Open Journal of Preventive Medicine, 2026, 16(6), 107-124
https://www.scirp.org/journal/ojpm
ISSN Online: 2162-2485
ISSN Print: 2162-2477

Role of Machine Learning Algorithms and Artificial Intelligence in the Evaluation of Insomnia in Adolescents

Arnav Sehgal[1], Sonya Jayakar[1], Narayan P. Verma[2*]

[1]BG Tricounty Neurology and Sleep Clinic, Warren, USA
[2]OUWB School of Medicine, Warren, USA
Email: *narayangod@aol.com

How to cite this paper: Sehgal, A., Jayakar, S. and Verma, N.P. (2026) Role of Machine Learning Algorithms and Artificial Intelligence in the Evaluation of Insomnia in Adolescents. *Open Journal of Preventive Medicine*, **16**, 107-124.
https://doi.org/10.4236/ojpm.2026.166007

Received: April 13, 2026
Accepted: June 26, 2026
Published: June 29, 2026

Abstract

Insomnia (the persistent inability to fall or stay asleep) is one of the most commonly diagnosed sleep disorders in teenagers today. However, the troubling fact is that the way doctors currently diagnose it is almost entirely based on what patients say about their own sleep. They fill out questionnaires and keep sleep diaries. Research shows that these self-reports are frequently inaccurate, and in some cases, patients who are sleeping normally are being diagnosed with insomnia (sometimes even receiving powerful prescription sleep medications as a result). This paper argues that consumer wearable devices, the smartwatches and fitness trackers millions of teenagers already own, combined with artificial intelligence (AI), could fundamentally change this. By measuring heart rate variability (the small changes between heartbeats), movement, and other physiological signals, these devices can track sleep in ways that are objective, continuous, and surprisingly accurate. This review looks at how these technologies work, how accurate they are, what AI tools are being used to analyze their data, and why using objective, wearable-based measurement as an adjunct to clinical assessment could be one of the most important upgrades in sleep medicine in decades.

Keywords

Adolescent Insomnia, Wearable Devices, Artificial Intelligence, Heart Rate Variability (HRV), Objective Sleep Assessment, Diagnostic Accuracy

1. Introduction

Teenagers need between eight and ten hours of sleep per night [1]. Yet, study after study shows they are among the most sleep-deprived individuals in modern society. About 77% of U.S. teenagers do not get adequate sleep for proper brain development [2]. A 2015 CDC report found that, nationwide, the prevalence of short sleep duration among high school students was 72.7% [1]. More recent data from the CDC's 2023 Youth Risk Behavior Survey show the problem is continuing, with only 23% of students reported achieving the recommended eight hours of sleep on an average school night [3]. A 2026 JAMA analysis of long-term CDC data found that the proportion of teenagers getting the recommended eight hours dropped from over 30% in 2007 to fewer than 25% by 2023, while the number sleeping fewer than five hours increased substantially across nearly every demographic [4].

The consequences are serious: poor academic performance, worsened mental health, increased anxiety and depression, and weakened immune function [2] [4].

It is important to distinguish between insufficient sleep (not getting enough hours) and insomnia disorder, which is defined by persistent difficulty initiating or maintaining sleep despite adequate opportunity, accompanied by daytime impairment. The statistics above describe insufficient sleep duration broadly. The following figures describe insomnia disorder specifically.

Insomnia is one of the most common sleep disorders contributing to this crisis. Studies estimate that between 4% and 39% of adolescents experience insomnia symptoms depending on how insomnia is defined and measured, with a commonly cited prevalence of 18.5% using DSM-5 criteria in older adolescents aged 16 - 18. Importantly, insomnia tends to be chronic once established, with 88% of adolescents with a history of insomnia reporting current insomnia [5].

As smartphones and social media have become central to teenage life, the problem has grown far worse. Light-emitting screens, particularly those emitting blue-wavelength light, suppress melatonin (the hormone the brain releases to promote sleep) and disrupt the circadian rhythms. Adolescents in early to mid-puberty appear particularly sensitive to the melatonin-suppressing effects of evening light compared to adults [6]. Social media keeps the brain stimulated and alert when it should be winding down [7], and the fear of missing out (FOMO) drives teenagers to check their phones throughout the night, disrupting sleep repeatedly [8].

However, there is a greater problem that is rarely discussed. That is, even when teenagers seek help for insomnia, the tools physicians use to diagnose it are fundamentally flawed. The diagnosis relies almost entirely on self-reported information, that is, what the patient says about their own sleep. And as this paper will show, people are surprisingly poor at knowing how much they actually sleep.

2. What Is Insomnia and How Is It Currently Diagnosed?

2.1. Defining Insomnia

Insomnia is defined as persistent difficulty falling asleep, staying asleep, or waking up too early, occurring at least three nights per week for at least three months, and causing problems during the day (such as fatigue, trouble concentrating, mood changes, or poor performance at school or at work) [5]. This is called chronic insomnia disorder.

There is also a milder form called subclinical insomnia, meaning the person has sleep problems but does not quite meet the full diagnostic criteria. Clinicians use the Insomnia Severity Index (ISI), a validated questionnaire, to measure how severe someone's sleep problems are. Scores of 0 - 7 indicate no clinically significant insomnia, scores of 8 - 14 indicate subclinical insomnia, and scores of 15 or higher indicate moderate to severe clinical insomnia [5].

2.2. The Gold Standard: Polysomnography

The most accurate way to measure sleep right now is called polysomnography (PSG), often referred to as a sleep study. During a PSG, a patient spends the night in a specialized sleep lab wearing sensors that measure brain waves (electroencephalography, or EEG), eye movements, muscle activity, heart rate, breathing, and blood oxygen levels. PSG is the gold standard because it provides a complete, objective picture of what is happening during sleep. However, it is expensive, inconvenient, requires an overnight stay in an unfamiliar environment, and captures only one or a small number of nights. This makes it impractical for routine monitoring or long-term tracking, especially in teenagers. Portable or home-based polysomnography may also be inaccurate.

2.3. How Insomnia Is Actually Diagnosed in the Real World

Because PSG is expensive, inconvenient, and impractical for routine use, insomnia in clinical practice is almost always diagnosed based on self-reported information. Patients fill out questionnaires such as the ISI or the Pittsburgh Sleep Quality Index (PSQI). They may also be asked to keep a sleep diary for two weeks. Doctors then review this information in a clinical interview [9]. This approach has an obvious problem. It relies entirely on the patient's memory, perception, and honesty. As the next section will show, each of these can fail in ways that lead to serious diagnostic errors.

3. The Problem with Self-Reported Sleep Data

3.1. People Are Bad at Knowing How Much They Sleep

Multiple large studies have shown that self-reported sleep duration and objective sleep duration, measured by actigraphy (a wrist-worn sensor that tracks movement), frequently disagree, sometimes dramatically.

One large study comparing self-reported and actigraphy-based sleep in community-dwelling older adults found an intraclass correlation of only 0.18, indicating very low agreement between what people said they slept and what devices actually recorded. People who reported trouble falling asleep or waking too early underreported their total sleep time by an average of 2.2 to 2.3 hours compared to what the actigraphy recorded. In other words, people who believed they were barely sleeping were actually sleeping significantly more than they had thought [10].

A separate study involving 10,268 participants found that self-reported sleep duration was systematically lower than objectively estimated sleep by an average of 42 minutes, with agreement limits spanning more than two hours in either direction. These errors were not random. They were influenced by factors such as social jetlag (a misalignment between the body's natural sleep schedule and social demands like school start times), insomnia symptoms, and even the specific wording of the questions being asked [11].

Research has shown that the way sleep questions are phrased matters enormously. Asking someone when they "get to bed" versus when they "fall asleep" produces two very different answers because people who read or use phones in bed before sleeping systematically overestimate their sleep time. Week-to-week and night-to-night variability in sleep also cannot be captured by a one-time survey. It requires sleep diaries or actigraphy for accurate measurement [12].

A study specifically in young people with mental health problems found that those reporting insomnia significantly underestimated their objective sleep duration by an average of 31 minutes, while those without insomnia showed little bias. This suggests that the perception of poor sleep, not just poor sleep itself, creates measurement error [13].

3.2. Insomnia Identity: When People Who Sleep Well Believe They Have Insomnia

One of the most interesting discoveries in insomnia research is the existence of what researchers call "insomnia identity" (a condition in which people believe they have chronic insomnia despite objective evidence that their sleep is within normal limits) [14].

Research presented at the annual meeting of the American Academy of Sleep Medicine found that about 30% of people who complained of poor sleep quality were actually sleeping well, as documented by both self-reported daily sleep diaries and objective measurements. These individuals reported all the classic symptoms of insomnia (feeling unrested, tired during the day, anxious about sleep), but their actual sleep was perfectly fine. Clinicians, the researchers noted, often make a diagnosis of insomnia based on a patient's self-report at a single clinical visit. As a result, some patients receive prescriptions for medications to correct sleep problems they do not actually have [14].

Cultural factors appear to play a role also. African American patients who slept poorly complained about their sleep less than Caucasian patients, and white patients were more likely to meet criteria for insomnia identity. This suggests that the perception and reporting of insomnia are influenced by cultural norms. This means that even the same physiological state of poor sleep may result in very different diagnostic outcomes depending on who the patient is. This is a form of diagnostic bias baked into the current system [14].

3.3. Paradoxical Insomnia: The Brain That Thinks It Is Awake

Possibly the strangest and most important discovery in insomnia research is what clinicians call paradoxical insomnia (also known as sleep state misperception). This is a condition where a person feels they have barely slept at all, despite objective sleep studies showing normal or near-normal sleep architecture (the structure and pattern of sleep stages) [15].

A patient with paradoxical insomnia might report sleeping only one or two hours per night. They describe a constant awareness of their surroundings while lying in bed, active thinking throughout the night, and hearing every sound in the house. Yet a PSG study shows they slept six or seven hours with relatively normal sleep stages [15]. This phenomenon was first observed in a French case report in 1959, and it was found that patients' subjective reports of total sleep time differed substantially from their objectively measured sleep duration, with the researchers commenting that pre-sleep cognitive arousal and worry about sleep may create a false impression of wakefulness during sleep [16].

Recent research using 256-electrode brain monitoring systems (a setup with 256 sensors applied to the scalp to measure brain activity, far more sensitive than the 6 to 20 electrodes used in standard sleep clinics) found that people with subjective insomnia show bursts of fast, wake-like brain wave activity, particularly during REM sleep (Rapid Eye Movement sleep, the stage when most dreaming occurs). When woken from these periods, patients reported having been awake and thinking about everyday tasks rather than dreaming [17]. More importantly, younger age has been shown to be an independent risk factor for overestimating how long it takes to fall asleep and underestimating total sleep time. This makes paradoxical insomnia more relevant for teenagers [15].

Research has also found that watching a digital clock while trying to fall asleep worsens sleep state misperception. People who monitor a clock overestimate how long it takes them to fall asleep compared to those watching a display showing random digits, because tracking time increases the amount of information the brain processes per unit of time, making the wait feel longer [18].

3.4. The Drug Problem: When a Bad Diagnosis Leads to Dangerous Treatment

The consequences of misdiagnosed insomnia are not just academic. Insomnia is a major driver of prescription sleep medication use. Research using nationally representative samples has found that insomnia symptoms prospectively predict the development of prescription opioid and benzodiazepine misuse, and vice versa. This creates a two-way vicious cycle. Prescription opioids contribute to the loss of about 15,000 lives a year in the U.S., and benzodiazepine misuse contributes to the loss of over 10,000 lives annually [19].

The clinical challenge is exacerbated by the difficulty of distinguishing between "therapy-seeking" (a patient who is actually suffering from insomnia) and "drug-seeking" (a patient exaggerating or fabricating symptoms to obtain prescription medications). Clinical researchers explicitly acknowledge in the literature that drug-seeking and therapy-seeking behavior "can become closely intertwined," making them difficult to differentiate [20]. A doctor relying solely on a patient's self-report has no objective tool to identify this distinction. People who stop taking benzodiazepines often experience "rebound insomnia," where sleep problems become significantly worse than before they started the medication, making it even harder to stop taking the drug. This becomes a cycle that, in many cases, begins with a potentially avoidable prescription. Sleeping medications such as Zolpidem (Ambien) and Eszopiclone (Lunesta) carry significant potential for physical dependence and misuse [20].

No existing published review has explicitly argued that consumer wearable data, combined with AI, should be integrated into the clinical diagnosis process for insomnia specifically as a tool to address the diagnostic failures created by insomnia identity, paradoxical insomnia, and drug-seeking behavior.

An objective measure like a wearable device recording physiological sleep data over weeks would provide the verification tool that current clinical practice lacks. A patient claiming severe insomnia whose wearable data shows normal sleep patterns would prompt a very different clinical response than one whose data confirms actual sleep disruptions. The potential to objectively verify insomnia claims before prescribing controlled substances could meaningfully reduce both misdiagnosis and medication misuse.

4. How Consumer Wearables Track Sleep

4.1. What Wearables Measure

Consumer wearable devices (including smartwatches and fitness trackers from companies like Apple, Fitbit, Garmin, Oura, and WHOOP) use a set of sensors to estimate sleep.

Accelerometers measure physical movement. When you are asleep, you move far less than when you are awake. By detecting movement patterns, a wearable can estimate when you fall asleep, how often you shift during the night, and when you wake up. The clinical version of this technique is called actigraphy, and it has been used in sleep research for decades [21].

Photoplethysmography (PPG) is the technology that measures heart rate. PPG sensors shine a small green LED light through the skin on your wrist. Blood absorbs this wavelength of light differently depending on how much blood is present at any given moment, which changes with each heartbeat. By measuring these tiny changes in light absorption, the sensor counts heartbeats and detects the intervals between them [22].

Heart Rate Variability (HRV) is calculated from PPG data and represents the tiny variations in time between consecutive heartbeats. These variations are driven by the autonomic nervous system (the system that controls automatic or involuntary body functions like breathing and heart rate). Higher HRV usually represents a relaxed state, while lower HRV indicates a stressed or activated state. During sleep, especially deep sleep, HRV typically increases. In people with insomnia, HRV during sleep is often reduced, reflecting the underlying physiological hyperarousal (a state of being too physiologically alert) that characterizes the disorder [23].

Skin temperature naturally drops slightly when you fall asleep, and some wearables, particularly the Oura Ring, incorporate temperature sensors as an additional signal for sleep staging [24].

4.2. How Accurate Are Consumer Wearables?

The accuracy of consumer wearable devices has been studied extensively in recent years by comparing their outputs to PSG, the gold standard mentioned in topic 2.2.

A 2024 study compared six consumer wearables (including the Fitbit Charge 5, Fitbit Sense, Apple Watch Series 8, Garmin Vivosmart 4, WHOOP 4.0, and Withings Scanwatch) against PSG in 62 adults across multiple overnight laboratory sessions. The study found that all devices detected more than 90% of actual sleep epochs (meaning periods of time when the person was truly asleep). This shows high sensitivity for detecting sleep. However, specificity (the ability to correctly identify wakefulness) was much lower, ranging from only 29% to 52%. In practical terms, the devices are good at knowing when you are sleeping, but they frequently misclassify quiet wakefulness as sleep. This is especially relevant for insomnia patients, who often lie still in bed while awake, potentially causing the devices to overestimate how much they are sleeping [25].

A 2024 Brigham and Women's Hospital study comparing the Oura Ring Gen3, Fitbit Sense 2, and Apple Watch Series 8 found that all three devices achieved sensitivity above 95% for detecting sleep versus wakefulness. The Oura Ring, which uses body temperature alongside HRV and PPG, showed the lowest proportional bias for sleep staging [26]. A comprehensive multicenter study evaluating 11 different consumer sleep trackers (including smartwatches, mattress pads, and phone apps) found that the Oura Ring 3 and Fitbit Sense 2 showed the best improvement in sleep stage detection compared to earlier device generations [24].

For HRV measurement specifically (one of the most important signals for insomnia detection), accuracy varies considerably by device. A validation study comparing five wearables against electrocardiogram (ECG) (a test that records the heart's electrical activity) across 536 nights found that the Oura Gen 4 achieved near-perfect HRV accuracy (concordance correlation of 0.99, meaning almost identical to the gold standard). At the same time, Garmin and Polar showed substantially lower accuracy, and WHOOP showed moderate accuracy [22].

An important limitation: Most wearable validation studies have been conducted on healthy adults in laboratory settings for single nights. Studies specifically validating consumer wearables in insomnia patients (who lie still while awake, creating the exact pattern of misclassification described above) are still limited. The researchers themselves note that "insomniacs often lie still in bed while trying to sleep, even though they are actually awake" and that "these periods of wakefulness can be misinterpreted as sleep" [24]. This is a meaningful gap in the current literature, and it suggests that accuracy estimates for general users may overstate performance in clinical insomnia populations.

5. How Artificial Intelligence Uses Wearable Data to Detect Insomnia

5.1. From Raw Data to Prediction

A smartwatch collecting data overnight generates thousands of data points (like a heart rate reading every few seconds, movement data continuously, HRV calculated across the night). This is far too much information for any human to interpret directly in an accurate and efficient manner. AI and machine learning (systems that learn patterns from large amounts of data) are what make this data clinically useful [23].

The approach works like this: researchers collect wearable data from a large group of people, some of whom have confirmed insomnia and some of whom do not. They then train an AI model to recognize the patterns in the wearable data that distinguish insomnia from healthy sleep. Once trained, the model can analyze new data from a new person and predict whether they have insomnia (without requiring a PSG) [21] (**Figure 1**).

Figure 1. Wearable AI collects data from different sources, and AI algorithms interpret and provide feedback.

5.2. Machine Learning Algorithms Used for Insomnia Detection

Random Forest is an algorithm that creates many decision trees (each tree asks a series of yes/no questions about the data). Each tree produces a prediction, and the algorithm takes the majority vote [27] (**Figure 2**). Random forests are popular in sleep research because they handle noisy, real-world wearable data well. Studies have found that random forest models achieve accuracy rates between 90% and 95% for detecting insomnia from physiological data [28] (**Table 1**).

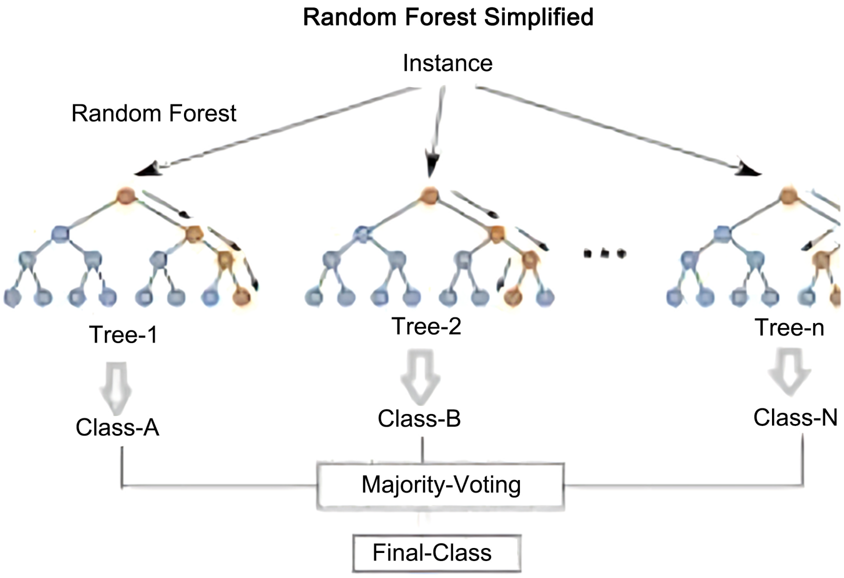

Figure 2. Visualization showing data from many trees or models synthesizing data from the same starting data point, together providing different conclusions ("class"), which result in one majority conclusion that is more robust and accurate than individual ones.

Table 1. Advantages and disadvantages of random forest classifier in machine learning.

Advantages	Disadvantages
Can handle both numerical and categorical data without preprocessing.	Requires more computational resources and memory.
Reduces overfitting through ensemble learning.	Training takes longer than simple models.
Works well even with missing data.	Less interpretable than single decision trees
Provides feature importance rankings automatically.	Trained models can take up
Enables parallel processing for faster computation.	Cannot predict beyond the range of training data
Does not require data scaling or normalization.	Not suitable for real-time predictions due to slow inference times
Handles imbalanced datasets effectively.	Requires careful parameter tuning (e.g., number of trees, depth) for optimal performance
Resistant to noisy data and can handle outliers effectively	Not the best choice for simple linear problems
Generally achieves high prediction accuracy.	May miss some complex feature interactions

Support Vector Machines (SVM) work by using a series of calculations to find the boundary that best separates two groups of data. SVMs function like a construction crew trying to build the widest possible road (margin) to separate two different cities (data classes). The algorithm identifies the buildings (data points) on the edges of the road, called support vectors, which determine its position. For datasets that are not easily separable by a straight line, the "kernel trick" allows the SVM to map the data into a more manageable, higher-dimensional space [29] (**Figure 3**). Applied to insomnia classification using heart rate and wearable data, SVM has been shown to achieve accuracy of around 89% to 91% in published studies that compared insomnia patients to healthy controls [28]. The main disadvantages of SVMs are that they are not suitable for large datasets, are sensitive to noise and outliers, and require careful parameter tuning [29].

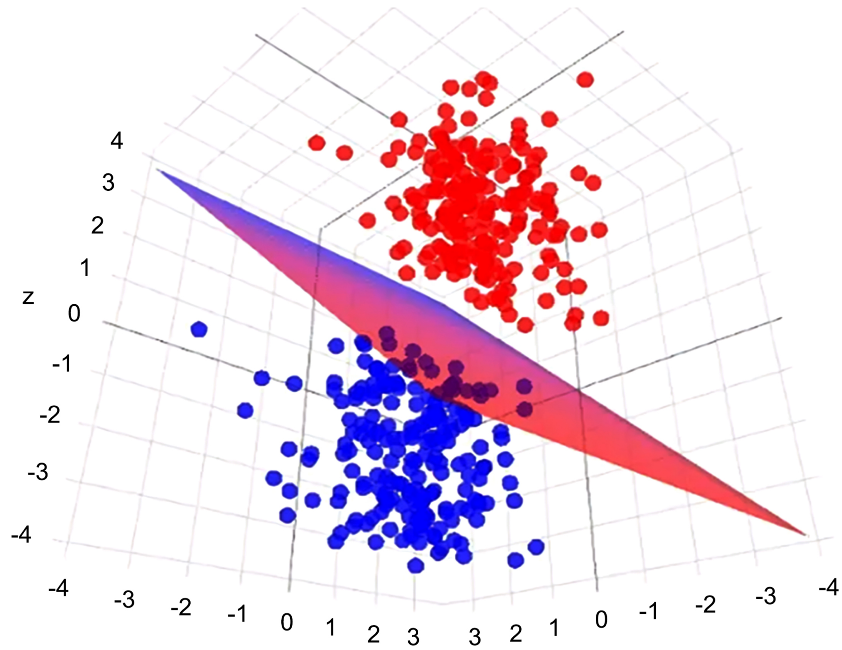

Figure 3. Scatter plot model showing how support vector machines (SVM) operate.

Extreme Gradient Boosting (XGBoost) builds a sequence of decision trees where each new tree specifically focuses on correcting the errors made by the previous ones. XGBoost has become one of the most popular algorithms in sleep research. It acts as a fast and efficient "booster", combining many weak decision trees into a strong ensemble. It is particularly known for handling large amounts of data quickly and automatically improving accuracy while reducing errors in both classification and regression projects [30]. In the context of sleep quality prediction (estimating whether an upcoming night's sleep will be good or poor), a 2024 study combining wearable sensor data (physical activity, light exposure, and HRV from a Galaxy Watch) found that XGBoost achieved an area under the curve (AUC) of 0.80 and 85% accuracy in predicting sleep quality, and performance improved further when HRV data was added to the model [31] (**Table 2**).

Table 2. Differences between traditional boosting and extreme gradient boosting.

Feature	Gradient Boosting (GBM)	Extreme Gradient Boosting (XGBoost)
Approximation	First-order gradient	Second-order Taylor expansion (Gradient + Hessian)
Regularization	Typically none	Regularized to prevent overfitting
Speed	Sequential only	Parallelized processing
Missing Values	Needs pre-processing	Built-in handling

Deep Learning (specifically Long Short-Term Memory networks (LSTMs)) is a type of artificial neural network designed to recognize patterns in sequential data over time [32]. It is particularly powerful for analyzing data that change over time, which is exactly what wearable sleep data look like. An LSTM can learn how HRV evolves across a night, recognizing patterns in the temporal sequence that simpler models miss. In the context of sleep quality prediction in adolescents, a study tracking 92 adolescents for a week using wrist actigraphy found that LSTM networks outperformed traditional machine learning methods for predicting sleep efficiency (whether sleep was good or poor) from daytime activity data [21].

Separately, for sleep stage classification (determining which sleep stage a person is in at any given moment), a review of machine learning algorithms using HRV for sleep stage classification concluded that HRV-based AI models achieve "high accuracy, sensitivity, and specificity" and represent "an active area of research with broad application potential" [33] (**Figure 4**).

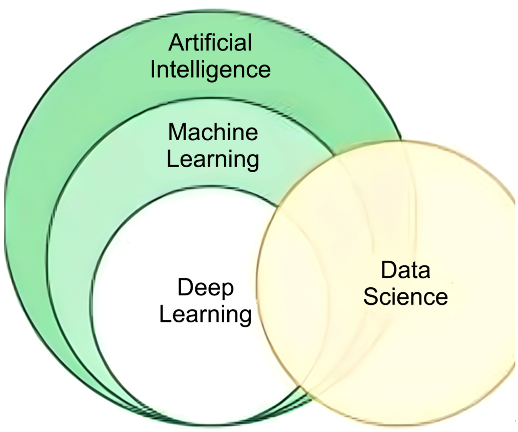

Figure 4. Diagram showing how artificial intelligence (AI), machine learning (ML), deep learning (DL), and data science overlap. AI is the broadest field; ML is a subset of AI, and DL is a subset of ML. Data science, a separate field, overlaps with the former.

5.3. HRV as the Key Indicator

Of all the signals measured by consumer wearables, HRV has emerged as the most promising for detecting and predicting insomnia. The reasoning is grounded in established physiology. Insomnia is a disorder of hyperarousal, meaning the nervous system is too activated at night. HRV directly measures the balance between activation (sympathetic nervous system) and relaxation (parasympathetic nervous system), making it a natural proxy for this physiological state [23].

A 2024 study using Samsung Galaxy Watch data and multiple machine learning algorithms (including ARIMA, Random Forest, XGBoost, and deep learning) found that daily HRV metrics collected during wakefulness were better predictors of the following night's sleep fragmentation than the ISI questionnaire. This finding shows that continuous physiological data from a smartwatch worn during the day predicted that night's sleep quality better than the standard clinical questionnaire used to diagnose insomnia [23].

Another study tracking 49 healthy individuals for four weeks with a Samsung Galaxy Watch found that daytime HRV patterns were linked to sleep quality on following nights. This suggests that the body's autonomic state during waking hours carries forward into sleep [32]. Deep learning models using HRV data streams of just five minutes achieved classification accuracies of up to 83% for predicting mental health measures such as stress and anxiety. These factors are closely related to insomnia [34].

6. Why This Could Revolutionize Insomnia Diagnosis

6.1. Connecting the Dots: New Framework for Objective Diagnosis

The existing literature reviewed above establishes several individual findings.

- ***Self-reported sleep is systematically inaccurate and biased* [11] [12] [27].**
- ***About* 30% *of people who present with insomnia say that they are sleeping normally* [14].**
- ***Paradoxical insomnia* (*when the brain misperceives sleep as wakefulness*) *is more common among younger populations* [15].**
- ***Consumer wearables can measure HRV with clinically meaningful accuracy, particularly the Oura Ring* [22].**
- ***AI models can predict sleep quality from wearable data with* 85% - 95% *accuracy in research settings* [28] [31].**
- ***HRV measured during wakefulness predicts that night's sleep quality better than standard questionnaires* [23].**

An important boundary must be acknowledged here. Wearable devices can measure physiological signals such as HRV, movement, and heart rate, but they cannot directly capture daytime impairment such as fatigue, difficulty concentrating, or mood disturbance, or the subjective distress about sleep, which are necessary components of a formal insomnia diagnosis. This means that wearable data alone is insufficient for diagnosis. The framework proposed here positions wearable-derived data as an adjunct to clinical assessment, not a replacement for it.

This paper proposes that these findings, taken together, form the basis for a new clinical framework. That is, wearable-based physiological data are used as an adjunct to standard clinical evaluation of insomnia. Rather than asking a patient

"How much do you sleep?" (a question research shows they cannot reliably answer), a clinician could review weeks of wearable data showing HRV trajectories, sleep timing variability, and nocturnal heart rate patterns, interpreted by an AI model trained on confirmed insomnia cases. This turns diagnosis from a subjective conversation into an evidence-based assessment. This specific argument (that wearable AI data should replace or supplement self-report in clinical insomnia diagnosis to reduce misdiagnosis and prevent avoidable prescriptions of controlled substances) does not appear explicitly in the current literature and is a novel contribution of this review.

A concrete proposed clinical workflow could proceed as follows: a clinician suspects insomnia based on an initial interview and asks the patient to wear a consumer device (such as an Oura Ring or Apple Watch) continuously for a minimum of two to four weeks. Core signals collected include nightly HRV trajectory, sleep timing variability, and nocturnal heart rate patterns. An AI model (such as those described in sections 5.2a and 5.3) then classifies the patient's physiological profile as low, moderate, or elevated risk for insomnia disorder. When wearable data conflict with self-report (for example, a patient reports severe insomnia but the device shows normal HRV and sleep patterns across multiple weeks), the clinician uses this discrepancy for further evaluation. This would include assessment of paradoxical insomnia, daytime functioning, and psychological factors before prescribing a medication. Again, the wearable data do not replace clinical judgment; they simply inform it.

6.2. What an Objective Diagnostic Tool Could Catch

We will consider three clinical scenarios that current diagnostic practice handles poorly.

Scenario 1: Paradoxical insomnia—A teenager presents with severe insomnia complaints. A wearable device recording their sleep over four weeks shows normal total sleep time and HRV within healthy ranges. The AI model classifies their data as consistent with healthy sleep. This objectively suggests paradoxical insomnia (a condition requiring reassurance and cognitive behavioral therapy instead of sleep medication).

Scenario 2: Unrecognized sleep-risk patterns—This is the opposite of Scenario 1. A teenager reports sleeping fine, but wearable data shows progressively declining HRV, increasing sleep timing variability, and elevated nocturnal heart rate across several weeks. The AI model flags these as physiological patterns associated with elevated insomnia risk. Rather than constituting a diagnosis, this pattern would prompt a more thorough clinical evaluation that might otherwise never occur.

Scenario 3: Subclinical insomnia that is preventable—A teenager's wearable data shows an early trajectory toward insomnia while their ISI score is still healthy. The AI model detects the trajectory before clinical symptoms emerge, enabling early behavioral intervention before the insomnia becomes established and dangerous to personal well-being.

6.3. Why Teenagers Specifically

Teenagers are the ideal population for this approach for several reasons. First, younger age is a risk factor for sleep state misperception and paradoxical insomnia. This alone makes objective measurement essential in this group [15]. Second, teenagers are at a developmental stage where establishing healthy sleep habits can have lifelong benefits. Insomnia is associated with depression and other psychiatric disorders, and is a risk factor for suicidality and substance abuse in adolescents. Third, insomnia in adolescence tends to be chronic. Eighty-eight percent of adolescents with a history of insomnia report current insomnia. This means that catching it early matters enormously [5].

7. Limitations and Challenges

7.1. The Wakefulness Detection Problem

The most significant technical limitation is what researchers call the wakefulness detection problem. Consumer wearables are very good at detecting sleep (they correctly identify 90%+ of sleep epochs). However, they are much worse at detecting wakefulness, with specificity rates of only 29% - 52% [25]. For insomnia patients, who lie still while awake, this means wearables will overestimate their sleep, potentially creating a new form of measurement bias, just in the opposite direction from self-report. This limitation needs to be addressed quickly before wearables can fully replace self-report as a clinical diagnostic tool.

7.2. Individual Baselines and the Need for Longitudinal Data

No single night of wearable data is diagnostically meaningful. Insomnia is defined as a persistent pattern over at least three months6. What matters for AI-based diagnosis is the trajectory of that reading over time. The 2024 XGBoost study found that prediction accuracy improved with more days of data [31]. For clinical use, this means wearable-based diagnosis requires a period of data collection before any assessment can be made. However, the approach of wearable devices is far more accessible than the traditional route [35].

7.3. Algorithmic Bias and Representation

Most AI models for sleep detection have been trained on data from adults in laboratory settings. This can sometimes (often) lack diversity. Research has shown that differences between self-reported and objective sleep measures vary by race and ethnicity [12]. Consumer wearable accuracy may also vary by skin tone, since PPG sensors using light to measure blood flow may perform differently across skin tones [25]. Studies validating consumer wearables (specifically in adolescent populations) are also limited. One of the few studies that tracked adolescents specifically found that deep learning models outperformed traditional methods for that age group, but the field needs far more targeted research [21].

7.4. Privacy and Data Ownership

Continuous physiological monitoring of teenagers can raise serious privacy concerns. Who owns this data? How is it stored? Could it be shared with insurers, employers, or schools? Clear regulatory frameworks governing the use of health data from consumer wearables do not yet fully exist. Before wearable-based diagnosis is integrated into clinical practice, these questions need answers.

8. Conclusions

Insomnia is among the most common health problems affecting teenagers, and its consequences are profound. Yet, the current system for diagnosing it is susceptible to significant limitations. It relies on patients accurately reporting how much they sleep, and decades of research show they cannot. The result is misdiagnosis in both directions. People who sleep normally are diagnosed with insomnia and prescribed powerful medications, while people with genuine insomnia who do not complain are left undetected and untreated.

Consumer wearable devices (already on millions of teenagers' wrists) offer a path out of this diagnostic failure. By continuously measuring HRV, movement, heart rate, and other physiological signals, these devices capture objective data that do not depend on memory, perception, or honesty. AI models can analyze these data with 85% - 95% accuracy, and in some studies, outperform standard clinical questionnaires in predicting sleep quality [23] [31].

The strange and underappreciated phenomenon of paradoxical insomnia makes objective measurement even more important. Research confirms it is more prevalent in younger people, making teenagers the population most in need of an objective diagnostic check. A teenager with paradoxical insomnia who presents to a doctor is likely to be diagnosed with insomnia and potentially prescribed sleep medication. In addition, they could be fabricating their sleep diaries to obtain drugs. However, a wearable device would tell a very different story, preventing such occurrences from happening again.

This review proposes that integrating wearable-derived physiological data into the clinical diagnosis of insomnia, and using AI to interpret that data, represent the most significant potential advancement in insomnia diagnosis in decades.

If done so correctly, it can also address three underappreciated failures of the current system: insomnia identity, paradoxical insomnia, and the challenges of objectively verifying insomnia claims before prescribing controlled substances. The technology exists. The AI tools exist. The data are, in many cases, already being collected. What remains is the clinical will to use it, the research to validate it in diverse adolescent populations, and the regulatory frameworks to make it safe and equitable.

The next generation of insomnia diagnosis should not begin with the question, "How is your sleep?" It should begin with 30 days of data from the watch already on the patient's wrist.

Another potential use of wearable devices could be accurately assessing the incidence and prevalence of subclinical and clinical insomnia in adolescents, which can potentially lead to chronic disorders such as substance abuse, attention deficit disorder (ADD), or increased criminal behavior.

Acknowledgements

Claude AI by Anthropic was used to improve readability and flow. Google Gemini was used to help generate some images and tables.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Wheaton, A.G., Jones, S.E., Cooper, A.C. and Croft, J.B. (2018) Short Sleep Duration among Middle School and High School Students—United States, 2015. *MMWR. Morbidity and Mortality Weekly Report*, **67**, 85-90. https://doi.org/10.15585/mmwr.mm6703a1

[2] CDC (2024) FastStats: Sleep in High School Students | Sleep. https://www.cdc.gov/sleep/data-research/facts-stats/high-school-students-sleep-facts-and-stats.html

[3] Krewson, C. (2025) CDC: Most US Teens Fail to Get Recommended 8 Hours of Sleep, Report Finds. Contemporary Pediatrics. https://www.contemporarypediatrics.com/view/cdc-most-us-teens-fail-to-get-recommended-8-hours-of-sleep-report-finds

[4] Bommersbach, T.J., Olfson, M. and Rhee, T.G. (2026) Insufficient Sleep among US Adolescents across Behavioral Risk Groups. *JAMA*, **335**, 1173-1176. https://doi.org/10.1001/jama.2026.1417

[5] de Zambotti, M., Goldstone, A., Colrain, I.M. and Baker, F.C. (2018) Insomnia Disorder in Adolescence: Diagnosis, Impact, and Treatment. *Sleep Medicine Reviews*, **39**, 12-24. https://doi.org/10.1016/j.smrv.2017.06.009

[6] Höhn, C., Hahn, M.A., Gruber, G., Pletzer, B., Cajochen, C. and Hoedlmoser, K. (2024) Effects of Evening Smartphone Use on Sleep and Declarative Memory Consolidation in Male Adolescents and Young Adults. *Brain Communications*, **6**, fcae173. https://doi.org/10.1093/braincomms/fcae173

[7] Silvani, M.I., Werder, R. and Perret, C. (2022) The Influence of Blue Light on Sleep, Performance and Wellbeing in Young Adults: A Systematic Review. *Frontiers in Physiology*, **13**, Article 943108. https://doi.org/10.3389/fphys.2022.943108

[8] Alam, M. (2024) Impacts of Blue Light Exposure from Electronic Devices on Circadian Rhythm and Sleep Disruption in Adolescent and Young Adult Students. chronobiolo-gyinmedicine.org. https://www.chronobiologyinmedicine.org/journal/view.php?number=167&viewtype=pubreader

[9] Moul, D.E., Hall, M., Pilkonis, P.A. and Buysse, D.J. (2004) Self-Report Measures of Insomnia in Adults: Rationales, Choices, and Needs. *Sleep Medicine Reviews*, **8**, 177-198. https://doi.org/10.1016/s1087-0792(03)00060-1

[10] Scarlett, S. (2021) Discrepancies in Self-Reported and Actigraphy-Based Sleep Duration Are Associated with Self-Reported Insomnia Symptoms in Community-Dwell-

ing Older Adults. *Sleep Health*, **7**, 83-92.
https://www.sleephealthjournal.org/article/S2352-7218(20)30173-X/fulltext

[11] Korman, M., Zarina, D., Tkachev, V., Merikanto, I., Bjorvatn, B., Bjelajac, A.K., *et al.* (2024) Estimation Bias and Agreement Limits between Two Common Self-Report Methods of Habitual Sleep Duration in Epidemiological Surveys. *Scientific Reports*, **14**, Article No. 3420. https://doi.org/10.1038/s41598-024-53174-1

[12] Robbins, R., Quan, S.F., Barger, L.K., Czeisler, C.A., Fray-Witzer, M., Weaver, M.D., *et al.* (2021) Self-Reported Sleep Duration and Timing: A Methodological Review of Event Definitions, Context, and Timeframe of Related Questions. *Sleep Epidemiology*, **1**, Article ID: 100016. https://doi.org/10.1016/j.sleepe.2021.100016

[13] Biddle, D.J., Robillard, R., Hermens, D.F., Hickie, I.B. and Glozier, N. (2015) Accuracy of Self-Reported Sleep Parameters Compared with Actigraphy in Young People with Mental Ill-Health. *Sleep Health*, **1**, 214-220.
https://doi.org/10.1016/j.sleh.2015.07.006

[14] Lamberg, L. (2018) Self-Reports of Poor Sleep Not Always Accurate. *Psychiatric News*, **53**. https://doi.org/10.1176/appi.pn.2018.9a11

[15] Lindert, T. (2025) Paradoxical Insomnia. MedLink Neurology.
https://www.medlink.com/articles/paradoxical-insomnia

[16] Joo, E.H., Altier, H.R., Selai, C., Gratton, M.K., Kim-Dahl, A., Allen, H., *et al.* (2025) Neurobiological Mechanisms of Sleep State Misperception in Insomnia Disorder: A Theoretical Review. *Sleep Medicine Reviews*, **81**, Article ID: 102096.
https://doi.org/10.1016/j.smrv.2025.102096

[17] Russo, F. and Stix, G. (2024) Some People with Insomnia Think They're Awake When They're Asleep. Scientific American.
https://www.scientificamerican.com/article/some-people-with-insomnia-think-theyre-awake-when-theyre-asleep/

[18] Harvey, A.G. and Tang, N.K.Y. (2012) (Mis)perception of Sleep in Insomnia: A Puzzle and a Resolution. *Psychological Bulletin*, **138**, 77-101.
https://doi.org/10.1037/a0025730

[19] Short, N.A., Austin, A.E. and Naumann, R.B. (2023) Associations between Insomnia Symptoms and Prescription Opioid and Benzodiazepine Misuse in a Nationally Representative Sample. *Addictive Behaviors*, **137**, Article ID: 107507.
https://doi.org/10.1016/j.addbeh.2022.107507

[20] Roehrs, T.A. and Roth, T. (2015) Sleep Disturbance in Substance Use Disorders. *Psychiatric Clinics of North America*, **38**, 793-803.
https://doi.org/10.1016/j.psc.2015.07.008

[21] Sathyanarayana, A., Joty, S., Fernandez-Luque, L., Ofli, F., Srivastava, J., Elmagarmid, A., *et al.* (2016) Sleep Quality Prediction from Wearable Data Using Deep Learning. *JMIR mHealth and uHealth*, **4**, e125. https://doi.org/10.2196/mhealth.6562

[22] Dial, M.B., Hollander, M.E., Vatne, E.A., Emerson, A.M., Edwards, N.A. and Hagen, J.A. (2025) Validation of Nocturnal Resting Heart Rate and Heart Rate Variability in Consumer Wearables. *Physiological Reports*, **13**, e70527.
https://doi.org/10.14814/phy2.70527

[23] Lee, H., Cho, M., Lee, S.W. and Park, S. (2025) Predicting Sleep Quality with Digital Biomarkers and Artificial Neural Networks. *Frontiers in Psychiatry*, **16**, Article 1591448. https://doi.org/10.3389/fpsyt.2025.1591448

[24] Lee, T., Cho, Y., Cha, K.S., Jung, J., Cho, J., Kim, H., *et al.* (2023) Accuracy of 11 Wearable, Nearable, and Airable Consumer Sleep Trackers: Prospective Multicenter

Validation Study. *JMIR mHealth and uHealth*, **11**, e50983. https://doi.org/10.2196/50983

[25] Schyvens, A., Peters, B., Van Oost, N.C., Aerts, J., Masci, F., Neven, A., *et al.* (2025) A Performance Validation of Six Commercial Wrist-Worn Wearable Sleep-Tracking Devices for Sleep Stage Scoring Compared to Polysomnography. *Sleep Advances*, **6**, zpaf021. https://doi.org/10.1093/sleepadvances/zpaf021

[26] Robbins, R., Weaver, M.D., Sullivan, J.P., Quan, S.F., Gilmore, K., Shaw, S., *et al.* (2024) Accuracy of Three Commercial Wearable Devices for Sleep Tracking in Healthy Adults. *Sensors*, **24**, Article 6532. https://doi.org/10.3390/s24206532

[27] IBM: What Is Random Forest? https://www.ibm.com/think/topics/random-forest

[28] Islam, M.M., Masum, A.K.M., Abujar, S. and Hossain, S.A. (2020) Prediction of Chronic Insomnia Using Machine Learning Techniques. 2020 11*th International Conference on Computing, Communication and Networking Technologies* (*ICCCNT*), Kharagpur, 1-3 July 2020, 1-7. https://doi.org/10.1109/icccnt49239.2020.9225570

[29] IBM: What Are Support Vector Machines (SVMs)? https://www.ibm.com/think/topics/support-vector-machine

[30] Extreme Gradient Boosting. Science Direct. https://www.sciencedirect.com/topics/computer-science/extreme-gradient-boosting

[31] Park, K.M., Lee, S.E., Lee, C., Hwang, H.D., Yoon, D.H., Choi, E., *et al.* (2024) Predicting Sleep Based on Physical Activity, Light Exposure, and Heart Rate Variability Data Using Wearable Devices. *Annals of Medicine*, **56**, Article ID: 2405077. https://doi.org/10.1080/07853890.2024.2405077

[32] Baigutanova, A., Park, S., Constantinides, M., Lee, S.W., Quercia, D. and Cha, M. (2025) A Continuous Real-World Dataset Comprising Wearable-Based Heart Rate Variability Alongside Sleep Diaries. *Scientific Data*, **12,** Article No. 1474. https://www.nature.com/articles/s41597-025-05801-3

[33] Yu, R., Li, Y., Zhao, K. and Fan, F. (2024) A Review of Automatic Sleep Stage Classification Using Machine Learning Algorithms Based on Heart Rate Variability. *Sleep and Biological Rhythms*, **23**, 113-125. https://doi.org/10.1007/s41105-024-00563-8

[34] Coutts, L.V., Plans, D., Brown, A.W. and Collomosse, J. (2020) Deep Learning with Wearable Based Heart Rate Variability for Prediction of Mental and General Health. *Journal of Biomedical Informatics*, **112**, Article ID: 103610. https://doi.org/10.1016/j.jbi.2020.103610

[35] Schyvens, A., Van Oost, N.C., Aerts, J., Masci, F., Peters, B., Neven, A., *et al.* (2024) Accuracy of Fitbit Charge 4, Garmin Vivosmart 4, and WHOOP versus Polysomnography: Systematic Review. *JMIR mHealth and uHealth*, **12**, e52192-e52192. https://doi.org/10.2196/52192