

Assessment of the Test-Retest Reliability of Level III Polysomnography: Comparison of Results Using Difference of Apnea-Hypopnea Index Score and Severity of Obstructive Sleep Apnea

Byung-Hyun Han, Ji Min Yun, and Ick Soo Choi 

Department of Otorhinolaryngology-Head and Neck Surgery, Inje University College of Medicine, Ilsan Paik Hospital, Goyang, Korea

Level III 수면검사의 Test-Retest 결과: Apnea-Hypopnea Index Score 차이와 Severity of Obstructive Sleep Apnea 차이를 이용한 결과 비교

한병현 · 윤지민 · 최익수

인제대학교 의과대학 일산백병원 이비인후과학교실

Received February 20, 2020

Revised September 9, 2020

Accepted December 1, 2020

Address for correspondence

Ick Soo Choi, MD, PhD

Department of Otorhinolaryngology-

Head and Neck Surgery,

Inje University College of Medicine,

Ilsan Paik Hospital,

170 Juhwa-ro, Ilsanseo-gu,

Goyang 10380, Korea

Tel +82-31-910-7250

Fax +82-31-910-7518

E-mail leochoics@gmail.com

Background and Objectives Although studies of test-retest reliability of the standard polysomnography has been reported numerous times, studies of portable sleep devices have been reported in very few cases to date. The purpose of this study is to investigate the test-retest reliability of level III devices, to examine the characteristics of patients with tendency of underestimation and to determine whether there are factors that can predict when retest is needed.

Subjects and Method We enrolled 74 patients for this study from April 2014 to March 2019. Two indicators were used to assess night to night variability. If the difference of apnea-hypopnea index (AHI) score was less than 10 or there is no difference of severity of obstructive sleep apnea (OSA) between two nights, we decided that the results have a reliability.

Results Seventeen out of 57 patients with an AHI difference of less than 10 showed unreliable results based on severity, and 3 out of 17 patients with the AHI difference more than 10 showed reliable results based on severity. Eight out of 74 patients were diagnosed at Night 2 but not at Night 1. In other words, about 11% of the subjects needed retest. Based on the change in severity of OSA, subjective sleep scores showed a statistically significant difference, indicating that the subjective sleep score should be considered when deciding a retest. AHI differed by more than 10 in about 23% of the subjects in this study, falling within “15%–32%,” which has been reported in the literature on the level I test.

Conclusion This suggests that there is no significant difference in reliability between level I and III. Korean J Otorhinolaryngol-Head Neck Surg 2021;64(7):479-85

Key Words Polysomnography · Reproducibility of results · Sleep apnea, obstructive

서론

폐쇄성 수면무호흡증후군(obstructive sleep apnea syn-

drome)은 여러 보고에 따라 다양한 편차를 보이기는 하지만 성인에서는 1%~4%, 65세 이상에서는 24%까지도 보고되고 있으며, 30~60세 남자 4%, 여자 2%에서 발생한다고 알려진 비교적 흔한 질환 중의 하나이다.¹⁻⁶⁾ 폐쇄성 수면무호흡은 수면 중 발생하는 상기도 폐쇄로 인하여 무호흡이나 저호흡이 발생하고 각성에 의해 다시 호흡이 회복되는 것이 반복되는

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

현상으로, 부정맥, 고혈압, 허혈성 심장질환, 뇌혈관질환, 대사 장애 등의 합병증을 유발할 뿐만 아니라 사회적으로도 문제를 일으킬 수 있다. 주간 졸림이 유발되어 일상생활에 상당한 불편을 줄 수 있으며, 교통사고 및 작업장에서의 사고율을 높이고 작업 능력 및 삶의 질이 저하된다. 또한 지속적인 수면 박탈로 인한 인격 변화, 우울증 및 성생활 장애와 같은 문제도 유발할 수 있다.^{3,4,7-10} 최근 수면무호흡증 환자가 크게 증가하는 추세이지만, 중등도 이상의 폐쇄성 수면무호흡증 환자의 경우 남성의 82%, 여성의 93% 가량이 진단되지 않은 상태로 살아간다고 보고되어 있다.^{6,11}

수면무호흡증후군의 진단을 위해서 표준 수면다원검사(standard polysomnography or level I polysomnography)가 현재 표준 검사로 인정되고 있다. 표준 수면다원검사는 코와 입을 통한 공기의 출입, 가슴과 복부의 호흡운동, 혈중 산소 포화도, 뇌파(electroencephalogram), 안구 운동(electrooculogram), 근전도(electromyogram) 등을 측정하여 정확한 수면 시간과 단계를 파악할 수 있으며, 검사자가 검사 시간 동안 관찰을 하고 필요 시 양압기 적정 압력 처방(titration)을 시행할 수 있는 장점이 있다. 하지만 검사의 비용이 높을 뿐 아니라, 평소 수면환경과는 다른 낯선 곳에서 적응 기간 없이 단 1회의 검사로 시행됨으로 인하여 제대로 수면을 취하지 못해 정확한 진단이 되지 않을 가능성의 단점도 있다.¹² 이와 같은 첫날밤 효과(first night effect)는 실제 검사-재검사(test-retest) 연구를 통해 약 15%~32% 정도에서 발생할 수 있는 것으로 보고되고 있다.¹³⁻¹⁵ 이러한 연구 결과는 표준 수면다원검사가 비록 가장 유용한 검사법임에도 불구하고 약 15%~32%의 피검사자에서 수면무호흡 정도의 과소평가 등으로 인해 적절한 치료를 하지 못할 잠재적 위험성이 있다는 것을 의미한다고 할 수 있다. 수면검사기기 중 이동식 수면검사기기(portable sleep monitoring device) 또는 level III 수면다원검사는 뇌파, 안구운동, 근전도 등을 측정할 수 없으므로, 정확한 수면 시간과 수면 단계를 파악하기 어렵다는 단점이 있으나 무호흡, 저호흡 및 관련 수치는 표준 수면다원검사와 비교하였을 때 크게 차이가 없는 것으로 보고되고 있는데,^{1,16-18} 수면 질환 자체가 아닌 수면무호흡에 국한된 검사에서는 표준 수면다원검사를 대체할 수 있는 유용한 검사라고 할 수 있다. 이동식 수면검사기기는 표준 수면다원검사와 비교하였을 때 경제성, 검사 시기 결정의 유연성, 재검사의 용이성 등의 장점이 있다고 할 수 있다. 또한 자가에서 시행이 가능하여 평소의 수면 환경과 동일하다는 장점으로 인해 표준검사에 비해 첫날밤 효과가 낮을 것으로 기대된다.¹⁶⁻¹⁸

수면검사의 검사-재검사 신뢰도(test-retest reliability) 및 첫날밤 효과에 관한 연구는 1981년 Coates 등¹⁹의 연구를 시

작으로 현재까지 계속 보고되어 왔으며, level I 검사는 level III 검사에 비해 단일 검사 신뢰도가 비교적 떨어지고 첫날밤 효과가 더 크다는 보고는 있으나, level III의 검사-재검사 연구는 level I 검사에 비해 현재까지 매우 제한적으로 보고되었다.²⁰ 또한 기존의 문헌들에서는 두 검사 간의 apnea-hypopnea index(AHI) 차이가 10 미만일 때를 기준으로 하여 검사-재검사 신뢰도가 있다고 판단하고 있으나,²¹ 이러한 기준은 실제 임상에서 진단과 치료의 기준이 되는 중증도(severity)를 기준으로 하였을 때 그 결과가 달라질 가능성이 있다. 그러므로 본 연구에서는 폐쇄성 수면무호흡의 진단에 유용하게 사용될 수 있는 level III 기기의 검사-재검사 신뢰도를 두 검사 간의 AHI 차이와 중증도 차이를 통해 알아보고 그 결과를 비교하고자 하며, 재검사에서 중증도가 상승한 환자들을 분석하여 첫 검사에서 과소평가될 가능성이 높은 환자들에 대한 예측 인자를 확인하고자 한다.

대상 및 방법

대 상

본 연구는 2014년 4월~2019년 3월까지 코골이 또는 수면무호흡, 과다 주간 졸림증 등을 주소로 인제대학교 일산백병원 이비인후과 수면클리닉을 방문하여, level III 수면검사기기를 이용하여 수면 검사를 시행한 환자들 중 검사 결과에 대해 환자가 불만족하여 재검을 원하거나, 수술적 치료 전 증상 지속 여부를 확인하기 위해 2달 이내에 재검을 시행한 74명을 대상으로 후향적 연구를 시행하였다. 연구 대상자는 나이 20세 이상, 체질량지수(body mass index, BMI) 30 이하, 현재 양압기 치료를 하지 않고 있는 환자를 포함하였으며 천식이나 만성 폐쇄성 폐질환과 같은 호흡기 질환을 가진 환자는 연구 대상에서 제외하였다. 진료 당시 주간 활동 시간에 졸리는 정도를 측정하는 주간기면지수(Epworth Sleepiness Scale, ESS)²²를 통해 설문조사를 시행하였고, 수면검사 당일의 수면 만족도를 0점(잠을 설쳤다), 1점(잘 못 잤다), 2점(비교적 잘 잔 편이다), 3점(평소처럼 잤다)의 주관적인 수면 만족 점수로 측정하여 기록하였다. 인제대학교 일산백병원 임상윤리위원회(IRB No. 2019-02-012-003)의 심의를 통과하였고, 후향적 차트 분석을 통해 연구를 진행하였다. 검사기기를 착용하고 수면을 하는 경우, 본인의 평소 수면 습관과 동일하게 하도록 설명하였고 평소와 동일한 환경을 조성하도록 하였다.

방 법

본 연구에 사용된 Embletta X10(Natus Medical Inc., San carlos, CA, USA) 모델은 9 채널 모델로 비강 압력 변환기,

온도 감지기, 2개의 호흡 운동 측정기(가슴, 복부), 산소 포화도 측정기, 자세 감지기, 코골이 감지기 등으로 구성되어 있다. 환자 또는 보호자에게 잠들기 전에 시작 버튼을 누르고 잠에서 깬 때 멈춤 버튼을 누르도록 교육을 하였으며, 수면 시간이 300분 이상일 때를 충분한 수면을 취했다는 기준으로 삼았다. 무호흡(apnea)은 기류(air flow)의 90% 이상의 감소가 최소 10초 이상 지속될 때를 기준으로, 저호흡(hypopnea)은 기류의 50% 이상의 감소가 최소 10초 이상 지속될 때를 자동으로 기록하여 계산한다. 무호흡-저호흡 지수(AHI)는 시간당 무호흡과 저호흡을 합한 값으로 정의되어 자동으로 계산된다.

무호흡-저호흡 지수(AHI), 무호흡 지수(apnea index, AI), 저호흡 지수(hypopnea index, HI), 평균 산소 포화도(average oxygen saturation), 최저 산소 포화도(lowest oxygen saturation) 등을 두 번의 검사를 통해 기록하였다. 폐쇄성 수면무호흡의 진단은 AHI값이 5 이상일 때를 기준으로 하였으며, AHI값을 기준으로 중증도를 구분하였다. AHI값이 5 미만인 경우 정상군(normal), 5 이상 15 미만인 경우 경증군(mild), 15 이상 30 미만인 경우 중등증군(moderate), 30 이상인 경우를 중증군(severe)으로 정의하였다. 또한 검사 시행 전 ESS 설문지를 작성하여 기록하였고, 첫 번째 검사 후 환자의 주관적 수면 만족 점수를 기록하게 하였다.

두 검사 간 AHI 결과의 야간 변동성을 평가하기 위해 두 가지 기준을 이용하였다. 첫 번째는 일반적으로 흔히 사용되는 방법으로 AHI 차이가 10 미만인 군과 AHI 차이가 10 이상인 군으로 분류하여 AHI 차이가 10 미만일 때를 검사-재검사 신뢰도가 있다고 판단하였다. 두 번째는 두 밤 간의 중증도 변화가 없는 군과 변화가 있는 군으로 분류하여 중증도 변화가 없는 군을 검사-재검사 신뢰도가 있다고 판단하였다.

첫 검사에서 실제 수면무호흡 정도에 비해 검사 결과가 과소평가되는 환자들을 예측하기 위하여 두 가지 인자를 고안하였다. AI와 HI 중 AI의 기준 역치가 더 높으므로 HI에 비해 재검 시 변동성이 더 작을 것이라는 가정 하에 AI/AHI값을 계산하여 두 번째 검사에서 중증도가 상승한 군과 그대로인 군을 비교해보았다. 또, 환자가 호소하는 주관적 증상에 비해 초기 검사 결과가 낮게 나오는 군이 과소평가되었을 가능성이 높으므로 AHI에서 ESS를 뺀 값(AHI-ESS)을 계산하였고 이를 범주화하여 두 군의 분포를 비교해보았다.

통계학적 분석은 IBM SPSS Statistics 23(IBM Corp., Armonk, NY, USA)를 이용하였다. 정규성 검증을 통과하지 못하는 경우 비모수 검정으로 Mann-Whitney U test를 시행하였으며, 독립 표본 t 검정을 통해 1차 검사와 2차 검사의 AHI, AI, HI, oxygen saturation 등을 비교하여 두 검사 간의 야간

변동성(night to night variability)을 평가하였고, 이 때 통계적인 유의성은 $p < 0.05$ 일 때로 간주하였다.

결 과

총 74명의 대상자에서 연령분포는 13~69세였으며 평균 나이는 44.65 ± 13.21 세였다. 성별 분포는 남자가 62명, 여자가 12명이었고 평균 체질량지수(BMI)는 25.31 ± 4.01 (kg/m^2)이었다. 초기 검사에서 정상군 28명(37.8%), 경증군 25명(33.8%), 중등증군 13명(17.6%), 중증군 8명(10.8%)이 연구에 포함되었다(Table 1).

두 검사 간 검사-재검사 신뢰도는 AHI 차이 기준과 중증도 기준으로 각각 57명(77.0%), 43명(58.1%)에서 신뢰성을 보였다(Table 2). AHI 차이 기준 10 미만으로 신뢰성을 보인 57명 중 17명(29.8%, 74명 중 23.0%)은 중증도를 기준으로 할 때 신뢰성이 없는 결과를 보이며, AHI 차이 기준 신뢰성이 없는 17명 중 3명(17.6%, 74명 중 4.1%)에서 중증도를 기준으로 할 때 신뢰성 있는 결과를 보였다. 두 기준을 모두 만족하며 신뢰성을 보인 경우는 40명(54.1%), 두 기준을 모두 만족하지 않는 환자는 14명(18.9%)으로 나타났다(Table 3). 이 중 첫 번째 검사에서 정상 소견을 보였다가 재검에서 경증 소견을 보인 환자는 7명, 경증에서 중등증 소견은 7명, 중등증에서 중증 소견은 2명, 정상에서 중등증 소견은 1명, 경증에서 중

Table 1. Characteristics of 74 patients in the study population

Age, years	44.65 ± 13.21
Sex	
Male	62 (83.8)
Female	12 (16.2)
Body mass index, kg/m^2	25.31 ± 4.01
Severity (night 1)	
Normal	28 (37.8)
Mild	25 (33.8)
Moderate	13 (17.6)
Severe	8 (10.8)

Data are presented as the mean ± standard deviation or n (%).

Table 2. Number of patients classified by test-retest reliability

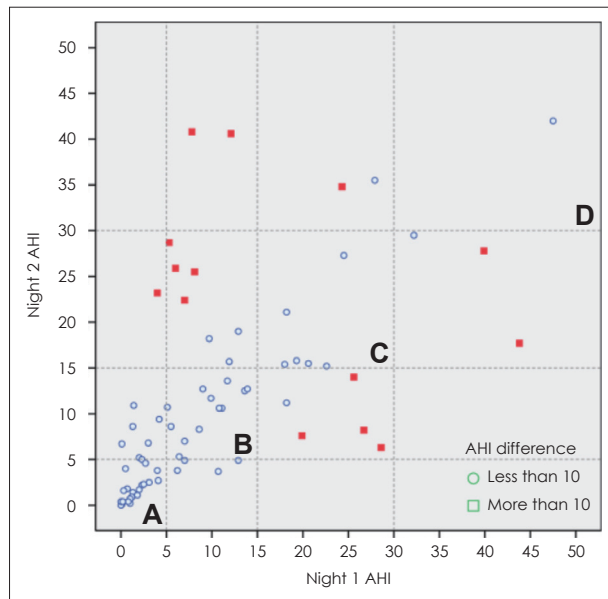
	Test-retest reliability	Total (n=74)
AHI difference		
Less than 10	Present	57 (77.0)
More than 10	Absent	17 (22.9)
Severity of OSA		
Same	Present	43 (58.1)
Different	Absent	31 (41.9)

Data are presented as n (%). AHI: apnea-hypopnea index, OSA: obstructive sleep apnea

Table 3. Number of patients classified by AHI difference and severity change

AHI difference	Severity of OSA	Total (n=74)
Less than 10 (n=57)	Same	40 (70.2/54.1)
	Different	17 (29.8/23.0)
More than 10 (n=17)	Same	3 (17.6/4.1)
	Different	14 (82.4/18.9)

Data are presented as n (%). AHI: apnea-hypopnea index, OSA: obstructive sleep apnea

**Fig. 1.** Correlation plot of the AHI (according to AHI difference) between two nights [AHI range: 0–55, normal to normal (A), mild to mild (B), moderate to moderate (C), severe to severe (D)]. AHI: apnea-hypopnea index.

중 소견은 2명으로 나타났다. Night 1과 Night 2 AHI값 분포를 도표로 나타내었다(Fig. 1). 그림에서 AHI의 범위는 0~55로 제한하여 두 검사 중 한번이라도 AHI값이 55를 초과하는 4명의 환자들은 아웃라이어 데이터로 간주하여 제외하였다. 면적 A는 Night 1에서 정상을 보인 환자가 Night 2에서도 정상인 경우를 나타내고, 면적 B는 경증에서 경증, 면적 C는 중등증에서 중등증, 면적 D는 중증에서 중증을 보인 경우를 나타낸다. 또한 각각의 정사각형에서 외의 점들은 중증도가 두 검사에서 차이를 보이는 경우를 나타낸다.

두 검사 간 AHI 차이는 10 미만이나 중등도 기준으로는 차이를 보이는 17명 중 정상에서 경증은 7명, 경증에서 중등증은 3명, 중등증에서 중증은 1명, 더 경한 소견으로 변화가 있었던 환자는 6명으로 나타났다.

AHI 변화와 중증도 변화에 대한 각각의 군에서 Night 1의 검사 결과 수치의 평균값을 분석하였다(Tables 4 and 5). AHI 차이가 10 미만인 군과 10 이상인 군의 AHI 평균은 각각 $10.24 \pm 16.06/25.56 \pm 19.99$ ($p=0.002$), AI 평균은 $6.45 \pm 12.76/16.92$

Table 4. Night 1 parameters of AHI difference groups (mean±standard deviation)

AHI difference	Less than 10 (n=57)	More than 10 (n=17)	p
AHI	10.24 ± 16.06	25.56 ± 19.99	0.002**
AI	6.45 ± 12.76	16.92 ± 6.50	0.001**
HI	3.70 ± 4.52	8.64 ± 7.03	0.003**
O ₂ Avg	95.56 ± 2.50	95.15 ± 2.50	0.396†
O ₂ Low	84.35 ± 9.22	79.88 ± 7.56	0.016**

*satisfying $p < 0.05$, †two sample t test, **Mann Whitney U test. AHI: apnea-hypopnea index, AI: apnea index, HI: hypopnea index, O₂Avg: average oxygen saturation, O₂Low: lowest oxygen saturation

Table 5. Night 1 parameters of severity change groups (mean±standard deviation)

Severity change	Same (n=46)	Different (n=28)	p
AHI	11.44 ± 18.94	16.97 ± 16.62	0.197*
AI	7.12 ± 14.45	11.25 ± 13.93	0.223*
HI	4.20 ± 5.60	5.71 ± 5.46	0.249*
O ₂ Avg	95.47 ± 2.76	95.45 ± 2.11	0.974*
O ₂ Low	84.56 ± 10.10	81.61 ± 7.07	0.167*

*two sample t test. AHI: apnea-hypopnea index, AI: apnea index, HI: hypopnea index, O₂Avg: average oxygen saturation, O₂Low: lowest oxygen saturation

Table 6. ESS, subjective sleep score and AI/AHI of severity change groups (mean±standard deviation)

	Same (n=43)	Increased (n=19)	p
ESS	9.33 ± 5.19	10.37 ± 3.53	0.361
Subjective sleep score	2.05 ± 1.02	0.79 ± 0.98	0.001*
AI/AHI	0.46 ± 0.38	0.55 ± 0.28	0.324

*satisfying $p < 0.05$. ESS: Epworth sleepiness scale, AI: apnea index, AHI: apnea-hypopnea index

± 6.50 ($p=0.001$), HI 평균은 $3.70 \pm 4.52/8.64 \pm 7.03$ ($p=0.003$)으로 AHI 차이가 10 이상인 군에서 평균값이 유의미하게 높았으며, 최저 산소포화도 수치는 각각 $84.35 \pm 9.22/79.88 \pm 7.56$ ($p=0.016$)으로 10 이상인 군에서 유의미하게 낮았다. 중등도 변화가 없는 군과 있는 군에서는 각 수치 간의 유의미한 차이가 없었다($p>0.05$).

두번째 검사에서 중증도 변화가 없는 군과 중증도가 상승한 군으로 나누어 ESS, 주관적 수면 점수, AI/AHI, AHI-ESS 값을 비교하였다. ESS는 중증도 변화가 없는 군과 상승한 군에서 각각 9.33 ± 5.19 , 10.37 ± 3.53 ($p=0.361$)로 두 군 간에 통계적으로 유의한 차이가 없었고, 주관적 수면 점수는 각각 2.05 ± 1.02 , 0.79 ± 0.98 ($p=0.001$)로 중증도가 상승한 군에서 유의하게 낮았다(Table 6). AI/AHI는 각각 0.46 ± 0.38 , 0.55 ± 0.28 ($p=0.324$)로 두 군 간에 유의한 차이가 없었다. AHI-ESS값은 전체 AHI의 평균값이 13.76 ± 18.10 , 전체 ESS의 평균값이 9.62 ± 4.82 이므로 각 평균값의 차인 4.14가 포함되는 군을 중

Table 7. AHI-ESS value of severity change groups

AHI-ESS subtypes	Same (n=43)	Increased (n=19)
AHI-ESS < -5	14 (32.56)	8 (42.10)
-5 ≤ AHI-ESS < 15	23 (53.49)	10 (52.63)
15 ≤ AHI-ESS	6 (13.95)	1 (6.67)

Data are presented as n (%). AHI: apnea-hypopnea index, ESS: Epworth sleepiness scale

간 범주로 하기 위해 -5 미만/-5 이상, 15 미만/15 이상으로 범주화 하였다. 중증도 변화가 없는 군에서 각 범주에 해당하는 환자 수는 14명(32.56%), 23명(53.49%), 6명(13.95%)이었으며, 중증도가 상승한 군에서 각각 8명(42.10%), 10명(52.63%), 1명(6.67%)으로 두 군 간에 유의한 분포 차이는 보이지 않았다(Table 7).

고 찰

수면검사의 검사-재검사 신뢰도에 관한 연구는 1981년 Coates 등¹⁹⁾이 12명의 정상 환자군, 12명의 불면증 환자군을 대상으로 이동식 level II 수면기기를 이용하여 연속적인 네 번의 검사를 통해, 두 군 모두 첫날밤과 뒤 이은 밤에서의 수면 적정성에는 큰 차이가 없음을 보여주었고, 1988년 Sharpley 등²³⁾이 다른 수면 관련 질환이 없는 환자 12명을 대상으로 level II 기기를 이용하여 집에서 세 번의 연속적인 검사를 하였고, 이를 통해 첫날밤 효과가 크지 않음을 보였다. Leven-dowski 등¹³⁾은 118명의 환자를 대상으로 level I 기기를 이용하여 2번의 표준 수면다원검사를 시행하였고, 치료 결과를 측정하는 데 있어 단일 수면다원검사의 신뢰도가 떨어지고 야간 변동성이 적절히 보강될 필요가 있으며, 1회 이상의 검사 데이터를 수집하고 평균하여 좋은 결과를 얻을 수 있을 것이라고 말했다. 또한 그는 37명의 환자를 대상으로 각각 level I 검사를 2회, 집에서 시행한 level III 검사를 2회 시행하였고, 집에서 시행한 검사의 첫날밤 효과가 더 적다는 것을 보여주었다.²⁰⁾ 현재까지 보고된 문헌들을 종합하면 level I 검사는 level III 검사보다 단일 검사 신뢰도가 떨어지고 첫날 밤 효과가 더 많이 존재한다는 것을 알 수 있다.

표준 수면다원검사와 이동식 수면검사를 한 환자에게 동시에 시행하여 두 검사 간의 상관관계를 보고, 이동식 수면검사가 선별검사로서 유용성이 있는지를 연구한 문헌은 국내에도 보고되어 있으며, 보고된 문헌에 따르면 무호흡, 저호흡 및 관련 수치는 표준 검사인 level I 수면다원검사와 비교하였을 때 크게 차이가 없는 것으로 나타났다. Hong 등¹⁾은 표준 수면다원검사와 이동식 수면검사기기인 MESAM IV와 Edentec을 동일한 환자에게 각각 1회씩 시행했을 때, 두 검사의 무호

흡 지수와 호흡장애 지수 사이에 통계적으로 유의한 상관관계가 있음을 보였고, Kim 등¹⁶⁾은 표준 수면다원검사와 이동식 수면검사기기인 Sleepstrip을 비교하여 그 진단적 유용성을 증명했다. 또한 Kim 등¹⁶⁾은 표준 수면다원검사와 이동식 수면검사기기인 ApneaLink의 비교에서 여러 가지 수면 지표의 높은 상관관계 및 민감도, 특이도를 보여줌으로써 수면무호흡증의 진단에 있어 ApneaLink의 임상적 유용성을 보여주었다. Chung 등¹⁸⁾은 표준 수면다원검사와 이동식 수면검사기기인 Embletta를 비교하여 두 검사 간의 AHI, AI, HI는 통계적으로 매우 유의한 상관관계가 있음을 보여주었다.

이동식 수면검사는 평소의 수면 환경과 동일하다는 점을 생각할 때 일반적으로 표준 수면다원검사에 비해 첫날밤 효과가 작아 재검사의 필요성이 낮을 것으로 추정할 수 있으나, 본 연구 결과 74명 중 8명(10.81%)에서 Night 1에서는 진단되지 않고 Night 2에서 진단되었다(normal to mild, 7; normal to moderate, 1). 즉 재검사가 필요한 대상은 약 11% 정도라고 볼 수 있다. 기존 문헌들을 조사한 바에 따르면, 현재까지 level III 수면검사의 검사-재검사 신뢰도에 대한 연구는 매우 제한적으로 보고되고 있고 재검사의 필요성에 대한 예측 인자를 분석한 문헌은 없었다. 또한 기존의 검사-재검사 신뢰도에 대한 연구에서는 검사 간의 AHI 차이로 신뢰도를 판단하였으나, 본 저자들은 수면 무호흡의 중증도가 실제 임상에서 진단과 치료의 기준이 되는 값이므로 중증도 변화 역시 신뢰도의 지표로 고려하였다. 따라서 본 연구에서는 level III 검사를 두 번 시행한 환자군을 대상으로 두 검사 간 AHI, 중증도 변화가 있었던 환자군과 없었던 환자군으로 나누어 각 군의 분포와 구별되는 특성을 파악하고자 하였다. 또한 실제로 수면 무호흡이 있지만 첫날밤 효과 등에 의해 과소평가되어 진단되지 않는 환자들을 예측할 수 있는 인자가 있는지 확인하고자 하였으며, 이를 위해 AI/AHI, AHI-ESS, 주관적 수면 점수를 고안하여 재검사에서 중증도 변화가 없는 환자군과 중증도가 상승한 환자군을 비교하였다.

두 검사 간 AHI 결과의 야간 변동성을 평가하기 위해 기존 연구들을 참고하여 AHI 차이가 10 미만일 때를 검사-재검사 신뢰도가 있다고 판단하였다. Level III 검사인 본 연구에서, 약 23%에서 AHI가 10 이상 차이가 났고(Table 2), 기존에 보고된 level I 검사의 문헌들에 따르면 두 밤 간에 10 이상의 AHI 차이를 보이는 환자는 15%~32%로 알려져 있다.^{17,18)} 이로 미루어 볼 때, 두 밤 간의 AHI 차이가 10 미만일 때 검사-재검사 신뢰도가 있다고 한다면, level I과 level III 사이에는 검사-재검사 신뢰도에 있어서 큰 차이가 없다고 볼 수 있다. 기존의 연구들과 차별점을 두기 위하여 본 연구에서는 두 밤 간의 중증도 변화를 기준으로 검사-재검사 신뢰도를

평가했으며(Table 2), AHI 차이와 중증도 변화를 동시에 평가했다(Table 3). 두 밤 간의 중증도 변화가 없었던 환자는 약 58%였고, AHI 차이가 10 미만이면 동시에 중증도 변화가 없는 환자는 약 54%로 나타났다. 따라서 AHI 차이만을 고려했을 때인 77%와 비교하여 중증도 변화를 함께 적용했을 때 좀 더 특이도 높은 검사-재검사 신뢰도 평가가 가능할 것으로 판단하였다. 두 검사 간 AHI 차이는 10 미만이나 중증도 기준으로는 차이를 보이는 17명 중 정상에서 경증은 7명, 경증에서 중등증은 3명, 중등증에서 중증은 1명으로 나타났다. Night 1에서 정상 소견을 보였던 28명 중 Night 2에서 AHI 차이는 10 미만이나 중증도 상승을 보인 7명(25%), Night 1 경증 25명 중 3명(12%), Night 1 중등증 13명 중 1명(8%)으로 비교해보면, 재검 시 AHI 기준으로는 두 검사 간 유의미한 결과 차이가 없으나 중증도 기준으로는 변화를 보일 가능성이 가장 높은 환자군은 정상 환자군이라고 해석할 수 있다. 따라서 향후 검사-재검사 신뢰도 연구에서 초기 검사에서 정상 소견을 보인 환자군에서는 중증도 변화를 신뢰도의 지표로 이용하는 것을 고려해 볼 수 있다.

검사 간 AHI 차이가 10 미만인 군과 10 이상인 군의 Night 1 검사 수치 중 AHI, AI, HI는 10 이상인 군에서 유의하게 높았고 최저 산소 포화도는 유의하게 낮았다. 이를 통해 Night 1에서 관련 수치들이 높게 나온 환자들은 재검 시 AHI 변동이 클 가능성이 높다고 해석할 수 있으나, 임상적으로는 첫 검사에서 환자의 주관적 증상에 비해 결과가 경하게 나오는 환자들에서 적절한 진단을 위해 재검을 고려하게 되므로 높은 AHI를 재검 필요성에 대한 예측 인자로 사용하는 것은 의미가 없다고 판단하였다. 이에 비해 중증도 변화를 기준으로 비교했을 때는 두 군 간의 유의미한 차이가 없었으므로 실제 임상을 더 잘 반영하는 신뢰도 지표는 중증도 변화라고 해석하였다.

주관적 수면 점수는 중증도 변화가 없는 군(same)과 상승한 군(increased)에서 각각 2.05 ± 1.02 , 0.79 ± 0.98 ($p=0.001$)로 두 군 간에 통계적으로 유의한 차이를 보였으며, 중증도 변화가 없는 군에서 더 높은 것을 볼 수 있었다(Table 6). 따라서, 이동식 수면검사 시 환자가 평소의 수면 상태에 비해 충분히 수면을 취하지 못했다고 주관적으로 판단할 경우 재검사가 필요할 수 있다. 저자들이 고안한 AI/AHI, AHI-ESS는 중증도 변화가 없는 군(same)과 상승한 군(increased)에서 유의미한 차이를 보이지 않았으므로 재검 시 중증도 상승에 대한 예측 인자로 이용할 수 없음을 확인하였다(Tables 6 and 7).

결론적으로, 본 연구에서 이동식 수면검사의 검사-재검사 신뢰도를 평가하기 위해 두 밤 간의 AHI 차이와 중증도 변화

두 가지 요소를 이용하였고, AHI 차이만을 이용했을 때보다 더 특이도 높은 검사-재검사 신뢰도 평가가 가능하다고 판단했다. 첫 검사에서 실제 수면 무호흡 정도에 비해 검사 결과가 과소평가되는 환자들을 예측하기 위한 인자로써 주관적 수면 점수를 이용할 수 있다.

Acknowledgments

None.

Author Contribution

Conceptualization: Ick Soo Choi. Data curation: Ick Soo Choi. Formal analysis: Byung-Hyun Han, Ji Min Yun. Investigation: Byung-Hyun Han. Methodology: Byung-Hyun Han. Project administration: Byung-Hyun Han. Resources: Ji Min Yun. Software: Ji Min Yun. Supervision: Ick Soo Choi. Validation: Ick Soo Choi. Visualization: Ji Min Yun. Writing—original draft: Byung-Hyun Han. Writing—review & editing: Ick Soo Choi, Ji Min Yun.

ORCID

Ick Soo Choi

<https://orcid.org/0000-0003-3538-7252>

REFERENCES

- 1) Hong SK, Park MH, Kim CN, Yoon HJ, Eum JH, Lee IK. The clinical efficacy of portable monitoring devices (MESAM IV and Edentec) in evaluating sleep apnea syndrome: A comparative study with simultaneously recorded standard polysomnography. *Korean J Otolaryngol* 1997;40(10):1404-9.
- 2) Min YG, Jung HW, Kim IT, Rhee CS, Jang YJ, Shim YS, et al. Evaluation of polysomnographic findings on diagnosis of OSA. *Korean J Otolaryngol* 1992;35(1):88-94.
- 3) Emsellem HA, Corson WA, Rappaport BA, Hackett S, Smith LG, Hausfeld JN. Verification of sleep apnea using a portable sleep apnea screening device. *South Med J* 1990;83(7):748-52.
- 4) Hida W, Shindoh C, Miki H, Kikuchi Y, Okabe S, Taguchi O, et al. Prevalence of sleep apnea among Japanese industrial workers determined by a portable sleep monitoring system. *Respiration* 1993;60(6):332-7.
- 5) Sullivan CE, Issa FG. Obstructive sleep apnea. *Clin Chest Med* 1985;6(4):633-50.
- 6) Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328(17):1230-5.
- 7) Ross SD, Sheinait IA, Harrison KJ, Kvasz M, Connelly JE, Shea SA, et al. Systematic review and meta-analysis of the literature regarding the diagnosis of sleep apnea. *Sleep* 2000;23(4):519-32.
- 8) Lam JC, Ip MS. An update on obstructive sleep apnea and the metabolic syndrome. *Curr Opin Pulm Med* 2007;13(6):484-9.
- 9) Kim HS. Analysis of snoring and obstructive sleep apnea using a portable sleep apnea recorder. *Korean J Rhinol* 1996;3(1):37-42.
- 10) Fairbanks DNF. Snoring: An overview with historical perspectives. In: Fairbanks DNF, Fujita S, editors. *Snoring and obstructive sleep apnea*, 2nd ed. New York: Raven Press;1994. p.1-16.
- 11) Young T, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep* 1997;20(9):705-6.
- 12) Standards of Practice Committee of the American Sleep Disorders Association. Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. *Sleep* 1994;17(4):372-7.
- 13) Levendowski DJ, Zack N, Rao S, Wong K, Gendreau M, Kranzler J, et al. Assessment of the test-retest reliability of laboratory

- polysomnography. *Sleep Breath* 2009;13(2):163-7.
- 14) Gouveris H, Selivanova O, Bausmer U, Goepel B, Mann W. First-night-effect on polysomnographic respiratory sleep parameters in patients with sleep-disordered breathing and upper airway pathology. *Eur Arch Otorhinolaryngol* 2010;267(9):1449-53.
 - 15) Chediak AD, Acevedo-Crespo JC, Seiden DJ, Kim HH, Kiel MH. Nightly variability in the indices of sleep-disordered breathing in men being evaluated for impotence with consecutive night polysomnograms. *Sleep* 1996;19(7):589-92.
 - 16) Kim JH, Yoo KS, Choi SH, Kim JC, Lee YS. Comparative analysis of sleepstrip versus laboratory polysomnography in the diagnosis of sleep disordered breathing. *Korean J Otolaryngol* 2001;44(9):940-5.
 - 17) Kim SH, Eun YG, Kim MG, Cho JS, Kim SW. Clinical value of ApneaLinkTM in the diagnosis of obstructive sleep apnea syndrome. *Korean J Otorhinolaryngol-Head Neck Surg* 2008;51(10):878-82.
 - 18) Chung H, Shin SY, Lee KH, Cho JS, Kim SW. Validation of embietta for diagnosis of obstructive sleep apnea. *Korean J Otorhinolaryngol-Head Neck Surg* 2010;53(8):480-5.
 - 19) Coates TJ, George JM, Killen JD, Marchini E, Hamilton S, Thorensen CE. First night effects in good sleepers and sleep-maintenance insomniacs when recorded at home. *Sleep* 1981;4(3):293-8.
 - 20) Levendowski D, Steward D, Woodson BT, Olmstead R, Popovic D, Westbrook P. The impact of obstructive sleep apnea variability measured in-lab versus in-home on sample size calculations. *Int Arch Med* 2009;2(1):2.
 - 21) Chediak AD, Acevedo-Crespo JC, Seiden DJ, Kim HH, Kiel MH. Nightly variability in the indices of sleep-disordered breathing in men being evaluated for impotence with consecutive night polysomnograms. *Sleep* 1996;19(7):589-92.
 - 22) Johns MW. A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep* 1991;14(6):540-5.
 - 23) Sharpley AL, Solomon RA, Cowen PJ. Evaluation of first night effect using ambulatory monitoring and automatic sleep stage analysis. *Sleep* 1988;11(3):273-6.