Effectiveness of sequential automatic-manual home respiratory polygraphy scoring

**Juan F. Masa*,#, Jaime Corral*,#, Ricardo Pereira*,#, Joaquin Duran-Cantolla#,¶,
Marta Cabello+, Luis Hernández-Blasco§, Carmen Monasterioƒ,
Alberto Alonso-Fernandez#,**, Eusebi Chiner##, Francisco-José Vázquez-Polo¶¶,
Jose M. Montserrat#,++ and the Spanish Sleep Group§§**

ABSTRACT: Automatic home respiratory polygraphy (HRP) scoring functions can potentially confirm the diagnosis of sleep apnoea-hypopnoea syndrome (SAHS) (obviating technician scoring) in a substantial number of patients. The result would have important management and cost implications. The aim of this study was to determine the diagnostic cost-effectiveness of a sequential HRP scoring protocol (automatic and then manual for residual cases) compared with manual HRP scoring, and with in-hospital polysomnography.

We included suspected SAHS patients in a multicentre study and assigned them to home and hospital protocols at random. We constructed receiver operating characteristic (ROC) curves for manual and automatic scoring. Diagnostic agreement for several cut-off points was explored and costs for two equally effective alternatives were calculated.

Of 366 randomised patients, 348 completed the protocol. Manual scoring produced better ROC curves than automatic scoring. There was no sensitive automatic or subsequent manual HRP apnoea–hypopnoea index (AHI) cut-off point. The specific cut-off points for automatic and subsequent manual HRP scorings (AHI >25 and >20, respectively) had a specificity of 93% for automatic and 94% for manual scorings. The costs of manual protocol were 9% higher than sequential HRP protocol; these were 69% and 64%, respectively, of the cost of the polysomnography.

A sequential HRP scoring protocol is a cost-effective alternative to polysomnography, although with limited cost savings compared to HRP manual scoring.

KEYWORDS: Automatic scoring, portable monitor, respiratory polygraphy, sleep apnoea

The prevalence of sleep apnoea–hypopnoea syndrome (SAHS) is about 2–5% in the adult population [1]. Several studies have shown associations with arterial hypertension [2, 3], cardiovascular mortality [4] and traffic accidents [5, 6].

Demand for SAHS diagnostic studies has increased in recent years, but access to subsequent diagnostic testing has been limited [7, 8]. Therefore, there has been increasing interest in alternative diagnostic processes [9].

The gold standard for SAHS diagnosis is in-hospital polysomnography (PSG), but it is time-consuming and expensive. Manual scoring of home respiratory polygraphy (HRP) is an accepted [10] and cost-effective [11–14] alternative for SAHS diagnosis in selected patients. Respiratory polygraphy involves a type 3 portable monitoring device [15], which includes sensors for airflow, respiratory effort measured with bands, and pulse oximetry recordings.

Most HRP devices offer the option of performing automatic scoring. Several studies have assessed the agreement of automatic *versus* manual scoring [16] or PSG [17–23]. The studies which assessed the agreement of automatic scoring with PSG found good apnoea–hyopnoea index (AHI) cut-off points to confirm SAHS diagnosis in a substantial number of patients, but did not find good cut-off points for ruling out the condition [17, 20]. Thus, one potential use of automatic scoring is to reduce the need for manual scoring, if automatic scoring is done first [18]. If SAHS is not confirmed, a manual HRP scoring should be done in the remaining cases. This sequential

approach could save the cost of a scoring technician in an important number of HRP tests, improving the cost-effectiveness of HRP. However, previously published studies did not have large samples or multicentre approaches, and in particular no overall cost-effectiveness studies have been carried out.

To shed light on this topic, as well as to obtain more accurate and definitive data to allow us to optimise HRP use, we performed a multicentre, randomised, blinded crossover study to determine, in a large sample, the following objectives: 1) different aspects of the agreement of automatic HRP scoring in comparison with manual HRP scoring, and both in comparison with PSG in the hospital setting; and 2) the agreement and cost-effectiveness of a sequential HRP protocol (automatic and then manual scorings) compared with both manual HRP scoring and PSG in the hospital setting.

Prior studies of diagnostic cost-effectiveness and the therapeutic efficacy of therapeutic decision-making between manual HRP and PSG in this cohort have been published [11, 24].

METHODS
Subjects
We included patients between 18 and 70 years old, referred to pulmonary clinics at eight hospitals in Spain for suspected SAHS, due to snoring, observed apnoeas, sleepiness (Epworth sleepiness scale >10) or morning fatigue. Patients with other suspected sleep disorders were not included. We excluded patients with severe heart disease, those who were unable to set up the HRP instrument in a trial and those who refused to participate in the study. The ethics committees of the eight participating centres approved the study. All patients provided written informed consent.

Protocol
All patients underwent PSG and HRP in a random order (fig. 1a). PSG and HRP scorings were done separately and the technicians and physicians were blinded to any identifying information about patients as well as any previous results. Once the first test was begun, the second test was scheduled for within the next 3 days.

Home respiratory polygraphy
Our HRP (Breas SC20; Breas Medical AB, Mölnlycke, Sweden) measurements included: oxygen saturation (model 8000 J; Nonin Medical; Plymouth, MN, USA), airflow through a nasal cannula, and thoracic and abdominal movements measured by piezoelectric bands (Pro-Tech reference 1295; Respironics, Pittsburgh, PA, USA), which also measured body position.

All patients were instructed on home use of the HRP device by a technician in the hospital setting before randomisation. Trained personnel from continuous positive airway pressure (CPAP) service companies in each hospital area, acting as transport companies, moved the HRP instruments from home to home. No additional assistance was provided by the transport services to help the patients set up the HRP devices. The raw data files were telematically transmitted from the home to the hospital [11].

The same technician in each centre scored the raw data, following manual and automatic scoring protocols. In the manual scoring, the total number of apnoeas and hypopnoeas was divided by the recording time, excluding ''invalid time'' (time with a bad signal that prevented scoring). For automatic scoring, the total number of apnoeas and hypopnoeas was divided by recorded time with no exclusions.

PSG in the hospital
We used the American Academy of Sleep Medicine 2007 recommendations regarding configuration, filters and signal sampling rates [10]. The neurological variables were electroencephalogram, electrooculogram and electromyogram. Flow tracing was provided by a nasal cannula and thoracoabdominal motion by piezoelectric bands. Oxygen saturation was measured with a finger pulse oximeter. The polysomnographic studies were analysed manually, according to the Rechtschaffen and Kales and the American Sleep Disorders Association 1992 criteria for sleep periods and arousals and according to the Spanish Sleep Network rule for respiratory scoring (see below) [25–27].

Definitions
A valid PSG or HRP had at recorded period of least 3 h. In addition, a valid HRP had at least 3 h of flow or band and oximetry measurements for scoring. An invalid recording could be repeated up to two times.

For PSG, an apnoea was defined as the absence of airflow ($\geqslant$90% reduction) for $\geqslant$10 s and a hypopnoea as a discernible airflow or band reduction ($\geqslant$30% and <90%) of $\geqslant$10 s duration with a $\geqslant$3% drop in oxygen saturation or final arousal [27]. For HRP, apnoeas and hypopnoeas were defined in the same way, but without the final arousal criteria for hypopnoeas. For automatic scoring, we predicted apnoea/hypopnoea events with both flow reduction and desaturation detection, using a previously published regression equation [16]. The number of apnoeas and hypopnoeas was divided by recording time for HRP and sleep time for PSG.

Since our principal objective was cost-effectiveness, we chose the polysomnographic AHI cut-off point ($\geqslant$5, $\geqslant$10 or $\geqslant$15) for SAHS diagnosis that produced the lowest negative post-test probability, and consequently had better capacity to rule out disease; this was $\geqslant$15, based on a recent publication with the same group of patients [11].

Cost analysis
We carried out a cost analysis for two equally effective alternatives [28]. It consisted of summing the costs of all tests needed to reach a final diagnosis (SAHS or no SAHS) in both arms of the study (PSG and HRP). Figure 1b contains the procedure for calculating the costs of both alternatives. We estimated costs in the following cost groups.

1) Test costs for PSG and HRP (manual and sequential protocols): hospital costs included the following expenditures: personnel salaries (technicians, physicians and administrative staff), linear 5-year depreciation of equipment (taking into account the number of recordings done in this period in each hospital as output), fungible material, and the proportional burden of the sleep laboratory in the general budget of the hospital. The costs of PSG and HRP were divided by the number of patients with a valid recording to obtain the cost per patient.
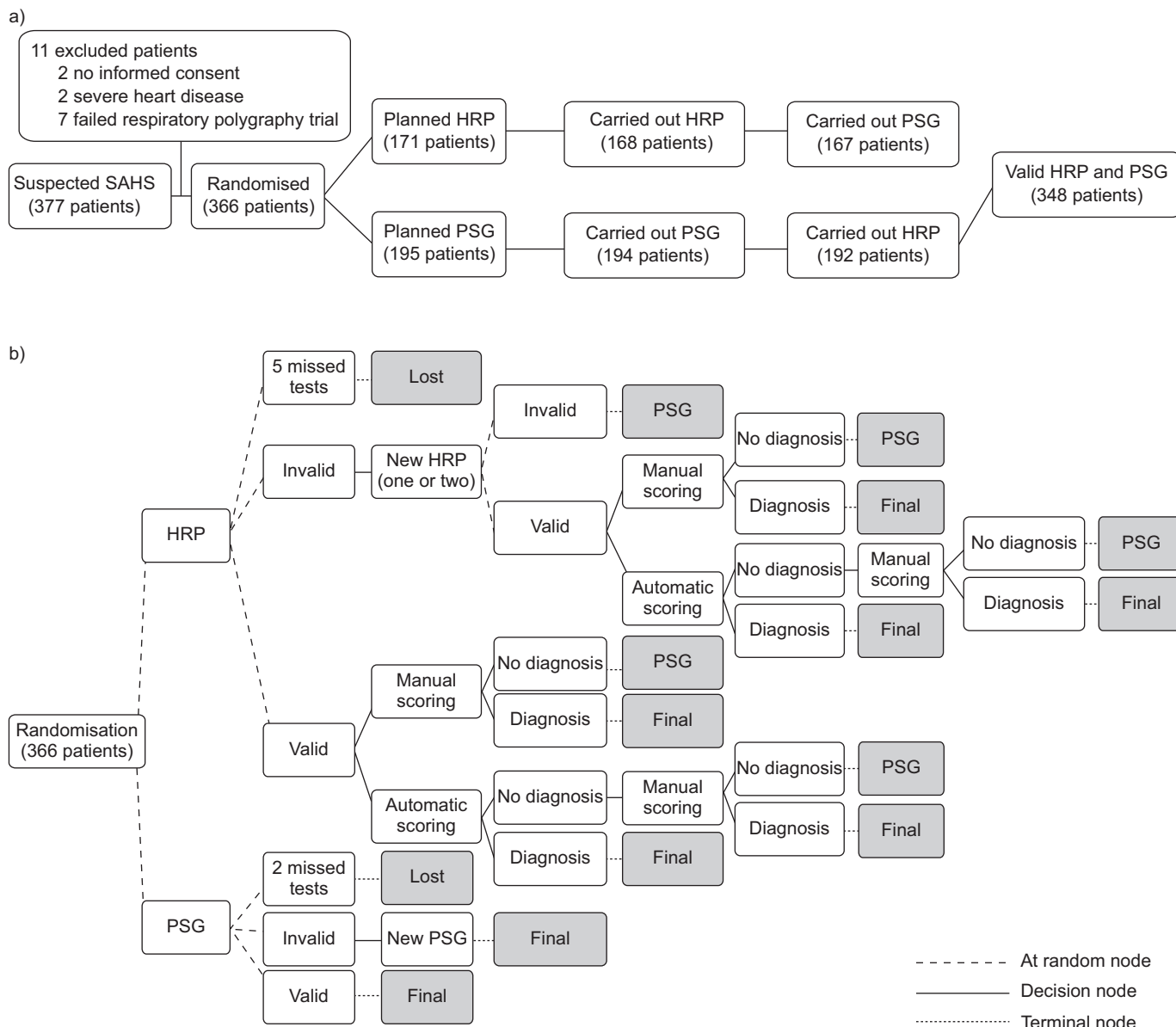
a)

b)

FIGURE 1. a) General flow chart of patients in the study; 359 patients completed both branches of the study. b) More detailed study flow chart to calculate the costs for the two alternatives. Decision nodes refer to alternatives by protocol, random nodes include alternatives determined at random and terminal nodes mark the end of the branch and additional costs. In the home respiratory polygraphy (HRP) branch five patients were lost without adding costs. Patients with invalid recordings underwent one or two new HRPs with two options: 1) patients with repeated invalid recordings reached the terminal node of the HRP arm; the costs of the consequent full polysomnography (PSG) and that of HRP repetitions were added to the final cost; 2) patients with a valid recording. In the latter group of patients, manual and automatic scorings were performed with two potential alternatives: diagnosis established or no diagnosis. Patients with manual or automatic HRP diagnosis (true positive or negative) reached the terminal node of the HRP branch and the burden of HRP (without a technician reading the raw-data in automatic scoring) and repetitions were included. Patients without a manual HRP diagnosis (grey zone and false positives and negatives) reached the terminal node and both HRP and PSG costs were included. In patients without an automatic HRP diagnosis, manual scoring was carried out with two possible options: diagnosis established or no diagnosis. Patients with a diagnosis reached the terminal node and the costs of HRP and repetitions were included. Patients without a diagnosis also reached the terminal node and both HRP and PSG costs were included. Patients with an initial valid HRP followed the same procedure as patients with a valid HRP after repetitions. In the PSG branch, two patients were lost to follow-up without additional costs. In patients with an invalid PSG, the test was repeated and they reached the terminal node of this arm, with the cost of two PSG procedures added to the total PSG costs. Patients with a valid PSG contributed the cost of one PSG. SAHS: sleep apnoea–hypopnoea syndrome.

2) Patient costs for PSG and HRP: the cost of moving from home to hospital and back, per kilometre. The cost per kilometre was calculated for each hospital. The sum of the patients' costs was divided by the number of patients with a valid recording to derive the cost per patient.

3) Total cost: the sum of test (PSG and HRP) and patient costs.

4) Costs of HRP and PSG for equal diagnostic efficacy: For manual and sequential HRP protocols, the costs were the sum of the HRP costs and: a) the costs of PSG in patients with

invalid recordings, after repetitions; b) the costs of repeated HRPs due to invalid recordings, in patients with a valid final recording; c) the cost of PSG in patients with indeterminate results (''grey zone''); and d) the cost of PSG in patients with false negative and positive results. For PSG, we added the initial test costs to the cost of repeated PSGs due to invalid recordings. The cost per patient for equal diagnostic efficacy (HRP protocols and PSG) was calculated by dividing this total cost by the number of randomised patients that completed the protocol.

5) Patient costs for equal diagnostic efficacy: the patient costs plus the burden due to transportation caused by repetitions.

6) Total cost for equal diagnostic efficacy: the sum of test costs (HRP protocols and PSG) and patient costs for equal diagnostic efficacy.

Statistical analysis

Agreement of automatic HRP
To determine the agreement in AHI measurements between automatic and manual scorings, and the agreement of both with PSG, we created Bland–Altman plots, comparing: 1) PSG and manual HRP; 2) PSG and automatic HRP; and 3) manual and automatic HRP.

To assess the diagnostic agreement of automatic HRP and sequential HRP protocols (first automatic and then manual scoring) with PSG we carried out the following analysis.

First, we constructed receiver operating characteristic (ROC) curves for manual and automatic HRP scoring, assuming a polysomnographic AHI score $\geqslant15$ as the criterion for SAHS diagnosis.

Secondly, we determined the exclusive (sensitive) and the confirming (specific) HRP cut-off points for SAHS diagnosis by means of: 1) sensitivity and specificity; 2) negative (1-sensitivity/specificity) and positive (sensitivity/1-specificity) likelihood ratios and 3) the post-test probability of having disease (SAHS) when the test was positive or negative, based on the pre-test probability (prevalence) and positive and negative likelihood ratios [29]. These analyses were carried out using the two HRP scoring protocols, manual and sequential. For the latter we first analysed automatic scoring and then manual scoring in the population undiagnosed by automatic scoring.

To find the optimal exclusionary and confirming HRP cut-off points, we tested the previous parameters each five HRP AHI points, starting with a value of five, in manual and sequential protocols.

Due to the high prevalence (pre-test probability) of SAHS in our study population, it is very probable to be able to identify an effective HRP cut-off point that rules in SAHS in our sample with intermediate and high clinical probabilities of disease. We defined one as follows: a positive likelihood ratio close to 10 and a post-test probability clearly greater than 90%, which means that less than one out of 10 patients with positive HRP, did not have the disease.

Since it is improbable that one could identify an effective HRP cut-off point to rule out SAHS in our patients, we defined one

as follows: a negative likelihood ratios <0.1 and a post-test probability <17%, which means that less than two out of 10 patients with negative HRP have the disease [30].

Cost-effectiveness analysis
Once the HRP cut-off points to confirm (specific) or rule out (sensitive) SAHS diagnosis were determined for manual and sequential protocols, we imputed the cost for PSG and HRP (manual and sequential protocols) arms with the finality of obtaining a definitive diagnosis using HRP or PSG.

RESULTS

Diagnostic efficacy of HRP
Initially, 377 patients were selected. 11 were excluded (fig. 1a). Of the 366 randomised patients, 15 could not produce a valid HRP and PSG (4.9%). Four of these patients (three starting with HRP and one with PSG) did not come to their scheduled appointments. Another rejected the PSG test after a valid HRP and two rejected HRP after a valid PSG.

Of the 359 patients who completed the protocol (both branches of study), PSGs were repeated once in nine patients (2.5%) and the HRP 52 times (once or twice per patient). Of these 52 HRP repetitions, 15 were due to the patients' failure to toggle the device on or off and 37 to invalid or poor quality time registers. Finally, eight patients could not produce a valid HRP after repetitions.

The clinical and anthropometric characteristics of the 348 patients with valid PSG and HRP results are shown in table 1 and data from sleep studies in table 2. AHI from automatic scorings were lower than manual and PSG scorings.

Figure 2 shows Bland–Altman plots for PSG and manual HRP, for PSG and automatic HRP and for manual and automatic HRP. AHI from manual HRP showed better agreement with PSG than automatic HRP, which showed important variability that could be clinically significant for an individual patient.

TABLE 1	Characteristics of the studied population
Males	75.6
Age years	48.7±11.8
Snoring	95
Observed apnoeas	75
Body mass index kg·m⁻²	31.0±6.6
Obesity	48.3
Subjective sleep time per day h	6.9±1.4
Subjective nap time h	0.5±0.6
Epworth sleepiness scale	11.6±5.0
Morning fatigue	86
Alcohol g	8.9±22.3
Smokers	23.9
Depression/anxiety	23.3
Hypertension	30.7
Cardiovascular events	5.6
Systolic pressure mmHg	130.5±17.2
Diastolic pressure mmHg	76.0±12.1
Starting home	47.7

Data are presented as % or mean±SD. n=348.

TABLE 2	Sleep studies			
		PSG	Manual HRP	Automatic HRP
Recording time min		443 ± 43	428 ± 82	428 ± 82
Sleep time or valid time min		372 ± 64	404 ± 71	
NREM stages 1 and 2 %		64 ± 22		
NREM stages 3 and 4 %		17 ± 12		
REM stage %		17 ± 13		
Arousal index		39 ± 25		
AHI		38 ± 29	31 ± 24	25 ± 22
AHI $\geqslant 5$		90	90	84
AHI $\geqslant 10$		83	80	68
AHI $\geqslant 15$		75	70	58
S_{aO_2} <90% of sleep or recording time		9.9 ± 19	6.5 ± 14	6.4 ± 13
Desaturation index		27 ± 29	29 ± 24	28 ± 24

Data are presented as % or mean$\pm$SD. PSG: polysomnography; HRP: home respiratory polygraphy; (N)REM: (non)rapid eye movement; AHI: apnoea–hypopnoea index; S_{aO2}: arterial oxygen saturation.
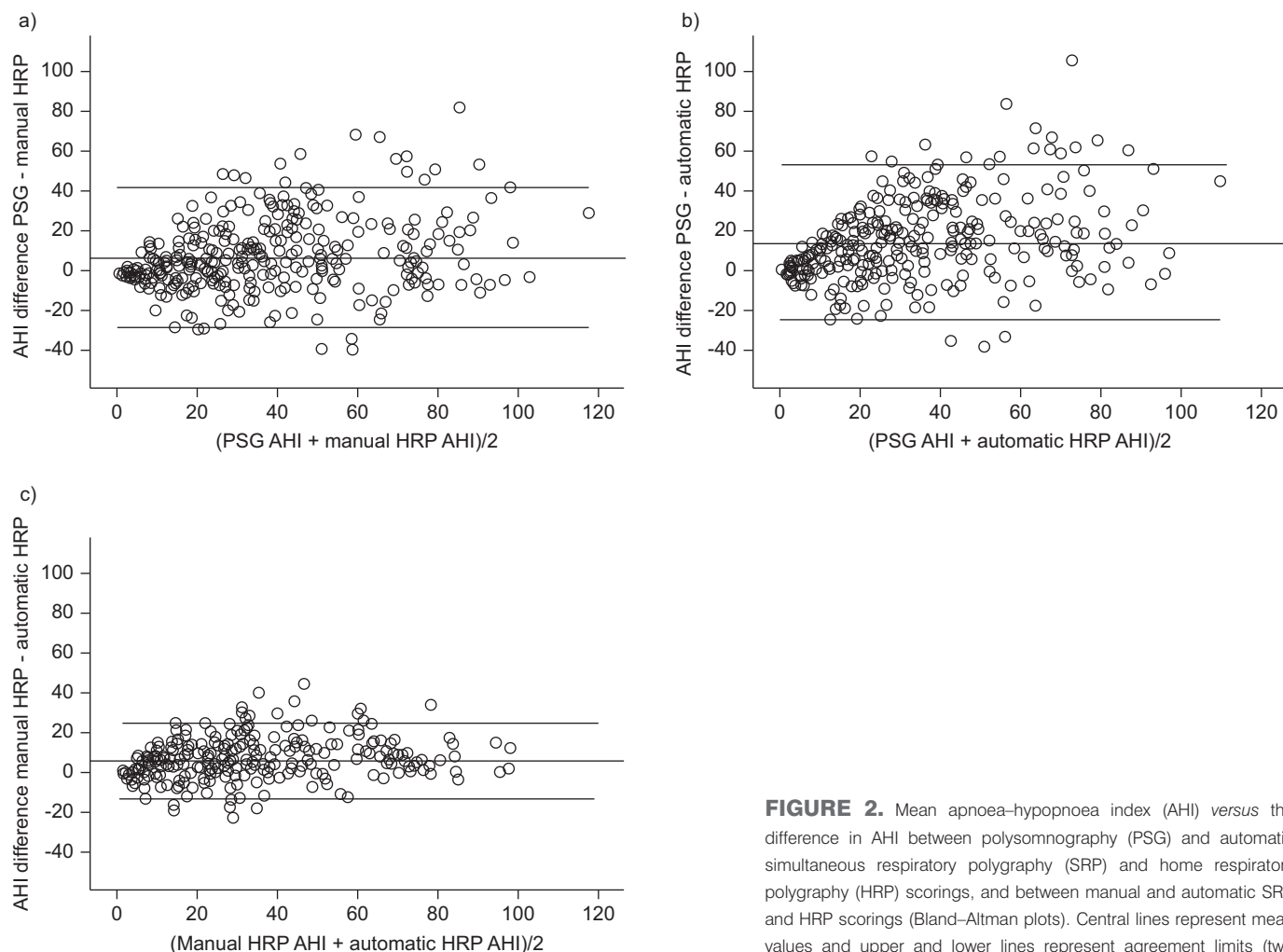

FIGURE 2. Mean apnoea–hypopnoea index (AHI) *versus* the difference in AHI between polysomnography (PSG) and automatic simultaneous respiratory polygraphy (SRP) and home respiratory polygraphy (HRP) scorings, and between manual and automatic SRP and HRP scorings (Bland–Altman plots). Central lines represent mean values and upper and lower lines represent agreement limits (two standard deviations).
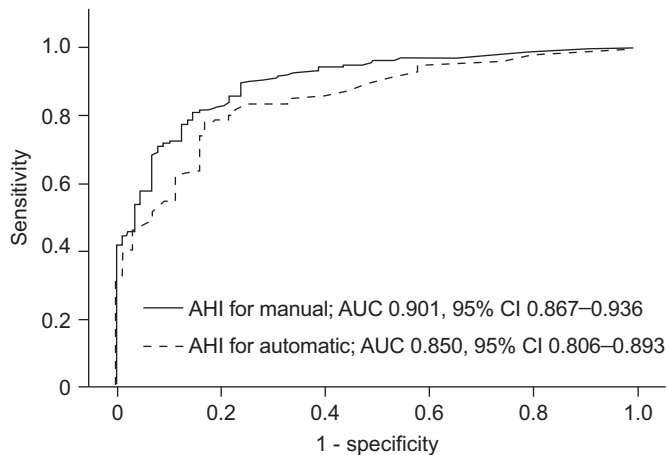
FIGURE 3. Receiver operating characteristic curves for the automatic and manual home respiratory polygraphy (HRP) scorings based on an apnoea–hypopnoea index (AHI) $\geqslant$15 polysomnographic cut-off point for sleep apnoea and hypopnoea syndrome diagnosis. AUC: area under the curve.

However, good AHI agreement was observed between manual and automatic HRP.

Figure 3 shows the SAHS diagnosis (PSG AHI $\geqslant$15) ROC curves for HRP (manual and automatic scorings). The area under the curve was better for manual AHI than for automatic AHI (p<0.001).

Table 3 shows the diagnostic agreement between manual and sequential HRP protocols. A manual HRP cut-off point $\geqslant$5 could effectively rule out a SAHS diagnosis, based on our previous definition for ruling out SAHS (see statistical analysis). An acceptable HRP AHI cut-off point for confirming SAHS would be 25, based on our previous definition for ruling in SAHS (table 3). The percentage of patients with a positive result (true or false positives) would be 90% for an HRP AHI $\geqslant$5 and 51% for

an HRP AHI $\geqslant$25. Therefore, the indeterminate result range (grey zone) requiring PSG represented 136 patients (39% of 348).

For the sequential protocol (table 3) an AHI <5 from automatic HRP could not effectively rule out SAHS. However, an AHI $\geqslant$25 could be a confirmation cut-off for HRP. 214 patients (61%) would remain undiagnosed and available for manual HRP scoring.

In the aforementioned 214 patients, a manual HRP AHI <5 could not rule out SAHS. These values, although close to our definition, do not meet our criteria. An effective HRP confirmation cut-off point could be $\geqslant$20. Therefore, the grey zone requiring PSG would include 135 patients (39% of 348).

Cost analysis

Table 4 shows the mean cost per patient of PSG and HRP with manual and sequential protocols for the calculated cost groups from the eight hospitals. The cost of manual HRP was six times lower than PSG and sequential HRP was seven times lower than PSG. The cost of the sequential HRP was 1.2 times lower than manual HRP due to reductions in the costs of a scoring technician.

To estimate the cost of manual HRP for diagnostic efficacy equal to PSG using the manual protocol, we chose HRP AHI cut-off points to effectively rule out (<5) and to effectively confirm ($\geqslant$25) SAHS, with indeterminate scores (cases needing further PSG assessment) in 39% of cases (table 3). In the case of the sequential protocol, both automatic and subsequent manual scorings did not have effective cut-off points for ruling out SAHS, but both had effective confirmation cut-off points ($\geqslant$25 for automatic and $\geqslant$20 for manual). Finally, 39% of the 348 needed PSG due to indeterminate scores.

The cost of achieving a sequential HRP efficacy equal to that of PSG was four times higher than the test cost without equal efficacy, but substantially lower than the cost of PSG.

TABLE 3	Apnoea–hypopnoea index home respiratory polygraphy (HRP) cut-off points to rule out or confirm sleep apnoea–hypopnoea syndrome (SAHS) diagnosis, for manual and sequential protocols						
Method and cut-off	Sensitivity	Specificity	LR + (95% CI)	Post-test probability + (95% CI)	LR - (95% CI)	Post-test probability - (95% CI)	HRP +/- %
Manual scoring[#]							
$\geqslant$5	98	31	1.42 (1.23–1.64)	81 (79–83)	0.06 (0.03–0.16)	15 (8–32)	90/10
$\geqslant$25	67	92	8.36 (4.09–17)	96 (92–98)	0.36 (0.30–0.44)	52 (47–57)	51/49
Sequential protocol							
Automatic scoring[#]							
$\geqslant$5	93	41	1.58 (1.32–1.89)	83 (80–85)	0.18 (0.11–0.29)	35 (25–47)	84/16
$\geqslant$25	49	93	7.11 (3.25–16)	96 (91–98)	0.55 (0.48–0.62)	62 (59–65)	39/61
Manual scoring[¶]							
$\geqslant$5	95	33	1.43 (1.22–1.68)	70 (67–73)	0.14 (0.06–0.31)	19 (9–34)	85/15
$\geqslant$20	56	94	9.01 (3.80–21)	94 (86–97)	0.47 (0.39–0.58)	44 (39–49)	37/63

LR: likelihood ratio; post-test probability +: the post-test probability of having disease (SAHS) when the test (HRP) was positive; post-test probability -: the post-test probability of having disease (SAHS) when the test (HRP) was negative; HRP +/-: proportion of patients with a positive/negative diagnosis based on HRP. [#]: for the total sample, N=348, pretest probability 75%; [¶]: for a sub-sample, N=214, pretest probability 62%.

TABLE 4	Cost per patient of polysomnography (PSG) and home respiratory polygraphy (HRP) (manual and sequential protocols), for the different cost groups in the eight hospitals		
	PSG	**HRP**	
		Manual	**Sequential**
Test costs	540.60	91.35	76.95
Patient costs	22.21	44.42	44.42
Total costs	562.81	135.77	121.37
Cost for equal diagnostic efficacy	554.08	348.37	320.87
Patient costs for equal diagnostic efficacy	23.13	49.98	49.98
Total cost for equal diagnostic efficacy	577.21	398.35	370.85

Costs are given in 2009 euros.

Considering the total cost of PSG as 100% (fig. 4), the cost of manual HRP was 69% and the cost of sequential HRP was 64%. The savings from sequential HRP in comparison with manual protocols was only 9%, primarily caused by the reduction in the cost of a scoring technician.

DISCUSSION

To our knowledge, this study has the largest sample of any published HRP study to date [17], as well as being the only multicentre study. The principal results in our selected population with intermediate or high SAHS suspicion were: 1) manual HRP scoring had better agreement than automatic HRP scoring; 2) the sequential HRP protocol is a cost-effective alternative to PSG; and 3) the cost savings of the sequential HRP protocol is low in comparison to the manual HRP protocol.

Automatic scoring is a tool available in most HRP devices. Several studies have tested the agreement with simultaneous
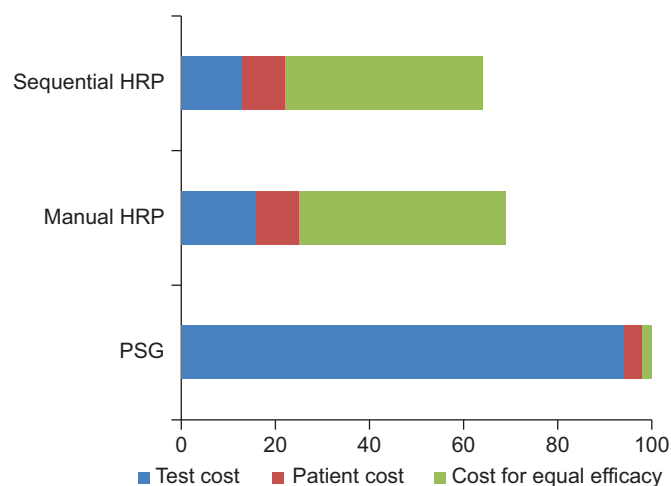

FIGURE 4. Percentages of total home respiratory polygraphy (HRP) (manual and sequential approaches) costs and their distribution in three groups (test cost, patients' cost and cost for equal efficacy) compared to polysomnography (PSG) costs, which were considered to be 100%.

PSG in sleep labs [17–21]. The common finding has been that it underestimates the average AHI, and shows worse AHI agreement with PSG in Bland–Altman plot. In addition, automatic scoring has lower diagnostic agreement than manual scoring. However, most of the cited studies had good confirmation cut-off points for the diagnosis of SAHS but did not find good cut-off points for ruling out the diagnosis. Only one study has compared automatic HRP and manual HRP scorings with PSG [23]. The AHI agreement for the Bland–Altman plots was better with automatic than with manual scorings. A study comparing automatic *versus* manual HRP scorings showed good AHI agreement for the Bland–Altman plots and similar diagnostic agreement [16]. In this study, the hypopnoea definition included desaturation, while only flow reduction was used in the previously mentioned home study comparing automatic and manual scorings with PSG.

Based on the above data, automatic HRP scoring could be useful for confirming a SAHS diagnosis in positive cases with the remaining cases referred for manual HRP scoring [18]. Technician time and subsequent costs could be saved for 39% of patients. If we had used another cut-off point for polysomnographic SAHS diagnosis instead of AHI $\geqslant 15$, the savings in comparison with PSG could be somewhat different. However, the difference in savings between manual and sequential protocols would remain similar, because the difference between both protocols in area under ROC curves for AHI $\geqslant 5$ and $\geqslant 10$ were similar to the area under the curve for AHI $\geqslant 15$ (0.06 for $\geqslant 5$, 0.06 for $\geqslant 10$ and 0.05 for $\geqslant 15$).

As mentioned, a constant in previously published studies and the present study is the underestimation of AHI by automatic scoring in comparison with manual scorings [17, 18, 20, 21], which can make it difficult to identify a cut-off point that rules out SAHS. This happens primarily with the hypopnoea number, indicating deficient algorithms for automatically identifying hypopnoeas. Therefore, future improvements in software may produce better automatic scoring.

Several studies have calculated the costs of HRP in comparison with PSG based on simulated hypothetical cohorts of patients, including diagnosis and treatment [31–34], with contradictory results. Our study had more adequate patient inclusion, since it was performed on a large cohort of real patients, but our cost estimation was limited to diagnosis. Although the diagnostic method seems to be the most important factor for finding cost differences in these hypothetical cohort studies, cost-effectiveness studies including diagnosis and CPAP treatment based on real patients needing SAHS diagnosis (intermediate and high clinical probability) are necessary.

The fact that manual and sequential scoring protocols had lower difference in cost-effectiveness could result in limited interest in the sequential protocol on the part of specialised sleep centres. However, the eventual usefulness of our results could be in an integrated network of tertiary and non-tertiary hospitals. Generally, the tertiary hospitals have complete sleep laboratories (PSG and HRP) and very specialised physicians and technicians. In contrast, in non-tertiary hospitals the frequent scenario is no complete sleep laboratory (PSG unavailable; option to get a HRP device), no highly specialised physicians (*i.e.* pulmonologists with general training) and no specialised

technicians. These physicians receive patients from general practitioners and they can select patients to be screened with HRP based on guidelines, similar to the procedure in tertiary hospitals. Instead of referring patients to tertiary centres for diagnosis, HRP with automatic analysis could be done instead. Patients susceptible to an accurate diagnosis could be diagnosed and be recommended for CPAP or not based on the guidelines. The recordings from the remaining patients could be telematically transmitted to a tertiary hospital for manual scoring, providing a diagnosis for an additional number of patients. Finally, patients without a diagnosis would need to be moved to a tertiary hospital for a PSG. Regarding the cost, the main saving from applying this model would be decreasing the patient's direct cost by eliminating patient travel to the tertiary hospital and the intangible cost from the inconvenience of moving. Nevertheless, specific studies to validate these potential results seem necessary.

In summary, sequential HRP scoring is a significantly lower cost alternative to PSG for the diagnosis of patients with suspected SAHS. However, the low cost savings of sequential *versus* manual scoring protocols could not suppose an interesting alternative to manual scoring until more refined software for automatic scoring becomes available.

STATEMENT OF INTEREST
None declared.

REFERENCES
1 Durán J, Esnaola S, Rubio R, *et al.* 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: 685–689.

2 Bazzano LA, Khan Z, Reynolds K, *et al.* Effect of nocturnal nasal continuous positive airway pressure on blood pressure in obstructive sleep apnea. *Hypertension* 2007; 50: 417–423.

3 Barbé F, Durán-Cantolla J, Capote F, *et al.* Spanish Sleep and Breathing Group. Long-term effect of continuous positive airway pressure in hypertensive patients with sleep apnea. *Am J Respir Crit Care Med* 2010; 181: 718–26.

4 Marin JM, Carrizo SJ, Vicente E, *et al.* Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnea with or without treatment with continuous positive airway pressure: an observational study. *Lancet* 2005; 365: 1046–1053.

5 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: 847–851.

6 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: 1407–1412.

7 Flemons WW, Douglas NJ, Kuna ST, *et al.* Access to diagnosis and treatment of patients with suspected sleep apnea. *Am J Respir Crit Care Med* 2004; 169: 668–672.

8 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: 188–198.

9 Kuna ST, Badr MS, Kimoff RJ, *et al.* ATS/AASM/ACCP/ERS Committee on Ambulatory Management of Adults with OSA. An official ATS/AASM/ACCP/ERS workshop report: Research priorities in ambulatory management of adults with obstructive sleep apnea. *Proc Am Thorac Soc* 2011; 8: 1–16.

10 Iber C, Ancoli-Israel S, Chesson A, *et al.* The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. 1st Edn. Westchester, American Academy of Sleep Medicine, 2007.

11 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: 567–573.

12 Whittle AT, Finch SP, Mortimore IL, *et al.* Use of home sleep studies for diagnosis of the sleep apnoea/hypopnoea syndrome. *Thorax* 1997; 52: 1068–1073.

13 Parra O, García-Esclasans 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: 1720–1724.

14 Alonso Alvarez Mde L, Terán Santos J, Cordero Guevara J, *et al.* Reliability of home respiratory polygraphy for the diagnosis of sleep apnea-hypopnea syndrome: analysis of costs. *Arch Bronconeumol* 2008; 44: 22–28.

15 Ferber R, Millman R, Coppola M, *et al.* Portable recording in the assessment of obstructive sleep apnea. ASDA standards of practice. *Sleep* 1994; 17: 378–392.

16 Ruiz-López FJ, Fernández-Suárez B, Guardiola-Martínez J, *et al.* Quality control of the ambulatory polygraphy using automatic analysis. *Chest* 2009; 135: 194–200.

17 Calleja JM, Esnaola S, Rubio R, *et al.* Comparison of a cardiorespiratory device *versus* polysomnography for diagnosis of sleep apnoea. *Eur Respir J* 2002; 20: 1505–1510.

18 Fietze I, Glos M, Röttig J, *et al.* Automated analysis of data is inferior to visual analysis of ambulatory sleep apnea monitoring. *Respiration* 2002; 69: 235–241.

19 Verse T, Pirsig W, Junge-Hülsing B, *et al.* Validation of the POLY-MESAM seven-channel ambulatory recording unit. *Chest* 2000; 117: 1613–1618.

20 Carrasco O, Montserrat JM, Lloberes P, *et al.* Visual and different automatic scoring profiles of respiratory variables in the diagnosis of sleep apnoea–hypopnoea syndrome. *Eur Respir J* 1996; 9: 125–130.

21 Zucconi M, Ferini-Strambi L, Castronovo V, *et al.* An unattended device for sleep-related breathing disorders: validation study in suspected obstructive sleep apnoea syndrome. *Eur Respir J* 1996; 9: 1251–1256.

22 Ayappa I, Norman RG, Seelall V, *et al.* Validation of a self-applied unattended monitor for sleep disordered breathing. *J Clin Sleep Med* 2008; 4: 26–37.

23 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: 253–259.

24 Masa JF, Corral J, Pereira R, *et al.* 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: 964–971.

25 Rechtschaffen A, Kales A, eds. 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.

26 Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. EEG arousals: scoring rules and examples. *Sleep* 1992; 15: 174–184.

27 Grupo Español de Sueño (GES). Consenso Nacional sobre el síndrome de apneas-hipopneas del sueño. [Spanish Sleep Group. National consensus on sleep apnea-hypopnea syndrome.]. *Arch Bronconeumol* 2005; 41: Suppl. 4, 3–110.

28 Drummond M, O'Brien B, Stoddart GL, *et al*. Methods for the Economic Evaluation of Health Care Programmes. Oxford, Oxford University Press, 1997.

29 Flemons WW, Littner MR, Rowley JA, *et al.* Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. *Chest* 2003; 124: 1543–1579.

30 Collop NA, Anderson WM, Boehlecke B, *et al.* 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: 737–747.

31 Chervin RD, Murman DL, Malow BA, *et al.* Cost-utility of three approaches to the diagnosis of sleep apnea: polysomnography, home testing, and empirical therapy. *Ann Intern Med* 1999; 130: 496–505.

32 Reuveni H, Schweitzer E, Tarasiuk A. A cost-effectiveness analysis of alternative at home or in laboratory technologies for the diagnosis of obstructive sleep apnea syndrome. *Med Decis Making* 2001; 21: 451–458.

33 Deutsch PA, Simmons MS, Wallace JM. Cost-effectiveness of split-night polysomnography and home studies in the evaluation of obstructive sleep apnea syndrome. *J Clin Sleep Med* 2006; 2: 145–153.

34 Pietzsch JB, Garner A, Cipriano LE, *et al.* An integrated health-economic analysis of diagnostic and therapeutic strategies in the treatment of moderate-to-severe obstructive sleep apnea. *Sleep* 2011; 34: 695–709.