Simultaneous laboratory-based comparison of ResMed Autoset™ with polysomnography in the diagnosis of sleep apnoea/hypopnoea syndrome


ABSTRACT: ResMed Autoset™ (AS) is a simplified diagnosis system for obstructive sleep apnoea/hypopnoea syndrome (OSAS) based on the respiratory flow/time relationship by pressure variation measured through simple nasal prongs.

A multicentre prospective trial was used to compare AS and polysomnography (PSG) for diagnosing 95 patients, with suspected OSAS. Physicians gave a pretest probability of the patient having OSAS. The apnoea/hypopnoea index (AHI) was compared between the two methods of diagnosis for the whole population and for subgroups according to the pretest probability.

Twenty-four patients had AHI <15 events·h⁻¹ on PSG and 19 AHI 15–30, and 52 patients had AHI ≥30. Correlation between AHI assessed by AS and PSG was r=0.87 for total sleep time (TST), p<0.0001. A Bland and Altman plot gave an agreement between the two methods of ±40%. For a threshold of AHI ≥15 events·h⁻¹ to diagnose OSAS, AS has a sensitivity of 92%, specificity of 79%, positive predictive value of 93% and negative predictive value of 76%. With a pretest probability ≥80%, sensitivity and positive predictive value were 98 and 100% respectively. Of six false negative, 93% and negative predictive value of 76%. With a pretest probability ≤80%, sensitivity measurement. Studies examining simplified systems of OSA diagnosis need to assess the sensitivity, specificity and predictive value of each system compared with complete PSG. Furthermore the most reliable combinations of parameters need to be identified. Autoset™ (AS) (ResMed, Sydney, Australia) is a continuous positive airway pressure (CPAP) delivery system which can be used in diagnostic mode [5–8]. If AS could be used to diagnose or exclude OSA with reliability then the utilization of sleep laboratory facilities could be improved and the number of unnecessary sleep studies reduced.

The usefulness of any diagnostic test depends on the purpose for which the test is being used. If a test is used for epidemiological screening then it must have a high sensitivity with few negative results. For diagnostic purposes high specificity (few false positive (FP)) and good predictive value would also be expected. The predictive value of a test is dependent on its sensitivity and specificity, but also on the pretest likelihood of disease or prevalence of the disease in the population studied [9–12]. Investigating a high prevalence population will necessarily lead to higher numbers of true positives (TP) without any major change in FP and, thus, any test will have a higher positive predictive value. Thus, to increase the positive predictive value of a test one may restrict its use to patients with high pretest probability, which selects a group with higher prevalence. Such a strategy has been used successfully for the diagnosis of pulmonary embolism with pulmonary scintigraphy.
The application of any new diagnostic test must take into account the likelihood or prevalence of disease in the population studied in order to relate its usefulness to other diagnostic tests.

The aim of this study was to examine the sensitivity, specificity and predictive value of AS in the diagnosis of OSAS compared with conventional PSG when taking into account the pretest probability of OSAS. The usefulness of inspiratory flow limitation in distinguishing primary snoring from upper airway resistance syndrome (UARS) was not assessed in this study. A prospective, comparative assessment of AS and conventional PSG was designed in the setting of six different hospital sleep laboratories to assess its usefulness in the diagnosis of sleep apnoea.

Methods

Patients

Ninety-five consecutive patients referred for snoring or suspected OSAS to one of six participating sleep laboratories were included in the study. A protocol questionnaire about symptoms associated with OSAS was filled in by the patient. A clinical examination was performed with measurements of body mass index (BMI; kg·m$^{-2}$), blood pressure, neck circumference and spirometry. The protocol conformed to the ethical guidelines of the participating centres.

Determination of pretest probability

The clinician reviewing the patient made an estimate of the likelihood of OSAS in percentage terms between 1 and 99% certainty. The clinician was told to take into account all of the clinical information available before the PSG and to mark on a percentage scale provided the pretest likelihood of OSAS [12]. The questionnaire and examination protocol emphasized points which have known predictive value for OSAS screening, such as a history of high blood pressure [10, 14], the presence of snoring [10, 15], witnessed apnoea [10, 14] and neck size [10, 16]. Sleepiness was assessed by the Epworth Sleepiness Score (ESS) [17]. In addition, a subjective impression of OSAS presence was used to define a group of patients with higher prevalence of OSAS, expecting a higher positive predictive value of AS in this subpopulation. An 80% or greater pretest probability of OSAS was considered as a high pretest likelihood. This was not a validated assessment but an attempt to define the usefulness of the AS in context of the likelihood of disease [11, 12].

Investigation

Full polysomnography attended by a technician was recorded. Signals included arterial oxygen saturation ($S_{\text{aO}}$), cardiac frequency, chest and abdominal movement, airflow by a pneumotachograph attached to a nasal mask associated with airflow thermistors, submental electromyogram (EMG), electro-oculogram (EOG) (two-channel) and three-channel electroencephalogram (EEG). Neither a snoring microphone nor leg electrodes were systematically used. The studies were read by experienced sleep technicians and doctors. Sleep was staged according to the criteria of Rechtschaffen and Kales [18]. Hypopnoea was defined as a 50% reduction in airflow compared to the previous cycles associated with an arousal or desaturation of ≥2%. The recordings were read without knowledge of the results of the other investigations or the pretest likelihood of disease.

The AS device consists of a computer-controlled nasal CPAP device which uses an oximeter and nasal prongs in diagnostic mode. The inspiratory airflow pattern is detected by monitoring pressure changes at the nares and is compared to a standard reference sinusoidal curve so that flattening of the curve is recognized. Nasal airflow is thus measured semiquantitatively in arbitrary units (range 0–20). Hypopnoea, in the AS study, is defined as a reduction in flow amplitude of ≥50% compared to the five previous breaths and which lasts for >10 s.

The data from the AS are transferred via a cable to a 486DX IBM-compatible computer (New York, USA) containing a specific A/D card assembly so that the signal can be analysed by dedicated software. The final report shows the compressed raw data of oxygen saturation, ventilation, snoring, apnoea and hypopnoea, where each apnoeic event is represented as well as the apnoea/hypopnoea index (AHI) on an hour-by-hour basis. Summary graphs of the percentage of the study time versus each parameter are also provided and overall apnoea and hypopnoea results per hour of study time given.

In this study the AS nasal prongs were placed under the nasal mask of the pneumotachograph used in the simultaneous PSG. No adjustments were made to the AS by the technicians during the study. A preliminary study was carried out to validate the use of a nasal mask over the nasal cannulae. A mannequin airway model (Laederle Airway Management Trainer, Stavanger, Norway) was used. Nasal cannulae connected to AS were applied to the nose of the model both with and without the nasal mask attached to the pneumotachograph. The left and right main bronchi of the model were attached to the expiratory and inspiratory lines of a volume cycled ventilator (Bear 33; Bear Medical Systems, Riverside, CA, USA). Five runs of >10 breaths each at different volumes (1,000, 500, 300 and 250 mL) were conducted at a flow rate of 50 L·min$^{-1}$. The tidal volume ($V_T$) by the ventilator was always underestimated by AS but measures were more accurate when the nasal prongs were under the nasal mask. The mean±SD volume measured with the nasal prongs under the nasal mask was 761.6±7.3 mL of a 1,000 mL delivery compared to 462.4±7.7 mL with the nasal prongs alone without the mask. However, a reduction in $V_T$ delivered (e.g. 50%) was recognized equally whether or not the nasal prongs were under the mask, being 53.0±0.4 versus 53.8±0.7%, respectively. In this model reduction of the $V_T$ from 1,000 to 500 mL did not lead to detection of hypopnoea by AS either with or without the mask. A larger reduction in $V_T$ from 1,000 to 250 mL, was required for all runs to be recognized as hypopnoeas with the nasal prongs alone, while four out of five runs were recognized by AS as hypopnoeas with the nasal prongs under the mask. At lower volumes a reduction in $V_T$ from 500 to 150 mL led to a recognition of hypopnoea in three of five runs. Thus using a pneumotachograph and nasal mask in the PSG study for identifying respiratory events was justified as it
did not change the ability of AS to recognize decreases in \( V_t \) and hypopnoeas.

Comparisons with PSG were performed by Pearson correlation coefficients and also by Bland and Altman plots for comparison between diagnostic methods [19].

**Results**

**Population**

Ninety-five patients were studied of whom 79 were males. Demographic characteristics are shown in table 1. The mean±SD BMI was 30.7±7.3 kg·m\(^{-2}\) and the ESS 11.5±4.9. As the study was carried out in established sleep laboratories the estimated pretest probability of OSAS was high, at 67±32% pretest likelihood. This was confirmed by the results of PSG, as the prevalence of OSA (AHI >15) was 75% (71/95). Fifty-two patients had an AHI >30 events·h\(^{-1}\); 24 patients had AHI <15 events·h\(^{-1}\), of whom six had fewer than 5 events·h\(^{-1}\) and 19 had mild-to-moderate OSAS (AHI 15–30).

**Sleep data**

Patients generally slept well, with sleep time ranging 169–498 min (median 351 min). As expected in this population there was a loss of slow-wave sleep; the median percentage of total sleep time (TST) being 5%.

**Recognition of apnoea and hypopnoea**

The mean±SD AHI by AS was 34.4±24.4 events·h\(^{-1}\) compared with 43.3±33.4 (p<0.05) from PSG per hour of sleep (TST) and 32.3±24.6 events·h\(^{-1}\) (nonsignificant (NS)) from PSG per hour in bed (TBT). The apnoea index (AI) from AS was 20.5±21.7 versus 27.3±31.5 (p<0.05) from PSG-TST and 20.4±23.5 from PSG-TBT (NS). The hypopnoea index (HI) from AS was 13.6±9.3 versus 16.0±14.5 (NS) from PSG-TST and 11.8±10.9 (NS) for PSG-TBT.

Significant mouth breathing was present in some patients, but in the important area of AHI between 10 and 30 there was no difference among centres in the proportion of apnoeas to hypopnoeas.

| Table 1. – Demographic and apnoea/hypopnoea index (AHI) results |
|------------------|------------------|
| **Sex M/F**      | 79/16            |
| **Age yrs**      | 52.5±11.3        |
| **BMI kg/m\(^2\)** | 30.7±7.3         |
| **Epworth Sleepiness Scale** | 11.5±4.9         |
| **Pretest probability %** | 67±32            |
| **AHI PSG (TST)·h sleep**   | 43.3±33.4        |
| **AHI AS·h recording**      | 34.4±24.4        |
| **AHI <5·h**       | 6                |
| **50AHI<15·h**     | 18               |
| **150AHI<30·h**    | 19               |
| **AHI>30·h**      | 52               |

*: number; #: mean±SD; M: male; F: female; BMI: body mass index; PSG: polysomnography; TST: total sleep time; AS: Autoset™.
transformation of these data showed a 95% confidence agreement of ±40% between the systems. Thus, at AHI of 5 events·h⁻¹ by AS the result will be ±2 events·h⁻¹ from PSG and at AHI of 15 events·h⁻¹ will be ±6 events·h⁻¹ and at AHI of 20 events·h⁻¹ will be ±8 events·h⁻¹, etc. AS was in agreement (95% CI) with PSG-TBT, within a mean of 1.9 (-20.7–24.6) AHI, -0.12 (-14.4–14.1) AI and 1.9 (-19.9–23.7) HI, which is a more precise comparison of event detection between the two methods.

### Recognition of obstructive sleep apnoea

Table 2 shows the sensitivity (TP/TP + false negative (FN)) and specificity (true negative (TN)/TN + FP) and positive (TP/TP+FP) and negative (TN/TN+FN) predictive values of AS compared with PSG-TST for diagnosing sleep apnoea at various cut-off thresholds of the AHI index commonly used. As the threshold increases from >15 to >30, sensitivity decreases from 92 (CI 86–97) to 79% (71–87), while specificity increases from 79 (71–87) to 93% (88–98). However, positive predictive value remains stable at 93%. This is explained by the greater number of TP (65 versus 41) and increased prevalence of OSA (75 versus 55%) for an AHI >15 compared to AHI >30.

At a threshold >15 events·h⁻¹ only six patients had an FN result from AS. Only two of these had an Epworth Score of <10 and a low pretest probability of OSA, 40 and 20% respectively. These two patients had an AHI of 16.7 and 19.6 events·hour of sleep⁻¹. Three of the mildly apnoeic FN patients had a high index of microarousals and probably had additional UARS. At this cut-off these patients had a high index of microarousals >16 (16)

Table 2. – Sensitivity, specificity, positive and negative predictive value (PV) of Autoset™ compared with polysomnography for the diagnosis of obstructive sleep apnoea at various cut-off levels of the apnoea/hypopnoea index (AHI).

<table>
<thead>
<tr>
<th>AHI</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive PV</th>
<th>Negative PV</th>
<th>FN/FP</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>97</td>
<td>50</td>
<td>97</td>
<td>50</td>
<td>3/3</td>
</tr>
<tr>
<td>15</td>
<td>92</td>
<td>79</td>
<td>93</td>
<td>76</td>
<td>6/5</td>
</tr>
<tr>
<td>20</td>
<td>86</td>
<td>86</td>
<td>93</td>
<td>76</td>
<td>9/4</td>
</tr>
<tr>
<td>30</td>
<td>79</td>
<td>93</td>
<td>93</td>
<td>78</td>
<td>11/3</td>
</tr>
</tbody>
</table>

FN: false negative; FP: false positive.

### Pretest probability

Table 3 shows the sensitivity, specificity, positive and negative predictive values of AS for a fixed threshold of OSA (AHI >15) but in different subsets of the population according to their pretest probability or Epworth score. In the group with a pretest probability between 10 and 40%, AS had a positive predictive value of 70%, while in the group with a pretest probability >80% the positive predictive value increased to 100%. This is explained by the higher number of TP (41 versus 7), the lower number of FP (0 versus 3) and the increased prevalence of OSA (91 versus 45.5%) in the group with >80% pretest probability compared with the group with lower pretest probability. There was no increase in the positive predictive value of AS in relation to the Epworth score as there were no FP results at a low Epworth score.

### Discussion

This study has shown that a simplified diagnostic system (AS) can be used to recognize confidently OSA in a quantitative manner, particularly in patients with a high pretest probability of OSA. The AS is a type IV device according to American Sleep Disorders Association (ASDA) criteria, with two channels and automated scoring which is not modifiable [21]. A pneumotachograph was used in these PSG studies to assess hypopnoeas with confidence and it was verified that this procedure did not change the recognition of hypopnoeas by AS. Significant mouth leaks were not a problem in this analysis.

The guidelines for the investigation of patients with possible OSAS differ from one country to another. The British guidelines state that oximetry alone or oximetry plus video recording may be sufficient to diagnose the condition in some patients [22]. Conversely American recommendations state that PSG should be performed in all

<table>
<thead>
<tr>
<th>Pretest probability (n)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive PV</th>
<th>Negative PV</th>
<th>FN/FP</th>
</tr>
</thead>
<tbody>
<tr>
<td>10–40% (22)</td>
<td>45.5</td>
<td>70</td>
<td>75</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>41–79% (19)</td>
<td>63</td>
<td>83</td>
<td>71</td>
<td>83</td>
<td>71</td>
</tr>
<tr>
<td>&gt;80% (46)</td>
<td>91</td>
<td>98</td>
<td>100</td>
<td>100</td>
<td>80</td>
</tr>
<tr>
<td>Epworth Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4–7 (18)</td>
<td>83</td>
<td>87</td>
<td>100</td>
<td>100</td>
<td>60</td>
</tr>
<tr>
<td>8–11 (28)</td>
<td>79</td>
<td>96</td>
<td>67</td>
<td>91</td>
<td>80</td>
</tr>
<tr>
<td>12–15 (29)</td>
<td>62</td>
<td>83</td>
<td>82</td>
<td>88</td>
<td>75</td>
</tr>
<tr>
<td>&gt;16 (16)</td>
<td>81</td>
<td>100</td>
<td>67</td>
<td>93</td>
<td>100</td>
</tr>
</tbody>
</table>

Note the increase in the prevalence of OSA as well as positive PV as pretest probabilities increase. FN: false negative; FP: false positive.
patients [23]. Although financial factors were not assessed directly in the present study, using AS as a first-line test in patients with a high pretest probability, >80% (n=46), followed by a PSG only when AS was negative (n=6), would have saved 41 polysomnographies or 43% of all those performed in this study. In the overall population, for an AHI >15 defining OSA, AS gave a negative result in 25 patients, who should have a PSG, but 70 polysomnographies would be saved.

AS was not evaluated for the diagnosis of UARS [24] because measurement of oesophageal pressure to detect intrathoracic pressure changes was not systematically performed. Fi,DREMON and RIMMERS [25] proposed using portable monitoring only in patients with a moderate to high probability of OSA, followed by complete PSG only when the first study is negative. Since in the present study four of the six FN patients had a pretest probability >60% or an Epworth Score >10, the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OAS diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.

With a threshold of AHI >15, five patients were identified as FP, which means that those patients will undergo an unnecessary treatment trial. However, three of the five FP patients had a pretest probability >60% or an Epworth Score >10, so the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OSA diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.

With a threshold of AHI >15, five patients were identified as FP, which means that those patients will undergo an unnecessary treatment trial. However, three of the five FP patients had a pretest probability >60% or an Epworth Score >10, so the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OAS diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.

With a threshold of AHI >15, five patients were identified as FP, which means that those patients will undergo an unnecessary treatment trial. However, three of the five FP patients had a pretest probability >60% or an Epworth Score >10, so the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OAS diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.

With a threshold of AHI >15, five patients were identified as FP, which means that those patients will undergo an unnecessary treatment trial. However, three of the five FP patients had a pretest probability >60% or an Epworth Score >10, so the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OAS diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.

With a threshold of AHI >15, five patients were identified as FP, which means that those patients will undergo an unnecessary treatment trial. However, three of the five FP patients had a pretest probability >60% or an Epworth Score >10, so the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OAS diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.

With a threshold of AHI >15, five patients were identified as FP, which means that those patients will undergo an unnecessary treatment trial. However, three of the five FP patients had a pretest probability >60% or an Epworth Score >10, so the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OAS diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.

With a threshold of AHI >15, five patients were identified as FP, which means that those patients will undergo an unnecessary treatment trial. However, three of the five FP patients had a pretest probability >60% or an Epworth Score >10, so the strategy of performing PSG in negative patients with moderate-to-high pretest probability or an Epworth Score >10 would have left only two FN, both with an AHI <20. Thus, if AS is to be used for OAS diagnosis, a negative result in a patient with a moderate-to-high pretest probability should be followed by PSG. Using this strategy in the present population would have an estimated sensitivity of 97%, specificity of 92%, positive predictive value of 97% and negative predictive value of 92%.
the printout to be taken as definitive. Caution should be applied in such circumstances as the greater reliability of AS is in the more severe cases. AS should be used preferentially in this group of patients to confirm disease. The determination of pretest probability using clinical prediction rules [14] or the items used in these rules, as in the present study, is a good way by which to select the population of patients who should benefit from AS. However, this means that the doctor using AS should have knowledge and experience in sleep disorders. No data are available on the subsequent management of these patients. Further studies are needed to analyse the treatment outcomes of patients diagnosed using simplified systems such as AS.

In conclusion, the Autoset™ in diagnostic mode may be a useful simplified diagnostic device for obstructive sleep apnoea if the clinical situation of the patient is taken into account. Cases where there is a strong likelihood of obstructive sleep apnoea should be restudied by polysomnography when an unanticipated negative result is found.

Acknowledgements: The authors acknowledge the help of the Sleep Laboratories participating in the study: Angers (J.L. Racineux), Grenoble (C. Feuerstein, C. Brambilla), Paris St Antoine (B. Fleury), Poitiers (J. Paquier), Marseille, Hôpital Nord (D. Charpin) and Nancy (J.M. Polo). The authors thank C. Deschaux for help with statistical analysis and M. Coisy for invaluable assistance.

References