Agreement Of Apnea-Hypopnea Indexes Based On Visual And Automatic Detection

THE siesta GROUP

Michael Woertz[1], Peter Anderer[2,3], Georg Gruber[2], Silvia Parapatics[1], Roman Rosipal[1], Bernd Saletu[3], Georg Dorffner[1,2,4]

[1] **Austrian Research Institute for Artificial Intelligence, Vienna, Austria**
[2] **The Siesta Group Schlafanalyse GmbH, Vienna, Austria**
[3] **Department of Psychiatry, Medical University of Vienna, Vienna, Austria**
[4] **Department of Medical Cybernetics and Artificial Intelligence, Medical University of Vienna, Vienna, Austria**

Introduction: The study investigates the reliability of a recently developed automatic detection system for breathing disturbances in sleep. The quality of the automatic apnea detection software, which is a modular part of the Somnolyzer24x7 framework, is determined using both, correlation measures and the Bland-Altman method. Results are based on data recorded in the SIESTA project.
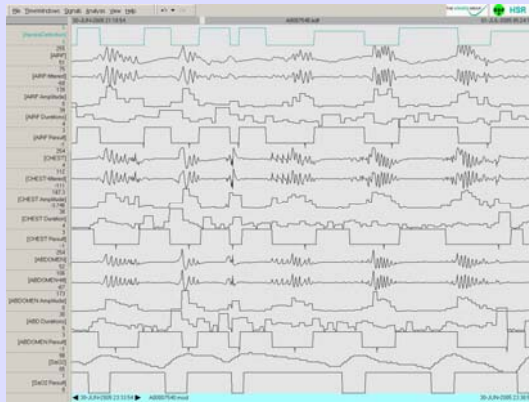

Figure 1: Example of Desaturation, Apnea & Hypopnea detection

Method: The apnea detection algorithm uses 4 polysomnographic signals: oxygen saturation, nasal airflow, movement of the chest wall and of the abdomen. Oxygen desaturations are detected by filtering and peak extraction. Intervals of decreased airflow are calculated based on the determination of single breaths. From both breathing effort signals, intervals with events are extracted similarly (see figure 1). A combined effort channel is calculated to detect breathing events, which cannot be seen in the airflow. The final detection of central, mixed and obstructive apneas as well as hypopneas is then accomplished with a decision tree. Calculated results are summarized in the Somnolyzer 24x7 spot report (see figure 2).

Up to now, there is obviously no commonly accepted standard for scoring respiratory events (see table 1). The presented automatic system has been designed to allow for an easy adaptation of detection criteria.

Table 1: Different apnea scoring criteria.

Def.	Apnea	Hypopnea
1	At least 90% reduction in airflow for 10 seconds. *Obstructive*: paradoxical effort. *Central*: no effort. *Mixed*: starting central and ending obstructive	At least 30% reduction in effort or airflow for 10 seconds, together with a 4% desaturation. If more than 50% reduction, a desaturation is no longer required
2	*Obstructive*: at least 30% reduction in effort or airflow for 10 seconds, together with a 3% desaturation or an arousal. If more than 50% reduction, a desaturation or arousal is no longer required. *Central*: same as obstructive, but with a clear parallel reduction in effort	Abnormal respiratory event lasting at least 10 seconds with at least 30% reduction in effort movement or airflow and with at least 4% desaturation
3	A cessation of airflow for at least 10 seconds	Abnormal respiratory event lasting at least 10 seconds with at least 30% reduction in effort movement or airflow and with at least 4% desaturation

Data: For the present comparison, polysomnograms of 51 subjects (44 males and 7 females, aged 51+/-10 years) were investigated. The visual inspection of the data was done by human scorers from different European sleep labs.
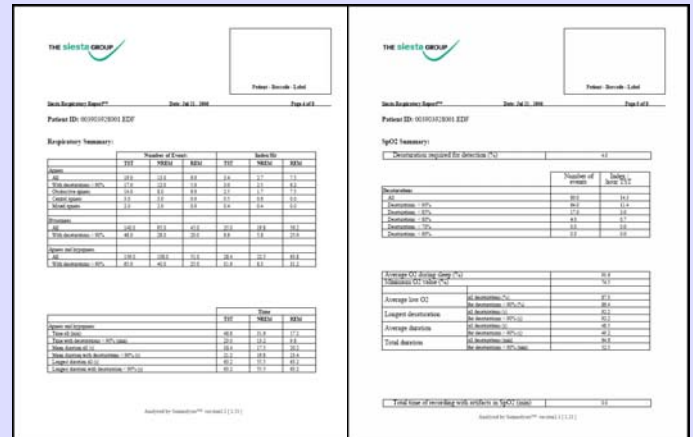

Figure 2: Respiratory related parts of the Somnolyzer 24x7 spot report (pages 4 and 5, respiratory and SpO_2 summary).

Results: Subjects had a human scored AHI of 45+/-31 (mean +/- standard deviation) during the adaptation night and 41+/-26 in the second night. The automatic detection resulted in an AHI of 42+/-26 and 41+/-22. The correlation between human scorers and the automatic method was r=0.94 for the first night and r=0.92 for the second night. Following the suggestions of Bland and Altman (Lancet, 1986; pp. 307-310), that correlation alone cannot express the agreement of two methods properly, the analysis showed absence of significant bias and the plots suggested equivalence of human scoring and the automatic apnea detection (compare figure 3).
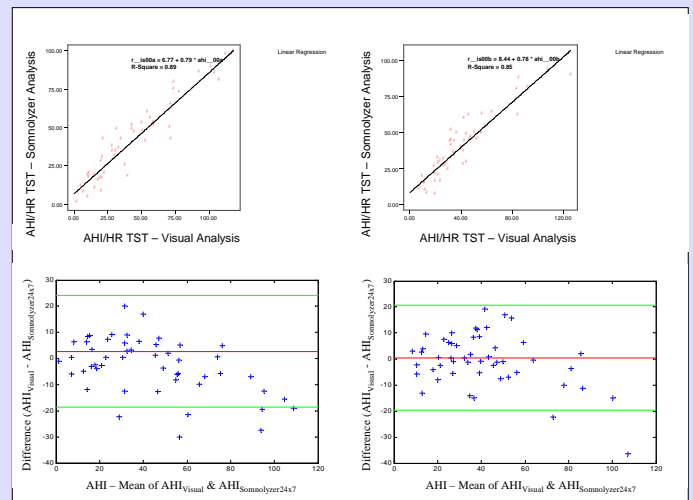

Figure 3: Left panel (first nights): the upper plot shows the correlation between human scoring and automatic apnea detection. The lower plot shows the Bland and Altman measurement of differences. The right panel displays the results for the second nights.

Conclusion: *A high correlation between human scorers and the developed apnea-hypopnea detection has been found. The Bland and Altman measurement of agreement did not reveal a significant bias between both approaches, which justifies the further usage of the automatic apnea-hypopnea detection. The system itself, being a part of Somnolyzer 24x7, is modular and flexible, which allows the easy adaptation to lab-specific apnea-hypopnea scoring rules.*