https://doi.org/10.5664/jcsm.8762

JCSM | Journal of Clinical Sleep Medicine

COMMENTARY

Autopilot and algorithms: accidents, errors, and the current need for human oversight

Commentary on Ioachimescu OC, Allam JS, Samarghandi A, et al. Performance of peripheral arterial tonometry–based testing for the diagnosis of obstructive sleep apnea in a large sleep clinic cohort. *J Clin Sleep Med*. 2020;16(10):1663–1674. doi:10.5664/jcsm.8620

Douglas Kirsch, MD, FAAN, FAASM

Sleep Medicine, Atrium Health, Charlotte, North Carolina; University of North Carolina School of Medicine, Chapel Hill, North Carolina

Tesla, arguably one of the most technologically advanced automobile manufacturers, offers Autopilot. This neural-network driving assistance program theoretically enhances driver safety. But how far would you trust it? Even Tesla cautions that: "Autopilot…[is] intended for use with a fully attentive driver, who has their hands on the wheel and is prepared to take over at any moment"[1] Accidents have been reported with Autopilot engaged, particularly with reportedly distracted drivers.[2] Driving is a complex task where assistive technology can reduce errors; however, drivers are still required to be focused and studying their surroundings to minimize their risk.

Unique in its nonairflow-based technology and proprietary algorithm to diagnose obstructive sleep apnea (OSA), the WatchPAT home sleep apnea testing (HSAT) technology has been heavily studied since its introduction to the marketplace. The device output is calculated primarily via algorithmic scoring with the option of rescoring by the interpreting clinician. The strength of the large volume of research data demonstrating diagnostic accuracy of the device and algorithm ultimately led to the PAT signal being included in *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications* as an approved option for HSAT assessment of OSA.[3]

Many of the research studies conducted with the WatchPAT technology have relied solely on the device algorithm in the absence of clinician rescoring for diagnosis of OSA; however, recent data suggest that individual patient differences may have an effect on the accuracy of WatchPAT device output. For example, Kinoshita and colleagues published in 2018 that differences in arterial stiffness might impact precision of the WatchPAT device outcomes.[4] In this issue of the *Journal of Clinical Sleep Medicine*, Ioachimescu and colleagues performed concurrent laboratory-based polysomnography and WatchPAT-200 studies comprising 500 veterans, demonstrating that differences in severity levels between the two studies occurred in a sizeable number of patients.[5] The authors noted that 5% of patients with a WatchPAT-based diagnosis of moderate or severe OSA who did not have OSA on laboratory-based polysomnography and the 20% of patients with WatchPAT-diagnosed mild OSA who had moderate or severe OSA on laboratory-based polysomnography. It is of concern that these erroneous diagnoses can lead to poor patient outcomes, either related to potential mistreatment or nontreatment.

Several practical considerations can be derived from the above-mentioned article, some of which are relevant to any HSAT device and others that are specific to the WatchPAT 200 device:

1. As stated in the American Academy of Sleep Medicine (AASM) diagnostic testing for OSA clinical practice guideline, "…if a single home sleep apnea test is negative, inconclusive, or technically inadequate, polysomnography be performed for the diagnosis of OSA."[6] Performing an in-laboratory test in a patient with a high pretest probability for OSA who has a normal HSAT result may be challenging for patients with higher insurance deductibles or those who wish to avoid in-laboratory polysomnography for other reasons. As always, patient history is the clearest guide to appropriate clinical decision making.
2. Clinical treatment decisions for patients with HSAT-diagnosed OSA should be made with the understanding that there will be some level of uncertainty about severity when compared with in-laboratory polysomnography. Whereas this level of uncertainty is manageable given current treatment options, the evolution of clinical practice to precision-based treatment approaches for OSA will require increasing levels of accuracy.
3. Although not specifically called out in Ioachimescu et al, artifact in WatchPAT studies (or any HSAT study, for that matter) can distinctly impact the study findings; ensuring exclusion of artifact and/or repeat of the HSAT study if artifact levels are high enough is a necessary step. Epoch-by-epoch study review remains essential and should guide the decision about the study quality.[7] This author has seen an artifact-filled WatchPAT study demonstrate an apnea-hypopnea index of 60+ events/h when a (nonconcurrent) laboratory-based polysomnogram on the same patient a few days later demonstrated an AHI of only 6 events/h.

4. The WatchPAT analytic technology allows for respiratory event scoring with either a 3% or a 4% oxygen desaturation criteria. In Ioachimescu et al, the WatchPAT 200 using the 4% criteria tended to have a higher specificity and improved the negative predictive value versus the 3% criteria (compared with polysomnography, which used the 3% desaturation and arousal scoring for hypopneas). This finding suggests that the 4% criteria may be a more accurate default setting for typical clinical use.

Technology moves fast in our current times; devices and their software evolve quickly while scientific articles that study them often lag. The HSAT device evaluated in this study is the WatchPAT 200 device; newer models, such as the WatchPAT 300, may have differing accuracy and options. It is a challenge for clinical practice that by the time a device-related article is published, the next generation device may already be in use, making up-to-date comparisons nearly impossible.

Proprietary algorithms and artificial intelligence will increasingly be a larger element in sleep medicine practice. We have already begun to see them in sleep laboratory autoscoring, patient OSA treatments (autotitrating continuous positive airway pressure, for example), and consumer-wearable sleep devices. The AASM has recently released a position statement on artificial intelligence, outlining some of the ways that that artificial intelligence may change (and hopefully improve) our clinical practices in the future.[8] Yet, Tesla's Autopilot warning and Ioachimescu et al's results remind us that for now, keeping your eyes on the road (or on the sleep study) is still a necessity for optimal outcomes.

CITATION

Kirsch D. Autopilot and algorithms: accidents, errors, and the current need for human oversight. *J Clin Sleep Med*. 2020; 16(10):1651–1652.

REFERENCES

1. Autopilot and Full Self-Driving Capability. Tesla website. https://www.tesla.com/support/autopilot. Accessed August 12, 2020.
2. Chokshi N. Tesla Autopilot System Found Probably at Fault in 2018 Crash. The New York Times website. https://www.nytimes.com/2020/02/25/business/tesla-autopilot-ntsb.html. Published February 25, 2020. Accessed August 12, 2020.
3. Berry RB, Quan SF, Abreu AR, et al; for the American Academy of Sleep Medicine. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Version 2.6. Darien, IL: American Academy of Sleep Medicine; 2020.
4. Kinoshita T, Yahaba M, Terada J, et al. Impact of arterial stiffness on WatchPAT variables in patients with obstructive sleep apnea. *J Clin Sleep Med.* 2018;14(3): 319–325.
5. Ioachimescu OC, Allam JS, Samarghandi A, et al. Performance of peripheral arterial tonometry based testing for the diagnosis of obstructive sleep apnea in a large sleep clinic cohort. *J Clin Sleep Med.* 2020;16(10):1663–1674.
6. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, Harrod CG. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of Sleep Medicine clinical practice guideline. *J Clin Sleep Med.* 2017;13(3):479–504.
7. AASM Facility Standards for Accreditation. American Academy of Sleep Medicine website. http://aasm.org/wp-content/uploads/2019/05/AASM-Facility-Standards-for-Accreditation-8.2020.pdf. Updated August 2020. Accessed August 12, 2020.
8. Goldstein CA, Berry RB, Kent DT, et al. Artificial intelligence in sleep medicine: an American Academy of Sleep Medicine position statement. *J Clin Sleep Med*. 2020; 16(4):605–607.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication August 18, 2020
Submitted in final revised form August 18, 2020
Accepted for publication August 18, 2020
Address correspondence to: Douglas Kirsch, MD, FAAN, FAASM, Atrium Health Sleep Medicine, 1601 Abbey Place, Bldg 2, Suite 200, Charlotte, NC 28209; Tel: (704) 512-2980; Fax: (704) 512-2970; Email: Douglas.Kirsch@atriumhealth.org

DISCLOSURE STATEMENT

All authors have seen and approved the manuscript. Work for this study was performed at Atrium Health, Charlotte, North Carolina. This was not an industry-supported study. The author reports no financial conflicts of interest.