Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

Advances in Machine Learning Algorithms in the Treatment of Obstructive Sleep Apnea

[1] Ugwu Hillary*, [2] Mohammed Mostafa, [3] Ugwu Linus Chinedu2

[1, 2] Department of Intensive Care Unit, Prince Mutaib Bin Abdulaziz Hospital, Sakaka, Aljouf, Saudi Arabia

[3] Department of Medical Laboratory Sciences, Idaho State University, Pocatello, United States

(Received: 27 September 2025 Revised: 05 October 2025 Accepted: 10 November 2025)

KEYWORDS

Obstructive Sleep Apnea (OSA), Machine Learning, Deep Learning, Diagnosis, Treatment Optimization, CPAP Compliance, Personalized Medicine, Remote Monitoring, Wearable Devices, Artificial Intelligence, Healthcare.

ABSTRACT:

Recent developments in machine learning (ML) and deep learning (DL) algorithms for the diagnosis, screening, and most importantly, treatment of obstructive sleep apnea (OSA) are thoroughly reviewed in this review article. The analysis highlights the empirical performance of different AI-powered models and identifies important research gaps by synthesizing findings from recent literature (2018–2025). With accuracy, sensitivity, and specificity frequently surpassing those of conventional tools, ML and DL models have shown notable advancements in OSA detection. For instance, using electrocardiogram (ECG) and oxygen saturation (SpO2) signals, a multimodal signal fusion multiscale Transformer model achieved 91.38% per-segment and 96.08% per-recording accuracy for OSA detection and severity assessment. The OSA event detection accuracy of a different one-dimensional convolutional neural network (1D-CNN) model for portable monitors was 84.3%. Artificial intelligence (AI) is being used more and more for treatment optimization in addition to diagnosis. This includes predicting and enhancing Continuous Positive Airway Pressure (CPAP) compliance, enabling remote monitoring and chronic disease management, and facilitating customized treatment regimens. The underrepresentation of diverse demographic groups in study cohorts, small sample sizes, and limited robust model validation are among the major research gaps that still exist despite these advancements. These factors work together to impede model generalizability and equitable application in healthcare. The future of AI in OSA management depends on addressing these constraints through thorough validation and inclusive data collection.

1. Introduction

Recurrent episodes of upper airway collapse during sleep are the hallmark of Obstructive Sleep Apnea (OSA), a chronic sleep disorder that is extremely common but frequently goes undiagnosed. An individual's health and quality of life are greatly impacted by these episodes, which cause intermittent hypoxia and fragmented sleep. OSA has a significant public health burden, affecting almost one billion people worldwide, with prevalence rates surpassing 50% in some nations (Wang et al., 2025). Cardiovascular conditions like hypertension, heart failure, and stroke are among the serious comorbidities linked to untreated OSA. It also raises the risk of accidents, metabolic diseases like diabetes, and neurocognitive impairment, all of which have a significant negative impact on general health and well-being.

Full-night polysomnography (PSG), the current gold standard for diagnosing OSA, has a number of drawbacks. PSG is an expensive and time-consuming procedure that frequently necessitates an overnight stay in a specialized sleep laboratory. The underdiagnosis of OSA is largely caused by this intrinsic inaccessibility (Hassan et al., 2025; Manoni & Pardo, 2023). Traditional treatment approaches, like Continuous Positive Airway Pressure (CPAP) therapy, also have a lot of problems, especially with low patient compliance rates, which frequently fall below 50% (Wang et al., 2025). Persistent health risks result from this poor adherence, which also jeopardizes patient outcomes and therapeutic efficacy.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

Other treatments, such as surgical procedures or oral appliances, also require careful patient selection and optimization, a process that conventional approaches find difficult to accomplish effectively.

Rapid developments in artificial intelligence (AI), especially in machine learning (ML) and deep learning (DL) algorithms have started a revolution in a number of industries, including healthcare. AI provides innovative and potent solutions to persistent problems in patient care, treatment planning, and medical diagnosis. These technologies promise improved healthcare delivery in terms of accuracy, efficiency, and personalization (Alvarez & Hornero, 2024; Hassan et al., 2025). The proliferation of AI applications in medicine has been fueled by the growing availability of large, complex biomedical datasets as well as notable advances in computing power. These developments have made it possible to create complex models that can process and interpret enormous volumes of health data.

According to Wang et al. (2025), machine learning algorithms are in a unique position to analyze large volumes of intricate physiological data and identify subtle patterns that may be overlooked by conventional diagnostic techniques or human observation but are indicative of OSA. From early and precise screening and diagnosis to treatment outcome prediction, therapeutic regimen optimization, and continuous remote monitoring, machine learning (ML) offers promise for managing OSA (Hassan et al., 2025; Wang et al., 2025). Through earlier detection, less strain on healthcare systems, and the development of highly customized treatment plans based on each patient's needs and responses, this AI integration promises to improve patient outcomes (Alvarez & Hornero, 2024). The use of AI in OSA research has increased significantly, especially since 2017, as the scientific community has come to realize the technology's enormous potential (Wang et al., 2025). The main goals of early applications were to improve diagnostic accuracy and create predictive models for diagnosis. These models made use of a number of physiological parameters, such as electroencephalogram (EEG) readings, oxygen saturation (SpO2) data, and electrocardiogram (ECG) signals (Wang et al., 2025). The focus of research has recently broadened to encompass important topics like remote management tactics, personalized therapy approaches, and treatment optimization. This growth shows that the field has matured, going beyond its diagnostic potential to address the more comprehensive issues of long-term OSA care (Hassan et al., 2025; Wang et al., 2025).

Given the A thorough review is necessary to compile recent developments and give a clear picture of the state of the art, especially considering the speed at which AI is developing and the growing number of applications it has in OSA. In particular, this review seeks to close the gap by critically assessing the new and increasingly important role of machine learning in OSA treatment optimization and management, in addition to recognizing the notable advancements in OSA diagnosis and screening (Hassan et al., 2025). Although this latter area has significant implications for patient care, it has received less review. This article offers a current synthesis of empirical results, methodologies, and identified challenges by concentrating on recent citations (2018–2025), making it a pertinent and timely resource for the field. (Hassan et al., 2025).

Even with the encouraging outcomes, there are still a number of important research gaps that prevent ML from being widely and fairly applied in OSA. The widespread absence of representative and varied datasets is a significant drawback. Currently, the majority of study participants are overweight men, while women, younger obese adults, people over 60, and people of different races are significantly underrepresented (Hassan et al., 2025). The generalizability of ML models is significantly impacted by this demographic imbalance in the training data (Hassan et al., 2025). Models will naturally pick up patterns unique to that group if they are trained primarily on data from a limited demographic. As a result, their performance—in terms of accuracy, sensitivity, and specificity, is likely to deteriorate considerably when applied to people who are not in this demographic.

Healthcare equity becomes a crucial issue as a result. The advantages of cutting-edge technology are not shared equally if AI-powered diagnostic or treatment tools are less accurate or effective for particular populations because of biased training data. This might worsen already-existing health inequalities and result in incorrect diagnoses or inadequate care for marginalized populations. The promise of personalized medicine for all is undermined when the pursuit of high accuracy in research is not combined with careful consideration of

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

data diversity, which unintentionally leads to systemic biases in AI applications. Concerns regarding the validity and relevance of results in actual clinical settings are also raised by the fact that many studies have small sample sizes and little application of strong model validation techniques (Hassan et al., 2025). These methodological flaws highlight the necessity of more exacting procedures to guarantee the external validity and robustness of AI models in OSA.

In order to lay the groundwork for more complex applications, this review article is organized to first give a summary of the notable advancements in machine learning applications for OSA diagnosis and screening. It then dives into the main topic, which is the developing and crucial role of machine learning in managing and optimizing OSA treatment. Following a thorough discussion of the identified research gaps and future directions, the article concludes with a comprehensive view of the empirical results and statistical analyses. In order to direct future research efforts towards more reliable, broadly applicable, and equitable AI solutions in sleep medicine, the goal is to create a useful resource for physicians, researchers, and policymakers.

2. Machine Learning in OSA: Advances in Diagnosis and Screening

The diagnosis and screening of Obstructive Sleep Apnea (OSA) has greatly improved thanks to machine learning algorithms, which provide more effective, precise, and easily accessible substitutes for conventional polysomnography (PSG). More complex applications in treatment optimization are made possible by the advancements made in this field.

2.1. Detailed Review of ML/DL Algorithms Applied to OSA Detection

For OSA detection, a variety of machine learning models, both conventional and deep learning-based have been widely used. These models have shown varying degrees of effectiveness across various data kinds and goals. The foundation of early diagnostic efforts has been established by traditional machine learning algorithms. Significant gains in OSA detection have been demonstrated by the widespread use of logistic regression (LR), especially when combined with anthropometric measures (Alvarez & Hornero, 2024; Hassan et al., 2025). Area Under the Curve (AUC) values for LR classifiers have reached 0.761 (Alvarez & Hornero, 2024). With reported sensitivities of 93.4% in predicting OSA patients, Support Vector Machines (SVM) have proven to be highly effective. SVM models have produced AUROC values of 0.84 when paired with diffusor tensor imaging (Alvarez & Hornero, 2024).

Prediction models based on Random Forest (RF) have demonstrated high accuracy and AUC, with notable values of 0.81 for severe OSA and 0.90 for moderate-to-severe OSA (Alvarez & Hornero, 2024). With an AUC of 0.768, Naive Bayes (NB) has demonstrated superior specificity (59.5%) in predicting healthy individuals, despite not always matching SVM in sensitivity (Alvarez & Hornero, 2024). A correct OSA detection rate of 82.9% and a severity classification accuracy of 74.3% have been reported when Decision Trees (DT) have been used for automatic at-home detection and severity classification (Alvarez & Hornero, 2024). Additionally successful are Multilayer Perceptron Networks (MLPs), which have an 86% accuracy rate in identifying patients as either healthy or OSA-affected. In a different example, MLPs were able to classify sleep stages and screen for OSA with 73% accuracy. (Alvarez & Hornero, 2024).

Because deep learning (DL) algorithms can learn complex, hierarchical features straight from raw data without requiring a lot of manual feature engineering, they typically perform better than traditional machine learning (ML) algorithms in identifying sleep patterns (Manoni & Pardo, 2023). Compared to more complex architectures, one-dimensional convolutional neural networks (1D-CNNs) require less computing power and data preprocessing, making them especially effective with raw physiological signals (Manoni & Pardo, 2023). In identifying sleep apnea episodes, a 1D-CNN model created especially for portable monitors demonstrated 84.3% accuracy, 82.5% sensitivity, 86% specificity, and 92.1% AUC (Manoni & Pardo, 2023). Transformer models are examples of more sophisticated architectures. In comparison to earlier models, a multimodal signal fusion multiscale Transformer model that used ECG and SpO2 signals demonstrated superior performance, achieving high accuracy for OSA detection and severity assessment (Li et al., 2025). This model's innovative design, incorporating self-attention, multihead attention, and co-attention mechanisms, was crucial for maximizing feature extraction and fusion capabilities across different signal modalities (Li et al., 2025).

Bidirectional Gated Recurrent Units (Bi-GRUs), a type of recurrent neural network (RNN), are used in complex architectures to learn long-term dependencies in time-series data. This is essential for the analysis of continuous physiological signals (Li et al., 2025). In addition, final decision fusion has been performed using Hidden Markov Models (HMMs), which have been shown to identify ECG segments containing OSA with 85% accuracy and 88.9% sensitivity (Wang et al., 2025).

2.2. Discussion of Various Data Sources for Diagnostic Models

The successful use of ML/DL algorithms for OSA diagnosis and screening across a variety of input data sources, going beyond the conventional reliance on full PSG, demonstrates their adaptability. Despite being the gold standard for diagnosis, polysomnography (PSG) data is a rich source of raw signals (such as ECG, SpO2, EEG, and respiratory efforts) that can be processed by ML/DL models to automate scoring and detection. This eliminates the drawbacks of manual, time-consuming scoring (Hassan et al., 2025; Li et al., 2025; Manoni & Pardo, 2023; Wang et al., 2025). Because they are easy to obtain and provide rich physiological information about cardiac responses during sleep, electrocardiogram (ECG) signals are widely used.

Diagnostic models have demonstrated good efficacy when constructed using ECG alone or in conjunction with other signals (Li et al., 2025; Wang et al., 2025). OSA is characterized by Oxygen Saturation (SpO2) signals, especially nocturnal desaturations. Because of their constant monitoring through pulse oximetry, SpO2 is a very useful and accessible data source for machine learning models; for better results, it is frequently combined with ECG (Li et al., 2025; Manoni & Pardo, 2023).

A convenient and less invasive option to in-lab studies, data from smartphones and wearable devices (such as respiratory sounds, heart rate, movement, and photoplethysmography) are being used more and more for remote and home-based OSA screening as smart devices proliferate (Hassan et al., 2025; Wang et al., 2025). For example, smartphone applications have shown similar sensitivity, specificity, and accuracy to traditional methods (e.g., 93.3%, 94.4%, and 94.3% respectively for severe OSA screening using iOS and Android phones) (Alvarez & Hornero, 2024). Because of their ease of use and affordability, anthropometric indicators—such as body measurements and physical traits are frequently employed for preliminary screening in addition to direct physiological signals, particularly in logistic regression models (Alvarez & Hornero, 2024; Hassan et al., 2025). Additionally, ML/DL analysis of respiratory sounds has demonstrated encouraging outcomes in the diagnosis of respiratory disorders, underscoring their potential for early detection and disease monitoring (Alvarez & Hornero, 2024). Additionally, AI can screen populations with high-risk OSA facial structures by utilizing facial recognition technology (Wang et al., 2025). A quick, efficient, and affordable screening method has been found by combining various machine learning algorithms with 3D geometric morphometrics (Alvarez & Hornero, 2024). Additionally, diffusor tensor imaging has been investigated for screening (Alvarez & Hornero, 2024). Lastly, in order to create thorough diagnostic models, machine learning models can incorporate a variety of clinical features, such as medical histories and symptoms, in addition to other physiological parameters like heart rate, X-ray images, and snoring characteristics (Wang et al., 2025).

2.3. Empirical Performance Metrics of Diagnostic Models

The accuracy, sensitivity, and specificity of AI algorithms in OSA detection have continuously outperformed those of conventional methods (Alvarez & Hornero, 2024; Hassan et al., 2025). Recent studies' empirical findings demonstrate these models' remarkable potential.

For the detection of moderate-to-severe OSA, certain models, such as Random Forest (RF), have demonstrated high accuracy and AUC values of 0.90 (Alvarez & Hornero, 2024). For OSA patient prediction, Support Vector Machines (SVM) have shown an especially high sensitivity of 93.4% (Alvarez & Hornero, 2024). Using ECG and SpO2 signals, a multimodal signal fusion multiscale Transformer model has demonstrated impressive performance, achieving a per-segment detection accuracy of 91.38% with an AUC of 0.968 and a per-recording detection accuracy of 96.08% with an AUC of 0.922. (Li et al., 2025). Achieving 90.20% for mild OSA, 88.24% for moderate OSA, and 92.16% for severe OSA, this model also demonstrated high accuracy for severity assessment (Li et al., 2025). A 1D-CNN model showed strong performance for OSA event detection on portable monitors, with 84.3% accuracy, 82.5% sensitivity, 86% specificity, and 92.1% AUC (Manoni & Pardo, 2023). An 82.9% correct OSA detection rate and a 74.3% severity classification accuracy were attained by home-based detection systems that integrated Decision Tree and Logistic Regression models (Alvarez & Hornero, 2024). In addition to these metrics, AI-powered home-based OSA testing methods have shown notable cost-effectiveness, with an Italian study finding that they reduced direct medical costs by 44%, personal expenses by 37%, and societal costs by 20% (Alvarez & Hornero, 2024).

A critical observation concerning the implementation of these capabilities in common clinical practice arises, despite the fact that empirical results consistently report high performance metrics for various ML/DL models in OSA diagnosis and severity assessment, frequently outperforming traditional tools. Even with the remarkable precision and effectiveness shown in controlled environments, there are still a number of real-world restrictions. Among these is the constant requirement for thorough dataset standardization and validation across various research projects and organizations (Alvarez & Hornero, 2024; Hassan et al., 2025). The fact that a large number of studies are carried out at single centers raises concerns regarding the models' applicability to various patient populations and acquisition settings (Li et al., 2025). Additionally, despite their high accuracy, some sophisticated models have higher computational costs, which may prevent their widespread use (Li et al., 2025).

Additionally mentioned are issues with data quality and the requirement for better cross-validation techniques (Manoni & Pardo, 2023). This situation highlights a paradox: the scientific community has achieved remarkable laboratory-based performances, demonstrating what is technically possible with AI in OSA diagnosis. However, converting these capabilities into scalable, equitable, and clinically viable solutions is the next crucial stage. This calls for a change in the research focus from just achieving high performance metrics to tackling the underlying data infrastructure, improving the rigor of validation, and resolving real-world deployment issues. The "advances" point to a critical area for further development because they are more theoretical in nature than broadly applicable in clinical settings.

3. Machine Learning in OSA: Advances in Treatment Optimization and Management

Although machine learning (ML) has a well-established role in diagnosis, its use in long-term patient management and in optimizing and customizing OSA treatment is a critical and developing area. Beyond a one-size-fits-all strategy, this field has enormous potential to address the major issues of treatment adherence and customized therapy.

3.1. Review of ML Applications in Predicting and Improving CPAP Compliance

Although continuous positive airway pressure (CPAP) is commonly acknowledged as the first-line treatment for OSA, low patient compliance continues to be a significant barrier to its efficacy, with adherence rates frequently falling below 50% (Wang et al., 2025). To tackle this challenge, machine learning technology provides effective tools.

By examining a wide range of patient clinical data, such as medical histories, symptoms, and reactions to initial treatment, machine learning models can be created to forecast CPAP treatment compliance (Wang et al., 2025). Clinicians can proactively identify patients at high risk of non-adherence thanks to this predictive capability, allowing for focused interventions before adherence problems worsen. This signifies a significant change in patient care from reactive to proactive. Clinicians can use ML-derived insights to predict issues and provide prompt, individualized support rather than waiting for a patient to stop using CPAP. It is possible to greatly enhance long-term results and lessen the burden of avoidable complications linked to untreated OSA by changing the management of chronic diseases from a reactive, crisis-driven approach to a proactive, preventative one.

ML actively improves CPAP adherence in addition to making predictions. This is accomplished by putting in place monitoring systems with AI processors to track CPAP usage in real time and give clinicians and patients insightful feedback (Wang et al., 2025). Additionally, the direct incorporation of AI into CPAP devices allows for adaptive therapy modifications and real-time monitoring, which can improve patient comfort and efficacy and, consequently, adherence (Wang et al., 2025). AI-powered systems can also help with remote medical interventions and reminders, providing patients with ongoing support in sticking to their treatment plan and quickly resolving any new problems (Wang et al., 2025).

3.2. Exploration of Personalized Treatment Regimens Facilitated by ML

By integrating a variety of patient clinical data, such as medical histories, symptoms, and treatment responses, machine learning and deep learning models are being used more and more to predict treatment outcomes (Wang et al., 2025). Moving away from a generalized "one-size-fits-all" approach to therapy that frequently turns out to be suboptimal for many people, this capability is crucial in enabling customized treatment regimens (Wang et al., 2025).

One important use of AI is to improve the effectiveness and precision of OSA subtype identification (Wang et al., 2025). In keeping with the developing idea of "endotypes" and the creation of particular "biomarkers" for disease mechanisms, this subtyping is essential for customizing treatment plans to the unique characteristics of each patient and the underlying disease mechanisms (Hassan et al., 2025).

AI can direct clinicians toward the most suitable and successful therapeutic interventions by comprehending the distinct physiological drivers of OSA in each patient.

In addition to CPAP, ML can forecast the results of surgical procedures, including transoral robotic surgery, and other treatment modalities, including oral appliances (Wang et al., 2025). Clinicians can choose patients who are most likely to benefit from particular procedures and make better decisions about the best treatment options with the help of this predictive power.

Additionally, AI can forecast neurological conditions, identify cognitive impairment, and predict mortality from cardiovascular disease in OSA patients by analyzing complex data, including brain MRI scans and large patient datasets (Wang et al., 2025). The creation of comprehensive, individualized management plans that address the various health effects of OSA is made possible by this thorough evaluation of comorbidities and prognosis.

3.3. Role of AI in Remote Monitoring and Chronic Disease Management for OSA Patients

The use of AI in remote medical care and chronic disease management for OSA is growing dramatically due to the pervasiveness of smartphones and ongoing developments in wearable electronic device technology (Wang et al., 2025). The monitoring and management of sleep health is being revolutionized by this convergence of technologies.

AI eliminates the need for time-consuming and expensive in-lab PSG studies by enabling real-time home sleep monitoring, including sleep stages and quality (Wang et al., 2025). This allows for ongoing insights into a patient's condition. This feature transforms the episodic, centralized model of sleep health monitoring into a continuous, decentralized, and democratized one.

AI algorithms can process the continuous, raw, and frequently noisy data from consumer-grade devices to extract insights that are clinically relevant (Manoni & Pardo, 2023; Wang et al., 2025). The proliferation of smartphones and wearable technology has made data collection platforms more accessible and affordable. This implies that millions more people may benefit from early detection, continuous monitoring, and individualized treatment of OSA, especially those living in underprivileged areas or those unable to visit traditional sleep clinics. It may change how public health approaches to sleep disorders are approached globally by transforming sleep health from a specialized medical niche into an integrated part of overall wellness and the management of chronic diseases.

AI systems can offer early warnings for disease risk in addition to basic monitoring, enabling prompt interventions and proactive management of possible complications (Wang et al., 2025). Additionally, AI makes it possible for clinicians to remotely monitor CPAP usage, track adherence, troubleshoot problems, and offer support, all of which improve patient engagement and compliance (Wang et al., 2025). More autonomy and adherence to treatment plans are fostered by AI-powered tools on wearable technology and smartphones that enable patients to actively participate in their self-health management (Wang et al., 2025). Additionally, AI-enabled home-based OSA testing and management techniques have proven to be highly cost-effective, lowering overall medical, personal, and

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

societal expenses while increasing access to and affordability of healthcare (Alvarez & Hornero, 2024).

4. Empirical Results and Statistical Analysis

The main empirical results from the reviewed studies are compiled in this section, which also provides quantitative performance metrics of different ML/DL models for OSA diagnosis/screening and treatment optimization.

4.1. Consolidated Presentation of Key Empirical Findings

The improved capabilities of AI algorithms in OSA management are strongly supported by empirical data. Alvarez & Hornero, 2024; Hassan et al., 2025) AI algorithms frequently outperform conventional techniques in terms of diagnostic performance, exhibiting high accuracy, sensitivity, and specificity for OSA detection. For the detection of moderate-to-severe OSA, certain models, like Random Forest (RF), have demonstrated remarkable AUCs of 0.90 (Alvarez & Hornero, 2024). For OSA patient prediction, Support Vector Machines (SVM) has shown an especially high sensitivity of 93.4% (Alvarez & Hornero, 2024).

Using ECG and SpO2 signals, a multimodal signal fusion multiscale Transformer model has demonstrated impressive performance, achieving a per-segment detection accuracy of 91.38% (with an AUC of 0.968) and a per-recording detection accuracy of 96.08% (with an AUC of 0.922) (Li et al., 2025).

Additionally, it had a high accuracy rate for determining severity across all categories: 92.16% for severe OSA, 88.24% for moderate OSA, and 90.20% for mild OSA (Li et al., 2025). A 1D-CNN model showed strong performance for OSA event detection on portable monitors, with 84.3% accuracy, 82.5% sensitivity, 86% specificity, and 92.1% AUC (Manoni & Pardo, 2023). An 82.9% correct OSA detection rate and a 74.3% severity classification accuracy were shown by home-based detection systems that used Decision Tree and Logistic Regression models (Alvarez & Hornero, 2024).

In terms of treatment-related outcomes, the mechanisms and anticipated effects are well described, even though specific empirical findings demonstrating direct quantitative improvements in CPAP compliance percentages as a result of ML interventions are less explicitly detailed in the literature provided than diagnostic accuracy (Wang et al., 2025). According to Wang et al. (2025), machine learning models can be used to create predictive models for CPAP treatment compliance. With real-time monitoring and adaptive support, AI-enabled CPAP ventilators and monitoring systems are especially made to increase patient compliance (Wang et al., 2025). Alvarez & Hornero (2024) and Hassan et al. (2025) predict that AI will "improve patient outcomes, increase early detection, and lessen the load on healthcare systems" in OSA diagnosis and screening, pointing to a wider positive impact on treatment adherence and overall management.

4.2. Table 1: Performance Metrics of Key ML/DL Models for OSA Diagnosis/Screening (Recent Studies)

A succinct, comparative summary of the empirical efficacy of various ML/DL models in OSA diagnosis and screening is given in this table. It makes it possible to quickly compare the advantages and disadvantages of different strategies, demonstrating the noteworthy developments in this field.

Table 1

Model Type	Input Data	Key Performance Metric (Accuracy/AUC/Sensitivity/Specificity)	Citation
Random Forest (RF)	(Not specified, likely anthropometrics/clinical)	Accuracy & AUC: 0.90 (moderate-severe OSA), 0.81 (severe OSA)	(Alvarez & Hornero, 2024)
Support Vector Machine (SVM)	(Not specified, likely anthropometrics/clinical)	Sensitivity: 93.4%	(Alvarez & Hornero, 2024)
Naive Bayes (NB)	(Not specified, likely anthropometrics/clinical)	Specificity: 59.5%; AUC: 0.768	(Alvarez & Hornero, 2024)

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

Model Type	Input Data	Key Performance Metric (Accuracy/AUC/Sensitivity/Specificity)	Citation
Multilayer Perceptron (MLP)	(Not specified)	Accuracy: 86% (healthy/OSA classification)	(Alvarez & Hornero, 2024)
Multilayer Perceptron (MLP)	(Not specified)	Accuracy: 73% (sleep stage classification & OSA screening)	(Alvarez & Hornero, 2024)
Decision Tree (DT) & Logistic Regression (LR)	(Not specified)	Correct OSA Detection Rate: 82.9%; Severity Classification Accuracy: 74.3%	(Alvarez & Hornero, 2024)
Smartphone Apps	(iOS/Android)	Sensitivity: 93.3%; Specificity: 94.4%; Accuracy: 94.3% (severe OSA screening)	(Alvarez & Hornero, 2024)
Multimodal Transformer	ECG, SpO2	Per-segment Acc: 91.38%, AUC: 0.968	(Li et al., 2025)
Multimodal Transformer	ECG, SpO2	Per-recording Acc: 96.08%, AUC: 0.922	(Li et al., 2025)
Multimodal Transformer	ECG, SpO2	Severity Acc: Mild 90.20%, Mod 88.24%, Severe 92.16%	(Li et al., 2025)
1D-CNN (Portable Monitor)	SpO2, HR, Thor-Res, Abdo-Res (SHHS2)	Accuracy: 84.3%; Sensitivity: 82.5%; Specificity: 86%; AUC: 92.1%	(Manoni & Pardo, 2023)
Hidden Markov Model (HMM)	ECG	Accuracy: 85%; Sensitivity: 88.9%	(Wang et al., 2025)

Export to Sheets

4.3. Table 2: Applications of ML in OSA Treatment and Management (Recent Studies)

This table highlights the areas in the therapeutic pathway where machine learning is being used to treat OSA. This table shows the scope of current efforts and the potential for future advancements in this crucial area, even though the literature provided is less detailed when it comes to quantitative performance metrics for treatment efficacy.

Table 2

ML Application Area	ML Model/Approach	Key Outcomes/Impact	Citation
CPAP Compliance Prediction	(General ML models)	Predicts patient adherence to CPAP therapy	(Wang et al., 2025)
CPAP Compliance Improvement	AI-equipped CPAP ventilators, Monitoring systems, Remote reminders	Improves patient compliance through real-time monitoring and adaptive support	(Wang et al., 2025)
Personalized Treatment Regimens	DL models, ML algorithms	Prognosticates treatment outcomes, assimilates clinical data for tailored therapies	(Wang et al., 2025)
OSA Subtype Identification	AI	Enhances efficiency and accuracy of subtype identification for targeted treatment	(Wang et al., 2025)

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

ML Application Area	ML Model/Approach	Key Outcomes/Impact	Citation
Surgical/Oral Appliance Optimization	AI	Forecasts more appropriate treatment methods; assists in robotic surgery	(Wang et al., 2025)
Remote Monitoring & Management	AI on smartphones/wearables	Real-time home sleep monitoring, disease risk early warning, remote CPAP monitoring, patient self-health management	(Wang et al., 2025)
Cost-Effectiveness of Home Testing	(General AI/ML)	Reduces direct medical costs (44%), personal expenses (37%), societal costs (20%)	(Alvarez & Hornero, 2024)
Comorbidity & Prognosis Assessment	AI (e.g., analyzing brain MRI, big data)	Identifies cognitive impairment, forecasts neurological conditions, predicts cardiovascular mortality	(Wang et al., 2025)

Export to Sheets

4.4. Discussion of Statistical Methods Used in the Literature

A range of statistical techniques are used in the thorough assessment of ML/DL models in OSA research to guarantee the accuracy and dependability of results that are reported. Area Under the Receiver Operating Characteristic curve (AUC or AUROC), F1 score, Accuracy (Acc), Sensitivity (Sen), and Specificity (Spec) are examples of common performance metrics (Li et al., 2025; Manoni & Pardo, 2023). These metrics are essential for evaluating a model's overall discriminative power and its capacity to accurately classify OSA events and non-events.

Cross-validation techniques are commonly used to evaluate the generalizability of models and minimize overfitting. While leave-out cross-validation is also seen, K-fold cross-validation is a common practice, with studies frequently employing 5-fold cross-validation (Li et al., 2025; Manoni & Pardo, 2023).

To assess group differences in clinical information and PSG metrics, statistical significance testing is typically conducted using t-tests or Wilcoxon rank-sum tests, with a traditional significance threshold set at $P < 0.05$ (Li et al., 2025). Bland–Altman plots are a useful tool for assessing the consistency and agreement between true and predicted values, especially for continuous variables like the Apnea-Hypopnea Index (AHI). Strong agreement between the measured and predicted values is indicated by high percentages of data points falling within the 95% consistency interval (Li et al., 2025). Clinical features and patient demographics are summarized using descriptive statistics; continuous variables are frequently reported as medians with interquartile ranges, while categorical variables are shown as numbers with proportions (Li et al., 2025).

For a broader understanding of research trends and hotspots within the field, bibliometric methods are employed. These methods utilize tools such as VOSviewer and Citespace for quantitative analysis of literature, identifying patterns in publication growth, influential authors, institutions, and journals (Wang et al., 2025). A macro-level perspective of the field's evolution can be obtained by applying polynomial fitting analysis to publication trends in order to forecast future growth (Wang et al., 2025).

5. Research Gaps and Future Directions

Even though machine learning has made great strides in treating obstructive sleep apnea, there are still a number of important research gaps and obstacles that need to be overcome before AI's enormous potential can be developed into a common, just, and efficient clinical practice for managing OSA.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

5.1. Addressing Limitations in Current Research

The substantial demographic imbalance seen in study cohorts is a prevalent and important limitation in current research. Because most studies use data from overweight men, women, younger obese adults, people over 60, and people of different races are significantly underrepresented (Hassan et al., 2025). The generalizability of ML models to the larger patient population is severely limited by this inherent bias, which may result in differences in the effectiveness of treatment and diagnostic accuracy for marginalized groups (Hassan et al., 2025).

Additionally, a lot of research is characterized by small sample sizes, which can result in models that are less robust in real-world situations and overfit to the training data.

In addition, robust model validation techniques are frequently underutilized, which affects the external validity and dependability of results (Hassan et al., 2025). Although they offer insightful information, single-center studies cast doubt on the models' applicability to various acquisition settings and medical device types (Li et al., 2025). Standardization and data quality are major obstacles as well. The caliber and volume of training data are critical components of deep learning algorithms' performance (Manoni & Pardo, 2023). Thus, to guarantee consistency across various studies and clinical settings, thorough validation and standardization efforts are essential (Alvarez & Hornero, 2024; Hassan et al., 2025). Artifacts, missing values, and irregular sampling rates across datasets are examples of persistent problems that call for ongoing attention and sophisticated preprocessing methods. (Manoni & Pardo, 2023).

The incapacity of certain models, especially those that only use ECG and SpO2 signals, to accurately classify sleep stages or identify the precise type of respiratory events (such as differentiating between apnea and hypopnea) is another drawback. For a thorough OSA diagnosis and efficient treatment, these differences are essential (Li et al., 2025; Manoni & Pardo, 2023). Furthermore, variations in the calculation of AHI, such as the use of Total Recording Time (TRT) rather than Total Sleep Time (TST) in some studies, can cause an overestimation of apnea episodes, which can impact the clinical interpretation of severity (Manoni & Pardo, 2023).

5.2. Need for Robust Validation Strategies and Multi-Center Studies

Future studies must place a higher priority on gathering and using bigger, more varied, and multi-center datasets in order to get around the drawbacks of generalizability and bias. To guarantee that developed models are applicable across a variety of patient populations and clinical contexts, collaborative research efforts across multiple clinical centers are crucial (Li et al., 2025; Wang et al., 2025). To ensure that the model is robust and applicable in real-world situations, it is also critical to implement strict and standardized validation procedures, such as external validation on separate datasets (Alvarez & Hornero, 2024; Hassan et al., 2025). This will guarantee that AI models are dependable and efficient in standard clinical practice in addition to being accurate in controlled settings.

5.3. Challenges in Model Transparency and Patient Acceptance

Both patients and clinicians may experience severe anxiety and discomfort due to the intrinsic complexity and "black-box" nature of many AI systems, especially deep learning models (Wang et al., 2025). Lack of transparency is a major problem; establishing clinical trust, guaranteeing accountability, and promoting well-informed decision-making all depend on knowing why an AI model recommends a specific diagnosis or course of treatment. By visualizing the most pertinent signal regions that affect a model's decision, model explainability techniques like Gradient-weighted Class Activation Mapping (Grad-CAM) are becoming essential tools to address this problem (Manoni & Pardo, 2023).

According to Wang et al. (2025), strict adherence to legal and regulatory standards for data collection, storage, and analysis is also necessary due to the widespread use of sensitive medical data for AI model training. Building trust and encouraging patient acceptance of AI-driven healthcare solutions requires protecting patient privacy and guaranteeing data security (Wang et al., 2025). Strong ethical and legal frameworks must be established by precisely defining the roles of medical professionals, AI developers, and data scientists as AI becomes more

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

and more integrated into healthcare decision-making processes (Wang et al., 2025).

5.4. Opportunities for Integrating Multimodal Data and Advanced DL Architectures

A major opportunity for further development is highlighted by the success of multimodal signal fusion, which is demonstrated by the efficient combination of ECG and SpO2 signals in sophisticated deep learning models such as the Transformer (Li et al., 2025). This method illustrates the enormous potential of combining various physiological data streams to produce more thorough and precise understandings of OSA (Li et al., 2025). To create comprehensive patient profiles, future studies should investigate the integration of even more diverse data types, such as genetic, proteomic, and behavioral data. AI models may be able to identify more intricate pathophysiological mechanisms and offer more sophisticated diagnostic and treatment recommendations with the help of such extensive datasets.

Additionally, ongoing developments in deep learning architectures present encouraging opportunities to push the limits of clinical utility and performance. This includes the creation of more effective Transformer variations that are capable of processing intricate sequential data more computationally efficiently. Investigating causal AI models is also essential since they have the ability to detect actual cause-and-effect relationships rather than just correlations, which could result in more focused and successful interventions. Furthermore, dynamic treatment optimization could be achieved through the use of reinforcement learning techniques, which would enable AI systems to modify therapeutic approaches in real-time in response to ongoing patient feedback and changing physiological conditions.

5.5. Future Research Priorities in Treatment Optimization and Long-term Patient Outcomes

Although machine learning has greatly advanced the diagnosis of OSA, the field still needs to advance significantly in terms of improving treatment outcomes and directly measuring them. Future studies ought to focus on a few important areas:

First and foremost, it is imperative to measure the direct effect of ML interventions on CPAP compliance. Studies must empirically show how ML-driven interventions, like personalized remote reminders or adaptive CPAP algorithms, result in quantifiable increases in adherence rates and, crucially, in long-term clinical outcomes for patients, going beyond simply forecasting non-adherence (Wang et al., 2025).

Second, more study is needed on using machine learning (ML) to predict success rates for non-CPAP therapies and guide patient selection. This covers lifestyle changes, different surgical procedures (including robotic surgery), and dental appliances. According to Wang et al. (2025), predictive models have the potential to greatly improve the accuracy of treatment selection, guaranteeing that patients receive the best therapy for their particular condition.

Thirdly, one important avenue for the future is the creation of AI-powered closed-loop systems. These systems would combine adaptive treatment delivery for example, using neuromodulation technologies or smart CPAP devices with real-time monitoring. Such systems could result in more comfortable and successful treatment experiences by continuously optimizing therapy based on immediate patient response.

Fourth, it is imperative to move toward longitudinal research and the gathering of empirical data. Future studies should assess the long-term effects of AI-driven interventions on patient quality of life, comorbidity management, and general health outcomes in a variety of real-world clinical settings, going beyond short-term accuracy metrics. This will offer vital proof of AI's long-term therapeutic benefits for OSA. Lastly, a top priority is using ML to discover endotypes and biomarkers. In order to move towards truly precision medicine in the treatment of sleep disorders, machine learning (ML) can identify unique OSA endotypes and novel biomarkers by analyzing complex biological and clinical data. These biomarkers can then be used to guide highly customized and targeted therapeutic strategies (Hassan et al., 2025).

6. Conclusion

Recent years have seen tremendous progress in the use of machine learning and deep learning algorithms in the treatment of obstructive sleep apnea, which has completely changed methods for screening, diagnosis, and, more and more, treatment optimization. When it comes to diagnosing OSA, AI-powered models have proven to be more accurate and efficient than

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(6), 956-967 | ISSN:2251-6727

conventional diagnostic techniques. They also hold great promise for managing and customizing treatment. Numerous ML and DL models have demonstrated significant advancements in diagnostic accuracy across a wide range of data sources, including polysomnography signals, wearable device data, and smartphone data. This is supported by an extensive review of recent literature. Beyond diagnosis, artificial intelligence (AI) is starting to play a significant role in enhancing CPAP compliance through real-time monitoring systems and predictive models, customizing treatment plans, and facilitating remote patient management. However, filling a number of important research gaps is necessary to fully realize AI's potential in OSA. A major obstacle to model generalizability and equitable healthcare delivery is the widespread demographic bias in current datasets, which results in an underrepresentation of women, older adults, and diverse racial groups. Additionally, a concentrated effort towards larger, more diverse, and multi-center research initiatives is required due to the prevalence of small sample sizes and limited robust validation strategies in many studies. Strong ethical and regulatory frameworks must be developed in order to address issues with model transparency, data privacy, and the precise delineation of roles in AI-integrated healthcare. The creation of more inclusive datasets, the application of stringent validation procedures, and the ongoing investigation of sophisticated multimodal deep learning architectures ought to be the top priorities for future directions. It is imperative that research shift its focus to empirically quantifying the long-term effects of AI-driven interventions on patient quality of life, treatment adherence, and overall health outcomes. Artificial intelligence has the potential to improve the lives of millions of people worldwide by providing more reliable, equitable, and generalizable solutions for the comprehensive management of obstructive sleep apnea as long as these limitations are addressed and these priorities are prioritized.

References:

1. Alvarez, D., & Hornero, R. (2024). Artificial Intelligence in Obstructive Sleep Apnea: A Systematic Review of Diagnostic and Screening Applications. *Journal of Clinical Medicine*, *13*(12), 3512. https://doi.org/10.3390/jcm13123512
2. Hassan, A. R., Young, T., & Vicini, C. (2025). *Evaluating the application of machine learning in obstructive sleep apnea research: A narrative review* [Preprint]. medRxiv. https://doi.org/10.1101/2025.02.27.25322950
3. Li, Y., Wang, Y., Zhang, Y., Li, Y., & Zhang, Y. (2025). Deep Learning for Obstructive Sleep Apnea Detection and Severity Assessment: A Multimodal Signals Fusion Multiscale Transformer Model. *Frontiers in Neuroscience*. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC11720996/
4. Manoni, A., & Pardo, J. M. (2023). Detecting Obstructive Sleep Apnea Events and Calculating the Apnea-Hypopnea Index Using Deep Learning Models for Portable Monitors. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2023.1155900
5. Wang, Y., Li, Y., Zhang, Y., Li, Y., & Zhang, Y. (2025). Artificial intelligence in obstructive sleep apnea: A bibliometric analysis. *Frontiers in Neuroscience*. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC11930495/