31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Revised Recommendations For Computer-based Sleep Recording And Analysis

Thomas Penzel *Member IEEE*, Martin Glos, Christoph Schöbel, Michaela Sebert, Beate Diecker, and Ingo Fietze

Abstract— **Sleep recording is the quantitative method used in sleep centers in order to assess sleep disorders and to quantify pathological events occurring during sleep. Cardiorespiratory polysomnography is the method used for sleep recording. Standards for sleep recording stem back to 1968 and were compiled by Rechtschaffen and Kales with specifications for recording and analysis of the sleep EEG. An update considering digital technology and possibly computer based analysis techniques was needed. As the result of a two year process with a large committee in 2007 a revised manual was published. The manual includes for the first specifications for digital signals acquisition and some descriptions for developing computer based sleep analysis. The new manual was prepared using the formal tools of evidence based medicine. The method for preparing the new manual are presented critically the results relevant for computer based sleep recording and analysis are presented in detail and perspectives for computer based sleep analysis are pointed out.**

I. Introduction

DIGITAL sleep recording and sleep analysis is the diagnostic basis for all sleep medicine centres. Polysomnography is implemented on computers and uses modern concepts of digital signal acquisition and analysis. Computer based sleep recording has replaced almost everywhere analogue paper based sleep recording over the past 20 years. Initially, digital recording and analysis were restricted by technical limitations. With current technology, the technical limitations of computer acquisition, data storage, or analysis are less constraining. Thus recommendations for the digital recording and computer supported scoring of sleep can be guided by the underlying physiological and medical needs.

A new guideline which overcomes the limitation of previous guidelines was needed. Previous guidelines originate in 1968 and are based on a four channel sleep recording on paper. Technical constrains at that time resulted in four channels which were one sleep - electroencephalogram (EEG), two electrooculograms (EOG) and one submental electromygram (EMG). The actual sleep staging was explained based on recording examples stemming from healthy volunteers. This manual was developed in order to have a common understanding between sleep researchers.

Only later when the prevalence of sleep disorders became more apparent and when more and more sleep centers opened this manual became the basis to stage sleep and to diagnose sleep disorders. Based on the increased knowledge about sleep disorders additional papers with standards of practice were published giving advice to selected problems such as central nervous activation, called arousals, or respiratory disorders during sleep, such as apneas and hypopneas, or movement disorders such as the periodic limb movement disorder. A comprehensive and uniform update of these papers based on the methods of evidence based reviews is the main motivation of the American Academy of Sleep Medicine to work and compile this new manual.

II. Methods and Procedures

With the development of a new standard of practice for polysomnography by a committee run by the American Academy on Sleep Medicine the issue of digital acquisition, display, and analysis has been addressed for the first time [2]. In order to develop recommendations and specifications, a systematic literature search, an evidence review, and a standardized consensus process has been undertaken. Five questions regarding computer-assisted sleep recording and analysis were addressed. These questions are:

1) the reliability of computerized scoring of sleep stages,
2) the analysis of elemental events and waveforms,
3) the physiological and/or clinical significance of digitally-analyzed signals,
4) the importance of proposed changes in standardized scoring that could incorporate digital analysis,
5) the potential advantages and disadvantages of computerized sleep recordings.

In order to answer these questions systematic literature searches were conducted on medical databases (Medline) and all resulting papers were checked and evaluated using previously agreed criteria for medical evidence. Since normal evidence grading in medicine is formally tied to clinical trials with medical interventions, such as pharmacological treatment or surgical treatment, and not linked to diagnostic approaches, specifically involving an

Manuscript received April 21, 2009. This work was supported in part by the European Union under Grant EU FP7 FET OPEN 222079.

All authors are with the Sleep Medicine Center, Depart. of Cardiology, Charité University Hospital, Berlin, Germany. (corresponding author is Dr. Thomas Penzel, phone: +49-30-450513013; fax: +49-30-450513906; e-mail: thomas.penzel@ charite.de).

evaluation of biosignals, new evidence criteria had to be developed. The new criteria had to be modified in many aspects in order to be adequate for diagnostic studies and to be adequate for the specific type of recording and analysis done in sleep medicine. The new evidence criteria derived in this way are of major importance, because they outline how future validation studies have to be designed. Thus the criteria can be regarded as a cookbook for future studies in this field.

TABLE 1
THE NEW EVIDENCE LEVELS DEVELOPED FOR THE EVALUATION OF COMPUTER BASED SLEEP ANALYSIS AND VALIDATION STUDIES.

Evidence Levels	Study Design
I	Sample size > 64; sequential or representative; event by event or epoch by epoch comparison
II	Sample size > 32; sequential or representative; event by event or epoch by epoch comparison
III	Sample size > 16; sequential or representative; event by event or epoch by epoch comparison
IV	Sequential or representative sample; event by event or epoch by epoch comparison
V	Case series

The new evidence criteria as listed in table 1 relate mainly to the sample size of the studies used to validate sleep analysis systems. They also say that an epoch by epoch scoring of sleep stages must be performed in oder to consider the study as a serious addition to the field.

Since evidence criteria as listed in table 1 is one important issue and since content related criteria are the next important issue also a content evaluation was developed. The contents which have to be considered in high class sleep analysis software papers are then listed in table 2. The criteria are again specified in detail. The criteria should be folloewd carefully if a new study is designed which wants to develop and validate computer based sleep analysis software which will be accepted by the sleep centers and by sleep researchers in general.

TABLE 2
EVIDENCE GRADING FOR SLEEP ANALYSIS SOFTWARE. HERE CRITERIA TO BE FOLLOWED WHEN DESIGNING A NEW STUDY ARE SPECIFIED. THESE ARE THE FACTORS FOR EVALUATION THE GRADING IN A PARTICULAR EVIDENCE LEVEL.

Number	Factor Description
1	Normal controls and clinically relevant group used in sample
2	No recording selection for quality and/or discarded < 5% of records
3	Clinical standard used for group classification (e.g. ICSD, DSM IV)
4	Standard used for sleep data scoring (if relevant) and recordings provide data to properly apply standard (e.g. Central EEG, Left & right EOG, and EMG submentalis for R&K)
5	Blind, independent scoring
6	Multiple human scorers used to set comparison standard
7	Entire recordings used
8	Description adequate for replication

III. RESULTS

The evidence review suggested that automated sleep analysis is still in the stage of development. It has been well recognized that many methods like frequency analysis of the electroencephalography to identify delta, theta, alpha, beta waves, and feature extraction methods to identify sleep spindles, K-complexes, vertex waves, had been applied with high success. In order to use these parameters for automatic sleep staging combination method were used like artificial neural networks, fuzzy logic, or linear discriminant analysis. Finally all efforts were limited by the limited accuracy provided by the visual classification of sleep stages. It is the expectation that the new classification is able to simplify and homogenise the visual classification in such a way that the visual accuracy improves. Then automatic methods have a much better chance to come close to the visual classification. This only will allow a successful development of computer based sleep scoring.

For many technical specification decisions, little or no direct evidence was found in the literature. Although basic engineering principles or standard practices provided good rationale for certain specifications. This was utilized to develop the recommendations. Recommendations included sampling rate for signals, filter settings, user interface specifications, and report specifications. These new recommendations can now be used to guide all subsequent device design and should be regarded as established rules.

TABLE 3
RECOMMENDED SIGNALS TO BE RECORDED IN CARDIORESPIRATORY POLYSOMNOGRAPHY. HERE THE FUNCTIONS TO BE INVESTIGATED ARE LISTED TOGETHER WITH THE RELATED SIGNALS, THE SENSOR METHOD MOST APPROPRIATE, THE OPTIMAL SAMPLING RATE AND THE FILTER SETTINGS.

function	signal	method	optimal sampling rate	filter setting [Hz]
sleep	EEG, EOG	Electrodes	500 Hz	0.3 - 35
	EMG	Electrodes	500 Hz	10 - 100
respiration	airflow	nasal pressure, thermal probes	100 Hz	0.1 - 15
	respiratory effort	inductive plethysmography	100 Hz	0.1 - 15
	Oxygen saturation	SpO2	25 Hz	-
	carbon dioxide	tcPaCO2	25 Hz	-
	snoring	microphone	500 Hz	-
cardial	EKG	electrodes	500 Hz	0.3 - 70
movement	EMG tibialis	electrodes	500 Hz	10 - 100
	body position	position sensor	1 Hz	-
	Video	video camera	5 Hz	-

All these recommendations were subject to the subsequent UCLA/Rand standardized consensus process which finally resulted in the new specifications. The new specifications had been reviewed companies producing PSG equipment

and their questions and comments were considered to adapt the final recommendations to practical needs and manufacturers experience.

The work to develop the new manual was performed in dedicated task forces which were lead by a steering committee. The individual task forces had to work on dedicated scoring topics: digital requirements [3], visual scoring [5], respiratory event scoring [4], cardiac event scoring, movement event scoring, and arousal scoring [1]. Two task forces did serve questions related to all other task forces. These were pediatric issues in sleep scoring and age related issues in sleep scoring. Both these task groups did send one or two representatives in the other task groups. Each task group prepared and published a review paper listing the results of the systematic evidence based literature review. The systematic literature review was performed on previously agreed terms using the Medline abstract database. The task forces also answered the a priori posed questions. Each task group had to draft recommendations for the new manual. These consisted of mandatory guidelines, standards of practice, options and weaker recommendations. These recommendations were either clear results of the evidence provided by the literature or where the literature was not clear enough the formal consensus discussions and decisions did provide the final result.

The new guidelines include some detailed recommendations for a computer based sleep system user interface. These are:

1) External calibration must be possible for all signal channels.
2) Invert channels individually in terms of graphical presentation. EEG is typically presented with negative values upward and respiration with negative values downward.
3) Change offset of channels individually
4) Show 5 seconds only and up to entire night recording in order to zoom in for graphical events in the sleep EEG and in order to get a quick overview on underlying trends in signals such as oxygen saturation.
5) Synchronized video with signals
6) Save profiles for recording and scoring.
7) Indicate all recognized patterns such as sleep spindles, K-complexes, delta waves as a first step towards automatic sleep analysis.

The profiles mentioned here should reflect the actual appearance of the biosignals on the computer screen during the time of recording and also during the time of analysis. Gain, offset, filter settings and time scale settings should be included in these parameters. Since it is possible to hide some channels during recording and to adjust gain and offset the technician might have seen or overseen some physiological or pathological events which occurred during the night attendance. The same may be true for the person doing the visual analysis of the recorded data. In order to track this, the committee felt, that the screen and channel settings should be saved in a profile for later access.

IV. Conclusions

The new rules based on literature review, evidence, and consensus discussions include the following main issues:

- Rules extended by aspects of cardiorespiratory polysomnography
- Recommendation for sampling rates and filter settings
- No automatic sleep analysis, only a marking of detected events and patterns
- Recommendation for PSG reporting
- Recommendation for user interfaces of computer assisted systems

The new evidence based manual for sleep scoring is an important step towards a computer based sleep recording and analysis. It is advantageous to have requirements for signal acquisition listed in detail. This can be used as a cookbook for the development of new computer based sleep analysis. The requirements should be carefully followed when preparing validation studies for new sleep stage analysis software.

Acknowledgment

This work had been supported by the European Union grant HIVE.

References

[1] M. H. Bonnet, K. Doghramji, T. Roehrs, E. J. Stepanski, S. Sheldon, A. S. Walters, M. S. Wise, and A. L. Chesson, "The scoring of arousal in sleep: reliability, validity, and alternatives," *J. Clin. Sleep Med.*: 3: 133-145, 2007.

[2] S. M. Caples, C. L. Rosen, W. K. Shen, A. S. Gami, W. Cotts, M. Adams, P. Dorostkar, K. Shivkumar, V. K. Somers, T. L. Morgenthaler, E. J. Stepanski, and C. Iber, "The scoring of cardiac events during sleep", *J. Clin. Sleep Med.* 3: 147-154, 2007.

[3] C. Iber, S. Ancoli-Israel, A. Chesson, and S. F. Quan for the American Academy of Sleep Medicine, *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, 1st ed.: Westchester, Illinois: American Academy of Sleep Medicine, 2007.

[4] T. Penzel, M. Hirshkowitz, J. Harsh, R. D. Chervin, N. Butkov, M. Kryger, B. Malow, M. V. Vitiello, M. H. Silber, C. A. Kushida, and A. L. Chesson, "Digital analysis and Technical Specifications," *J. Clin. Sleep Med.* 3: 109-120, 2007.

[5] A. Rechtschaffen, and A. Kales, *A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects.* US Department of Health, Education, and Welfare Public Health Service – National Institute of Health, 1968.

[6] S. Redline, R. Budhiraja, V. Kapur, C. L. Marcus, J. H. Mateika, R. Mehra, S. Parthasarthy, V. K. Somers, K. P. Strohl, L. G. Sulit, D. Gozal, M. S. Wise, and S. F. Quan, "The scoring of respiratory events in sleep: reliability and validity", *J. Clin. Sleep Med.* 3: 169-200, 2007.

[7] M. H. Silber, S. Ancoli-Israel, M. H. Bonnet, S. Chokroverty, M. M. Grigg-Damberger, M. Hirshkowitz, S. Kapen, S. A. Keenan, M. H. Kryger, T. Penzel, M. R. Pressman, C. Iber, "The Visual Scoring of Sleep in Adults", *J. Clin. Sleep Med.* 3: 121-131, 2007.

[8] A. S. Walters, G. Lavigne, W. Hening, D. L. Picchietti, R. P. Allen, S. Chokroverty, C. A. Kushida, D. L. Bliwise, M. W. Mahowald, C. H. Schenck, and S. Ancoli-Israel, "Movements in sleep: A review by the movement task force of the american academy of sleep medicine", *J. Clin. Sleep Med.* 3: 155-167, 2007.