







Article

Machine Learning Models to Enhance the Berlin Questionnaire Detection of Obstructive Sleep Apnea in at-Risk Patients

Luana Conte ^{1,2}, Giorgio De Nunzio ^{1,2,*}, Francesco Giombi ^{3,4}, Roberto Lupo ⁵, Caterina Arigliani ⁶, Federico Leone ⁷, Fabrizio Salamanca ^{3,7}, Cosimo Petrelli ⁸, Paola Angelelli ⁹, Luigi De Benedetto ¹⁰ and Michele Arigliani ¹¹

- ¹ Laboratory of Biomedical Physics and Environment, Department of Mathematics and Physics “E. De Giorgi”, University of Salento, Via per Monteroni, 73100 Lecce, Italy; luana.conte@unisalento.it
 - ² Laboratory of Advanced Data Analysis for Medicine (ADAM) at the Laboratory of Interdisciplinary Research Applied to Medicine (DReAM), University of Salento and Local Health Authority (ASL) Lecce, Piazza Filippo Muratore, 73100 Lecce, Italy
 - ³ Department of Biomedical Sciences, Humanitas University, Via Rita Levi Montalcini 4, 20090 Pieve Emanuele, Milan, Italy; francesco.giombi@humanitas.it (F.G.); fabrizio.salamanca@sanpiox.humanitas.it (F.S.)
 - ⁴ Otorhinolaryngology Unit, IRCCS Humanitas Research Hospital, Via Manzoni 56, 20089 Rozzano, Milan, Italy
 - ⁵ Unit of Admitting and Emergency Medicine and Surgery, “San Giuseppe da Copertino” Hospital, Local Health Authority (ASL) Lecce, Via Carmiano, 73043 Copertino, Lecce, Italy; robertolupo_2015@libero.it
 - ⁶ Unit of Anesthesia, Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro del Portillo, 00128 Rome, Italy; cate.arigliani10@gmail.com
 - ⁷ Otorhinolaryngology Unit, Snoring & OSA Research Center, “Humanitas San Pio X” Hospital, Via Francesco Nava 31, 20159 Milan, Italy; federico.leone88@gmail.com
 - ⁸ Unit of Internal Medicine, “San Giuseppe da Copertino” Hospital, Local Health Authority (ASL) Lecce, Via Carmiano, 73043 Copertino, Lecce, Italy; drpetrellicosimo@virgilio.it
 - ⁹ Department of Experimental Medicine, College ISUFI, Ecotekne, Via per Monteroni s.n., 73100 Lecce, Italy; paola.angelelli@unisalento.it
 - ¹⁰ Unit of Integrated Therapies in Otolaryngology, Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro del Portillo, 00128 Rome, Italy; luigidebenedetto@yahoo.it
 - ¹¹ Unit of Otorhinolaryngology, “Vito Fazzi” Hospital, Local Health Authority (ASL) Lecce, Piazza Filippo Muratore, 73100 Lecce, Italy; michele.mariano.arigliani@gmail.com
- * Correspondence: giorgio.denunzio@unisalento.it



Citation: Conte, L.; De Nunzio, G.; Giombi, F.; Lupo, R.; Arigliani, C.; Leone, F.; Salamanca, F.; Petrelli, C.; Angelelli, P.; De Benedetto, L.; et al. Machine Learning Models to Enhance the Berlin Questionnaire Detection of Obstructive Sleep Apnea in at-Risk Patients. *Appl. Sci.* **2024**, *14*, 5959. <https://doi.org/10.3390/app14135959>

Academic Editor: Christos Bouras

Received: 15 June 2024

Revised: 29 June 2024

Accepted: 4 July 2024

Published: 8 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The Berlin questionnaire (BQ), with its ten questions, stands out as one of the simplest and most widely implemented non-invasive screening tools for detecting individuals at a high risk of Obstructive Sleep Apnea (OSA), a still underdiagnosed syndrome characterized by the partial or complete obstruction of the upper airways during sleep. The main aim of this study was to enhance the diagnostic accuracy of the BQ through Machine Learning (ML) techniques. A ML classifier (hereafter, ML-10) was trained using the ten questions of the standard BQ. Another ML model (ML-2) was trained using a simplified variant of the BQ, BQ-2, which comprises only two questions out of the total ten. A 10-fold cross validation scheme was employed. Ground truth was provided by the Apnea–Hypopnea Index (AHI) measured by Home Sleep Apnea Testing. The model performance was determined by comparing ML-10 and ML-2 with the standard BQ in the Receiver Operating Characteristic (ROC) space and using metrics such as the Area Under the Curve (AUC), sensitivity, specificity, and accuracy. Both ML-10 and ML-2 demonstrated superior performance in predicting the risk of OSA compared to the standard BQ and were also capable of classifying OSA with two different AHI thresholds ($AHI \geq 15$, $AHI \geq 30$) that are typically used in clinical practice. This study underscores the importance of integrating ML techniques for early OSA detection, suggesting a direction for future research to improve diagnostic processes and patient outcomes in sleep medicine with minimal effort.

Keywords: obstructive sleep apnea; OSA; Berlin questionnaire; machine learning; artificial intelligence

1. Introduction

Obstructive Sleep Apnea (OSA) is a syndrome characterized by the partial or complete obstruction of the upper airways during sleep. This blockage leads to frequent awakenings to reopen the airway, disrupting sleep, causing excessive daytime sleepiness, and triggering a stress response in the body. The obstruction can also result in lowered blood oxygen levels during sleep [1], increased carbon dioxide levels, and potential damage to the cardiovascular system. OSA is also linked to a variety of health issues including stroke, high blood pressure, and even death [2–6]. These health problems are especially pronounced in individuals who are overweight and vary based on gender and age.

The occurrence of OSA, estimated to be between 9% and 38% of the Italian population, varies widely, with a higher likelihood in older adults, men, and those who are obese [1,7,8]. Among older individuals, its prevalence may rise up to 84% [1]. Despite an increase in research and medical attention towards OSA in recent years, it remains a condition that is frequently not diagnosed. This underdiagnosis can be attributed to the lack of biomarkers capable of identifying the disease [9–13].

In 2019, CER GAS (Research Center on Health and Social Care Management at the Bocconi University) released data estimating that the annual costs associated with OSA in Italy are approximately 31 billion euros. On average, the cost for each patient with severe OSA was calculated to be around 3850 euros. Despite having an estimated 12 million people with moderate to severe OSA, only about 460,000 individuals in Italy have been formally diagnosed, and merely half of these diagnosed patients have received treatment. This situation places Italy at the bottom among major countries in terms of the number of individuals diagnosed with OSA [14]. Considering that each patient is diagnosed many years after the onset of the disease, the direct and indirect healthcare costs impose a significant burden for the National Health System (NHS), which affects every single citizen. Prevention and early diagnosis are the only ways to achieve an improved quality of life and cost containment [9,15].

For the diagnosis of OSA, polysomnography (PSG) is considered the gold standard, and the severity of OSA is typically measured using the apnea–hypopnea index (AHI), with thresholds set at $\geq 5/h$ for OSA diagnosis, $\geq 15/h$ for moderate to severe OSA, and $\geq 30/h$ for severe OSA [1]. However, this method is expensive [9] and requires the patient to be monitored continuously by healthcare professionals [16], leading to a scarcity of available testing and, consequently, delays in diagnosis and an increase in the burden of disease [17–19]. Therefore, Home Sleep Apnea Testing (HSAT) is often used as an alternative. HSAT offers several advantages over traditional PSG. One of the foremost benefits of HSAT is the convenience it provides; patients can undergo testing in the familiar and comfortable setting of their own home. This not only reduces the anxiety and discomfort often associated with spending a night in an unfamiliar sleep lab environment, but also removes the logistical challenges of arranging for an overnight stay away from home. Furthermore, HSAT stands out for its cost-effectiveness. Generally costing less than laboratory-based PSG, it becomes a more accessible option for a broader range of patients, breaking down financial barriers to obtaining a diagnosis.

Recent advancements in software technologies and Machine Learning (ML) methods have significantly enhanced the development of effective predictive and diagnostic tools, becoming increasingly prevalent in various fields of medical research and applications, including for OSA [11,12,20–27]. The prediction models described in existing research primarily utilize clinical data, such as demographic information (age and gender), comorbid conditions, anthropometric measures (Body Mass Index (BMI), waist and neck circumferences), symptoms of OSA, and physiological parameters (blood pressure, overnight pulse oximetry, and lung function tests). The effectiveness of these models in predicting OSA, as indicated by an AHI $\geq 5/h$, has shown sensitivity rates ranging from 66% to 100% and specificity rates ranging from 30.8% to 76.2%. For predicting more severe OSA (AHI $\geq 15/h$), the sensitivity ranges from 60.3% to 92.7%, with the specificity ranging between 33.3% and 90.7% [24]. The variability in these models' ability to discriminate

between cases may be due to factors such as the complexity of the models, sample size, OSA prevalence, and the proportion of cases with different severities of OSA. It is noted that most OSA prediction models prioritize higher sensitivity over specificity to facilitate early diagnosis, although this approach may result in a higher rate of false positives and potentially lead to unnecessary PSG testing [24].

The Berlin questionnaire (BQ) [28] stands out as one of the simplest and most widely implemented non-invasive screening tools for diagnosing OSA, demonstrating a sensitivity of 86% and a specificity of 95% for OSA diagnosis. Originally introduced in the United States (US), the BQ consists of a concise set of questions focused on the risk factors and symptoms associated with OSA, aimed at identifying patients at high risk who might benefit from undergoing PSG to facilitate increased diagnosis rates. While the standard BQ comprises 10 questions, we previously introduced a streamlined questionnaire version by using a trained classifier [22], reducing the questionnaire to just two questions (“simplified Berlin questionnaire”, or BQ-2). This abbreviated version has been shown to achieve results comparable to the original BQ, offering an efficient means of rapidly screening high-risk OSA patients.

The main aim of this research was to enhance the sensitivity, specificity, and accuracy of the conventional BQ by incorporating ML techniques. For this purpose, we developed an ML-enhanced BQ model (ML-10) capable of predicting the risk of OSA using the BQ items as model features. Additionally, we explored a simplified version of ML-10, called ML-2, based on BQ-2 [22], to determine whether it yields comparable results. The predictive performance of these models was evaluated against the conventional BQ approach, which does not incorporate ML techniques. Furthermore, we utilized the ML-10 and the ML-2 models to identify patients with OSA at two different AHI thresholds: $\geq 15/h$, and $\geq 30/h$, thereby assessing their efficacy across a spectrum of OSA severity.

In conclusion, the integration of an ML algorithm into the conventional BQ demonstrated a significant enhancement in the ability to predict the risk of OSA across various severity thresholds. This advancement underscores the potential of ML-enhanced diagnostic tools in improving the early detection of OSA. The findings of this research validate the application of innovative ML approaches in enhancing the diagnostic processes for OSA, potentially leading to more timely and effective interventions for this widely prevalent but underdiagnosed condition.

The remaining sections of this paper are organized as follows: Section 2 details the participants and methods used in this study, including the study design, OSA diagnosis process, and ML predictive models. Section 3 presents the results of our experiments, comparing the performance of the conventional BQ, the ML-10 model, and the simplified ML-2 model. Section 4 discusses the implications of our findings, situates our work within the broader context of existing research, and outlines the limitations of our study. Finally, Section 5 concludes the paper with a summary of our contributions and suggestions for future research.

2. Participants and Methods

2.1. Design

From January to December 2023, an observational multicenter study was conducted across two Italian hospitals: the Otorhinolaryngology Unit at the “Vito Fazzi” Hospital in Lecce and the Otorhinolaryngology Head & Neck Surgery Unit at the IRCCS Humanitas Research Hospital in Milan. A total of 462 subjects, including 112 from Lecce and 350 from Milan, were screened due to suspected symptoms of OSA and underwent HSAT.

2.2. Participants

The inclusion criteria for this study were as follows: (1) participants aged ≥ 18 years and (2) who had undergone a HSAT recording. Before the HSAT examination, a baseline screening questionnaire was used to assess each participant’s basic information, medication

history, and surgical history. The participants were measured for height, weight, and BMI (kg/m^2) [28] at the time of registration.

2.3. OSA Diagnosis

All the sleep-related signals were obtained using a HSAT device (Embletta Gold Portable Testing Device[®], RemLogicE[®] Software v3.4.4 (2015), Embla System Inc., Broomfield, CO, USA, used in Lecce, and the Embletta[®] Multi Parameter Recorder-Polygraph (MPR-PG), RemLogicE[®] 3.4.1, Embla Systems, Kanata, ON, Canada, used in Milan). This study adhered to the guidelines set forth by the American Academy of Sleep Medicine (AASM) [29,30].

2.4. The Berlin Questionnaire and the Simplified Berlin Questionnaire

The BQ [28] is structured into three categories that assess the risk of sleep apnea. Patients are classified as either high risk or low risk for OSA based on their responses to individual items and their cumulative scores within these symptom categories. Category 1, comprising five items, focuses on snoring behaviors. Category 2, with three items, investigates daytime somnolence. Category 3 consists of a single item that evaluates the presence of hypertension. A positive score in the first two categories requires frequent symptom occurrence, defined as more than 3–4 times per week. In contrast, a positive score in the third category results from either a history of hypertension or a BMI greater than $30 \text{ kg}/\text{m}^2$ [28]. The overall assessment is based on the collective responses across these categories, with patients categorized as high risk for OSA if they have positive scores in two or more categories; otherwise, they are deemed low risk [28].

Our previous research showed that, among the ten questions in the standard BQ, two questions were sufficient to closely approximate the BQ output using a trained classifier. Further details are available in [22]. In summary, the first critical question assesses high blood pressure, asking, “Do you have high blood pressure?”. This inquiry is followed by one of two options regarding fatigue: “How often do you feel tired or fatigued after your sleep?” or “During your waking time, do you feel tired, fatigued or not up to par?” These questions are designed to be selected independently yet provide insightful data for OSA risk assessment. Despite their independence, we arbitrarily opted to utilize the first fatigue-related question (“How often do you feel tired or fatigued after your sleep?”). This decision was based on the observation that the models using one or the other yielded comparable results when applied independently, suggesting that favoring one fatigue-related question over the other offers no significant advantage in the context of our study.

2.5. Statistical Analysis

The baseline characteristics and BQ items for all participants, encompassing patients with confirmed OSA and those without, underwent descriptive statistical analysis. Continuous variables were summarized using the mean and standard deviation (SD), whereas categorical variables were described using frequencies and percentages. Fisher’s exact test was employed to explore the associations between two categorical variables. Additionally, the Mann–Whitney U-test was utilized to assess the statistical significance of differences between the distributions of two continuous variables among participants categorized on the basis of their AHI values, specifically those who are not at risk of OSA ($\text{AHI} < 5$) and those who are ($\text{AHI} \geq 5$), according to the threshold defined in the BQ [28]. A p -value of less than 0.05 was considered statistically significant. The scoring of the BQ and all statistical analyses, including evaluations of both qualitative and quantitative variables, were performed using Matlab software, version 2023b.

2.6. Machine Learning Predictive Value

Calculating group statistics is crucial in establishing the statistical relevance of variables within a diagnostic context, allowing for the assessment of risk factors and relationships with comorbidities. However, it is widely recognized that statistical relevance does

not equate to discriminant power, which is more critical for classification and prediction tasks. Variables that are statistically significant in a model do not necessarily guarantee superior prediction performance, and attributes deemed non-significant might be predictive. Therefore, we opted to investigate the predictive capabilities of the BQ using ML techniques. To this end, six distinct classifiers were evaluated for their suitability in the predictive task: Naive Bayes, Support Vector Machine (SVM), Decision Trees, Error-correcting Output Codes (ECOCs), Discriminant Analysis, Ensemble of decision trees, and Artificial Neural Networks (ANNs). Among these, the Ensemble of decision trees demonstrated the best performance. This model was initially trained with the ten responses from the standard BQ and then separately with only the two responses from the simplified version, BQ-2, independently, resulting in the development and evaluation of two distinct models designated as ML-10 and ML-2, respectively.

We employed a 10-fold cross-validation (CV) approach for the training and quality assessment. For both models, features were normalized to a 0–1 range using min–max normalization on the training dataset in each CV iteration, with identical normalization parameters applied to the corresponding validation set.

The Receiver Operating Characteristic (ROC) curve was used to illustrate the diagnostic capability of the models at various decision thresholds, providing a graphical representation of the trade-off between sensitivity (true positive rate) and 1-specificity (false positive rate). Initially, we identified the specific operating point in the ROC space corresponding to the conventional BQ, indicating the combined sensitivity (ability to correctly identify cases at high risk of OSA) and specificity (ability to correctly identify low or non-OSA cases) achieved without integrating ML techniques. Subsequently, we compared this point with the performance of the ML-enhanced models (both ML-10 and ML-2) at equal specificity and equal sensitivity, by vertically and horizontally adjusting them from the BQ point until the ROC curve of the ML-10 model was intersected. This approach allowed us to evaluate how ML-10 and ML-2 could enhance sensitivity while maintaining the specificity of the conventional BQ, and vice versa.

Subsequently, we extended our analysis to evaluate the ML-10 and ML-2 models across two different AHI thresholds ($AHI \geq 15$, and $AHI \geq 30$) referenced in the literature to classify OSA as moderate to severe, or severe, respectively [1]. For this purpose, the ROC curve was utilized to assess the classifier performance and to determine an “optimal” prediction threshold that maximizes accuracy. Binary classifiers were derived from this optimal operating point. Performance metrics including the Area Under the Curve (AUC), accuracy, sensitivity, and specificity were used to measure the models’ effectiveness. All computational analyses were performed using MATLAB software, version R2023b.

2.7. Ethical Considerations

The experimental protocol received approval from the Bioethics Committees of the Local Health Authorities of Lecce (Protocol Number 74, dated 22 April 2022) and Milan (Protocol Number CET Lombardia 5-PIO X-153 /23, dated 19 September 2023). Conducted in full compliance with the Helsinki Declaration for Human Research, this study ensured the ethical treatment and protection of all participants. Written informed consent was secured from each subject who agreed to partake in the study, underscoring our commitment to ethical research practices. The ethical considerations of the study were meticulously outlined in the questionnaire introduction, designed in alignment with the principles established by the Italian Data Protection Authority (DPA). Participants were informed of their right to voluntary participation, with the explicit option to withdraw from the study at any point should they choose to. The process of obtaining informed consent was structured to emphasize the voluntary nature of participation, while highlighting that the confidentiality and anonymity of all collected information would be ensured. This approach ensured that participants were fully aware of their rights and the ethical standards of the study, fostering an environment of trust and respect for individual autonomy.

3. Results

3.1. Sample Demographics

The baseline characteristics of the participants were analyzed. Overall, 460 subjects who had undergone HSAT were enrolled in this study. Of these, 141 were women, 257 were over 60 years old and 310 (67%) had an AHI ≥ 5 , therefore being considered positive. The median BMI was 27.38 kg/m² (range 13–53 kg/m²).

Clinical features were compared between patients with and without suspicion of OSA (cutoff AHI ≥ 5) and among the subgroups at three cutoffs (AHI ≥ 10 , AHI ≥ 15 , AHI ≥ 30). The results are reported in Table 1. Compared to patients without OSA, those with suspected OSA were older (<0.001 ***), more obese (<0.001 ***), sleeper (<0.001 ***), and more likely to be men.

Table 1. Baseline characteristics of the cohort. Data are expressed as mean \pm standard deviation. The *p*-value represents the comparison between the AHI ≥ 5 (positive) and AHI < 5 (negative) groups. A *p*-value < 0.05 was considered statistically significant and labeled with asterisks.

	Negative (AHI ≤ 5)	Positive (AHI ≥ 15)	Positive (AHI ≥ 30)	<i>p</i> Value (Comparison between Positive and Negative, Cutoff AHI ≥ 5)
<i>n</i>	150	181	83	
Age > 60 <i>n</i> (%)	60 (40)	117 (64)	53 (63)	<0.001 ***
Female <i>n</i> (%)	52 (35)	39 (22)	11 (13)	0.19
Height (cm)	172.4 \pm 9.8	174.1 \pm 10.1	174.7 \pm 8.06	0.94
Weight (kg)	77.6 \pm 16.3	88.3 \pm 16.9	93.4 \pm 15.9	<0.001 ***
BMI (kg/m ²)	26.0 \pm 4.6	29.2 \pm 5.6	30.7 \pm 5.6	<0.001 ***
<i>HSAT</i>				
AHI	2.5 \pm 1.4	32.8 \pm 15.7	45.7 \pm 14.5	<0.001 ***
ODI	2.6 \pm 1.8	31.3 \pm 18.7	42.2 \pm 14.8	<0.001 ***
LOS	92.6 \pm 11.1	88.1 \pm 10.3	85.5 \pm 11.8	<0.001 ***
SpO ₂ mean (%)	89.8 \pm 8.1	82.6 \pm 7.7	81.1 \pm 8.8	<0.001 ***

BMI = Body Mass Index; ODI = Oxygen Desaturation Index; LOS = Length of Stay.

3.2. Berlin Questionnaire Score and Metrics

The BQ was administered to all the participants, and the collected answers were analyzed. Before delving into the specifics of the BQ scores, we categorized the subjects into low versus high OSA risk groups based on the cutoff utilized in connection with the BQ, which considers an AHI ≥ 5 as positive (high risk) [28]. Consequently, Table 2 compares the high versus low OSA risk groups as determined by this BQ cutoff. It is important to clarify that this initial categorization uses the ‘ground truth’ based on the AHI ≥ 5 threshold, rather than the metrics derived from the questionnaire itself. The latter will be examined subsequently to assess how well the BQ scores align with the established ‘ground truth’. This approach allows for a direct comparison between the questionnaire categorization and the clinical benchmark, providing insight into the BQ’s effectiveness in identifying patients with varying levels of OSA risk.

Table 2. Differences between low vs. high OSA risk groups. Statistical significance was determined by the Mann–Whitney test and labeled with asterisks.

Items	Z-Value	Rank Sum	p-Value
<i>Snoring category</i>			
History of snoring	−3.31	32,084	<0.001 ***
Very loud snoring	−2.30	31,200	0.02
Snoring every night	−0.08	3.4004×10^4	0.92
Bothersome snoring	−1.70	3.2403×10^4	0.08
Interrupting night breathing	4.77	40,315	<0.001 ***
<i>Symptoms category</i>			
Tired upon awakening	4.25	39,622	<0.001 ***
Tired while daytime	3.29	38,386	<0.001 ***
Dozing off while driving	−5.39	29,418	<0.001 ***
Frequency of dozing off	4.33	37,442	<0.001 ***
<i>Hypertension category</i>			
High blood pressure	−7.34	2.5476×10^4	<0.001 ***

In our sample, the high-risk OSA group had a significantly larger proportion of respondents reporting frequent snoring compared to the low-risk group ($p < 0.001$). The high-risk group also reported more breathing interruptions than the low-risk subjects ($p < 0.001$). Fatigue and somnolence upon awakening and during the daytime were also significantly more present in the high-risk group compared to the low-risk group ($p < 0.001$). High blood pressure was also highly reported in subjects with a high risk of OSA, and this difference was statistically significant ($p < 0.001$). Following the administration of the BQ and the subsequent data collection, we calculated the BQ scores as prescribed by its guidelines. The outcomes of this analysis, including the accuracy, sensitivity, and specificity of the BQ, are detailed in Table 3. The ROC space is shown in Figure 1, where a red point indicates the BQ position. Notably, the classic BQ is positioned in the upper right corner of the evaluation plot, approaching the point (1, 1), which represents the maximum sensitivity and minimum specificity. This characteristic reflects the aim of the BQ to function as a high-sensitivity screening tool, intended to minimize false negatives even at the cost of accepting a higher number of false positives.

Table 3. Comparing the performance of the standard BQ, the BQ enhanced through Machine Learning (ML-10) and the simplified BQ (i.e., BQ-2) enhanced through ML (ML-2), using metrics such as the AUC, Sensitivity, and Specificity. Note that the Sensitivity and Specificity for ML-10 and ML-2 were obtained with the procedure described in the text (i.e., by preserving BQ Specificity or Sensitivity, respectively), at specific points A10/A2 and B10/B2 (see Figure 1).

	BQ	ML-10	ML-2
AUC	(not applicable)	86%	78%
Sensitivity	82%	93% (Figure 1, point B10)	88% (Figure 1, point B2)
Specificity	53%	73% (Figure 1, point A10)	54% (Figure 1, point A2)

AUC = Area Under the (ROC) Curve.

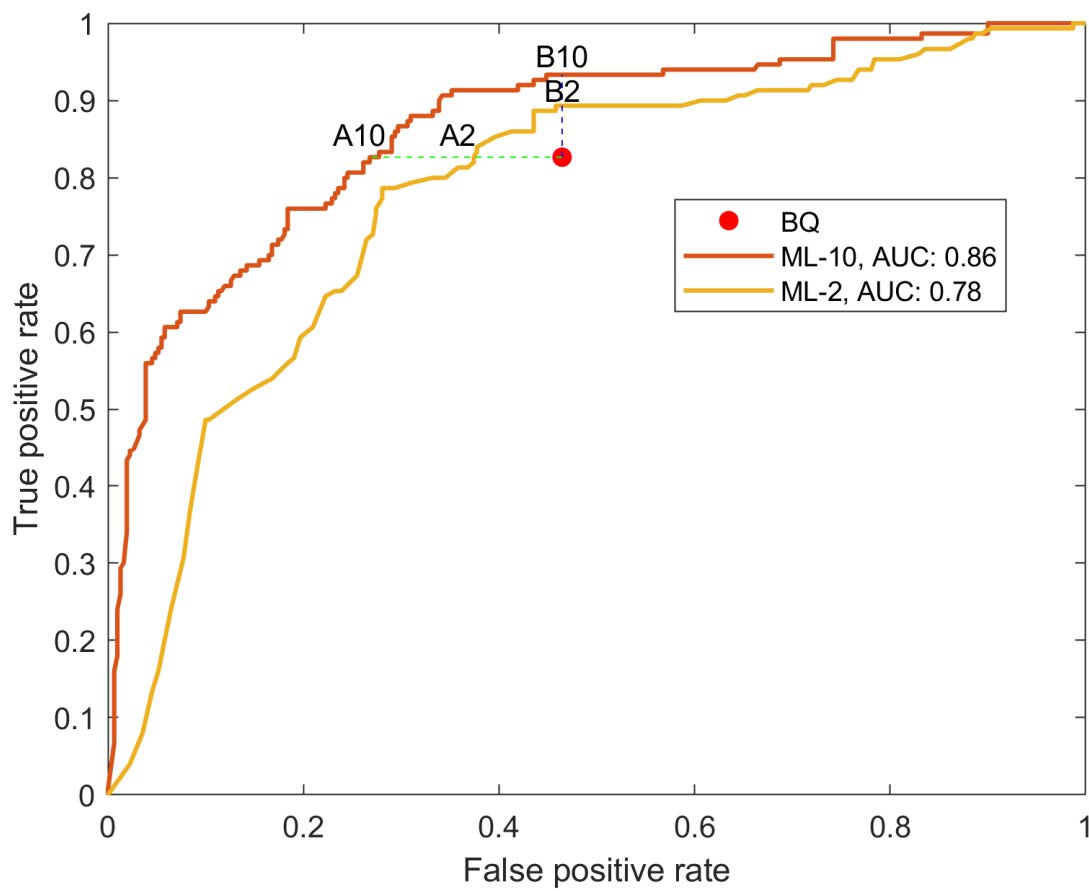


Figure 1. ROC comparison among the standard BQ, the BQ enhanced through Machine Learning (ML-10), and the simplified BQ (BQ-2, [22]) enhanced through Machine Learning (ML-2). $AHI \geq 5$ was used as the cutoff. Although some statistical software also associates AUC values with classifiers with binary output (when just one point exists in the ROC space), we preferred to neglect this feature and only drew the particular BQ working point (red point in the plot) that gives the BQ performance in terms of fixed sensitivity and specificity. A2 and A10 are the intersection points of a horizontal line from BQ, with the ML-2 and ML-10 ROC curves. B2 and B10 are the same for a vertical line (see text for details).

3.3. The ML-10 Model

To determine whether the ML-10 and the reduced version ML-2 outperform the traditional BQ in predicting patients at a high versus low risk of OSA, we conducted a comparative analysis using the same threshold used in the standard BQ ($AHI \geq 5$). By employing the same $AHI \geq 5$ threshold across all models, we ensured a consistent basis for comparison, enabling a clear understanding of the potential advantages offered by integrating ML techniques into the traditional BQ assessment. Figure 1 presents the ROC space comparing the three models (BQ, ML-10, and ML-2), and Table 3 reports the calculated metrics. Consider Figure 1. By maintaining the BQ level of sensitivity (through a horizontal displacement in the ROC space from BQ coordinates to point A10), the ML-10 model showcased a remarkable specificity of 73%, significantly outperforming the BQ (53%). This improvement indicates the ML-10 model's enhanced capability to correctly identify individuals without OSA at a fixed sensitivity, thereby reducing the incidence of false positives. Conversely, when aligning with the BQ specificity level (via a vertical displacement from BQ to point B10), the ML-10 model demonstrated a sensitivity of 93%, a substantial increase from the BQ (82%). This indicates the ML-10 model's superior ability to accurately detect individuals at a high risk of OSA at a fixed specificity, lowering the risk of overlooking affected patients.

Additionally, ML-2 also showed (smaller) improvements when compared to the conventional BQ, and (as expected) its figures of merit were smaller than those of ML-10 (points A2 and B2 in Figure 1).

Recognizing the importance of a nuanced clinical evaluation, we expanded our analysis to investigate how well the ML-10 and ML-2 models distinguish between patients across different levels of OSA severity. This step involved utilizing two AHI thresholds ($\text{AHI} \geq 15$, and $\text{AHI} \geq 30$) commonly referenced in the literature to categorize OSA severity as moderate to severe, and severe, respectively [1]. The outcomes of this comprehensive evaluation are presented in Figure 2 and Table 4. Before all, we realized that the AUCs for ML-10 (0.85 and 0.88) were (slightly) better than the AUCs for ML-2 (0.82 and 0.87), being $\text{AHI} \geq 15$ and $\text{AHI} \geq 30$, respectively. Then, by arbitrarily selecting the ones that yield the highest accuracy as the optimal thresholds for the classifier output (therefore precisely defining working points in the ROC space), the performance of the ML-10 model consistently remained high across the two AHI thresholds and larger than that of ML-2, except in one case where, compared to the ML-10 model, ML-2 had higher sensitivity when assessing moderate to severe OSA with a cutoff of $\text{AHI} \geq 15$ (88% vs. 70%, at similar Accuracy): this is not particularly relevant because we also remark that the corresponding specificity is lower than ML-10's.

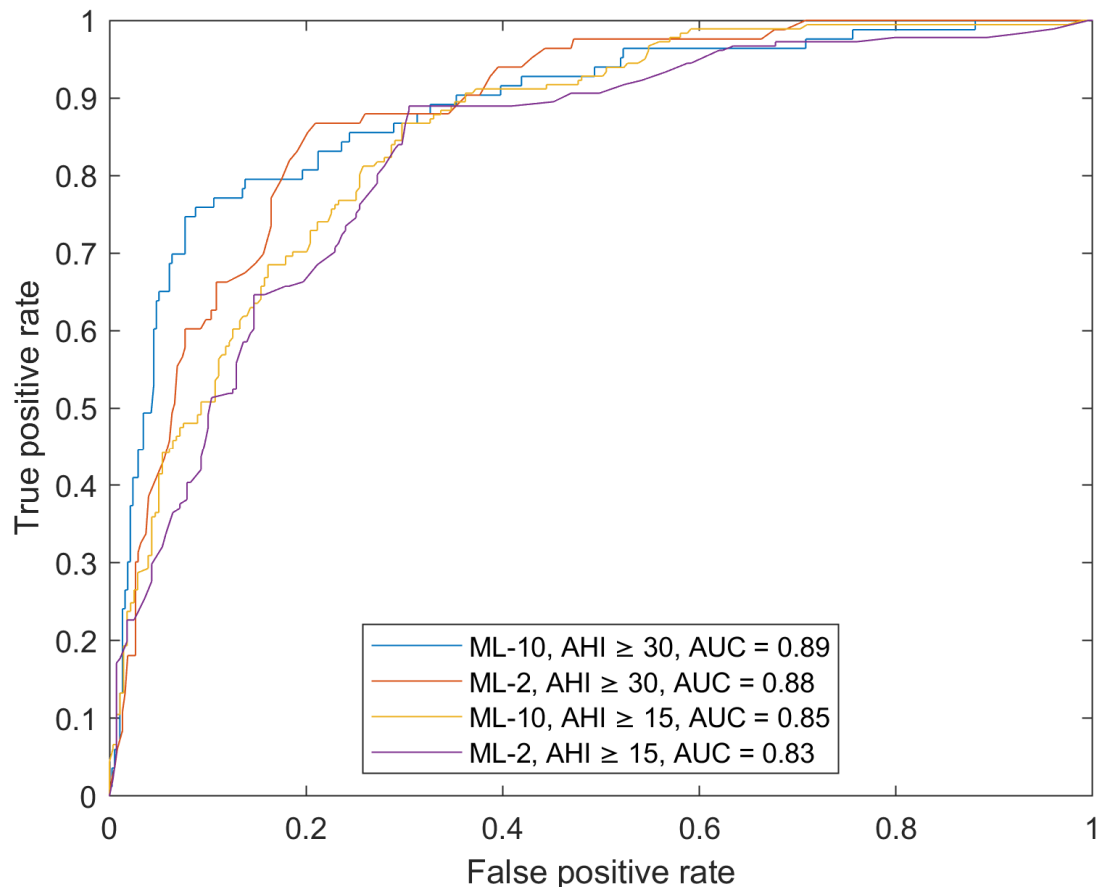


Figure 2. ROC curve comparison between the ML-10 and the ML-2 models in assessing OSA severity at two cutoffs ($\text{AHI} \geq 15$, $\text{AHI} \geq 30$).

Table 4. Metrics comparison among ML-10 model and BQ-2 model in assessing OSA severity with the two cutoffs ($AHI \geq 15$, $AHI \geq 30$).

	$AHI \geq 15$	$AHI \geq 30$
ML-10		
AUC	85%	89%
Sensitivity	70%	69%
Specificity	81%	93%
Accuracy	77%	89%
ML-2		
AUC	83%	88%
Sensitivity	88%	60%
Specificity	69%	92%
Accuracy	76%	86%

4. Discussion

OSA is increasingly recognized as a significant concern within global health and economic contexts, underlining the importance of its early detection and diagnosis in the realm of preventive medicine [1,17,31]. The prompt identification of OSA is essential for initiating timely interventions, which can mitigate a broad range of associated health risks and enhance patient outcomes. Given that the standard diagnostic test for OSA, namely in-laboratory PSG, is expensive and often subject to long wait times due to high demand, there is a clinical imperative to identify the key factors and develop a simple yet reliable tool for estimating the OSA risk [17,18]. In general, BQ has an expectedly high sensitivity, as this tool has been developed for the identification of patients at a high risk of OSA in primary care settings. Despite this advantage, the BQ's low specificity and consequent high misclassification rate reveal its limited discriminatory capability, rendering its utility comparable to subjective clinical judgments [30,32]. In the quest for a straightforward questionnaire to ascertain OSA risk, clinicians are demanding enhancements to existing tools. Arunsurat et al. [33] posited that with certain modifications, the BQ could serve effectively as an OSA screening instrument. Furthermore, Stelmach-Mardas et al. [34] added to the growing body of evidence indicating the BQ's inadequacy in distinguishing between high- and low-risk patients, suggesting the need for the development of alternative protocols to heighten the diagnostic precision for such individuals.

In this research, we sought to advance the capabilities of the traditional BQ through the integration of ML techniques. Our research integrates ML models with the standard BQ to harness Artificial Intelligence capabilities for analyzing patterns and correlations in data that might not be immediately apparent to human evaluators. This method facilitates a more detailed assessment of risk factors, potentially identifying the subtle signs of OSA risk overlooked by conventional approaches. To determine whether our ML-10 and the simplified two-item version ML-2 outperform traditional BQ in predicting patients at a high versus low risk of OSA, we conducted a comparative analysis using the established threshold used in the standard BQ ($AHI \geq 5$) and by comparing points in the ROC space. The findings underscore the efficacy in terms of sensitivity and specificity of the ML-10 model when contrasted with conventional BQ. A sensitivity of 93% at the same specificity as conventional BQ indicates that the model can correctly identify 93% of individuals (at low or high risk of OSA), operating with the same TN-rate. This result is significant as it demonstrates that, while maintaining the same rate of false alarms ($1 - \text{Specificity}$), the ML-10 model is more effective in detecting OSA risk cases compared to conventional BQ. On the other hand, a specificity of 73% at the same sensitivity as conventional BQ emphasizes that the ML-10 model reduces the number of false positives (healthy individuals erroneously identified as at risk of OSA) compared to the conventional BQ, while still correctly detecting 82% of true positives. In this way, the ML-10 model shows excellent performance in identifying non-risk cases, surpassing conventional BQ.

These results indicate that the ML-10 model surpasses conventional BQ both in terms of sensitivity (when specificity is maintained) and specificity (when sensitivity is maintained). This implies that, depending on clinical or screening needs, the ML-10 model can be adjusted to optimize the ability to detect OSA risk cases (by maximizing sensitivity) or the ability to reduce false positives (by maximizing specificity), offering a more flexible and accurate approach in the diagnosis of OSA.

In the comparative evaluation between conventional BQ and the classifier based on its simplified version, the results indicate that ML-2, despite the significant reduction in the number of questions to only two, slightly outperforms BQ in terms of sensitivity and specificity (fixing one of the two variables at the BQ value). Additionally, the use of ML-2 offers the flexibility needed to adjust the operating point on the ROC curve depending on the specific needs of clinical or screening applications, thus providing a potential advantage in terms of customizing the diagnostic approach.

After assessing the ML-10 and ML-2 performance against traditional BQ using a single cutoff, we expanded our analysis to include two clinically relevant AHI cutoffs. This step involved utilizing two AHI thresholds ($AHI \geq 15$, and $AHI \geq 30$) commonly used in the literature to categorize the OSA severity as moderate to severe, and severe, respectively [1]. The decision to employ these specific AHI thresholds is rooted in their widespread acceptance and use in clinical practice and research for defining the severity of OSA. Such a differentiated approach allows for a more detailed assessment of the models' performance, providing insights into their predictive capabilities across a spectrum of OSA severity. This is particularly relevant for clinicians and healthcare providers seeking to tailor interventions and management strategies based on the severity of the condition. By choosing the optimal threshold for maximum accuracy, the ML-10 model performance consistently demonstrated its strength at both AHI thresholds.

These results highlight the potential for a more streamlined and efficient screening process. By examining whether a simplified model can retain or surpass the full BQ predictive accuracy, this study suggests the possibility of more accessible and less cumbersome OSA screening approaches. This is especially pertinent in primary care environments or areas with limited access to specialized sleep medicine services, where a rapid and dependable screening tool could significantly improve the early detection of individuals at risk of OSA. However, we should consider that using only two questions likely makes the test sensitive but not specific, as various diseases could present with the same broad symptoms.

The present study is subject to several limitations that merit consideration. Firstly, the participant cohort was drawn exclusively from two hospitals in Italy, limiting the data set representativeness of the broader population. Consequently, the predictive model developed herein might not possess widespread generalizability, potentially limiting its applicability to populations beyond the initial study setting or to diverse ethnic groups [24,35]. Secondly, this observational study did not account for undiagnosed medical conditions commonly associated with OSA, such as neurological, cardiovascular, and pulmonary disorders. The absence of these variables could impact the model's predictive accuracy. Furthermore, our model lacked detailed anthropometric imaging or measurements, which might have restricted its ability to identify disease-specific causes of OSA accurately.

In light of these limitations, there is a clear need for further research to enhance the model's robustness and applicability. To this end, we are planning a prospective clinical trial aimed at evaluating ML-10 and ML-2 across a more representative sample of the general population. This forthcoming trial is expected to address the current study limitations by incorporating a broader range of demographic and clinical variables, thereby improving the model's predictive performance and generalizability.

5. Conclusions

Given the substantial proportion of individuals still undiagnosed with OSA, coupled with the current absence of definitive diagnostic biomarkers for the condition, there is a pressing need for improved screening methodologies. The BQ, when enhanced with ML

techniques, stands out as a significant advancement in this regard. This study discovered that the ML-10 model was particularly effective in identifying individuals at risk of OSA with greater accuracy than the traditional BQ. By integrating ML techniques, we achieved a notable improvement in sensitivity and specificity, highlighting the potential of ML to refine diagnostic processes. This suggests that the ML-10 model can more effectively distinguish between high-risk and low-risk individuals, thereby reducing the likelihood of false positives and negatives. Furthermore, ML-2, with its reduced question set, also showcased its utility by maintaining slightly better diagnostic accuracy than the full BQ while offering a more streamlined and accessible screening tool. This adaptation could facilitate wider screening efforts, particularly in primary care settings or areas with limited access to sleep medicine specialists. Additionally, the flexibility of the classifier allows for adjustments across different operating points, enabling the selection of an optimal threshold that best balances sensitivity and specificity for the targeted population. This adaptability is crucial in tailoring the screening process to diverse clinical environments and patient needs, optimizing the early detection and management of OSA.

Moreover, the application of the ML-10 model extends beyond the commonly used AHI threshold of the standard BQ ($AHI \geq 5$), demonstrating its utility across other clinically relevant AHI thresholds, specifically ≥ 15 and ≥ 30 , which are frequently used in the literature to categorize the severity of OSA as moderate to severe, and severe, respectively. This versatility underscores the model's ability to adapt to varying clinical requirements, offering a nuanced approach to diagnosing OSA across its spectrum. Such adaptability ensures that the ML-10 model is not only a tool for preliminary screening but also a significant asset in stratifying OSA severity, thus enhancing the precision of diagnostic decisions and subsequent management plans.

By leveraging these insights, healthcare professionals can better stratify individuals based on their risk levels, paving the way for more tailored diagnostic and management strategies for sleep apnea. ML-10 embodies the potential to transform the approach to diagnosing OSA, offering a more individualized assessment of risk. Looking forward, the insights gained from this research could serve as a foundation for further innovations in the field, ultimately leading to earlier detection, improved patient outcomes, and a reduction in the healthcare burden associated with OSA. These results can be achieved with minimal effort, because no modification to the BQ itself is necessary. The approach does not necessitate developing new questions or methodologies; instead, it leverages AI techniques to optimize an existing, widely used tool. This means that new screenings could achieve greater accuracy, and previously administered questionnaires could be easily re-examined using the ML-10 model. Consequently, more cases of OSAS could be identified, and more healthy individuals could be correctly reassured. In the end, this study underscores the value of combining traditional clinical assessment tools with cutting-edge technology to address complex health challenges, marking a significant stride towards the future of personalized medicine in sleep health.

Author Contributions: Conceptualization, L.C., G.D.N., M.A., F.G., F.L., R.L., F.S. and C.P.; methodology, G.D.N., L.C., P.A., L.D.B. and C.A.; software, L.C. and G.D.N.; validation, G.D.N., L.C., M.A., L.D.B. and P.A.; formal analysis, G.D.N. and L.C.; investigation, F.G., F.L., F.S., L.D.B. and M.A.; data curation, L.C., G.D.N., F.G., F.L., F.S., L.D.B. and M.A.; writing—original draft preparation, L.C., G.D.N., C.A., C.P. and L.D.B.; writing—review and editing, L.C., G.D.N., M.A., L.D.B., F.G., F.L., F.S., and R.L.; visualization, L.C., G.D.N., M.A., F.G., F.L., F.S., R.L., C.A., C.P., L.D.B. and P.A.; supervision, G.D.N., L.C., M.A. and L.D.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Bioethics Committees of the Local Health Authorities of Lecce (Protocol Number 74, dated 22 April 2022) and of Milan (Protocol Number CET Lombardia 5-PIO X-153 /23, dated 19 September 2023). See also Section 2.7.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Senaratna, C.V.; Perret, J.L.; Lodge, C.J.; Lowe, A.J.; Campbell, B.E.; Matheson, M.C.; Hamilton, G.S.; Dharmage, S.C. Prevalence of Obstructive Sleep Apnea in the General Population: A Systematic Review. *Sleep Med. Rev.* **2017**, *34*, 70–81. [[CrossRef](#)] [[PubMed](#)]
2. Marin, J.M.; Carrizo, S.J.; Vicente, E.; Agusti, A.G. Long-Term Cardiovascular Outcomes in Men with Obstructive Sleep Apnoea-Hypopnoea with or without Treatment with Continuous Positive Airway Pressure: An Observational Study. *Lancet* **2005**, *365*, 1046–1053. [[CrossRef](#)] [[PubMed](#)]
3. Dyken, M.E.; Im, K. Bin Obstructive Sleep Apnea and Stroke. *Chest* **2009**, *136*, 1668–1677. [[CrossRef](#)] [[PubMed](#)]
4. Toraldo, D.M.; De Benedetto, M.; Conte, L.; De Nuccio, F. Statins May Prevent Atherosclerotic Disease in OSA Patients without Co-Morbidities? *Curr. Vasc. Pharmacol.* **2017**, *15*, 5–9. [[CrossRef](#)] [[PubMed](#)]
5. Garbarino, S.; Scoditti, E.; Lanteri, P.; Conte, L.; Magnavita, N.; Toraldo, D.M. Obstructive Sleep Apnea with or without Excessive Daytime Sleepiness: Clinical and Experimental Data-Driven Phenotyping. *Front. Neurol.* **2018**, *9*, 505. [[CrossRef](#)] [[PubMed](#)]
6. Vicini, C.; Cannavici, A.; Cioccioloni, E.; Meccariello, G.; Cammaroto, G.; Gobbi, R.; Sanna, A.; Toraldo, D.M.; Bonetti, G.A.; Passali, F.M.; et al. Treatment. In *Obstructive Sleep Apnea*; Springer International Publishing: Cham, Switzerland, 2023; pp. 85–104.
7. Benjafield, A.V.; Ayas, N.T.; Eastwood, P.R.; Heinzer, R.; Ip, M.S.M.; Morrell, M.J.; Nunez, C.M.; Patel, S.R.; Penzel, T.; Pépin, J.-L.; et al. Estimation of the Global Prevalence and Burden of Obstructive Sleep Apnoea: A Literature-Based Analysis. *Lancet Respir. Med.* **2019**, *7*, 687–698. [[CrossRef](#)] [[PubMed](#)]
8. Fietze, I.; Laharnar, N.; Obst, A.; Ewert, R.; Felix, S.B.; Garcia, C.; Gläser, S.; Glos, M.; Schmidt, C.O.; Stubbe, B.; et al. Prevalence and Association Analysis of Obstructive Sleep Apnea with Gender and Age Differences—Results of SHIP-Trend. *J. Sleep Res.* **2019**, *28*, e12770. [[CrossRef](#)] [[PubMed](#)]
9. Toraldo, D.M.; Passali, D.; Sanna, A.; De Nuccio, F.; Conte, L.; De Benedetto, M. Cost-Effectiveness Strategies in OSAS Management: A Short Review. *Acta Otorhinolaryngol. Ital.* **2017**, *37*, 447–453. [[CrossRef](#)] [[PubMed](#)]
10. Conte, L.; Greco, M.; Toraldo, D.M.; Arigliani, M.; Maffia, M.; De Benedetto, M. A Review of the “OMICS” for Management of Patients with Obstructive Sleep Apnoea. *Acta Otorhinolaryngol. Ital.* **2020**, *40*, 164–172. [[CrossRef](#)] [[PubMed](#)]
11. Arigliani, M.; Toraldo, D.M.; Montevicchi, F.; Conte, L.; Galasso, L.; De Rosa, F.; Lattante, C.; Ciavolino, E.; Arigliani, C.; Palumbo, A.; et al. A New Technological Advancement of the Drug-Induced Sleep Endoscopy (DISE) Procedure: The “All in One Glance” Strategy. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4261. [[CrossRef](#)] [[PubMed](#)]
12. Arigliani, M.; Toraldo, D.M.; Ciavolino, E.; Lattante, C.; Conte, L.; Arima, S.; Arigliani, C.; Palumbo, A.; De Benedetto, M. The Use of Middle Latency Auditory Evoked Potentials (MLAEP) as Methodology for Evaluating Sedation Level in Propofol-Drug Induced Sleep Endoscopy (DISE) Procedure. *Int. J. Environ. Res. Public Health* **2021**, *18*, 2070. [[CrossRef](#)] [[PubMed](#)]
13. Arigliani, C.; Arigliani, M.; Ciavolino, E.; Conte, L.; Toraldo, D.M.; Passariello, S.; Arima, S.; Palumbo, A.; De Benedetto, M. Polygraphic Findings in Simplified Barbed Reposition Pharyngoplasty (BRP) as a Treatment for OSA Patients. *J. Interdiscip. Res. Appl. Med.* **2021**, *5*, 19–26. [[CrossRef](#)]
14. Armeni, P.; Borsoi, L.; Costa, F.; Donin, G.; Gupta, A. *Cost-of-Illness Study of Obstructive Sleep Apnea Syndrome (OSAS) in Italy*; Bocconi University: Milan, Italy, 2019.
15. Toraldo, D.M.; Toraldo, S.; Conte, L. The Clinical Use of Stem Cell Research in Chronic Obstructive Pulmonary Disease: A Critical Analysis of Current Policies. *J. Clin. Med. Res.* **2018**, *10*, 671–678. [[CrossRef](#)] [[PubMed](#)]
16. Gottlieb, D.J.; Punjabi, N.M. Diagnosis and Management of Obstructive Sleep Apnea. *JAMA* **2020**, *323*, 1389. [[CrossRef](#)] [[PubMed](#)]
17. Knauert, M.; Naik, S.; Gillespie, M.B.; Kryger, M. Clinical Consequences and Economic Costs of Untreated Obstructive Sleep Apnea Syndrome. *World J. Otorhinolaryngol. Head Neck Surg.* **2015**, *1*, 17–27. [[CrossRef](#)] [[PubMed](#)]
18. Stewart, S.A.; Skomro, R.; Reid, J.; Penz, E.; Fenton, M.; Gjevre, J.; Cotton, D. Improvement in Obstructive Sleep Apnea Diagnosis and Management Wait Times: A Retrospective Analysis of a Home Management Pathway for Obstructive Sleep Apnea. *Can. Respir. J.* **2015**, *22*, 167–170. [[CrossRef](#)] [[PubMed](#)]
19. Kapur, V.K.; Auckley, D.H.; Chowdhuri, S.; Kuhlmann, D.C.; Mehra, R.; Ramar, K.; Harrod, C.G. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. *J. Clin. Sleep Med.* **2017**, *13*, 479–504. [[CrossRef](#)]
20. Davenport, T.; Kalakota, R. The Potential for Artificial Intelligence in Healthcare. *Futur. Healthc. J.* **2019**, *6*, 94–98. [[CrossRef](#)] [[PubMed](#)]
21. Yu, K.-H.; Beam, A.L.; Kohane, I.S. Artificial Intelligence in Healthcare. *Nat. Biomed. Eng.* **2018**, *2*, 719–731. [[CrossRef](#)] [[PubMed](#)]
22. De Nunzio, G.; Conte, L.; Lupo, R.; Vitale, E.; Calabrò, A.; Ercolani, M.; Carvello, M.; Arigliani, M.; Toraldo, D.M.; De Benedetto, L. A New Berlin Questionnaire Simplified by Machine Learning Techniques in a Population of Italian Healthcare Workers to Highlight the Suspicion of Obstructive Sleep Apnea. *Front. Med.* **2022**, *9*, 866822. [[CrossRef](#)] [[PubMed](#)]

23. Kuan, Y.-C.; Hong, C.-T.; Chen, P.-C.; Liu, W.-T.; Chung, C.-C. Logistic Regression and Artificial Neural Network-Based Simple Predicting Models for Obstructive Sleep Apnea by Age, Sex, and Body Mass Index. *Math. Biosci. Eng.* **2022**, *19*, 11409–11421. [[CrossRef](#)] [[PubMed](#)]
24. Huang, W.-C.; Lee, P.-L.; Liu, Y.-T.; Chiang, A.A.; Lai, F. Support Vector Machine Prediction of Obstructive Sleep Apnea in a Large-Scale Chinese Clinical Sample. *Sleep* **2020**, *43*, zsz295. [[CrossRef](#)] [[PubMed](#)]
25. Kirby, S.D.; Eng, P.; Danter, W.; George, C.F.; Francovic, T.; Ruby, R.R.; Ferguson, K.A. Neural Network Prediction of Obstructive Sleep Apnea from Clinical Criteria. *Chest* **1999**, *116*, 409–415. [[CrossRef](#)] [[PubMed](#)]
26. Zerah-Lancner, F.; Lofaso, F.; D'Ortho, M.P.; Delclaux, C.; Goldenberg, F.; Coste, A.; Housset, B.; Harf, A. Predictive Value of Pulmonary Function Parameters for Sleep Apnea Syndrome. *Am. J. Respir. Crit. Care Med.* **2000**, *162*, 2208–2212. [[CrossRef](#)] [[PubMed](#)]
27. Zou, J.; Guan, J.; Yi, H.; Meng, L.; Xiong, Y.; Tang, X.; Su, K.; Yin, S. An Effective Model for Screening Obstructive Sleep Apnea: A Large-Scale Diagnostic Study. *PLoS ONE* **2013**, *8*, e80704. [[CrossRef](#)] [[PubMed](#)]
28. Netzer, N.C.; Stoohs, R.A.; Netzer, C.M.; Clark, K.; Strohl, K.P. Using the Berlin Questionnaire to Identify Patients at Risk for the Sleep Apnea Syndrome. *Ann. Intern. Med.* **1999**, *131*, 485. [[CrossRef](#)] [[PubMed](#)]
29. Oku, Y.; Okada, M. Periodic Breathing and Dysphagia Associated with a Localized Lateral Medullary Infarction. *Respirology* **2008**, *13*, 608–610. [[CrossRef](#)] [[PubMed](#)]
30. Sert Kuniyoshi, F.H.; Zellmer, M.R.; Calvin, A.D.; Lopez-Jimenez, F.; Albuquerque, F.N.; van der Walt, C.; Trombetta, I.C.; Caples, S.M.; Shamsuzzaman, A.S.; Bukartyk, J.; et al. Diagnostic Accuracy of the Berlin Questionnaire in Detecting Sleep-Disordered Breathing in Patients with a Recent Myocardial Infarction. *Chest* **2011**, *140*, 1192–1197. [[CrossRef](#)] [[PubMed](#)]
31. Salman, L.A.; Shulman, R.; Cohen, J.B. Obstructive Sleep Apnea, Hypertension, and Cardiovascular Risk: Epidemiology, Pathophysiology, and Management. *Curr. Cardiol. Rep.* **2020**, *22*, 6. [[CrossRef](#)] [[PubMed](#)]
32. Cowan, D.C.; Allardice, G.; Macfarlane, D.; Ramsay, D.; Ambler, H.; Banham, S.; Livingston, E.; Carlin, C. Predicting Sleep Disordered Breathing in Outpatients with Suspected OSA. *BMJ Open* **2014**, *4*, e004519. [[CrossRef](#)] [[PubMed](#)]
33. Arunsurat, I.; Luengyosuechakul, S.; Pratephoung, K.; Siripaupradist, P.; Khemtong, S.; Jamcharoensup, K.; Thanapatkaiporn, N.; Limpawattana, P.; Laohasiriwong, S.; Pinitsoontorn, S.; et al. Simplified Berlin Questionnaire for Screening of High Risk for Obstructive Sleep Apnea among Thai Male Healthcare Workers. *J. UOEH* **2016**, *38*, 199–206. [[CrossRef](#)] [[PubMed](#)]
34. Stelmach-Mardas, M.; Iqbal, K.; Mardas, M.; Kostrzewska, M.; Piorunek, T. Clinical Utility of Berlin Questionnaire in Comparison to Polysomnography in Patients with Obstructive Sleep Apnea. *Adv. Exp. Med. Biol.* **2017**, *980*, 51–57. [[PubMed](#)]
35. Kim, Y.J.; Jeon, J.S.; Cho, S.-E.; Kim, K.G.; Kang, S.-G. Prediction Models for Obstructive Sleep Apnea in Korean Adults Using Machine Learning Techniques. *Diagnostics* **2021**, *11*, 612. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.