

## REVIEW ARTICLE

# Computer based sleep recording and analysis

Thomas Penzel and Regina Conradt

*Medical Policlinic of Philipps-University, Marburg, Germany*

---

*Sleep analysis is based on polysomnography. Modern polysomnographic systems are computer based. Visual and automatic analysis of sleep and respiration is supported by most computer based systems. Four functions can be distinguished in computer based polysomnography: recording, documentation during the recording, automatic and visual analysis and report generation. This review compiles the minimal requirements for digital sleep recording, documentation, analysis and reporting. The basic principles of automatic sleep analysis are reported. The requirements and the basic principles for the analysis of non-electroencephalography (EEG) signals, such as respiration, snoring, oxygen saturation, electrocardiography (ECG) and options are reported. New developments in sleep EEG processing are discussed to enlighten how computer based sleep analysis can add quantitative parameters to the rules for visual sleep staging established by Rechtschaffen and Kales 30 years ago. This helps to extend our understanding of sleep.*

© 2000 Harcourt Publishers Ltd

**Key words:** polysomnography, sleep recording, sleep analysis, sleep scoring, EEG analysis, computerized sleep analysis, respiration analysis.

---

## Introduction

Sleep scoring is the fundamental basis of sleep research and sleep medicine. In 1998 we can look back on 30 years of sleep scoring based on rules which were compiled by a committee chaired by Rechtschaffen and Kales [1]. Today it is widely recognized that these rules have severe limitations which were not foreseen 30 years ago [2–4]. Nevertheless, these rules have survived all criticism raised in the past. The unbeaten advantage of these rules is that they extended the earlier observations of Loomis [5], who described sleep stages from A to E, with the observations from Aserinsky and Kleitman describing rapid eye movement (REM) sleep [6]. Another reason for their continuous use is that the guidelines give clear directives for most situations and they can be learned with minimum effort. The limitations of these rules are discussed extensively by Himanen and Hasan [4].

Attempts to develop automatic sleep scoring using computer technology are as old as the rules of Rechtschaffen and Kales (R&K). Computerized electroencephalography (EEG) analysis is even older [7]. The development of automatic sleep scoring was

---

Correspondence to be addressed to: Dr Thomas Penzel, Medical Policlinic of Philipps-University, Baldingerstr. 1, D-35033 Marburg, Germany. Tel: +49-6421-286435; Fax: +49-6421-285450; E-mail: [penzel@mail.uni-marburg.de](mailto:penzel@mail.uni-marburg.de)

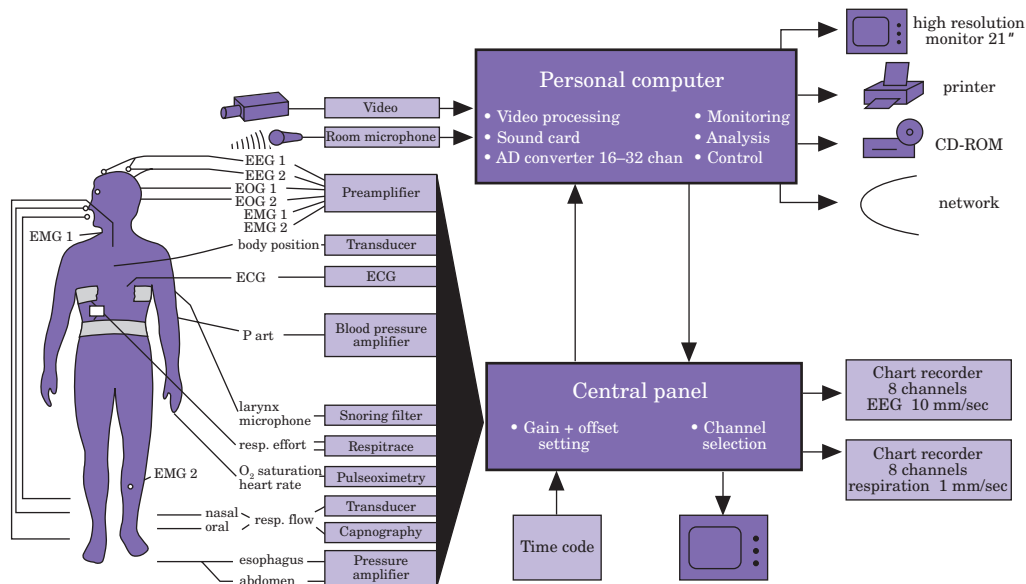
driven by two objectives. Firstly, visual sleep scoring is very time consuming, especially in patients with disturbed sleep; automatic sleep scoring should help to reduce the time needed by an experienced sleep scorer. Secondly, the limitations of the R&K rules in describing certain sleep disorders have become clearer during the past decade, with increased hope that automatic sleep scoring is able to give new quantitative measures which correspond better to the extent of these disorders [8]. Today, both goals have been reached in part, but the results of automatic sleep scoring systems currently available are not good enough to replace the established R&K rules. In this review the current concepts and results of computerized polysomnography and automatic sleep scoring systems are presented and discussed.

The review is based on a number of fundamental papers on sleep scoring and a number of related reviews. In addition a MEDLINE search was performed on the terms "sleep scoring", "sleep analysis" and "polysomnography". We found two earlier reviews which also used MEDLINE searches. Searches were performed using the keywords "Sleep—computer—scoring" in one review [9] and "polysomnography" in the other review [10]. The MEDLINE search on sleep—computer—scoring covered the years 1966 to the beginning of 1993, and the intersection of the terms provided 56 references. The evaluation revealed that 16 papers were on sleep stage scoring in humans, and only six papers assessed system function of computerized sleep scoring systems. The other papers were related to apnoea scoring, use of computers in sleep research and sleep studies in animals. The new search covered the years 1966 to September 1998 and identified 77 papers for sleep analysis and 44 papers for sleep scoring. The other paper presenting a MEDLINE search for polysomnography covered the years 1991 to January 1997 and identified 1614 references [10]. The new search using this term found 2731 papers. These searches were only of limited value because the most important papers used for this review were not included.

The review also takes into account our own experience in the development of automatic sleep analysis systems since 1982 and experience with a number of different commercial sleep analysis computers [11]. In 1997 a new international initiative was started to develop a new standard for automatic sleep analysis. Within this European project normal volunteers of different age groups and patients with selected sleep disorders were recorded using a common protocol. The recordings are evaluated using modern tools of signal processing to gain new parameters which describe sleep [12, 13]. The parameters are then validated against reaction time, vigilance, performance and other psychometric measures obtained during extensive testing the next day. The SIESTA project is currently funded by the European Union. Some of the basic ideas are also included in this review.

#### **Four functions of digital polysomnography**

In general, automatic sleep scoring should fulfill four major tasks [14]. The first task is that systems should replace conventional paper chart recorders, in a similar way to that in which computer based systems allow paperless EEG recording in neurophysiology today. The intention of this function is to produce less paper and minimize space requirements for archiving without losing the raw data. The second task is documentation. With the computer based system a technician should be able to enter all additional notes and observations made during the nocturnal recording, that have previously been documented on paper. The third task is evaluation of sleep and



**Figure 1.** A schematic overview of an automatic sleep scoring system. A computer based system consists of sensors, amplifiers, filters, analogue-digital converters and a computer with peripheral devices. On the right side of the diagram all peripherals essential for documentation can be seen.

cardiorespiratory functions. An automatic sleep scoring system should use its computational power to support sleep evaluation. The system should analyse EEG, electrooculography (EOG) and electromyography (EMG) in terms of sleep stages; respiration, snoring and oxygen saturation in terms of sleep related breathing disorders; and EMG tibialis in terms of movement disorders. Other parameters recorded in a sleep laboratory, such as body temperature, ECG, blood pressure, capnography and oesophageal pressure may require additional analysis. The system should support visual evaluation as an alternative to automatic analysis and it should allow the editing of the results of automatic analysis. The fourth task is reporting. The computer based system should help in the generation of a final report of the investigation, and should include an advanced filing system to archive the report as well as the data in a structured way on CD-ROM or other high capacity media. This will enable the sleep laboratory to keep track of patients and to recall reports when needed. This option enables the review of former polysomnographic recordings, which is seldom used today due to the difficult access to old paper recordings, in conventional paper archives. (Fig. 1.) The specifications for the four functions mentioned are now given in detail.

### Chart writer

To fulfil the needs of a chart recorder for polysomnography, a computer based system should be able to record all signals without introducing noise or artefacts. These requirements are mainly related to the quality of the amplifiers and the digital conversion process of the signals. Minimal criteria for digital recording of polysomnographies are given in Table 1 [14].

**Table 1** Minimum requirements for digital polysomnography as a basis for automatic sleep scoring

Function	Signal	Sampling rate	Digital resolution
Sleep	EEG	100 Hz	1 $\mu$ V/Bit
	EOG	100 Hz	1 $\mu$ V/Bit
	EMG	100 Hz	0.5 $\mu$ V/Bit
Respiration	oronasal airflow	25 Hz	n.a.
	respiratory movements	25 Hz	n.a.
	oesophageal pressure	100 Hz	0.5 mmHg/Bit
	capnography	25 Hz	0.1%/Bit
	oxygen saturation	1 Hz	1 %/Bit
	transcutaneous $PO_2$ , $PCO_2$	1 Hz	0.1 mmHg/Bit
	snoring integrated	1 Hz	n.a.
	snoring raw	100 Hz	n.a.
Cardiovascular	ECG	100 Hz	10 $\mu$ V/Bit
	heart rate	1 Hz	1 bpm
	blood pressure	100 Hz	1 mmHg/Bit
Movement	EMG	100 Hz	0.5 $\mu$ V/Bit
	body position	1 Hz	n.a.

The digital amplitude resolution is chosen according to the measurement precision of the underlying instrument. (n.a.: non applicable.)

### ***Sampling rate requirements***

Several papers of recommendation have discussed the need for a sufficient sampling rate for EEG signals [9,14–17]. For a computer based frequency analysis of EEG, signals at twice the highest frequency of interest are required according to Nyquist's law. This requires a sampling frequency of 32 Hz for 16 Hz sleep spindles. However, if small changes in frequencies (i.e. shifts of 0.25 Hz in sleep spindles) are of interest then very high sampling frequencies such as 1024 Hz are needed for quantitative signal analysis [15]. Currently 100 Hz is regarded as the minimum acceptable sampling rate for EEG, EOG and EMG in sleep recordings when using 35 Hz high filters [17], a sampling rate of 200 Hz is required when using 70 Hz high filters [17]. International recommendations for ECG require a sampling rate of at least 250 Hz for long-term ECG recordings [18]. These recommendations also require that at least two channels of ECG are recorded. If the ECG recorded in a sleep laboratory is only used to estimate the heart rate and not to investigate heart rate variability, then a sampling rate of 100 Hz may be sufficient. This results in an accuracy of 10 msec when calculating R-R intervals which is sufficient for clinical interpretation [14]. A sampling rate of 100 Hz for oesophageal pressure is recommended in order to allow a detailed analysis of this signal. According to our observations small and very rapid swings in oesophageal pressure indicate snoring. This is only visible with a high sampling rate and no high filters for this signal.

### ***Digital amplitude resolution requirements***

Few recommendations discussed the amplitude resolution necessary for sleep recordings [9,14,16,17]. The initial amplitude resolution selected is essential because

all later software based amplification cannot go beyond the resolution chosen during the initial digital conversion. Previous studies have shown that in order to obtain at least a 40 dB signal to noise ratio at the output of digital filters, a 12 bit analogue–digital conversion is needed [19]. To resolve EEG to a resolution of  $0.5\ \mu\text{V}$  a 12 bit A/D conversion is required [17]. Therefore a 12 bit resolution is the minimum acceptable for sleep recordings, providing a range of values between  $-2048$  and  $2047$ . A 16 bit resolution improves software based amplification and reaches the limits of electric noise in the electronic EEG amplifiers [14,16].

### ***Scroll back mode***

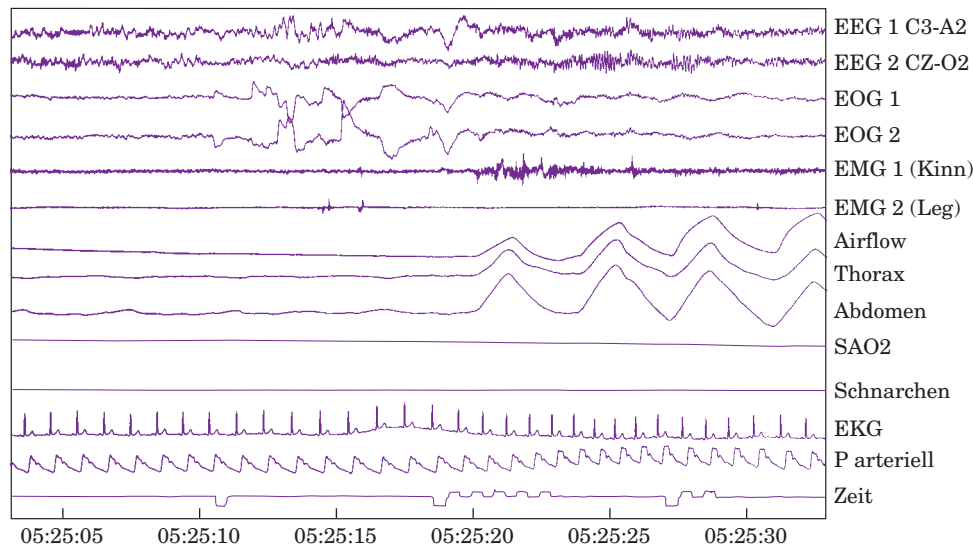
Another important requirement for chart recorders in polysomnography is the ability to review the preceding minutes and hours of the recording as recording continues [9, 20]. Previously, there has been no need to explicitly mention this because the paper of the ongoing recording was fully available for supervision. Computer based systems usually display only one screen showing the last 30 sec of the recording. Due to this severe limitation, the computer program should offer a mode to scroll back in the recording without interruption of ongoing data acquisition. This option becomes very important if there is a need to compare acute changes of EEG or ECG with the signals recorded earlier on the same night to observe acute seizures or cardiac arrhythmias.

### **Documentation**

A full visual scoring of the recording according to the rules of R&K must be possible. For this purpose the computer screen for review of the recording should have a sufficiently high resolution [9]. The computer screen must be at least 20 inches in size and display a resolution of  $1280 \times 1024$  pixels flickerfree (i.e. 75 Hz monitor scan rate) [14]. Compared to paper scoring the main advantage of digital sleep scoring is that the scorer is able to zoom in or out during the review of the recording to improve the evaluation of patterns. Unfortunately, there are no studies which compare paralleled visual computer screen based scoring with paper based scoring to show the effectiveness of this approach. At this time there also have been no studies which determine inter- and intra-rater variability of visual computer screen based scoring. One early study studied inter-rater variability using an Oxford Medilog 9000 screen based on 16 sec epochs (agreement 85.5%) and compared this with inter-rater variability for conventional paper scoring (agreement 87.8%) [21]. Hard copies of the recording on paper have to preserve the high resolution of signals as offered by 300 dpi printers (Fig. 2).

### ***Archiving of sleep recordings***

The hardware of computer systems changes rapidly these days, as well as the operating systems and the application software. Digital documentation and storage of raw data only improve accessibility if new systems are still able to work with the media and data formats used, while storing the information from a couple of years ago. This is not a trivial issue if one remembers the numerous types of optical disks and the various formats of floppy disks (i.e. 8 inch; single-sided; double-density) which are no



**Figure 2.** An example of digital polysomnography with sufficient resolution in sampling rate and amplitude.

longer supported by systems used in sleep laboratories. In addition, magnetic media degrade over the 10 years in archive [17]. Therefore, standards for media type and data formats are very important for documentation, and in addition they are a prerequisite for data exchange between laboratories. The most common media for storing digital polysomnographies are presently CD-ROMs and they are sufficiently standardized. Depending on the number of channels and the sampling rate, one to 20 recordings can be stored on one disk. CD-ROMs are produced at low cost and a CD-ROM drive for reading data is very inexpensive and widely available. For these reasons, the CD-ROM is the current media of choice for storing digital polysomnographs.

The most common format for exchange of digital polysomnographies today is the "EDF format", which is simple in structure and allows storage of raw signals sampled at different sampling rates, together with their calibration information and some general information on the recording conditions [22]. Additional comments or markings have to be stored in other non-standardized files. In order to add this information and the results of analysis a new file format is under preparation. This new file format will be worked out together with companies and will become an international standard after its general introduction. It can be foreseen that the EDF format will be a subset of the new format to allow backward compatibility.

### ***Video recording of sleep***

The ability to observe the patient's behaviour and to monitor any vocalization or snoring during sleep is a crucial and integral part of polysomnography [23]. This is most easily done through video monitoring [20,23]. The video recorder, video monitor and camera control should be at the desk to permit closeup views of events from the technician [23]. Video tape recording is the most common technique still in use. Some advanced computer based sleep systems have integrated video boards which allow



simultaneous recording of digital video with sufficient resolution in time and space. These recordings use a lot of space on the computer hard disk. Today, a compromise between storage capacity and resolution of the video has to be taken. Often only short segments of digital video are actually saved. With increasing capacity of new storage media this will become less important in the very near future. The new DVD media (digital versatile disk) will be appropriate for storing digital videos of sleep recordings.

### **Automatic analysis**

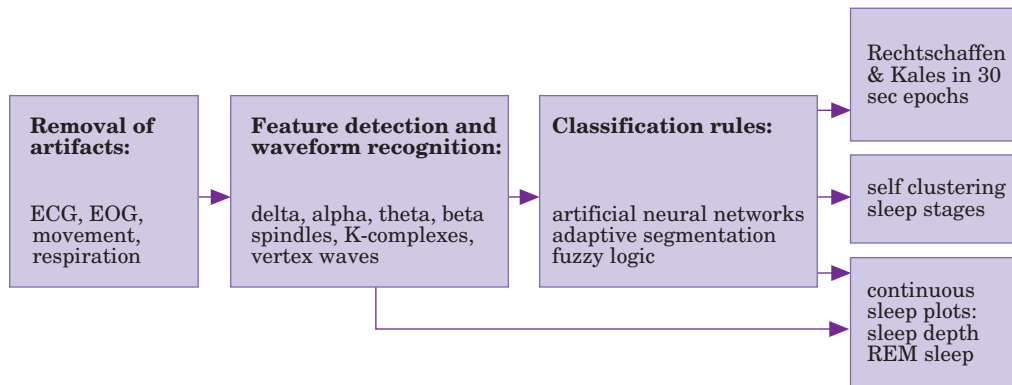
At first automatic analysis of sleep focuses on EEG, EOG and EMG. Besides this traditional core sleep analysis, modern systems also have to provide evaluation of respiration, snoring, oxygen saturation, ECG, heart rate and limb movements. In addition, body position must be considered for the evaluation of respiration. Options for body temperature, capnography, blood pressure and oesophageal pressure are desirable.

### ***Principles of automatic sleep scoring***

Many publications are available on the different principles used for automated sleep analysis as reflected by the Medline search. Some reviews present a comprehensive overview of different methodologies [11,16,24,25]. Many studies are technically oriented, which means a new principle is developed and tested with a limited number of recordings, usually between six and a few dozen subjects. These studies usually investigate the performance of the new analysis using recordings of normal volunteers, or sometimes using selected patient groups. Only a few methods have been tested in subjects with different sleep disorders and the results compared.

Automatic sleep analysis always consists of a number of consecutive steps (Fig. 3). The first step is signal preprocessing. The purpose of preprocessing is to reduce the vast amount of raw data to an amount which can be managed by statistical tools. Typically, the resolution of 100 Hz for raw data is reduced to a resolution of 1 Hz as the result of preprocessing [8,16]. The preprocessing yields a number of so called features. These are the amplitude or the power of regular waves, such as beta, alpha, theta and delta waves, and are also specific patterns such as K-complexes, sleep spindles and vertex sharp waves. In EOG the specific patterns are rapid eye movements and slow eye movements which have to be recognized and distinguished from each other [26]. An important objective for preprocessing is to consider artefact rejection for all signals being analysed. The most important artefacts in the EEG signal are ECG, EOG, EMG, body movement and electrode artefacts due to movement or sweating. Artefact rejection can be achieved by digital filtering or by correlation analysis.

The second step of the analysis is the combination of the extracted features and waveforms to a limited number of sleep stages. The decision of a sleep stage based on the features is often done using logic rules. The rules may be set once during the development of the program or may be adapted, taking into account new recorded data to be continuously improved. A very popular technique of building rules is an artificial neural network. In this case the rules are constructed by feeding classified data into a learning system which is explained later. In systems used for clinical purposes the final result of automatic sleep scoring has to come close to a visual R&K



**Figure 3.** Schematic diagram of steps involved in automatic sleep scoring. The result of automatic sleep scoring can be different. The results indicated may even coexist as simultaneous outputs of automatic sleep analysis. The different results may show advantages under different conditions.

scoring. In addition, automatic sleep analysis systems can also provide measures of the number of delta waves, alpha waves, sleep spindles and other variables which may give more details about sleep. Some basic techniques of preprocessing commonly used and techniques for the combination of features are reported.

### ***Preprocessing of EEG signals***

Spectral analysis has the longest tradition of all EEG analysis techniques because it can be applied to quantify the different frequency contents of the signal similar to visual analysis, according to Loomis [5]. Spectral analysis is easily performed with the help of fast fourier transformation algorithms (FFT). Principal problems associated with this analysis are that the FFT assumes that the signal is stationary in a mathematical sense while the signal segment is being analysed. Therefore, the choice of the duration of the signal segment is essential. The duration of the appropriate segment or epoch actually depends on the EEG activities to be detected and to be allocated in time. The frequency components identified by FFT can no longer be attributed to one specific moment within the epoch analysed. Usually this is not of interest for alpha, delta, theta and beta activity using an epoch duration of 30 sec. However, having found sigma activity in an epoch of 30 sec does not allow the estimation of the number of sleep spindles nor their location within this epoch—which in contrast might be of interest. To allocate sleep spindles the signal segment for FFT should not be longer than 2 sec [27].

Together with the sampling rate, the signal segment determines the frequency resolution of the FFT. Whereas half of the sampling rate is the highest possible frequency found by FFT, the inverse of the signal segment defines the frequency resolution: a segment duration of 2 sec yields a frequency resolution of 0.5 Hz (this originates from Nyquist's law). Zero-padding, which is a computational technique to artificially increase the segment duration being analysed, may overcome this problem with increased computational effort [11]. Short-time fourier transformation is another technique based on a two dimensional time-frequency transformation to overcome



this problem without increasing the computational effort. Another new approach is the fast time frequency transformation which has proved its usefulness in a new sleep analysis system [28].

Another problem using FFT is that some definitions for wave classification, such as the criteria for delta waves ( $75 \mu\text{V}$ ), are amplitude dependent. As pointed out, the calculated power given by the FFT does not give the amplitude of any specific wave in the signal segment analysed.

Spectral analysis can be done either by FFT, as described above, or by autoregressive filtering, the latter allowing continuous calculation of spectral power [13,29]. Using this approach, spectral analysis is no longer bound to consecutive segments of fixed durations and values can be updated every second [30].

Calculating the power in frequency bands is in general problematic due to fixed frequency boundaries. Frequency bands can be too narrow or can be defined with unsuitable frequency boundaries. One approach for overcoming this limitation is to use relative percentages of the total power and determine the corresponding frequencies [27]. These variable frequencies have to be interpreted as defined percentiles of a varying frequency distribution (i.e. 95 percentiles or 66 percentiles).

In contrast to all frequency analysis techniques mentioned so far, period analysis is a preprocessing method which comes closer to the visual analysis performed by human sleep scorers [19]. Period analysis is a time domain analysis implemented either as zero crossing or as peak detection of the EEG waveform. Period analysis is also popular due to the fact that it needs minimal computational effort. A systematic comparison of period amplitude analysis and spectral analysis showed that both methods are equivalent for low delta frequencies and spectral analysis is superior for higher frequencies [31].

As all preprocessing methods have their specific advantages and disadvantages it seems to be best to choose a combination of different methodologies for the ideal automatic sleep scoring system.

As rapid fluctuations in vigilance may occur not only in patients with sleep fragmentation, preprocessing at the start of analysis has to be performed in short epochs to avoid a loss of time resolution. Adaptive segmentation is superior to analysis based on fixed epoch duration [32]. To accommodate this, the concept of adaptive segmentation was introduced by Praetorius and Bodenstein [32]. Studies which used this concept for sleep analysis [30] and for drowsiness and sleep onset [33] demonstrated advantages compared to analysis based on fixed epoch duration. It was possible to follow better sleep stage changes and vigilance fluctuations during sleep onset.

### ***Combination of features derived from EEG analysis***

The next step in automatic sleep analysis is the combination of features of each segment according to rules used to determine final sleep stages. This step is a lot more difficult than signal preprocessing because it is no longer based on quantitative signal analysis. Subjective decisions influence the result because the classification of some segments remains ambiguous. The final result of this step is evaluated by counting the agreement between automatic analysis and visual analysis. A number of publications are available on automatic sleep scoring in normals [13,31,34,35]. Reliability varies between 70% and 90%. Fewer evaluations have been published about man-machine agreement in

disturbed sleep, usually in selected sleep disorders [36–39]. In these studies reliability varies between 65% and 87%.

Since the end of the 80s, neural networks have been used as a new technique for improving the automatic classification of sleep stages based on prior feature detection [40]. Artificial neural networks are computational tools composed of a large number of connected elementary processors, called neurons. The information stored in each neuron represents mathematical weight factors which describe the computational influence of all connecting neurons. The basic idea is that a global coherent behaviour can evolve from such a network. In the process of decision making each neuron sums up the weighted inputs it receives from other neurons and generates an output value which is passed to further neurons. An individual network can be characterized by the number of neurons and the connections between the neurons. Typically, the neurons are grouped in layers and the most popular model is a multilayer network with one input layer, one output layer and one or more hidden layers in between. The weight factors of the connections are adjusted during a learning phase. Supervised and unsupervised learning can be distinguished. In both cases data of preclassified sleep stages are presented to the network. Unsupervised learning tries to catch regularities in the input data. Supervised learning tries to present selected reference patterns to the network to adjust the weight factors of the connections. Several implemented systems were developed for sleep classification in animal research, some systems only distinguish some sleep stages. Systems developed by two groups try to mimic the complete R&K sleep staging. Schaltenbrand *et al.* trained their system based on twelve night recordings and tested it on eleven other recordings [41]. Agreement between visual consensus scoring and the system was 80.6% with lowest agreement for sleep stage 1 and 3. The authors further improved their system and presented a bigger study including different disorders [38]. Interexpert agreement in 60 subjects was 87.5% and expert-machine agreement was 82.3% on average. Based on the analysis of nine healthy subjects Roberts and Tarassenko concluded that, although they had derived an automated hypnogram, this is not the best method for detailed investigations of the sleep process because of the poor temporal resolution and the limitation imposed by having discrete stages [42]. Later they presented a system based on autoregressive modelling for feature extraction and an artificial neural network [13]. In their study Pardey *et al.* follow the concerns about imitating R&K scoring with computers as raised earlier [42]. They now present a new approach by giving continuous measures in parallel with automatically calculated sleep stages [13]. The background is given in the next paragraph.

### ***New approaches in automatic sleep processing***

Because all of the methods, except the very last one, described so far try to mimic the rules of R&K they all have to struggle with the well recognized limitations [2,4,13,16]. A recognized weakness of R&K scoring is the non-reflected microstructure of sleep [3, 43]. By introducing adaptive segmentation, quantitative analysis of EEG waves and continuous parameters, several groups have tried to overcome the limitations of R&K. These new parameters extend the value of automatic sleep scoring and help us in a better understanding of the sleep process. In 1970 Hjorth introduced the continuous parameters, activity, mobility and complexity, which were statistical measures of the EEG, to obtain “sleep profiles” [44]. Later, Hausteiner *et al.* used EEG filtering and EMG

analysis to derive two parameters which describe sleep depth and allow the recognition of REM sleep [45]. Both methods were based on 10 sec intervals. Continuous parameters were refined with respect to resolution in time. The proposal of Kemp consists of a continuous sleep depth scale ranging from 0–100% and reflecting the NREM sleep similar to delta power and an on/off switch for REM sleep on the basis of 1 sec intervals [8]. This proposal assumes that a NREM sleep process and a REM sleep process can be simultaneously active. The system developed by Pardey *et al.* [13] which has already been mentioned, provides three continuous output variables derived from a neural network. The three variables range from 0–1.0 and reflect probabilities. The variables correspond to “wakefulness”, “REM/light sleep” and “deep sleep”. It is possible to calculate conventional sleep stages from these variables, but the rapid fluctuations of these variables allow a better estimation of arousals and sleep fragmentation.

The search for useful variables has also attracted the tools of chaos theory to EEG analysis. The main question in applying this theory is whether the EEG is the result of a stochastic process with changing conditions or whether the EEG is the result of non-linear deterministic processes such as “wave generators”. Achermann *et al.* applied chaos theory to sleep EEG to investigate this question [46]. By calculating the correlation dimension they conclude that sleep EEG is not generated by a chaotic attractor, which means that deterministic processes have a relevant influence on the sleep EEG. Röscke *et al.* calculated the first positive Lyapunov exponent and the correlation dimension in 15 healthy male subjects [47]. They found that as the amount of delta waves during slow wave sleep increases the correlation dimension value decreases. This uncovered a reduced complexity of the EEG during slow wave sleep which is regarded as a non-trivial result using this methodology [47]. Another paper investigates both parameters using sliding windows of 1024 sec duration and their relation to sleep stages in five normal subjects [48]. The authors correlate the two parameters with sleep stages and are optimistic that the method uncovers underlying neural processes in the brain. These studies show that new intriguing methods can provide new promising variables, which then have to prove their usefulness in larger clinical trials.

It is clear at this point that automatic systems which try to imitate visual scoring according to R&K cannot be improved very much beyond the results given today as long as the agreement between different scorers is in the same range as the agreement found between man and machine.

### Practice Points

To summarize the principles of automatic sleep scoring, the strengths are:

1. automatic removal of artefacts;
2. good quantitative evaluation of delta waves (spectral power and peak analysis);
3. if signal quality is good, automatic analysis over-rules visual analysis in terms of precision and reliability.

Unfortunately some weaknesses remain:

1. sleep stage 1 and REM sleep are difficult to distinguish due to similar EEG; EOG is indispensable here;
2. wakefulness and REM sleep are difficult to distinguish because this depends heavily on the quality of the EMG signal—and this is often a problem due to electrode fixation;

3. sleep stage 2 may be difficult to define, if the person has only few sleep spindles or if the spindle frequency is outside of the normal values.

## Principles of automatic non-EEG analysis

### Respiration

The automatic analysis of respiration is highly dependent on the type of signal and this depends on the type of transducers and on the fact that some signals are not calibrated. It is known that thermistors and thermocouples do not correlate well with ventilated airflow quantified by a pneumotachograph. It is also recognized that piezo strain gauges and other length dependent measures of respiratory movement are very sensitive to body position and movement artifacts during sleep. Respiratory inductive plethysmography (RIP) appears to be the most reliable non-invasive measure of respiratory effort. Therefore, the quality of automatic analysis is much more dependent on equipment than EEG and EOG analysis. One recent study presented criteria for automated detection and classification of sleep disordered breathing [49]. To overcome the limitations mentioned Taha *et al.* started with the detection of desaturation and then analysed the sum of RIP. They defined hypopnoea as a reduction in the sum of RIP of at least 20% for three breaths. Other recommendations use a reduction of at least 50% for hypopnoea [20,50]. Unfortunately, no standards have been set until recently [51]. Apnoeas are defined as periods of at least 10 sec duration with no airflow [49,50]. Using indirect measurements of breathing, and often uncalibrated signals, it is difficult to obtain a zero flow signal during obstructive apnoeas. Therefore, Taha *et al.* allowed a flow rate tolerance of 25 ml per second to the baseline [49]. Penzel *et al.* recommended defining an apnoea as a reduction in the sum of RIP of at least 90% for 10 sec [20]. Finally, a general problem with all respiratory amplitude definitions remains, and that is the definition of baseline values. These values are derived from regular breathing intervals prior to apnoeic or periodic breathing and may be adapted over night. The analysis of the breathing amplitude can be achieved by the analysis of the first derivative of the signal. Additionally, the second derivative of the signal may be used to improve recognition [52].

### Snoring

Microphones are either attached to the skin or are placed somewhere near the head. The difficulty in measuring and quantifying snoring using objective criteria is that snoring is a subjective perception of the listener [53]. The analysis of snoring is highly dependent on the type of signal recorded. Sometimes the output of a microphone is recorded in a similar manner to EMG signals [54]. Usually only increases and decreases in total power are considered. The use of a snoring filter with selective frequency amplification has proved to be less sensitive to environmental noise, thus allowing a more sensitive estimation of snoring [55]. Interval analysis of identified snores may allow the separation of patients with apnoea from those with regular snoring [55]. A relative evaluation of loudness with a sound level meter may be interesting, but as

long as the relation between loudness of snoring and upper airway obstruction remains unsolved the relevance of such a measure is unclear [56].

### ***Oxygen saturation***

Drops in oxygen saturation are taken into account if the drop is at least 2% at a rate of 0.1% per second [49] or 4% from baseline during regular breathing [20,50]. Despite the same problem with proper baseline definition as has been discussed in the respiration section, regarding respiratory amplitude definitions, this analysis is very straightforward and certainly is more accurate than visual analysis. The drops in oxygen saturation have to be associated with apnoeic and hypopnoeic events [49]. An oxygen desaturation index gives the number of events per hour of sleep. Overall level of oxygen saturation can be a useful measure of the respiratory situation if partial pressures of oxygen and carbon dioxide are available.

### ***Electrocardiogram and heart rate***

Automatic evaluation of ECG [18] and heart rate [57] is well established and standardized by the American Heart Association and the European Society of Cardiology. In cardiology this measure only differentiates between night-time and day-time. Together with sleep stage evaluation, sophisticated measures of heart rate variability can be related to sleep stages to check for abnormalities during REM sleep or NREM sleep. Also, new techniques which analyse beat-to-beat intervals (i.e. Poincare plots) are used and provide insight into the different regulation of the autonomous nervous system in wakefulness, NREM sleep and REM sleep [58]. Cyclic variation of heart rate can be used to support the diagnosis of sleep related breathing disorders [55,59].

### ***Leg movements***

The analysis of leg movements is part of standard polysomnography evaluation. Scoring rules were set by an atlas task force of the American Sleep Disorders Association [60]. The rules set are straightforward and can be implemented in computer based polysomnographic systems. Leg movements are defined as 0.5–5 sec bursts of anterior tibialis activity with an amplitude above 25% of initial calibration movements. To score a periodic leg movement sequence, four or more leg movements separated by 5–90 sec have to be found in any one of the sleep stages or wakefulness [20,60].

### ***Optional non-EEG signals***

The computer based evaluation of blood pressure is dependent on the type of signal recorded. Non-invasive intermittent blood pressure disturbs sleep and is therefore of limited use. Non-invasive continuous blood pressure is very useful but only available in a few sleep laboratories. This method does not induce additional arousals in normal volunteers or in patients with hypersomnia. Invasive continuous blood pressure can only be recorded under special research conditions due to its invasive nature and the

risks associated with measurement. Standardized evaluation techniques for this signal have not been established until now. Systolic, diastolic and mean pressure values, as well as the pulse rate, can be obtained for each heart beat with limited computational effort [61,62]. The traces of systolic, mean and diastolic pressure can be used to follow blood pressure changes during different sleep stages. REM related hypertension and cardiovascular consequences of sleep related breathing disorders can be quantified and documented. Further evaluations, such as pulse transit time, may add information to detect microarousals and explain hypersomnolence [63].

Computer based analysis of capnometry is now under development. It can be expected that recommendations will soon be available. Oesophageal pressure seems to be a difficult signal for automatic processing. The application of different digital filters allows separating of respiratory components and snoring components in a satisfactory way. Artefacts in the signal due to swallowing and movement have to be considered prior to automatic analysis of the filtered signal. Body temperature is often influenced by numerous artefacts. Therefore, this signal needs careful editing prior to statistical analysis. Time of minimum temperature and the absolute difference between lowest and highest value are of diagnostic importance.

### **Report generation**

After analysis the computer based sleep scoring system should document all results relevant for diagnosis and should present them in a structured way. Systematic overviews of relevant parameters are given in several publications [10,20]. It should be clear whether or not the results presented were calculated automatically, were edited in addition or were the result of visual evaluation. Graphic hypnograms and standard sleep measures have to be included. Sleep measures are total number of minutes and percentages of time spent in sleep stages, sleep efficiency, latency to sleep stages and number of periods of wakefulness. The criteria for these measures should be stated in the documentation, i.e. whether sleep onset is defined as onset of sleep stage 2 or sleep stage 1. Total sleep time, sleep period time and time in bed are other standard sleep measures. A table of sleep stage shifts is also useful in sleep disturbed patients. The number of arousals has to be presented. The analysis of respiratory parameters has to include the number of hypopnoeas and apnoeas. They have to be separated into obstructive, mixed and central events. The number of apnoeas and hypopnoeas should be counted separately for different body positions. Drops in oxygen saturation level have to be listed. The percentage of time spent with oxygen saturation below 90%, below 80% and so on has to be presented in a table and possibly in graphical form [50]. In addition, mean oxygen saturation per sleep stage may be of interest. The number of periodic leg movements has to be stated if applicable. It should be tabulated whether they were found together with arousal, awakening or respiratory events [60].

Modern sleep scoring systems should ultimately offer a handy filing system linked to report generation. This allows a systematic archival of polysomnographic recordings on numbered CD-ROMs and fast access to former reports. In the near future such systems with report archiving capabilities will be integrated in hospital information systems, which will come to be integrated digital patient records as have already been introduced in other hospital departments.



### Concluding remarks and perspective

Automatic sleep scoring is now at a stage which allows its use in clinical practice for the diagnosis of sleep disorders, even if it cannot yet replace visual R&K scoring. When using such systems one has to be aware that these systems have their limitations as explained in detail. The results always depend on the quality of the recorded signals. The analysis can never be better than the analysis of an experienced sleep scorer. Artefacts can produce wrong results. Therefore, the results of automatic sleep scoring need careful inspection by a trained polysomnographer.

Additionally, automatic sleep scoring systems offer the computational power to calculate additional variables which overcome the limitations of Rechtschaffen and Kales caused by discrete sleep stages and fixed epoch duration. Unfortunately, no standard variables have been defined until now. Therefore, one has to investigate many different variables provided by the system, such as adaptive segmentation, delta power measures, colour spectral arrays or continuous traces to get a better insight in sleep physiology and in the disturbed sleep of a particular patient. Ongoing projects try to develop and to establish rules for these new measures. Networking between sleep laboratories helps to bring the experiences of many sleep laboratories with different groups of patients together, to improve such new measures for sleep physiology and disturbed sleep.

### Practice points

1. A computer based sleep system should have a fast processor to minimize analysis time to less than 30 min. It should have a 20 inch high-resolution monitor and fast display capabilities which allow rapid scrolling during data review. It should store data on CD-ROM for archival purposes. Software to keep track of archived data should be included. It should allow visual and manual scoring in independent modes.
2. Do not trust automatic analysis, without being fully acquainted with the limitations.
3. Learn about the limitations of your preferred system by local validation in normal sleep and in disturbed sleep recorded in your own laboratory using your own recording equipment.
4. Repeat visual evaluation of patterns once in a while to remember the limitations and specific characteristics of your particular system.

### Research agenda

1. Validation studies of computer based sleep scoring are needed. The validation studies should be undertaken in normals of different age groups and in patients with various sleep disorders. Validations should be undertaken as multicentre studies because one specific analysis may be tuned according to local sleep classification practice.
2. Validation studies of computer based analysis of other polysomnographic signals is needed for different disorders.
3. New parameters have to be developed which overcome the limitations of Rechtschaffen and Kales in the diagnosis of sleep disorders. At the same time

the added value of these new parameters has to be proven in clinical studies with patient groups suffering from different sleep disorders.

4. Automatic recognition and classification of arousals is needed and has to be validated with the inclusion of autonomous nervous system parameters.

## Acknowledgement

This work was funded in part by the European Commission, DG XII, Biomed-2 project no. BMH4-CT97-2040 (SIESTA). Monique Aarts helped by correcting the English language in this manuscript.

## References

- 1 Rechtschaffen A, Kales A. *A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects*. Los Angeles, CA: BIS/BRI, Univ. of California, 1968.
- 2 Kubicki S, Herrmann WM, Höller L. Critical comments on the rules by Rechtschaffen and Kales concerning the visual evaluation of EEG records. In: Kubicki S, Herrmann WM (eds). *Methods of sleep research*. Stuttgart New York: Fischer, 1985: 19–35.
- 3 Stradling JR. Sleep studies for sleep-related breathing disorders. Consensus report. *J Sleep Res* 1992; **1**: 265–273.
- 4 Himanen SL, Hasan J. Limitations of Rechtschaffen and Kales. *Sleep Medicine Reviews* 2000; **4**: 149–167.
- 5 Loomis AL, Harvey EN, Hobart GAI. Cerebral stages during sleep, as studied by human brain potentials. *J Exp Psychol* 1937; **21**: 127–144.
- 6 Aserinsky E, Kleitman N. Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. *Science* 1957; **118**: 273–274.
- 7 Knott JR, Gibbs FA, Henry CE. Fourier transform of the electroencephalogram during sleep. *J Exp Psychol* 1942; **31**: 465–477.
- 8 Kemp B. A Proposal for Computer-Based Sleep/Wake Analysis. Consensus report. *J Sleep Res* 1993; **2**: 179–185.
- \*9 Hirshkowitz M, Moore CA. Issues in computerized polysomnography. *Sleep* 1994; **17**: 105–112.
- \*10 Bloch KE. Polysomnography: a systematic review. *Technology and Health Care* 1997; **5**: 285–305.
- \*11 Penzel T, Stephan K, Kubicki S, Herrmann WM. Integrated sleep analysis, with emphasis on automatic methods. *Epilepsy Res* 1991 (Suppl.) **2**: 177–204.
- 12 Makeig S, Bell AJ, Jung TP, Sejnowski TJ. Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems* 1996; **8**: 145–151.
- 13 Pardey J, Roberts S, Tarassenko L. A new approach to the analysis of the human sleep/wakefulness continuum. *J Sleep Res* 1996; **5**: 201–210.
- 14 Penzel T, Brandenburg U, Fischer J *et al.* Empfehlungen zur computergestützten Aufzeichnung und Auswertung von Polygraphien. *Somnologie* 1998; **2**: 42–48.
- 15 Smith JR. Automated EEG analysis with microcomputers. *Med Instr* 1980; **14**: 319–321.
- \*16 Hasan J. Past and future of computer-assisted sleep analysis and drowsiness assessment. *J Clin Neurophysiol* 1996; **13**: 295–313.
- 17 American Electroencephalographic Society. *Guidelines for recording clinical EEG on digital media*. American EEG society, Bloomfield, CT; 1991.
- 18 Bailey JJ, Berson AS, Garson A *et al.* Recommendations for Standardization and Specifications

---

\* The most important references are denoted by an asterisk.

- in Automated Electrocardiography: Bandwidth and Digital Signal Processing. *Circulation* 1990; **81**: 730–739.
- 19 Principe JC, Smith JR. SAMICOS—a sleep analyzing microcomputer system. *IEEE Trans Biomed Eng* 1986; **33**: 935–941.
  - 20 Penzel T, Hajak G, Hoffmann RM *et al*. Empfehlungen zur Durchführung und Auswertung polygraphischer Ableitungen im diagnostischen Schlaflabor. *Z EEG EMG* 1993 **24**: 65–70.
  - 21 Hoelscher TJ, McCall WV, Powell J, Marsh GR, Erwin CW. Two Methods of Scoring Sleep with the Oxford Medilog 9000: Comparison to Conventional Paper Scoring. *Sleep* 1989 **12**: 133–139.
  - 22 Kemp B, Värri A, Rosa AC *et al*. A Simple Format for Exchange of Digitized Polygraphic Recordings. *Electroenceph Clin Neurophysiol* 1992; **82**: 391–393.
  - 23 American Electroencephalographic Society. Guidelines for polygraphic assessment of sleep-related disorders (polysomnography). *J Clinical Neurophysiol* 1992; **9**: 88–96.
  - 24 Hasan J. Automatic analysis of sleep recordings: A critical review. *Ann Clin Res* 1985; **17**: 280–287.
  - 25 Pardey J, Roberts S, Tarassenko L. A review of parametric modelling techniques for EEG analysis. *Med Eng Phys* 1996; **18**: 2–11.
  - 26 Värri A, Hirvonen K, Häkkinen V, Hasan J, Loula P. Nonlinear eye movement detection method for drowsiness studies. *Int J Biomed Comput* 1996; **43**: 227–242.
  - 27 Penzel T, Petzold J. A new method for the classification of subvigil stages, using the Fourier transform, and its application to sleep apnea. *Comput Biol Med* 1989; **19**: 7–34.
  - 28 Martens WLJ. The fast time frequency transform (FTFT): a novel on-line approach to the instantaneous spectrum. Proceedings of 14th Annual International Conference IEEE Engineering in Medicine and Biology Society 1992; 2594–2595.
  - 29 Nielsen KD, Drewes AM, Svendsen L, Bjerregard K, Taagholt S. Ambulatory recording and power spectral analysis by autoregressive modelling of polygraphic sleep signals in patients suffering from chronic pain. *Methods Inf Med* 1994; **33**: 76–78.
  - 30 Jansen BH, Hasman A, Leuten R. Piece-wise EEG analysis: an objective evaluation. *Int J Biomedical Comput* 1981; **12**: 17–27.
  - 31 Geering BA, Achermann P, Eggimann F, Borbely AA. Period-amplitude analysis and power spectral analysis: a comparison based on all-night sleep EEG recordings. *J Sleep Res* 1993; **2**: 121–129.
  - 32 Praetorius HM, Bodenstein G. Adaptive Segmentation of EEG Records: A new Approach to automatic EEG analysis. *Electroenceph Clin Neurophysiol* 1977; **42**: 84–94.
  - 33 Hasan J, Hirvonen K, Värri A, Häkkinen V, Loula P. Validation of computer analysed polygraphic patterns during drowsiness and sleep onset. *Electroenceph Clin Neurophysiol* 1993; **87**: 117–127.
  - 34 Kubicki S, Höller L, Berg I, Pastelak-Price C, Dorow R. Sleep EEG Evaluation: A Comparison of Results Obtained by Visual Scoring and Automatic Analysis with the Oxford Sleep Stager. *Sleep* 1989; **12**: 140–149.
  - 35 Todorova A, Hofmann HC, Dimpfel W. A new frequency based automatic sleep analysis—description of the healthy sleep. *Eur J Med Res* 1997; **2**: 185–197.
  - 36 Biernacka H, Douglas NJ. Evaluation of a computerised polysomnography system. *Thorax* 1993; **48**: 280–283.
  - 37 White DP, Gibb TJ. Evaluation of a computerized polysomnographic system. *Sleep* 1998; **21**: 188–196.
  - 38 Schaltenbrand N, Lengelle R, Toussaint M, Luthringer R, Carelli G, Jacqmin A, Lainey E, Muzet A, Macher JP. Sleep Stage Scoring Using the Neural Network Model: Comparison Between Visual and Automatic Analysis in Normal Subjects and Patients. *Sleep* 1996; **19**: 26–35.
  - 39 Sforza E, Vandi S. Automatic Oxford-Medilog 9200 sleep staging scoring: comparison with visual analysis. *J Clin Neurophysiol* 1996; **13**: 227–233.
  - \*40 Robert C, Guilpin C, Limoge A. Review of neural network applications in sleep research. *J Neurosci Meth* 1998; **79**: 187–193.
  - 41 Schaltenbrand N, Lengelle R, Macher JP. Neural Network Model: Application to Automatic Analysis of Human Sleep. *Comp Biomed Res* 1993; **26**: 157–171.
  - 42 Roberts S, Tarassenko L. New method of automated sleep quantification. *Med Biol Eng* 1992; **30**: 509–517.

- 
- \*43 Kubicki S, Herrmann WM. The future of computer-assisted investigation of the polysomnogram: sleep microstructure. *J Clin Neurophysiol* 1996; **13**: 285–294.
  - 44 Hjorth B. EEG analysis based on time domain properties. *Electroenceph Clin Neurophysiol* 1970; **29**: 306–310.
  - 45 Hausteiner W, Pilcher J, Klink J, Schulz H. Automatic analysis overcomes limitations of sleep stage scoring. *Electroenceph Clin Neurophysiol* 1986; **64**: 364–374.
  - 46 Achermann P, Hartmann R, Gunzinger A, Guggenbühl W, Borbely AA. All-night sleep EEG and artificial stochastic control signals have similar correlation dimensions. *Electroenceph Clin Neurophysiol* 1994; **90**: 384–387.
  - 47 Rösche J, Fell J, Beckmann P. The calculation of the first positive Lyapunov exponent in sleep EEG data. *Electroenceph Clin Neurophysiol* 1993; **86**: 348–352.
  - 48 Pradhan N, Sadasivan PK. The nature of dominant lyapunov exponent and attractor dimension curves of EEG in sleep. *Comput Biol Med* 1996; **26**: 419–428.
  - \*49 Taha BH, Dempsey JA, Weber SM, Badr MS, Skatrud JB, Young TB, Jacques AJ, Seow KC. Automated detection and classification of sleep-disordered breathing from conventional polysomnography data. *Sleep* 1997; **20**: 991–1001.
  - 50 American Thoracic Society. Indications and standards for cardiopulmonary sleep studies. *Am Rev Respir Dis* 1989; **139**: 559–568.
  - 51 Redline S, Sanders M. Hypopnea, a floating metric: implications for prevalence, morbidity estimates and case finding. *Sleep* 1997; **20**: 1209–1217.
  - 52 Korten JB, Haddad GG. Respiratory waveform pattern recognition using digital techniques. *Comput Biol Med* 1989; **19**: 207–217.
  - 53 Hoffstein V, Mateika S, Nash S. Comparing perceptions and measurements of snoring. *Sleep* 1996; **19**: 783–789.
  - 54 Perez-Padilla JR, Slawinski E, DiFrancesco LM, Feige RR, Remmers JE, Whitelaw WA. Characteristics of the snoring noise in patients with and without occlusive sleep apnea. *Am Rev Respir Dis* 1993; **147**: 635–644.
  - 55 Penzel T, Amend G, Meinzer K, Peter JH, von Wichert P. MESAM: a heart rate and snoring recorder for detection of obstructive sleep apnea. *Sleep* 1990; **13**: 175–182.
  - 56 Gavriely N, Jensen O. Theory and measurements of snores. *J Appl Physiol* 1993; **74**: 2828–2837.
  - 57 Task Force of The European Society of Cardiology. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. *Eur Heart J* 1996; **17**: 354–381.
  - 58 Raetz SL, Richard CA, Garfinkel A, Harper RM. Dynamic Characteristics of Cardiac R-R Intervals during Sleep and Waking States. *Sleep* 1991; **14**: 526–533.
  - 59 Guilleminault C, Connolly S, Winkle R, Melvin R. Cyclical variation of the heart rate in sleep apnea syndrome: mechanisms and usefulness of 24 h electrocardiography as a screening technique. *Lancet* 1984; **I**: 126–131.
  - 60 The ASDA Atlas Task Force. Recording and scoring leg movements. *Sleep* 1993; **16**: 749–759.
  - 61 Värri A, Grote L, Penzel T, Cassel W, Peter JH, Hasan J. A new method to study blood pressure, heart rate and EEG as a function of reaction time. *Meth Inform Med* 1993; **33**: 314–316.
  - 62 Penzel T. Blood pressure analysis. *J Sleep Res* 1995; **4** (Suppl. 1): 15–20.
  - 63 Stradling JR, Davies RJO, Pitson DJ. New approaches to monitoring sleep-related breathing disorders. *Sleep* 1996; **19**: S77–S84.