

Comparison of a limited computerized diagnostic system (ResCare Autoset™) with polysomnography in the diagnosis of obstructive sleep apnoea syndrome

J.L. Kiely*, C. Delahunty*, S. Matthews*, W.T. McNicholas*

Comparison of a limited computerized diagnostic system (ResCare Autoset™) with polysomnography in the diagnosis of obstructive sleep apnoea syndrome. J.L. Kiely, C. Delahunty, S. Matthews, W.T. McNicholas. ©ERS Journals Ltd 1996.

ABSTRACT: The increasing numbers of patients referred for evaluation of suspected obstructive sleep apnoea (OSA) places a growing burden on available sleep laboratory resources. A number of limited diagnostic systems have been developed in an effort to cope with this clinical problem.

In this study, the diagnostic capabilities of one limited diagnostic system (ResCare Autoset™) were compared with full polysomnography (PSG), using the Oxford SAC™ computerized system. Thirty six patients with suspected OSA had simultaneous studies performed both with the Autoset and Oxford PSG systems. The apnoea plus hypopnoea index (AHI) (events·h⁻¹) scored by the Autoset system was compared with the AHI scored from the PSG raw tracings by an experienced sleep technician.

There were highly significant correlations between the Autoset AHI and the AHI scored by the manual PSG scoring method ($r=0.92$; $p<0.001$). The positive predictive value for diagnosis of OSA for the Autoset was 86% when compared with manual PSG scoring, based on an AHI threshold for OSA of 15 events·h⁻¹. However, the agreement between Autoset and PSG was poor in severe cases of OSA, although not sufficiently so as to result in mistaken diagnosis in any of these cases.

We conclude that the Autoset system is a sensitive and easy to use system, which facilitates screening for obstructive sleep apnoea with a reasonable degree of accuracy.

Eur Respir J., 1996, 9, 2360–2364.

*Respiratory Sleep Laboratory, Dept of Respiratory Medicine, St. Vincent's Hospital, Dublin, Ireland.

Correspondence: W.T. McNicholas
Dept of Respiratory Medicine
St. Vincent's Hospital
Elm Park
Dublin 4
Ireland

Keywords:
Limited computerized diagnostic system
obstructive sleep apnoea
polysomnography
sleep studies

Received: March 27 1996
Accepted after revision June 27 1996

JLK is the recipient of an Irish Health Research Board Postgraduate Fellowship grant.

Obstructive sleep apnoea (OSA) is now recognized to be a common clinical condition, which affects an estimated 2–6% of middle-aged males [1, 2]. The condition is associated with an increased cardiovascular and cerebrovascular morbidity and mortality [3–5], and is also associated with impaired daytime cognitive function. Sleep apnoea is also a recognized co-factor in the aetiology of road traffic accidents [6].

The increased recognition of OSA has inevitably resulted in a substantial increase in the number of referrals to specialist sleep centres, which leads to lengthening waiting times for diagnostic sleep studies. Traditionally, polysomnography (PSG) has been the "gold standard" for investigating such patients, which involves monitoring of sleep stage in addition to respiration, gas exchange and other variables, continuously during sleep. However, this investigation is time-consuming, labour intensive, and requires continuous supervision of the patient and recorder during the course of the sleep study to achieve optimum results [7]. Furthermore, the equipment used for PSG studies is expensive, to which must be added the cost of the consumables, such as electrodes, gel, and data storage media (paper or optical disc) used in each sleep study. There may also be a diminution of sleep time and quality due to the dis-

comfort caused by the numerous electrodes attached to the patient.

A less costly and simpler diagnostic system, that required less time to set up and less supervision during the course of the study, would be a valuable tool, but only if the accuracy of the system was adequate for the correct diagnosis and grading of OSA severity. A further advantage would be if the data analysis was less time-consuming. We therefore evaluated a new limited diagnostic system, the ResCare Autoset™ with version 3.03 software (ResCare Ltd, Sydney, Australia) in patients referred for evaluation of suspected OSA. The device has been in clinical use in this department for 18 months, but earlier software provided only an analysis of apnoeas, and not hypopnoeas [8, 9]. The latest software (version 3.03) reports apnoea index (AI) and apnoea plus hypopnoea index (AHI), which provides a more accurate analysis of OSA severity [7].

Methods

Forty one consecutive patients, who were scheduled to have clinical sleep studies for evaluation of suspected OSA, participated in the study, of whom 36 had technically satisfactory overnight recordings with both

study devices. Patients with significant chronic obstructive pulmonary disease (COPD) and/or baseline awake hypoxaemia due to other causes were excluded from recruitment. Overnight sleep studies were performed on each patient, during which Autoset and PSG measurements were recorded simultaneously. Each system was set up by an experienced sleep technician. All patients completed an Epworth Sleepiness Score to quantify the level of daytime sleepiness [10].

Autoset system

The Autoset has two functional modes, diagnostic and treatment. In the diagnostic mode, the Autoset measures pressure generated in the anterior nares through standard nasal cannulae. The internal pressure transducer and the system software allow estimation of nasal airflow and the flow/time relationship [11]. Snoring is qualitatively assessed by analysing the flow signal following band-pass filtering. Using an external digital oximeter (Biox 3700e, Ohmeda, Essex, UK), nocturnal arterial oxygen saturation is measured and recorded simultaneously by the Autoset, although oximetry data are not used in defining a respiratory event. The Autoset uses two independent detectors, one that detects apnoeas and the other detects apnoeas and hypopnoeas combined. The device does not separately record hypopnoeas. These variables are calculated per hour of total study time as this system does not record sleep stage. The apnoea detector is triggered if there is a decrease in the ventilation signal by 75% or more, averaged over each 2 s interval, for at least 10 contiguous seconds. The apnoea and hypopnoea detector is triggered if there is a decrease in ventilation by 50% or more averaged over any 10 s interval. Since the two detectors are independent, it is not always possible to assume that the frequency of hypopnoeas can be calculated by simply subtracting the apnoea index from the apnoea and hypopnoea index. In clinical practice, we have recorded occasional studies using the Autoset device where the AI is reported as having exceeded the AHI. This phenomenon may occur where an apnoea is directly followed by hypopnoea and then another apnoea. In this latter scenario, the AI detector will score two events, the AHI detector will score only one (personal communication, M. Berthon-Jones, ResMed Ltd, Australia). We have only encountered this phenomenon in very severe cases of OSA. Since oral respiration is not monitored by the Autoset, periods of mouth-breathing might also be a potential cause for this feature.

In review mode, a real time reproduction of the nasal airflow pattern, snoring parameters and oximetry record allows assessment of the study for technical defects related to poor airflow signal input. In order to make satisfactory comparisons between the Autoset and PSG systems, technically poor Autoset studies were excluded after visual inspection of the raw data for defects by an experienced staff member. Two Autoset studies, along with their corresponding PSG studies, were excluded on this basis. In one case, computer malfunction was responsible as the system does not warn the operator whether there is sufficient space to record a full study, and when the computer memory is full, the system "crashes". The other excluded study resulted from

the operator being unable to keep the nasal cannulae positioned in the nares without repeatedly waking the patient.

Polysomnography

Full overnight PSG studies were performed using the Oxford Medilog SAC 847™ Sleep Analysing Computer system (Oxford Instruments PLC, Oxford, UK), which measures standard parameters of electroencephalography (EEG), electro-oculograms (EOG), submental electromyogram (EMG), electrocardiography (ECG), digital pulse oximetry, chest and abdomen movement by inductance plethysmography, body position, oral and nasal airflow. Particular care was taken to ensure that there was no interference between the nasal airflow signals of the Autoset device and the Oxford system by careful positioning of the PSG thermal airflow sensors and the Autoset nasal cannulae. All PSG sleep stage data were manually scored according to the criteria of RECHTSCHAFFEN and KALES [12] by an experienced staff member, who was blinded to the results from the Autoset system. Respiratory events were defined on the basis of airflow and respiratory effort; if airflow was <20% of baseline the event was classified as an apnoea, if >20% and <50% the event was classified as a hypopnoea. The Oxford system used chest and abdominal movement measured by impedance plethysmography to assess respiratory effort in order to further classify the nature of these respiratory events as obstructive or central. Technically poor PSG studies related to poor electrode signal quality and the corresponding Autoset study were excluded from further evaluation. This resulted in a total of three PSG studies which had to be excluded on the grounds that accurate sleep staging would not be possible.

In all cases, the reasons for patient exclusion were technical rather than patient-related, and no special characteristics were noted in the patients whose Autoset or PSG studies were unsatisfactory.

Polysomnographic scoring of AHI was based on total sleep time rather than total study time, and therefore intervening periods of wakefulness during the night were excluded in the PSG analysis, but not in the Autoset analysis. An AHI level of 15 events·h⁻¹ was used in this study to define clinically significant OSA.

Statistical analysis

Data analysis indicated that the AHI values were not normally distributed in each of the methods used, and subsequent correlation analysis was performed using nonparametric methods (Spearman's rank order correlation). BLAND and ALTMAN [13] plots of the difference in AHI score against the average AHI score were constructed to compare the two methods.

Results

The 36 patients studied consisted of 27 males, mean (SD) age 45 (13) yrs, with a body mass index (BMI) of 28.0 (5.3) kg·m⁻² and Epworth scores of 10.3 (5.8),

Table 1. – Summary of results in 36 patients

		Correlation	Limits of agreement
AHI events·h ⁻¹			
Autoset	18.3 (20.1)	0.92	-15.5 to 13.0
PSG	19.4 (24.7)		
Total A + H n			
Autoset	134.2 (152.8)	0.91	-54 to 112
PSG	105.2 (137.0)		
AI events·h ⁻¹			
Autoset	11.2 (20.6)	0.85	-15.1 to 10.5
PSG	8.9 (16.6)		
Total A n			
Autoset	91.8 (156.0)	0.86	-178.4 to 108.4
PSG	48.2 (83.5)		

Values are presented as mean, and SD in parenthesis. A: apnoeas; H: hypopnoeas; AHI: apnoea plus hypopnoea index; AI: apnoea index; PSG: polysomnography using Oxford SAC system.

and 9 females, aged 41 (10) yrs, with a BMI 28.3 (5.0) kg·m⁻² and Epworth scores of 10.4 (6.2).

The data for AHI, AI and total apnoeas and hypopnoeas, both by PSG and Autoset, as well as the calculated limits of agreement, are presented in table 1. There was a highly significant correlation ($r=0.92$; $p<0.001$) between the Autoset system and PSG for AHI estimation (fig. 1a). The Autoset tended to underscore the AHI, particularly at higher levels of AHI, but the agreement between the methods was otherwise acceptable when assessed by the method of BLAND and ALTMAN [13] (fig. 1b). When only the patients with AHIs less than 50 events·h⁻¹ by PSG were considered ($n=31$), the limits of agreement were -8.8 to 10.3 events·h⁻¹.

The AHI for the PSG studies was calculated for total sleep time (*i.e.* excluding intervening periods of wakefulness), whereas the AHI for Autoset studies was calculated for total study time. This point is particularly relevant since, on average, patients slept for only 76% of the total study time as assessed by manual sleep stage analysis of the PSG records. In order to exclude any

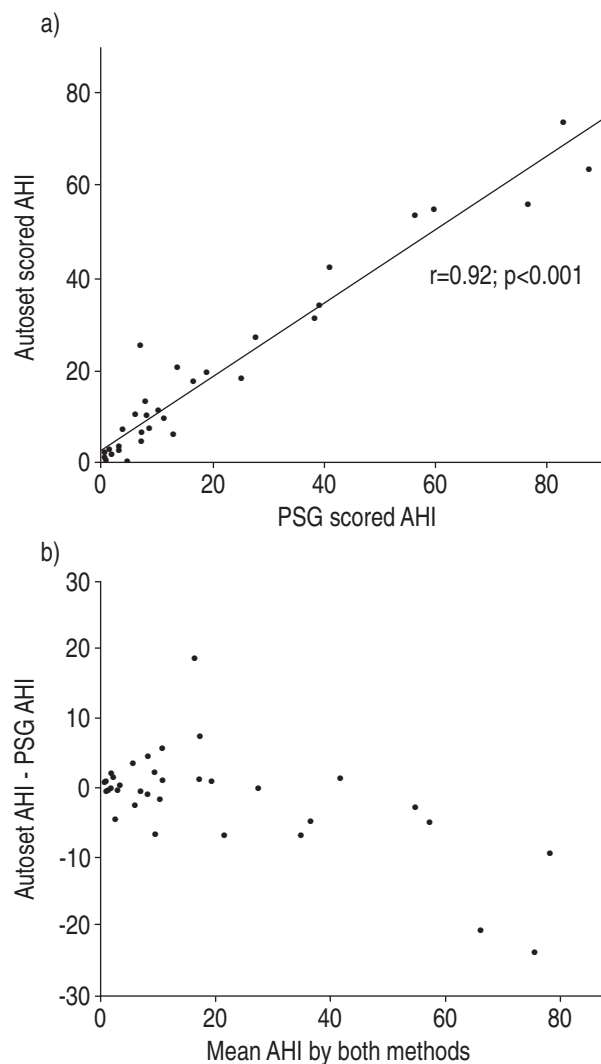


Fig. 1. – a) Comparison of AHI scored by the Autoset and standard PSG methods; and b) estimation of the limits of agreement by plotting the difference against the mean for both methods (BLAND and ALTMAN [13] technique). AHI: apnoea/hypopnoea index; PSG: polysomnography.

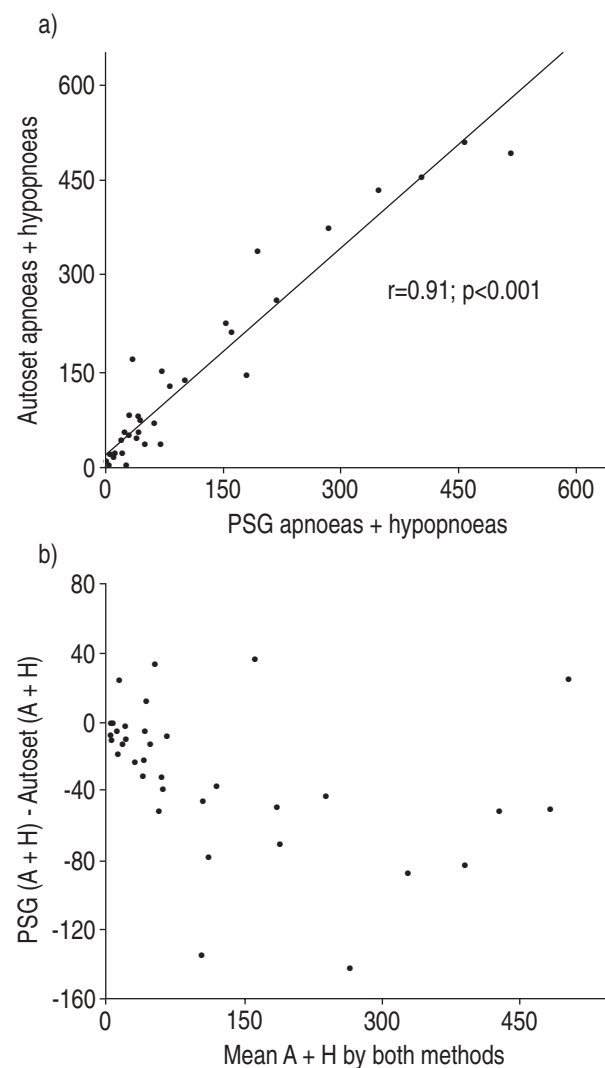


Fig. 2. – a) Relationship of the sum of polysomnographic (PSG) scored apnoeas (A) and hypopnoeas (H) to the sum of Autoset scored apnoeas and hypopnoeas using the total study time in each case; and b) estimation of the limits of agreement between the two methods.

Table 2. – Sensitivity, specificity and predictive values for the diagnosis of OSA by the Autoset system compared to manually scored PSG using different AHI thresholds

OSA diagnostic level events·h ⁻¹	Sensitivity %	Specificity %	PPV %	NPV %
10	85	87	79	91
15	100	92	86	100
20	88	93	78	96

OSA: obstructive sleep apnoea; AHI: apnoea/hypopnoea index; PSG: polysomnography; PPV: positive predictive value; NPV: negative predictive value.

possible bias related to the different methods of measuring apnoea and hypopnoea frequency, total numbers of apnoeas and hypopnoeas for the two systems were compared over the total study time. These data are presented in figure 2, and also show a close relationship between the two systems, although the tendency for the Autoset to overscore total apnoeas and hypopnoeas is particularly evident by this method of analysis and the limits of agreement are quite wide (table 1).

The mean AHI for PSG studies, when expressed in terms of total study time rather than total sleep time fell to 14.5 (18.6), reflecting the fact that intervening periods of wakefulness are included, which represented 24% of the total study time. However, once again, a highly significant correlation was found between the two AHIs by this method of analysis ($r=0.91$; $p<0.001$).

Sensitivity, specificity and predictive values for the Autoset were calculated for three levels of AHI, which reflect different cut-off levels for the diagnosis of OSA used in various centres throughout the world (table 2). At the most commonly used diagnostic threshold for AHI of 15 events·h⁻¹, the negative predictive value was 100%. Although the positive predictive value was not quite as good, it indicates that the Autoset device did not fail to detect any patient with OSA based on an AHI criterion of 15 events·h⁻¹ determined by PSG, but did record some patients as having OSA who had an AHI <15 events·h⁻¹ on PSG.

Discussion

The present data indicate that the Autoset device provides a reasonably accurate method of assessing the presence or absence of OSA. However, the Autoset tended to overestimate the total number of apnoeas and hypopnoeas compared to PSG, particularly in severe cases. One potential source of this discrepancy is that the Autoset measures only nasal airflow and may mistakenly classify periods of oral breathing as apnoeas or hypopnoeas. However, because the AHI measured by the Autoset was expressed in terms of total study time, rather than total sleep time, the resultant apnoea/hypopnoea index was reduced by this time factor, since intervening periods of wakefulness were included in the analysis. This factor tended to compensate for the over-detection of total events, and, therefore, the discrepancy in AHI comparison between the two systems was less than the discrepancy for total events.

Polysomnographic analysis expresses AHI in terms of time spent asleep, excluding intervening periods of wakefulness, and thus provides a more precise assessment of sleep apnoea severity. The discrepancy between Autoset and PSG was most pronounced in severe cases of OSA, which should not be surprising since such patients have numerous arousals and consequent poor sleep quality. Thus, the effect of over-scoring apnoeas and hypopnoeas by the Autoset is diluted by using total study time, and results in overall lower Autoset AHI scores compared to PSG AHI scores.

A second factor that may contribute to the overscoring of apnoeas by the Autoset is the apnoea detection threshold, which scores an apnoea if the airflow is less than 25% of baseline, as opposed to less than 20% for the Oxford PSG system. This difference in the threshold for apnoea detection results in greater apnoea detection by the Autoset than the Oxford system (table 1). However, both systems use a threshold detection criterion of less than 50% of baseline airflow for hypopnoea, and, therefore, the values for total apnoeas and hypopnoeas and for AHI from the Autoset show better agreement with the corresponding PSG AHI values (table 1).

The Autoset is a user-friendly dual purpose device that facilitates both diagnosis of OSA and automated nasal continuous positive airway pressure (nCPAP) titration when used in either diagnostic or treatment modes, respectively. The device requires little time to set up for a patient study, and has the added advantage that considerably less skill is required to link the device to the patient than for a full PSG system. It can sometimes be difficult to achieve reliable and "noise-free" EEG data during polysomnography, and repositioning displaced or noisy electrodes can in itself cause sleep disruption to the patient. Several different electrodes are required in standard PSG and the monitoring of the signal quality from these is largely removed in the Autoset system, where only nasal airflow and oximetry are recorded. This can reduce operator workload during the study, freeing time for other tasks. The results of an Autoset study can be analysed more rapidly than manually scoring a PSG study and this time saving can be considerable when several studies are performed simultaneously. The device is also less uncomfortable for the patient in the diagnostic mode, which may result in better sleep quality during the study.

The data generated by a single study on a 486DX PC-based system amounts to approximately 18 Mbytes and can be easily stored on a backup tape unit. The diagnostic sensitivity and specificity of this system is superior to that of oximetry alone [8], and therefore appears preferable as an initial screening procedure for suspected OSA. The highly significant positive correlation ($r=0.92$; $p<0.001$) between the AHI derived from the comparison of the Autoset system and the manually scored PSG indicates that the Autoset system is likely to be of benefit in reducing waiting lists for sleep studies and treatment, particularly by excluding those without OSA, and also by diagnosing those with moderately severe OSA (AHI >25). However, the true diagnostic ability of the Autoset may be more accurately reflected by the correlation between total apnoeas and hypopnoeas scored by the two systems, which was 0.86 ($p<0.001$).

Concern remains regarding the ability of the Autoset to reliably classify patients with borderline or mild OSA, since our data indicate that the Autoset is not accurate enough to reliably predict the presence or absence of significant OSA in such patients, as seen in the estimation of the relatively wide limits of agreement by the BLAND and ALTMAN [13] plots (figs. 1 and 2). Therefore, patients with AHI scores of 5–25 events·h⁻¹ by the Autoset device may need full PSG studies for definitive diagnosis. On the other hand, the present data indicate that no patient with an Autoset-scored AHI of <5 or >25 events·h⁻¹ was incorrectly classified as OSA or non-OSA when compared to full PSG studies.

The latest version (Version 3.03) software in the Autoset system provides measurement of apnoeas and hypopnoeas, whereas previous software versions measured only apnoeas. Nevertheless, our data agree with the two other recent reports that have compared the Autoset system with full PSG using earlier software versions [8, 9]. An important limitation of the software in the Autoset is that it does not allow editing of data and, therefore, a poor signal from nasal cannulae or oximetry probe dislodgement may lead to erroneous results. Data obtained for analysis in the present study were reviewed by experienced sleep technicians to determine the adequacy of signal acquisition quality, and studies were excluded if the signal quality did not reach technically adequate levels as already described. However, only a small number of studies were excluded by these criteria in the present study.

We conclude that, despite certain limitations, the Autoset system is a reliable and reasonably accurate system for assessing patients with suspected obstructive sleep apnoea. Our findings indicate that the Autoset system would be an appropriate device for use as an initial screening procedure in centres with limited experience of polysomnography. Referral to a specialized sleep centre should continue to be recommended when there is a persisting doubt about the level of severity of obstructive sleep apnoea, particularly in borderline cases.

References

1. Young T, Palta M, Dempsey J, Skatrud J, Webber S, Bader S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993; 328: 1230–1235.
2. Partinen M. Epidemiology of sleep disorders. *In*: Kryger MH, Roth T, Dement WC, eds. *Principles and Practice of Sleep Medicine*. 2nd edn. Philadelphia, W.B. Saunders, 1994; pp. 437–452.
3. Koskenvuo M, Kaprio J, Telakivi K, Partinen M, Heikkila K, Sarna S. Snoring as a risk factor for ischaemic heart disease and stroke in men. *BMJ* 1987; 294: 16–19.
4. He J, Kryger MH, Zorick FJ, Conway W, Roth T. Mortality and apnea index in obstructive sleep apnea: experience in 385 male patients. *Chest* 1988; 94: 9–14.
5. Partinen M, Jamieson A, Guilleminault C. Long-term outcome for obstructive sleep apnea patients. *Chest* 1989; 94: 1200–1204.
6. Findley LJ, Umversat ME, Suratt PM. Automobile accidents involving patients with obstructive sleep apnea. *Am Rev Respir Dis* 1988; 138: 337–340.
7. Douglas NJ, Thomas S, Jan MA. Clinical value of polysomnography. *Lancet* 1992; 339: 347–350.
8. Gugger M, Mathis J, Bassetti C. Accuracy of an intelligent CPAP machine with inbuilt diagnostic abilities in detecting apnoeas: a comparison with polysomnography. *Thorax* 1995; 50: 1199–1201.
9. Bradley PA, Mortimore IL, Douglas NJ. Comparison of polysomnography with ResCare Autoset in the diagnosis of the sleep apnoea/hypopnoea syndrome. *Thorax* 1995; 50: 1201–1203.
10. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep* 1992; 15(4): 376–381.
11. Berthon-Jones M. Feasibility of a self-setting CPAP machine. *Sleep* 1993; 16: S120–S123.
12. Rechtschaffen A, Kales A, eds. *A Manual of Standardised Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Bethesda, National Institute of Neurological Disease and Blindness, 1968; NIH publication No. 204.
13. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; i: 307–310.