air flow to the lungs.

Y axis a measure of IBI



UNIVERSITY OF CAMBRIDGE

Penzel, T., Kantelhardt, J., Lo, CC. *et al.* Dynamics of Heart Rate and Sleep Stages in Normals and Patients with Sleep Apnea. *Neuropsychopharmacol* **28** (Suppl 1), S48–S53 (2003).

Apps vs PSG: not very good!

Application	n	Polysomnography								
		Sleep efficiency	Wake	Light sleep	Deep sleep	Sleep onset	Wake time %	Light sleep %	Deep sleep %	Snore time
Good Sleep!	45	0.058					0.034			
MotionX	64	−0.004	−0.04	0.186	−0.01					
Sleep Analyzer	64	0.222	0.153	0.129	0.121		0.268*	0.021	0.055	
Sleep Cycle	52	−0.262								
Sleep Time	65	0.006	−0.11	0.07	0.000		−0.131	-0.371*	−0.302*	
Smart Sleep	48				−0.12				−0.195	
SnoreClock	47									0.483***
SnoreLab	54									0.475**
WakeApp Pro	56	0.02	0.018	0.058	-0.06	0.307*				

* $p < 0.05$
** $p \leq 0.01$
*** $p \leq 0.001$

UNIVERSITY OF CAMBRIDGE

Kim K, Park DY, Song YJ, Han S, Kim HJ. Consumer-grade sleep trackers are still not up to par compared to polysomnography. Sleep Breath. 2021 Nov 5.

Actigraphy

- Use of accelerometers for measuring movement while sleeping
- Movement patterns associated with awake/sleep
  - Unable to detect motionless awake

PPG!

- Wearables with accelerometers and PPG can measure motion as well as HR, PRW (pulse rate variability).
- PPG-based wearables identify wake and sleep with a performance similar to, or better than, research-grade actigraphy devices.
- Sleep stages monitoring is still accurate enough…

UNIVERSITY OF CAMBRIDGE

PPG on wearables vs PSG

Wearable	*n*	PSG					
		TST	Light sleep	Deep sleep	REM sleep	WASO	SE
Mi Band2	55	r = 0.367 ** ICC = 0.297*	r = 0.032 ICC = 0.024	r = 0.116 ICC = 0.095		r = 0.383** ICC = 0.148	
Gear Fit2	54	r = 0.307 * ICC = 0.209					r = 0.18 ICC = 0.048
Fitbit Alta HR	61	r = 0.466 *** ICC = 0.205	r = 0.179 ICC = -0.19	r = 0.372** ICC= 0.301**	r = 0.310* ICC= 0.323**	r = 0.425*** ICC = 0.213*	

Light sleep: sum of N1 and N2 sleep

Deep sleep: N3 sleep

TST: Total sleep time

REM: Rapid eye movement

WASO: Wake time after sleep onset

SE: Sleep efficiency

r: Spearman's rho

ICC: Intraclass correlation coefficient

* $p < 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

The deep sleep duration, REM sleep duration, and WASO measured by the Fitbit Alta HR significantly correlated with PSG results, although Fitbit Alta HR underestimated the duration of light sleep when compared to PSG (253 min vs 287 min).

UNIVERSITY OF CAMBRIDGE

How do we calculate Sleep Stages from PPG? An example

- Sleep epochs of 30s.
- Motion features
  - Activity count over epoch (e.g. integrated area under the accelerometer signal).
  - Accelerometer magnitude.
  - Time since last significant movement.
  - Time till next significant movement.

Estimation of sleep stages in a healthy adult population from optical plethysmography and accelerometer signals. Z Beattie, Y Oyang, A Statan, A Ghoreyshi, A Pantelopoulos, A Russell and C Heneghan.. 2017 Institute of Physics and Engineering in Medicine
Physiological Measurement, Volume 38, Number 11.

Heart Rate Variability Features

- Inter Beat Interval to calculate Pulse Rate Variability:
- High Frequency (eg through DFT) 0.15–0.4 Hz
- Low Frequency 0.04–0.15 Hz
- VLF power (0.015–0.04 Hz)
- RMSSD: Root mean square of successive differences of IBI
- pNN50: proportion of successive IBIs that differ more than 50ms over total IBIs
- Delta IBIs
- Mean heart rate
- 90th percentile heart rate
- 10th percentile heart rate

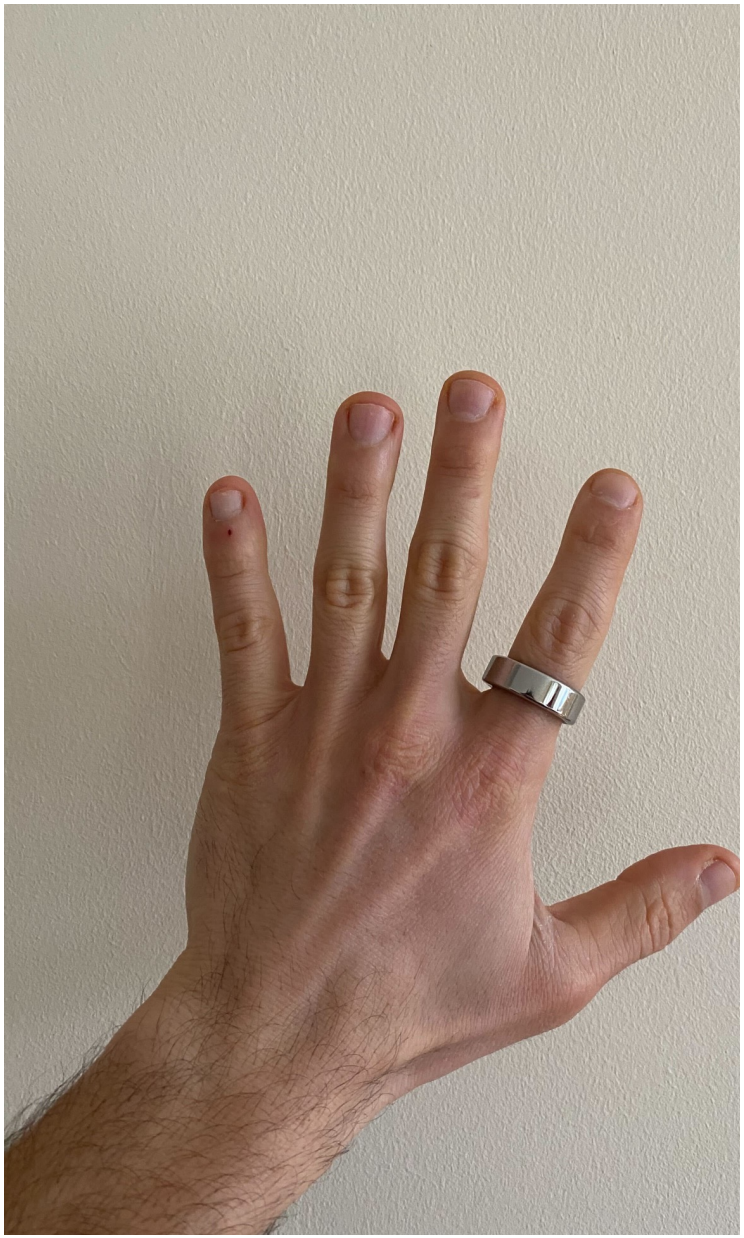
Breathing Features

1s breathing sample: take the frequency spectrum (and limit the power of frequency to plausible breathing frequencies).

- HF power (0.15–0.4 Hz)
- LF power (0.04–0.15 Hz)
- VLF power (0.015–0.04 Hz)

New wearables and sleep…

- Oura Ring
- Oura underestimated TST and overestimated WASO.
- Oura significantly underestimated REM sleep and light sleep (stage N1+N2), and overestimated time spent in deep sleep (stage N3)



Dovepress
open access to scientific and medical research
Nature and Science of Sleep
Dove Medical Press | This Article | Subscribe | Submit a Manuscript | Search | Follow

Nat Sci Sleep. 2021; 13: 177–190.
Published online 2021 Feb 15. doi: 10.2147/NSS.S286070
PMCID: PMC7894804
PMID: 33623459

Multi-Night Validation of a Sleep Tracking Ring in Adolescents Compared with a Research Actigraph and Polysomnography

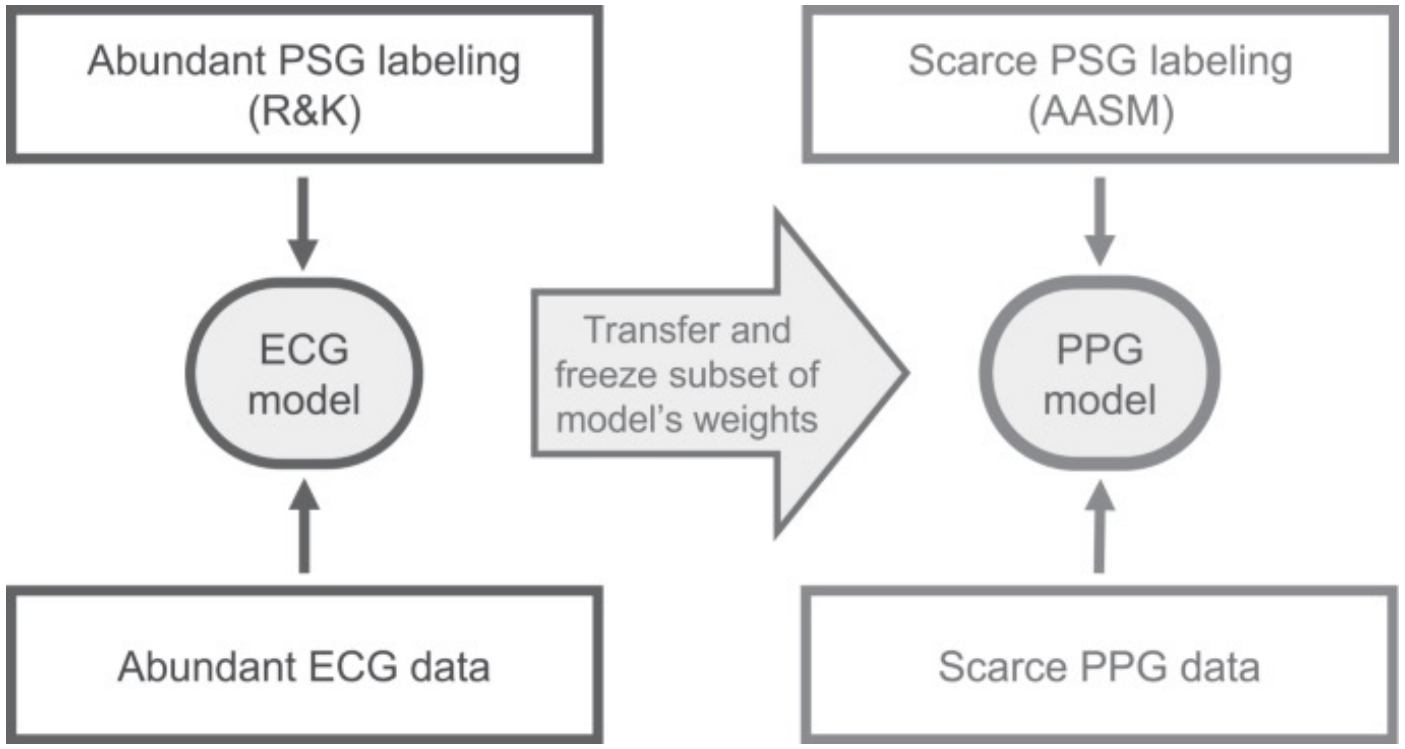
Nicholas I Y N Chee,[#1,*] Shohreh Ghorbani,[#1,*] Hosein Aghayan Golkashani,[1] Ruth L F Leong,[1] Ju Lynn Ong,[1] and Michael W L Chee[1]

Nat Sci Sleep

Deep learning approaches over PPG Sleep

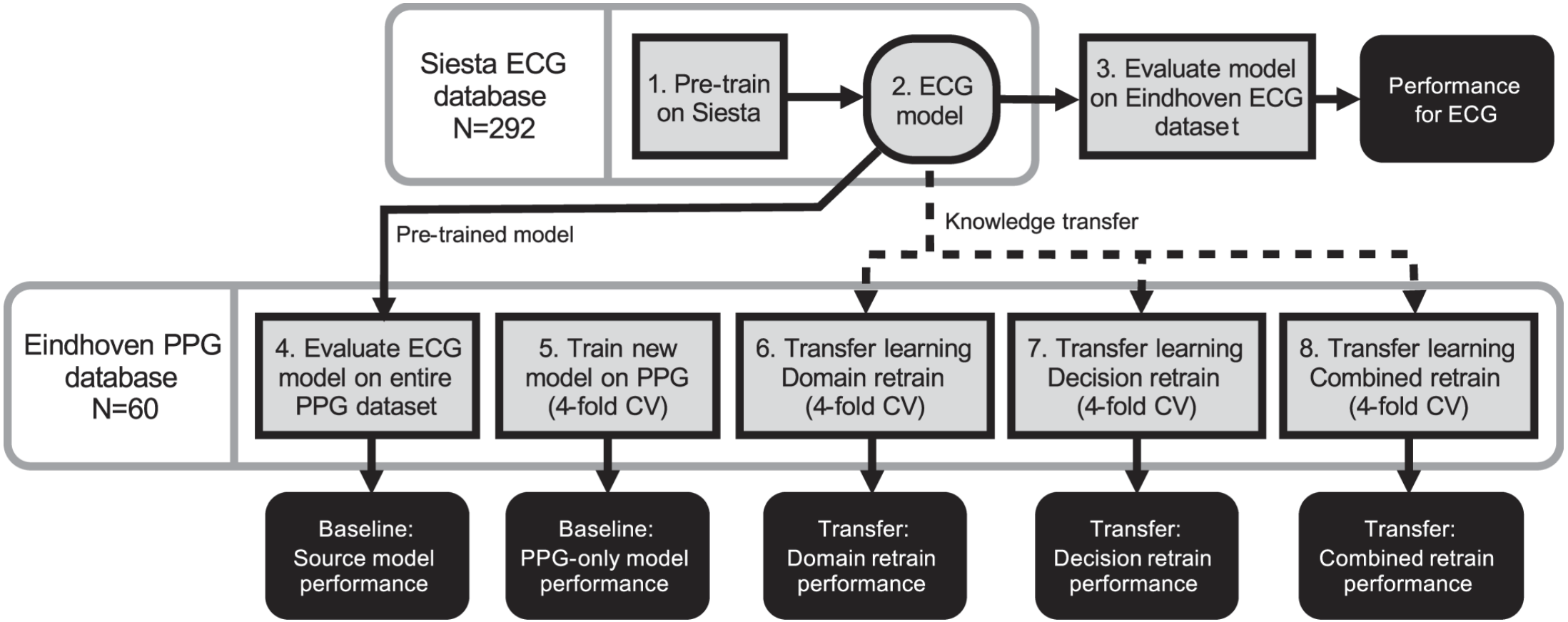
- Works that use **transfer learning**: model trained on a large database of heart rate variability (HRV) measures and then fine-tuned to a smaller database of pulse rate variability (PRV) measures derived from the IBIs detected on the PPG.
- ECG can be used to calculate HRV and HRV can be correlated with sleep stages (LSTM models seem good)
- There is a lot of ECG data: train on that!
- Then transfer to lower data regime

Deep learning approaches over PPG Sleep

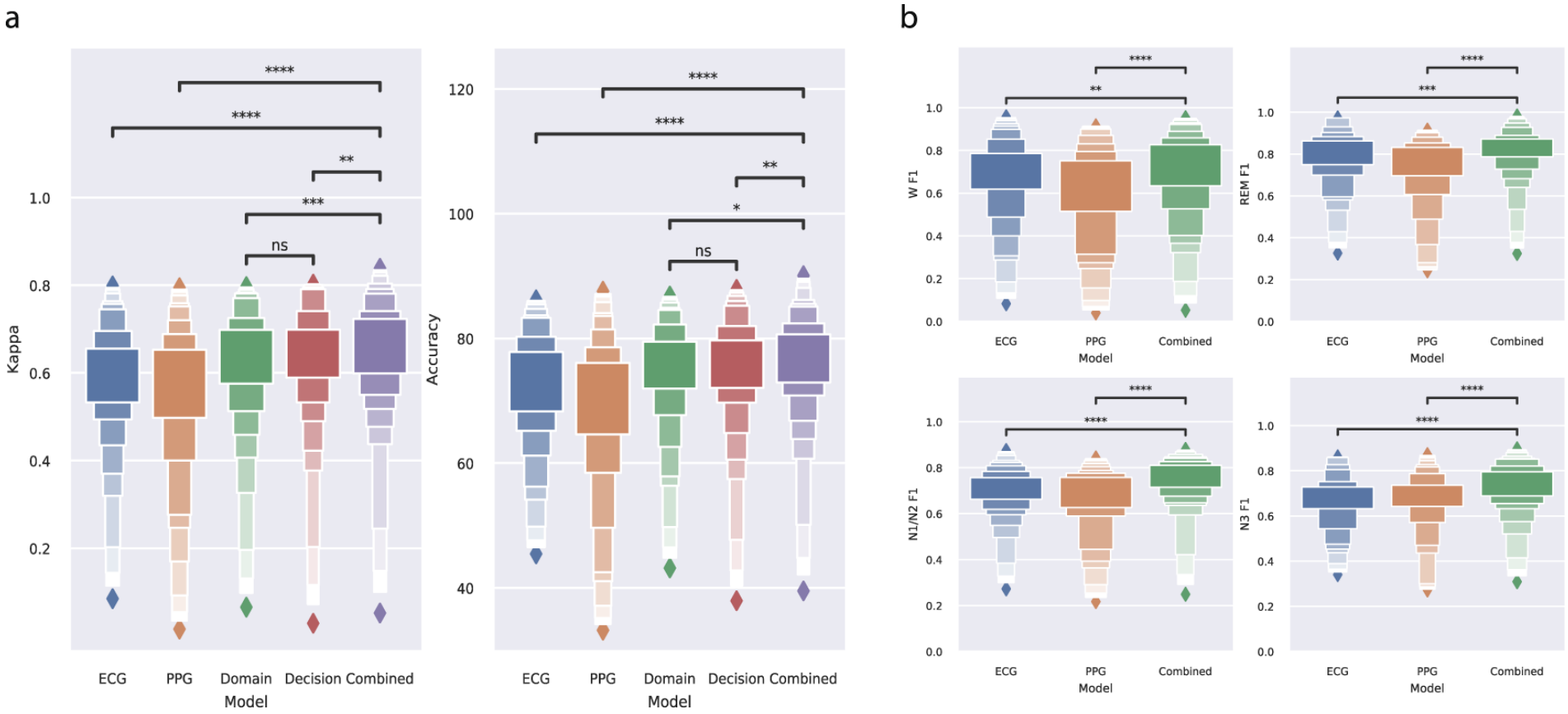


UNIVERSITY OF CAMBRIDGE

Radha, M., Fonseca, P., Moreau, A. et al. A deep transfer learning approach for wearable sleep stage classification with photoplethysmography. npj Digit. Med. 4, 135 (2021).


Siesta ECG database N=292
1. Pre-train on Siesta
2. ECG model
3. Evaluate model on Eindhoven ECG dataset
Performance for ECG
Pre-trained model
Knowledge transfer
Eindhoven PPG database N=60
4. Evaluate ECG model on entire PPG dataset
5. Train new model on PPG (4-fold CV)
6. Transfer learning Domain retrain (4-fold CV)
7. Transfer learning Decision retrain (4-fold CV)
8. Transfer learning Combined retrain (4-fold CV)
Baseline: Source model performance
Baseline: PPG-only model performance
Transfer: Domain retrain performance
Transfer: Decision retrain performance
Transfer: Combined retrain performance

Performance

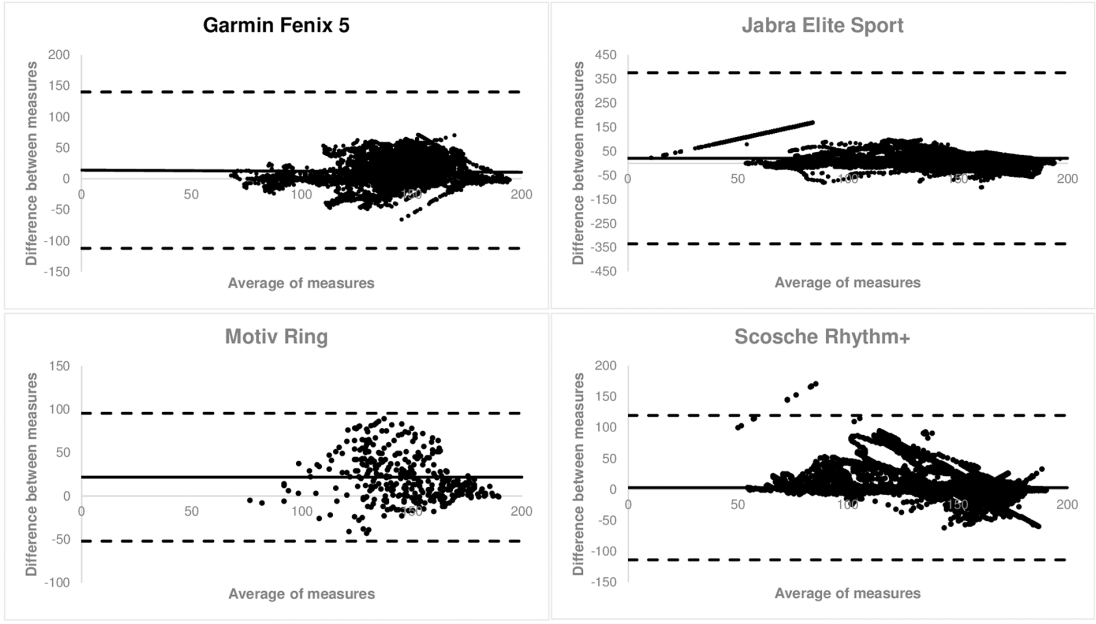


UNIVERSITY OF CAMBRIDGE

Cohen's kappa: how much is the agreement of each method with PSG?

PPG. When does it not work: Motion… Bland-Altman Plots

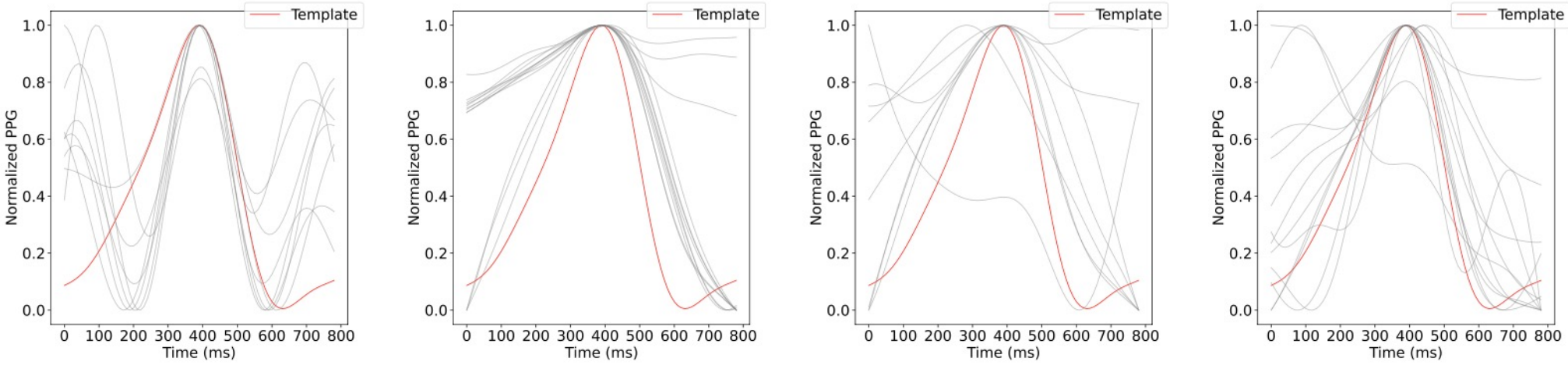
Polar 7 strap as ground truth



Navalta JW, Montes J, Bodell NG, Salatto RW, Manning JW, DeBeliso M (2020) Concurrent heart rate validity of wearable technology devices during trail running. PLoS ONE 15(8).

UNIVERSITY OF CAMBRIDGE

PPG on Earables. When does it not work: Micro motion…



(a) Shake. (b) Brow Raiser. (c) Lip Puller. (d) Mouth Stretch.

Red line: PPG in stationary case. Gray line:(same user) in various motion sessions for that movement

PPG. When it does not work: Skin colour

- Paper [1] discusses:
  - That Paper [2] found no significant difference in accuracy across skin tones but did find differences by devices in response to changes in activity.
  - Previously reported studies [3] finding wearables using green light technology had larger errors rates in tracking heart rate and energy expenditure for individuals with darker skin tones especially if exercising.
  - Racial biases and limitations of Fitzpatrick Skin Type Scale: originally used for propensity skin to burn :)
  - too few people with the darkest skin tones were included (*n* = 9 in FST Type 6) in paper [2].

References on the next slide

[1] Colvonen, P.J. Response To: Investigating sources of inaccuracy in wearable optical heart rate sensors. npj Digit. Med. 4, 38 (2021).

References from previous page

[1] Colvonen, P.J. Response To: Investigating sources of inaccuracy in wearable optical heart rate sensors. npj Digit. Med. 4, 38 (2021).

[2] Bent, B., Goldstein, B.A., Kibbe, W.A. *et al.* Investigating sources of inaccuracy in wearable optical heart rate sensors. *npj Digit. Med.* **3**, 18 (2020).

[2] Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, Wheeler MT, Ashley EA. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. J Pers Med. 2017 May 24;7(2):3.

Expert insight into current research

News & views

Racism in science

Forum: Medical devices

Skin colour affects oxygen-sensor accuracy

COVID-19 broadened the use of pulse oximeters for rapid blood-oxygen readings, but it also highlighted the fact that skin pigmentation alters measurements. Two groups of researchers analyse this issue, and its effects on people with dark skin.

Matthew D. Keller & Brandon Harrison-Smith
Pulse-oximetry errors affect patient outcomes

Since Sjoding and colleagues' report, several large retrospective studies have confirmed that darker-skinned people (those self-identifying as Black, Asian, Hispanic or a combination of these) are more likely than white people to experience occult hypoxaemia[2–5]. In one study of people with COVID-19, had equivalent arterial blood-gas values[3]. A more comprehensive analysis showed that, even when baseline health conditions are taken into account, people with occult hypoxaemia are prone to organ dysfunction and in-hospital mortality, and that Black people in this group have the worst organ dysfunction[5].

Although clinical reports of skin-colour bias in pulse oximetry were not widespread until the COVID-19 pandemic, evidence for this issue has been accumulating for decades[6,7]. A comparison reported in February found that pulse-oximeter readings from nine devices were consistently less accurate for darker-skinned people than for lighter-skinned people[8]. But the study also found that testing healthy individuals under carefully controlled laboratory conditions resulted in fewer cases of occult hypoxaemia than are measured in hospitals. In fact, none of the 491 people who were tested by the authors had readings consistent with occult hypoxaemia, whereas Sjoding and colleagues tallied 187 cases out of 3,527 measurements from a

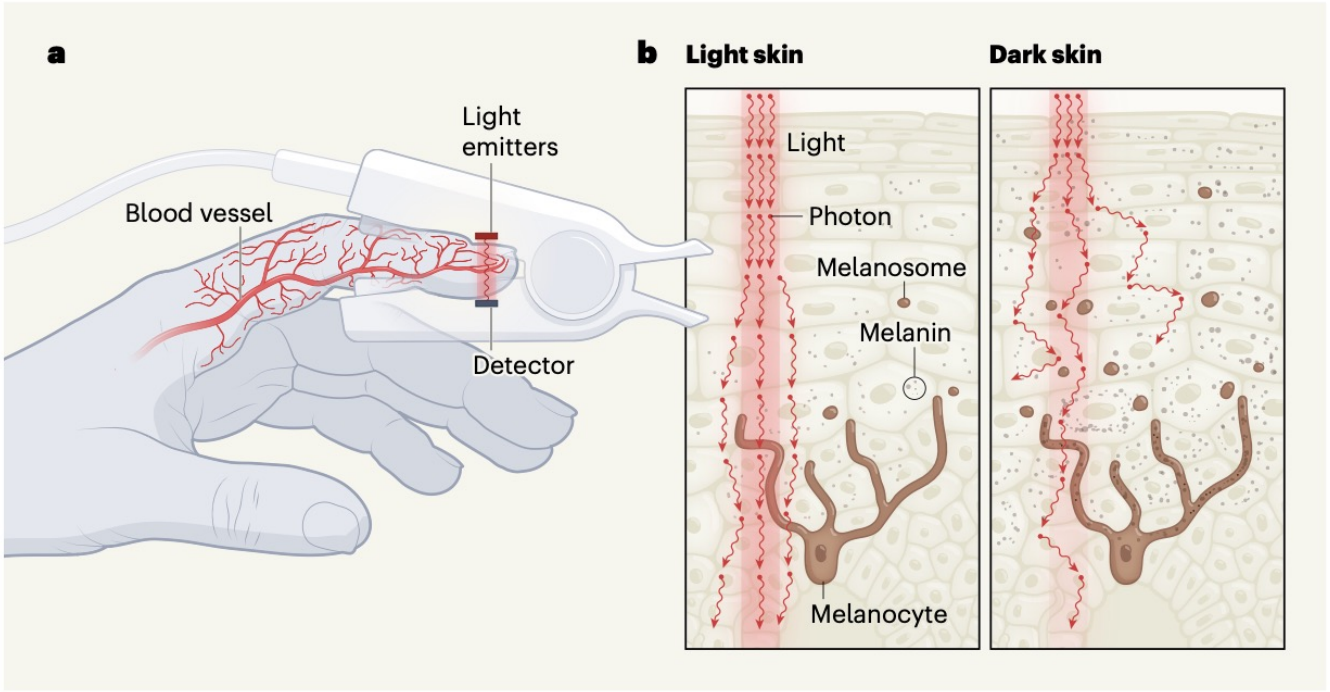


Figure 1 | Pulse-oximetry accuracy varies with skin tone. a, Devices known as pulse oximeters estimate the oxygen concentration in a person's blood by shining red and infrared light through their fingertip. Oxygenated haemoglobin absorbs infrared light more efficiently than it does red light, whereas the opposite is true for deoxygenated haemoglobin. **b**, These signals are affected by melanin, which is distributed through the skin in structures, known as melanosomes, that are produced by cells called melanocytes. Melanosomes in dark skin are both larger and more numerous than are those in light skin. Long-standing oximetry theory does not fully account for the way in which photons are scattered by the biomolecular content and structure of the tissue, and thus imprecisely corrects for the effect of pigmentation. Calibration studies compound this problem, because they typically oversample light-skinned people. This has led to overestimation of the oxygen concentration in some Black individuals' blood, and therefore to missed diagnoses of dangerously low oxygen levels.

Nature **610**, 449-451 (2022)

Questions