Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Desaturation Depth Index and Excessive Daytime Sleepiness in Severe Obstructive Sleep Apnea: A Retrospective Polysomnography Study with Exploratory Sex-Stratified Analyses

barış çil

drbariscil@hotmail.com

Mardin Training and Research Hospital

Mehmet Sinan Bodur

Diyarbakır Gazi Yaşargil Training and Research Hospital

Mehmet Kabak

Mardin Artuklu University

Research Article

Keywords: obstructive sleep apnea, excessive daytime sleepiness, desaturation depth index, polysomnography, nocturnal hypoxemia, sex differences, phenotype

Posted Date: October 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10785067/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Excessive daytime sleepiness (EDS) is absent in a substantial proportion of patients with severe obstructive sleep apnea (OSA), suggesting that apnea–hypopnea frequency alone does not capture symptom burden. We evaluated the Desaturation Depth Index (DDI), a pragmatic summary-derived metric calculated as mean nocturnal SpO_2 minus nadir SpO_2, and asked whether it provides information about EDS beyond conventional clinical and polysomnographic measures and whether this association differs by sex.

Methods

This retrospective, single-centre, cross-sectional study included 290 adults with severe OSA (AHI ≥ 30 events/h) identified from 694 consecutive full-night polysomnography records obtained between January 2024 and January 2025. EDS was defined as Epworth Sleepiness Scale (ESS) > 10. Analyses included group comparisons, Spearman correlations, ROC curves, multivariable logistic regression, an exploratory nested-model comparison with and without DDI, bootstrap internal validation, and exploratory sex-stratified analyses. The dataset coding was 0 = male and 1 = female. N2/REM was analysed only when REM% was > 0, and oxygen-band analyses excluded records with missing or internally implausible band totals.

Results

Of 290 patients, 163 (56.2%) had EDS. DDI was higher in EDS than in non-sleepy patients (median 18.0 [IQR 13.0–23.5] vs 13.0 [9.5–17.8] percentage points; $p < 0.001$), correlated with ESS ($\rho = 0.358$; $p < 0.001$), and showed moderate standalone discrimination (AUC 0.677, 95% bootstrap CI 0.616–0.742). In the adjusted model, DDI remained independently associated with EDS (OR 1.626 per SD, $p = 0.005$), together with AHI, neck circumference, and female sex. Adding DDI to the otherwise identical clinical model increased apparent AUC from 0.787 to 0.810 (ΔAUC 0.023) and improved model fit (likelihood-ratio $\chi^2=8.66$, $p = 0.003$); the optimism-corrected AUC of the full model was 0.793. EDS prevalence was 71.7% in women and 40.7% in men. DDI discrimination was numerically higher in men than women (AUC 0.730 vs 0.686), but neither the DDI×sex interaction ($p = 0.124$) nor the bootstrap AUC difference ($p = 0.507$) was significant.

Conclusions

The principal new finding is that DDI, despite being calculated from only two routinely reported oxygenation values, retained an independent association with EDS and provided modest additional

information when added to a model containing AHI and clinical covariates. DDI did not outperform established measures when used alone; its potential value is therefore pragmatic and complementary rather than substitutive. These findings identify DDI as a hypothesis-generating marker of the sleepy severe-OSA phenotype that warrants prospective external validation before clinical thresholds or sex-specific applications are considered.

1. Introduction

Obstructive sleep apnea (OSA) is a common sleep-related breathing disorder with a substantial global burden [1]. Excessive daytime sleepiness (EDS) is clinically important because it is associated with impaired functioning and increased accident risk [2]. Nevertheless, EDS is not universal even among patients with severe OSA, and symptom-based OSA phenotypes remain heterogeneous [3–5]. This discordance indicates that the apnea–hypopnea index (AHI), although central to diagnosis and severity classification, does not capture all physiological dimensions relevant to daytime symptoms.

Nocturnal oxygenation provides additional information beyond event frequency. Event-specific hypoxic burden has been associated with cardiovascular outcomes [6], and multiple approaches have been proposed for characterising oximetry patterns in OSA [7]. Desaturation-related indices have also been linked to sleepiness [8]. However, many routinely available polysomnography reports contain only summary oxygenation variables rather than event-level hypoxic-burden calculations. We therefore derived an exploratory Desaturation Depth Index (DDI) from two universally available summary values: mean nocturnal SpO_2 and nadir SpO_2. DDI was defined as the arithmetic difference between these values and was intended as a pragmatic descriptor of the maximum fall from the patient's mean nocturnal oxygenation level, not as a replacement for event-specific hypoxic burden.

Sex-related differences in OSA presentation are well recognised [9–11]. These differences motivated an exploratory assessment of whether DDI–sleepiness associations varied between women and men. The primary aim was to determine whether DDI discriminated EDS from non-sleepy severe OSA and retained an independent association with EDS after adjustment for established clinical variables. Secondary aims were to compare DDI performance with conventional polysomnographic and anthropometric measures, examine a clinically interpretable DDI cut-off, and explore sex-stratified patterns using formal interaction and AUC-difference testing.

2. Methods

2.1 Study design and population

This was a retrospective, single-centre, cross-sectional study conducted at a tertiary sleep disorders centre. Consecutive adults referred for suspected sleep-disordered breathing who underwent attended full-night laboratory polysomnography between 1 January 2024 and 1 January 2025 were screened. The source dataset contained 694 records. The analysis cohort comprised 290 patients meeting the pre-

specified definition of severe OSA (AHI ≥ 30 events/h) [4]. Patients with central sleep apnea, prior positive-airway-pressure therapy, neuromuscular disease, or unavailable ESS data were excluded according to the original study protocol. All eligible records in the defined period were included; therefore, no a priori sample-size calculation was performed.

Sex was analysed according to the study data dictionary (0 = male, 1 = female). Patient names were not used to infer or alter sex. Analyses were performed on coded, de-identified study variables.

2.2 Polysomnography

Full-night polysomnography included electroencephalography, bilateral electrooculography, chin and leg electromyography, electrocardiography, oronasal airflow, thoracic and abdominal effort, and pulse oximetry. Sleep staging and respiratory-event scoring followed American Academy of Sleep Medicine criteria [4, 12]. Apnea was defined as cessation of airflow for at least 10 s; hypopnea was defined as a ≥ 30% reduction in airflow for at least 10 s associated with ≥ 3% oxygen desaturation or an arousal. AHI was calculated as the number of apneas plus hypopneas per hour of sleep.

Sleep architecture variables included N1%, N2%, N3%, and REM%. The N2/REM ratio was calculated as N2% divided by REM%. Because division by zero is undefined, N2/REM was treated as missing when REM%=0. In the severe-OSA cohort, 61 patients had REM%=0 and one additional patient had missing REM%, leaving 228 patients with an analysable N2/REM ratio. This complete-case subset was used only for analyses involving N2/REM.

2.3 Oxygenation variables and DDI

Mean nocturnal SpO_2 and nadir SpO_2 were extracted directly from the PSG record. DDI was derived for this study as DDI = mean nocturnal SpO_2 − nadir SpO_2 and is expressed in percentage points. This exploratory metric reflects the distance between average overnight oxygenation and the lowest recorded saturation; it does not measure event frequency, desaturation area, or cumulative exposure.

The source spreadsheet contained percentage-of-sleep-time fields for SpO_2 90–100%, 80–89%, 70–79%, and 60–69%, but no separate < 60% band. Therefore, the sum of the 80–89%, 70–79%, and 60–69% fields was analysed as recorded hypoxemic time in the 60–89% range (H60–89), rather than labelled as conventional T90. Decimal separators in these fields were normalised during reproducibility checks. Three severe-OSA records had missing or internally implausible oxygen-band totals (> 101% after summing available bands) and were excluded only from H60–89-specific analyses; they remained in all analyses based on other valid variables. Thus, H60–89 analyses included 287 patients.

2.4 Excessive daytime sleepiness and clinical variables

Daytime sleepiness was assessed using the Epworth Sleepiness Scale (ESS), an eight-item self-report measure ranging from 0 to 24 [13]. EDS was defined as ESS > 10, consistent with established clinical use [4, 13]. Patients were classified as EDS (n = 163) or non-sleepy (ESS ≤ 10; n = 127). Recorded clinical

variables included age, BMI, neck circumference, waist circumference, diabetes mellitus, hypertension, chronic kidney disease, smoking, and alcohol use.

2.5 Statistical analysis

Continuous variables were assessed for distributional shape and are reported as median (IQR) because several key variables were non-normally distributed. Between-group comparisons used the Mann–Whitney U test; categorical variables used Pearson's chi-square test. The Shapiro–Wilk framework was used for normality assessment [14]. Spearman rank correlation quantified monotonic associations with ESS [15]. Because non-parametric group comparisons were used, Cohen's d was not reported.

ROC analysis evaluated discrimination of EDS. For variables inversely associated with EDS (mean and minimum SpO_2, REM%), lower values were treated as the positive direction. AUC 95% confidence intervals were obtained by non-parametric bootstrap resampling. Optimal cut-offs were selected using the maximum Youden index [16]. DDI was additionally dichotomised at the overall Youden-optimal cut-off for exploratory clinical stratification.

The multivariable logistic model included DDI, AHI, BMI, age, neck circumference, and sex. Continuous predictors were z-standardised; sex was retained as a binary variable (0 = male, 1 = female), so its odds ratio represents women versus men. Mean SpO_2, nadir SpO_2, and H60–89 were not entered simultaneously with DDI because they are conceptually and statistically overlapping oxygenation constructs. Variance-inflation factors for model predictors were all < 1.60. To examine whether DDI provided information beyond the same clinical covariates, an exploratory nested-model comparison was performed between the full model and an otherwise identical model excluding DDI; model fit was compared using the likelihood-ratio test and the apparent change in AUC was reported. Full-model discrimination was reported as apparent AUC, and internal optimism was estimated using 1,000 bootstrap resamples. Nagelkerke R^2, apparent classification accuracy at a 0.50 probability threshold, and Brier score were also calculated.

Exploratory sex-stratified analyses were performed separately in women and men. To avoid interpreting subgroup differences solely from separate significance tests, a DDI×sex interaction term was added to the adjusted logistic model. The difference between women's and men's DDI AUCs was evaluated using 5,000 stratified bootstrap resamples. All tests were two-sided with $p < 0.05$ considered statistically significant. The original analyses were documented as being conducted in IBM SPSS Statistics 29; revision-stage reproducibility and bootstrap analyses were independently recomputed in Python using SciPy, statsmodels, and scikit-learn.

2.6 Ethics

The study received approval from the Clinical Research Ethics Committee of University of Health Sciences Gazi Yaşargil Training and Research Hospital (approval date: 14 August 2026; document/decision no. 377) and was conducted in accordance with the Declaration of Helsinki [18]. The ethics approval document states that no ethical objection was identified. The supplied approval letter

does not separately document an individual informed-consent waiver; therefore, no waiver decision is attributed to the committee in this manuscript.

3. Results

3.1 Study population and baseline characteristics

Among 694 screened PSG records, 290 patients (41.8%) had severe OSA and formed the study cohort (Fig. 1). EDS was present in 163 patients (56.2%), whereas 127 (43.8%) were non-sleepy. The cohort contained 145 women and 145 men. EDS prevalence was 71.7% (104/145) in women and 40.7% (59/145) in men ($p < 0.001$).

Compared with non-sleepy patients, the EDS group had higher AHI, BMI, neck circumference, waist circumference, DDI, and H60–89, together with lower mean and minimum SpO_2 (Table 1). Age, N1%, N2%, N3%, and heart-rate range did not differ significantly. N2/REM was higher in EDS patients in the 228-patient analysable subset.

Table 1
Baseline characteristics by sleepiness status (N = 290)

Variable	Non-sleepy (n = 127)	EDS (n = 163)	p value
Age, years	49.0 (40.0–56.5)	50.0 (41.0–57.0)	0.503
Female sex, n (%)	41 (32.3%)	104 (63.8%)	< 0.001
BMI, kg/m²	32.4 (28.7–36.7)	34.5 (30.4–38.5)	0.008
Neck circumference, cm	39.0 (36.0–42.0)	42.0 (40.0–43.0)	< 0.001
Waist circumference, cm	109.0 (101.0–114.0)	112.0 (110.0–115.0)	< 0.001
AHI, events/h	40.7 (34.6–55.4)	53.7 (42.1–71.9)	< 0.001
ESS score	6.0 (4.0–8.0)	16.0 (14.0–19.0)	< 0.001
DDI, percentage points	13.0 (9.5–17.8)	18.0 (13.0–23.5)	< 0.001
Mean SpO_2, %	91.0 (89.5–92.0)	89.0 (84.0–91.0)	< 0.001
Minimum SpO_2, %	77.0 (72.5–82.0)	71.0 (58.0–77.0)	< 0.001
H60–89, %*	19.1 (7.0–41.0)	40.7 (18.0–66.1)	< 0.001
N1, %	3.5 (1.1–7.7)	2.3 (1.1–7.3)	0.125
N2, %	56.9 (45.2–66.3)	59.4 (47.8–71.0)	0.139
N3, %	26.7 (19.4–36.8)	27.1 (16.0–42.0)	0.756
REM, %†	7.2 (0.2–15.2)	5.0 (1.2–10.1)	0.083
N2/REM ratio‡	5.37 (3.51–10.57)	8.66 (4.66–22.08)	< 0.001
Mean HR, bpm	70.0 (64.0–78.0)	71.0 (65.0–78.5)	0.247
HR range, bpm	46.0 (38.0–54.0)	46.0 (38.0–55.0)	0.841

*Values are median (IQR) unless otherwise stated. *H60–89 available in 126 non-sleepy and 161 EDS patients after oxygen-band quality control. †REM% available in 126 non-sleepy and 163 EDS patients. ‡N2/REM available in 94 non-sleepy and 134 EDS patients. Mann–Whitney U test for continuous variables; Pearson chi-square for sex.*

3.2 Correlation and univariable discrimination

DDI correlated positively with ESS (ρ = 0.358; p < 0.001). Mean SpO_2, minimum SpO_2, AHI, H60–89, and neck circumference showed larger absolute univariable correlations with ESS than DDI, indicating that DDI should be interpreted as a complementary rather than dominant marker (Table 2).

Table 2
Spearman correlations with ESS score

Variable	ρ	p value	n
Mean SpO_2, %	−0.445	< 0.001	290
Minimum SpO_2, %	−0.430	< 0.001	290
AHI, events/h	+ 0.417	< 0.001	290
H60–89, %	+ 0.416	< 0.001	287
Neck circumference, cm	+ 0.412	< 0.001	290
DDI, percentage points	+ 0.358	< 0.001	290
Waist circumference, cm	+ 0.317	< 0.001	290
N2/REM ratio	+ 0.256	< 0.001	228
BMI, kg/m²	+ 0.208	< 0.001	290
REM, %	−0.095	0.107	289
Age, years	+ 0.090	0.127	290
N3, %	+ 0.049	0.405	290
HR range, bpm	−0.007	0.911	290

Variables are ordered by absolute Spearman ρ. H60–89: recorded percentage of sleep time in SpO_2 bands 60–89%.

In ROC analysis, neck circumference had the highest univariable AUC (0.714), followed by minimum SpO_2 (0.703), mean SpO_2 (0.692), AHI (0.687), DDI (0.677), and H60–89 (0.673). The overall DDI Youden-optimal cut-off was ≥ 19 percentage points, with sensitivity 0.472 and specificity 0.819 (Table 3).

Table 3
ROC discrimination of EDS

Predictor	AUC (95% bootstrap CI)	Sensitivity	Specificity	Youden cut-off	n
Neck circumference	0.714 (0.645–0.773)	0.828	0.535	≥ 40.0 cm	290
Minimum SpO_2	0.703 (0.638–0.761)	0.595	0.717	≤ 73.0%	290
Mean SpO_2	0.692 (0.633–0.758)	0.558	0.764	≤ 89.0%	290
AHI	0.687 (0.628–0.744)	0.724	0.575	≥ 43.9/h	290
DDI	0.677 (0.616–0.742)	0.472	0.819	≥ 19.0	290
H60–89	0.673 (0.609–0.732)	0.584	0.690	≥ 33.22%	287
N2/REM ratio	0.655 (0.582–0.727)	0.866	0.372	≥ 4.04	228
Waist circumference	0.643 (0.576–0.706)	0.773	0.551	≥ 110 cm	290
BMI	0.591 (0.527–0.659)	0.877	0.268	≥ 29.0	290
REM%	0.559 (0.487–0.628)	0.767	0.397	≤ 10.6%	289

For mean/minimum SpO_2 and REM%, lower values were treated as the positive direction. Cut-offs are exploratory and internally derived.

3.3 Multivariable model and internal validation

In the adjusted model, AHI, female sex, neck circumference, and DDI were independently associated with EDS, whereas BMI and age were not. DDI remained significant after adjustment (OR 1.626 per 1-SD increase; p = 0.005). In the exploratory nested-model comparison, the model containing AHI, BMI, age, neck circumference, and sex but not DDI had an apparent AUC of 0.787. Adding DDI increased apparent AUC to 0.810 (ΔAUC 0.023) and significantly improved model fit (likelihood-ratio χ^2=8.66, 1 df; p = 0.003). The full model had Nagelkerke R^2=0.379, classification accuracy = 0.700, and Brier score = 0.174. Mean bootstrap optimism in full-model AUC was 0.0175, yielding an optimism-corrected AUC of 0.793. Because the nested comparison was performed in the derivation cohort, the observed ΔAUC should be interpreted as exploratory rather than as evidence of established clinical utility.

Table 4
Multivariable logistic regression for EDS

Variable	β	OR	Bootstrap 95% CI	p value
AHI	+ 0.750	2.116	1.567–3.251	< 0.001
Female sex (1 vs 0)	+ 1.449	4.257	2.264–10.533	< 0.001
Neck circumference	+ 0.551	1.735	1.148–2.762	0.005
DDI	+ 0.486	1.626	1.094–2.728	0.005
Age	+ 0.081	1.084	0.808–1.547	0.595
BMI	−0.068	0.934	0.612–1.389	0.703

Continuous variables are z-standardised; ORs are per 1-SD increase. Sex is binary and not standardised: 0 = male, 1 = female. Bootstrap CIs are based on 1,000 resamples. Apparent AUC = 0.810; optimism-corrected AUC = 0.793; Nagelkerke R^2=0.379; accuracy = 0.700; Brier score = 0.174.

3.4 Exploratory stratification by DDI cut-off

One hundred patients (34.5%) had DDI ≥ 19 and 190 (65.5%) had DDI < 19. EDS prevalence was 77.0% in the high-DDI group versus 45.3% in the low-DDI group (p < 0.001). High DDI co-occurred with more severe nocturnal oxygenation abnormalities, higher AHI, higher BMI, lower REM%, and a modestly wider heart-rate range (Table 5). Because the cut-off was selected and evaluated in the same cohort, these findings are descriptive and require external validation.

Table 5
Clinical profile by DDI cut-off

Variable	DDI ≥ 19 (n = 100)	DDI < 19 (n = 190)	p value
ESS score	15.5 (11.0–19.0)	9.0 (6.0–15.8)	< 0.001
H60–89, %*	52.7 (29.3–72.7)	19.3 (7.0–41.1)	< 0.001
Mean SpO_2, %	87.0 (81.9–89.0)	91.0 (89.5–92.4)	< 0.001
Minimum SpO_2, %	61.0 (52.8–67.3)	78.0 (74.0–82.0)	< 0.001
AHI, events/h	57.6 (45.8–75.9)	43.0 (35.5–58.5)	< 0.001
BMI, kg/m²	35.2 (30.8–41.7)	32.4 (29.3–36.4)	< 0.001
REM, %†	3.3 (0.0–9.8)	6.5 (1.7–13.4)	0.011
HR range, bpm	49.0 (39.0–57.0)	45.0 (37.0–53.0)	0.020
EDS prevalence, n (%)	77 (77.0%)	86 (45.3%)	< 0.001

*Values are median (IQR) unless stated. *H60–89 available in 99 high-DDI and 188 low-DDI records. †REM% available in 99 high-DDI and 190 low-DDI records. Cut-off was internally derived and should not be interpreted as an established clinical threshold.*

3.5 Exploratory sex-stratified analyses

Among women (n = 145), 104 (71.7%) had EDS; among men (n = 145), 59 (40.7%) had EDS. DDI was higher in sleepy than non-sleepy patients within both sex strata. DDI AUC was 0.730 in men (optimal cut-off ≥ 19; sensitivity 0.576, specificity 0.826) and 0.686 in women (optimal cut-off ≥ 13; sensitivity 0.721, specificity 0.634). The observed AUC difference (men minus women, 0.044) was not statistically significant (bootstrap 95% CI − 0.083 to 0.179; p = 0.507). In the adjusted model, the DDI×sex interaction was also non-significant (p = 0.124). Therefore, subgroup-specific cut-offs are presented as exploratory observations rather than evidence for different sex-specific clinical thresholds.

Table 6
Exploratory sex-stratified comparisons

Variable	Men NS (n = 86)	Men EDS (n = 59)	p	Women NS (n = 41)	Women EDS (n = 104)	p
DDI	14.0 (10.0–18.0)	20.5 (15.0–27.5)	< 0.001	12.0 (9.0–17.0)	15.5 (12.0–22.0)	< 0.001
AHI	48.3 (37.3–57.9)	54.9 (45.4–75.3)	0.001	36.2 (31.3–41.5)	53.6 (41.2–68.6)	< 0.001
BMI	32.2 (28.7–36.7)	36.4 (31.5–46.6)	< 0.001	33.2 (30.0–36.6)	33.1 (30.3–36.4)	0.979
Neck circumference	38.0 (34.0–40.0)	41.0 (39.0–42.0)	< 0.001	42.0 (40.0–42.0)	42.0 (41.0–43.0)	0.048
H60–89, %*	21.4 (8.0–40.2)	54.0 (31.8–79.2)	< 0.001	14.8 (5.6–42.8)	32.5 (12.5–56.3)	0.017
REM, %†	6.1 (0.0–13.7)	3.3 (0.8–9.2)	0.227	10.7 (1.5–16.3)	5.7 (1.4–11.6)	0.052
N2/REM ratio‡	6.77 (3.78–11.99)	12.71 (4.88–27.39)	0.001	4.50 (2.75–7.33)	8.12 (4.57–16.88)	0.001

*Values are median (IQR). *H60–89 excludes records failing oxygen-band quality control. †One man in the non-sleepy stratum had missing REM%. ‡N2/REM sample sizes: men NS n = 62, men EDS n = 49, women NS n = 32, women EDS n = 85. Separate subgroup p values do not establish effect modification; formal DDI×sex interaction p = 0.124.*

4. Discussion

This study identifies a clinically relevant dissociation between conventional OSA severity and daytime symptom burden and introduces a simple way to capture one dimension of that dissociation from routine PSG summaries. Three findings are particularly important. First, 43.8% of patients remained non-sleepy despite severe OSA, reinforcing that AHI alone does not define the symptomatic phenotype. Second, DDI was higher in sleepy patients, correlated with ESS, and remained independently associated with EDS after adjustment for AHI, anthropometry, age, and sex. Importantly, adding DDI to the otherwise identical clinical model increased apparent AUC from 0.787 to 0.810 and improved model fit, although the magnitude of improvement was modest and requires external validation. Third, DDI discrimination differed numerically by sex but formal interaction and AUC-difference testing did not support sex-specific effects. The central novelty is therefore not that DDI outperforms established OSA metrics, which it did not, but that a reproducible descriptor derived from only two routinely available oxygenation values captured additional information about the sleepy severe-OSA phenotype without event-level signal processing.

4.1 Interpretation of DDI and nocturnal hypoxemia

DDI combines two routinely available oxygenation summaries into a simple measure of the distance between mean and nadir SpO_2. This distinguishes its intended use from event-level measures such as desaturation depth, desaturation area, or hypoxic burden, which require identification and processing of individual respiratory events or oxygen-saturation trajectories [6, 7]. The methodological novelty of DDI is therefore pragmatic rather than physiologically comprehensive: it can be reconstructed from standard PSG summary outputs even when raw oximetry signals or event-level hypoxic-burden calculations are unavailable. A high DDI may reflect a large downward excursion from the patient's average overnight oxygenation; however, because it is sensitive to a single nadir and ignores duration, area, frequency, position, and stage, it should not be interpreted as a substitute for event-specific hypoxic burden.

The source dataset also illustrates why terminology matters. Because no < 60% saturation-time band was available, the summed 60–89% bands cannot be assumed to equal conventional T90. We therefore report H60–89 rather than T90. H60–89 remained strongly associated with ESS and EDS after data-quality filtering, but its interpretation is limited to the recorded bands. This distinction improves reproducibility and prevents conflation with established hypoxic-burden constructs [6, 7].

4.2 AHI, anthropometry, and the integrated model

AHI remained the strongest standardised continuous predictor in the adjusted model, while neck circumference and DDI also retained independent associations and BMI did not. The nested comparison provides a more direct answer to whether DDI added information beyond these conventional covariates: inclusion of DDI increased apparent AUC by 0.023 (0.787 to 0.810) and improved likelihood-based model fit ($p = 0.003$). This improvement is statistically informative but clinically modest, and it should not be interpreted as proof of improved decision-making. Rather, it supports the hypothesis that the depth of the lowest overnight oxygen excursion relative to average oxygenation contains symptom-related information that is not fully represented by AHI and anthropometry. Obesity remains a major contributor

to OSA pathophysiology [21], but BMI is an imprecise surrogate for regional anatomy. Bootstrap correction reduced the full-model AUC from 0.810 to 0.793, underscoring the need for external validation before DDI is incorporated into prediction tools or clinical pathways.

4.3 Sex-stratified findings

EDS was more common in women than men in this cohort (71.7% vs 40.7%), underscoring the marked sex-related heterogeneity of symptomatic OSA presentation [9–11, 22, 23]. DDI discrimination was numerically higher in men, and men with EDS also showed larger differences in BMI and H60–89 than non-sleepy men. In women, AHI showed a particularly marked difference between sleepy and non-sleepy groups, while REM% showed a borderline reduction. These patterns suggest that the physiological pathways linked to perceived sleepiness may not be identical across sex strata; however, REM-related and other stage-specific mechanisms cannot be resolved by DDI alone [24].

However, the formal evidence did not support a statistically established sex-specific DDI effect: the DDI×sex interaction was non-significant, and the men–women AUC difference included zero. Accordingly, the different internally derived DDI cut-offs (≥ 19 in men and ≥ 13 in women) should not be presented as sex-specific clinical thresholds. They are hypothesis-generating observations that require substantially larger external cohorts and formal validation.

4.4 N2/REM ratio and missingness

N2/REM was higher in EDS patients overall and within both sex strata, but this analysis was necessarily restricted to 228 patients because 61 had REM%=0 and one had missing REM%. Patients with no recorded REM may represent a clinically distinct subgroup, so complete-case N2/REM analyses may be subject to selection bias. For this reason, N2/REM findings should be interpreted as supportive exploratory observations rather than central evidence. Future studies should examine REM absence itself, total REM duration, REM-specific AHI, and stage-specific hypoxemia rather than relying on a ratio that becomes undefined when REM is absent.

5. Limitations

This study has several limitations. First, its retrospective, single-centre, cross-sectional design precludes causal inference and limits generalisability. The sex variable was analysed according to the source data dictionary (0 = male, 1 = female) and was not independently cross-verified against individual clinical records; therefore, isolated individual-level coding errors cannot be entirely excluded. Second, EDS was defined using ESS alone without objective sleepiness tests such as the Multiple Sleep Latency Test or Maintenance of Wakefulness Test. Third, potentially important contributors to sleepiness—including depression, sedating medications, shift work, habitual sleep duration, insomnia, and circadian factors—were not available in the supplied dataset, leaving residual confounding. Fourth, DDI depends on a single-night global nadir and does not capture event-level desaturation area, within-night variability, position, or stage-specific physiology.

Fifth, the available oxygen-band data did not include a < 60% field and contained three records that were missing or internally implausible after band summation; H60–89 analyses therefore used a quality-controlled subset of 287 patients and should not be equated with conventional T90. Sixth, N2/REM was analysable in only 228 patients, creating potential selection bias. Seventh, multiple exploratory comparisons and sex-stratified analyses increase the risk of chance findings. Finally, the model and ROC cut-offs were derived in the same cohort; bootstrap internal validation reduces but does not eliminate optimism and cannot replace external validation.

6. Conclusions

In severe OSA, DDI was independently associated with EDS after adjustment for AHI, neck circumference, BMI, age, and sex. The study's main contribution is the demonstration that a metric calculated solely from mean and nadir nocturnal SpO_2 can capture additional symptom-related information: adding DDI to the same clinical model increased apparent AUC from 0.787 to 0.810 and improved model fit, while the optimism-corrected AUC of the full model remained 0.793. This added value was modest, and DDI remained inferior to several established variables as a standalone discriminator; accordingly, DDI should be viewed as a pragmatic complementary phenotype marker rather than a replacement for AHI, nadir SpO_2, or event-specific hypoxic-burden measures. The sex-stratified findings were hypothesis-generating only, because neither the DDI×sex interaction nor the between-sex AUC difference was significant. Prospective multicentre validation should determine whether DDI is reproducible, whether it improves prediction beyond established metrics in external cohorts, and whether any clinically useful threshold can be justified.

Abbreviations

AHI
apnea–hypopnea index
AUC
area under the ROC curve
BMI
body mass index
CI
confidence interval
DDI
Desaturation Depth Index
EDS
excessive daytime sleepiness
ESS
Epworth Sleepiness Scale
H60–89
recorded percentage of sleep time with SpO_2 in the 60–89% bands

HR
heart rate
IQR
interquartile range
OSA
obstructive sleep apnea
PSG
polysomnography
REM
rapid eye movement
ROC
receiver operating characteristic
SD
standard deviation
SpO_2
peripheral oxygen saturation

Declarations

Ethical Approval:

Approved by the Clinical Research Ethics Committee of University of Health Sciences Gazi Yaşargil Training and Research Hospital (14 August 2026; document/decision no. 377). The study was conducted in accordance with the Declaration of Helsinki.

Informed Consent:

The study used retrospectively collected, de-identified clinical records. The available ethics approval letter does not separately document a waiver of individual informed consent; accordingly, no waiver decision is attributed to the ethics committee in this manuscript.

Declarations

on the Use of Generative Artificial Intelligence: During manuscript preparation, generative-AI tools were used to assist with language editing, consistency checking, and reproducibility review. The authors reviewed and edited the resulting content and take full responsibility for the manuscript.

Funding:

No funding was received for this research.

Conflict of Interest: The authors declare that they have no conflicts of interest.

Data Availability: The de-identified dataset used for the analyses may be made available by the corresponding author subject to institutional and ethical requirements.

Author Contribution

B.Ç. conceptualized and designed the study, performed the statistical analyses, interpreted the data, and drafted the manuscript. M.S.B. contributed to data interpretation and critically revised the manuscript for important intellectual content. M.K. contributed to the literature review, interpretation of findings, and manuscript editing. All authors reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Data Availability

The de-identified data that support the findings of this study are available from the corresponding author upon reasonable request, subject to institutional and ethical requirements.

References

1. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ et al (2019) Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. Lancet Respir Med 7(8):687–698. 10.1016/S2213-2600(19)30198-5. PMID:31300334
2. Tregear S, Reston J, Schoelles K, Phillips B (2009) Obstructive sleep apnea and risk of motor vehicle crash: systematic review and meta-analysis. J Clin Sleep Med 5(6):573–581. 10.5664/jcsm.27662. PMID:20465027
3. Garbarino S, Scoditti E, Lanteri P, Conte L, Magnavita N, Toraldo DM (2018) Obstructive sleep apnea with or without excessive daytime sleepiness: clinical and experimental data-driven phenotyping. Front Neurol. ;9:505. 10.3389/fneur.2018.00505. PMID:29997573
4. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K et al (2017) Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea. J Clin Sleep Med 13(3):479–504. 10.5664/jcsm.6506. PMID:28162150
5. Zinchuk A, Yaggi HK (2020) Phenotypic subtypes of OSA: a challenge and opportunity for precision medicine. Chest 157(2):403–420. 10.1016/j.chest.2019.09.002. PMID:31539538
6. Azarbarzin A, Sands SA, Stone KL, Taranto-Montemurro L, Messineo L, Terrill PI et al (2019) The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the Osteoporotic

Fractures in Men Study and the Sleep Heart Health Study. Eur Heart J 40(14):1149–1157. 10.1093/eurheartj/ehy624. PMID:30376054

7. Terrill PI (2020) A review of approaches for analysing obstructive sleep apnoea-related patterns in pulse oximetry data. Respirology 25(5):475–485. 10.1111/resp.13635. PMID:31246376
8. Jha S, Chakrabarti S, Talukdar T, Kumar R, Kaushik R, Gupta N et al (2025) Oxygen Desaturation Index in patients of OSA: correlation with sleepiness and hypertension. Sleep Vigil 9(1):75–84. 10.1007/s41782-025-00304-y
9. Wimms A, Woehrle H, Ketheeswaran S, Ramanan D, Armitstead J (2016) Obstructive sleep apnea in women: specific issues and interventions. Biomed Res Int 2016:1764837. 10.1155/2016/1764837. PMID:27699167
10. Theorell-Haglöw J, Miller CB, Bartlett DJ, Yee BJ, Openshaw HD, Grunstein RR (2018) Gender differences in obstructive sleep apnoea, insomnia and restless legs syndrome in adults: what do we know? A clinical update. Sleep Med Rev 38:28–38. 10.1016/j.smrv.2017.03.003. PMID:28495359
11. Shepertycky MR, Banno K, Kryger MH (2005) Differences between men and women in the clinical presentation of patients diagnosed with obstructive sleep apnea syndrome. Sleep 28(3):309–314. 10.1093/sleep/28.3.309
12. Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, Quan SF et al (2017) AASM scoring manual updates for 2017 (Version 2.4). J Clin Sleep Med 13(5):665–666. 10.5664/jcsm.6576. PMID:28416048
13. Johns MW (1991) A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 14(6):540–545. 10.1093/sleep/14.6.540. PMID:1798888
14. Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52(3–4):591–611. 10.1093/biomet/52.3-4.591
15. Spearman C (1904) The proof and measurement of association between two things. Am J Psychol 15(1):72–101. 10.2307/1412159
16. Youden WJ (1950) Index for rating diagnostic tests. Cancer. ;3(1):32–35. 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3. PMID:15405679
17. Nagelkerke NJD (1991) A note on a general definition of the coefficient of determination. Biometrika 78(3):691–692. 10.1093/biomet/78.3.691
18. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA (2013) ;310(20):2191–2194. 10.1001/jama.2013.281053. PMID:24141714
19. Bouloukaki I, Vouis T, Velidakis A, Moniaki V, Mavroudi E, Stathakis G et al (2025) Polysomnography differences between sleepy and non-sleepy obstructive sleep apnea (OSA) patients. Healthc (Basel) 13(5):478. 10.3390/healthcare13050478. PMID:40077040
20. Trzepizur W, Blanchard M, Ganem T, Balusson F, Feuilloy M, Girault JM et al (2022) Sleep apnea-specific hypoxic burden, symptom subtypes, and risk of cardiovascular events and all-cause

mortality. Am J Respir Crit Care Med 205(1):108–117. 10.1164/rccm.202105-1274OC. PMID:34648724

21. Meyer EJ, Wittert GA (2024) Approach the patient with obstructive sleep apnea and obesity. J Clin Endocrinol Metab 109(3):e1267–e1279. 10.1210/clinem/dgad572. PMID:37758218
22. Martins FO, Conde SV (2021) Gender differences in the context of obstructive sleep apnea and metabolic diseases. Front Physiol 12:792633. 10.3389/fphys.2021.792633
23. Tzinas A, Karkala A, Kotoulas SC, Kalamaras G, Mouloudi E, Massa E et al (2026) Age and sex matter: phenotypic heterogeneity, diagnostic gaps, and screening tool performance in obstructive sleep apnea—a 10-year sleep clinic cohort study. Diagnostics (Basel) 16(7):983. 10.3390/diagnostics16070983. PMID:41975696
24. Bonsignore MR, Mazzuca E, Baiamonte P, Bouckaert B, Verbeke W, Pevernagie DA et al (2024) REM sleep obstructive sleep apnoea. Eur Respir Rev 33(171):230166. 10.1183/16000617.0166-2023. PMID:38355150

Figures

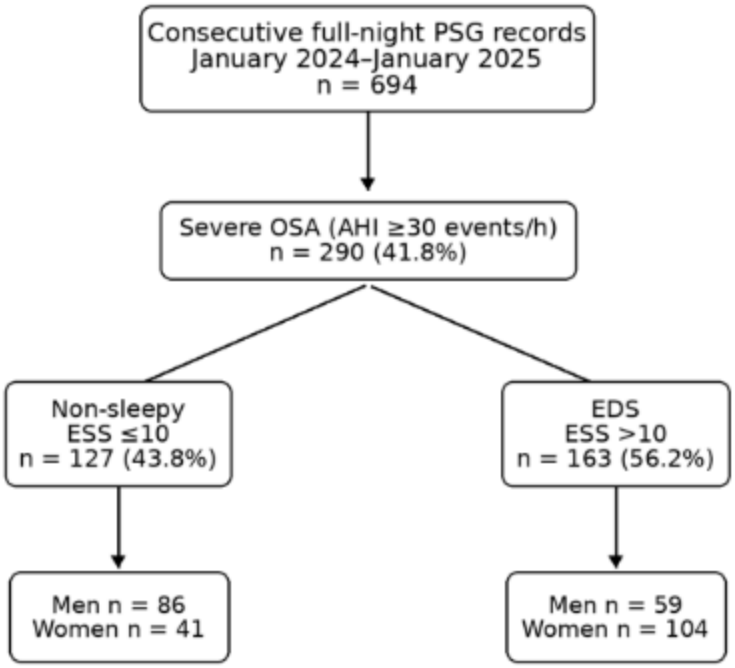


Figure 1

Study flow and sleepiness/sex distribution in the severe-OSA cohort.

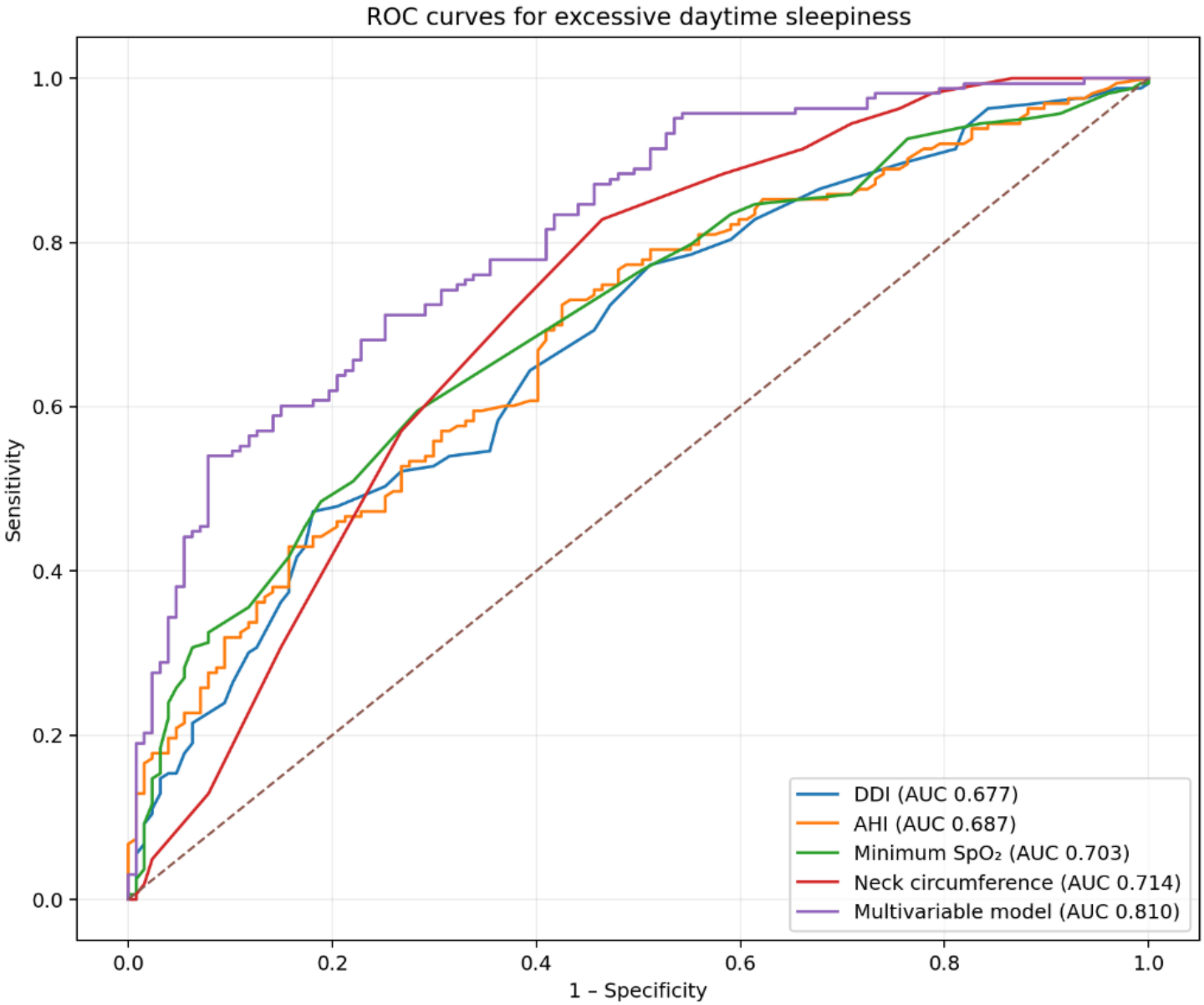


Figure 2

ROC curves for selected univariable predictors and the multivariable model. Model AUC is apparent (derivation-sample) performance.

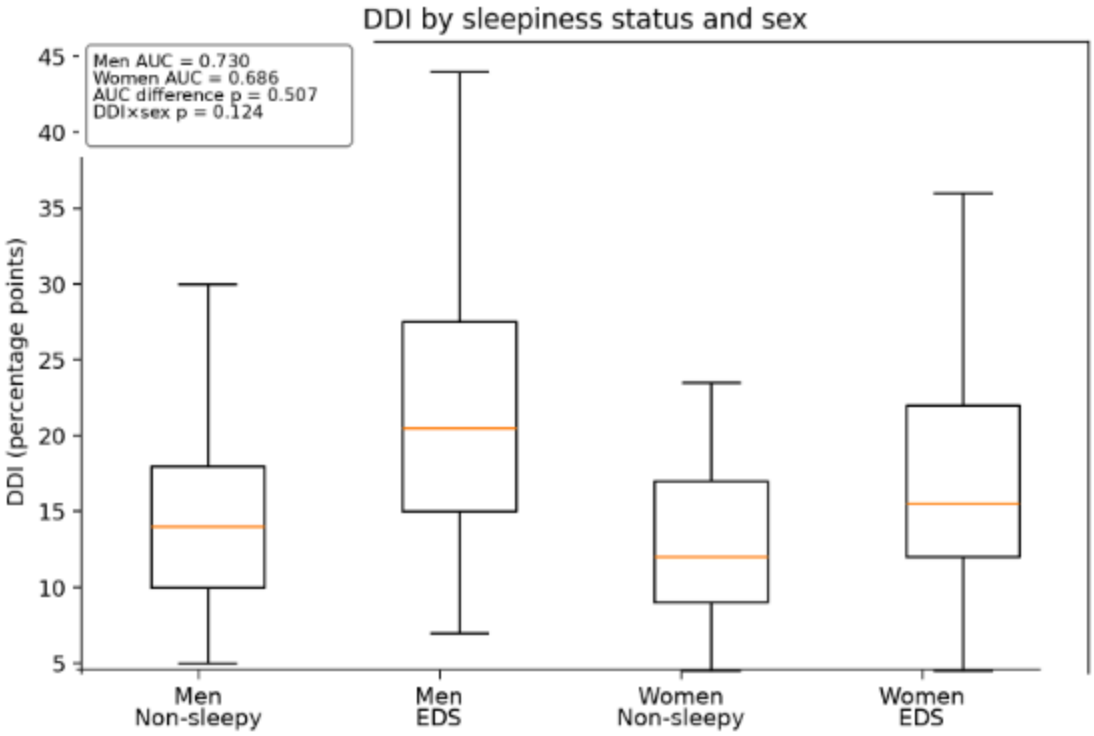


Figure 3

DDI distributions by sleepiness status and sex. Men showed numerically higher DDI discrimination than women, but the between-sex AUC difference and DDI×sex interaction were not statistically significant.