

Significance of Including a Surrogate Arousal for Sleep Apnea-Hypopnea Syndrome Diagnosis by Respiratory Polygraphy

Juan F. Masa, MD^{1,9}; Jaime Corral, MD^{1,9}; Javier Gomez de Terreros¹; Joaquin Duran-Cantolla, MD^{2,9}; Marta Cabello³; Luis Hernández-Blasco⁴; Carmen Monasterio, MD^{5,9}; Alberto Alonso, MD^{6,9}; Eusebi Chiner, MD⁷; Felipe Aizpuru^{2,9}; Jose Zamorano¹; Ricardo Cano¹; Jose M. Montserrat, MD^{8,9}; Collaborating group: Estefania Garcia-Ledesma¹; Ricardo Pereira¹; Laura Cancelo²; Angeles Martinez³; Lirios Sacristan⁴; Neus Salord^{5,9}; Miguel Carrera⁶; José N. Sancho-Chust⁷; Cristina Embid⁸

¹San Pedro de Alcantara Hospital, Cáceres, Spain; ²Txagoritzu Hospital, Vitoria, Spain; ³Valdecilla Hospital, Santander, Spain; ⁴General Universitario Hospital, Alicante, Spain; ⁵Belvitge Hospital, Barcelona, Spain; ⁶Son Dureta Hospital, Palma de Mallorca, Spain; ⁷San Juan Hospital, Alicante, Spain; ⁸Clinic Hospital, Barcelona, Spain; ⁹CIBER de Enfermedades Respiratorias (CIBERES), Madrid, Spain

Rationale: Respiratory polygraphy is an accepted alternative to polysomnography (PSG) for sleep apnea/hypopnea syndrome (SAHS) diagnosis, although it underestimates the apnea-hypopnea index (AHI) because respiratory polygraphy cannot identify arousals.

Objectives: We performed a multicentric, randomized, blinded crossover study to determine the agreement between home respiratory polygraphy (HRP) and PSG, and between simultaneous respiratory polygraphy (respiratory polygraphy with PSG) (SimultRP) and PSG by means of 2 AHI scoring protocols with or without hyperventilation following flow reduction considered as a surrogate arousal.

Methods: We included suspected SAHS patients from 8 hospitals. They were assigned to home and hospital protocols at random. We determined the agreement between respiratory polygraphy AHI and PSG AHI scorings using Bland and Altman plots and diagnostic agreement using receiver operating characteristic (ROC) curves. The agreement in therapeutic decisions (continuous positive airway pressure treatment or not) between HRP and PSG scorings was done with likelihood ratios and post-test probability calculations.

Results: Of 366 randomized patients, 342 completed the protocol. AHI from HRP scorings (with and without surrogate arousal) had similar agreement with PSG. AHI from SimultRP with surrogate arousal scoring had better agreement with PSG than AHI from SimultRP without surrogate arousal. HRP with surrogate arousal scoring had slightly worse ROC curves than HRP without surrogate arousal, and the opposite was true for SimultRP scorings. HRP with surrogate arousal showed slightly better agreement with PSG in therapeutic decisions than for HRP without surrogate arousal.

Conclusion: Incorporating a surrogate arousal measure into HRP did not substantially increase its agreement with PSG when compared with the usual procedure (HRP without surrogate arousal).

Keywords: Sleep apnea, portable monitor, respiratory polygraphy, and CPAP treatment

Citation: Masa JF; Corral J; Gomez de Terreros J; Duran-Cantolla J; Cabello M; Hernández-Blasco L; Monasterio C; Alonso A; Chiner E; Aizpuru F; Zamorano J; Cano R; Montserrat JM. Significance of including a surrogate arousal for sleep apnea-hypopnea syndrome diagnosis by respiratory polygraphy. *SLEEP* 2013;36(2):249-257.

INTRODUCTION

Sleep apnea and hypopnea syndrome (SAHS) is a prevalent disorder,¹ which increases the risk of hypertension,² cardiovascular mortality,³ and traffic accidents.^{4,5} Although polysomnography (PSG) is assumed to be the gold standard for SAHS diagnosis, respiratory polygraphy using type 3 portable monitors⁶ is an accepted cost-effective alternative for at-home diagnosis.^{7,8} The latter is less time-consuming, favoring better diagnostic access and reducing waiting lists.^{9,10}

The main difference between PSG and respiratory polygraphy is that the latter does not include neurological variables, and therefore: (1) the number of apneas and hypopneas must be divided by recording time instead of sleep time (the method used with PSG); this results in a systematic underestimation of the AHI; and (2) respiratory polygraphy cannot identify arousal and consequently cannot incorporate this criterion into the hypopnea definition.¹¹

The American Academy of Sleep Medicine (AASM) recommends 2 alternative definitions of hypopnea, either including or excluding an arousal.⁷ Some large studies of adding an arousal measure to the hypopnea definition resulted in an increase up to 300% in the hypopnea number.^{12,13} The definition that includes an arousal is predominant in PSGs performed in clinical practice; therefore, including a surrogate arousal measure in respiratory polygraphy AHI scoring may improve agreement with PSG. However, limited research has been carried out on this issue.^{14,15}

In a PSG recording, a hypopnea is habitually followed by a hyperventilation episode concurrent with arousal. We hypothesized that using this hyperventilation as a marker of surrogate arousal might increase the agreement between PSG and respiratory polygraphy.

We previously studied the diagnostic cost-effectiveness and efficacy of therapeutic decision-making with manual home respiratory polygraphy (HRP) versus PSG in a large multicenter, randomized, blinded crossover study. We found good diagnostic cost-effectiveness⁸ and agreement in therapeutic decision-making, but, for the latter, only in the subpopulation with an AHI > 30.¹⁶ Using this cohort, we have now attempted to determine different aspects of the agreement between HRP and in-hospital PSG, and between simultaneous respiratory polygraphy (respiratory polygraphy with PSG) (SimultRP) and in-hospital PSG by means of two AHI scoring protocols—one

Submitted for publication March, 2012

Submitted in final revised form August, 2012

Accepted for publication August, 2012

Address correspondence to: Juan F. Masa, MD, San Pedro de Alcantara Hospital, Cáceres, Spain; Tel: 0034927256296; Fax: 0034927256297; E-mail: fmasa@separ.es

including final hyperventilation defined as an surrogate arousal, and the other not including this surrogate arousal.

METHODS

Subjects

We sequentially included patients between 18 and 70 years old referred to 8 hospitals in Spain for suspected SAHS because of snoring, observed apneas, sleepiness (Epworth Sleepiness Scale score > 10), or morning tiredness. Patients with other suspected sleep disorders were not included. We excluded patients with severe and unstable heart disease, those who were incapable of setting up the respiratory polygraphy device in a trial, and those who refused to participate in the study. The ethics committees of the 8 participating centers approved the study protocol. All patients provided written informed consent.

Protocol

All patients performed at-home and hospital protocols in a random order. PSG and respiratory polygraphy scorings were completed separately and the technicians, and physicians were blinded to any identifying patient information and any previous results.

Respiratory Polygraphy at Home

The HRP (BreastSC20; Breast Medical AB; Mölnlycke, Sweden) measured: oxygen saturation (model 8000 J; Nonin Medical; Plymouth, MN); airflow by nasal cannula; thoracic and abdominal movements by piezoelectric bands (Pro-Tech reference 1295; Respironics; Pittsburgh, PA); and body position.

Before randomization, a technician provided instructions for home use of the HRP device, and a test was performed in the hospital setting. The HRP instruments were moved between residences by staff from CPAP service companies in each hospital area, acting as transport companies. These transport services did not provide any help to the patients in setting up the HRP devices. The raw data files were telematically sent from the home to the hospital.⁸

The hospital technician downloaded the data file from the trial web site and manually scored the raw data using 2 scoring protocols: HRP-desaturation (HRP-D) and HRP-surrogate arousal (HRP-SA) (see definitions below).

PSG and SimultRP In-Hospital

Polysomnographic recordings were analyzed manually at each participating center, based on standard criteria.¹⁷⁻¹⁹ Neurological variables were electroencephalogram, electrooculogram, and electromyogram. Flow tracing was measured using a nasal cannula and thoracoabdominal movement with thoracic and abdominal bands. Oxygen saturation was assessed by a pulse oximeter. The nasal cannula was divided by a Y connection to PSG and SimultRP. Independent sensors measured the other SimultRP parameters. The scoring of respiratory events in SimultRP was done with the same criteria as in HRP, which included 2 scoring protocols: SimultRP-desaturation (SimultRP-D) and SimultRP-surrogate arousal (SimultRP-SA) (see definitions below).

Variables and Definitions

We collected information on the following variables using questionnaires and direct measurements: age, sex, body mass

index, neck circumference, systolic and diastolic blood pressure, comorbidities, job, alcohol intake, and tobacco consumption. Subjective nocturnal sleep time and napping time were ascertained based on recall of the previous 4 weeks. Episodes of subjective asphyxia, apneas and snoring frequency, nocturia, morning headache, morning tiredness, and sleepiness while driving⁵ were collected in 4 degrees of intensity (never, sometimes, frequently and always) based on the previous 4 weeks. Epworth and ASDA (American Sleep Disorders Association)²⁰ sleepiness scales were measured based on the previous 4 weeks. PSG reports included recording time, sleep time, sleep periods, AHI (total sleep time and sleep time in the supine position), arousal and desaturation indexes, and time with SatO₂ < 90%; and HRP data included recording and valid recording time, AHI (total valid recording time and valid time in the supine position), desaturation index, and time with SatO₂ < 90%.

PSGs or respiratory polygraphies were considered valid if they contained ≥ 3 recorded hours. In addition, a valid respiratory polygraphy included ≥ 3 h of flow or band and oximetry measurements for scoring because we excluded recording time with a bad signal from these measurements. An invalid recording could be repeated up to two times.

For PSGs, an apnea was defined as the absence of airflow ≥ 10 sec and a hypopnea as a discernible airflow or band reduction > 30% and < 90%, of ≥ 10 sec duration and with $\geq 3\%$ drop in oxygen saturation or final arousal.²¹ The definition of apnea for the respiratory polygraphy was the same, but hypopneas were scored using 2 definitions. For in-home or simultaneous “respiratory polygraphy-desaturation” scorings, we considered hypopneas with desaturation; and for in-home or simultaneous “respiratory polygraphy-surrogate arousal” scorings, we considered hypopneas with desaturation or surrogate arousal. Surrogate arousal for HRP or SimultRP was a clear resolution of airflow or band reduction by a sudden increase in amplitude and frequency ≥ 2 breaths (Figure 1). The criteria for hypopnea were the same amplitude reduction and duration on flow or bands as for PSG and $\geq 3\%$ drop in oxygen saturation (desaturation criterion) or surrogate arousal.

In PSG, the number of apneas and hypopneas was divided by sleep time to obtain the AHI. In HRP and SimultRP, the number of apneas and hypopneas was divided by valid time (recording time excluding invalid time).

Post Hoc Analysis

Therapeutic Decision for HRP

When each patient completed the 2 branches of the study, we made a therapeutic decision based on PSG results. At the end of the study we carried out a blinded post hoc analysis to determine the agreement in therapeutic decisions based on PSG and HRP scorings (HRP-D and HRP-SA). No information about the “real” therapeutic decision was taken into account for this post hoc analysis.

Therapeutic decisions (CPAP, no CPAP, or impossible decision) for HRP-D, HRP-SA, and PSG were made by an investigator at each center (always the same investigator), based on the same set of clinical variables collected from each patient at baseline, shown in Table 1 (less apnea and snoring frequency) as well as PSG and HRP scoring variables included in Table 2.¹⁶

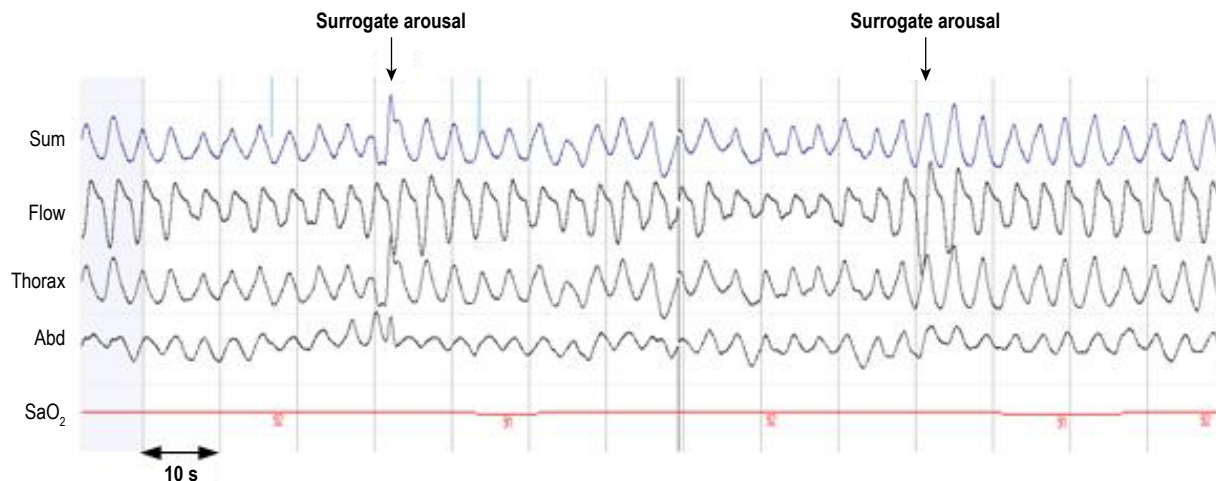


Figure 1—Polysomnographic fragment from a woman patient included in the study. The arrows mark the hyperventilation after clear hypoventilation longer than 10 s. (hypopnea) without oxygen desaturation criterion. With visible neurological channels, both episodes coincide with cortical arousals. Sum, sum of bands; Flow, nasal pressure; Thorax, thoracic band; Abd, abdominal band; SaO₂, oxygen saturation.

Patients and diagnostic methods (PSG, HRP-D, and HRP-SA) were extracted at random from an electronic database. When the researcher selected 1 of the 3 options (CPAP, no CPAP, or impossible decision), data from another patient were offered at random to repeat the process. Each patient was presented 3 times (with the PSG, HRP-D or HRP-SA information) at random and non-consecutively. Participant identification numbers for patients and other data were hidden. After a month, the same therapeutic decision procedure was repeated.

The criteria for recommending CPAP were $AHI \geq 30$, or AHI between 5 and 30 with significant symptoms or consequences according to the Spanish Sleep Network guidelines.^{16,21}

Post-Flow Reduction Hyperventilation as an Arousal

In order to determine the agreement between our surrogate arousal measure in the respiratory polygraphy scorings and cortical arousal scoring from PSG, at the end of the study each center re-analyzed the PSG results from 5 patients chosen at random, to identify and compare hypopneas identified with surrogate arousal only (without desaturation), based on only flow or bands channels, and hypopneas identified only with cortical arousal (without desaturation) based on flow or bands and neurologic channels with the following protocol: (a) the marks of previous apneic events were moved to a neurological channel (i.e., C2-A1) and deleted from the respiratory channels; hypopneas with arousal criterion only (without desaturation) were identified and marked in another neurological channel; (b) then, all neurological channels were hidden (the scoring of the sleep periods in each epoch was maintained); and (c) new scoring of hypopneas with the surrogate arousal criterion (without the desaturation criterion) was carried out with respiratory channels only (similar to respiratory polygraphy).

Statistical Analysis

The agreement between PSG and the 4 respiratory polygraphy scorings (HRP-D, HRP-SA, SimultRP-D, and SimultRP-SA) was determined with the following analysis:

1. To determine the agreement of AHI scores between the PSG and both respiratory polygraphy (HRP and SimultRP) diagnostic methods and scorings (desaturation and surrogate arousal), we used Bland and Altman plots.
2. We constructed separate receiver operating characteristic (ROC) curves including the 4 scorings (HRP-D, HRP-SA, SimultRP-D, and SimultRP-SA) for the different polysomnographic AHI cutoff points (≥ 5 , ≥ 10 , and ≥ 15) of the SAHS diagnosis to determine the best diagnostic agreement.
3. To evaluate the agreement in the therapeutic decision between PSG and both HRP-D and HRP-SA scorings we used: (a) sensitivity and specificity; (b) negative (1-sensitivity/specificity) and positive (sensitivity/1-specificity) likelihood ratios (LR); (c) we computed the post-test probability of obtaining a true positive when the test was positive (probability of recommending CPAP in agreement with PSG) or a true negative when the test was negative (probability of not recommending CPAP in agreement with PSG), based on the pre-test probability of recommending CPAP (percentage of the total cases with CPAP recommendation based on PSG) and positive and negative LRs²²; and (d) agreement level (100 minus the sum of false positives and negative percentages). We utilized the same analysis with the 2 therapeutic decisions (before and after a month) based on PSG results. These PSG results were considered the “Reference” values, as they state the variability of the gold standard (PSG) and the “ideal” outcome for HRP scorings.¹⁶

In the post hoc analysis, the agreement between hypopneas with cortical arousal and surrogate arousal scorings in PSG was assessed using Bland and Altman plots.

RESULTS

Initially, 377 patients were selected. Eleven were excluded due to lack of informed consent (2), severe heart disease (2), or a failed respiratory polygraphy trial (7). Of the 366 randomized patients, 24 (7%) had no valid PSG or respiratory polygraphy tests or scorings.

Table 1—Characteristics of the studied population

		N = 342
Male, %		75.1
Age, years; mean (SD)		48.7 (11.8)
Body mass index; mean (SD)		31.0 (6.6)
Neck circumference; cm (SD)		40.9 (4.2)
Alcohol, g; mean (SD)		8.8 (22.1)
Smokers, %		24.4
Subjective sleep time/day, h; mean (SD)		6.9 (1.4)
Subjective nap time, h; mean (SD)		0.5 (0.6)
Epworth Sleepiness Scale; mean (SD)		11.7 (5.0)
Job with accident risk, %		10.6
ASDA scale, %	Absence	13.1
	Mild	28.8
	Moderate	43.8
	Severe	14.4
Sleepiness while driving, %	Never	61.1
	Sometimes	23.9
	Frequently	12.7
	Always	1.3
Snoring, %	Never	4.9
	Sometimes	7.5
	Frequently	21.9
	Always	65.7
Observed apneas, %	Never	24.8
	Sometimes	30.4
	Frequently	20.9
	Always	23.9
Subjective asphyxia episodes, %	Never	45.4
	Sometimes	36.6
	Frequently	13.4
	Always	4.6
Nocturia, %	Never	30.7
	Sometimes	31.7
	Frequently	16.7
	Always	21.9
Morning headache, %	Never	47.7
	Sometimes	31.4
	Frequently	12.4
	Always	8.5
Morning tiredness, %	Never	14.4
	Sometimes	24.5
	Frequently	26.8
	Always	34.3
Depression, %		23.5
Hypertension, %		30.9
Cardiac disease, %		3.2
Cerebrovascular disease		1.5
Systolic pressure, mm Hg (mean and SD)		130.2 (17.3)
Diastolic pressure, mm Hg (mean and SD)		75.9 (12.2)

SD, standard deviation.

The clinical and anthropometric characteristics of the 342 patients with valid PSG, HRP, and SimultRP results are shown in Table 1. The sample was predominantly male, middle-aged,

obese, and with excessive sleepiness. Respiratory polygraphy AHI values were lower than PSG values, except for SimultRP-SA scoring (Table 2); the surrogate arousal scoring AHI values were higher than desaturation scorings for both HRP and SimultRP scoring protocols.

The difference in AHI values was lower between PSG and SimultRP scorings than between PSG and HRP scorings (Figure 2). HRP-SA and SimultRP-SA scorings had lower mean AHI differences with PSG than did HRP-D and SimultRP-D scorings. However, there were similar agreement limits (2 SD) between PSG and both HRP scorings (HRP-D and HRP-SA) and slightly lower agreement limits for SimultRP-SA compared to SimultRP-D scorings. Similar results were found in the subpopulation with respiratory polygraphy AHI between 5 and 30.

Figure 3 shows the difference in AHI between HRP-SA and HRP-D (Panel A) as well as between SimultRP-SA and SimultRP-D (Panel B), all versus disease severity (AHI from PSG). These differences tended to be higher for low and intermediate AHI values as well as for patients who had any change in their AHI once hypopneas with surrogate arousal were added to the hypopnea with desaturation definition. This effect tended to be more appreciable for SimultRP than for HRP.

Figure 4 shows the ROC curves for the 4 respiratory polygraphy scorings with different polysomnographic AHI cutoff points (≥ 5 , ≥ 10 , and ≥ 15) for SAHS diagnosis. The best areas under the curve were obtained with SimultRP scorings rather than HRP scorings. Better area under the curve was observed for SimultRP-SA compared to SimultRP-D, and marginally worse area under the curve was seen for HRP-SA than HRP-D. Similar results were found in the subpopulation with respiratory polygraphy AHI scores between 5 and 30.

Therapeutic Decision-Making for HRP

The “impossible decision” situation was not observed with either PSG or HRP scorings. In the total sample, the probability of recommending CPAP in agreement with PSG was similar in HRP-D and HRP-SA since the pre-test to post-test probabilities increased from 67% to 88% in both (Table 3). The probability of not recommending CPAP in agreement with PSG was slightly better for HRP-SA, decreasing from 67% (pre-test) to 33% for HRP-SA, and to 40% for HRP-D. Consequently, the agreement levels were similar, although with a tendency to higher agreement for HRP-SA (Figure 5). Both HRP-SA and HRP-D were very different from the Reference values.

In the subsamples with AHI > 30 for HRP-D and HRP-SA (Table 3), the probability of recommending CPAP in agreement with PSG was similar, increasing from 94% (pre-test) to 96% for HRP-D, and from 92% to 93% for HRP-SA. The probability of not recommending CPAP in agreement with PSG was slightly better for HRP-SA, decreasing from 92% (pre-test) to 50% (post-test), and from 94% to 66% for HRP-D. The agreement level for HRP-SA was only 2 points higher than for HRP-D, and both were close to the Reference (Figure 5).

In patients with HRP-D and HRP-SA AHIs between 5 and 30 (Table 3), the probability of recommending CPAP in agreement with PSG and the probability of not recommending CPAP in agreement with PSG changed in a similar way, and the differences between the agreement levels and the Reference were similar.

Table 2—Sleep study variables for PSG, HRP, and SimultRP (desaturation and surrogate arousal scorings)

	PSG	HRP		SimultRP	
		Desaturation	Surrogate Arousal	Desaturation	Surrogate Arousal
Time recording, min	443 (43)	428 (81)	—	446 (48)	—
Sleep/valid time, min	371 (65)	402 (701)	—	433 (49)	—
Light sleep time, %	64 (22)	—	—	—	—
Deep sleep time, %	17 (12)	—	—	—	—
REM sleep time, %	17 (13)	—	—	—	—
Arousal index	39 (25)	—	—	—	—
AHI	38 (29)	31 (24)	34 (24)	34 (27)	37 (27)
AHI supine	44 (32)	38 (28)	40 (28)	37 (29)	39 (29)
CT 90%	9.9 (19)	6.5 (14)	—	6.5 (12)	—
Desaturation index	27 (29)	29 (24)	—	30 (26)	—

PSG, polysomnography; HRP, home respiratory polygraphy; SimultRP, simultaneous respiratory polygraphy; AHI, apnea-hypopnea index; CT 90%, percentage of sleep or register time under 90% oxygen saturation.

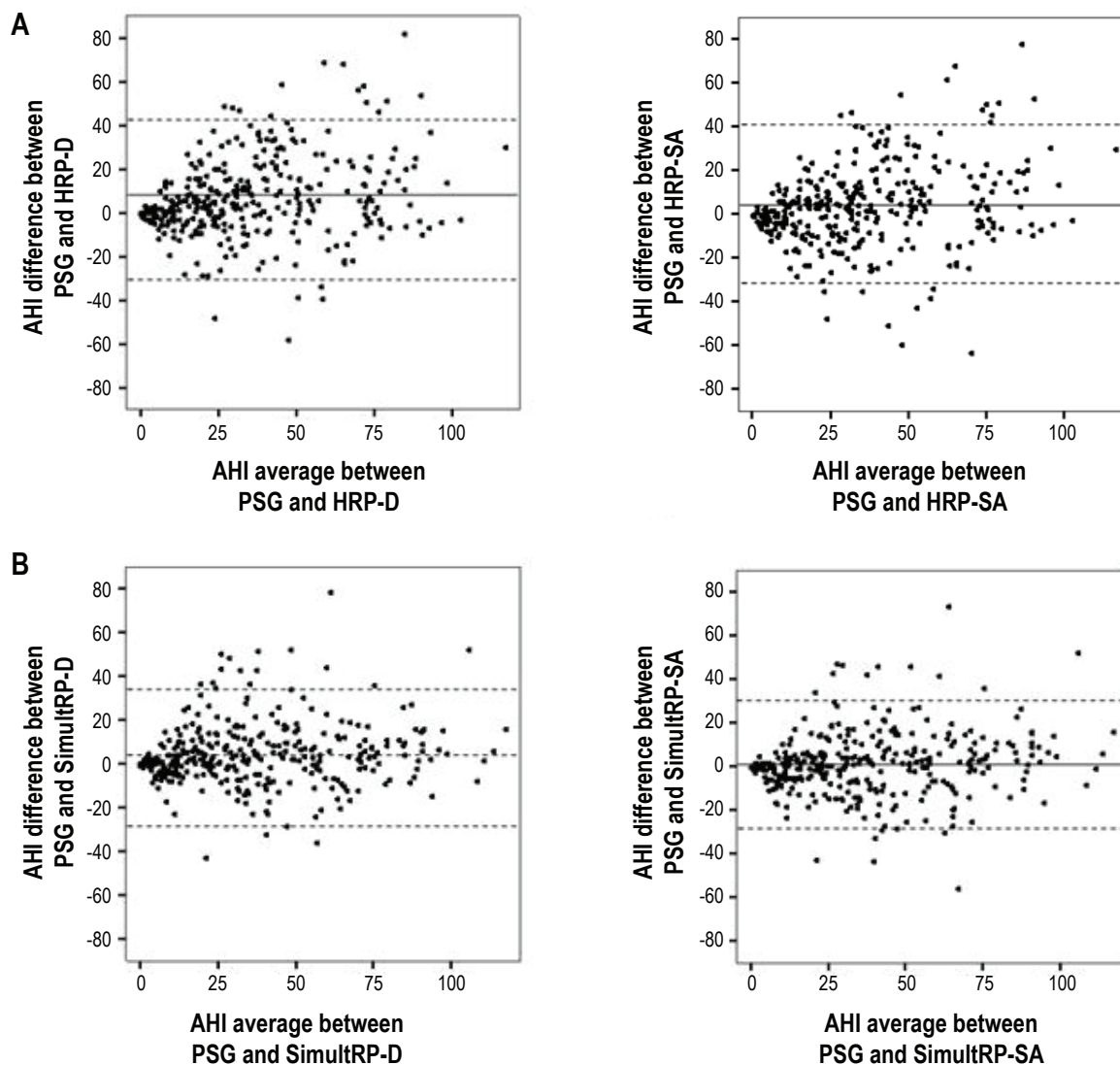
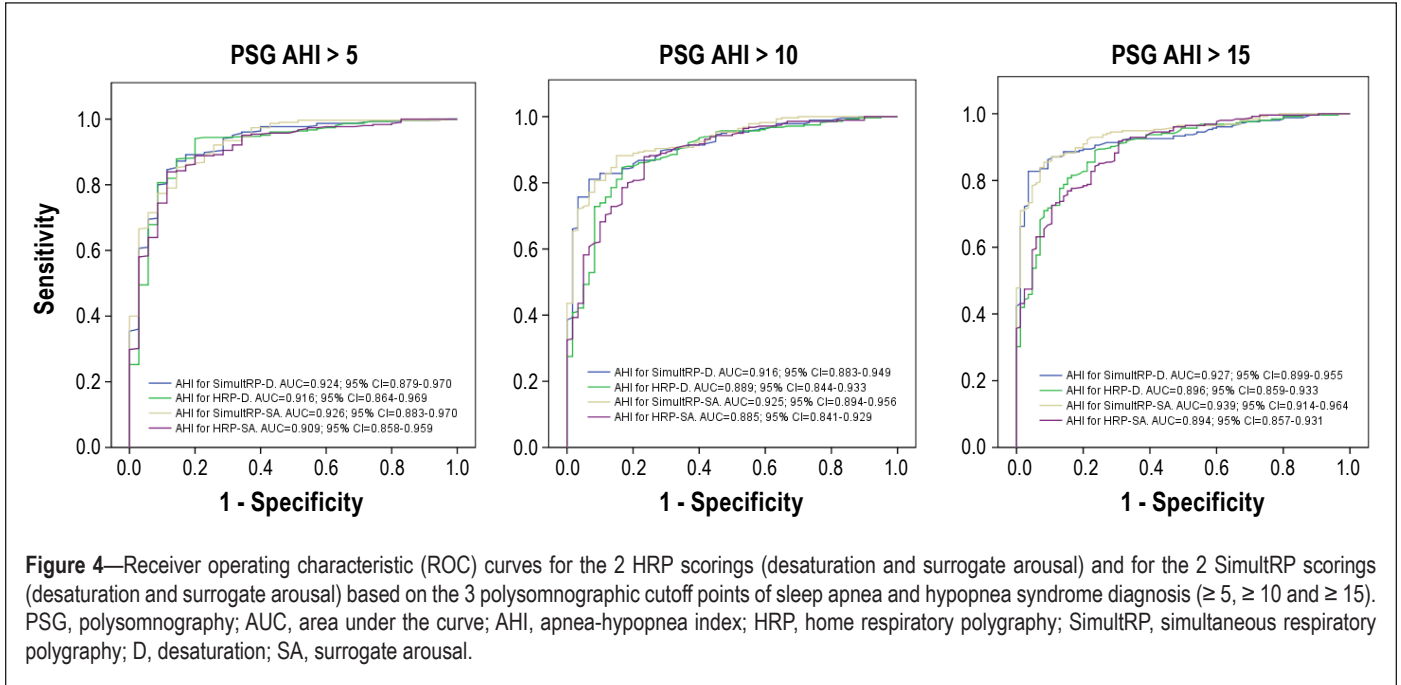
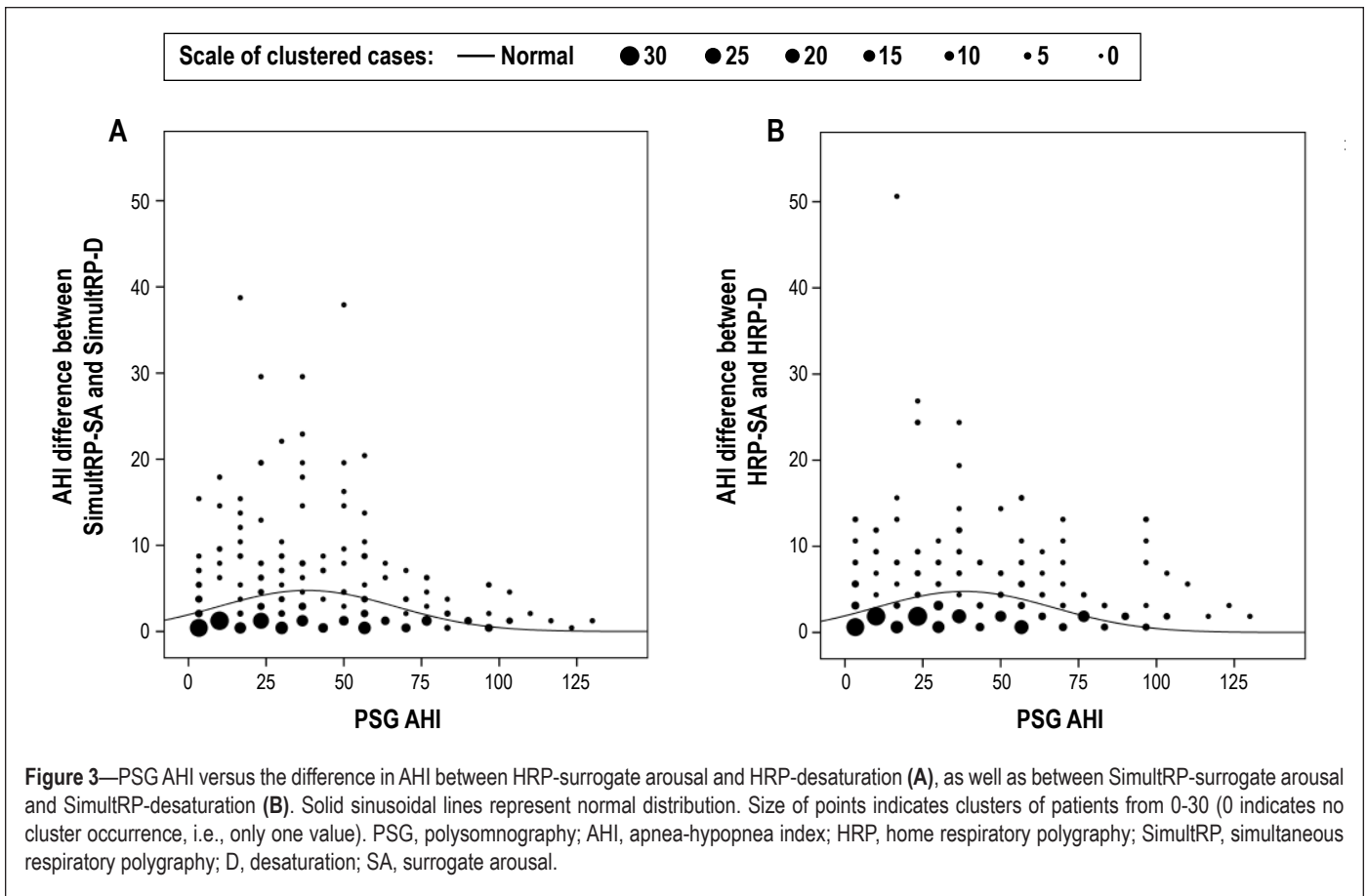


Figure 2—Mean AHI versus the difference in AHI between PSG and both HRP (desaturation and surrogate arousal scorings) (A) and SimultRP (desaturation and surrogate arousal scoring) protocols (B) (Bland-Altman plots). Central lines represent mean values and upper and lower discontinuous lines represent agreement limits (2 SD). PSG, polysomnography; AHI, apnea-hypopnea index; HRP, home respiratory polygraphy; SimultRP, simultaneous respiratory polygraphy; D, desaturation; SA, surrogate arousal.



Post-Hypopnea Hyperventilation as an Arousal

Anthropometric, clinical, and polysomnographic characteristics from 40 patients included in this post hoc analysis are shown in Table S1 (supplemental material). The number of PSG apneas and hypopneas were 3,397 and 4,866, respectively (desaturation or cortical arousal criterion). Some 1,143 hypopneas with the cortical arousal criterion only and 1,417 with the

surrogate arousal criterion only were identified. The hypopnea index between the 2 hypopnea scorings (cortical and surrogate arousals) showed good agreement (Figure 6)

DISCUSSION

To our knowledge, this cohort is the largest sample of all studies of agreement between PSG and respiratory polygraphy.

Table 3—Sensitivity, specificity, likelihood ratios, post-test probabilities, and agreement levels between HRP-desaturation (HRP-D) and HRP-surrogate arousal (HRP-SA) and PSG, and between PSG before and after a month (Reference)

	Se	Sp	LR + (95% CI)	Post-test prob. + (95% CI)	LR – (95% CI)	Post test prob. – (95% CI)	AL%
Total sample (N = 342; pretest probability 67%)							
HRP-D	74	79	3.59 (2.48-5.21)	88 (84-92)	0.32 (0.25-0.41)	40 (34-46)	76
HRP-SA	81	76	3.44 (2.45-4.84)	88 (84-91)	0.24 (0.18-0.23)	33 (27-41)	80
Ref	93	89	8.57 (5.02-15)	95 (91-97)	0.08 (0.05-0.13)	14 (9-21)	92
HRP-desaturation AHI ≥ 30 (N = 143; pretest probability 94%)							
HRP-D	94	44	1.69 (0.94-3.04)	96 (93-98)	0.13 (0.05-0.36)	66 (43-84)	91
Ref	98	78	4.40 (1.30-15)	98 (95-100)	0.03 (0.01-0.09)	31 (13-57)	97
HRP-surrogate arousal AHI ≥ 30 (N = 168; pretest probability 92%)							
HRP-SA	98	21	1.25 (0.95-1.64)	93 (91-95)	0.09 (0.02-0.41)	50 (18-82)	93
Ref	95	79	4.42 (1.62-12)	98 (95-99)	0.07 (0.03-0.14)	44 (25-61)	97
HRP-desaturation AHI between 5 and 30 (N = 166; pretest probability 55%)							
HRP-D	50	77	2.18 (1.37-3.46)	73 (63-81)	0.65 (0.51-0.82)	45 (39-50)	62
Ref	86	88	7.06 (3.81-13)	90 (83-94)	0.13 (0.10-0.24)	17 (11-25)	87
HRP-surrogate arousal AHI between 5 and 30 (N = 152; pretest probability 49%)							
HRP-SA	49	81	2.53 (1.52-4.21)	71 (60-80)	0.63 (0.49-0.81)	38 (32-44)	65
Ref	89	90	8.60 (4.44-17)	89 (81-94)	0.12 (0.06-0.23)	10 (6-18)	89

AHI, apnea-hypopnea index; HRP-D, home respiratory polygraphy desaturation; HRP-SA, home respiratory polygraphy with surrogate arousal; Ref, Reference; PSG, polysomnography; Se, sensitivity; Sp, specificity; LR, likelihood ratio; CI, confidential interval; Post test prob. +, the post-test probability of obtaining a true positive diagnosis when the test (home respiratory polygraphy) was positive; Post test prob. –, the post-test probability of obtaining a true positive diagnosis when the test (home respiratory polygraphy) was negative; AL, agreement level (100 minus the sum of false positives and negative percentages).

It is the only multicenter study, and the effect of adding a surrogate arousal to respiratory polygraphy hypopnea scoring has not been extensively investigated. The principal results were: (a) AHI from SimultRP scorings had lower differences from PSG than HRP scorings. AHI from SimultRP-SA scoring has slightly lower differences from PSG than SimultRP-D scoring. However, HRP-SA and HRP-D scorings had similar AHI differences; (b) SAHS diagnostic agreement was slightly better for respiratory polygraphy-surrogate arousal scoring than for respiratory polygraphy-desaturation scoring in the SimultRP protocol and similar in the HRP protocol; (c) therapeutic decisions were similar with HRP-SA and HRP-D scorings, although HRP-SA tended to be closer to Reference; and (d) in PSG, hypopneas with surrogate arousal agreed well with hypopneas with cortical arousal.

By consensus, the Spanish Sleep Network includes an arousal in the PSG hypopnea definition in clinical practice, but logically, the respiratory polygraphy-hypopnea definition does not include arousal.²¹ This apparent incongruence, also present in guidelines from other countries, led us to investigate this issue in a large sample of patients by studying different aspects of the agreement between PSG and respiratory polygraphy.

In a study, Ayappa et al. ascertained the agreement between a PSG and respiratory polygraphy scoring, with hypopnea defined as an airflow reduction of ≤ 50% and desaturation of 1% plus some surrogate arousal indicators (head movement, changes in snoring, or change in pulse rate).¹⁴ Agreement between PSG in AHI seen in the Bland and Altman plots was very similar to the agreement in our study for the SimultRP and HRP protocols (diagnostic and therapeutic decision agreements were not tested). Another study compared SimultRP and PSG (HRP was not studied) and used a definition of surrogate arousal similar to

the Ayappa study.¹⁵ AHI differences between PSG and the two types of SimultRP scoring (with and without surrogate arousal) were similar when examined with Bland and Altman plots. Diagnostic agreement was slightly worse for SimultRP-SA scoring (therapeutic decision agreement was not tested). Our study had methodological differences with both studies mentioned above. Our study had five¹⁴ and more than two¹⁵ times the sample size, respectively, a multicenter design, and checked SimultRP and HRP with several agreement measures as well as different SA definitions. Unlike previous studies, we tested and found that our surrogate arousal showed good agreement with polysomnographic cortical arousal. However, our results showed slightly better agreement between PSG and SimultRP-SA than between PSG and SimultRP-D; these results apparently contradict the results of the last study mentioned, but they do not have very distinct meanings—both indicate that adding surrogate arousal to respiratory polygraphy does not produce substantially different results from respiratory polygraphy-desaturation scoring.

Different studies have compared the agreement between HRP (desaturation scoring) and PSG. No clear differences in agreement between studies including polysomnographic hypopneas with a cortical arousal criterion^{8,23-25} and others without a cortical arousal criterion^{14,26-28} were observed on Bland and Altman plots. Similar results were found in studies comparing SimultRP (desaturation scoring) with PSG using hypopneas with a cortical arousal criterion^{15,24,25,30-32} and others without a cortical arousal criterion.^{14,27,28} Our present results, including a surrogate arousal in respiratory polygraphy scoring that shows only a small benefit, are consistent with these previous findings.

Adding hypopneas identified by the surrogate arousal criterion decreased the mean AHI differences between PSG and RP (HRP and SimultRP) scorings (Figure 1). This was the ex-

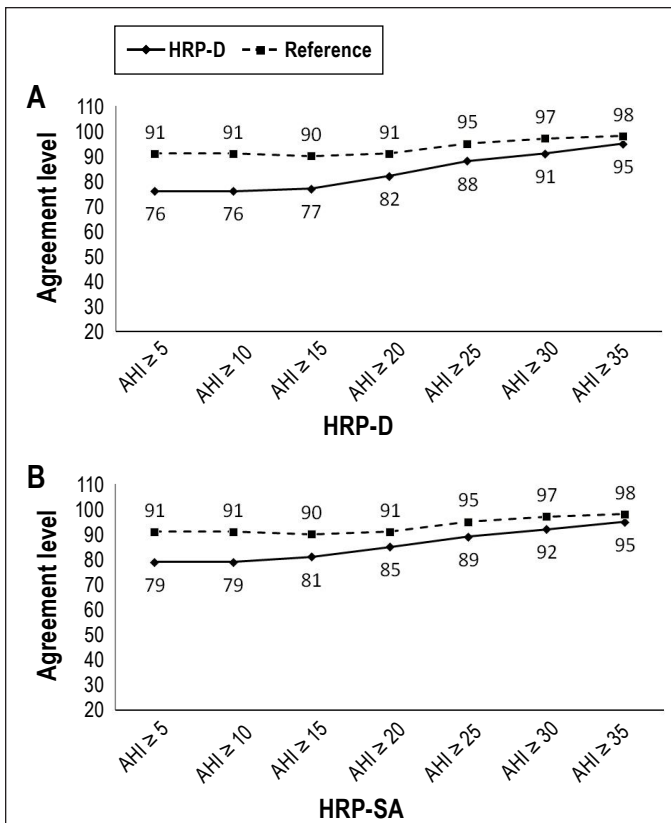


Figure 5—(A) Evolution of the agreement level in therapeutic decisions made (100 - sum of true positives and negatives) between HRP-D and PSG and the Reference, selecting patients by increasing severity according to HRP-D AHI. (B) Evolution of the agreement level in therapeutic decisions (100 - sum of true positives and negatives) made between HRP-SA and PSG and the Reference, selecting patients by increasing severity according to HRP-SA AHI. HRP-D, home respiratory polygraphy with desaturation criterion; HRP-SA, home respiratory polygraphy with surrogate arousal criterion; PSG, polysomnography; AHI, apnea plus hypopnea index; Reference, agreement between before and after a month of the PSG.

pected result, since hypopneas with surrogate arousal criterion were added to hypopneas-desaturation scoring, and the mean AHI from respiratory polygraphy is normally lower than the mean AHI from PSG. However, the agreement limits of the AHI difference between PSG and HRP-D and between PSG and HRP-SA were similar, and slightly better between PSG and SimultRP-SA when compared to PSG and SimultRP-desaturation. A similar result was found when examining diagnostic efficacy (Figure 4). The main difference between the PSG and SimultRP and between the PSG and HRP tests was that the HRP was performed at a different time (night) and in a different setting. This may have produced less home sleep time in a supine position^{14,24,29} and better home sleep quality and stability, leading to, among other factors, different number of obstructive events.³³ These differences could make the benefits of adding an surrogate arousal in HRP seems less evident or even absent. Our interpretation is that, although adding surrogate arousal to HRP scoring can produce modest improvement, the magnitude of the change is not sufficient to overcome other factors that produce AHI differences between PSG and HRP, resulting in no impact on clinical practice.

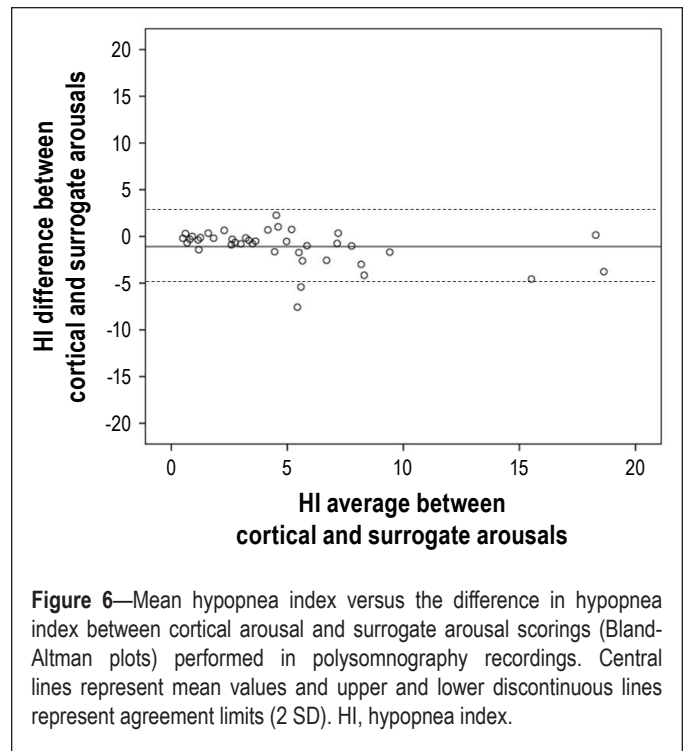


Figure 6—Mean hypopnea index versus the difference in hypopnea index between cortical arousal and surrogate arousal scorings (Bland-Altman plots) performed in polysomnography recordings. Central lines represent mean values and upper and lower discontinuous lines represent agreement limits (2 SD). HI, hypopnea index.

Several studies have focused on the significance of adding arousal to the hypopnea definition in PSG. Some of them did not find a relevant increase in the number of hypopneas,³⁴⁻³⁶ but others with larger samples showed up to 300% increases in the hypopnea number when arousal was added to the definition, especially in mild and moderate SAHS patients, who can have frequent, subtle hypopneas without desaturation due to, in part, a lower body mass index.^{12,13} Effectively, in the present study, the change in AHI occurred primarily in patients with low and intermediate values of PSG-AHI. However, the Bland and Altman analysis and diagnostic agreement for patients with low PSG-AHI values (i.e., from 5 to 30) produced similar results for the entire population, probably due to the small size of the AHI changes and the potential inclusion of false hypopneas. On the other hand, our study population had higher AHI and BMI scores than previous studies, which found large numbers of hypopneas with cortical arousal and without desaturation in PSG registers. This may explain the lower number of additional hypopneas when adding a surrogate arousal criterion in our study.

We have previously shown that therapeutic decision-making (CPAP or other treatments) for HRP is a different diagnostic process with a different agreement with PSG.¹⁶ For this reason, we have added this analysis to more globally assess the agreement between HRP-SA and PSG. Like HRP-D, HRP-SA has acceptable agreement with PSG for patients with an HRP AHI > 30. However, a pending question from the previous analysis was whether adding a surrogate arousal criterion to HRP-D scoring could improve the agreement in therapeutic decisions between PSG and HRP in the population with an HRP AHI between 5 and 30.¹¹ The present results do not show a significant improvement.

In summary, this study has showed that incorporating a surrogate arousal criterion into the definition of hypopnea in HRP did not substantially increase agreement with PSG in comparison with normal management in clinical practice (PSG scoring

with cortical arousal and HRP scoring without surrogate arousal). Therefore, no procedural modifications seem necessary.

ACKNOWLEDGMENTS

The authors are indebted to John-Paul Glutting, Verónica Rodríguez, and Vanessa Iglesias for their assistance in the preparation of the manuscript and Asunción Martín, Elena Sandoval, Soledad Guillen, Trinidad Amigo, Pablo Mejias, and Carmen Lorenzana for technical assistance in the sleep laboratory.

DISCLOSURE STATEMENT

This was not an industry supported study. Funded by: Instituto de Salud Carlos III (Fondo de Investigaciones Sanitarias, Ministerio de Sanidad y Consumo), Spanish Respiratory Society (SEPAR), Telefonica SA (Spain), Air Liquide (Spain), and Breas Medical (Spain) The authors have indicated no financial conflicts of interest.

REFERENCES

1. Durán J, Esnaola S, Rubio R, Iztueta A. Obstructive sleep apnea-hypopnea and related clinical features in a population-based sample of subjects aged 30 to 70 yr. *Am J Respir Crit Care Med* 2001;163:685-9.
2. Durán-Cantolla J, Aizpuru F, Montserrat JM, et al; Spanish Sleep and Breathing Group. Continuous positive airway pressure as treatment for systemic hypertension in people with obstructive sleep apnoea: randomised controlled trial. *BMJ* 2010;341:c5991. <http://dx.doi.org/10.1136/bmj.c5991>.
3. Marin JM, Carrizo SJ, Vicente E, Agusti AG. 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;365:1046-53.
4. Terán-Santos J, Jiménez-Gómez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents. Cooperative Group Burgos-Santander. *N Engl J Med* 1999;340:847-51.
5. Masa JF, Rubio M, Findley LJ. Habitually sleepy drivers have a high frequency of automobile crashes associated with respiratory disorders during sleep. *Am J Respir Crit Care Med* 2000;162:1407-12.
6. Ferber R, Millman R, Coppola M, et al. Portable recording in the assessment of obstructive sleep apnea. *ASDA standards of practice. Sleep* 1994;17:378-92.
7. Collop NA, Anderson WM, Boehlecke B, et al; Portable Monitoring Task Force of the American Academy of Sleep Medicine. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 2007;3:737-47.
8. Masa JF, Corral J, Pereira R, et al. Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. *Thorax* 2011;66:567-73.
9. Flemons WW, Douglas NJ, Kuna ST, Rodenstein DO, Wheatley J. Access to diagnosis and treatment of patients with suspected sleep apnea. *Am J Respir Crit Care Med* 2004;169:668-72.
10. Masa Jiménez JF, Barbé Illa F, Capote Gil F, et al; Working Group. Resources and delays in the diagnosis of sleep apnea-hypopnea syndrome. *Arch Bronconeumol* 2007;43:188-98.
11. Kimoff RJ. To treat or not to treat: can a portable monitor reliably guide decision-making in sleep apnea? *Am J Respir Crit Care Med* 2011;184:871-2.
12. Masa JF, Corral J, Teran J, et al. Apneic and obstructive non-apneic sleep respiratory events (ONEs). *Eur Respir J* 2009;34:156-61.
13. Cracowski C, Pépin JL, Wuyam B, Lévy P. Characterization of obstructive nonapneic respiratory events in moderate sleep apnea syndrome. *Am J Respir Crit Care Med* 2001;164 6:944-8.
14. Ayappa I, Norman RG, Seelall V, Rapoport DM. Validation of a self-applied unattended monitor for sleep disordered breathing. *J Clin Sleep Med* 2008;4:26-37.
15. To KW, Chan WC, Chan TO, Tung, et al. Validation study of a portable monitoring device for identifying OSA in a symptomatic patient population. *Respirology* 2009;14:270-5.
16. Masa JF, Corral J, Pereira R, et al; Spanish Sleep Network. Therapeutic decision-making for sleep apnea and hypopnea syndrome using home respiratory polygraphy: a large multicentric study. *Am J Respir Crit Care Med* 2011;184:964-71.
17. Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. EEG arousals: scoring rules and examples. *Sleep* 1992;15:174-84.
18. Iber C, Ancoli-Israel S, Chesson A, Quan S. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. 1st ed. Westchester, IL: American Academy of Sleep Medicine, 2007.
19. Rechtschaffen A, Kales A. editors. A manual of standardized terminology and scoring system for sleep stages of human subjects. Los Angeles: Brain Information Service/Brain Research Institute, University of California at Los Angeles; 1968.
20. The international classification of sleep disorders: diagnostic and coding manual diagnostic classification steering committee. American Sleep Disorders Association. Rochester, MN; 1990.
21. Grupo Español de Sueño (GES). Consenso Nacional sobre el síndrome de apneas-hipopneas del sueño. *Arch Bronconeumol* 2005;41 Supl 4:3-110.
22. Flemons WW, Littner MR, Rowley JA, et al. Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review co-sponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. *Chest* 2003;124:1543-79.
23. Parra O, García-Escasas N, Montserrat JM, et al. Should patients with sleep apnoea/hypopnoea syndrome be diagnosed and managed on the basis of home sleep studies? *Eur Respir J* 1997;10:1720-4.
24. García-Díaz E, Quintana-Gallego E, Ruiz A, et al. Respiratory polygraphy with actigraphy in the diagnosis of sleep apnea-hypopnea syndrome. *Chest* 2007;131:725-32.
25. Santos-Silva R, Sartori DE, Truksinas V, et al. Validation of a portable monitoring system for the diagnosis of obstructive sleep apnea syndrome. *Sleep* 2009;32:629-36.
26. Kushida CA, Littner MR, Hirshkowitz M, et al; American Academy of Sleep Medicine. Practice parameters for the use of continuous and bilevel positive airway pressure devices to treat adult patients with sleep-related breathing disorders. *Sleep* 2006;29:375-80.
27. Dingli K, Coleman EL, Vennelle M, et al. Evaluation of a portable device for diagnosing the sleep apnoea/hypopnoea syndrome. *Eur Respir J* 2003;21:253-9.
28. Reichert JA, Bloch DA, Cundiff E, Votteri BA. Comparison of the Nova-Som QSG, a new sleep apnea home-diagnostic system, and polysomnography. *Sleep Med* 2003;4:213-8.
29. Yin M, Miyazaki S, Ishikawa K. Evaluation of type 3 portable monitoring in unattended home setting for suspected sleep apnea: factors that may affect its accuracy. *Otolaryngol Head Neck Surg* 2006;134:204-9.
30. Calleja JM, Esnaola S, Rubio R, Durán J. Comparison of a cardiorespiratory device versus polysomnography for diagnosis of sleep apnoea. *Eur Respir J* 2002;20:1505-10.
31. Candela A, Hernández L, Asensio S, et al. Validation of a respiratory polygraphy system in the diagnosis of sleep apnea syndrome. *Arch Bronconeumol* 2005;41:71-7. Spanish.
32. Ballester E, Solans M, Vila X, et al. Evaluation of a portable respiratory recording device for detecting apnoeas and hypopnoeas in subjects from a general population. *Eur Respir J* 2000;16:123-7.
33. Stepnowsky CJ Jr, Orr WC, Davidson TM. Nightly variability of sleep-disordered breathing measured over 3 nights. *Otolaryngol Head Neck Surg* 2004;131:837-43.
34. Tsai WH, Flemons WW, Whitelaw WA, Remmers JE. A comparison of apnea-hypopnea indices derived from different definitions of hypopnea. *Am J Respir Crit Care Med* 1999;159:43-8.
35. Manser RL, Rochford P, Pierce RJ, Byrnes GB, Campbell DA. Impact of different criteria for defining hypopneas in the apnea-hypopnea index. *Chest* 2001;120:909-14.
36. Redline S, Kapur VK, Sanders MH, et al. Effect of varying approaches for identifying respiratory disturbances on sleep apnea assessment. *Am J Respir Crit Care Med* 2000;161:369-74.

Table S1—Anthropometric, clinical, and polysomnographic characteristics from 40 patients included in the post hoc analysis

	N = 40
Male, %	70
Age, years, mean (SD)	46.7 (11.2)
Body mass index, mean (SD)	30.7 (6.5)
Epworth sleepiness scale, mean (SD)	11.0 (4.6)
Depression, %	10.0
Hypertension, %	32.5
Total sleep time, min	379.4 (68.3)
Apnea-hypopnea index	33.8 (25.1)
Apnea index	14.2 (21.3)
Hypopnea index	19.6 (13.82)
Hypopnea index with only cortical arousal	4.6 (4.0)
Hypopnea index with only surrogate arousal	5.6 (4.8)