JCSM
Journal of Clinical Sleep Medicine

Digital Analysis and Technical Specifications

Thomas Penzel, Ph.D.[1]; Max Hirshkowitz, Ph.D.[2]; John Harsh, Ph.D.[3]; Ron D. Chervin, M.D.[4]; Nic Butkov[5]; Meir Kryger, M.D.[6]; Beth Malow, M.D.[7]; Michael V. Vitiello, M.D.[8]; Michael H. Silber, M.B.ChB.[9]; Clete A. Kushida, M.D., Ph.D.[10]; Andrew L. Chesson, Jr., M.D.[11]

[1]Charité Universitätsmedizin Berlin, Berlin, Germany; [2]Baylor College of Medicine, Houston, TX; [3]University of Southern Mississippi, Hattiesburg, MS; [4]University of Michigan, Michigan School of Medicine, Ann Arbor, MI; [5]School of Clinical Polysomnography, Medford, OR; [6]University of Manitoba, Winnipeg, Canada; [7]Vanderbilt University, Nashville, TN; [8]University of Washington, Seattle, WA; [9]Mayo Clinic College of Medicine, Rochester, MN; [10]Stanford University, Stanford, CA; [11]LSU Health Science Center – Shreveport, Department of Neurology and Office of the Dean, Shreveport, LA

Abstract: Digital acquisition and analysis of sleep data has become more common over the past 20 years. Many investigators have developed strategies to record and analyze sleep in a quantitative way. Initially, digital recording and analysis were restricted by technical limitations. With current technology, the technical limitations of computer acquisition, data storage, and analysis are less constraining, and the development of recommendations for the specifications and scoring of sleep can be more clearly guided by the goal of characterizing physiologic phenomena. In order to develop recommendations and specifications regarding digital acquisition and analysis, a literature search, evidence review, and standardized consensus process focused on 5 questions regarding computer-assisted sleep recording and analysis. These questions included: 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, and 5) the potential advantages and disadvantages of computerized sleep recordings. Of 154 studies identified by the search, 119 were found to be suitable for evidence review. The evidence review suggested that computer scoring and quantitative analysis of sleep is still in the formative stage of development. For many technical specification decisions, little or no direct evidence was found, although basic engineering principles or standard practices provided some rationale which was utilized to develop the recommendations formulated during the subsequent UCLA/Rand standardized consensus process.

Keywords: Digital polysomnography, computerized scoring, sleep stage scoring, computerized sleep recording

Citation: Penzel T; Hirshkowitz M; Harsh J et al. Digital analysis and technical specifications. *J Clin Sleep Med* 2007:3(2);109-120

Disclosure Statement

This is was not an industry supported study. Dr. Penzel is on the board of directors of Advanced Sleep Research Berlin and has participated in speaking engagements for and has received research support from Respironics and Weinmann. Dr. Hirshkowitz is a member of the speakers bureaus for Sanofi-Aventis, Takeda, and Cephalon; has lectured for Exxon Mobile, Atlanta School of Sleep Medicine, Palo Alto School of Sleep Medicine, and the American Academy of Sleep Medicine; and is the principal investigator on research managed by Baylor College of Medicine including contract research from Evotec, Neurogen, Sanofi-Aventis, GlaxoSmithKline, Merck, Takeda, NBI, Respironics, and Organon. Dr. Harsh is on the advisory board for Cephalon. Dr. Chervin is on the scientific advisory board of Pavad Medical and is a consultant for Alexa Pharmaceuticals. Mr. Butkov has financial interests in the School Of Clinical Polysomnography, Inc. and Synapse Media. Dr. Vitiello is on the speakers bureau for Takeda. Dr. Kushida is an investigator on research managed by Stanford University including contract research from GlaxoSmithKline, Boehringer-Ingelheim, XenoPort, Inc., Schwarz Pharmaceuticals, and Kyowa Pharmaceuticals and has received travel grants from XenoPort, Inc. for oral presentations. Drs. Kryger, Malow, Silber, and Chesson have reported no financial conflicts of interest.

Submitted for publication February 1, 2007
Accepted for publication March 15, 2007

Address correspondence to: Thomas Penzel, Ph.D., Center for Sleep Medicine, CCM13, Charité Universitätsmedizin Berlin, Luisenstrasse 13, D-10117 Berlin, Germany, Fax: 49 30 450513906; E-mail: thomas.penzel@charite.de

1.0 HISTORICAL PERSPECTIVE

In 1935, Loomis and colleagues recorded the first continuous all-night polysomnogram.[1] They developed a data reduction scheme and classified sleep into stages. Sleep stage scoring evolved over the next 35 years. Nathaniel Kleitman, the founder of American sleep research, developed a modified sleep stage system that was later revised after the discovery of REM sleep by Eugene Aserinsky in 1953.[2] This Dement-Kleitman system was rivaled by one used by a second American group (Williams-Karacan system) and by a European system which categorized sleep as synchronized and desynchronized. In 1968 Rechtschaffen and Kales (R & K) developed a system that has been used since that time as the standardized system for sleep staging.[3] In 2004, the American Academy of Sleep Medicine (AASM) initiated a development process[4] to revise R & K sleep staging and to address digital methodology as well as the scoring of arousals, respiratory events, sleep related movement disorders, and cardiac abnormalities, with consideration of pediatric and geriatric age groups. This process evolved under the supervision of a steering committee and 8 task forces, resulting in evidence review papers published in the *Journal of Clinical Sleep Medicine* and recommended scoring rules and specifications published in a new sleep scoring manual.[5] Recommendations published in the manual reflect both evidence review and, when evidence was limited, a standardized consensus process.

2.0 METHODS

2.1 Timeline

Between January 2005 and April 2006, the Digital Task Force (see page 117) met by conference call on 12 occasions and face-to-face once to discuss the process of evidence review and develop consensus on the issues of terminology, technical specifications, and possible scoring rules for digital analysis. The first part of this review paper covers the literature review. The second part describes the rationale for consensus for recommendations for which peer-reviewed evidence-based literature was not available. The third part of this paper outlines remaining issues and options for addressing these issues in the future.

The completed paper was sent for outside review and was then approved by the steering committee and by the AASM Board of Directors.

2.2 Search Terms

The first step for reviewing the relevant literature was to formulate questions that were potentially amenable to the evidence-based medicine analytic approach. To this end, the following questions were proposed:

1. Can computerized polysomnographic analysis score sleep stages reliably and accurately (using manual scoring as a reference standard)?
2. Can computerized polysomnographic analysis detect and analyze standardized elemental sleep waveforms (slow waves, delta, theta, alpha, sigma, and beta) and events (spikes, sharps, K complexes, vertex sharp waves) consistently and accurately?
3. Do any of the automatically detected or analyzed waveforms or events have underlying physiological correlates, theoretical significance, or provide sensitive markers or outcome variables for clinical, pharmacological, or age related changes? Which, if any, of these markers might be clinically useful?
4. Are there proposed changes to R & K involving computerized analysis that might be important for characterization of sleep?
5. What are the advantages and disadvantages of computerized polysomnography? How can advantages be promoted? What remedies, if any, can be applied to the disadvantages?

Search terms to address these questions were developed and provided to American Academy of Sleep Medicine staff who conducted a PubMed literature search of references published between 1968 and 2004. The search terms *Scoring* and *Digital* were combined for the searches as shown in Table 1. Limiters included *English, Clinical Trial*, and *Humans*.

In addition to the citations found using these search terms, other relevant papers were gathered through *pearling*, secondary searches, and expert consultation. After the initial search, the identified article titles, authors, citation, and abstracts were prereviewed to determine relevance and culled to those related to the topic under study. There were 154 relevant articles identified as needing formal review and scoring. Of those, 119 were found suitable for evidence tables to address questions prepared by the task force. Articles were restricted to those published in English dealing with human sleep scoring in adults.

2.3 Evidence Grading and Rationale

The classification of evidence used for the evidentiary articles is listed in Table 2. These utilize a special grading system for performance evaluation studies**.** The Digital Task Force reviewed the relevant literature, scored the evidence, and assembled evidence s. In many areas of technical specifications, published articles directly addressing these issues were not available. When published articles were not available, the task force developed rationales for scoring decisions based on logical analysis, common practices, and practical implementation. The evidence analyses of published articles and the rationales for technical specifications became the basis for this paper and for the construction of consensus ballots for scoring rule development.

Table 1—Search terms used

Scoring search terms	(combined with)	Digital search terms
REM sleep		Digital
Stage 4 sleep		Score
Stage 3 sleep		Sleep stage scoring
Stage 2 sleep		Digital scoring
Stage 1 sleep		Automated
Alpha AND sleep		Automatic
Delta AND sleep		Computerized
Drowsiness AND normal		Computer score
Eye blinks		Computerized score
Eye movements AND sleep		Computerized scoring
K complexes		Computer scoring
Spindles		Computer sleep stage scoring
Sleep staging		Computerized polysomnography
Sleep scoring		
Apnea		
Hypopnea		
Leg movements		

2.4 Consensus Decisions

In some cases the task force recommendations included in this paper are based on evidence published in peer-reviewed journals. However, where scientific data are absent, insufficient, or inconclusive, recommendations are based upon consensus agreement using the RAND/UCLA Appropriateness Method.[6] This consensus method incorporates an initial round 1 ballot without interaction and a subsequent round 2 ballot after task force discussion. Further ballots were utilized when needed, following this process. The first round of Rand balloting took place January 30, 2006, and balloting was completed March 6, 2006. Numerous content areas required serial reballoting when agreement could not initially be reached or when clarification was needed. The results of the consensus balloting and evidence review are incorporated into the scoring manual.

3.0 BACKGROUND

Visual sleep stage scoring is a time consuming and labor-intensive process. Additionally, an element of *scorer judgment* is involved, resulting in concerns about interrater reliability. This concern increases as demand for rapid data reduction rises, resulting in time pressure on a scorer. With the advent of the laboratory computer capable of signal processing, researchers and engineers began exploring ways to automate sleep stage scoring. Various systems to mimic human scoring have been described, tested, revised, and retested. Rigor and sophistication of designs and selection of benchmarks have varied widely. Analytic procedures, initially developed to solve engineering problems, were applied

Table 2—Evidence grading for performance evaluation studies

Level	Performance Criteria
1	1. Sample size >64 (number of recordings) without nesting within subject (unless it is systematic, e.g. 2 from each) 2. Sequential or representative sample 3. Event x Event or Epoch x Epoch Comparison
2	1. Sample size >32 (number of recordings) without nesting within subject (unless it is systematic, e.g. 2 from each) 2. Sequential or representative sample 3. Event x Event or Epoch x Epoch Comparison
3	1. Sample size >16 (number of recordings) without nesting within subject (unless it is systematic, e.g. 2 from each) 2. Sequential or representative sample 3. Event x Event or Epoch x Epoch Comparison
4	1. Sequential or representative sample 2. Event x Event or Epoch x Epoch Comparison
5	Case series

Notes: Evaluation Factors for Grading within 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

to electroencephalographic, electrooculographic, and electromyographic signals. Fourier transforms, period-amplitude analysis, compressed spectral array calculation, and complex demodulations were just a few of items in the computer's repertoire. These techniques were used to explore differences between groups or outcomes. Thus, the literature contains descriptions of EEG waveform changes associated with drug administration, different psychiatric conditions, and aging.

Today, computerized recording and storage systems have all but replaced the paper-based analog polysomnograph recordings and paper storage. One consequence of this is that computerized polysomnography has developed in the absence of professionally endorsed and uniform standards. Market forces dictated who survived and who did not. Concern is now emerging that undesirable compromises were made related to signal acquisition, conditioning, and quality.

3.1 Specifications, Terminology and Processes

Computerized polysomnography involves recording, analyzing, displaying, scoring, tabulating, distilling, and storing sleep studies. From a data processing standpoint 5 basic and distinct processes can be defined:

1. data acquisition (recording)
2. data display (viewing)
3. data manipulation (scoring and editing)
4. data reduction (parameterization for reporting)
5. data filing (storage)

Currently, no uniform standards exist for any of these processes. The information sources serving as a basis for developing standards have varied depending on the process. For some processes, an evidence-based medicine approach will be possible (e.g. scoring validation against a standard). For other processes, the basis for standards will be largely engineering principles (e.g., data acquisition and filing). Finally, there are processes (e.g., parameterization for reporting) that are based on expert preferences and needs.

For some technical and digital specifications, or identifications of parameters to be reported, recommendations have been provided in the literature. These have generally been procedural and reporting guidelines based on common practices or recommendations from groups of experts or associations, and often have not been based on high levels of directly applicable evidence. These groups have included clinical neurophysiologists.[7-9]

Some parameters recommended for inclusion in polysomography reports have been published by the Standards of Practice Committee of the American Academy of Sleep Medicine.[10] Some practical technical specifications have also been recommended for use for recorded analysis of ECG data. Recommendations for standardization and specifications in automated electrocardiology (bandwidth and digital signal processing) have been published.[11] Minimum technical requirements for performing clinical EEG and guidelines for recording clinical EEG on digital media are available.[8]

4.0 EVIDENCE FOR RELIABILITY OF RULES, SPECIFICATIONS, OR MEASURES

4.1. Question: Can computerized polysomnographic analysis score sleep stages reliably and accurately?

The search terms used resulted in 45 relevant citations. Table 3 (which can be accessed on the web at www.aasmnet.org), lists these citations, identifying study type, number of subjects, results, and evidence level. Twenty-five related to computer-human reliability for scoring sleep stages or some other general feature of sleep. The remaining 20 were reviews, demonstrations of an analytic process, or instructional papers. Two studies were graded at Level 1; two were graded at Level 2; nine at Level 3; five at Level 4; and 27 at Level 5.

The first of the two Level 1 studies listed in Table 3 is a recent large-scale, multicenter, well-designed validation study performed in Europe as part of the "Siesta Project."[12] Five hundred ninety recordings, split between development and validation samples, were used. This group of researchers found 80% agreement between their computer system scoring and human scoring. Test-retest reliability with the computer system was very close to complete agreement.

The second Level 1 study analyzed recordings made on 200 men and women in a study of sleep and aging.[13] Epoch-by-epoch validation yielded a 74% agreement for the system with a kappa coefficient of 0.57, which is statistically good agreement. Using polysomnographic recordings from 60 subjects, one of the Level

2 studies found 87.5% agreement between two expert scorers and 82%-90% agreement with an automated system.[14] By contrast, the other Level 2 study reported substantial errors in sleep staging when full automatic mode was employed.[15]

A Level 3 comparison between human and computer scoring that used 30 patient polysomnograms[16] concluded that computerized scoring was acceptable. Comparable concordance was found in agreement ratios for computer vs. human scoring and human vs. human scoring. In an older study, Stanus and coworkers[17] developed and applied a stochastic staging method. They scored 15 patient records and 15 control subject records and compared the results to human scoring, finding 70%-75% concordance. A study comparing the Morpheus system to two human scorers using 31 sleep recordings found 78% and 73% agreement between human and machine, and 82% agreement between human scorers.[18] Another Level 3 study compared Fast Fourier Transform based scoring of recordings obtained from 18 subjects (9 male and 9 female). They found high reproducibility in characterizing the results of polysomnography; however, the authors did not report on computer vs. human sleep stage differences.[19] Mykytyn and colleagues[20] examined 20 patient records and concluded that portable sleep recordings were scorable and interpretable but would benefit from more durable attachment of monitoring devices. Andreas and colleagues applied computerized scoring to 27 consecutive patients with sleep apnea.[21] They found considerable agreement but warned that "Automated analysis should only be used by those who are able to perform a visual analysis." Another study used a hybrid (analog-digital) system and examined samples from younger and older normal subjects, as well as individuals with alcoholism. They found agreement levels between computer and human scoring of 80% for young normal subjects, 77% for older normal subjects, and 75% for alcoholics.[22]

4.1.1 Summary

While the studies above considered various aspects of computerized classification of sleep stages, discussion related to accuracy and precision should consider these data along with associated factors.

1. Overall, the literature concerning sleep stage validation appears to provide evidence that human and computer agreement with some systems, circumstances of training, and human editing intervention has reached the level of results of human scoring agreement between different laboratories. However, the level of agreement can not be generalized from one system to another, and automated use of such systems show conflicting Level 2 data on reliability. Most of these studies were on selected patients in sophisticated laboratories and utilized highly trained research technicians and sleep experts. Furthermore, changes being made in the human scoring system for sleep stages as part of the current overall R & K revision process will necessitate reprogramming, debugging, and revalidation of existing codes.
2. Classification accuracy of any given system must be evaluated in both normal samples and appropriate pathophysiological samples of recordings. Age-related alterations in waveforms in healthy and sleep disordered individuals must be differentiated. Studies must be adequately powered, since failure to find statistically significant differences in an underpowered design does not mean equivalence between methods in stage classification or event detection.
3. Technologist training, electrode application, and overall polysomnographic recording quality are critical, perhaps even more for computerized systems than for visually scored recordings. The powerful manipulation possible using computerized polysomnography, the signal conditioning, and nearly unlimited filtering power do not obviate the need for good recordings. Some computerized systems, for example, can manipulate the signal to the extent that complete open channel non-recordings can pass for bioelectrical signals. The age-worn acronym GIGO (Garbage In-Garbage Out) is especially appropriate for computerized polysomnography. Procedures to assure input signal quality are crucial and cannot be overemphasized.

4.2 Question: Can computerized polysomnographic analysis detect and analyze the elemental sleep waveforms (slow waves, delta, theta, alpha, sigma, and beta) and events (spikes, sharps, K complexes, vertex sharp waves) consistently and accurately?

The task force reviewed 26 published articles describing computerized systems for detecting EEG waveforms and artifacts routinely seen in polysomnographic recordings (Table 4—which can be accessed on the web at www.aasmnet.org). Of these, 22 presented data, 1 was a review,[23] and 3 were descriptions of system developments.[24-26] After grading, one report reached Level 1, one was rated Level 2, seven qualified as Level 3, eight met criteria for Level 4, and the remaining nine were Level 5.

The Level 1 study by Declerck and colleagues[27] compared 3 spindle detectors. Little variation was found using a sample of 500 recordings, indicating that there are several reliable techniques for automatically detecting sleep spindle activity. In another study, weighing in at Level 2,[28] wakefulness was detected during polysomnography using an alpha-slow wave index. Data from 16 young adult normal subjects and 16 elderly subjects with insomnia were compared for computer-human agreement. Agreement rate was 94% in the elderly and almost 97% in healthy young adults.

A host of Level 3 studies explored a range of analytic techniques to identify differing EEG features. Fuzzy logic performed well for identifying alpha activity.[29] Power spectral analysis has revealed changes over the time course of sleep.[30,31] Schlogl and coworkers[33] observed EEG nonlinearities in terms of amplitude distributions and suggested using signal range saturation values in order to detect artifacts in recorded signals. In patients with recent strokes, traditional sleep stage measures did not differ significantly from the control group.[34] However, computerized indices of slow wave activity across sleep and wakefulness differentiated groups.

Other data presented in Level 3 and 4 studies provided evidence supporting computer detection of sleep spindles,[35] EEG characteristics of hypoglycemia,[36] paroxysmal EEG activity characterizing nocturnal seizures,[37] and tonic REM sleep.[38] Additionally, Mann and colleagues[39] found EEG root mean square values help characterize the dynamics of the sleep process. In a completely different approach, EEG resonance to event-related potentials was found to differ between stage 1, the other NREM stages, and REM sleep.[40] The oscillations produced consisted of alpha, delta, and theta activity, respectively. Finally, in an investigation of the effects of eroding display resolution, decreases exceeding 1/9th produced significantly reduced scoring accuracy.[41] Kupfer and coworkers[42] supplemented traditional sleep stage scoring with automated REM and delta sleep analyses to provide more precise description of REM activity and delta wave patterns in young and old patients with depression.

Using measures of the cyclic alternating pattern to index sleep fragmentation, greater sensitivity was found compared with traditional sleep stage parameters[43] in a group of patients with parasomnias compared to controls.

4.2.1 Summary

The literature review indicates that waveform detection can be and has been performed; however, documenting the validity of such systems has taken a back seat to their research applications. Published prospective validation of existing detectors for particular patterns is needed. The role of artifact in the recording remains a critical and inadequately addressed problem. Recording artifact is regarded as creating the weakest link between bioelectrical activity and consistently interpretable physiologic processes. Systems detecting waveforms and frequency signatures need more field trials, including multicenter applications.

5.0 EVIDENCE FOR VALIDITY OF RULES, SPECIFICATIONS, OR MEASURES

5.1 Question: Do any of the automatically detected or analyzed waveforms or events have underlying physiological correlates, theoretical significance, or provide sensitive markers or outcome variables for age related, pharmacological, or clinical differences? Which, if any, of these markers might be clinically useful?

Advocates for computerized polysomnographic analysis have long touted the potential for greater insight into the sleep process achievable by directly measuring waveforms. The argument typically depicts sleep stage classification as a generalization within a time domain (one epoch) that discards a major proportion of available data in order to achieve data reduction. With computer analysis, this information can be retained and can potentially provide a more direct understanding of underlying physiological processes. For example, given the evidence that sleep spindles originate by reciprocal interaction between reticular thalamic and thalamocortical neurons,[44-48] monitoring spindle frequency changes associated with GABA-A receptor agonist drugs would provide optimal sensitivity. By contrast, if beta activity arises from cortex (excitatory and inhibitory reciprocal interactions), drug-related or depression-related alterations in beta activity could reflect frontal and central brain activation and desynchronization in NREM and REM sleep.[27,49]

Seven papers were found concerning computerized polysomnography and aging; three were Level 1, three were Level 2, and one was Level 3 (Table 5—which can be accessed on the web at www.aasmnet.org). Data from 200 men and women found computerized indices to be sensitive markers for age-related changes, especially in the delta bandwidth.[13] In another report, computer analyzed data from a large sample revealed increases in spindle frequency and decreases in spindle amplitude accompanying advancing age.[50] By contrast, no age-related changes were found in spindle duration. In an early paper, Smith and coworkers[51] used the sleep analyzing hybrid computer (SAHC) to establish age-related trends in basic sleep EEG waveforms. Alpha frequency was stable over the lifespan (except in the very youngest group), REM-related beta and theta activity differed between adults and preteens, and spindle frequency was lower in younger groups. Level 2 findings included a decrease in delta activity associated with aging.[52] Reynolds and colleagues[53] had previously reported utility of slow wave activity in differentiating elderly patients with and without dementia. Eye movement density was found to be substantially reduced in the elderly according to a study graded at Level 2.[54] Computer assisted quantification of EEG and REM activity was also reported using power spectral analysis in an article graded at Level 3.

Table 6 (which can be accessed on the web at www.aasmnet.org), lists eight identified studies that used computer indices to investigate drug effects. Evidence that computerized polysomnography is sensitive to drug administration has been consistently demonstrated in studies with grading of Level 1,[55] Level 2,[56,57] and Level 3.[58,59]

Our search for computerized polysomnographic applications to psychiatric and neurological conditions yielded 15 citations, listed in Table 7 (which can be accessed on the web at www.aasmnet.org). Twelve of the 15 articles focused on sleep microarchitectural and waveform alterations associated with major depressive disorder (MDD). Of the total set of 15 reports, one was graded at Level 1, six were Level 2, six met criteria for Level 3, one was Level 4, and one was Level 5. MDD is the focus of several studies examining eye movements (two Level 3 reports[60,61]), slow wave activity (four Level 2 articles[53,61-64]), and REM and slow wave activity (Level 1[65] and Level 3[66-68]). Schizophrenia was investigated with the help of computerized sleep analysis in two Level 3 studies[58,68] and one Level 4 study.[67]

5.1.1 Summary

The literature in this area is in a formative stage. Much research being conducted with computerized analysis is searching for sensitive indices for exploring underlying mechanisms associated with age, drugs, and comorbidity. This approach has added to our knowledge and understanding of sleep process. However, applying such techniques has not been standardized as a useful part of routine clinical practice.

6.0 RATIONALES FOR TECHNICAL SPECIFICATIONS

Digital polysomnography systems have largely replaced analog equipment, yet no formal standards have been set regarding their design and function. As a result, present-day recording systems may lack basic features necessary for accurate and reproducible polysomnographic data. Some recording functions have been added or modified without formal consensus or input from sleep field professionals. A particular concern is excessive digital filtering. Other issues relate to the omission of customary polysomnograph features, such as visual calibration displays, electrode selector devices, and impedance measurements. The implementation of standards in digital polysomnography is critical to ensure accuracy and reproducibility of sleep recordings in both clinical and research arenas.

In addition to information derived from published literature described in section 5, some questions required a panel of experts to reach consensus because published data are lacking or insufficient. Furthermore, some critical specifications for computerized polysomnography must necessarily derive from engineering principles (e.g. sampling rates for bioelectrical signals), while others are a matter of preference (e.g., the summary parameters displayed on the computer reports). A series of questions can be posed concerning data acquisition and display.

6.1 Questions for rationale development

1. Should standard digital filter design for PSG data collection replicate the signal characteristics of conventional analog-style filters, preserving the nuances of physiological data within the selected frequency range, without excessive signal filtering, and allowing artifacts to be seen when the input signal is of poor quality? Should standard specifications be developed for digital amplifiers and filters to ensure compatibility and reproducibility of data among digital recording systems?
2. Should digital PSG recording systems be required to provide a visual (on-screen) standard –50 microvolt DC calibration signal for all channels to demonstrate polarity, amplitude, and time constant settings for each recorded parameter?
3. Should digital PSG recording systems provide a conventional electrode selector panel (or reasonable facsimile thereof) for choosing and/or changing input signal derivations without relying on a common reference electrode?
4. Should digital PSG systems provide separate 50/60 Hz filter controls for each channel?
5. Should digital PSG systems provide the capability of selecting appropriate sampling rates for each channel?
6. Should digital PSG systems provide a method of measuring actual (individual) electrode impedance against a reference (the latter may be the sum of all other applied electrodes)?
7. Should digital PSG systems provide the capability of retaining and viewing the data as it was recorded by the attending technologist (i.e., retain and display all derivation changes, sensitivity adjustments, filter settings, temporal resolution)? Should the same capability be available for the data as they appeared when the study was scored by the scoring technologist?

Review of the literature clarified there was no direct evidence to address these questions based on studies related to sleep or on comparisons to established "gold standards." This paper's analyses and suggestions are based on application of engineering principles, other groups' guidelines, (supported at times by practice experience), and recommendations from the RAND/UCLA Appropriateness Method ballot outcomes for consensus.

6.2 Digital filters

Present-day digital filters tend to overprocess the input signal. As a result, the recorded data may appear "clean," but lack essential detail.

The sharp cut-off design of some digital filters essentially removes all fast and slow frequency artifacts, regardless of input signal quality. Consequently, the attending technologist may not be aware of input signal degradation (i.e., electrodes may be detached from the patient and continue to generate a narrow bandwidth of nonphysiological frequencies resembling the EEG).

Digital filters can produce practically any kind of frequency response. Digital filters can simulate conventional analogue circuit filter frequency response curves. Recommendations for the filter design can help ensure the reproducibility of data generated by different brands of equipment and preserve continuity with the previous gold standard, as applied to paper-based recordings.

6.3 Calibration circuit

Some present-day systems have eliminated the calibration circuit, resulting in an inability to demonstrate channel configuration settings. Substitute internal calibrations offered by these systems do not provide a visual depiction of the calibration wave. Furthermore, some systems are sold with default inverted polarity settings. Without the ability to verify signal polarity (i.e., negative G1 voltage = upward deflection), inverted EEG recordings may be generated without the attending technologist's awareness. It should also be noted that a visual demonstration of the calibration signals before and after each study is a sleep center accreditation requirement.

A conventional DC calibration wave applied at the electrode input and not internally would allow the technologist and the reading physician to verify and document the following:

- Correct signal amplitude settings by examining the height of the wave relative to a calibration grid on the computer screen.
- Appropriate filter settings for each channel by examining the corresponding time constant of each calibration wave.
- Correct signal polarity according to the international polarity convention for EEG and PSG recordings.

6.4 Electrode selector panel

Most present-day digital equipment offers system referencing as a substitute for a conventional electrode selector panel, whereby all input signals are referenced to a single electrode, such as Cz. System referencing is useful for digital EEG recordings, providing post-recording "remontaging" capabilities. However, in standard polysomnography, this method is not ideal. During PSG data collection, the primary reason for changing input derivations is to eliminate recording artifact. If the system reference becomes the source of artifact, all channels are affected.

In contrast, a conventional electrode selector panel allows the attending technologist to change input derivations without relying on a single electrode. Furthermore, the use of post-recording remontaging tends to shift the burden of artifact recognition and correction to those who read the study, rather than addressing recording problems as they occur.

It should also be noted that the use of system referencing in polysomnography has never been formally examined or approved. Remontaging is an offshoot of EEG technology that has become a very important option in EEG, particularly in sleep centers conducting expanded EEG montage recordings. However, it should not be a substitute for a conventional electrode selector panel equivalent, a feature that was previously an essential component of every PSG recording system. Remontaging is not a regular option in polysomnography systems derived from respiratory recordings. In any case, it is essential to document what the technician performing the recording has seen and has done regarding the montage.

6.5 Individual 50/60 Hz filters

Some present-day digital systems do not provide separate 50/60 Hz filtering capabilities for each individual channel. This poses a problem, because 50/60 Hz filters are generally used only on an as-needed basis. It is not standard practice to

use 50/60 Hz filters for EEG recordings. The inappropriate use of 50/60 Hz filters may easily mask a poor quality input signal (similar to the principle of using overefficient high and low frequency filters). Individual 50/60 Hz filter controls should be useful for any sleep recording system, but their use should not be a substitute for proper electrode application and recording techniques. Digital filters for the line interference can be designed with sharp edges.

6.6 Analog to digital conversion sampling rates

Present-day computers are more than capable of offering the necessary processing speeds and storage capabilities to accommodate higher sampling rates. Channel-by-channel selection allows the operator to fine-tune the recording in the same manner as selecting appropriate high and low frequency filters for each parameter. Minimum settings for signals are no longer compromises to technology but are based on frequency content according to signal physiology. Except for ECG, there are few published consensus documents available; therefore, recommendations on minimum and desirable sampling rates were determined using the Rand/UCLA appropriateness process. Any selection of sampling rates was chosen for sleep scoring and may not necessarily be appropriate for other diagnostic tests (e.g. seizure detection, evoked potentials, cardiac ischemia assessment). The EEG and EOG should be treated similarly due to similar physiological signal content. The EMG is mainly interesting for the evaluation of muscle tone and twitches. The actual EMG signal contains frequency components higher than the specified upper limit. For the evaluation of sleep in clinical sleep laboratories, only the envelope of amplitude changes is considered. Therefore sampling rates similar to EEG and EOG are still appropriate.

Respiratory signals such as airflow, nasal pressure, effort belts, and esophageal pressure derive signals reflecting respiration and should enable the recognition of rapid and shallow breathing as well as artifacts. Sampling rate for all these signals can be similar. Oximetry may provide a pulse wave which should be treated as a cardiovascular signal and provides oxygen saturation as a calculated value which is updated at each heart beat at best. Therefore oxygen saturation does not require a high sampling rate. The actual update interval has to be documented with the technical specifications of the system. Body position changes do not require a high sampling rate either. Snoring sound recording is similar to EMG recording. In order to evaluate the actual waveform a sampling rate appropriate for audio signals would be needed. In sleep labs practitioners are more interested in sound intensity and the rough equivalent (which is amplitude) is sufficient. As a consequence, the sampling of snoring can be treated in a manner similar to EMG. The specific minimal and desirable rates are identified for each parameter in the scoring manual.

In each case, when selecting sampling rates a basic engineering principle has to be considered. When a waveform needs to be reconstructed, no signal with a frequency exceeding half the sampling rate is allowed at the input of the analog-digital converter (the so-called Nyquist law). This is taken care of with the help of "anti-aliasing filters." For a possible sampling rate of 200 Hz, the anti-aliasing filter will remove everything higher than 100 Hz prior to the digitization. In order to display waveforms similar to paper recordings, the sampling rate should be at least four to five times the maximum desired visible frequency, as an engineering rule of the thumb (e.g., to identify 30 Hz EEG waveforms the desired minimum sampling rate for a good display should be at least 150 Hz).

6.7 Impedance measurements

Some present-day digital systems offer an impedance *imbalance* measurement between pairs of electrodes. This does not provide information about the actual impedance levels of individual electrodes. Thus, both electrodes may have equally high impedances, yet show a low reading. A separate (hand-held) impedance meter is useful during the electrode application procedure, allowing corrections to be made before the patient is put to bed. An alternative may be a built-in meter within the electrode jackbox/amplifier, provided that the information reflects individual electrode measurements rather than imbalance readings.

6.8 Data files

Corrections and adjustments to the recording should be made during data collection by the attending technologist to ensure high quality sleep studies. Subsequent viewing of the data in the manner it was recorded is essential for streamlining the scoring and interpretation process (i.e., the scorer/interpreter should not have to duplicate the task performed by the recording technologist). Also, the recorded data should be protected against any data manipulation and changes to ensure data integrity. Any changes such as changed filters or changes to the data should result in new files. The new files may just store filter and signal settings and may not duplicate all recorded data.

The use of multiple files provides the capability of making additional changes without altering the initial recording and offers the option of performing scoring comparisons without altering the initial scoring file. For optimal scoring the reader may benefit from being able to view the PSG data in the format used to view it by the scorer.

To allow the exchange of recorded digital data between sleep centers using different digital systems, a common data format for digital polysomnography is required. Current efforts try to harmonize existing data formats in order to achieve one digital polysomnography exchange format. The desired specification should include patient identification, sleep center identification, all recorded signals and settings, as well as annotated events during the recording. The EDF data format is commonly used for the exchange of digitized signal data.[69] This data format is supported by most digital systems. In contrast to many previous file formats, it supports different sampling rates for simultaneous signals. The EDF data file is a binary file and thereby uses efficient digital storage space. Data are stored in continuous blocks of a given duration of 1 second or a multiple thereof. There are no limitations for the duration of the recording and the number of channels. It also specifies general recording information and separate information for each recorded channel. The exact encoding of this information is not standardized, and therefore there are frequent problems when importing EDF data recorded with another digital system. The biggest disadvantage is the inability to record events and annotations. In order to overcome this limitation, an improved version of the format, the EDF+ has been developed. This format is now implemented by many digital systems.

7.0 UNRESOLVED ISSUES AND FUTURE RESEARCH

7.1 Are there proposed changes to R & K that might be important for characterization of sleep by computerized analysis?

Many researchers have suggested alterations to the standardized R & K system that would presumably improve our description and understanding of sleep. Sleep fragmentation is not well represented by current visual sleep scoring. Computerized analysis can provide scoring in smaller epochs than visual scoring would allow.[70] Adaptive segmentation of epoch duration is suggested to score sleep stages. This approach would loosen the rigid time constraint of 30-second epochs and allow a sleep stage to begin and end more precisely where its component activity starts or finishes. In this manner, the 12 seconds of stage 2 occurring in a 30-second epoch of wakefulness would be counted as stage 2 rather than obscured by that epoch's majority of wake activity. There is little doubt that a more accurate representation of the night's sleep would be obtained; however, our search did not reveal any studies with gradable data sources demonstrating this assertion based on computerized analysis, nor any studies investigating whether clinical outcomes would differ with such a scoring approach.

7.2 What are the advantages and disadvantages of computerized polysomnography? How can the advantages be promoted? What remedies if any, can be applied to the disadvantages

Based on all the literature reviewed, the computerized approach has proven to be helpful for researchers to characterize and differentiate groups of subjects differing in age and psychiatric comorbidity. These often are studies which look at comparisons or interventions rather than studies focused on accuracy, reliability, or validity in comparison to visual scoring. Key benefits are that some measures of waveforms that would be too time consuming to assess manually are clearly sensitive to drug ingestion and may provide an assay for understanding pharmacodynamics and mechanisms.

In the clinical realm, automatic scoring of central nervous system arousals could be very helpful. One avenue with potential for providing this information on a routine basis to clinicians is automatic analysis of the cyclic alternating pattern (CAP). Three computer system validation studies were found for CAP as part of our literature search (Table 8—which can be accessed on the web at www.aasmnet.org). Two studies met criteria for Level 4[71,72] while one was a Level 5.[73] Further analysis of CAP characteristics can also be found in the companion review paper on arousals.[74]

As a research tool, parameters that can be easily calculated by the computer show potential to both aid our understanding of sleep mechanisms and provide useful indices for practice. Table 9 (which can be accessed on the web at www.aasmnet.org), presents extractions from 6 papers that used computerized methods to explore sleep-related processes and 9 reports about altered waveforms associated with sleep disorders, sleepiness, or both. The 6 papers exploring sleep-related processes looked at a wide variety of such processes. Slow wave activity has long been conceptualized as a marker for the homeostatic process governing sleep. Its occurrence early in the night and exponential decay over time make it a good physiological candidate to contrast with the circadian biological markers of core body temperature and DLMO.[75,76] More sensitive computerized measures suggest EEG theta correlates with sleepiness in a Level 2 study.[77] Sex differences were also found in a Level 3 study showing women have higher spectral densities than men in NREM and REM sleep.[78] Finally, in an ingenious and novel design, Schabus and colleagues[79] found spindle activity varied as a function of the amount of successfully newly acquired memorized information (Level 3).

The rest of the papers in Table 9 took the approach of investigating possible relationships between sleep mechanisms assessed by computer in subjects with sleep disorders. Our literature search revealed 9 papers concerning computerized polysomnographic measures and sleep-disordered breathing (three Level 2, two Level 3, and one Level 4), 2 papers about insomnia (both Level 3), and one concerning sleepwalking (Level 3).

These Level 2 studies were analyzed. Bennett and coworkers[80] found computerized measures of movement and sleep depth reflected therapeutic improvement after positive airway pressure therapy for sleep related breathing disorders (SRBD). Less slow wave activity was found among patients with sleep apnea, and EEG power in the 7-9 Hz band was greater in patients with airway resistance syndrome than in patients with sleep apnea.[81] Theta and sigma are also reported diminished in patients with SRBD compared to controls.[82] Lower slow wave activity was also reported in a Level 3 report.[83] One of the more intriguing papers, albeit graded only at Level 4 due to a small sample size, examined post-respiratory event resumption of breathing in patients with increased airway resistance after an identifiable arousal and in the absence of an arousal.[84] Spectral analysis revealed increased alpha, sigma, and beta activity just before esophageal pressure drops, even when no arousal occurred. Investigators examining high-frequency components in the EEG reported an association between beta/gamma activity and primary insomnia.[85,86] This association was present not only when comparing subjects with insomnia to controls, but also when comparing groups with insomnia from primary vs. secondary etiologies. In an unrelated Level 3 study, sleepwalking was associated with slow wave activity increases in a 2-minute window preceding the event.

7.3 Computer interpretation or redefining of sleep

Visual scoring (R & K) has been a fundamental underpinning of the sleep field; and initial interest in computer applications focused on whether computer reproduction of this task was valid and reliable. Progress is being made in computer use to characterize and compare sleep states, conditions, and interventions. Many issues remain to be addressed; the degree of reliability in needed for clinical applications, for instance, has not been determined.

R & K focused on EEG, EKG, and EOG recordings and identified practical parameters for use in visual scoring (30-sec epochs, stage continuity rules, etc), but limitations of visual scoring and possibilities of computerized scoring have become increasingly clear. Computers, for instance, can provide greater differentiation of amplitude and background wave forms than can the human eye. Finally, it is not clear whether the goal of computer-based scoring should be epoch-by-epoch matching with visual scoring

The consideration of whether computer scoring can provide more useful information than visual scoring depends on studies of outcomes rather than studies of scoring concordance. Perhaps shorter or longer epochs are physiologically better. Perhaps slow wave sleep should be divided into more subsets based on small differences in amplitude. The outcomes of more detailed analysis that computers can provide has yet to be applied broadly enough

to answer the question of whether it may be better, irrespective of whether it meets the goal of equaling visual scoring. Assessment of computer/digital capacity to mimic the well-trained visual scorer may be useful, but research is still needed to determine whether this technology will contribute new methods for understanding sleep and its disorders.

While these findings suggest discernible relationships between sophisticated computerized EEG parameters and both normal and abnormal sleep, our understanding is minimal at this time. Much more work is needed in this area before the sleep specialist will have an acceptable clinical tool. There are myriad possibilities, but validation remains a question. It is critical that the measurements reflect details about extant phenomena and not artifact. For example, in spectral analysis, a basic assumption of the analytic technique concerns the nature of the signal as stationary. The continued variation and non-sinusoidal nature of EEG can, in some cases, generate information that does not really exist.

Another issue for the future is how to combine advances in computer technology with other advances in sleep medicine. Is our goal to update the new scoring manual using its current format? It may be that computer-based analysis can provide more knowledge about sleep physiology by using techniques that computers can do better than human scorers: quantitative EEG measures, fast Fourier transformation (FFT), period amplitude analysis (PAA), spectral analysis, zero crossings for frequency bands, etc. This will require outcomes data.

Some computer based systems allow users to modify analysis settings. If users with limited knowledge modify these settings, it is not clear whether scoring results will be valid. Studies of software provided by manufacturers, independent investigators, and by national funding sources will be needed to assure clinically valid information.

DIGITAL TASK FORCE MEMBERS

The Digital Task Force members participating in evidence review and/or consensus decisions to derive digital scoring rules and specifications included: Thomas Penzel, chair, Max Hirshkowitz, co-chair, Nick Butkov, Ronald D. Chervin, Meir Kryger, Clete A. Kushida, Beth A. Malow, Michael H. Silber, Michael V. Vitiello, John Harsh, and Andrew L. Chesson Jr.

REFERENCES

1. Loomis AL, Harvey EN, Hobart GA. Cerebral states during sleep, as studied by human brain potentials. J Exp Psychol 1937;21:127–144.
2. Aserinsky E, Kleitman N. Regularly occurring periods of eye motility and concomitant phenomena during sleep. Science 1953;118:273-4.
3. Rechtschaffen A, Kales A**.** A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects Washington, DC: Public Health Service, U.S. Government Printing Office; 1968.
4. Iber C. Development of a new manual for characterizing sleep. Sleep 2004;27:190-2.
5. Iber C, Ancoli-Israel S, Chesson AL, Quan SF**.** 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.
6. Fitch F, Bernstein SJ, Aguilar MS, et al. The RAND/UCLA Appropriateness Method User's Manual:RAND Corporation; 2001.
7. American Electroencephalographic Society guidelines for polygraphic assessment of sleep-related disorders (polysomnography). J Clin Neurophysiol 1992;9:88-96.
8. American Clinical Neurophysiology Society. Guideline 8: Guidelines for recording clinical EEG on digital media. J Clin Neurophysiol 2006;23:122-4.
9. Penzel T, Conradt R. Computer based sleep recording and analysis. Sleep Med Rev 2000;4:131-48.
10. Kushida CA, Littner MR, Morgenthaler T, Alessi CA, Bailey D, Coleman J. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. Sleep 2005;28:499-521.
11. Bailey JJ, Berson AS, Garson A, Jr., et al. Recommendations for standardization and specifications in automated electrocardiography: bandwidth and digital signal processing. A report for health professionals by an ad hoc writing group of the Committee on Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiology, American Heart Association. Circulation 1990;81:730-9.
12. Anderer P, Gruber G, Parapatics S, et al. An E-health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the Somnolyzer 24 x 7 utilizing the Siesta database. Neuropsychobiology 2005;51:115-33.
13. Prinz PN, Larsen LH, Moe KE, Dulberg EM, Vitiello MV. C STAGE, automated sleep scoring: development and comparison with human sleep scoring for healthy older men and women. Sleep 1994;17:711-7.
14. Schaltenbrand N, Lengelle R, Toussaint M, et al. Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients. Sleep 1996;19:26-35.
15. White DP, Gibb TJ. Evaluation of a computerized polysomnographic system. Sleep 1998;21:188-96.
16. Sangal RB, Semery JP, Belisle CL. Computerized scoring of abnormal human sleep: a validation. Clin Electroencephalogr 1997;28:64-7.
17. Stanus E, Lacroix B, Kerkhofs M, Mendlewicz J. Automated sleep scoring: a comparative reliability study of two algorithms. Electroencephalogr Clin Neurophysiol 1987;66:448-56.
18. Pittman SD, MacDonald MM, Fogel RB, et al. Assessment of automated scoring of polysomnographic recordings in a population with suspected sleep-disordered breathing. Sleep 2004;27:1394-403.
19. Todorova A, Hofmann HC, Dimpfel W. A new frequency based automatic sleep analysis - description of the healthy sleep. Eur J Med Res 1997;2:185-97.
20. Mykytyn IJ, Sajkov D, Neill AM, McEvoy RD. Portable computerized polysomnography in attended and unattended settings. Chest 1999;115:114-22.
21. Andreas S, von Breska B, Magnusson K, Kreuzer H. Validation of automated sleep stage and apnoea analysis in suspected obstructive sleep apnoea. Eur Respir J 1993;6:48-52.
22. Hasan J. Differentiation of normal and disturbed sleep by automatic analysis. Acta Physiol Scand Suppl. 1983;526:1-103.
23. Watanabe A. Cerebral changes in hepatic encephalopathy. J Gastroenterol Hepatol 1998;13:752-60.
24. Powell TE, Harding GF. Twenty-four hour ambulatory EEG monitoring: development and applications. J Med Eng Technol 1986;10:229-38.
25. Kumar A, Hofman W, Campbell K. An automatic spindle analysis and detection system based on the evaluation of human ratings of the spindle quality. Waking Sleeping 1979;3:325-33.
26. Huupponen E, Varri A, Himanen SL, Hasan J, Lehtokangas M, Saarinen J. Autoassociative MLP in sleep spindle detection. J Med Syst. 2000;24:183-93.
27. Declerck AC, Martens WL, Wauquier W. Sleep spindle detection and its clinical relevance. Eur Neurol 1986;25 Suppl 2:56-60.
28. Jobert M, Escola H, Poiseau E, Gaillard P. Automatic analysis of sleep using two parameters based on principal component analysis

of electroencephalography spectral data. Biol Cybern 1994;71:197-207.
29. Huupponen E, Himanen SL, Hasan J, Varri A. Sleep depth oscillations: an aspect to consider in automatic sleep analysis. J Med Syst 2003;27:337-45.
30. Feinberg I. Importance of both amplitude and incidence measures in time-domain analysis. Sleep Med Rev 1988;11:571-2.
31. Feinberg I, March JD, Fein G, Aminoff MJ. Log amplitude is a linear function of log frequency in NREM sleep eeg of young and elderly normal subjects. Electroencephalogr Clin Neurophysiol 1984;58:158-60.
33. Schlogl A, Kemp B, Penzel T, et al. Quality control of polysomnographic sleep data by histogram and entropy analysis. Clin Neurophysiol 1999;110:2165-70.
34. Muller C, Achermann P, Bischof M, Nirkko AC, Roth C, Bassetti CL. Visual and spectral analysis of sleep EEG in acute hemispheric stroke. Eur Neurol 2002;48:164-71.
35. Huupponen E, Saastamoinen A, Niemi J, et al. Automated frequency analysis of synchronous and diffuse sleep spindles. Neuropsychobiology 2005;51:256-64.
36. Gade J, Rosenfalck A, Bendtson I. Detection of EEG patterns related to nocturnal hypoglycemia. Methods Inf Med 1994;33:153-6.
37. Burr W, Stefan H. Computerized analysis of epileptic activity and sleep in mobile long-term EEG monitoring. Eur Neurol 1986;25 Suppl 2:61-5.
38. Larsen LH, Prinz PN, Moe KE. Quantitative analysis of the EEG during tonic REM sleep—methodology. Electroencephalogr Clin Neurophysiol 1992;83:24-35.
39. Mann K, Backer P, Roschke J. Dynamical properties of the sleep EEG in different frequency bands. Int J Neurosci 1993;73:161-9.
40. Roschke J, Aldenhoff J. The dimensionality of human's electroencephalogram during sleep. Biol Cybern 1991;64:307-13.
41. Kawada T, Kiryu Y, Naganuma S, Aoki S, Suzuki S. Effect of the continuity and sampling rate of polygraphy data on sleep stage by a computerized scoring system—a reliability study. Nippon Eiseigaku Zasshi 1992;47:826-30.
42. Kupfer DJ, Ulrich RF, Coble PA, et al. Electroencephalographic sleep of younger depressives. Comparison with normals. Arch Gen Psychiatry. 1985;42:806-10.
43. Zucconi M, Oldani A, Ferini-Strambi L, Smirne S. Arousal fluctuations in non-rapid eye movement parasomnias: the role of cyclic alternating pattern as a measure of sleep instability. J Clin Neurophysiol 1995;12:147-54.
44. von Krosigk M, Bal T, McCormick DA. Cellular mechanisms of a synchronized oscillation in the thalamus. Science 1993;261:361-4.
45. Steriade M, McCormick DA, Sejnowski TJ. Thalamocortical oscillations in the sleeping and aroused brain. Science 1993;262:679-85.
46. Steriade M. Cellular substrates of brain rhythms. In: Niedermeyer E, Lopes da Silva F, eds. Electroencephalography. Baltimore, MD: Williams and Wilkins; 1999.
47. McCormick DA, Bal T. Sleep and arousal: thalamocortical mechanisms. Annu Rev Neurosci 1997;20:185-215.
48. Contreras D, Destexhe A, Sejnowski TJ, Steriade M. Spatiotemporal patterns of spindle oscillations in cortex and thalamus. J Neurosci 1997;17:1179-96.
49. Doman J, Detka C, Hoffman T, et al. Automating the sleep laboratory: implementation and validation of digital recording and analysis. Int J Biomed Comput 1995;38:277-90.
50. Principe JC, Smith JR. Sleep spindle characteristics as a function of age. Sleep 1982;5:73-84.
51. Smith JR, Karacan I, Yang M. Automated measurement of alpha, beta, sigma, and theta burst characteristics. Sleep 1979;1:435-43.
52. Reynolds CF 3rd, Monk TH, Hoch CC, et al. Electroencephalographic sleep in the healthy "old old": a comparison with the "young old" in visually scored and automated measures. J Gerontol 1991;46: M39-46.
53. Reynolds CF 3rd, Kupfer DJ, Taska LS, Hoch CH, Sewitch DE, Grochocinski VJ. Slow wave sleep in elderly depressed, demented, and healthy subjects. Sleep 1985;8:155-9.
54. Darchia N, Campbell IG, Palagini L, Feinberg I. Rapid eye movement density shows trends across REM periods but is uncorrelated with NREM delta in young and elderly human subjects. Brain Res Bull 2004;63:433-8.
55. Hirshkowitz M, Thornby JI, Karacan I. Sleep spindles: pharmacological effects in humans. Sleep 1982;5:85-94.
56. Smith JR, Karacan I, Keane BP, Yang M. Automated sleep EEG analysis applied to the evaluation of drugs: illustration by study of clorazepate dipotassium. Electroencephalogr Clin Neurophysiol 1976;41:587-94.
57. Hirshkowitz M, Thornby JI, Karacan I. Sleep pharmacology and automated EEG analysis. Psychiatry Annals. 1979;9:510-20
58. Kajimura N, Kato M, Okuma T, Sekimoto M, Watanabe T, Takahashi K. A quantitative sleep-EEG study on the effects of benzodiazepine and zopiclone in schizophrenic patients. Schizophr Res 1995;15:303-12.
59. Johnson LC, Hanson K, Bickford RG. Effect of flurazepam on sleep spindles and K-complexes. Electroencephalogr Clin Neurophysiol 1976;40:67-77.
60. McPartland RJ, Kupfer DJ, Coble P, Shaw DH, Spiker DG. An automated analysis of REM sleep in primary depression. Biol Psychiatry 1979;14:767-76.
61. Kupfer DJ, Shaw DH, Ulrich R, Coble PA, Spiker DG. Application of automated REM analysis in depression. Arch Gen Psychiatry 1982;39:569-73.
62. Kupfer DJ, Reynolds CF 3rd, Ulrich RF, et al. Comparison of automated REM and slow-wave sleep analysis in young and middle-aged depressed subjects. Biol Psychiatry 1986;21:189-200.
63. Kupfer DJ, Ulrich RF, Coble PA, et al. Application of automated REM and slow wave sleep analysis: II. Testing the assumptions of the two-process model of sleep regulation in normal and depressed subjects. Psychiatry Res 1984;13:335-43.
64. Kupfer DJ, Ulrich RF, Coble PA, et al. Application of automated REM and slow wave sleep analysis: I. Normal and depressed subjects. Psychiatry Res 1984;13:325-34.
65. Reynolds CF 3rd, Kupfer DJ, Thase ME, et al. Sleep, gender, and depression: an analysis of gender effects on the electroencephalographic sleep of 302 depressed outpatients. Biol Psychiatry 1990;28:673-84.
66. Roschke J, Fell J, Beckmann P. Nonlinear analysis of sleep EEG in depression: calculation of the largest lyapunov exponent. Eur Arch Psychiatry Clin Neurosci 1995;245:27-35.
67. Roschke J, Mann K. The sleep EEG's microstructure in depression: alterations of the phase relations between EEG rhythms during REM and NREM sleep. Sleep Med 2002;3:501-5.
68. Roschke J, Mann K, Fell J, Roschke J, Mann K, Fell J. Nonlinear EEG dynamics during sleep in depression and schizophrenia. Int J Neurosci 1994;75:271-84.
69. Kemp B, Värri A, Rosa AC, Nielsen KD, Gade J. A simple format for exchange of digitized polygraphic recordings. Electroencephalogr Clin Neurophysiol 1992;82:391-3.
70. Black JE, Guilleminault C, Colrain IM, et al. Upper airway resistance syndrome. Central electroencephalographic power and changes in breathing effort. Am J Respir Crit Care Med 2000;162:406-11.
71. Ferrillo F, Gabarra M, Nobili L, et al. Comparison between visual scoring of cyclic alternating pattern (CAP) and computerized assessment of slow EEG oscillations in the transition from light to deep non-REM sleep. J Clin Neurophysiol 1997;14:210-6.
72. Ferri R, Bruni O, Miano S, Smerieri A, Spruyt K, Terzano MG. Inter-rater reliability of sleep cyclic alternating pattern (CAP) scoring and validation of a new computer-assisted CAP scoring method. Clin Neurophysiol. 2005;116:696-707.
73. Rosa AC, Parrino L, Terzano MG. Automatic detection of cyclic

alternating pattern (CAP) sequences in sleep: preliminary results. Clin Neurophysiol 1999;110:585-92.
74. Bonnet MH, Karl Doghramji K, Timothy Roehrs T, Stepanski EJ, Chesson A. The scoring of arousal in sleep: reliability, validity, and alternatives. J Clin Sleep Med: 2007:3:133-145.
75. Feinberg I, Floyd TC, March JD. Acute deprivation of the terminal 3.5 hours of sleep does not increase delta (0-3-Hz) electroencephalograms in recovery sleep. Sleep 1991;14:316-9.
76. Feinberg I, Baker T, Leder R, March JD. Response of delta (0-3 Hz) EEG and eye movement density to a night with 100 minutes of sleep. Sleep. 1988;11:473-87.
77. Wichniak A, Geisler P, Brunner H, et al. Spectral composition of NREM sleep in healthy subjects with moderately increased daytime sleepiness. Clin Neurophysiol 2003;114:1549-55.
78. Dijk DJ, Beersma DG, Bloem GM. Sex differences in the sleep EEG of young adults: visual scoring and spectral analysis. Sleep 1989;12:500-7.
79. Schabus M, Gruber G, Parapatics S, et al. Sleep spindles and their significance for declarative memory consolidation.[see comment]. Sleep 2004;27:1479-85.
80. Bennett LS, Langford BA, Stradling JR, Davies RJ. Sleep fragmentation indices as predictors of daytime sleepiness and nCPAP response in obstructive sleep apnea. Am J Respir Crit Care Med 1998;158:778-86.
81. Guilleminault C, Do Kim Y, Chowdhuri S, Horita M, Ohayon M, Kushida C. Sleep and daytime sleepiness in upper airway resistance syndrome compared to obstructive sleep apnoea syndrome.[comment]. Eur Respir J 2001;17:838-47.
82. Ondze B, Espa F, Dauvilliers Y, Billiard M, Besset A. Sleep architecture, slow wave activity and sleep spindles in mild sleep disordered breathing. Clin Neurophysiol 2003;114:867-74.
83. Morisson F, Decary A, Petit D, Lavigne G, Malo J, Montplaisir J. Daytime sleepiness and EEG spectral analysis in apneic patients before and after treatment with continuous positive airway pressure. Chest 2001;119:45-52.
84. Black JE, Guilleminault C, Colrain IM, Carrillo O. Upper airway resistance syndrome. Central electroencephalographic power and changes in breathing effort. Am J Respir Crit Care Med 2000;162:406-11.
85. Perlis ML, Smith MT, Andrews PJ, Orff H, Giles DE. Beta/Gamma EEG activity in patients with primary and secondary insomnia and good sleeper controls. Sleep 2001;24:110-7.
86. Perlis ML, Kehr EL, Smith MT, Andrews PJ, Orff H, Giles DE. Temporal and stagewise distribution of high frequency EEG activity in patients with primary and secondary insomnia and in good sleeper controls. J Sleep Res 2001;10:93-104.
87. Hoffmann R, Moffitt A, Wells R, Sussman P, Pigeau R, Shearer J. Quantitative description of sleep stage electrophysiology using digital period analytic techniques. Sleep 1984;7:356-64.
88. Sforza E, Vandi S. Automatic Oxford-Medilog 9200 sleep staging scoring: comparison with visual analysis. J Clin Neurophysiol 1996;13:227-33.
89. Albertario CL, Zendell SM, Hertz G, Maberino MM, Feinsilver SH. Comparison of a frequency-based analysis of electroencephalograms (Z-ratio) and visual scoring on the multiple sleep latency test. Sleep 1995;18:836-43.
90. Hasan J, Hirvonen K, Varri A, Hakkinen V, Loula P. Validation of computer analysed polygraphic patterns during drowsiness and sleep onset. Electroencephalogr Clin Neurophysiol 1993;87:117-27.
91. Haustein W, Pilcher J, Klink J, Schulz H. Automatic analysis overcomes limitations of sleep stage scoring. Electroencephalogr Clin Neurophysiol 1986;64:364-74.
92. Hirvonen K, Hasan J, Hakkinen V, Varri A, Loula P. The detection of drowsiness and sleep onset periods from ambulatory recorded polygraphic data. Electroencephalogr Clin Neurophysiol 1997;102:132-7.
93. Holler L, Riemer H. Comparison of visual analysis and automatic sleep stage scoring (Oxford Medilog 9000 System). Eur Neurol 1986;25 Suppl 2:36-45.
94. Agarwal R, Takeuchi T, Laroche S, Gotman J. Detection of rapid-eye movements in sleep studies. IEEE Trans Biomed Eng 2005;52:1390-6.
95. Agarwal R, Gotman J. Computer-assisted sleep staging. IEEE Trans Biomed Eng 2001;48:1412-23.
96. Boukadoum AM, Ktonas PY. Non-random patterns of REM occurrences during REM sleep in normal human subjects: an automated second-order study using Markovian modeling. Electroencephalogr Clin Neurophysiol 1988;70:404-16.
97. Ferri R, Ferri P, Colognola RM, Petrella MA, Musumeci SA, Bergonzi P. Comparison between the results of an automatic and a visual scoring of sleep EEG recordings. Sleep 1989;12:354-62.
98. Matsuoka S, Ishikawak T, Inoue K, Hatashi A. [Automatic determination system of human sleep stages on an experimental basis]. J Uoeh. 1986;8 Suppl:169-71.
99. Park HJ, Oh JS, Jeong DU, Park KS. Automated sleep stage scoring using hybrid rule- and case-based reasoning. Comput Biomed Res 2000;33:330-49.
100. Ray SR, Lee WD, Morgan CD, Airth-Kindree W. Computer sleep stage scoring—an expert system approach. Int J Biomed Comput 1986;19:43-61.
101. Agarwal R, Gotman J. Digital tools in polysomnography. J Clin Neurophysiol 2002;19:136-43.
102. Hasan J. Past and future of computer-assisted sleep analysis and drowsiness assessment. J Clin Neurophysiol 1996;13:295-313.
103. Hirshkowitz M, Moore CA. Issues in computerized polysomnography. Sleep 1994;17:105-12.
104. Ktonas PY. Automated analysis of abnormal electroencephalograms. Crit Rev Biomed Eng 1983;9:39-97.
105. Kubicki S, Herrmann WM. The future of computer-assisted investigation of the polysomnogram: sleep microstructure. J Clin Neurophysiol 1996;13:285-94.
106. Roberts S, Tarassenko L. New method of automated sleep quantification. Med Biol Eng Comput 1992;30:509-17.
107. Smith JR. Automated EEG analysis with microcomputers. Med Instrum 1980;14:319-21.
108. Gath I, Bar-on E. Computerized method for scoring of polygraphic sleep recordings. Comput Programs Biomed. 1980;11:217-23.
109. Jansen BH, Dawant BM. Knowledge-based approach to sleep EEG analysis—a feasibility study. IEEE Trans Biomed Eng 1989;36:510-8.
110. Kuwahara H, Higashi H, Mizuki Y, Matsunari S, Tanaka M, Inanaga K. Automatic real-time analysis of human sleep stages by an interval histogram method. Electroencephalogr Clin Neurophysiol 1988;70:220-9.
111. Smith JR, Karacan I. EEG sleep stage scoring by an automatic hybrid system. Electroencephalogr Clin Neurophysiol 1971;31:231-7.
112. Baumgart-Schmitt R, Herrmann WM, Eilers R, Bes F. On the use of neural network techniques to analyse sleep EEG data. First communication: application of evolutionary and genetic algorithms to reduce the feature space and to develop classification rules. Neuropsychobiology. 1997;36:194-210.
113. Courtney P, Noton D. A hybrid computer system for unsupervised scoring of sleep records. Biomed Sci Instrum 1972;9:161-7.
114. Escourrou P, Luriau S, Rehel M, Nedelcoux H, Lanoe JL. Needs and costs of sleep monitoring. Stud Health Technol Inform 2000;78:69-85.
115. Hoffmann RF, Moffitt AR, Shearer JC, Sussman PS, Wells RB. Conceptual and methodological considerations towards the development of computer-controlled research on the electro-physiology of sleep. Waking Sleeping. 1979;3:1-16.
116. Kapfhammer G, Raber W. A computerized system for acquisition and evaluation of polysomnographic recordings. Int J Clin Monit

Comput 1992;9:111-6.
117. Schaltenbrand N, Lengelle R, Macher JP. Neural network model: application to automatic analysis of human sleep. Comput Biomed Res 1993;26:157-71.
118. Huupponen E, Himanen SL, Varri A, et al. Fuzzy detection of EEG alpha without amplitude thresholding. Artif Intell Med 2002;24:133-47.
119. Nakata M, Mukawa J, Fromm GH. Evaluation of human consciousness level by means of "Automated Fluctuation Analysis" of high frequency electroencephalogram fitted by double Lorentzians. Integr Physiol Behav Sci 1993;28:343-52.
120. Hoffmann R, Jeakins D. Computer quantification of delta activity in sleep EEG. Comput Biomed Res 1987;20:366-72.
121. Roschke J, Aldenhoff JB. Excitability and susceptibility of the brain's electrical activity during sleep: an analysis of late components of AEPs and VEPs. Int J Neurosci 1991;56:255-72.
122. Brunner DP, Vasko RC, Detka CS, Monahan JP, Reynolds CF 3rd, Kupfer DJ. Muscle artifacts in the sleep EEG: automated detection and effect on all-night EEG power spectra. J Sleep Res 1996;5:155-64.
123. Ktonas PY, Osorio PL, Everett RL. Automated detection of EEG artifacts during sleep: preprocessing for all-night spectral analysis. Electroencephalogr Clin Neurophysiol 1979;46:382-8.
124. Ktonas P, Nygren A, Frost J Jr. Two-minute rapid eye movement (REM) density fluctuations in human REM sleep. Neurosci Lett 2003;353:161-4.
125. Michel CJ. A computer method for identifying patterns in electroencephalogram signals. J Med Eng Technol 2003;27:267-75.
126. Reynolds CF 3rd, Hoch CC, Buysse DJ, et al. Sleep in late-life recurrent depression. Changes during early continuation therapy with nortriptyline. Neuropsychopharmacology. 1991;5:85-96.
127. Ehlers CL, Kupfer DJ. Effects of age on delta and REM sleep parameters. Electroencephalogr Clin Neurophysiol 1989;72:118-25.
128. Astrom C, Trojaborg W. Effect of growth hormone on human sleep energy. Clin Endocrinol 1992;36:241-5.
129. Borbely AA, Akerstedt T, Benoit O, Holsboer F, Oswald I. Hypnotics and sleep physiology: a consensus report. European Sleep Research Society, Committee on Hypnotics and Sleep Physiology. Eur Arch Psychiatry Clin Neurosci 1991;241:13-21.
130. Rosadini G, De Carli F, Ferrillo F. An EEG computerized system for the evaluation of hypnotic drugs. Int J Psychophysiol 1992;13:291-7.
131. Liscombe MP, Hoffmann RF, Trivedi MH, Parker MK, Rush AJ, Armitage R. Quantitative EEG amplitude across REM sleep periods in depression: preliminary report. J Psychiatry Neurosci 2002;27:40-6.
132. Tekell JL, Hoffmann R, Hendrickse W, Greene RW, Rush AJ, Armitage R. High frequency EEG activity during sleep: characteristics in schizophrenia and depression. Clin EEG Neurosci 2005;36:25-35.
133. Keshavan MS, Reynolds CF 3rd, Haas G, Miewald JM, Montrose DM. Sleep EEG changes in psychotic disorders: gender and age effects. Neuropsychobiology 1995;32:1-8.
134. Bendtson I, Gade J, Thomsen CE, Rosenfalck A, Wildschiodtz G. Sleep disturbances in IDDM patients with nocturnal hypoglycemia. Sleep 1992;15:74-81.
135. Wichniak A, Geisler P, Brunner H, et al. Spectral composition of NREM sleep in healthy subjects with moderately increased daytime sleepiness. Clin Neurophysiol 2003;114:1549-55.
136. Armitage R, Hoffmann RF. Sleep EEG, depression and gender. Sleep Med Rev 2001;5:237-46.
137. Guilleminault C, Poyares D, Aftab FA, Palombini L, Abat F. Sleep and wakefulness in somnambulism: a spectral analysis study.[erratum appears in J Psychosom Res 2001 Oct;51:621 Note: Abat F [corrected to Aftab FA]. J Psychosom Res 2001;51:411-6.
138. Heinzer R, Gaudreau H, Decary A, et al. Slow-wave activity in sleep apnea patients before and after continuous positive airway pressure treatment: contribution to daytime sleepiness.[comment]. Chest 2001;119:1807-13.
139. Black JE, Guilleminault C, Colrain IM, et al. Upper airway resistance syndrome. Central electroencephalographic power and changes in breathing effort. Am J Respir Crit Care Med 2000;162:406-11.
140. Espa F, Ondze B, Deglise P, Billiard M, Besset A. Sleep architecture, slow wave activity, and sleep spindles in adult patients with sleepwalking and sleep terrors. Clin Neurophysiol 2000;111:929-39.

Table 3—Evidence table related to general computer application and validation of sleep staging

Author/ref	Year	Grade	Study Type	Subjects	Results
Anderer/(12)	2005	1	Multicenter	590 recordings	Normal healthy subjects, aged 20-95 yr and patients with sleep disorders. Overall epoch-by-epoch agreement was 80% (Cohen's kappa: 0.72) between the Somnolyzer 24 x 7 and the human expert scoring, compared with an inter-rater reliability of 77% (Cohen's kappa: 0.68) between 2 human experts scoring the same dataset.
Prinz/(13)	1994	1	Consecutive sample	200 men & women.	Epoch-by-epoch comparisons in the validation sample revealed a mean proportion of agreement of 0.74 and a mean kappa coefficient of 0.57, indicating the 2 methods provide reasonable agreement on an epoch-by-epoch basis. C STAGE is a valid sleep/waking scoring system for healthy older adults.
Schaltenbrand/(14)	1996	2	Cross-sectional	60	Average agreement rate of 87.5% between 2 experts on basis of 30-second epochs.automatic/expert agreement grew from 82.3% to 90%
White/(15)	1998	2	Clinical series	50 patients	When allowed to autoscore, ALICE 3 produced substantial errors in sleep staging (REM sleep time 56.4 + 4.9 minutes vs 73.2 + 8.4 minutes for paper versus computer). However, with editing, it produced results similar to those obtained with paper.
Sangal/(16)	1997	3	Consecutive clinical sample	30 patients	The ratios of computer-human agreement descriptors to human-human agreement descriptors indicate that computerized analysis of abnormal human sleep offers reasonable results with savings in technologist time and work, but not in physician time and work.
Stanus/(17)	1987	3	Cross sectional sample	15 patients & 15 controls	Deterministic and stochastic sleep staging (DSS and SSS) methods were compared with expert visual analysis; found 70-75% concordance
Pittman/(18)	2004	3	Clinical sample	31 PSGs	Morpheus I Sleep Scoring System was tested for sleep staging against 2human scorers. Overall agreement was 78% with one scorer and 73% with the other. The scorers agreement with one another was 82%
Todorova/(19)	1997	3	Consecutive sample	18, 9 of each sex	The characterization of the recorded bioelectrical signals is based on the spectral frequency analysis by Fast Fourier Transformation (FFT). A new sleep-staging algorithm was developed, which allows for an objective presentation of the sleep phenomena. This system is characterized by a high reproducibility of the individual polysomnographic data throughout the recording with no significant difference between different all-night recordings of the subject.
Mykytyn/(20)	1999	3	Clinical sample	20 patients	Portable polysomnography is a viable alternative to laboratory-based polysomnography and may be improved further by better sensor attachment
Andreas/(21)	1993	3	Consecutive sample	27 patients	The mean difference between visual scoring and automated analysis was -1, 111, -140, -3, 1 and 27 min, for sleep stages awake, 1, 2, 3, 4, and REM, respectively.
Hasan/(22)	1983	3	Representative sample	24 patients	Authors developed and tested an automatic hybrid system in 3 groups. The computer vs. human scoring percentage agreement obtained for the 3 groups: young normals 80%, older normals 77%, and anonymous alcoholics 75%.
Doman/(49)	1995	5	Instructional	n/a	Information about setting up, calibrating, and validating a computerized laboratory
Hoffmann/(87)	1984	3	Unsystematic sample	24 nights from 8 subjects	The results indicate a high predictive accuracy (91.05%), supporting the contention that the computer-quantified data set includes the variance normally captured by stage scoring
Sforza/(88)	1996	3	Clinical sample	30 patients	Computerized scoring significantly underestimated TST and stage 2 and overscored stage 1. No significant differences were found for SWS and REM sleep scoring.
Albertario/(89)	1995	4	Clinical sample	13 patients	Used Z-ratio to classify as awake or asleep on MSLT recordings. agreed with human scorers on approximately 80% of epochs

Hasan/(90)	1993	4	Clinical sample	9 patient MSLT recordings	The agreements between the computer and visual scores were relatively good for 5 subjects having a prominent occipital alpha activity during wakefulness (range 70-79%) but less promising (range 64-70%) for the other 4 subjects with “poor” occipital alpha activity.
Haustein/(91)	1986	4	Clinical sample	12 PSG recordings	The results show that the EEG parameters agree with the traditional sleep scoring method. Detailed information is not available.
Hirvonen/(92)	1997	4	Clinical sample	7 patients with OSA	Epoch-by-epoch agreement between visual and computer scoring for wake vs. sleep was over 90% and the agreement for NREM was 64%.
Holler/(93)	1986	4	Clinical sample	4 PSG recordings	The Oxford Medilog 9000 System with Sleep Stager was tested. The automatic analysis consistently scored less REM than the visual assessment. More testing is needed.
Agarwal/(94)	2005	5	Case series	5 subjects	REM sleep detected with an overall sensitivity of 67.2% and specificity of 77.5%.
Agarwal/(95)	2001	5	Case series	12 records	Records of varying types (normal, abnormal, male, female, varying age groups used and they found 80.6% agreement with manual scoring of 20-s epochs
Boukadoum/(96)	1988	5	Case series	6 controls	There may exist 2 aspects to REM generation, a relatively unstable phasic aspect, quantified by first-order Markovian parameters, and a more stable tonic aspect, quantified by second- and possibly higher-order parameters
Ferri/(97)	1989	5	Case series	uncertain	Data are compared with the results of an automatic scoring of the same recordings. The validity of this automatic method of scoring is discussed.
Matsuoka/(98)	1986	5	Case series	4 divers over extended time	This system was very useful for the analysis of the tremendous volume of sleep records during a nearly one-month experiment.
Park/(99)	2000	5	Case series	3 normals and 3 patients with OSA	A sleep stage scoring system was developed that used a hybrid rule- and case-based reasoning. Average agreement rate in normal recordings was 87.5% and case-based scoring enhanced the agreement rate by 5.6%.
Ray/(100)	1986	5	Case series	7 sleep recordings	This expert system approach reported an overall 89.6% accuracy in classifying sleep stages compared to human scoring
Agarwal/(101)	2002	5	Review	n/a	Automatic sleep staging methods is emphasized with some illustrative results on inter-scorer variations.
Hasan/(102)	1996	5	Review	n/a	Review of computerized sleep and drowsiness detection.
Hirshkowitz/(103)	1994	5	Review	n/a	Issues facing computerized polysomnography were discussed.
Ktonas/(104)	1983	5	Review	n/a	Discusses automated EEG analysis and abnormal event (e.g. spikes) detection
Kubicki/(105)	1996	5	Review	n/a	Past and future of computers in sleep medicine discussed
Roberts/(106)	1992	5	Review	n/a	A method of analyzing the EEG that does not require rules and aims to give some indication of the dynamics of sleep in humans is proposed
Smith/(107)	1980	5	Review	n/a	Describes microcomputer technique and approach to EEG analysis
Gath/(108)	1980	5	Development project	n/a	A new method for automatic analysis of polygraphic sleep recordings was described. The output available is the hypnogram, describing the sequence of sleep stages during the night, graphs describing the relative power changes of the main EEG activities (delta, theta, alpha, and sigma), and 3-dimensional plots of the spectral changes during the night.

Gath/(108)	1980	5	Instructional	n/a	3-dimensional plots of the spectral changes during the night. In addition, a few quantitative measures related to the relative amount of the various sleep stages and to the number of rapid eye movements are calculated.
Jansen/(109)	1989	5	Development project	n/a	The system detected K complexes and sleep spindles. Its performance indicates that the approach followed is feasible and can become a powerful tool for automated EEG interpretation
Kuwahara/(110)	1988	5	Development project	n/a	Overall agreement (89.1%) between the computer and human scorers; only 3% less than the agreement (92.1%) among the human scorers. The primary areas of disagreement occurred in stages 1, 2, and REM.
Smith/(111)	1971	5	Development Project	n/a	Design information and goals for the sleep analyzing hybrid computer system.
Baumgart-Schmitt/(112)	1997	5	Instructional	n/a	The goal was to automate sleep stage scoring. The system sleep analysis system to challenge innovative artificial networks (SASCIA) was developed and implemented. The profiles generated by SASCIA were in reasonable agreement with the sleep stages scored by experts
Courtney/(113)	1972	5	Instructional	n/a	A hybrid computer system for completely automated scoring of sleep records is described.
Escourrou/(114)	2000	5	Instructional	n/a	Discusses needs and cost of equipment in a sleep laboratory emphasizing computerized polysomnograph.
Hoffmann/(115)	1979	5	Instructional	n/a	This paper includes a critical discussion of the visual scoring approach to the measurement of sleep electrophysiology, details some theoretical shortcomings of that procedural model. An alternative approach employing high-speed, general purpose digital computers is presented.
Kapfhammer/(116)	1992	5	Instructional	n/a	Created a system running on a microcomputer emulating all functions of a conventional chart recorder and offering the advantage of making computer assisted evaluations.
Schaltenbrand/(117)	1993	5	Instructional	n/a	Neural network models were used for sleep staging and simulated on a digital computer. First, automatic sleep stage scoring was performed using a multilayer feedforward network. Second, supervision of the automatic decision was achieved using ambiguity rejection and artifact rejection. Third, numerical analysis of sleep was carried out using all-night spectral analysis for the background activity of the EEG and sleep pattern detectors for the transient activity.

Table 4—Evidence table related to waveform detection & analysis

Author/ref	Year	Grade	Study Type	Subjects	Results
Watanabe/(23)	1998	5	Review	n/a	Describes cerebral changes in hepatic encephalopathy.
Powell/(24)	1986	5	Development	n/a	Automated analysis techniques were used to score spike and wave activity and sleep stages.
Kumar/(25)	1979	5	Instructional	n/a	A complex demodulation-based spindle detector was compared with both the human visual scoring and with a second automatic system employing phaselocked loop techniques.
Huupponen/(26)	2000	5	Instructional	n/a	Authors describe a spindle detection system allowing spindle detection without an amplitude threshold that can be used for automatic decision making of whether or not a sleep spindle is present in the EEG at a certain point of time
Declerck/(27)	1986	1	Archived sample	500 records	Compared output of 3 spindle detectors with manual validation.
Jobert/(28)	1994	2		16 elderly insomniacs and 16 young healthy subjects	A computerized method for detecting episodes of wakefulness during sleep based on the alpha slow-wave index (ASI). The rate of agreement between the computerized procedure and the visual scoring of wakefulness was 94.0% for the insomniacs and 96.9% for the healthy subjects.
Huupponen/(29)	2003	5	Selected	?	Quantitative analysis of these oscillations was done in this work via a mean frequency measure and FFT. Overall characteristics of these oscillations were studied, focusing on the waves with period times of 5-150 s.
Roizenblatt/(32)	1997	3	Sequential case series	34 children with fibromyalgia; 11 with diffuse pain	Significant concordance was observed regarding FM diagnosis in children and their mothers. Sleep complaints and polysomnography findings were less prominent in affected children compared to mothers with FM. A significant correlation between polysomnographic indexes, sleep anomalies, and pain manifestations in children and their mothers was observed.
Muller/(34)	2002	3	Clinical series and control group	20 patients and 10 controls	TST and time spent in SWS and REM sleep were lower in stroke patients, but differences were not significant. The slow wave activity (SWA) ratio NREM sleep/wakefulness was lower in patients than in controls ($p < 0.05$).
Huupponen/(35)	2005	4	Unsystematic sample		The spindle detector had a 76.17% true positive rate and 0.93% false-positive rate. Pure central spindles were faster and pure frontal spindles were slower than diffuse spindles measured simultaneously from both locations.
Gade/(36)	1994	4	Case series	?	The aim of the project was to detect specific EEG patterns related to hypoglycemia. The rate of these specific EEG patterns was below 5% in normal nights. In patients who were known to have no or a reduced glucagon response to hypoglycemia, the rate increased to 20%-80%.
Burr/(37)	1986	4	Clinical series	?	Tested algorithm for paroxysmal EEG activity to detect seizures.
Larsen/(38)	1992	4	Review	n/a	Robust time series analysis techniques and a modified power spectral analysis (Z-spectra) are used to suppress artifactual information and to automatically select samples of tonic REM sleep EEG. The spectra (amplitude vs. frequency relationship) of this specific EEG state is then assessed for diagnostically relevant information.
Kawada/(41)	1992	4	Unsystematic sample	5 records from 6 young normal males	Scoring reliability declined as the sampling resolution decreased from continuous to 1/3rd, 1/9th, and 1/15th resolution with 1/9th and 1/15th reaching significance.
Zucconi/(43)	1995	3	Clinical series with control	21 patients and 6 controls	Classic sleep parameters were no different in the patients and controls. However, compared with the controls, patient sleep microstructure showed increases in CAP rate, the number of the CAP cycles, and the number of arousals with EEG synchronization.
Nakata/(119)	1993	3	Unsystematic sample	74 records from 20 normal adults	power spectral density of high frequency EEG activity was composed of double Lorentzian fluctuations, and the power distribution of S1 value in topographical display was frontal dominant. This pattern of S1 value disappeared and S2 value became lower during sleepiness and the second Lorentz disappeared during sleep.

Hoffmann/ (120)	1987	4	Clinical series	94, 25-sec epochs of data from 2 subjects	A comparison of inter-scorer & scorer/computer agreements showed good correspondence between them, with the human/ computer contrasts being equivalent to the human/human.
Roschke/ (121)	1991	4	Unsystematic sample	10 healthy males	Evoked potentials considered as transfer functions in an oscillating system. a clear alpha resonance during stage 1, a pronounced delta resonance during stages 2, 3, and 4, and a theta resonance during REM sleep.
Brunner/(122)	1996	5	case series	Adults	Artifact detection and its effects on all night automatic spectral analysis
Ktonas/(123)	1979	5	Case series	Adults	Artifacts could be detected and skipped over in subsequent spectral analysis.
Ktonas/(124)	2003	5	Case series	five healthy young adults	Autoregressive (AR) spectral analysis of the REM density time series estimated periodicities in the range of 1.7-2.4 min for 14 of the 16 REM periods analyzed, regardless of duration or number of REMs.
Michel/(125)	2003	5	Case series	?	Application of computerized sleep analysis of the healthy adult shows a periodicity modulo 10 in all derivations. A possible neurophysiological meaning is presented in the discussion.

Table 5—Evidence table related to automated analysis of age-related sleep changes

Author/ref	Year	Grade	Study Type	Subjects	Results
Prinz/(13)	1994	1	Consecutive sample	200 men & women.	See table 3
Principe/(50)	1982	1	Nonrandom sample	5 age groups: 3-5; 13; 25-34; 42-53; 67-79 yd	The frequency increased with increasing age. No age differences were found in spindle duration. Spindle amplitude reached a peak in the group I subjects, then decreased with increasing age.
Smith/(51)	1979	1	Cohort	5 normal subjects in 5 age groups: 3-5; 13; 25-34; 43-53; and 67-79 y	No age-related changes were found in the alpha frequency (except for the younger group). During REM sleep, the average beta and theta frequencies of the 2 youngest groups were significantly different from those of the 3 older groups. The average frequency of stage 2 sleep spindles of the two youngest groups was less than that in the middle group; average spindle frequency of this group was significantly less than that of the 2 older groups.
Reynolds/(53)	1985	2	Cohort	Elderly pts with dementia and controls	Computerized analysis of SWA was able to differentiate between groups in terms of decrease delta power.
Darchia/(54)	2004	2	Cohort	19 young normal adults and 19 elderly normal adults	Digitized electrooculograms were analyzed with the extensively validated zero-cross period-amplitude module of PASS PLUS software. Incidence of eye movements during REM sleep is substantially reduced in the elderly.
Reynolds/(126)	1991	2	Cohort	groups of 60, 70, and 80 year old subjects	Sleep efficiency and REM sleep were stable across 3 decades of late life, but a slight decline of slow wave sleep in the 80-year-olds (decreased total delta wave counts). Women showed better preservation of slow wave sleep than men.
Ehlers/(127)	1989	3	Nonrandom sample	24 healthy men	Measured sleep stages, computer-assisted delta and REM quantification, and power spectral analysis.

Table 6—Evidence table related to drug study applications

Author/ref	Year	Grade	Study Type	Subjects	Results
Hirshkowitz/(55)	1982	1	RCT	100+	Spindle frequency, amplitude, and duration could be quantified and differed in response to different drugs studied in a variety of randomized controlled dose-response clinical and experimental trials.
Smith/(56)	1976	2	RCT	10 subjects, 18 consec nights	Results suggest that sleep EEG waveform descriptors are sensitive indicators of drug activity and that beta activity in particular may be useful in the detailed description of various drug effects.
Hirshkowitz/(57)	1979	2	RCT	100+	Waveform analysis performed for a series of randomized controlled drug trials.
Kajimura/(58)	1995	3	RCT	6 adult male schizo-phrenic patients	All-night sleep stage scoring was carried out by visual analysis, and computerized period-amplitude analysis of sleep EEG was also performed. Computer analysis was more sensitive than hand scoring, finding significant differences between waveforms parameters where not sleep stage differences were found.
Johnson /(59)	1976	3	RCT	5 subjects for 7 nights each. 4 female, 1 male, young adults	Compressed spectral analysis yielded a computer-generated somnogram on each of the 15 nights of sleep, and an automated spindle detector was used to count and measure the duration of spindle bursts with frequencies of 12.25-15.5 c/sec on baseline nights 3 and 4, drug nights 1, 2, 3, and 7, on the 3rd withdrawal night, and on the 4-6 week follow-up record.
Astrom/(128)	1992	5	Case Series	9 patients	Power spectrum analysis showed that when circulating GH was elevated the energy in the REM sleep and delta sleep (stage 3 + 4) were higher.
Borbely/(129)	1991	5	Review	n/a	The effects of hypnotics on descriptive and functional aspects of electrophysiological sleep parameters are assessed in this report. computer-aided methods of EEG analysis have become increasingly important for recording and interpreting pharmacological effects on sleep. Of particular interest are the changes of EEG slow wave activity.
Rosadini/(130)	1992	5	Demonstration	n/a	Fourier time course spectral descriptors are used to study the cyclic patterns of sleep and to compute the DSRI (Delta Sleep Regularity Index), a synthetic measure of sleep quality. Compared to convention analysis, topographical distribution of spectral parameters allows detection of regional variations of EEG in specific pharmacological conditions.

Table 7—Evidence table related to using waveform analysis to investigate psychiatric, medical & neurological disorders.

Focus	Author/ref	Year	Grade	Study Type	Subjects	Results
PSY	Kajimura/(58)	1995	3	Clinical series	6 adult male schizophrenic patients	See table 5
MDD	McPartland/(60)	1979	3	Clinical series	23 Patients with MDD	Eye movements could be quantified automatically, changes in REM latency & REM density revealed
MDD	Kupfer/(62)	1986	2	Cohort	Young & middle age pts with MDD	Distinct findings from automated analyses were noted in the distribution of REM and delta sleep throughout the night.
MDD	Kupfer/(63)	1984b	2	Cohort	Pt with MDD and Controls	Declining SWA across night (Process S), especially in controls and continued ultradian and circadian pattern (Process C).
MDD	Kupfer/(64)	1984a	2	Cohort	Pts with MDD and Controls	Decreased slow wave activity and increased EM density early in the night.
MDD	Reynolds/(65)	1990	1	Cohort	302 pts with MDD	REM sleep latency,1st REM period duration, sleep efficiency, and early morning awakening showed robust age effects, but no main effects for sex or sex-by-age interactions. Sex effects on slow wave sleep and delta wave counts in depression parallel sex effects seen in healthy aging.
MDD	Roschke/(66)	1995	3	Clinical series with controls	15 MDD and 13 healthy controls, matched for sex, age, & education	Lyapunov exponents L1 of EEG segments corresponding to sleep stages 1, 2, 3, 4, and REM, according to Rechtschaffen and Kales, for the lead positions CZ and PZ. Significant decreased values of L1 during sleep stage 4 in depressives compared with a healthy control group.
MDD	Roschke/(67)	2002	4	Clinical series with controls	13 unmedicated MDD and 13 controls	Clear difference between REM and NREM sleep cycles at certain frequency bands. The most impressive changes occurred for the delta/beta and theta/beta correlations, which change their signs between NREM (negatively correlated) and REM (positively correlated) sleep cycles. Following an analysis of variance model with repeated measurement design, a statistically significant group effect (P=0.024) between depressives and controls. was observable during NREM sleep for the delta/beta (P=0.010) and theta/beta (P=0.018) interactions.
MDD	Roschke/(68)	1994	3	Clinical series with controls	9 MDD 11 schizophrenic 13 healthy controls	Results point to altered nonlinear brain dynamics mainly during slow wave sleep in depression and during REM sleep in schizophrenia.
PSY	Keshavan/(133)	1995	2	Clinical series	38 male and 23 female patients with functional psychoses	No sex effects were seen for any sleep parameters. However, older psychotic males had less slow wave sleep than older psychotic females.
MED	Bendtson/(134)	1992	5	Case series	8 patients with DM	Scoring was based on the color density spectral array of the EEG. Blood glucose values below 2.0 mmol/l were observed in some of the patients accompanied by EEG changes with increased theta and delta activity.

Table 8—Evidence table related to CAP detection

Author/ref	Year	Grade	Study Type	Subjects	Results
Ferrillo/(71)	1997	4	Case series	10 healthy adults, age 20-30 years	Using a zero-crossing technique, the number of total power oscillations were evaluated. Data suggest that slow rhythmic oscillations expressed by CAP can be detected by means of spectral analysis and their dynamics appear to be related to the EEG synchronization processes.
Ferri/(72)	2005	4	Case series	11 normal records scored by 4 raters	CAP scoring has good inter-rater reliability. The computer-assisted CAP scoring works well for general parameters such as CAP rate but more editing is necessary for more specific parameters.
Rosa/(73)	1999	5	Case series	4 middle-aged adults	High agreement between the detector and visual scoring was found fro fully automated scoring; however, more exhaustive evaluation is needed.

Table 9—Evidence table related to using waveform analysis to investigate sleep mechanisms and disorders.

Focus	Author/ref	Year	Grade	Study Type	Subjects	Results
MED	Feinberg/(75)	1991	4	Clinical series	9 young men	Visually scored delta and computer-measured 0-3-Hz EEGs increase above the baseline levels only if there has been loss of stage 3/4 EEG (or of sleep) from the first 2 NREMs.
MED	Feinberg/(76)	1988	5	Case series	?	The neurophysiological implications of a response of EEG amplitude as opposed to wave density are briefly considered; separate measurement of these variables is more readily accomplished with period-amplitude than with spectral analysis.
MED	Dijk/(78)	1989	3	Clinical series	28 patients	Visual scoring of EEGs revealed no significant differences between the sexes in the amounts of SWS and REM sleep. Spectral analysis, however, detected significantly higher power densities during NREM sleep over a wide frequency range (0.25-11.0 Hz) in the female versus male subjects. Also, during REM sleep, power densities were higher in the females.
MED	Schabus/(79)	2004	3	Random sample	24 healthy subjects	Participants performed a cued recall in the evening after learning (160 word pairs) as well as in the subsequent morning after 8 hours of undisturbed sleep with full polysomnography. Overnight change in the number of recalled words, but not absolute memory performance, correlated significantly with increased spindle activity. Time spent in each sleep stage could not account for this relationship.
SRBD	Bennett/(80)	1998	2	Cohort	41 patients (36 men, 5 women)	Patients ranged from nonsnorers to severe OSA. All had polysomnography with microarousal scoring, computerized EEG analysis, autonomic arousal detection, and body movement analysis. Multiple regression analysis showed the best predictor of nCPAP related improvement in objective sleepiness was body movement index (explaining 43% of the variance). Variability in EEG sleep depth, quantified from computerized EEG analysis, was the only other predictor.
SRBD	Ondze/(82)	2003	2	Clinical sample	18 patients with OSA and 18 controls.	Spectral analysis revealed slow wave and theta band activity. Slow wave and theta activity decreased across the night in both groups. Theta and sigma activity was lower in patients than controls. The sleep fragmentation produced by the breathing disorder is thought to reduce spindling.
SRBD	Morisson/(83)	2001	3	Clinical sample	14 patients with OSA and 10 controls	Untreated OSA patients showed EEG slowing in frontal and central cortical regions during both wakefulness and during REM sleep compared to healthy control subjects. This EEG slowing was found to be independent of time spent with arterial oxygen saturation <90% or severity of OSA. CPAP treatment was found to correct the EEG slowing for both REM sleep and wakefulness.
INS	Perlis/(85)	2001	3	Clinical sample	9 pts with primary insomnia, 9 pt with MDD-related insomnia, 9 controls	High frequency activity is inversely correlated with delta activity. These data are consistent with the hypothesis that high frequency activity is related to CNS arousal to the extent that greatest Beta/Gamma activity occurs during light sleep and in patients with primary insomnia.
INS	Perlis/(86)	2001	3	Clinical sample	9 pts with primary insomnia, 9 pt with MDD-related insomnia, 9 controls	Subjects with Primary Insomnia exhibited more average NREM activity for Beta-1 (14-20Hz), Beta-2 (20-35Hz) and Gamma activity (35-45Hz) than the other two groups (p.<.01). Group differences were also suggestive for Omega activity (45.0-125Hz) (p<.10), with MDD subjects tending to exhibit more activity than the other groups.

MED	Wichniak/(135)	2003	2	Clinical sample	19 subjects, 2 nights each	Power spectra were compared between subjects with MSLT sleep latencies < 10 and ≥10 minutes and ESS scores < 6 and ≥ 6. Subjects with short MSLT sleep latencies showed a reduced theta EEG activity. There was no evidence of reduced synchronization of sleep EEG in subjects with high ESS scores.
MED	Armitage/(136)	2001	5	Review	24 healthy subjects	There is an inherent increased vulnerability to depression in women that arises out of basic sex differences in brain organization and state regulation, particularly in response to a "biological challenge" during sleep. It is argued that the inherent properties of organization and regulation of sleep EEG in healthy men and women, elicited under challenge conditions, show sex-specific vulnerability to organizational abnormalities that model homeostatic abnormalities in depressed men and women and contribute to the genesis of depression. Sophisticated EEG techniques can reveal these differences
SRBD	Guilleminault/(137)	2001	2	Clinical sample	12 pts with UARS, 12 pts with OSA, and 12 controls	EEG spectral analysis was used to sleepiness. OSA patients had less slow wave activity than patients with UARS. UARS patients had a significantly higher absolute power in the 7-9 Hz bandwidth than OSAS patients. The absolute delta power over the different sleep cycles was also different between controls and patients, and between UARS and OSAS patients.
SRBD	Heinzer/(138)	2001	3	Clinical sample	10 patients with OSA and 10 controls	Slow wave activity (SWA) was defined as the power in the 0.75- to 4.5-Hz frequency band. A positive correlation between SWA of the first cycle and the MSLT ($r = 0.56$; $p = 0.045$) was found before treatment. SWA in the first 2 NREM/REM cycles improved in OSA patients after treatment.
SRBD	Black/(139)	2000	4	Clinical sample	15 patients with UARS	Esophageal pressure usually drops when respiratory arousals occur. However, sometimes a drop occurs when there is no EEG alpha arousal. Spectral analysis reveals increase alpha, sigma, and beta activity associated with pressure change even when an EEG arousal is absent.
PAR	Espa/(140)	2000	3	Clinical series with control	11 sleepwalkers and 11 controls	The slow wave activity (SWA) absolute values averaged during the 2 min immediately preceding an episode of parasomnia were significantly higher than the SWA averaged during 2 min in the same stage 10 min before an episode of parasomnia. Moreover, SWA was higher in the slow wave sleep episodes.

Key: CAP= cyclic alternating pattern; CPAP=continuous positive airway pressure; ESS=Epworth Sleepiness Scale; FFT=fast Fourier transformation; FM=fibromyalgia; GH=growth hormone; HF=high frequency; MDD=major depressive disorders; MED=medical; MSLT=multiple sleep latency test; OSA=obstructive sleep apnea; PAR=parasomnias; PSY=psychiatric; RCT=randomized controlled trial; SWA=slow wave activity; SWS=slow wave sleep; SRBD=sleep-related breathing disorder; SZ=schizophrenic; UARS=upper airway resistance syndrome