Erratum
September 28, 2018. DOI: 10.1159/000493518

Review

Neuropsychobiology

Neuropsychobiology 2013;67:127–167
DOI: 10.1159/000343449

Received: August 7, 2012
Accepted after revision: November 26, 2012
Published online: March 16, 2013

Guidelines for the Recording and Evaluation of Pharmaco-Sleep Studies in Man: The International Pharmaco-EEG Society (IPEG)

Marc Jobert[a] Frederick J. Wilson[b] Thomas Roth[c] Gé S.F. Ruigt[d] Peter Anderer[e]
Wilhelmus H.I.M. Drinkenburg[f] The IPEG Pharmaco-EEG Guidelines Committee

[a]International Pharmaco-EEG Society, Berlin, Germany; [b]Pharmatherapeutics Precision Medicine, Pfizer Ltd, Sandwich, UK; [c]Sleep Disorders and Research Center, Henry Ford Hospital, Detroit, Mich., USA; [d]Clinical Consultancy for Neuroscience Drug Development, Oss, The Netherlands; [e]Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria; [f]Neuroscience Discovery, Janssen Research and Development, Pharmaceutical Companies of Johnson and Johnson, Beerse, Belgium

Key Words
Actigraphy · Biomarkers · Central nervous system · Drug development · Electroencephalography · First-night effect · Guidelines · Pharmacodynamics · Pharmacokinetics · Pharmaco-sleep · Polysomnography · Source localisation · Translatable electroencephalographic biomarkers · Translational electroencephalographic endpoints

Abstract
The International Pharmaco-EEG Society (IPEG) presents guidelines summarising the requirements for the recording and computerised evaluation of pharmaco-sleep data in man. Over the past years, technical and data-processing methods have advanced steadily, thus enhancing data quality and expanding the palette of sleep assessment tools that can be used to investigate the activity of drugs on the central nervous system (CNS), determine the time course of effects and pharmacodynamic properties of novel therapeutics, hence enabling the study of the pharmacokinetic/pharmacodynamic relationship, and evaluate the CNS penetration or toxicity of compounds. However, despite the presence of robust guidelines on the scoring of polysomnography recordings, a review of the literature reveals inconsistent aspects in the operating procedures from one study to another. While this fact does not invalidate results, the lack of standardisation constitutes a regrettable shortcoming, es-

Co-authors (in alphabetical order):
Frederik W. Bes: Sleep Research and Clinical Chronobiology, Institute of Physiology, Charité – Universitätsmedizin Berlin, Germany; ***Martin Brunovsky:*** Psychiatric Center and Charles University, Prague, Czech Republic; ***Heidi Danker-Hopfe:*** Competence Center for Sleep Medicine, Charité – Universitätsmedizin Berlin, Germany; ***Jon Freeman:*** Brooklyn Strategic Consulting, N.Y., USA; ***Joop M.A. van Gerven:*** Centre for Human Drug Research, Leiden, The Netherlands; ***Georg Gruber:*** The Siesta Group, Vienna, Austria; ***Bob Kemp:*** MCH Hospital – Sleep Centre, Den Haag, The Netherlands; ***Gerhard Klösch:*** Department of Neurology, Medical University of Vienna, Austria; ***Junshui Ma:*** Merck Research Laboratories, Rahway, N.J., USA; ***Thomas Penzel:*** Sleep Medicine Center, Charité – Universitätsmedizin Berlin, Germany; ***Barry T. Peterson:*** Clinical Affairs, Philips Respironics, Bend, Oreg., USA; ***Hartmut Schulz:*** Department of Educational Science and Psychology, Free University, Berlin, Germany; ***Luc Staner:*** Centre Hospitalier de Rouffach, Rouffach, France; ***Bernd Saletu:*** Department of Psychiatry and Psychotherapy, Medical University of Vienna, Austria; ***Vladimir Svetnik:*** Merck Research Laboratories, Rahway, N.J., USA.

KARGER

E-Mail karger@karger.com
www.karger.com/nps

0302–282X/13/0673–0127$38.00/0

Dr. Marc Jobert
E-Mail publications@IPEG-society.org

pecially in the context of drug development programmes. The present guidelines are intended to assist investigators, who are using pharmaco-sleep measures in clinical research, in an effort to provide clear and concise recommendations and thereby to standardise methodology and facilitate comparability of data across laboratories.

Introduction

Pharmaco-sleep research concerns the description and the quantitative analysis of the effects of drugs on the central nervous system (CNS) by means of (neuro)physiological methods applied to subjects during a sleep period within the framework of clinical and experimental pharmacology, neurotoxicology, drug research, and associated disciplines. Such research can have one of a wide range of objectives, from providing evidence of efficacy to support registration of a new drug, such as a hypnotic, or evaluating CNS side effects of a new drug, to providing basic pharmacodynamic (PD) data at an early stage in the development of a new chemical entity. For the remainder of this article, the term 'pharmaco-sleep' strictly refers to human quantitative polysomnography (PSG) in the context of drug testing. Separate guidelines for the recording and evaluation of pharmaco-electroencephalography (pharmaco-EEG) data in man have recently been published by the International Pharmaco-EEG Society (IPEG) [1]. Separate guidelines for pharmacological studies in animals are in preparation for publication by the IPEG.

In the 80s and early 90s, several guidelines were published with the goal to standardise the acquisition and processing of data collected in pharmaco-EEG studies [2, 3] or to provide methodological recommendations for the recording and quantitative analysis of EEG activity in research contexts [4]. In parallel, several organisations published recommendations and guidelines for the use of EEG in various clinical fields [5–7] in an effort to improve standardisation and facilitate the proper utilisation of the technique in clinical practice.

For many years, the evaluation of PSG recordings has been conducted according to the rules presented in 1968 by a panel led by Rechtschaffen and Kales [8] in their publication entitled *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*, and this manual (known as R&K) provided the generally accepted method for the scoring of the adult human PSG tracings in clinical and research settings. Despite criticism expressed by several authors [9–11], the scoring criteria have been used as the 'gold standard' until May 2007 when *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification* was published by the American Academy of Sleep Medicine (AASM) [12], with the goal to replace the scoring rules published by Rechtschaffen and Kales in 1968.

However, neither the R&K nor the AASM manuals have focused on application to pharmaco-sleep studies. Hence, although these guidance documents have enabled a good level of standardisation in the evaluation of sleep recordings and many operational aspects, other inconsistencies remain in the conduct of pharmaco-sleep studies, including methodological differences (most importantly the number of consecutive nights studied) and the application of inclusion/exclusion criteria [13]. The current guidelines are intended to augment the AASM manual by providing additional guidance relating specifically to pharmaco-sleep studies and their specific challenges. To improve consistency, investigators using pharmaco-sleep methodology are urged to follow and reference these guidelines when designing and conducting studies, and in particular when reporting the methods used and results obtained.

Clinical trials must be conducted in compliance with good clinical practice (GCP), an international quality standard launched by the International Conference on Harmonisation (ICH), an international body defining standards that governments can transpose into regulations for the conduct of clinical trials involving human subjects and more specifically in relation to research for the registration of pharmaceuticals. GCP requires standard operating procedures (SOPs) to be defined and applied to all methods and procedures during drug development research, and while the scoring manuals (R&K or AASM) provide sound references for sleep scoring, there are other important aspects that must be taken into consideration. The purpose of the present guidelines is to provide clear and concise recommendations for pharmaco-sleep studies and thereby to standardise methodology, including study designs, and facilitate comparability of data across laboratories.

The EEG is a non-invasive method which reflects the spontaneous synchronised postsynaptic neuronal activity of the human cortex with high temporal resolution. While EEG parameters, including sleep measures, are among the CNS activity biomarkers with the highest heritability to the extent that they even constitute an individual fingerprint [14–18], they show at the same time a

very high sensitivity to changes in internal (state-modulated traits) as well as environmental factors. The sensitivity to such factors, which are extraneous to the objectives of many studies, means that a high degree of quality control and detailed standard operating procedures are required in order to decrease the effect of confounders in the recording and analysis of the data.

While pharmaco-sleep research has demonstrated its value in the development of CNS-active compounds in many instances, and while validated quantitative methods have been available for a long time to study the effects of drugs on brain functions in patients and volunteers [19–21], there is still reluctance to apply this method in large-scale clinical trials or for decision-making drug studies outside the area of sleep medications. There are a number of reasons contributing to this situation:

(1) While there is evidence indicating the putative utility and validity of EEG and sleep as biomarkers relevant to a range of drug classes covering several therapeutic indications, they have not yet been generally accepted as such, particularly in the case of non-hypnotics where their role is less easy to define. Further, the translatability of pharmaco-sleep signatures from animal to man is not universal across the spectrum of CNS-active drugs, but depends on the pharmacological mechanism and the preclinical species used. Hence, its use as a translatable biomarker for the preclinical screening of compounds and the development of new drugs requires careful interpretation [22].

(2) Despite the fact that the effects of drugs on sleep have been investigated in research laboratories for several decades now [23–28], operating procedures have not yet been standardised to an extent facilitating a reliable comparison of datasets and results across units, making it difficult (or even impossible) to share datasets between sites or to pool results from different clinical trials.

(3) This lack of standardisation constitutes a difficult obstacle for the design and interpretation of clinical trials due to the difficulties in comparing results across the literature.

(4) For a long time, the amount of data generated by recording PSG signals across multiple nights in different treatment periods quickly overwhelmed the storage capacity of computers, meaning that only derived data was stored rather than the raw EEG recording, and the processing techniques were constrained by central processing unit power, particularly when undertaking spectral analysis of the EEG from multiple electrodes. These limitations have now disappeared as a consequence of vastly increased computing power and storage capacities.

In this context, one of the crucial steps is to enhance the standardisation of the operating procedures, not only to improve the ability to compare datasets and results generated in different laboratories by reducing variance, but also to facilitate the creation of a centralised data repository where large numbers of records can be stored and shared. Such a repository would enable the following endeavours:

(1) constitute reference datasets (i.e. both the raw PSG signals and derived parameters) obtained under standardised environmental conditions from different studies using various drugs (with emphasis on reference drugs and including placebo) and study populations (healthy volunteers and various patient populations) under standardised behavioural conditions, enabling comparative analyses and development of novel signal-processing techniques (e.g. automated sleep scoring);

(2) identify PSG parameters and properties that could be exploited as potential (translatable) biomarkers and quantify their validity in large populations;

(3) confirm the utility and reproducibility of PSG as a sensitive assay providing PD data suitable for determining pharmacokinetic/pharmacodynamic (PK/PD) relationships (exposure/response curves) and dose response in both animals and humans;

(4) facilitate the transition of novel compounds from preclinical to clinical R&D programmes by enabling the early comparison of results obtained in preclinical screening and early clinical experiments, thereby improving the decision-making process as well as derisking and accelerating the development of new CNS-active compounds.

Pharmaco-Sleep Studies

Fundamentals

It is well established that control of physiological functioning during wake and sleep differ. In fact, within sleep there are differences in functioning during rapid eye movement (REM) and non-REM (NREM) sleep and even within REM sleep there are differences between phasic and tonic REM. For example, hypercapnic drive is blunted in NREM sleep relative to wake and is further blunted in phasic REM sleep. Importantly, the basis of sleep disorder medicine is the realisation that a given physiological function can be normal while awake and be pathological when

the person is asleep. Sleep apnoea and REM behaviour disorders are examples of this. Thus, it becomes critical to be able to evaluate physiological functioning across the sleep state. The PSG recording of sleep provides the researcher as well as the clinician with a unique insight into the nature of functioning during one third of a human's existence. The basis of PSG is the simultaneous recording of EEG, electromyography (EMG) and electrooculography (EOG) signals. In addition, a multitude of assays ranging from brain metabolism to endocrine function can be recorded simultaneously to answer specific questions. For example, clinicians routinely record a variety of cardiorespiratory parameters as well as several EMG locations to detect a variety of sleep disorders. Insomnia researchers often record autonomic nervous system parameters. These recordings have the potential to provide insights into both the efficacy as well as the mechanism of action of sleep-promoting drugs or drugs that interact in other ways with sleep (e.g. vigilance-enhancing drugs) or the sleep-wake cycle (e.g. chronobiotics such as melatoninergic substances). In some such cases, sleep measures can be used as biomarkers of pharmacological action even when the primary drug indication cannot be specifically related to the sleep effects (e.g. REM sleep suppression with antidepressants [20]). Importantly, PSG is not limited to defining sleep initiation simply by looking at sleep latency and sleep maintenance with wake time after sleep onset, but new parameters are being routinely used. For example, frequency and duration of sleep and wake bouts have routinely been used in animal research and are now being applied to the analysis of human sleep. Spectral analysis is another valuable tool in understanding sleep EEG data, especially the amount and regulation of slow-wave activity. These and other types of measures, like transient arousals, are enabling investigators to capture more of the richness of PSG data [29].

An essential part of understanding sleep quality is the comparison of PSG measures to those reported by subjects in post-sleep questionnaires. These comparisons tell us how patients' *perceptions* of how they sleep relate to how they *actually* sleep. For example, insomnia patients routinely overestimate the severity of their sleep problem as determined by PSG recordings. The pattern of these discrepancies is important to appreciate in drug development. For example, with benzodiazepine receptor agonists, patients may report greater efficacy than that suggested by PSG data [30]. In contrast, with other drug classes, such as serotonin antagonists and melatonin agonists, PSG results are more robust than patient reports [31]. For this reason, as well as other considerations, some regulatory bodies require both types of data for the evaluation of sleep-promoting agents. Generally PSG data are critical for the evaluation of drugs. PSG is uniquely able to show objectively measured dose-related changes in hypnotic activity that are independent of subject reports, both in terms of efficacy parameters as well as sleep stage distribution.

Currently, several new technologies are being developed and used in sleep research. Examples include actigraphy [32], which is discussed in detail in a separate section below, and various home-based sleep-monitoring systems. Some of these, such as that described by Shambroom et al. [33], are PSG systems modified to be suitable for use in the home, which retain EEG as the core biosignal. When considering the utility of such devices in pharmaco-sleep studies, the most important factor is therefore control of the environmental conditions, which is discussed further below. Other home sleep testing systems, however, use a reduced set of recording channels without EEG. Such systems have been proposed as cost-effective, patient-friendly and scientifically valid in the diagnosis of sleep-disordered breathing [34], although some controversy remains [35, 36]. Indeed, it has recently been proposed that obstructive sleep apnoea can be detected from the electrocardiography (ECG) signal alone [37]. In general, such reduced-channel systems are not suitable for use in pharmaco-sleep studies as their application is too narrowly focused on the respiratory aspects of sleep, and hence they do not enable a full examination of drug effects.

As part of the development of new techniques, there is invariably a validation study against PSG. While such validation studies are important, they somewhat miss the point. There is no replacement for the control and richness of laboratory-based PSG, while the laboratory condition equally cannot replace the home situation with respect to a number of the (often variable) environmental factors that affect a person's sleep. The real issue is what types of information these alternative methodologies can provide that will complement PSG data. If total sleep time across multiple weeks is an important piece of information, home actigraphy recordings are better suited for this than PSG recordings. However, if a more detailed analysis is needed for selected days, PSG is the ideal technique for that purpose. In general PSG data are the most comprehensive and reliable data of sleep. These guidelines are intended to help investigators and clinicians to maximise the use of this important tool.

Finally, all clinical trial programmes must follow the ICH guidelines aimed at 'ensuring that good quality, safe and effective medicines are developed and registered in

the most efficient and cost-effective manner. These activities are pursued in the interest of the consumer and public health, to prevent unnecessary duplication of clinical trials in humans and to minimize the use of animal testing without compromising the regulatory obligations of safety and effectiveness'. The ICH guidelines cover a broad range of activities related to drug development [38, 39].

Subjects

PSG studies have an important role in many research areas. The choice of subjects for a given study depends on the goal of that study. For example, in drug development, if the goal of the research is to define the pharmacological activity of the drug in terms of sleep stage distribution, degree and duration of sedative activity, or physiological activity during sleep, then normal healthy volunteers are the appropriate subject population. On the other hand, if the goal of the study is to determine efficacy, then the appropriate patient group needs to be studied. Studying both groups allows us to understand drug effects more fully. For example, a drug may have a pharmacological activity to suppress slow-wave sleep. However, if the subjects are patients with minimal slow-wave activity, this drug effect may be missed. Conversely, if a drug has the ability to help induce sleep but healthy subjects with normal sleep latency are studied, this effect may be missed.

Regardless of the subject population being studied, there are certain exclusion criteria which need to be kept in mind in all PSG studies. First, most medical and psychiatric disorders are associated with disturbed sleep. This can be evidenced by difficulty initiating and maintaining sleep in most psychiatric disorders and fragmented sleep seen in many medical disorders [40]. In addition, these clinical populations often take medications and it must be recognised that virtually any drug that penetrates the CNS has the potential to alter sleep. For example, respiratory stimulants are given for lung disease, but they fragment sleep. Similarly, first-generation H_1 antagonists are given for allergies, but they also promote sleep. Aside from prescription and over-the-counter medications, and herbal preparations, attention must be paid to the amount and timing of alcohol and caffeine consumption. Consumption of large amounts of these compounds precludes subjects from participation in studies as discontinuation can lead to rebound sleep disturbance in the case of alcohol or rebound sleepiness in the case of caffeine. Thus, only moderate users should be included, for instance by excluding subjects if they meet any of the following conditions within the previous 6 months:

- history of regular alcohol consumption exceeding 2–3 units/day for females and 3–4 units/day for males [41];
- history of regular use of tobacco- or nicotine-containing products exceeding the equivalent of 5 cigarettes/day;
- history of regular consumption of caffeine exceeding the equivalent of 4 cups of coffee/day, a level that approximates health-related criteria [42].

In addition, subjects should refrain from alcohol and caffeine for at least 24 h, and from tobacco or nicotine products for at least 4 and preferably 8 h, prior to a PSG recording. In healthy volunteer studies subjects using illicit drugs of abuse (including recreational use) should be excluded. Stricter exclusion criteria or restrictions (with respect to age, sex, body mass index, handedness, physical/mental/medical status/history including cardiac and laboratory parameters, visual/auditory function, ability to communicate, etc.) may be applied on a study-by-study basis where required.

Another critical variable to consider in selecting healthy volunteers for a pharmaco-sleep study is the regularity, duration and timing of sleep. Healthy adult volunteers should routinely spend 6.5–8.5 h in bed each night as less than 6.5 h can be associated with chronic sleep deprivation and/or restriction and more than 8.5 h can be indicative of a disorder. However, it may be necessary to alter this criterion in the case of adolescents or the elderly to account for age-related differences. Aside from duration, subjects who maintain regular bed times should be selected, with a suitable inclusion criterion being variations of less than 2 h over a 2-week period. Finally, subjects should sleep at traditional hours. Typically subjects are required to go to sleep routinely between 22.00 and 00.00 h (midnight), although this criterion may be varied according to the subject population and cultural differences and it may also, in some cases, be necessary to set different criteria for weekdays and weekends. These requirements should be controlled either by the use of sleep diaries and/or actigraphy over a period of at least 1 week during screening, prior to enrolment in the PSG study. Finally, night and shift workers, and individuals who need to fly across multiple time zones should also be excluded. Clearly, in the case of patients with sleep disorders, these criteria will need to be relaxed appropriately for the condition.

During the week preceding a PSG recording session, the subjects should be requested to keep to their usual circadian routine and to adhere to a regular sleep/wake pattern that matches the inclusion criteria set for the study, as discussed above. Sleep diaries should be used to

document compliance. Where adherence to a regular sleep/wake pattern is a particular concern, the use of actigraphy recordings as objective compliance verification should be considered.

Phenotyping and/or genotyping of the participants is acquiring increasing importance for safety and PK reasons; since drugs can be primarily metabolised by specific cytochrome P450 iso-enzymes, the metabolic status could be defined accordingly in order to avoid potential accumulation/fast elimination, during the washout phase, for example. If a drug is known to be subject to major genetic polymorphism, studies could be performed in panels of participants of known phenotype or genotype for the polymorphism in question.

The choice of subjects is a critical issue in all PSG studies. Investigators need to balance studying the most stable homogeneous subjects, so as to minimise variability, with the need to make the population broad enough so that the results can be generalised to the general population.

Environmental Conditions

There are many environmental factors affecting the function and activity of the central nervous system and, as a result, also affecting neurophysiological readouts of the wake/sleep process. It is therefore necessary to control these factors to the best possible extent. If deviations from normal, pre-existing or predefined conditions are observed, then these should be recorded as meta-data. In clinical trials, it is mandatory to predefine criteria by which data may be excluded from the analysis in the event of significant deviations, and to document the effect of applying these. In some cases altered environmental conditions can be used as models of a particular disorder. For example, noise, sleeping whilst sitting up or simply the first-night effect can be used as models of transient insomnia.

Recording Environment

To lower the environmental impact on wake/sleep propensity mechanisms and minimise the variability between and within subjects, PSG measurements should be recorded under strictly controlled conditions in a laboratory setting, thereby restraining the influence of most external and internal confounding factors.

The recording should occur in a separate, comfortable, darkened, sound-attenuated room with regulated temperature and humidity within the normal range for the geographical location (typically 18–22°C/65–72°F). Sleep-recording sessions should be started at a time that accommodates individual habitual sleep times (typically between 22.00 and 00.00 h, depending on the inclusion criteria) and last for 8 h from lights-out or until terminal awakening if sooner. The timing of drug administration should be fixed relative to the timing of lights-out. During the evening hours preceding recording, the subjects should refrain from excessive physical exercise or activities causing a high degree of mental stimulation such as vigorous video gaming. Also the timing and type of food for dinner must be carefully selected: While confined, the daily caloric intake per subject should not exceed approximately 3,200 kcal and the nutritional composition should be approximately 50% carbohydrate, 35% fat and 15% protein. Consumption of caffeine has to be controlled during the course of the study. In the case of studies with repeated measurements, it is important to ensure that environmental and confounding factors are consistent from one session to the next.

Ambulatory recordings in the home environment are technically feasible and have the advantage in disease studies (e.g. insomnia) that the effect of the drug can be evaluated in the surroundings in which the disorder is manifesting itself, taking into account that part of the disorder may be dependent on the home environment. However, such a setting requires specific planning and instructions to the subjects to control the study course knowing that the environmental conditions are uncontrolled per se. Due to the technical complexities of carrying out PSG recordings at home, actigraphy may be a viable alternative in some cases (see section 'Sleep Assessment Using Actigraphy').

Adaptation

The first-night effect (FNE) is a well-known phenomenon in sleep research. The major factors responsible for this effect are the unfamiliar surroundings of the sleep laboratory and the subjects' lack of adaptation to the discomfort induced by electrodes, cables and instruments applied to their body [43, 44]. Other factors influencing the magnitude of the FNE have been discussed and include the psychological effect of being the object of study [43–46] or personality traits, such as trait anxiety [47–49] or even a placebo effect [50]. The FNE can be quantitatively measured even when the subjects do not report subjective sleep quality degradation [47, 51].

Generally, the first night is consistently characterised by increased REM sleep latency and a reduction in the amount of time spent in REM, although increased sleep onset latency, shorter total sleep time and hence lower

overall sleep efficiency are often also observed [43–45, 47, 52]. In addition, the partial REM sleep deprivation in the first night may likely have consequences for the second night (REM rebound) [53]. The FNE phenomenon has been replicated in many studies with healthy participants [43, 53, 54], psychiatric patients [44, 52, 55, 56], epileptics [57], juvenile rheumatoid arthritics [58] and in subjects with chronic fatigue syndrome [59], although the magnitude of the effects varies between groups. In children and adolescents with sleep-disordered breathing, FNEs were observed in the sleep parameters but not in the respiratory parameters [60–62]. Consequently, it is important to be aware of this phenomenon when designing clinical sleep studies. The data obtained from the first night are, in most cases, discarded and not included in assessing a subject's sleep.

In the context of drug research, the results of statistical tests for significant differences between patients and controls, between pre- and postdrug conditions, or between drug-induced and placebo-induced changes are related to the observed variances. The pronounced FNE on group variances might thus lead to erroneous results. In order to obtain reliable and valid measures of sleep, the recording of an adaptation night is mandatory in most cases, independent of the subject's sex, age and diagnosis, and the data of the adaptation night should be discarded. An exception is when the FNE itself is of interest in studying the effect of the drug, for example when using FNE as a model of transient insomnia in healthy volunteers [63–65].

In the case in which subjects are required to sleep in the laboratory during several non-consecutive blocks of nights, typically for long-term observation with follow-up assessment or for a crossover study, the possibility of differences in (re-)adaptation effects between blocks must be taken into consideration. Current evidence suggests that the re-adaptation process is highly dependent on the nature of the subjects as some groups are more adaptive than others. In a study aimed at assessing the FNE for consecutive blocks of night recording in healthy young subjects [66], it was found that a relatively small FNE, which was detectable only in the REM-sleep-related variables, was present only on the first night of the whole study ('the very first night') and that the effect did not persist during the remaining first nights of the subsequent periods, even when the study was interrupted for a period of 1 month. Conversely, other studies have shown a significant FNE, particularly for REM-sleep-related variables, on the first nights of subsequent periods in young healthy volunteers [67], healthy older subjects [68] and insomniacs [69, 70], albeit that the magnitude of the effects was consistently found to be smaller in subsequent blocks than on 'the very first night'. Therefore, the decision as to whether an adaptation night is required in each block or only at the start of the study must be made on a case-by-case basis. Important factors to consider include the design of the study and statistical analysis plan (in particular, a symmetrical design is generally recommended in a crossover study, which necessitates the inclusion of an adaptation night in all periods or in none), the nature of the subjects, the endpoints being studied and the expected size of the drug effect.

While some ambulatory studies suggest that conducting home recording eliminates or reduces the FNE [71–73], others conducted with healthy participants [45], elderly individuals [74, 75] and patients [47] conclude that adaptation effects are still observed. Consequently, as home-based ambulatory PSG does not prevent a possible FNE, such study settings should always include an adaptation night.

Data Acquisition

Digital Recording

Digitising is the conversion of an analogue (continuous) signal into a digital (discrete) signal (i.e. a sequence of numbers). Modern analogue-to-digital converters (ADC) usually have a resolution of 16 bits, meaning that the analogue amplitude of each discrete point is rounded to the closest one of the available 65,536 (0 to $2^{16} - 1$) digital values. If the pre-amplifier gains are set to make these values cover a 2-mV range, then the nominal resolution is 5 bits/μV, meaning that the maximum round-off error is 0.015 μV. The sampling rate (T_S) corresponds to the time interval between two subsequent points and determines the resolution in time. The sampling frequency (F_S) expresses the number of samples digitised per second and is the reciprocal of the sampling rate ($F_S = 1/T_S$). For instance, with $F_S = 500$ Hz, the resolution in time corresponds to the sampling rate and is given by the reciprocal of F_S, i.e. 2 ms.

From a theoretical point of view, F_S must be at least twice the highest frequency present in the signal to be digitised (Nyquist-Shannon sampling theorem). Conversely, F_S may be set to at least twice the highest frequency interest, and then frequency components higher than $F_S/2$, also called Nyquist frequency, must be removed using analogue filters before digitising to avoid aliasing effects. Errors introduced in the digitised signal by aliasing

cannot be detected and corrected afterwards. Because of imperfections in the analogue filters, F_S is in practice at least fourfold the analogue filter (anti-aliasing) cut-off frequency. Analogue filters may also cause other problems due to distortions. This can be mitigated by sampling all biosignals at the highest possible frequency (e.g. 2,500 Hz) using a low-pass filter that rejects frequencies over 625 Hz. Thereafter, the signals can be downsampled to 500 Hz after applying a digital low-pass filter at 125 Hz or below.

From a practical point of view, the following applies to pharmaco-sleep studies: F_S must be at least 200 Hz (i.e. 200 samples/s). However, a F_S of 500 Hz or above is preferred to enable spectral analysis in the higher frequency range of the EEG, and is recommended by the AASM. The analogue-to-digital converter must have a digital resolution of at least 12 bits (16 bits is recommended) and have a round-off error below 0.2 μV (0.1 μV is recommended). Prior to sampling, an anti-aliasing low-pass filter (with a roll-off of at least 12 dB/octave) and a high-pass filter must be used. Ideally, any additional filters must only be applied post hoc. This enables the effect of the filtering step to be evaluated. In particular, the use of a notch-filter (50 or 60 Hz) should be avoided during recording as it can potentially disguise an electrode problem, while mains noise can be eliminated off-line at the data-processing stage. Due to the profound effect that different filtering procedures can have on the final results, it is imperative that full details of the process followed are reported alongside the results, so that meaningful comparisons can be made with other studies.

The electrode impedance (resistance) should conventionally be maintained below 5 kΩ. The pre-amplifier input impedance must be over 100 MΩ. Modern amplifiers with high internal resistances are able to record at higher scalp impedances, but it is still important to balance impedance across electrode sites. As the rejection of crosstalk between channels is important for coherence or other measures of relationship between electrodes, a crosstalk rejection of at least 90 dB is required and better is recommended.

To facilitate the export and import of PSG signals, several file formats have been suggested in the past. Of those formats, only European data format (EDF) has been, and is being, used successfully in many multi-centre research programmes and is nowadays supported by more than 50 hardware and software companies. Because of its simplicity, many research groups apply EDF in their proprietary analysis software. An enhanced EDF-compatible revision (called EDF+) has been released in 2003 [76]. In addition to handling PSG signals, the structure of the file format copes with evoked potentials, electrocardiography, annotations and actigraphy recordings.

Comprehensive overviews on technical aspects related to the digital recording of EEG signals have been published elsewhere and provide additional insight into specific details [77–80]. The guidelines for recording and evaluation of pharmaco-EEG studies in man recently published by the IPEG [1] also provide further information on the recording and processing of EEG signals.

Calibration

Recording accuracy (how far the sample varies from the 'true' signal value) is dependent upon system calibration. The calibration procedure is aimed at testing the performance of the entire hardware and must be carried out before each measurement. Calibration is also essential to achieve a reference potential of known voltage against the absolute amplitude of the recorded signals. To pick up possible time-dependent fluctuations of amplifiers, due to thermic effects for example, it is strongly recommended to recheck the calibration at the end of each measurement session.

Nowadays, most PSG machines have internal hardware calibration, and some will carry out a calibration check fully automatically. Verification is made that the same input signals (specifically sine waves with known amplitude and frequency) applied to all channels are present with the same amplitude at the output of the amplifiers and are subsequently correctly transmitted to the analogue-to-digital converter. If internal calibration is not available, then an external device should be used to generate stable test waves that are relayed through the electrode sockets.

Sleep Biosignals

REM and NREM sleep can be distinguished, and NREM sleep can be further subdivided into distinct stages according to the AASM guidelines [12], based on combinations of patterns in the EEG, chin EMG and EOG signals, which can be recorded using skin electrodes.

EEG

The number of electrodes used to record sleep EEG depends upon the nature of the scientific question under investigation, the electrodes being placed according to the international 10-20 system [81]. As a matter of principle, recording against one reference electrode is recommended to allow all conceivable montages (and of-

fline remontages). Such unipolar montages offer the advantage of post hoc rereferencing if any particular electrode becomes problematic or if a need for examining hemispheric asymmetries arises. In particular, while a 'linked-mastoids' reference offers the advantage of reducing common mode artefact in the EEG channels, this configuration may become problematic if the impedance of one of the reference electrodes varies differentially during a recording (for example if one of the electrodes becomes dislodged). Also, electrically linking the two brain regions can lead to distortion of the electric potential distribution. Thus, the 'linked-mastoids' reference should be avoided, especially when source localisation is applied later. The data must be stored in a format permitting conversion from the recording reference to any other reference (common average reference, current source derivation, other channels as reference, etc.).

According to the AASM guidelines, the recommended EEG derivations for sleep scoring include scalp EEG derivations from the frontal (F3/F4), central (C3/C4), and occipital (O1/O2) regions referenced to the contralateral mastoid (M1 or M2). The AASM guidelines (and item V.5 of the corresponding Scoring Manual FAQ posted on www.aasmnet.org) suggest derivations F4-M1, Fz-Cz, Fpz-E1 or C4-M1 (or the contralateral equivalents) for the assessment of NREM sleep depth. It is not clear which one of these derivations is to be preferred and it is likely that the choice matters because different derivations have different EEG amplitudes. Only the C4-M1 derivation is compatible with the old manual that was widely adopted in 1968 [8].

Should investigators want to examine scalp sites beyond the standard 21 locations identified by the 10-20 system (e.g. for topographic or localisation studies), then the international extended 10-20 electrode placement system (also known as the 10% system) should be utilised [82].

The most important EEG components during sleep differ from those during wakefulness and include sleep spindles (11–16 Hz), K complexes, slow waves (0.5–2 Hz) and sawtooth waves, but alpha and theta rhythms are also relevant [83]. They are described in the AASM scoring manual [12].

EOG

In sleep recordings, EOG is used to detect eye blinks, slow and rapid eye movements. The AASM scoring manual [12] suggests recordings from each outer canthus, both referred to the M2 mastoid, or bipolar derivations, measuring horizontal and vertical components separately. The latter montage has the advantage that EOG artefacts can be removed from the EEG using appropriate computerised algorithms, which is not possible with the referential montage. Nevertheless, one has to keep in mind that eye movements are not necessarily synchronised, and thus separate measurements for both eyes may be indicated [84].

Chin EMG

EMG activity is recorded with bipolar electrodes placed on the chin (mental or submental) and is required to detect REM sleep. As the decrement of the tonic EMG level during REM sleep may be very slight, the electrodes have to be applied carefully to shaved skin with 1 cm interelectrode distance. Three electrodes should be used – one above and two (one as a backup) below the inferior edge of the mandible – positioned to the right and left, respectively.

Electrocardiography (ECG)

At least one ECG lead should be recorded to derive heart rate for investigating changes during sleep (which indicate transient tachycardia or bradycardia) and detecting possible arrhythmias. However, while such a recording is useful to identify possible ECG artefacts in the EEG, it is not suitable for detecting PQRST complex abnormalities, and the conclusive diagnosis of cardiac pathologies (which have a high comorbidity with sleep apnoea) definitely requires more ECG channels.

Other Signals

Dependent upon the aim of the pharmaco-sleep study and/or the kind of disorder under investigation, further biosignals will be required to assess, for example, respiration, leg movements, snoring, blood pressure, indirect measurement of blood gases (oxygen saturation, transcutaneous CO_2), body temperature, body position, intercostal EMG, etc. [12, 85]. In most pharmaco-sleep studies a complete PSG recording, as specified in the AASM manual, will be required to enable a proper control of respiratory, motor or other disturbances, whether present at baseline (in the case of patients) or drug-induced. In some cases, continuous video recording may also be useful to enable physically observable events to be linked to observations of fluctuations in the other signals.

Minimum Requirements

The minimum requirements for pharmaco-sleep studies are reported in table 1, and compliance is required

Table 1. Minimum set of requirements for the recording of PSG signals in clinical trials

EEG	
Electrodes	• Sintered silver-silver chloride preferred; gold-plated cup electrodes may be used when frequencies below 0.5 Hz are not to be analysed • F4-M1, C4-M1 and O2-M1 (with backup electrodes placed at F3, C3, O1 and M2 for use if problems develop over the course of the recording with the right hemisphere selections) • Alternate montage: Fz-Cz or FPz-E1, Cz-Oz and C4-M1 (with backup electrodes at FPz, C3, O1 and M2) • For topographic and tomographic analysis, at least 21 electrodes placed according to the international 10-20 system (optional 48 placed according to the 10% system)
Sampling frequency (F_S) and frequency range	• F_S ≥500 Hz (recommended): – Anti-aliasing filter ≤125–150 Hz – High-pass filter ≤0.03 Hz – Frequency range 0.03–100 Hz • F_S <500 Hz (and ≥200 Hz): – Anti-aliasing filter 70 Hz – High-pass filter ≤0.03 Hz – Frequency range 0.03–40 Hz
Amplitude range	• ± 3,000 μV with 16-bit ADC (or ± 1,000 μV with 12-bit ADC)
Amplifier noise level	• Below 1 μV_{pp} (or <0.15 μV_{rms} at 5 kΩ resistance within 0.1–20 Hz)
Analogue notch filter	• To be avoided
EOG	
Number of channels	• Minimum: 2 Location: right outer canthus and left outer canthus Unipolar versus common reference • Alternate: both outer canthi referred to Fpz (bipolar) • Noise level: <2 μV_{pp}
Sampling frequency (F_S) and frequency range	• Same F_S as for EEG • For determination of eye position: DC
Amplitude range	• ± 2,000 μV
Amplifier noise level	• Below 2 μV_{pp}
Analogue notch filter	• To be avoided
EMG (sub)mental	
Number of channels	• Minimum: 1 (with 1 as backup) Location: mental or submental (2 cm interelectrode distance)
Sampling frequency (F_S) and frequency range	• Same F_S as for EEG • Frequency range: 10 Hz up to the analogue filter cut-off frequency
Amplitude range	• ± 3,000 μV with 16-bit ADC (or ± 200 μV with 12-bit ADC)
Amplifier noise level	• Below 0.1 μV_{pp} (or <0.02 μV_{rms} at 5 kΩ resistance within 10–150 Hz)
Analogue notch filter	• To be avoided
ECG	
Electrode site	• One lead modified from lead II to monitor the heartbeat during sleep, with an electrode placed in the right subclavicle area and left torso over the lower ribcage beneath the heart • Sampling rate same as for EEG – Anti-aliasing filter: as for EEG – High-pass filter: 0.5 Hz – Frequency range: 0.5–20 Hz

to ensure quality results and comparability between centres. The use of sintered silver-silver chloride electrodes is essential when EEG frequencies below 0.5 Hz are to be analysed [86]. Otherwise, gold-plated cup electrodes may be used.

While recording the PSG biosignals with F_S = 200 Hz is considered acceptable, it is recommended to operate with at least 500 Hz using the appropriate anti-aliasing analogue filter. In general, higher sampling frequencies enable the application of more signal processing and analysis methods [87].

Artefacts

One of the most crucial pitfalls of measuring brain electrical activity by means of EEG is its vulnerability to technical and biological artefacts. Artefacts in the EEG are defined as any interference caused by extracerebral sources. Artefacts can be subdivided into two categories: physiological and non-physiological artefacts [88].

Non-physiological artefacts typically relate to problems with electrodes (i.e. impedance and adherence), instrumental issues (i.e. ground loop, amplifier instability) or electrical power noise (i.e. 50 or 60 Hz mains noise). Each of these types of artefact will significantly distort the EEG signal to the point where signal accuracy is compromised.

Physiologic artefacts in sleep studies are caused by movements (head, eyes, body), cardiac pulses, sweating, and electrical activity of head muscles. The ECG artefact resulting from the QRS complex is problematic when ECG spikes are present in the EEG channels because the frequency spectra of EEG and ECG signals overlap (primarily in the range from 2 to 7 Hz). EOG artefacts are typically comprised of blinks, vertical and horizontal movements. Given the proximity to frontal EEG placements, there is a greater likelihood of eye movements leaking into the anterior EEG channels although eye movements can also be observed in central channels. This is particularly true for REM sleep and wakefulness where EOG activity is abundant.

Care should be taken to reduce these artefacts at the recording stage where possible and whenever quantitative results are analysed, careful attention should be paid to the quality of data acquisition, as problems in any one of these categories are sufficient to invalidate the end product analysis [87, 89]. In all cases, the artefact correction or reduction methods used should be clearly stated when reporting the results.

Visual Scoring

Equipment for Visual Scoring

While traditional pen-writing analogue (paper) systems are still acceptable to view PSG datasets, digital (computer monitor) display is the recommended method. However, it is important to keep in mind that visual display on a computer screen is also a digital process, images being drawn as discrete pixels. Screen resolution (vertical and horizontal) may be a limiting factor for a precise representation of the signals [80].

At the time of writing (2012), typical monitors have a resolution of 1,920 pixels horizontally by 1,080 pixels vertically and a diagonal size of 23 inches (58.5 cm) resulting in 0.266 mm/pixel. To display an epoch of 30-second duration implies that each second is allotted 64 pixels (17 mm). With F_S = 256 Hz, 4 recorded digital values must be represented per pixel, resulting in a compression (squeezing) and limiting the highest signal frequency that can be displayed. As a rule, the maximum frequency faithfully representable on a computer display is half the number of pixels per second. On the typical monitor described, this is 32 Hz (64/2). Conversely, to faithfully display a 30-second epoch with components up to 40 Hz, the computer monitor would need to have a horizontal resolution of at least 2,400 pixels.

With 1,080 pixels as vertical resolution and 20 channels, 54 pixels (14.4 mm) per trace are available and it is enough to display EEG signals ($\pm$50 μV; 1.85 μV per pixel) with a sensitivity of 7 μV/mm. In the event where more channels must be displayed, then the vertical monitor resolution must be increased (by using a larger monitor with, for example, 1,200 or 1,600 pixels).

The display of PSG signals on monitors with inadequate resolution can result in 'spatial aliasing' which is similar to violating the Nyquist-Shannon sampling theorem when choosing the recording sampling rate. Spatial aliasing artefacts will occur when viewing fast frequency activity exceeding the maximum frequency given by the monitor's horizontal resolution. When such spatial aliasing artefacts are suspected, the number of seconds displayed per page should be decreased. Throughout the duration of the visual scoring, epochs must be equal in length, as expanded views may affect the scorer's interpretation of the data. It is not acceptable to modify the start or endpoints of epochs in order to create new epochs.

In summary, 23-inch computer monitors with a resolution of 1,920 × 1,080 pixels, as commonly available at the time of writing, are sufficient for the scoring of standard PSG signals.

Visual Scoring of the Sleep Macrostructure

After Monroe [90] stressed in 1967 at a meeting of the Association for the Psychophysiological Study of Sleep (APSS) the lack of inter-rater reliability of the scoring of sleep records, a committee was formed to standardise the methods of assessment and evaluation of sleep recordings. This committee, which was spearheaded by Allan Rechtschaffen and Anthony Kales (R&K), came up with *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects* [8], which for almost 40 years served as the gold standard.

The R&K manual specified that the visual scoring of sleep records (which was originally performed using paper traces) should be based on information from 3 biosignals: (1) an EOG to record eye movements with one electrode placed 1 cm above and one 1 cm below and both slightly lateral to the outer canthus and referenced to the same ear or mastoid electrode; (2) an EMG to assess muscle tone with electrodes placed on and beneath the chin (mental and submental), and finally (3) an EEG to record brain activity with electrodes placed either at C4/A1 or C3/A2. Based on these 3 biosignals, 7 (sleep) stages are distinguished: wake, movement time, stage 1, stage 2, stage 3, stage 4 and REM sleep (stage REM). This standard was not changed with the transition from paper to digital recordings.

R&K pointed out that the proposal was designed for (healthy) adult humans. They were aware that there are, in some cases, individual variations, which may require further elaboration. This became apparent particularly when the standard was applied to subjects with sleep disturbances. They furthermore stressed that this 'handbook should be viewed as a working instrument rather than a statute. … Experience with the manual may suggest possible revisions. When these suggestions accumulate appreciably, it would seem in order to have a review of the manual' [8, p. 13]. With increasing knowledge (e.g. of developmental changes throughout the lifespan, the nature and importance of other sleep-related events and phenomena), several limitations of the R&K standard were indeed recognised over time and suggestions for modifications and amendments were made. Nevertheless, it took almost 40 years for a major revision to be published. The process of the development of a new manual was initiated in 2003 by the Board of Directors of the AASM. The manual was published in 2007 [12]. With regard to sleep stage scorings the major advantage is the introduction of technical and digital specifications, and more precise scoring rules. EEG is recommended to be recorded from 3 sites: frontal (F4/F3), central (C4/C3) and occipital (O2/O1) referenced against the contralateral mastoid. Furthermore, basic parameters to be reported for polysomnography and its definitions are specified. The number of stages which are distinguished is reduced to 5: stage W (wakefulness), stage N1, stage N2 and stage N3 (a combination of R&K stages 3 and 4) and stage R. There is an ongoing debate as to whether the new standard adds clinically relevant information [91, 92] or improves inter-rater reliability of sleep stage scoring [93] and, in particular, amongst European sleep medicine/sleep research societies as to whether they should adopt this standard.

When reporting the results of clinical trials, it should be clearly stated which standard was used and it should be kept in mind that results obtained with different standards are not directly comparable [94]. Finally besides the need for standardisation, which allows comparability, there should always be room for additional approaches, e.g. recording with more electrodes, to enable novel explorative data analyses to deepen further our understanding of the complexities of sleep.

Visual Scoring of the Sleep Microstructure

The current and widely accepted standard for the visual scoring of sleep microstructure has been defined in the AASM manual [12] and discussed in detail (e.g. [95]). To maintain consistency, the present pharmaco-sleep guidelines fully adhere to this standard and the definitions presented in *italics* are taken from the AASM manual.

Due to the complex nature of sleep microstructure measures, it is likely that the endpoints used will be study dependent in each case and hence specific recommendations for the quantitative measures to be used are not included in this section. Typically, endpoints will be count variables such as the number of occurrences of a particular event in a given period of sleep or the density (number of events per hour), or some measure of the intensity or frequency of the phenomenon.

In general, the definitions and morphology of the various events observed in the sleep microstructure are strongly dependent upon the recording techniques used (EEG locations and reference). For the remainder of this section, the given guidelines rely on the standard PSG montage (see section 'Data Acquisition').

Sleep spindle = A train of distinct waves with frequency 11–16 Hz (commonly 12–14 Hz) with a duration ≥0.5 seconds, usually maximal in amplitude using central derivations. Sleep spindles are rhythmic, sinusoidal waves characterised by progressively increasing, then gradually de-

creasing ('waxing and waning') amplitudes. Sleep spindle properties (density, amplitude and frequency) are affected by age [96, 97] or by sleep deprivation [98], and their frequency might be influenced by sleep disorders, neurological disorders, and hypnotic drugs [99, 100]. Sleep spindles seem to play a major role in memory consolidation [101] and might be used as a biomarker for general cognitive and learning abilities [102]. Sleep spindles show a large variability in their topographic distribution [103], but the majority of them appear in central regions [104], with a bimodal distribution of activity in the sigma range characterised by a slow sigma activity (around 12 Hz) predominant over the frontal areas and a fast activity (around 14 Hz) over the midline central and parietal areas [105–107]. As sleep spindle incidence is more pronounced in centro-parietal leads, scoring should include at least 1 central channel.

K complex = A well-delineated negative sharp wave immediately followed by a positive component standing out from the background EEG, with total duration ≥0.5 seconds, usually maximal in amplitude when recorded using frontal derivations. For an arousal to be associated with a K complex, it must commence no more than 1 second after termination of the K complex. K complexes represent a synchronised pattern consisting of alternating bursts of firing and silence within extended cortical networks during sleep, which trigger and synchronise other sleep activities in the thalamus [108] and are thought to relate to a sleep-protecting mechanism responsible for maintaining sleep [109]. Amplitude and frequency of occurrence of K complexes decrease with age [110] and under the influence of hypnotic drugs [99, 100]. Both evoked and spontaneous K complexes show a frontal maximum [104, 111, 112] and are usually bilaterally symmetrical [113]. It is therefore recommended to score from frontal channels, wherever possible.

Vertex sharp wave = Sharply contoured waves with duration <0.5 seconds maximal over the central region and distinguishable from the background activity. The vertex sharp waves are grapho-elements standing out from the background EEG. They become apparent during the sleep onset period [114] and mainly occur in late stage N1 and early stage N2, but also during REM sleep. They are observed in wide scalp areas with the maximal amplitude at Cz [115] and may become inconspicuous and poorly demonstrable in elderly subjects [116].

Sawtooth wave = Trains of sharply contoured or triangular, often serrated, 2-6 Hz waves maximal in amplitude over the central head regions and often, but not always, preceding a burst of rapid eye movements. Like vertex waves, the sawtooth waves are a typical EEG pattern of stage R predominantly observed in the central areas [115, 116] and should be assessed from central channels.

Rapid eye movements (REM) = Conjugate, irregular, sharply peaked eye movements with an initial deflection usually lasting <500 msec. REM are a key characteristic of stage R, which is altered in a number of sleep disorders [117]. They may also be observed during wakefulness when subjects scan the environment, but disappear when drowsiness increases.

Slow eye movements (SEM) = Conjugate, reasonably regular, sinusoidal eye movements with an initial deflection usually lasting >500 msec. SEM are a phenomenon typical of the sleep onset period and are a maker of sleepiness, but they are also found in REM sleep [118, 119].

Arousal = An abrupt shift of EEG frequency including alpha, theta and/or frequencies greater than 16 Hz (but not spindles) that lasts at least 3 seconds, with at least 10 seconds of stable sleep preceding the change. Scoring of arousal during REM requires a concurrent increase in submental EMG lasting at least 1 second. While the nature of arousals in sleep is still a matter of debate, there is evidence showing that arousals play a prominent role in the pathophysiology of sleep disorders [120] and constitute a suitable marker for both the diagnosis of primary insomnia and the evaluation of treatment efficacy [121]. In addition, there are significant changes in arousal threshold during the recovery sleep following sleep deprivation [122, 123].

Cyclic alternating pattern (CAP). The CAP is a long-lasting periodic activity consisting of arousal-related phasic events (phase A) that periodically interrupt the tonic theta/delta activities of NREM sleep (phase B), characterising two different functional states in the arousal control mechanism [124]. Clinically, CAP is a potentially useful EEG feature in the diagnosis of posttraumatic and other causes of coma [125] and may be a relevant marker in other conditions involving sleep fragmentation, such as fibromyalgia [126]. The physiological fluctuations of CAP activity during sleep are accompanied by changes in balance between the sympathetic and vagal components of the autonomic system [127]. Lastly, CAP measures are sensitive to pharmacological treatments and constitute valuable endpoints in insomnia research [128].

Visual Scoring of Additional Sleep-Associated Events

The scoring of sleep-associated events includes the scoring of movement events, which are defined in the AASM manual [12] and described further in a review pa-

per [129]. These events include twitches, brief muscle activations, leg and limb movements, increased and decreased muscle tone. Specific events and combinations of events are related to sleepwalking and movement disorders during sleep. They are distinguished as simple movements and complex movements, which may show some coordinated activations. Sleepwalking is a complex movement during sleep. Talking during sleep is an additional event involving some motor functions. Bruxism during sleep, which can be induced by various psycho-active medications [130], can be scored by video, muscle recordings, and sometimes even by audio events.

The respiratory events, which are scored in cardiorespiratory PSG, are apnoea, hypopnoea and hypoventilation events. Definitions are provided in the AASM manual [12] but it should be noted that there is still some uncertainty about the optimal definitions [131], which is reflected in the AASM manual by the inclusion of two different options for hypopnoeas, for example. Apnoea and hypopnoea events have a minimum duration of 10 s and are defined as a drop in respiratory amplitude by at least 90% and either 30 or 50%, respectively. Obstructive, central, and mixed apnoea events are distinguished based on the parallel scoring of oronasal airflow and respiratory effort recorded by respiratory movement sensors. Central apnoea is a cessation of airflow and no remaining respiratory effort. Obstructive apnoea is a cessation of airflow with remaining respiratory effort. A mixed apnoea shows both, initially no respiratory effort which then resumes during the second part of the event. As part of the hypopnoea definition, oxygen desaturation is evaluated. Two alternative definitions of a hypopnoea are used: (i) a drop in oxygen saturation by at least 4% and a reduction of nasal pressure by at least 30% for over 10 s and 90% of the event's duration or (ii) a drop in oxygen saturation by at least 3% and a reduction of airflow by at least 50% over the same time period. The first definition is recommended by the AASM and is actually the same as being used by some reimbursement regulations in the USA (Medicare). The second definition has been used in many publications and is therefore included as a reference to literature. A hypoventilation is defined by an increase in CO_2 above the normal increase which can be observed during sleep. The pathological increase in CO_2 shows that the lowering of airflow during sleep is more than it should be. Snoring is not scored regularly because it is regarded as a normal variant and not as a pathological event. No standard definitions for the scoring of snoring have been developed until now. Respiratory flow limitation has often been discussed as having pathological consequences but the definitions of flow limitation were variable between studies. The scoring of respiratory effort-related arousals is now defined as an option to solve this problem. The definition of respiratory effort-related arousal is based on the fact that some degree of flow limitation tends to lead to cortical arousals, and hence combines arousal criteria [132] with the requirement for 10 s of increasing respiratory effort of reduced airflow. For children the rules for respiratory events are revised and are, in general, based on shorter time intervals [133].

The scoring of cardiac events during sleep is limited due to the fact that usually only one ECG lead is recorded [134]. The interpretation of a single-lead ECG is necessarily reduced to heart rate changes and a coarse examination for detection of ectopic beats (detailed classification of ectopic beats requires more than one lead). Likely asystoles can be detected, but a multiple-lead ECG is required for confirmation. Therefore, in general, the scoring of cardiac events from a standard PSG is reduced to simple yes/no decisions about the likelihood of a particular abnormality, and multiple-lead diagnostic ECG recordings are required for a more detailed assessment and firm diagnosis.

The recording of other, unforeseen, events during sleep is important, because this leaves the door open for any kinds of known and unknown phenomena. These could be movement-related but outside the conventional definitions mentioned above, gastric reflux events, nocturnal asthma episodes, epileptic seizures or simply sleep-talking. From the perspective of a sleep scorer, it is important that such events are marked, as detection will rely on linking them to the sleep log taken by the sleep technician during the recording. In this context, video and audio recording are mandatory.

Scorer

A crucial aspect in the evaluation of sleep recordings by human scorers is the inter- and intra-individual reliability. This can be judged at different levels: (1) by visual inspection of hypnograms, (2) by comparison of (quantitative) sleep parameters derived from PSG studies, and (3) by comparing scorings epoch by epoch. Visual inspection is merely subjective and has no statistical approach to quantify the degree of agreement or disagreement. For quantitative sleep parameters the degree of agreement between 2 experts can statistically be assessed by a t test for dependent variables (for normally distributed traits) or independently of the distribution by mean of Wilcoxon's matched pairs signed-rank test. For more than 2 experts the corresponding tests are a univariate analysis of variance with repeated measures or the non-parametric

Table 2. Matrix of (sleep) stage scorings by two scorers according to the R&K (1968) standard, stages 3 and 4 are combined into SWS

	Scorer 2					Sum
	Wake	Stage 1	Stage 2	SWS	Stage REM	
Scorer 1						
Wake	***f_{11} = 33,045***	f_{12} = 4,345	f_{13} = 1,993	f_{14} = 52	f_{15} = 302	$f_{1.}$ = 39,737
Stage 1	f_{21} = 4,189	***f_{22} = 10,033***	f_{23} = 6,379	f_{24} = 60	f_{25} = 1,296	$f_{2.}$ = 21,957
Stage 2	f_{31} = 2,093	f_{32} = 7,378	***f_{33} = 65,114***	f_{34} = 4,193	f_{35} = 1,207	$f_{3.}$ = 79,985
SWS	f_{41} = 68	f_{42} = 55	f_{43} = 5,042	***f_{44} = 13,761***	f_{45} = 2	$f_{4.}$ = 18,928
Stage REM	f_{51} = 547	f_{52} = 2,837	f_{53} = 1,772	f_{54} = 3	***f_{55} = 22,800***	$f_{5.}$ = 27,959
Sum	$f_{.1}$ = 39,942	$f_{.2}$ = 24,648	$f_{.3}$ = 80,300	$f_{.4}$ = 25,607	$f_{.5}$ = 18,069	N = 188,566
	$\frac{f_{1.} \cdot f_{.1}}{N^2}$	$\frac{f_{2.} \cdot f_{.2}}{N^2}$	$\frac{f_{3.} \cdot f_{.3}}{N^2}$	$\frac{f_{4.} \cdot f_{.4}}{N^2}$	$\frac{f_{5.} \cdot f_{.5}}{N^2}$	Pr(e)
	0.045	0.015	0.181	0.010	0.020	0.270

The percentage of agreement between the two scorers results from the sum of epochs in the diagonal of the matrix (indicated in bold and italics). In this case the sum is 144,753 which is 76.8% of all epochs scored (N = 188,566):

$$\Pr(o) = \frac{f_{11} + f_{22} + f_{33} + f_{44} + f_{55}}{N} = 0.768$$

The κ-coefficient is a statistical measure of inter-rater agreement which is generally considered to be a more robust measure than simple agreement calculation because it takes into account the agreement occurring by chance. Cohen's κ is defined as the difference between the observed and the expected level of agreement, Pr(o) and Pr(e) respectively, expressed as a fraction of the maximum level of agreement exceeding chance:

$$\kappa = \frac{\Pr(o) - \Pr(e)}{1 - \Pr(e)}$$

with

$$\Pr(e) = \sum_{i=1}^{5} \left(\frac{f_{i.} \cdot f_{.i}}{N^2} \right)$$

Applying the formula, it follows that κ = (0.768 – 0.270)/(1 – 0.270) = 0.682

Friedman test. Another approach is to calculate intraclass correlations, although these carry the caveat that such correlations do not take into account systematic over- or underestimation by 1 scorer. For an epoch-by-epoch comparison of scorings from 2 or more scorers as well as intra-individual variation, the degree of agreement can be specified as percentages or using κ-statistics (Cohen's κ for 2 scorers and Fleiss' κ for more than 2 scorers). Since κ takes into account the agreement occurring just by chance, it is a more robust measure than percentage agreement.

The difference between the percentage of agreement and κ for 2 scorers is illustrated in table 2 based on a dataset from an EU-funded research project presented in Danker-Hopfe et al. [135]. The table summarises the scoring results for a sample of 196 nights from 98 patients with different (sleep) disorders. In this example, the simple percentage agreement between the 2 scorers is 76.8% and the Cohen's κ-coefficient is 0.682.

Empirical data show that the degree of agreement decreases with an increasing number of scorers and/or laboratories involved as well as with age of patients/subjects. Agreement is lower in patients than in healthy subjects, varies with sleep stages (best for REM/R sleep and worst for stage 1/N1 sleep), the number of stages distinguished (the lower the number of stages distinguished, the higher

the agreement), and medical condition/sleep disorder. Finally for slow-wave sleep (stages 3 and 4 and N3, respectively) agreement is lower for male than for female subjects [93, 135].

In the framework of the development of the AASM scoring standard, Silber et al. [95] pointed out: 'No visual based scoring system will ever be perfect, as all methods are limited by the physiology of the human eye and visual cortex, individual differences in scoring experiences, and the ability to detect events viewed using a 30-second epoch. Nevertheless, we believe it is possible to develop a rigorous, biologically valid scoring system that can be applied meaningfully in clinical and research settings. The new scoring system is presented as a step forward along this path' (p. 129). A study by Danker-Hopfe et al. [93] showed that there is no substantial improvement in the inter-rater reliability when scoring is done according to the AASM standard as compared to the R&K standard. The physiology of the human eye and visual cortex mentioned as limitation above can also be an advantage in certain clinical trials and especially in the PSG screening phase. The human eye might detect small changes in the EEG signal, e.g. seizure-like discharges or bruxism, which might escape detection in an automatic analysis.

A prerequisite for visual scoring in any case is that the scorer is well trained and has sufficient experience. Scorers should regularly participate in internal and external training/retraining activities with certification. In the context of clinical trials, the following requirements apply:

- In multi-centre studies, a central scoring laboratory is to be used due to the scoring variability generally observed between scorers from different sleep laboratories [135–137].
- When several experts are involved in the scoring for one trial, the inter-rater agreement (Cohen's κ or Fleiss' κ, as appropriate) should be reported alongside the results. In general, $\kappa \geq 0.75$ should be achievable in pharmaco-sleep studies although it is recognised that in some patient populations, this level of agreement may not be attained. Studies where $\kappa < 0.60$ should be interpreted with caution.
- Repeated recordings from the same subject should always be evaluated by the same expert.

Reporting Results

Table 3 summarises the sleep parameters extracted from visual scoring to quantify sleep induction and continuity and sleep architecture. This set of measures constitutes the minimum requirement for reporting results in the context of clinical trials, although additional parameters may have to be included dependent upon the study objectives. In many cases it may be useful to report the relevant measures for each hour and each quarter of the night, in addition to the whole night, to highlight any temporal variations. In addition to the natural variations in sleep architecture through the different phases of the night, such temporal variations are particularly relevant for pharmaco-sleep studies, since drug concentrations and the effects on sleep can vary considerably over time.

The representation of treatment effects is usually based on group statistics which correspond, for each recording session, to the average (mean) of the individual parameters within treatment groups. From a theoretical point of view, the mean may be inappropriate for describing the central location in the case of one-sided heavy-tailed distributions, since a single value can induce a bias which is considerably accentuated in case of a small sample size. Thus, if the distribution of the data is not symmetric, which is particularly the case for measurements of latencies, the mean results must be assessed with caution. On the other hand, the median might be too robust if the distribution is heavy tailed. An appropriate way of circumventing this problem is to use the trimean as a non-parametric measure to quantify the central location of the sample. This measure combines the median for robustness, with the quartiles suitable for reflecting asymmetries in the data distribution [138].

Consequently, when reporting group (i.e. treatment) results, the parameters listed in table 3 should include mean, standard deviation, median, first and third quartiles, and trimean for descriptive purposes. Further guidance on the reporting of the results of statistical testing is given in the section 'Reporting of the Results of Statistical Analyses'.

Digital Data Processing

Fundamentals

Representation in the Time Domain

In the time domain, the variations of potential after amplification are displayed as a function of time and signals are usually denoted by a function s(t) [or $s(k\,T_S)$ (with $k = 1..N$) in its digital form]. The representation in the time domain is used for the visual inspection of PSG traces and for the evaluation of events for which the position in time is relevant. Thus, time is considered as a variable of the observed phenomenon and the specifications in terms of amplitude and time resolution may differ dependent upon the information content of the recorded sig-

Table 3. Sleep scoring parameters to be reported when summarising the results of pharmaco-sleep studies. Additional parameters may have to be added depending upon the study goal. The reader should also refer to the AASM manual, in particular section II, when determining which additional parameters to report

Nomenclature	Definition
Sleep induction and continuity indices	
Total recording time (TRT), min	Lights-out to lights-on
Total sleep time (TST), min	Number of minutes asleep (in any sleep stage) from lights-out to the end of the recording
Sleep efficiency (SE), %	(TST/TRT) × 100
Sleep onset latency (SOL), min	Time from lights-out to sleep onset (i.e. the first epoch of any sleep stage)
Latency to persistent sleep (LPS), min	Time from lights-out to the start of persistent sleep (i.e. the first consecutive 10 min of sleep)
Wake after sleep onset (WASO)[1], min	Number of minutes of stage W between the start of persistent sleep and the end of the recording
Wake time during sleep (WTDS)[1], min	Number of minutes of stage W between the start of persistent sleep and terminal awakening
Wake time after sleep (WTAS)[1], min	Number of minutes of stage W between terminal awakening and the end of the recording
Stage N3 latency, min	Time from sleep onset to first epoch of stage N3
Stage R latency, min	Time from sleep onset to first epoch of stage R
Number of arousals	Number of arousals after sleep onset. Typically denoted NASO
Number of awakenings	Defined as wake periods of at least 1 or 2 epochs after sleep onset Typically denoted NAASO1 and NAASO2, respectively
Sleep architecture indices	
Time in stage N1, min	Number of minutes in stage N1
Stage N1, %	Percentage of total sleep time spent in stage N1
Time in stage N2, min	Number of minutes in stage N2
Stage N2, %	Percentage of total sleep time spent in stage N2
Time in stage N3 (delta sleep), min	Number of minutes in stage N3
Stage N3 (delta sleep), %	Percentage of total sleep time spent in stage N3
Time in stage R (REM sleep), min	Number of minutes in stage R
Stage R (REM sleep), %	Percentage of total sleep time spent in stage R

[1] Note that WASO = WTDS + WTAS.

nals. The detection of patterns or transient activities in the PSG signals usually relies on processing algorithms operating in the time domain.

Representation in the Frequency Domain

The transformation of a signal s(t) into the frequency domain using the fast Fourier transformation (FFT) implicitly assumes that s(t) can be split up as a finite sum of weighted sinusoidal waveforms [denoted as s(f)]. The number of sinusoidal waveforms is dependent upon the window size (i.e. the number of points of the input signal) subjected to FFT. The resulting graphical representation displays the spectral characteristics of s(t), which is then depicted by peaks in the frequency domain. Thus, while the representation of a signal in the frequency domain differs from its representation in the time domain, it is only another way to present the same information. This kind of display is generally used for the evaluation of spontaneous activity for which the position in time of events has no direct relevance.

The FFT (as an orthogonal transformation) is essentially a mathematical operation performed on time series data which does not alter the information content of the signal. Neither is any assumption made regarding the nature of the data or any interpretation implied. Within the limitations of computational accuracy, the full reversibility of the transformation is implicit and given only as long as numbers are retained in their complex form and not averaged.

The resolution in the frequency domain depends upon F_S and the number of sampling points (N, size of the signal window) subjected to FFT. This resolution (denoted here as ΔF) is given by the ratio ($\Delta F = F_S/N$). Accordingly, the longer the signal window N, the better the resolution of the frequency content.

When considering the results of an FFT applied to a signal window of N = 2,048 points (a number chosen because it corresponds to 2^{11} and facilitates a rounded-off length) with $F_S = 512$ Hz, then $\Delta F = 0.25$ Hz, which means that the frequency analysis can resolve 0.25 Hz (that is,

Table 4. Frequency ranges for EEG spectral analysis in pharmaco-sleep studies

Nomenclature	Frequency range, Hz	Unit of result
Delta	0.5 to <4.0	μV/Hz
Theta	4.0 to <8.0	μV/Hz
Alpha	8.0 to <12.0	μV/Hz
$Beta_1$ (sigma)	12.0 to <16.0	μV/Hz
$Beta_2$	16.0 to <20.0	μV/Hz
$Beta_3$	20.0 to <30.0	μV/Hz
Gamma	30.0 to <40.0	μV/Hz
Total	0.5 to <40.0	μV/Hz
Dominant frequency	4.0 to <12.0	Hz
ASI	$\frac{\text{alpha}}{\text{delta + theta}}$	–

The alpha slow-wave index (ASI) is defined as the ratio between alpha activity and the sum of the activity in the delta and the theta frequency ranges [144].

resolve 10.25 vs. 10.50 Hz directly). In this particular case, the signal window (called as epoch) will have a length of 4 s.

To reduce the broadband artefact, known as leakage, the signal window must be tapered toward zero at their initial and final data points (this tapering is usually done using a windowing function). When the FFT is applied on sequential epochs, then discarding a proportion of the signal through windowing can lead to differences in spectra depending on the starting point of the epoch series. An alternative that results in a spectrum less sensitive to the starting point is to use partially overlapping epochs so that all data is represented.

Spectral analysis via FFT was one of the first computerised techniques used for the parameterisation of the EEG during sleep [139, 140]. It continues to be the most common method of choice for the processing of EEG signals and the assessment of frequency versus energy variation as a function of time (i.e. overnight).

Non-Stationarity

Many signals, including EEG, are non-stationary, which means that they have a time-varying frequency spectrum, although they can be considered locally stationary over short segments in which the parameters of interest vary minimally.

In practice, the choice of the segment length is a trade-off between frequency resolution (which suggests a longer epoch) and ensuring quasi-stationarity (which suggests a shorter epoch). For the pharmaco-sleep EEG, epochs of 2–10 s duration are used.

Spectral Analysis of the Sleep EEG

The traditional parameterisation of pharmaco-sleep EEG activity is largely based on spectral analysis. To this end, the recorded signals are divided into epochs (2–10 s) which are subjected to spectral analysis using FFT. This transformation in the frequency domain and subsequent computation of the power spectrum allows a first data reduction.

The second step of data reduction consists of the extraction of spectral parameters. The frequency range is subdivided into frequency bands and the spectral performance (area under the curve) is computed for each of them and expressed in microvolts (square root of absolute power) or using another transformation (e.g. the natural logarithm) to better meet the assumption of normal distribution [141–143]. The transformation should preferably be carried out prior to any other manipulations, such as averaging spectral parameters across several epochs.

Substantial variability exists in the literature regarding frequency bands [83]. For quantitative pharmaco-EEG studies, the IPEG has recently published a definition of frequency ranges to be used in the context of drug testing [1]. These frequency bands have been defined on the basis of factorial analysis of EEG records. Table 4 provides a summary of the frequency ranges to be used in pharmaco-sleep studies. It does not mean that other frequency ranges shall not be used for specific purposes. However, to ensure that the results of a study can be compared with other published studies and that the results can thus provide useful reference material for other scientists, publications should always report the results obtained for this standard frequency band configuration (beside others if appropriate).

Absolute spectral EEG values are recommended as the primary outcome measures (endpoints) in the pharmaco-EEG and pharmaco-sleep profiling. Test-retest reliability investigations have shown that intra-individual EEG spectral measures can be treated as a stable trait [145]. Additional computed spectral parameters, such as relative values in frequency ranges, average and dominant frequency within a frequency band, peak skewness (asymmetry coefficient), peak kurtosis (peak shape), activity ratio between different frequency bands and log-transformed values [log(x/[1 – x]) where x = relative power in a frequency band] should be interpreted in the light of absolute values. These derived parameters may provide additional insights for the interpretation of the data.

When carrying out spectral analysis, decisions must be taken, according to the objectives of the study, over which time periods to average the data. In most cases the spectral properties of NREM and REM sleep should be analysed and reported separately. In many cases, in order to highlight any temporal variations, it may also be useful to report the relevant measures for each discrete period of NREM or REM sleep, or to divide the recording into hours and quarters of the night, as suggested above for PSG measures.

Digital Artefact Processing

Artefact identification and elimination is crucial for the proper quantitative analysis of EEG records. Artefacts can have various physiological origins (see section 'Artefacts') and can be identified and/or eliminated, either online during the recording or offline.

Rejection procedures omit segments with artefacts from analysis and are typically used in conjunction with experimental control, while correction procedures attempt to remove the effect of artefacts from the EEG signal. A number of algorithms have been developed for the automated rejection and correction of EOG artefacts [146–149], ECG artefacts [150–153] and EMG artefacts [154–157] present in the EEG. The techniques rely on a palette of numerical methods, such as autoregressive models, independent component analysis, regression analysis, general linear models, wavelet transformation and adaptive filtering. However, when computerised algorithms are used, a semi-automatic procedure that includes additional visual inspection is recommended, and while the approaches used for automated removal of interference caused by artefacts demonstrate good performance, they must be applied with caution.

Should the investigator doubt the validity of the procedure either because of the large percentage of EEG segments containing artefacts or because the kind of artefacts could be confused with the treatment effect, for example due to similar spectral content, then a comparative biometrical evaluation and assessment of the artefact-free and the complete data should follow [89].

Digital Scoring of the Sleep Macrostructure

Attempts to develop computer-assisted identification of sleep stages [158, 159] are as old as the rules of Rechtschaffen and Kales [8]. The main drivers for the development of an automated scoring were twofold: first, visual scoring is a very time-consuming, and hence costly, practice and automation would save the time spent by experienced sleep scorers and reduce the overall cost of the study; second, the assumption is made that automation would improve rating reliability and quality, thereby enhancing the understanding of the sleep process.

Over the years, many attempts of computerised sleep staging have been undertaken [160]. Various techniques have been applied, such as periodic and discriminant analyses [161], power spectral analysis [162, 163], deterministic and stochastic methods [164, 165], neural networks [166, 167], segmentation and clustering [168, 169], hidden Markov models based on data from a single EEG channel [170, 171], non-linear techniques combined with a gaussian mixture model classifier [172], and multi-dimensional analysis based on linear and non-linear parameters [173]. Several commercial systems have been developed [174–178]; however, it is generally admitted that routine use requires manual supervision and intervention (semi-automated scoring). Nevertheless, the scoring time can be reduced by a factor of at least five [179].

In most cases, the algorithms listed above have been constructed with the goal to replace (mimic) visual scoring and the performance has been tested by quantifying the match with the outcome from visual scoring with the R&K and, more recently, the AASM rules. Although the automated methods have demonstrated satisfactory concordance, the achievable agreement rates depend substantially on the patients' age and diagnosis. In addition, since all these methods have been validated with different recordings, a direct comparison of the results of the various approaches and their respective performance is not possible. Two factors with equal importance have to be taken into consideration [180]:

(1) the problems that have hindered a straightforward and generally accepted automated solution are mainly related to peculiarities of the scoring rules, artefact recognition, and individual differences in electrophysiological sleep signs;
(2) comparisons of the performance of visual and computerised sleep stage scoring have historically been a one-way street, since the results of visual scoring have been taken as the reference against which the automated scoring results should be compared.

When considering the inter- and intra-rater reliability generally observed when assessing visual scoring [93], whether with the R&K or the AASM rules, it is indeed questionable whether measuring the performance of a computerised system compared to such a reference is a valid approach. By default, the achievable outcome cannot be better than what is achieved based on a group of raters. Actually, the strength of a computerised solution

resides in the fact that intrarating reliability equals 1, and it is probable that the interrating reliability between systems based on an assessment of the same set of recordings would provide results similar to what is observed by visual scoring.

Nevertheless, for the time being, visual scoring remains the standard in the context of clinical trials, and unsupervised automatic analysis is not recommended. However, to reduce the influence of interindividual differences in visual sleep scoring, a semi-automated method with minimal visual editing should be applied in pharmaco-sleep studies in addition to the usual safeguards such as double-blind readers etc. Such an approach has shown its validity [181–183], even if caution and further validation are required [184, 185].

Digital Scoring of the Sleep Microstructure

As seen in the previous section, the use of a semi-automated method for analysing the sleep macrostructure (i.e. 'sleep staging') has been validated in the context of drug research and can therefore be used. However, sleep is a continuous process over time with transitions between different states characterised by specific varying activities, and staging is most certainly resulting in significant amounts of information being ignored. In addition, the following aspects have to be considered:

(1) During the process of visual evaluation of PSG curves, the decisions of the human scorer are implicitly influenced by the results of neighbouring epochs. In extreme cases, there is an ongoing change of the scoring criteria. Such a fuzzy procedure could be simulated by computers, provided the rules have been clearly defined so that they can be translated according to signal-processing principles.

(2) The assessment of the structural complexity of sleep stages using taxonomic statistics reveals that visually scored sleep stages are not homogeneous units, but rather agglomerations of various significant configurations (types) from physiological variables that represent different aspects of the momentary EEG and EMG activity [186]. Similar observations are made when investigating sleep stage 2 [187]. Thus, significant variations exist within the same stage and are indicated by changes in different sources and biosignals.

(3) While early scoring was restricted to a very limited number of recording channels, the recording of a variety of additional channels (respiratory, cardiac, additional EEG and EMG traces, and others) has become standard practice and, therefore, has been defined with recording techniques and standard parameter extraction in the new AASM scoring manual. The additional signal analysis according to arousal rules, cardiac rules, movement rules, and respiratory rules imposes a heavy load on the human visual analyser [180].

Digital processing allows the characterisation and quantification of one or several parameters of a signal and the display of its fluctuations over time with a high temporary resolution (e.g. with 1-second cycle). A simple example is the representation over the night of the continuous variations of slow waves (i.e. square root of the absolute power in the delta range) and sleep spindle activity (sigma), both showing plots highly correlated with the sleep stages of the hypnogram (e.g. fig. 1 in Kemp [87]). Similarly, the alpha slow-wave index can be used to detect episodes of wakefulness in sleep in both young healthy subjects and elderly insomniacs with accuracy [188]. Composite parameters based on EEG and EMG activities display synchronisation and desynchronisation as descriptors of the time course of sleep [189]. Principal component analysis applied to spectral parameters offers information on the time course dynamic of the sleep cycles [190] and allows the quantification of drug effects on sleep onset latency [191]. The combination of autoregressive modelling with the pattern recognition capabilities of an artificial neural network can be used to track the sleep dynamic and to pinpoint both micro-arousals and periods of severely disturbed sleep [192]. The assessment of the temporal evolution of coherence and spectral power activity overnight in the low delta, alpha and sigma frequency ranges displays the switch between NREM and REM sleep [193]. Period amplitude analysis has been applied and compared with power spectral analysis for all-night recordings [194]. The investigation of the correlation between spectral EEG activity and heart rate variability during sleep shows that sympathetic nervous system activity continuously fluctuates in accordance with sleep deepening and lightening [195]. Finally, by comparing the predicted rhythm to the actual EEG at each sample with the rhythm of the previous one in specific frequency ranges, it is possible to construct micro-continuity parameters with physiological implications sensitive to sleep and less sensitive to artefacts [196]. When concurrently applied to the signals recorded from several electrode locations (topographic analysis), various kinds of correlation or propagation of brain activity between locations can be observed [87].

The strength of the approaches presented above relies on the continuous representation of activities over the night. Applied simultaneously to multiple channels, they provide the possibility to assess correlations and to de-

scribe various states of the underlying phenomenon (i.e. sleep) without having to build another sleep staging system in full.

Another approach to the assessment of the microstructure of sleep involves the detection of short-lived patterns or changes in the EEG signal and the localisation of specific events in the time domain. Different techniques have been applied to detect various patterns, such as sleep spindles [197–199], alpha waves [200], K complexes [100, 201, 202], arousals [203, 204], CAPs [205], transient EEG events [206], or eye movements [207–209]. Time-frequency signal parameterisation methods, such as wavelets and matching pursuit, provide an elegant way to assess EEG recordings and to localise sleep patterns (e.g. sleep spindles, K-complexes, arousals) in the time-frequency plane with high precision [210–213].

All these methods provide useful information about micro-events. However, the main drawback is that they have been developed and tested using various different samples. Therefore, a direct comparison of their performance is not possible, even if each method taken separately seems to provide good results. Only if they were available on the same platform enabling a meaningful assessment against each other using the same large sample of recordings, could their real value be determined.

In the context of drug testing in clinical trials, the computerised assessment of sleep microstructure is still in its infancy and cannot be considered appropriate for use as a primary endpoint, although exploratory evaluation may be valuable. However, it is clear that, through combining techniques and including a large number of biosignals and physiological parameters, it ultimately has the potential to provide the tools necessary for a thorough investigation of the effects of compounds on sleep.

Digital Analysis of Additional Biosignals

Sleep recording today is based on cardiorespiratory PSG. This typically includes additional biosignals besides the standard EEG, EOG and EMG recordings. Firstly, EMG of the muscles of the extremities is recorded in addition to the standard EMG of the chin and submental muscles, in order to analyse movements. Many well-validated algorithms are available for computerised processing of such limb movements [214]. The biggest challenge here is the removal of artefacts, which is followed by an envelope calculation (signal rectification with previous subtraction of the mean value). Based on this analysis, short EMG spikes (twitches), limb movements (arousal associated) as well as changes in muscular tone (associated with changes of sleep stages) can be detected. Criteria have been set to classify clinically relevant limb movements [214, 215]. It is more difficult and generally unsuccessful to use muscle tone for discriminating sleep stages, largely because a reduction in EMG signal quality, arising from a decrease in impedance or an increase in noise, tends to give rise to an apparent increase in muscle tone. Movements such as bruxism- and parasomnia-related movements have not been successfully detected using computerised analysis and require manual interpretation.

The computerised analysis of ECG is straightforward, and several different algorithms have been developed to derive heart rate from ECG recordings during sleep [216]. In addition other features can be readily derived, such as R, T or S wave amplitude, and width of the QRS complex. More sophisticated ECG-derived parameters, such as the cardiac vector angle or ST segment values, cannot be determined reliably with the single ECG lead typically used in a sleep study. In addition, the sampling rate chosen for the ECG may be a limiting factor. The AASM minimal specification of $F_S = 200$ Hz is sufficient for the calculation of heart rate and the other amplitude values mentioned above. However, $F_S \geq 500$ Hz is required to provide a full quantitative ECG analysis. Nevertheless, the ECG parameters derived even from a recording with a lower F_S are still good enough to investigate heart rate variability changes with sleep stages, with arousals and sleep disorders. Many algorithms have been developed to recognise sleep-disordered breathing from the sleep ECG [216, 217]. Heart rate variability analysis is recognised as a useful tool to calculate sympathetic and parasympathetic activity in an indirect way [218]. Respiration can be derived from cyclic changes in heart rate and from amplitude changes of R [219], T and S waves.

Continuous recordings of blood pressure and the pulse wave are also sometimes analysed in sleep recordings [220, 221]. For blood pressure, the systolic, diastolic and mean values need to be calculated per beat in order to detect rapid changes. Rapid changes are found with arousals and with pathologies [217]. Moderate variations are found with changing sleep stages. For pulse waves, there are several new methods under development. Some derive pulse transit time from a combination of the ECG and pulse wave measures. This is used as a blood pressure surrogate [222]. Others derive pulse wave velocity just from the pulse wave contour, again in order to capture a blood pressure surrogate. Amplitude analysis and wavelet analysis are the preferred tools for pulse wave analysis [223].

Statistical Concerns

There are two categories of sleep-related measurements: (i) summary variables derived from the epoch-by-epoch sleep stage scoring (e.g. total sleep time) and (ii) variables derived directly from the raw biosignals (e.g. average power spectral density in a given EEG frequency band). This section focuses specifically on the statistical analysis of these pharmaco-sleep variables. Please refer to Ferber et al. [224] and Jobert et al. [1] for statistical topics related to general pharmaco-EEG. For general information on the statistical aspects of clinical trial design and clinical data analysis, please refer to the International Conference on Harmonisation tripartite guideline [225] and the Committee for Proprietary Medicinal Products Working Party [226].

Study Design

PSG measurements generally have a large inter-subject variability but a relatively smaller intra-subject variability. A crossover clinical trial design is thus preferred in the early phases of drug development studies where regulatory acceptance is not a priority, in order to achieve higher statistical power with a smaller sample size [227]. In contrast, a parallel group design is preferred in later stage confirmatory trials to avoid regulatory concerns over the potential carryover issues associated with crossover designs [227]. In this case, the larger inter-subject variability of PSG measurements must be compensated by increasing the sample size of the study.

The appropriate sample size for a pharmacological sleep study should be estimated based on the expected change and variability in the primary endpoints using standard methods [228]. Even when PSG measurements are not specified as the primary endpoints, calculating the statistical power for the chosen sample size is still recommended, in order to provide a better understanding of how reliably changes in the PSG endpoints will be detected.

Statistical Modelling, Hypothesis Testing, and Inference

Although Student's t test, or its non-parametric counterparts (e.g. the Wilcoxon-Mann-Whitney test [229]), is widely used for its simplicity, it is only recommended for hypothesis testing for a single endpoint in a study directly comparing two conditions, e.g. before versus after treatment, or drug versus placebo, without additional covariates. Non-parametric methods are recommended when the endpoint, or its transformation, is not normally distributed. A paired difference test (i.e. the paired t test, or the Wilcoxon signed-rank test) should be used in a study when all treatment situations are applied to each subject. Beyond this simple scenario, inference based on a statistical model is recommended to improve the statistical power by incorporating covariates and other relevant factors.

When baseline data is collected in a study, it is recommended to incorporate the baseline in the statistical model using the constrained longitudinal data analysis method [230, 231] under the assumption that the subjects are properly randomised to the different treatment groups in a parallel design, or to different treatment sequences in a crossover design.

In many cases sex and age significantly affect PSG parameters. When modelling measurements from either a parallel study or an unbalanced crossover study, sex and age should be incorporated as covariates. Since the relationship between the PSG measurements and age is generally non-linear, the age should be treated as a categorical covariate after it has been stratified into several age ranges.

In a crossover study, it is recommended to first check whether a carryover effect exists [227]. If the carryover effect is negligible, a mixed effect model [232] should be used. The random effect in the mixed model is typically the subject effects. When a large clinical study involves many clinical sites, the random effect is composed of the subject effects nested within the site effects (which may be reduced by using a centralised and blinded expert scoring site). Besides the treatment and other obvious factors, the fixed effects of the mixed model can include the period factor and ignore the treatment sequence factor, if the primary interest is the treatment effect. Please refer to Kenward and Jones [227] for general statistical issues regarding analysis of crossover studies.

When pharmaco-sleep EEG is used to study a drug's PD properties, such as dose response and time course, it is typical that the PSG measurements of interest, denoted as variable Y, are a group of endpoints related by a parameter, denoted as variable X. That is, the Ys can be considered as a function of X [Y = f(X)]. It is recommended to include the variable X in the model as a categorical covariate. If the variable X has too many levels, the functional data analysis approach [233] suggests that a smaller number of properly chosen basis functions can replace the categorical variable X in the model [234]. This reduces the number of model coefficients to be estimated. As for specifying the covariance matrix structure in the model, if the study size is large, an unstructured covari-

ance matrix over the variable X is preferred; if the study size is not large, a correlation matrix with an AR(1) or a compound symmetry structure is preferred [232]. Markov chain models can be used to describe sleep stage transition as a function of time after drug intake and time after last sleep stage change. The probability to change sleep stage can be employed for PK/PD modelling [235].

The estimate of the treatment effect on a PSG measurement, along with its standard error (SE), can be calculated from the fitted model using a properly constructed contrast matrix and the variance-covariance matrix of the model coefficients.

Due to the relatively large data size of clinical EEG-sleep studies in drug development and especially in phase III, Student's t distribution of the effect estimate is frequently approximated as a normal distribution to remove the need to estimate the degrees of freedom under various complex situations. The effect estimate and its SE thus become sufficient to conduct hypothesis testing and to calculate the associated p-value and confidence interval (CI).

Multiplicity adjustment becomes necessary when hypothesis testing on a group of related endpoints is conducted simultaneously. If the estimates of this group of endpoints are obtained from the same model, the correlation matrix of these estimates can be derived from the model. Under this condition, it is recommended to use the multiplicity adjustment method proposed by Hothorn et al. [236]. This method allows the calculation of both the adjusted p-values and the global simultaneous CIs of the estimates. Conversely, when the estimates are obtained from different models, their correlation matrix is not readily available. It is then recommended to use Hochberg's [237] step-up procedure to control the family-wise error rate.

Reporting of the Results of Statistical Analyses

Tables are recommended to present the detailed statistics of the pharmaco-sleep endpoints for each treatment situation. The table should typically include at least 8 fields: (1) number of data points; (2) arithmetic mean; (3) standard deviation or SE; (4) geometric mean and (5) its CI; (6) treatment effect estimate and (7) its p-value, and (8) CI. Typically, 95% CIs are used, but this can be varied depending on the objectives of the study.

Graphics are preferable to present statistics for a group of related endpoints, since their treatment effect estimates, region of significances (determined by the p-values), and the CIs can be simultaneously presented in one plot, thereby facilitating the visualisation of drug effects.

Normative and Reference Databases

For the development of new compounds, it would be particularly helpful to have access to an accurate normative repository of data collected using standardised methods, providing benchmarks from a large and representative population of individuals (healthy volunteers and various patient populations, e.g. with sleep, psychiatric or somatic disorders) and using various drugs (with emphasis on reference drugs and including placebo). However, whilst some individual pharmaceutical companies and contract research organisations may have some data available, no such comprehensive repository is publicly available at the moment. For early drug development it would be important also to include reference datasets for various non-human species, enabling the study of translation and discovery of translatable biomarkers (or surrogate markers) for the early selection of potentially interesting and viable compounds and for guiding the design of early clinical studies.

The results of a large meta-analysis of quantitative sleep parameters published in 2004 [238] demonstrated age-related changes in objectively recorded sleep patterns across the human lifespan (children, adolescents, adults, elderly and old elderly subjects). However, the authors observed that the effect sizes for the different sleep parameters were affected by the quality of subject screening. Another study published in 2005 assessed datasets from 198 healthy non-sleep-disturbed subjects in the age range of 20–95 years and provided normative data for a large set of selected sleep parameters [239]. More recently, a collaboration of 16 sleep laboratories in Germany compiled normative data from 209 children and adolescents aged 1–18 years, thus providing a clear picture of the development of sleep in normal children [240, 241]. While all three studies offer helpful reference datasets that can be used in the context of the design of clinical trials, the lack of standardised screening, differences in study designs (in particular, whether a habituation night was included) and the limited sample size of most studies that were included restrict the application and interpretation of these data.

Several databases hosting sleep EEG recordings are available for download and include datasets recorded in healthy subjects and patients with sleep disorders, as summarised in table 5. Such repositories are particularly useful in providing data to evaluate signal-processing algorithms. More importantly, performance comparisons between various techniques would be highly facilitated and the value of the results significantly improved, if algorithms were to be tested against the same reference datasets.

Table 5. Examples of published and available databases with open access to PSG recordings from healthy subjects and patients with sleep disorders

Name, web source and reference	Population	Database description
The Sleep-EDF Database – sleep recordings and hypnograms in European data format (EDF) http://www.physionet.org/physiobank/database/sleep-edf/ Kemp et al. [196]	Healthy subjects: 8 males and females Age 21–35 years	Publicly available sleep data in EDF EEG (FPz-Cz, Pz-Oz), EOG (horizontal), EMG (submental) Respiratory signals: oronasal Additional signals: body temperature Signal sample rates: 100 and 1 Hz R&K hypnograms of all recordings Additional subject information: no
The Sleep Heart Health Study Polysomnography Database http://www.physionet.org/pn3/shhpsgdb/ Quan et al. [242] Lind et al. [243]	Patients: Approx. 6,400 Age >40 years No previous sleep apnoea, tracheotomy or current home oxygen therapy	Portable home sleep recordings EEG (C4-A1, C3-A2), EOG (left, right), EMG (submental), ECG Respiratory signals: flow, thorax, abdomen SaO_2 Additional signals: body position, ambient light Signal sample rates; 125, 50, 10 and 1 Hz Signal available in EDF format Additional subject information: yes
St. Vincent's University Hospital/ University College Dublin Sleep Apnea Database http://www.physionet.org/pn3/ucddb/ Heneghan et al. [244]	Patients: 21 males and 4 females mean age: 50 ± 10 years, range 28–68 with probable sleep-disordered breathing	One-night PSG recordings EEG (C4-A1, C3-A2), EOG (left, right), EMG (submental and tibials anterior), ECG Respiratory signals: flow, thorax, abdomen SaO_2, pulse Additional signals: body position Signal sample rates: 256, 128, 64, 8 Hz R&K hypnograms of all recordings and reports of respiratory events Additional subject information: no
SIESTA Database http://www.ofai.at/siesta/database.html (access request to www.thesiestagroup.com free for non-commercial purposes) Danker-Hopfe et al. [239]	Healthy subjects: 94 males, 103 females Age 20–95 years Patients: 72 males, 26 females Age 21–75 years With sleep disorders (psychiatric, respiratory and neurological comorbidities)	Two consecutive PSG nights EEG (FP_2-A1, FP_1-A2, C4-A1, C3-A2, O2-A1, O1-A2) EOG (left, right), EMG (submental), ECG Respiratory signals: flow, thorax, abdomen SaO_2, pulse Additional signals: additional EEG signals, snoring, body temperature, body position Signal sample rates: 256, 200, 128, 64, 16, 8 Hz R&K hypnograms (consensus scorings) of all recordings and reports of respiratory events, arousals and periodic leg movements in sleep, database of artefacts, microstructure of sleep (spindles, K complexes) Additional subject information: yes

Pharmaco-Sleep Study-Related Topics

Measurement of Sleep Tendency

An important aspect of pharmaco-EEG studies is the evaluation of the degree of daytime sleepiness. The gold standard for the measurement of sleepiness is the multiple sleep latency test (MSLT) [245, 246]. The MSLT is based on the assumption that the greater the degree of sleepiness, the greater the rapidity of sleep onset. Thus, the MSLT is a standardised methodology to assess rapidity of sleep onset or sleep propensity.

The standard MSLT consists of 4 or 5 nap tests. In each nap test the subject is put to bed in a sleeping room, similar to that described for nocturnal PSG, the lights are put out and the subject is instructed to try to fall asleep. The montage used for the MSLT is essentially the same as for the basic nocturnal PSG. For MSLT in research, the subject is deemed to have fallen asleep, and the test is terminated, after 3 epochs of stage 1 sleep, or 1 epoch of another stage of sleep. Three epochs of stage 1 are required to ensure that the subject has truly achieved sleep. For clinical MSLTs the nap is allowed to continue for 20 min to establish the onset of REM sleep, which is used as a diagnostic for narcolepsy. One of the criteria for narcolepsy is 2 or more REM naps out of 5, although this can also occur very occasionally with other types of hypersomnia. The scoring of the MSLT consists of determining the latency to sleep onset (i.e. first 16 continuous seconds of any stage of sleep) for each nap. The mean sleep latency across the 4 or 5 naps is the primary endpoint. Multiple naps are

used for 2 reasons: first, since there are differences in sleep propensity as a function of time of day, and second, because multiple assessments of sleep onset latency reduce variability. The MSLT has been validated against sleep deprivation, sleep disorders, circadian time, and a variety of sedating and alerting dugs [246]. In addition to the standardised methodology, there are well-established reference norms available for healthy volunteers of different ages as well as for a variety of sleep disorders, and the MSLT has been adopted as part of the diagnostic criteria for some sleep disorders [247]. Its utility in the clinical context is both to define and objectively determine the severity of the symptoms of excessive daytime sleepiness, as well as to demonstrate multiple sleep onset REM periods to confirm a diagnosis of narcolepsy.

A modification of the MSLT is the maintenance of wakefulness test (MWT). Variants of the multiple sleep latency tests were originally developed with the hypothesis that the ability to fall asleep and the ability to stay awake represent different physiological states. There is currently no data to support this concept, but the MWT is nonetheless often used in clinical trials (e.g. continuous positive airway pressure, stimulants) [248]. Unlike the MSLT, which benefits from standardised procedures that have been extensively validated as well as the availability of norms, there are significant variations in the execution of the MWT. In the literature there are MWTs where nap tests with durations of 20, 30 or 40 min are performed. Increasingly the test is being performed using a 30-min nap. However, in healthy volunteers, 30 min without a sleep onset is not uncommon. Although originally some MWTs had the subjects sitting in chairs, the MWT is now increasingly performed with subjects lying in a bed with a 45-degree incline. Unlike the MSLT, which is performed in darkness, the MWT is carried out in 'dim light'. Importantly, for the MWT, the subjects are instructed to try to stay awake rather than to try to fall asleep. Despite all of these differences, both of these tests are, in principle, measures of sleep tendency, and qualitatively give similar results.

These tests of sleep tendency are critical in the development of stimulant medications as a measure of efficacy, and can be used for the assessment of side effects of sedatives, as they provide an objective assay. In addition, in clinical practice the MSLT serves as an essential element in the diagnosis of narcolepsy.

EEG Source Localisation

While several neuro-imaging techniques are available to explore brain dynamics in humans, including magnetoencephalography (MEG), positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), the scalp-recorded EEG combined with post hoc source localisation provides the only way to investigate non-invasively cortical activities during undisturbed sleep with high temporal resolution. The minimal requirement for EEG-based source localisation is the recording of the scalp potential field with at least 19 equally distributed electrodes.

The challenge for EEG source localisation methods is to provide a unique solution to the inverse problem, trying to find a physiologically meaningful distribution of putative generators without prior knowledge of the number, location or orientation of the actual sources of the scalp EEG activity. This inverse solution needs also to be independent of the arbitrary choice of the reference electrode.

Low-resolution brain electromagnetic tomography (LORETA) as devised by Pascual-Marqui et al. [249, 250] in 1994 was one of the first attempts to solve both the inverse problem and the reference electrode problem. Meanwhile LORETA has received considerable validation from studies combining it with other more established localisation methods, such as fMRI [251], structural MRI [252] and PET [253]. Further LORETA validation has been based on comparison with the localisation findings from implanted depth electrodes, in a number of studies in epilepsy [254] and on cognitive event-related potentials where the activated brain regions are known a priori [255]. One has to be aware that only cortical activities that make major contributions to the scalp-recorded EEG can be estimated by source localisation, and this means most of the deep active sources in the brain are not detectable. This limiting factor has to be considered not only in the interpretation of electrophysiological neuro-imaging results, but also in studies comparing different neuro-imaging methods or in multimodal neuro-imaging approaches.

In recent years, various new methods have been developed and validated, partly with simulated and partly with real data, to solve the ill-posed inverse problem. These source localisation methods comprise, in addition to LORETA, minimum norm estimates (MNE), weighted MNE, MNE with focal underdetermined system solution (FOCUSS), LORETA with FOCUSS, standardised LORETA (sLORETA), variable resolution electrical tomography (VARETA), quadratic regularisation and spatial regularisation (S-MAP) using dipole intensity gradients, spatiotemporal regularisation (ST-MAP), spatiotemporal modelling, the Backus-Gilbert method, weighted resolution optimisation (WROP), the local autoregressive aver-

age (LAURA), exact LORETA (eLORETA) and depth-weighted minimum norm solution as well as shrinking methods and multi-resolution methods such as S-MAP with iterative focusing, shrinking LORETA-FOCUSS, standardised shrinking LORETA-FOCUSS (SSLOFO) and adaptive standardised LORETA/FOCUSS. Several reviews describe the various methods and their differential source localisation properties [256–260]. In a study on source modelling of sleep slow-waves, LORETA, sLORETA, LAURA and the Bayesian minimum norm were applied to a high-density 256-channel sleep EEG recording [261]. Most interesting from a methodological point of view is the finding that all 4 source localisation methods revealed similar results. The similar findings for the various inverse solutions and the physiologically meaningful results obtained from source localisation based not only on high-density but also on rather low-density EEG recordings [103, 255, 262] clearly justify the utilisation of these methods in pharmaco-sleep studies. However, it needs to be stressed that full details of the applied source localisation method with all its underlying assumptions (e.g. head model, regularisation parameters, solution space and statistical approaches) needs to be presented in detail alongside the results to facilitate proper comparisons between studies.

Sleep Models

Enormous progress was made in research on sleep and chronobiology in the second half of the last century. Although sleep and circadian rhythms are closely interrelated, both fields developed separately at first because different methodologies were used to study them in the beginning. The primary technique used in sleep research is the continuous recording of the EEG and other physiological signals for whole nights. In contrast, the main technique of chronobiological research is the continuous recording of behaviour (activity/rest), combined with selected physiological variables, mainly deep body temperature. Early chronobiological models were based mainly on the interaction of several internal, self-sustained circadian oscillators [263]. A new modelling approach evolved when, in the 1970s and 1980s, long-term sleep recordings were performed for the first time under conditions of temporal isolation, the standard experimental setting of chronobiology. The combined use of sleep-EEG and circadian data led to the conceptualisation of more realistic models.

It is now widely accepted that the sleep-wake cycle can be considered as an active, neurobiologically regulated process [264] and that it can be modelled with two interacting components, one homeostatic and one circadian in nature [265, 266].

The homeostatic, or use-dependent, component can be described as a wake dependent growth function which dissipates during sleep. It reflects a build-up of sleep need during wakefulness and a reverse process during sleep. Its intensity can be estimated from parameters of low frequency EEG activity [265–267].

The circadian component can be described as a time-related sinusoidal function. It reflects a sleep-related process which is dependent on the phase of the central circadian pacemaker in the suprachiasmatic nuclei. Phase and amplitude can be estimated from hormonal parameters like the circadian melatonin profile, or the course of core body temperature [268].

This two-process concept proved to be so successful that it has been expanded gradually to include also modelling of circadian vigilance and performance variations [269–272].

Although the model is well accepted, there is less agreement on the nature of the interaction between the two components. Their combined effect on sleep has been described as a linear (additive) interaction [269]. Experiments with forced desynchrony protocols, which permit systematic dissociation of the circadian and homeostatic parameters [268] and studies with suprachiasmatic nucleus-lesioned animals [273] indicate that both components are to a large extent independent from each other, with an interaction that might be interpreted as being non-linear instead [270, 274].

Moreover, linear interaction fails to adequately model two abundantly described characteristics of the sleep-wake cycle, namely the afternoon nap (or performance dip) zone and the evening wake maintenance (or forbidden sleep) zone [275]. An explicit non-linear approach has been proposed with a multiplicative interaction, which successfully models the four main characteristics of sleep propensity across 24 h [276].

Modelling sleep with a circadian and a homeostatic component strongly influenced the development of new sleep- and vigilance-modulating drugs. Homeostatic parameters like slow-wave sleep or EEG slow-wave activity became a target variable for measuring the effect of hypnotics and other drugs on sleep. At the same time chronobiotic drugs (such as e.g. melatoninergic substances) were developed, which act quite selectively on circadian parameters. It has become increasingly clear that the two major components, which regulate the sleep-wake rhythm, can be influenced separately by pharmacological agents [277].

This modelling approach also impacts the development and selection of adequate study designs. Drugs with a potential influence on sleep, and especially on slow-wave sleep/slow-wave activity, can be tested in the usual setting of sleep studies with specific experimental modifications, if necessary (e.g. using sleep deprivation or sleep reduction). In contrast, the testing of specific effects of potential chronobiotic agents needs designs suitable for the study of chronobiology, such as sleep shift, constant routine, forced desynchrony and others. As an example, Kräuchi et al. [278] tested the effect of an early evening (18.00 h) dose of melatonin and S-20098 on the circadian phase of different physiological parameters under constant routine conditions. In their study the pharmacological agents induced an earlier dim-light melatonin onset as well as a phase shift in indicators of body temperature regulation. There was an earlier increase in distal skin temperature at sleep onset and an earlier decrease in body core temperature during sleep. Such effects are to be expected after the application of chronobiotic but not hypnotic agents. The results also underline the importance of the correct timing of administration of a chronobiotic agent for producing an optimal effect. In a follow-up study the effects of both active agents on body temperature were replicated and it was shown that both substances enhanced REM sleep while the EEG in NREM sleep, as measured by an analysis of EEG power density, remained unaffected [279]. This latter result is in line with the assumption that REM and NREM sleep are regulated by different mechanisms, as suggested by the models cited above.

Sleep Assessment Using Actigraphy

Activity-based sleep-wake monitoring using accelerometer-based devices (actigraphy) has become a major tool in sleep research and sleep medicine [280] as a potential alternative to the neurophysiological techniques described above in some pharmaco-sleep studies, particularly when there is a need to assess the rest-activity cycle over long time periods. Several reviews have established the use of actigraphy as a reliable and valid assessment method to document sleep-wake patterns [32, 281–283], and its strengths and limitations have been thoroughly discussed [280]. Compared to PSG, actigraphy measures the physical manifestations of sleep, rather than its neurological basis, and hence provides only estimates of sleep-related parameters such as SOL, TST, WASO and sleep efficiency in both healthy subjects [284, 285] and insomnia patients [286, 287]. Normative data over several age groups are available. Actigraphy has been shown to distinguish between clinical groups, to identify certain sleep-wake disorders, and to document the effects of various behavioural and medical interventions on sleep-wake patterns [288–290], while limitations have been observed in its ability to detect wakefulness during sleep, and hence to assess sleep patterns in clinical populations with highly fragmented sleep [291]. Finally, actigraphy has the advantage of providing objective information on sleep habits in the patient's natural sleep environment rather than a sleep laboratory [292]. This enables, for example, the examination of sleep patterns during adaptation and re-adaptation to different shift work schedules [293].

Modern actigraphs are based on a solid-state accelerometer that measures acceleration (m/s^2) in 1–3 dimensions. The firmware in the device samples the acceleration (typically 32 times/s) and mathematically summarises the data collected in each epoch (typically 15 s to 1 min). Different devices use different algorithms to calculate 'activity counts'. These include a simple mean of the magnitude of the acceleration above a specified threshold (to account for gravity), a count of the number of peaks or zero-crossings during the epoch, or some variation of these methods [32]. Although actigraphs can be worn anywhere on the body, sleep is usually assessed with a device worn on the non-dominant wrist. There are several different devices available commercially that have been certified as suitable for medical investigation (e.g. through CE marking as a medical device, or FDA medical device registration) and that can collect data for up to several months, depending upon the epoch length. Many consumer devices are also commercially available which do not carry medical certification. Although some consumer devices may offer good performance, such devices should, in general, be used with caution in clinical trials as their reproducibility, particularly between devices, may not be well characterised.

A methodological issue repeatedly raised is the lack of standard equipment, analytical methods and reporting, thereby impeding the comparison of findings and conclusions across studies. The major methodological challenge when using actigraphy concerns the procedures used for data sampling, processing and analysis [294]. This problem is worsened by the difficulty faced when trying to compare the performance of different devices and algorithms [285, 295]. To improve this situation, some investigators have used more direct activity endpoints, such as mean activity during the night, to avoid having to make assumptions about the subject's sleep/wake status [296, 297].

While actigraphy has predominantly been used in the field of sleep research and chronobiology, there is a growing body of evidence for its utility to measure drug effects in clinical trials, as reviewed by Stanley [298]. The effects of psycho-active drugs have been described over long time periods by comparing treatments [299] or by measuring nocturnal and daytime motor activity [300]. Actigraphy has been used to assess and quantify the daytime sedative effects of tricyclic antidepressants, which were positively correlated with both the subjective ratings of tiredness and the impaired cognitive and psychomotor performance [301], to investigate the effect of hypnotics on nocturnal motor activity [296], and to quantify the night-time sleep-promoting effects of zolpidem administered in a clinical research unit environment [297]. Actigraphy has been shown to be sensitive enough to detect sleep changes during periods of temazepam administration and withdrawal in patients suffering from insomnia [289], to trace low physical activity and altered sleep in patients with schizophrenia treated with olanzapine or risperidone [302] and to measure improvements in total sleep time in patients with restless leg syndrome treated with pregabalin [303]. The effects of melatonin and zopiclone in a placebo-controlled protocol using actigraphy with air crew members coping with adaptation by transatlantic flights and time zone changes confirmed the results previously obtained in a laboratory setting [304].

In the context of clinical trials, one of the major advantages of actigraphy lies in the fact that it is possible to continuously measure and monitor activity over several weeks (longitudinal monitoring). As sleep deficits alter sleep architecture and thereby interfere with drug effects, actigraphy is a suitable method to identify subjects with irregular sleep patterns and hence to exclude them from PSG studies. Therefore, where this is a particular concern, the use of actigraphy should be considered for at least 1 week prior to a stationary PSG recording session to document the habitual sleep-wake activities of the subjects involved and to detect possible underlying sleep problems, or non-compliance with the lifestyle guidelines defined in the study protocol.

Practice parameters have been developed by task forces commissioned by the AASM for the use of actigraphy in the assessment of sleep, circadian rhythms and sleep disorders [305, 306]. The recommendations are based on a comprehensive review of the literature and should serve as guidance for the appropriate use of actigraphy in pharmaco-sleep studies (clinical trials). To ensure that results of a study can be compared with those of other published studies, publications should report the set of metadata listed in table 6 as a minimum along with the study results.

Although PSG is the acknowledged 'gold standard' for assessing sleep, one needs to consider that its usefulness in some situations is limited by the complexity of the equipment, its cost and hence its inability to assess long-term sleep-wake patterns. Patient-reported outcomes or sleep questionnaires have the advantage of being low-cost and they can be used for many nights at home to record the subjects' perception of their sleep, but the disadvantage of being subjective. The objectivity of actigraphy combined with the ability to make long-term assessments of sleep-wake patterns in the home environment at a lower cost than PSG makes it an attractive method for assessing and quantifying the effects of treatments on sleep. However, actigraphy does not measure the neurological aspects of sleep but rather it measures immobility/mobility as physiological and behavioural manifestations of the sleep-wake cycle, and this leads to limitations. In the absence of the perfect sleep measurement system, investigators need to be mindful of the strengths and weaknesses of each assessment tool when choosing which method to use in a particular study.

EEG Biomarkers and Translational Medicine

According to the definition formalised by the Biomarkers Definitions Working Group [307], a biomarker is 'a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention', whereas a surrogate endpoint is 'a biomarker that is intended to substitute for a clinical endpoint'. In addition, the term endophenotype refers to 'a set of quantitative, heritable, trait-related, state-independent and family-associated characteristics of a disorder typically assessed by biochemical, endocrine, neuroanatomical, neurophysiological, neuropsychological, and other methods' [308]. In other words, biomarkers are objectively measured indices of pharmacological responses (or biological processes) that are quantifiable, precise, and reproducible. As such, biomarkers may be used to answer a number of questions in clinical research (markers for diagnosis, progression or staging of a disease, subtyping of patients) or in the drug development process (efficacy markers distinguishing between treatment responders and non-responders, effectiveness markers which can be used to monitor therapeutic improvement, markers that can be used for assessing bioavailability or central penetration of the drug, or for PK/PD modelling,

Table 6. Minimum set of requirements for the use of actigraphy in clinical trials

Instrumentation
• Registered trademark of the devices
• Version of the software embedded in the devices
• Placement of the actigraphs (worn on the wrist of the non-dominant hand with the monitor in the upside position is recommended)
– For long-time and repeated recordings a given subject should always use the same actigraph (recommended, unless data confirming sufficient interdevice consistency is available and appropriate quality control is included in the study)
– Actigraphs are classified as a class I medical device and must carry a certification (at least CE classified in Europe or FDA for the USA)
Sleep/wake diary
– Actigraphy recording always combined with a diary
– Time and date of every removal of the unit has to be documented (e.g. when taken off during bathing and showering)
Data sampling
• Duration of data collection (minimum should be 1 week)
• Epoch length (below 1 min, 30 s being recommended)
• Mode of data collection (e.g. ZCM, TAT, PIM, or TRI mode)
Data processing and analysis
– Validated procedure for initialising the devices, downloading datasets and storing files
– Quality control by visual inspection of the raw data to detect technical failures and other anomalies
– Documented procedure for handling missing data
• If the scoring is accomplished by different individuals, reliability procedure and percent agreement between scorers is to be reported
• Software (and version) used for data scoring
• Algorithm used for data processing (e.g. filtering, smoothing)
• Algorithm used to delineate sleep/wake and activity/rest segments
• Selected variables
Reporting
• Graphs should report days of the week and months (by name)
• Number of sleep and activity intervals included (single, multiple, intra-individual mean, group mean)
• If recordings were impacted by clock time change due to daylight saving and a manual correction made, this must be reported
Extracted endpoints (parameters)
• aTRT = Total recording time (min)
• aTST = Total total sleep time (min)
• aWASO = Time awake after sleep onset (min)
• aSE = Sleep efficiency (%) expressed as aTST/aTRT
• When reporting group (i.e. treatment) results, the endpoints should be reported using mean, standard deviation, median, first and third quartiles, and trimean

The summary includes items of procedural nature (–) and the set of information to be reported in publications (•).

ZCM = Zero Crossing Mode; TAT = Time Above Threshold; PIM = Proportional Integrating Measure; TRI = Trimode (ZCM, TAT and PIM). The prefix 'a' used to identify the actigraphy-extracted endpoints is aimed at ensuring that actigraphy- vs. PSG-derived parameters are easily identified.

markers reporting on the pharmacological mechanism of action or on potential safety risks of a drug). Biomarkers play an increasingly important role in the drug development process. If they have translational validity, meaning that they provide measures of drug action that are consistent across species boundaries, they are particularly useful in guiding the design of clinical study programmes based on preclinical findings.

Pharmaco-sleep studies are important in characterising the effect of CNS-active drugs since a favourable (or unfavourable) influence on sleep initiation, continuity and architecture has a major impact on the clinical use of any drug. Unwanted drug effects on sleep are not necessarily problematic for drug development, especially when the effects are transient and diminish with repeated dosing, provided this transient effect is understood. In any case, pharmaco-sleep effects are very suitable PD biomarkers for CNS-active drugs, which can be monitored easily in acute dosing studies. An important advantage of pharmaco-sleep studies is the highly preserved sleep ar-

chitecture across most mammalian species, despite differences in sleep duration and timing, as well as the consistency of most drug effects on sleep architecture, making translational pharmaco-sleep studies a first class provider of translational biomarkers for CNS-active drugs. Sleep-EEG characteristics prove to be very sensitive indices of a drug's CNS activity that offer interesting biomarker options for translational medicine including PK/PD relationships [309, 310]. Although sleep is disrupted in several psychiatric disorders (and their animal models), pharmaco-sleep biomarkers do not need to be disease specific in order to be valuable in drug discovery. Drug-induced changes in sleep-wake patterns and associated EEG spectra in healthy animals and human volunteers have been successfully used to characterise novel CNS-active drugs as potential therapeutic agents for different CNS disorders in a probabilistic manner (71–95%) [311, 312].

Over the past several decades there has been considerable effort in exploring the utility of sleep parameters and EEG measures of sleep as reliable and sensitive biomarkers in many areas, despite an inherent lack of specificity which necessitates careful interpretation [for a review, see 22, 313], including:

- major depressive disorder [314, 315]; the high sensitivity of sleep architecture changes in depression makes it one of the most important biological markers in psychiatry. In this way sleep changes have been used as an endophenotype and/or vulnerability biomarkers in family studies of depression [316–318];
- prediction of antidepressant treatment responses in major depressive disorder patients [314, 319, 320] and more generally as a biomarker for the effects of all classes of antidepressants [20, 321–323];
- posttraumatic stress disorder [324–326], schizophrenia [327, 328], and other psychiatric and neurodegenerative disorders [329, 330];
- self-administration of sleep-promoting drugs [331];
- genotype-dependent difference in the response to sleep deprivation [332, 333];
- effect of illicit recreational drugs (cocaine, ecstasy and marijuana) upon sleep [334];
- general cognitive and learning abilities [102].

It is clear that sleep disturbances are core symptoms of mood disorders, and as such they are an integral part of the diagnostic criteria for major depression. Several sleep-EEG characteristics have been examined in terms of their predictive utility and a number of studies have demonstrated strong evidence that one of the most important markers of major depressive disorder is a shorter REM latency, a prolonged first REM period, and an increased REM density [22]. Moreover, the persistence of these deviant measures was found to be a prognostic indicator of an unfavourable course of the disease [335]. These changes could be, therefore, considered intermediate phenotypes between the genetic and biochemical cause of the disease and the complex psychological and behavioural manifestations of the symptoms of depression. Most antidepressants suppress REM sleep in humans and animals, but REM suppression is not a prerequisite for a substance to act as an antidepressant. Thus, the identification of a sleep-EEG variable or a cluster of such variables which can be used to screen substances for their antidepressant potential and/or for the further characterisation of subgroups of depressed patients with different therapeutic responses to antidepressant drugs remain important research targets [314].

It should be noted, however, that not all results of clinical pharmaco-sleep studies are in agreement, mostly due to various methodological issues and diagnostic uncertainties in psychiatric patients: the number of consecutive nights patients are studied; whether they are studied in their home environment or in a sleep laboratory; the way that the time between sleep onset and REM sleep onset (REM sleep latency) is determined; the definition of increased REM density; whether concurrent use of psychotropic drugs is allowed and whether there are differences in the required duration of withdrawal from psychotropic medications; variation in sleep schedules; severity of the illness, and problems caused by potential differences in patient populations due to the overlap of symptoms among various disorders as described in the Research Diagnostic Criteria, the DSM-IV, as well as in other classification systems for mental disorders [13].

Animal models of normal and disordered sleep generally focus on the primary endpoints of clinical interest – sleep onset, sleep architecture, sleep EEG spectra – and sometimes also next-day effects, such as sleepiness, cognition, metabolism etc. Few comparative studies have quantified which sleep variables of psychotropic drugs are more or less similar between species. Recent studies have extended these analyses with PK/PD modelling, showing that sleep quantitative EEG and PSG characterisation of novel compounds in rodents and primates are useful models to predict PK/PD relationships for these biomarkers in humans [336].

Psychiatric and neurological disorders are associated with a constellation of characteristic pathophysiological changes. These can result in changes in EEG activity in the resting state or during certain phases of the sleep-

wake cycle, or in alterations to the brain's electrophysiological response to sensory stimuli or endogenous cognitive processing. These abnormal neurophysiological mechanisms may be considered as endophenotypes, and can be utilised for developing disease-relevant animal models and translatable biomarkers [337–340]. Transgenic animal models focused on sleep and circadian rhythms offer the potential to identify novel genes and genetic pathways underlying sleep disorders. Translatability of mouse mutants appears to be robust as many of the mutants expected to have circadian effects or sleep alterations, based on established human pharmacology, do in fact show effects that are consistent with pharmacology in man and animal models [336].

Successful application of translational research led to the identification of animal gene polymorphisms associated with altered circadian rhythmicity or sleep homeostasis which have subsequently been found to alter sleep in healthy people or have been associated with circadian rhythm sleep disorders, narcolepsy or restless leg syndrome [309, 341]. There are also several animal models potentially suitable for translational research of sleep disorders, including those for primary insomnia [342, 343], obstructive sleep apnoea [344], restless leg syndrome [345], narcolepsy [346], and in disease-related sleep disturbance such as Huntington's disease [347], Parkinson's disease [348, 349], Alzheimer's disease [350, 351], bipolar disorder [352, 353], and epilepsy [354].

In conclusion, sleep is the physiological manifestation of a complex interplay between several of the most important neurotransmitter systems in the brain. Consequently, sleep-EEG variables have proven to be suitable as translational biomarkers for sleep-related changes in neuropsychiatric disorders, prediction of treatment response and for the elucidation of the mechanism of action of drugs on the CNS in general, as well as on the sleep/wake architecture in particular. Though the application of sleep EEG is often restricted to studies with small numbers of patients, the similar physiology of sleep regulation in rodents and humans suggests that changes in sleep microarchitecture could be used also as translatable biomarkers for the evaluation of novel treatment targets [355].

However, despite the substantial progress in translational research of sleep disturbances in various disorders, and sleep disorders in particular, several obstacles remain for animal sleep research including:

- the need for surgical implantation of EEG and EMG electrodes in animals and the time-consuming interpretation of recordings;
- the lack of fully objective, automatic, high throughput and yet sensitive methods to assess sleep in hundreds of animals at a time;
- a lack of standardisation of animal EEG methodology and sleep classification across academic research centres and the pharmaceutical industry, but also across different animal species;
- limited potential for deriving spatial information from animal sleep EEG due to the small number of electrodes;
- the lack of fully automated and objective artefact detection and reduction algorithms in animal sleep EEG spectral analysis;
- the lack of specific guidelines for animal quantitative EEG and animal pharmaco-sleep EEG recording;
- the need for more research to characterise the sleep-disruptive effects of specific environmental factors or methodological procedures on animal sleep-waking behaviour;
- lack of a full description of the anatomical, pharmacological and genetic correlates of sleep and waking behaviour; despite some recent advances in this area, there is still much to be uncovered.

In this context, it becomes crucial that efforts are made in the standardisation of experimental conditions and in the development of protocols facilitating the comparison of data collected in both man and animals between different centres. Specific guidelines for preclinical, animal pharmaco-EEG recording and analysis are in preparation for publication by the IPEG.

Application of novel recording devices and advanced computational techniques in translational research will provide more detailed data on the micro-architecture of sleep and it is expected that more subtle sleep changes and accurate biomarkers will be detected with further optimisation of the translational value of pharmaco-sleep EEG in the near future.

Conclusion

Quantitative EEG and related methods have the potential to offer reliable biomarkers and will, in view of the recent developments of quantitative EEG technology, play an increasingly important role in preclinical research and in all phases of clinical drug development. The evaluation and quantification of drug effects using EEG, sleep and evoked potentials/event-related potentials (EP/ERP) provides a set of methods to capture the pharmacological activity, therapeutic benefits and potential

adverse effects that a drug induces in diverse patient populations. By combining various methods and their respective strengths, it is reasonable to argue that they will provide a more complete characterisation of the spectrum of pharmacological CNS responses of known and novel therapeutic drugs [22].

In this context, it is mandatory to enhance and standardise recording, analysis and study design methodologies to facilitate the comparability of data across laboratories both in academia and in industry. To this end, investigators using pharmaco-sleep methodology are urged to refer to and comply with the guidelines presented here when designing and conducting studies, and to reference the present paper when publishing study results.

Acknowledgements

The IPEG Pharmaco-Sleep Guidelines Committee wishes to thank the following for their critical review of this paper and helpful comments: **Angela J. Campbell** (WellSleep Sleep Investigation Centre, Department of Medicine, University of Otago, Wellington, New Zealand – Australasian Sleep Association); **Eric A. Nofzinger** (Sleep Neuroimaging Research Program, Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pa., USA); **Verne Pitman** (Clinical Sciences, Pfizer Primary Care, Groton, Conn., USA); **Timothy A. Roehrs** (Sleep Disorders and Research Center, Henry Ford Hospital, Detroit, Mich., USA); **Michael J. Thorpy** (Montefiore Medical Center, Bronx, N.Y., USA); **William Vennart** (Pharmatherapeutics Precision Medicine, Pfizer Ltd, Sandwich, UK); **James K. Walsh** (Sleep Medicine and Research Center, St. Luke's Hospital, Chesterfield, Mo., USA).

References

1 Jobert M, Wilson FJ, Ruigt GSF, Brunovsky M, Prichep LS, Drinkenburg WHIM, et al: Guidelines for the recording and evaluation of pharmaco-EEG data in man – International Pharmaco-EEG Society (IPEG). Neuropsychobiology 2012;66:201–220.

2 Stille G, Herrmann WM: Guidelines for pharmaco-EEG studies in man; in Herrmann WM (ed): Electroencephalography in Drug Research. Stuttgart, Fischer, 1982, pp 12–19.

3 Dumermuth G, Ferber G, Herrmann WM, Hinrichs H, Künkel H: Recommendations for standardization of data acquisition and signal analysis in pharmaco-electroencephalography. Neuropsychobiology 1987;17:213–218.

4 Pivik RT, Broughton RJ, Coppola R, Davidson RJ, Fox N, Nuwer MR: Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. Psychophysiology 1993;30:547–558.

5 Duffy FH, Hughes JR, Miranda F, Bernad P, Cook P: Status of quantitative EEG (QEEG) in clinical practice. Clin Electroencephalogr 1994;25:6–22.

6 Nuwer MR: Assessment of digital EEG, quantitative EEG, and EEG brain mapping: report of the American Academy of Neurology and the American Clinical Neurophysiology Society. Neurology 1997;49:277–292.

7 Winterer G, Dorn H, Herrmann WM, Gallhofer B, Bauer U, Hegerl U, Ihl R, Dierks T, Maurer K: The AMDP modules I-IV: recommendations for a standardized acquisition of EEG data in psychiatry. Association for Methodology and Documentation in Psychiatry. Neuropsychobiology 1997;36:100–106.

8 Rechtschaffen A, Kales A: A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Los Angeles, University of California, Brain Information Service/Brain Research Institute, 1968.

9 Kubicki S, Herrmann WM, Höller L: Critical comments on the rules by Rechtschaffen and Kales concerning the visual evaluation of EEG sleep records; in Kubicki S, Herrmann WM (eds): Methods of Sleep Research. Stuttgart, Fischer, 1985, pp 19–35.

10 Kubicki S, Herrmann WM: The future of computer-assisted investigation of the polysomnogram: sleep microstructure. J Clin Neurophysiol 1996;13:285–294.

11 Himanen SL, Hasan J: Limitations of Rechtschaffen and Kales. Sleep Med Rev 2000;4:149–167.

12 Iber C, Ancoli-Israel S, Chesson AL, Quan SF, for the American Academy of Sleep Medicine: The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. Westchester, American Academy of Sleep Medicine, 2007.

13 Franzen PL, Buysse DJ: Sleep in psychiatric disorders; in Chokroverty S (ed): Sleep Disorders Medicine: Basic Science, Technical Considerations, and Clinical Aspects, ed 3. Philadelphia, Elsevier Books, 2009, pp 538–549.

14 Finelli LA, Achermann P, Borbély AA: Individual 'fingerprints' in human sleep EEG topography. Neuropsychopharmacology 2001; 25(suppl):S57–S62.

15 Van Beijsterveldt CEM, van Baal GCM: Twin and family studies of the human electroencephalogram: a review and a meta-analysis. Biol Psychol 2002;61:111–138.

16 Smit DJ, Posthuma D, Boomsma DI, Geus EJ: Heritability of background EEG across the power spectrum. Psychophysiology 2005;42: 691–697.

17 Smit CM, Wright MJ, Hansell NK, Geffen GM, Martin NG: Genetic variation of individual alpha frequency (IAF) and alpha power in a large adolescent twin sample. Int J Psychophysiol 2006;61:235–243.

18 De Gennaro L, Marzano C, Fratello F, Moroni F, Pellicciari MC, Ferlazzo F, Costa S, Couyoumdjian A, Curcio G, Sforza E, Malafosse A, Finelli LA, Pasqualetti P, Ferrara M, Bertini M, Rossini PM: The electroencephalographic fingerprint of sleep is genetically determined: a twin study. Ann Neurol 2008; 64:455–460.

19 Obermeyer WH, Benca RM: Effects of drugs on sleep. Neurol Clin 1996;14:827–840.

20 Rijnbeek B, de Visser SJ, Franson KL, Cohen AF, van Gerven JM: REM sleep effects as a biomarker for the effects of antidepressants in healthy volunteers. J Psychopharmacol 2003;17:196–203.

21 De Martinis NA, Winokur A: Effects of psychiatric medications on sleep and sleep disorders. CNS Neurol Disord Drug Targets 2007;6:17–29.

22 Leiser SC, Dunlop J, Bowlby MR, Devilbiss DM: Aligning strategies for using EEG as a surrogate biomarker: a review of preclinical and clinical research. Biochem Pharmacol 2011;81:1408–1421.

23 Tissot R: The effects of certain drugs on the sleep cycle in man. Prog Brain Res 1965;18: 175–177.

24 Oswald I, Jones HS, Mannerheim JE: Effects of two slimming drugs on sleep. Br Med J 1968;i:796–799.

25 Evans JI, Lewis SA: Sleep studies in early delirium and during drug withdrawal in normal subjects and the effect of phenothiazines on such states. Electroencephalogr Clin Neurophysiol 1968;25:508–509.

26 Kupfer DJ, Spiker DG, Coble P, McPartland RJ: Amitriptyline and EEG sleep in depressed patients. I. Drug effect. Sleep 1978;1: 149–159.
27 Schlösser R, Röschke J, Rossbach W, Benkert O: Conventional and spectral power analysis of all-night sleep EEG after subchronic treatment with paroxetine in healthy male volunteers. Eur Neuropsychopharmacol 1998;8: 273–278.
28 Paterson LM, Nutt DJ, Ivarsson M, Hutson PH, Wilson SJ: Effects on sleep stages and microarchitecture of caffeine and its combination with zolpidem or trazodone in healthy volunteers. J Psychopharmacol 2009;23:487–494.
29 Stepanski E, Lamphere J, Badia P, Zorick F, Roth T: Sleep fragmentation and daytime sleepiness. Sleep 1984;7:18–26.
30 Walsh JK, Vogel GW, Scharf M, Erman M, William Erwin C, Schweitzer PK, Mangano RM, Roth T: A five week, polysomnographic assessment of zaleplon 10 mg for the treatment of primary insomnia. Sleep Med 2000; 1:41–49.
31 Roth T, Stubbs C, Walsh JK: Ramelteon (TAK-375), a selective MT1/MT2-receptor agonist, reduces latency to persistent sleep in a model of transient insomnia related to a novel sleep environment. Sleep 2005;28:303–307.
32 Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP: The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003;26:342–392.
33 Shambroom JR, Fábregas SE, Johnstone J: Validation of an automated wireless system to monitor sleep in healthy adults. J Sleep Res 2012;21:221–230.
34 Patel MR, Davidson TM: Home sleep testing in the diagnosis and treatment of sleep disordered breathing. Otolaryngol Clin North Am 2007;40:761–784.
35 Chediak AD: Why CMS approved home sleep testing for CPAP coverage. J Clin Sleep Med 2008;4:16–18
36 Collop NA: Home sleep testing: it is not about the test. Chest 2010;138:245–246.
37 Mendez MO, Corthout J, Van Huffel S, Matteucci M, Penzel T, Cerutti S, Bianchi AM: Automatic screening of obstructive sleep apnea from the ECG based on empirical mode decomposition and wavelet analysis. Physiol Meas 2010;31:273–289.
38 ICH guidelines. http://www.ich.org.
39 Wang T, Jacobson-Kram D, Pilaro AM, Lapadula D, Jacobs A, Brown P, Lipscomb J, McGuinn WD: ICH guidelines: inception, revision, and implications for drug development. Toxicol Sci 2010;118:356–367.
40 Budhiraja R, Roth T, Hudgel DW, Budhiraja P, Drake CL: Prevalence and polysomnographic correlates of insomnia comorbid with medical disorders. Sleep 2011;34:859–867.
41 Heather N: The importance of keeping regular: accurate guidance to the public on low-risk drinking levels. Alcohol Alcohol 2009; 44:226–228.
42 Higdon JV, Frei B: Coffee and health: a review of recent human research. Crit Rev Food Sci Nutr 2006;46:101–123.
43 Agnew HW, Webb WB, Williams RL: The first night effect: an EEG study of sleep. Psychophysiology 1966;2:263–266.
44 Mendels J, Hawkins DR: Sleep laboratory adaptation in normal subjects and depressed patients ('first night effect'). Electroencephalogr Clin Neurophysiol 1967;22:556–558.
45 Le Bon O, Staner L, Hoffman G, Dramaix M, San Sebastian I, Murphy JR, Kentos M, Pelc I, Linkowski P: The first-night effect may last more than one night. J Psychiatry Res 2001; 35:165–172.
46 Curcio G, Ferrara M, Piergianni A, Fratello F, De Gennaro L: Paradoxes of the first-night effect: a quantitative analysis of antero-posterior EEG topography. Clin Neurophysiol 2004;115:1178–1188.
47 Saletu B, Klösch G, Gruber G, Anderer P, Udomratn P, Frey R: First-night-effects on generalized anxiety disorder (GAD)-based insomnia: laboratory versus home sleep recordings. Sleep 1996;19:691–697.
48 Kajimura N, Kato M, Sekimoto M, Watanabe T, Takahashi K, Okuma T, Mizuki Y, Yamada M: A polysomnographic study of sleep patterns in normal humans with low- or high-anxiety personality traits. Psychiatry Clin Neurosci 1998;52:317–320.
49 Riedel BW, Winfield CF, Lichstein KL: First night effect and reverse first night effect in older adults with primary insomnia: does anxiety play a role? Sleep Med 2001;2:125–133.
50 Suetsugi M, Mizuki Y, Yamamoto K, Uchida S, Watanabe Y: The effect of placebo administration on the first-night effect in healthy young volunteers. Prog Neuropsychopharmacol Biol Psychiatry 2007;31:839–847.
51 Coates TJ, George JM, Killen JD, Marchini E, Hamilton S, Thorensen CE: First night effects in good sleepers and sleep-maintenance insomniacs when recorded at home. Sleep 1981;4:293–298.
52 Toussaint M, Luthringer R, Schaltenbrand N, Carelli G, Lainey E, Jacqmin A, Muzet A, Macher JP: First-night effect in normal subjects and psychiatric inpatients. Sleep 1995; 18:463–469.
53 Toussaint M, Luthringer R, Schaltenbrand N, Nicolas A, Jacqmin A, Carelli G, Gresser J, Muzet A, Macher JP: Changes in EEG power density during sleep laboratory adaptation. Sleep 1997;20:1201–1207.
54 Browman CP, Cartwright RD: The first-night effect on sleep and dreams. Biol Psychiatry 1980;15:809–812.
55 Toussaint M, Luthringer R, Staner L, Muzet A, Macher JP: Changes in EEG power density during sleep laboratory adaptation in depressed inpatients. Biol Psychiatry 2000;47: 626–633.
56 Woodward SH, Bliwise DL, Friedman MJ, Gusman FD: First night effects in post-traumatic stress disorder inpatients. Sleep 1996; 19:312–317.
57 Selwa LM, Marzec ML, Chervin RD, Weatherwax KJ, Vaughn BV, Foldvary-Schaefer N, Wang L, Song Y, Malow BA: Sleep staging and respiratory events in refractory epilepsy patients: is there a first night effect? Epilepsia 2008;49:2063–2068.
58 Ward TM, Brandt P, Archbold K, Lentz M, Ringold S, Wallace CA, Landis CA: Polysomnography and self-reported sleep, pain, fatigue, and anxiety in children with active and inactive juvenile rheumatoid arthritis. J Pediatr Psychol 2008;33:232–241.
59 Le Bon O, Minner P, Van-Moorsel C, Hoffmann G, Gallego S, Lambrecht L, Pelc I, Linkowski P: First-night effect in the chronic fatigue syndrome. Psychiatry Res 2003; 120:191–199.
60 Le Bon O, Hoffmann G, Tecco J, Staner L, Noseda A, Pelc I, Linkowski P: Mild to moderate sleep respiratory events: one negative night may not be enough. Chest 2000;118: 353–359.
61 Scholle S, Scholle H-C, Kemper A, Glaser S, Rieger B, Kemper G, Zwacka G: First night effect in children and adolescents undergoing polysomnography for sleep-disordered breathing. Clin Neurophysiol 2003;114: 2138–2145.
62 Verhulst SL, Schrauwen N, De Backer WA, Desager KN: First night effect for polysomnographic data in children and adolescents with suspected sleep disordered breathing. Arch Dis Child 2006;91:233–237.
63 Erman MK, Loewy D, Scharf MB: Comparison of temazepam 7.5 mg with temazepam 15 mg for the treatment of transient insomnia. Curr Med Res Opin 2004;20:441–449.
64 Rosenberg R, Roth T, Scharf MB, Lankford DA, Farber R: Efficacy and tolerability of indiplon in transient insomnia. J Clin Sleep Med 2007;3:374–379.
65 Roth T, Heith Durrence H, Jochelson P, Peterson G, Ludington E, Rogowski R, Scharf M, Lankford A: Efficacy and safety of doxepin 6 mg in a model of transient insomnia. Sleep Med 2010;11:843–847.
66 Lorenzo JL, Barbanoj MJ: Variability of sleep parameters across multiple laboratory sessions in healthy young subjects: the 'very first night effect'. Psychophysiology 2002;39: 409–413.
67 Thornby JI, Mauk JL, Karacan I: A further examination of the adjustment/readjustment-night phenomena. Sleep Res 1979;8:152.

68 Wyatt JK, Bootzin RR, Bell IR, Kuo TF, Anthony JL, Rider SP: First night effects across multiple 2-night recordings in older adults. Sleep Res 1995;24:503.
69 Scharf MB, Kales A, Bixler EO: Readaptation to the sleep laboratory in insomniac subjects. Psychophysiology 1975;12:412–415.
70 Stepanski E, Roehrs T, Saab P, Zorick F, Roth T: Readaptation to the laboratory in long-term sleep studies. Bull Psychonom Soc 1981; 17:224–226.
71 Sharpley AL, Solomon RA, Cowen PJ: Evaluation of first night effect using ambulatory monitoring and automatic sleep stage analysis. Sleep 1988;11:273–276.
72 Edinger JD, Fins AI, Sullivan RJ Jr, Marsh GR, Dailey DS, Hope TV, Young M, Shaw E, Carlson D, Vasilas D: Sleep in the laboratory and sleep at home: comparisons of older insomniacs and normal sleepers. Sleep 1997; 20:1119–1126.
73 Herbst E, Metzler TJ, Lenoci M, McCaslin SE, Inslicht S, Marmar CR, Neylan TC: Adaptation effects to sleep studies in participants with and without chronic posttraumatic stress disorder. Psychophysiology 2010;47:1127–1133.
74 Wauquier A, van Sweden B, Kerkhof GA, Kamphuisen HA: Ambulatory first night sleep effect recording in the elderly. Behav Brain Res 1991;42:7–11.
75 Edinger JD, Marsh GR, McCall WV, Erwin CW, Lininger AW: Sleep variability across consecutive nights of home monitoring in older mixed DIMS patients. Sleep 1991;14:13–17.
76 Kemp B, Olivan J: European data format 'plus' (EDF+), an EDF alike standard format for the exchange of physiological data. Clin Neurophysiol 2003;114:1755–1761.
77 Nilsson J, Panizza M, Hallett M: Principles of digital sampling of a physiologic signal. Electroencephalogr Clin Neurophysiol 1993; 89:349–358.
78 Nuwer MR, Comi G, Emerson R, Fuglsang-Frederiksen A, Guerit J-M, Hinrichs H, Ikeda A, Luccas FJC, Rappelsburger P: IFCN standards for digital recording of clinical EEG. Electroencephalogr Clin Neurophysiol 1998;106:259–261.
79 Jobert M: Recording EEG and EP data. Methods Find Exp Clin Pharmacol 2002; 24(suppl C):7–16.
80 Maus D, Epstein CM, Herman ST: Digital EEG; in Schomer DL, Lopes da Silva FH (eds): Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, ed 6, rev. Philadelphia, Lippincott Williams & Wilkins, 2011, pp 119–141.
81 Jasper HH: The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol 1958;10:371–375.
82 Chatrian GE, Lettich E, Nelson PL: Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. Am J EEG Technol 1985;25:83–92.
83 Rodenbeck A, Binder R, Geisler P, Danker-Hopfe H, Lund R, Raschke F, Weess H-G, Schulz H: A review of sleep EEG patterns. I. A compilation of amended rules for their visual recognition according to Rechtschaffen and Kales. Somnologie 2006;10:159–175.
84 Boukadoum AM, Ktonas PY: EOG-based recording and automated detection of sleep rapid eye movements: a critical review, and some recommendations. Psychophysiology 1986;23:598–611.
85 Chokroverty S, Radtke R, Mullington J: Polysomnography: technical and clinical aspects; in Schomer DL, Lopes da Silva FH (eds): Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, ed 6. Philadelphia, Lippincott Williams & Wilkins, 2011, pp 817–864.
86 Tallgren P, Vanhatalo S, Kaila K, Voipio J: Evaluation of commercially available electrodes and gels for recording of slow EEG potentials. Clin Neurophysiol 2005;116:799–806.
87 Kemp B: Measurement of sleep. Prog Brain Res 2010;185:21–35.
88 Dworetzky B, Herman ST, Tatum WO IV: Artifacts of recording; in Schomer DL, Lopes da Silva FH (eds): Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, ed 6. Philadelphia, Lippincott Williams & Wilkins, 2011, pp 239–266.
89 Anderer P, Roberts S, Schlögl A, Gruber G, Klösch G, Herrmann W, Rappelsberger P, Filz O, Barbanoj MJ, Dorffner G, Saletu B: Artifact processing in computerized analysis of sleep EEG – a review. Neuropsychobiology 1999;40:150–157.
90 Monroe LJ: Inter-rater reliability of scoring EEG sleep records. Paper read at the Association for Psychophysiological Study of Sleep Meeting, Santa Monica, California, April 1967 (abstract). Psychophysiology 1968;4:370–371.
91 Schulz H: Phasic or transient? Comment on the terminology of the AASM manual for the scoring of sleep and associated events. J Clin Sleep Med 2007;3:752.
92 Parrino L, Ferri R, Zucconi M, Fanfulla F: Commentary from the Italian Association of Sleep Medicine on the AASM manual for the scoring of sleep and associated events: for debate and discussion. Sleep Med 2009;10:799–808.
93 Danker-Hopfe H, Anderer P, Zeitlhofer J, Boeck M, Dorn H, Gruber G, Heller E, Loretz E, Moser D, Parapatics S, Saletu B, Schmidt A, Dorffner G: Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. J Sleep Res 2009;18:74–84.
94 Moser D, Anderer P, Gruber G, Parapatics S, Loretz E, Boeck M, Klösch G, Heller E, Schmidt A, Danker-Hopfe H, Saletu B, Zeitlhofer J, Dorffner G: Sleep classification according to AASM and Rechtschaffen & Kales: effects on sleep scoring parameters. Sleep 2009;32:139–149.
95 Silber MH, Ancoli-Israel S, Bonnet MH, Chokroverty S, Grigg-Damberger MM, Hirshkowitz M, Kapen S, Keenan SA, Kryger MH, Penzel T, Pressman MR, Iber C: The visual scoring of sleep in adults. J Clin Sleep Med 2007;3:121–131.
96 Gaillard JM, Blois R: Spindle density in sleep of normal subjects. Sleep 1981;4:385–391.
97 Crowley K, Trinder J, Kim Y, Carrington M, Colrain IM: The effects of normal aging on sleep spindle and K-complex production. Clin Neurophysiol 2002;113:1615–1622.
98 Knoblauch V, Martens WL, Wirz-Justice A, Cajochen C: Human sleep spindle characteristics after sleep deprivation. Clin Neurophysiol 2003;114:2258–2267.
99 Kubicki S, Haag-Wüsthoff C, Röhmel J, Herrmann WM, Scheuler W: The pharmacodynamic influence of three benzodiazepines on rapid eye movements, K-complexes and sleep spindles in healthy volunteers. Hum Psychopharmacol 1988;3: 247–255.
100 Jobert M, Poiseau E, Jähnig P, Schulz H, Kubicki S: Pattern recognition by matched filtering: an analysis of sleep spindle and K-complex density under the influence of lormetazepam and zopiclone. Neuropsychobiology 1992;26:100–107.
101 Schabus M, Gruber G, Parapatics S, Sauter C, Klösch G, Anderer P, Klimesch W, Saletu B, Zeitlhofer J: Sleep spindles and their significance for declarative memory consolidation. Sleep 2004;27:1479–1485.
102 Schabus M, Hödlmoser K, Gruber G, Sauter C, Anderer P, Klösch G, Parapatics S, Saletu B, Klimesch W, Zeitlhofer J: Sleep spindle-related activity in the human EEG and its relation to general cognitive and learning abilities. Eur J Neurosci 2006;23:1738–1746.
103 Anderer P, Klösch G, Gruber G, Trenker E, Pascual-Marqui RD, Zeitlhofer J, Barbanoj MJ, Rappelsberger P, Saletu B: Low-resolution brain electromagnetic tomography revealed simultaneously active frontal and parietal sleep spindle sources in the human cortex. Neuroscience 2001;103:581–592.

104 McCormick L, Nielsen T, Nicolas A, Ptito M, Montplaisir J: Topographical distribution of spindles and K-complexes in normal subjects. Sleep 1997;20:939–941.
105 Jobert M, Poiseau E, Jähnig P, Schulz H, Kubicki S: Topographical analysis of sleep spindle activity. Neuropsychobiology 1992; 26:210–217.
106 Werth E, Achermann P, Dijk DJ, Borbély AA: Spindle frequency activity in the sleep EEG: individual differences and topographic distribution. Electroencephalogr Clin Neurophysiol 1997;103:535–542.
107 Zeitlhofer J, Gruber G, Anderer P, Asenbaum S, Schimicek P, Saletu B: Topographic distribution of sleep spindles in young healthy subjects. J Sleep Res 1997;6:149–155.
108 Amzica F, Steriade M: The functional significance of K-complexes. Sleep Med Rev 2002;6:139–149.
109 Forget D, Morin CM, Bastien CH: The role of the spontaneous and evoked K-complex in good-sleeper controls and in individuals with insomnia. Sleep 2011;34:1251–1260.
110 Crowley K, Trinder J, Colrain IM: An examination of evoked K-complex amplitude and frequency of occurrence in the elderly. J Sleep Res 2002;11:129–140.
111 Bastien C, Campbell K: The evoked K-complex: all-or-none phenomenon? Sleep 1992; 15:236–245.
112 Riedner BA, Hulse BK, Murphy MJ, Ferrarelli F, Tononi G: Temporal dynamics of cortical sources underlying spontaneous and peripherally evoked slow waves. Prog Brain Res 2011;193:201–218.
113 Cote KA, de Lugt DR, Langley SD, Campbell KB: Scalp topography of the auditory evoked K-complex in stage 2 and slow wave sleep. J Sleep Res 1999;8:263–272.
114 Campbell KB, Colrain IM: Event-related potential measures of the inhibition of information processing. II. The sleep onset period. Int J Psychophysiol 2002;46:197–214.
115 Yasoshima A, Hayashi H, Iijima S, Sugita Y, Teshima Y, Shimizu T, Hishikawa Y: Potential distribution of vertex sharp wave and saw-toothed wave on the scalp. Electroencephalogr Clin Neurophysiol 1984;58:73–76.
116 Chang BS, Schomer DL, Niedermeyer E: Normal EEG and sleep: adults and elderly; in Schomer DL, Lopes da Silva FH (eds): Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, ed 6, rev. Philadelphia, Lippincott Williams & Wilkins, 2011, pp 183–214.
117 Abad VC, Guilleminault C: Review of rapid eye movement behavior sleep disorders. Curr Neurol Neurosci Rep 2004;4:157–163.
118 De Gennaro L, Ferrara M, Ferlazzo F, Bertini M: Slow eye movements and EEG power spectra during wake-sleep transition. Clin Neurophysiol 2000;111:2107–2115.
119 Pizza F, Fabbri M, Magosso E, Ursino M, Provini F, Ferri R, Montagna P: Slow eye movements distribution during nocturnal sleep. Clin Neurophysiol 2011;122:1556–1561.
120 Halász P, Terzano M, Parrino L, Bódizs R: The nature of arousal in sleep. J Sleep Res 2004;13:1–23.
121 Terzano MG, Parrino L, Spaggiari MC, Palomba V, Rossi M, Smerieri A: CAP variables and arousals as sleep electroencephalogram markers for primary insomnia. Clin Neurophysiol 2003;114:1715–1723.
122 De Gennaro L, Ferrara M, Bertini M: EEG arousals in normal sleep: variations induced by total and selective slow-wave sleep deprivation. Sleep 2001;24:673–679.
123 Sforza E, Chapotot F, Pigeau R, Paul PN, Buguet A: Effects of sleep deprivation on spontaneous arousals in humans. Sleep 2004;27:1068–1075.
124 Terzano MG, Parrino L, Sherieri A, Chervin R, Chokroverty S, Guilleminault C, Hirshkowitz M, Mahowald M, Moldofsky H, Rosa A, Thomas R, Walters A: Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. Sleep Med 2002;3:187–199.
125 Terzano MG, Mancia D, Salati MR, Costani G, Decembrino A, Parrino L: The cyclic alternating pattern as a physiologic component of normal NREM sleep. Sleep 1985;8: 137–145.
126 Rizzi M, Sarzi-Puttini P, Atzeni F, Capsoni F, Andreoli A, Pecis M, Colombo S, Carrabba M, Sergi M: Cyclic alternating pattern: a new marker of sleep alteration in patients with fibromyalgia? J Rheumatol 2004;31: 1193–1199.
127 Ferri R, Parrino L, Smerieri A, Terzano MG, Elia M, Musumeci SA, Pettinato S: Cyclic alternating pattern and spectral analysis of heart rate variability during normal sleep. J Sleep Res 2000;9:13–18.
128 Svetnik V, Ferri R, Ray S, Ma J, Walsh JK, Snyder E, Ebert B, Deacon S: Alterations in cyclic alternating pattern associated with phase advanced sleep are differentially modulated by gaboxadol and zolpidem. Sleep 2010;33:1562–1570.
129 Walters AS, Lavigne G, Hening W, Picchietti DL, Allen RP, Chokroverty S, Kushida CA, Bliwise DL, Mahowald MW, Schenck CH, Ancoli-Israel S: The scoring of movements in sleep. J Clin Sleep Med 2007;3: 155–167.
130 Winocur E, Gavish A, Voikovitch M, Emodi-Perlman A, Eli I: Drugs and bruxism: a critical review. J Orofac Pain 2003;17:99–111.
131 Redline S, Budhiraja R, Kapur V, Marcus CL, Mateika JH, Mehra R, Parthasarthy S, Somers VK, Strohl KP, Sulit LG, Gozal D, Wise MS, Quan SF: The scoring of respiratory events in sleep: reliability and validity. J Clin Sleep Med 2007;3:169–200.
132 Bonnet MH, Doghramji K, Roehrs T, Stepanski EJ, Sheldon SH, Walters AS, Wise M, Chesson AL Jr: The scoring of arousal in sleep: reliability, validity, and alternatives. J Clin Sleep Med 2007;3:133–145.
133 Grigg-Damberger M, Gozal D, Marcus CL, Quan SF, Rosen CL, Chervin RD, Wise M, Picchietti DL, Sheldon SH, Iber C: The visual scoring of sleep and arousal in infants and children. J Clin Sleep Med 2007;3:201–240.
134 Caples SM, Rosen CL, Shen WK, Gami AS, Cotts W, Adams M, Dorostkar P, Shivkumar K, Somers VK, Morgenthaler TI, Stepanski EJ, Iber C: The scoring of cardiac events during sleep. J Clin Sleep Med 2007; 3:147–154.
135 Danker-Hopfe H, Kunz D, Gruber G, Klösch G, Lorenzo JL, Himanen SL, Kemp B, Penzel T, Röschke J, Dorn H, Schlögl A, Trenker A, Dorffner G: Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. J Sleep Res 2004;13:63–69.
136 Norman RG, Pal I, Stewart C, Walsleben JA, Rapoport DM: Interobserver agreement among sleep scorers from different centers in a large dataset. Sleep 2000;23:901–908.
137 Collop NA: Scoring variability between polysomnography technologists in different sleep laboratories. Sleep Med 2002;3: 43–47.
138 Rosenberger JL, Gasko M: Comparing location estimators – trimmed means, medians, and trimean; in Hoaglin DC, Mosteller F, Tukey JW (eds): Understanding Robust and Exploratory Data Analysis. New York, Wiley & Sons, 1983, pp 297–338.
139 Dietsch G: Fourier-Analyse von Elektroencephalogrammen des Menschen. Pflügers Arch Eur J Physiol 1932;230:106–112.
140 Knott JR, Gibbs FA, Henry CE: Fourier transforms of the electroencephalogram during sleep. J Exp Psychol 1942;31:465–477.
141 Gasser T, Bächer P, Möcks J: Transformations towards the normal distribution of broad band spectral parameters of the EEG. Electroencephalogr Clin Neurophysiol 1982;53:119–124.
142 Koopman PAR, Wouters PAWM, Krijzer FNC: Mean power spectra from pharmacoelectrocorticographic studies: relative baseline correction and log transformation for a proper analysis of variance to assess drug effects. Neuropsychobiology 1996;33: 100–105.

143 Kiebel SJ, Tallon-Baudry C, Friston KJ: Parametric analysis of oscillatory activity as measured with EEG/MEG. Hum Brain Mapp 2005;26:170–177.
144 Matejcek M: Vigilance and the EEG; in Herrmann WM (ed): Electroencephalography in Drug Research. Stuttgart, Fischer, 1982, pp 405–508.
145 Gasser T, Bächer P, Steinberg H: Test-retest reliability of spectral parameters of the EEG. Electroencephalogr Clin Neurophysiol 1985;60:312–319.
146 Croft RJ, Barry RJ: Removal of ocular artifact from the EEG: a review. Clin Neurophysiol 2000;30:5–19.
147 Joyce CA, Gorodnitsky IF, Kutas M: Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. Psychophysiology 2004; 41:313–325.
148 Schlögl A, Keinrath C, Zimmermann D, Scherer R, Leeb R, Pfurtscheller G: A fully automated correction method of EOG artifacts in EEG recordings. Clin Neurophysiol 2007;118:98–104.
149 Delorme A, Sejnowski T, Makeig S: Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage 2007;34:1443–1449.
150 Larsen LH, Prinz PN: EKG artifacts suppression from the EEG. Electroencephalogr Clin Neurophysiol 1991;79:241–244.
151 Zhou W, Gotman J: Removal of EMG and ECG artifacts from EEG based on wavelet transform and ICA. Conf Proc IEEE Eng Med Biol Soc 2004;1:392–395.
152 Cho SP, Song MH, Park YC, Choi HS, Lee KJ: Adaptive noise canceling of electrocardiogram artifacts in single channel electroencephalogram. Conf Proc IEEE Eng Med Biol Soc 2007;2007:3278–3281.
153 Devuyst S, Dutoit T, Stenuit P, Kerkhofs M, Stanus E: Removal of ECG artifacts from EEG using a modified independent component analysis approach. Conf Proc IEEE Eng Med Biol Soc 2008;2008:5204–5207.
154 Crespo-Garcia M, Atienza M, Cantero JL: Muscle artifact removal from human sleep EEG by using independent component analysis. Ann Biomed Eng 2008;36:467–475.
155 Shackman AJ, McMenamin BW, Slagter HA, Maxwell JS, Greischar LL, Davidson RJ: Electromyogenic artifacts and electroencephalographic inferences. Brain Topogr 2009;22:7–12.
156 Asaduzzaman K, Reaz MB, Mohd-Yasin F, Sim KS, Hussain MS: A study on discrete wavelet-based noise removal from EEG signals. Adv Exp Med Biol 2010;680:593–599.
157 McMenamin BW, Shackman AJ, Maxwell JS, Bachhuber DR, Koppenhaver AM, Greischar LL, Davidson RJ: Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG. Neuroimage 2010;49:2416–2432.
158 Itil TM, Shapiro DM, Fink M, Kassebaum D: Digital computer classifications of EEG sleep stages. Electroencephalogr Clin Neurophysiol 1969;27:76–83.
159 Smith JR, Karacan I: EEG sleep stage scoring by an automatic hybrid system. Electroencephalogr Clin Neurophysiol 1971;31: 231–237.
160 Penzel T, Hirshkowitz M, Harsh J, Chervin RD, Butkov N, Kryger M, Malow B, Vitiello MV, Silber MH, Kushida CA, Chesson AL Jr: Digital analysis and technical specifications. J Clin Sleep Med 2007;3:109–120.
161 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;7:356–364.
162 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;17:711–717.
163 Todorova A, Hofmann HC, Dimpfel W: A new frequency based automatic sleep analysis – description of the healthy sleep. Eur J Med Res 1997;2:185–197.
164 Stanus E, Lacroix B, Kerkhofs M, Mendlewicz J: Automated sleep scoring: a comparative reliability study of two algorithms. Electroencephalogr Clin Neurophysiol 1987;66:448–456.
165 Kemp B, Gröneveld EW, Janssen AJMW, Franzen JM: A model-based monitor of human sleep stages. Biol Cybernet 1987;57: 365–378.
166 Schaltenbrand N, Lengelle R, Toussaint M, Luthringer R, Carelli G, Jacmin A, Lainey E, Muzet A, Macher JP: Sleep stage scoring using the neural network model: comparison between visual automatic analysis in normal subjects and patients. Sleep 1996; 19:26–35.
167 Baumgart-Schmitt R, Herrmann WM, Eilers R: On the use of neural network techniques to analyze sleep EEG data. Third communication: robustification of the classificator by applying an algorithm obtained from 9 different networks. Neuropsychobiology 1998;37:49–58.
168 Kaplan A, Röschke J, Darkhovsky B, Fell J: Macrostructural EEG characterization based on non-parametric change point segmentation: application to sleep analysis. J Neurosci Methods 2001;106:81–90.
169 Agarwal R, Gotman J: Digital tools in polysomnography. J Clin Neurophysiol 2002; 19:136–143.
170 Grube G, Flexer A, Dorffner G: Unsupervised continuous sleep analysis. Methods Find Exp Clin Pharmacol 2002;24(suppl D):51–56.
171 Flexer A, Gruber G, Dorffner G: A reliable probabilistic sleep stager based on a single EEG signal. Artif Intell Med 2005;33:199–207.
172 Acharya UR, Chua EC, Chua KC, Min LC, Tamura T: Analysis and automatic identification of sleep stages using higher order spectra. Int J Neural Syst 2010;20:509–521.
173 Krakovská A, Mezeiová K: Automatic sleep scoring: a search for an optimal combination of measures. Artif Intell Med 2011;53: 25–33.
174 Kubicki S, Höller L, Berg I, Pastelak-Price C, Dorow R: Sleep EEG evaluation: a comparison of results obtained by visual scoring and automatic analysis with the Oxford sleep stager. Sleep 1989;12:140–149.
175 Sforza E, Vandi S: Automatic Oxford-Medilog 9200 sleep staging scoring: comparison with visual analysis. J Clin Neurophysiol 1996;13:227–233.
176 White DP, Gibb TJ: Evaluation of a computerized polysomnographic system. Sleep 1998;21:188–196.
177 Pittman SD, MacDonald MM, Fogel RB, Malhotra A, Todros K, Levy B, Geva AB, White DP: Assessment of automated scoring of polysomnographic recordings in a population with suspected sleep-disordered breathing. Sleep 2004;27:1394–1403.
178 Anderer P, Gruber G, Parapatics S, Woertz M, Miazhynskaia T, Klösch G, Zeitlhofer J, Barbanoj MJ, Danker-Hopfe H, Himanen SL, Kemp B, Penzel T, Röschke J, Grözinger M, Kunz D, Rappelsberger P, Saletu B, Dorffner G: An E-health solution for automatic sleep classification according to echtschaffen and Kales: validation study of the Somnolyzer 24 × 7 utilizing the Siesta database. Neuropsychobiology 2005;51:115–133.
179 Anderer P, Moreau A, Woertz M, Ross M, Gruber G, Parapatics S, Loretz E, Heller E, Schmidt A, Boeck M, Moser D, Kloesch G, Saletu B, Saletu-Zyhlarz GM, Danker-Hopfe H, Zeitlhofer J, Dorffner G: Computer-assisted sleep classification according to the standard of the American Academy of Sleep Medicine: validation study of the AASM version of the Somnolyzer 24 × 7. Neuropsychobiology 2010;62:250–264.
180 Schulz H: Rethinking sleep analysis. J Clin Sleep Med 2008;4:99–103.
181 Saletu B, Prause W, Anderer P, Mandl M, Aigner M, Mikova O, Saletu-Zyhlarz GM: Insomnia in somatoform pain disorder: sleep laboratory studies on differences to controls and acute effects of trazodone, evaluated by the Somnolyzer 24 × 7 and the Siesta database. Neuropsychobiology 2005;51:148–163.

182 Barbanoj MJ, Clos S, Romero S, Morte A, Giménez S, Lorenzo JL, Luque A, Dal-Ré R: Sleep laboratory study on single and repeated dose effects of paroxetine, alprazolam and their combination in healthy young volunteers. Neuropsychobiology 2005;51: 134–147.
183 Svetnik V, Ma J, Soper KA, Doran S, Renger JJ, Deacon S, Koblan KS: Evaluation of automated and semi-automated scoring of polysomnographic recordings from a clinical trial using zolpidem in the treatment of insomnia. Sleep 2007;30:1562–1574.
184 Zammit GK: Insufficient evidence for the use of automated and semi-automated scoring of polysomnographic recordings. Sleep 2008;31:449–450.
185 Svetnik V, Ma J, Soper K, Doran S, Renger JJ, Deacon S, Koblan KS: Automated sleep scoring with human supervision adds value compared with human scoring alone: a reply to Zammit GK: Insufficient evidence for the use of automated and semi-automated scoring of polysomnographic recordings. Sleep 2008;31:451.
186 Müller B, Gäbelein WD, Schulz H: A taxonomic analysis of sleep stages. Sleep 2006; 29:967–974.
187 Brandenberger G, Ehrhart J, Buchheit M: Sleep stage 2: an electroencephalographic, autonomic, and hormonal duality. Sleep 2005;28:1535–1540.
188 Jobert M, Schulz H, Jähnig P, Tismer C, Bes F, Escola H: A computerized method for detecting episodes of wakefulness during sleep based on the alpha slow-wave index (ASI). Sleep 1994;17:37–46.
189 Haustein W, Pilcher J, Klink J, Schulz H: Automatic analysis overcomes limitations of sleep stage scoring. Electroencephalogr Clin Neurophysiol 1986;64:364–374.
190 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.
191 Jobert M, Escola H, Jähnig P, Schulz H: A comparison between visual and computer assessment of sleep onset latency and their application in a pharmacological sleep study. Sleep 1993;16:233–238.
192 Pardey J, Roberts S, Tarassenko L, Stradling J: A new approach to the analysis of the human sleep/wakefulness continuum. J Sleep Res 1996;5:201–210.
193 Achermann P, Borbély AA: Temporal evolution of coherence and power in the human sleep electroencephalogram. J Sleep Res 1998;7(suppl 1):36–41.
194 Uchida S, Feinberg I, March JD, Atsumi Y, Maloney T: A comparison of period amplitude analysis and FFT power spectral analysis of all-night human sleep EEG. Physiol Behav 1999;67:121–131.
195 Ako M, Kawara T, Uchida S, Miyazaki S, Nishihara K, Mukai J, Hirao K, Ako J, Okubo Y: Correlation between electroencephalography and heart rate variability during sleep. Psychiatry Clin Neurosci 2003;57: 59–65.
196 Kemp B, Zwinderman AH, Tuk B, Kamphuisen HA, Oberyé JJ: Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. IEEE Trans Biomed Eng 2000;47:1185–1194.
197 Declerck AC, Martens WL, Wauquier W: Sleep spindle detection and its clinical relevance. Eur Neurol 1986;25(suppl 2):56–60.
198 Huupponen E, Saastamoinen A, Niemi J, Virkkala J, Hasan J, Värri A, Himanen SL: Automated frequency analysis of synchronous and diffuse sleep spindles. Neuropsychobiology 2005;51:256–264.
199 Huupponen E, Gómez-Herrero G, Saastamoinen A, Värri A, Hasan J, Himanen SL: Development and comparison of four sleep spindle detection methods. Artif Intell Med 2007;40:157–170.
200 Huupponen E, Himanen SL, Värri A, Hasan J, Saastamoinen A, Lehtokangas M, Saarinen J: Fuzzy detection of EEG alpha without amplitude thresholding. Artif Intell Med 2002;24:133–147.
201 Da Rosa AC, Kemp B, Paiva T, Lopes da Silva FH, Kamphuisen HA: A model-based detector of vertex waves and K complexes in sleep electroencephalogram. Electroencephalogr Clin Neurophysiol 1991;78:71–79.
202 Devuyst S, Dutoit T, Stenuit P, Kerkhofs M: Automatic K-complexes detection in sleep EEG recordings using likelihood thresholds. Conf Proc IEEE Eng Med Biol Soc 2010;2010:4658–4661.
203 Shmiel O, Shmiel T, Dagan Y, Teicher M: Data mining techniques for detection of sleep arousals. J Neurosci Methods 2009; 179:331–337.
204 Sorensen GL, Jennum P, Kempfner J, Zoetmulder M, Sorensen HB: A computerized algorithm for arousal detection in healthy adults and patients with Parkinson disease. J Clin Neurophysiol 2012;29:58–64.
205 Mariani S, Grassi A, Mendez MO, Parrino L, Terzano MG, Bianchi AM: Automatic detection of CAP on central and frontocentral EEG leads via Support Vector Machines. Conf Proc IEEE Eng Med Biol Soc 2011;2011:1491–1494.
206 Saccomandi F, Priano L, Mauro A, Nerino R, Guiot C: Automatic detection of transient EEG events during sleep can be improved using a multi-channel approach. Clin Neurophysiol 2008;119:959–967.
207 Tsuji Y, Satoh H, Itoh N, Sekiguchi Y, Nagasawa K: Automatic detection of rapid eye movements by discrete wavelet transform. Psychiatry Clin Neurosci 2000;54: 276–277.
208 Tan X, Campbell IG, Feinberg I: A simple method for computer quantification of stage REM eye movement potentials. Psychophysiology 2001;38:512–516.
209 Agarwal R, Takeuchi T, Laroche S, Gotman J: Detection of rapid-eye movements in sleep studies. IEEE Trans Biomed Eng 2005; 52:1390–1396.
210 Jobert M, Tismer C, Poiseau E, Schulz H: Wavelets – a new tool in sleep biosignal analysis. J Sleep Res 1994;3:223–232.
211 Durka PJ, Blinowska KJ: Analysis of EEG transients by means of matching pursuit. Ann Biomed Eng 1995;23:608–611.
212 Richard C, Lengelle R: Joint time and time-frequency optimal detection of K-complexes in sleep EEG. Comput Biomed Res 1998; 31:209–229.
213 Durka PJ: From wavelets to adaptive approximations: time-frequency parametrization of EEG. Biomed Eng Online 2003; 2:1.
214 Zucconi M, Ferri R, Allen R, Baier PC, Bruni O, Chokroverty S, Ferini-Strambi L, Fulda S, Garcia-Borreguero D, Hening WA, Hirshkowitz M, Högl B, Hornyak M, King M, Montagna P, Parrino L, Plazzi G, Terzano MG, International Restless Legs Syndrome Study Group (IRLSSG): The official World Association of Sleep Medicine (WASM) standards for recording and scoring periodic leg movements in sleep (PLMS) and wakefulness (PLMW) developed in collaboration with a task force from the International Restless Legs Syndrome Study Group (IRLSSG). Sleep Med 2006;7:175–183.
215 Ferri R, Zucconi M, Manconi M, Plazzi G, Bruni O, Ferini-Strambi L: New approaches to the study of periodic leg movements during sleep in restless legs syndrome. Sleep 2006;29:759–769.
216 Penzel T, McNames J, de Chazal P, Raymond B, Murray A, Moody G: Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. Med Biol Eng Comput 2002;40: 402–407.
217 De Chazal P, Heneghan C, McNicholas WT: Multimodal detection of sleep apnoea using electrocardiogram and oximetry signals. Philos Transact A Math Phys Eng Sci 2009;367:369–389.
218 Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Eur Heart J 1996;17:354–381.

219 Moody GB, Mark RG, Bump MA, Weinstein JS, Berman AD, Mietus JE, Goldberger AL: Clinical validation of the ECG-derived respiration (EDR) technique. Comput Cardiol 1986;13:507–510.
220 Argod J, Pépin J, Lévy P: Differentiating obstructive and central sleep respiratory events through pulse transit time. Am J Respir Crit Care Med 1998;158:1778–1783.
221 Chua CP, McDarby G, Heneghan C: Combined electrocardiogram and photoplethysmogram measurements as an indicator of objective sleepiness. Physiol Meas 2008; 29:857–868.
222 O'Rourke MF, Pauca A, Jiang XJ: Pulse wave analysis. Br J Clin Pharmacol 2001;51: 507–522.
223 De Melis M, Morbiducci U, Rietzschel ER, De Buyzere M, Qasem A, Van Bortel L, Claessens T, Montevecchi FM, Avolio A, Segers P: Blood pressure waveform analysis by means of wavelet transform. Med Biol Eng Comput 2009;47:165–173.
224 Ferber G, Abt K, Fichte K, Luthringer R: IPEG guideline on statistical design and analysis for pharmacodynamic trials. Neuropsychobiology 1999;39:92–100.
225 ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. Stat Med 1999;18: 1905–1942.
226 Biostatistical methodology in clinical trials in applications for marketing authorizations for medicinal products. CPMP Working Party on Efficacy of Medicinal Products Note for Guidance III/3630/92-EN. Stat Med 1995;14:1659–1682.
227 Kenward M, Jones B: Design and Analysis of Cross-Over Trials, ed 3. Boca Raton, Chapman & Hall/CRC Press, 2012.
228 Julious S: Sample Sizes for Clinical Trials. Boca Raton, Chapman & Hall/CRC Press, 2009.
229 Gibbons J, Chakraborti S: Nonparametric Statistical Inference, ed 5. Boca Raton, Chapman & Hall/CRC Press, 2010.
230 Liu GF, Lu K, Mogg R, Mallick M, Mehrotra DV: Should baseline be a covariate or dependent variable in analyses of change from baseline in clinical trials? Stat Med 2009;28: 2509–2530.
231 Roger J, Kenward M: The use of baseline covariates in crossover studies. Biostatistics 2010;11:1–17.
232 Pinheiro JC, Bates DM: Mixed-Effects Models in S and S-Plus. New York, Springer, 2000.
233 Ramsay J, Silverman BW: Functional Data Analysis. New York, Springer, 2010.
234 Ma J, Wang S, Raubertas R, Svetnik V: Statistical methods to estimate treatment effects from multichannel electroencephalography (EEG) data in clinical trials. J Neurosci Methods 2010;190:248–257.
235 Karlsson MO, Schoemaker RC, Kemp B, Cohen AF, van Gerven JM, Tuk B, Peck CC, Danhof M: A pharmacodynamic Markov mixed-effects model for the effect of temazepam on sleep. Clin Pharmacol Ther 2000;68:175–188.
236 Hothorn T, Bretz F, Westfall P: Simultaneous inference in general parametric models. Biom J 2008;50:346–363.
237 Hochberg Y: A sharper Bonferroni procedure for multiple tests of significance. Biometrika 1988;75:800–802.
238 Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV: Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep 2004;27: 1255–1273.
239 Danker-Hopfe H, Schäfer M, Dorn H, Anderer P, Saletu B, Gruber G, Zeitlhofer J, Kunz D, Barbanoj MJ, Himanen SL, Kemp B, Penzel T, Röschke J, Dorffner G: Percentile reference charts for selected sleep parameters for 20- to 80-year-old healthy subjects from the SIESTA database. Somnologie 2005;9:3–14.
240 Scholle S, Beyer U, Bernhard M, Eichholz S, Erler T, Graness P, Goldmann-Schnalke B, Heisch K, Kirchhoff F, Klementz K, Koch G, Kramer A, Schmidtlein C, Schneider B, Walther B, Wiater A, Scholle HC: Normative values of polysomnographic parameters in childhood and adolescence: quantitative sleep parameters. Sleep Med 2011;12: 542–549.
241 Scholle S, Feldmann-Ulrich E: Polysomnographic Atlas of Sleep-Wake States during Development from Infancy to Adolescence, ed 2. Landsberg, Ecomed Medizin, 2012.
242 Quan SF, Howard BV, Iber C, Kiley JP, Nieto FJ, O'Connor GT, Rapoport DM, Redline S, Robbins J, Samet JM, Wahl PW: The Sleep Heart Health Study: design, rationale, and methods. Sleep 1997;20:1077–1085.
243 Lind BK, Goodwin JL, Hill JG, Ali T, Redline S, Quan SF: Recruitment of healthy adults into a study of overnight sleep monitoring in the home: experience of the Sleep Heart Health Study. Sleep Breath 2003;7:13–24.
244 Heneghan C, Chua CP, Garvey JF, de Chazal P, Shouldice R, Boyle P, McNicholas WT: A portable automated assessment tool for sleep apnea using a combined Holter-oximeter. Sleep 2008;31:1432–1439.
245 Carskadon M, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S: Guidelines for the Multiple Sleep Latency Test (MSLT): a standard measure of sleepiness. Sleep 1986;9:519–524.
246 Roth T, Roehrs TA: Measurement of sleepiness and alertness: Multiple Sleep Latency Test. Chokroverty S (ed): Sleep Disorders Medicine, ed 3. Philadelphia, Saunders Elsevier, 2009, pp 218–223.
247 American Academy of Sleep Medicine: International Classification of Sleep Disorders: Diagnostic and Coding Manual, ed 2 (ICSD-2). Rochester, American Sleep Disorders Association, 2005.
248 Harsh JR, Hayduk R, Rosenberg R, Wesnes KA, Walsh JK, Arora S, Niebler GE, Roth T: The efficacy and safety of armodafinil as treatment for adults with excessive sleepiness associated with narcolepsy. Curr Med Res Opin 2006;22:761–774.
249 Pascual-Marqui RD, Michel CM, Lehmann D: Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int J Psychophysiol 1994;18:49–65.
250 Pascual-Marqui RD, Lehmann D, Koenig T, Kochi K, Merlo MC, Hell D, Koukkou M: Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naïve, first-episode, productive schizophrenia. Psychiatry Res 1999;90:169–179.
251 Vitacco D, Brandeis D, Pascual-Marqui R, Martin E: Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing. Hum Brain Mapp 2002;17:4–12.
252 Worrell GA, Lagerlund TD, Sharbrough FW, Brinkmann BH, Busacker NE, Cicora KM, O'Brien TJ: Localization of the epileptic focus by low-resolution electromagnetic tomography in patients with a lesion demonstrated by MRI. Brain Topogr 2000;12: 273–282.
253 Zumsteg D, Wennberg RA, Treyer V, Buck A, Wieser HG: $H_2{}^{15}O$ or $^{13}NH_3$ PET and electromagnetic tomography (LORETA) during partial status epilepticus. Neurology 2005;65:1657–1660.
254 Zumsteg D, Friedman A, Wieser HG, Wennberg RA: Propagation of interictal discharges in temporal lobe epilepsy: correlation of spatiotemporal mapping with intracranial foramen ovale electrode recordings, Clin Neurophysiol 2006;117: 2615–2626.
255 Anderer P, Saletu B, Saletu-Zyhlarz G, Gruber D, Metka M, Huber J, Pascual-Marqui RD: Brain regions activated during an auditory discrimination task in insomniac postmenopausal patients before and after hormone replacement therapy: low-resolution brain electromagnetic tomography applied to event-related potentials. Neuropsychobiology 2004;49:134–153.
256 Pascual-Marqui RD: Review of methods for solving the EEG inverse problem. Int J Bioelectromagn 1999;1:75–86.
257 Grech R, Cassar T, Muscat J, Camilleri KP, Fabri SG, Zervakis M, Xanthopoulos P, Sakkalis V, Vanrumste B: Review on solving the inverse problem in EEG source analysis. J Neuroeng Rehabil 2008;5:25.

258 Yao J, Dewald JPA: Evaluation of different cortical source localization methods using simulated and experimental EEG data. Neuroimage 2005;25:369–382.
259 Pascual-Marqui RD: Discrete, 3D distributed, linear imaging methods of electric neuronal activity. 1. Exact, zero error localization. 2007. http://arxiv.org/pdf/0710.3341.
260 Pascual-Marqui RD, Lehmann D, Koukkou M, Kochi K, Anderer P, Saletu B, Tanaka H, Hirata K, John ER, Prichep L, Biscay-Lirio R, Kinoshita T: Assessing interactions in the brain with exact low-resolution electromagnetic tomography. Philos Transact A Math Phys Eng Sci 2011;369:3768–3784.
261 Murphy M, Riedner BA, Huber R, Massimini M, Ferrarelli F, Tononi G: Source modeling sleep slow waves. Proc Natl Acad Sci USA 2009;106:1608–1613.
262 Saletu M, Anderer P, Saletu-Zyhlarz GM, Mandl M, Arnold O, Zeitlhofer J, Saletu B: EEG-tomographic studies with LORETA on vigilance differences between narcolepsy patients and controls and subsequent double-blind, placebo-controlled studies with modafinil. J Neurol 2004;251:1354–1363.
263 Strogatz SH: The Mathematical Structure of the Human Sleep-Wake Cycle. Berlin, Springer, 1986.
264 Rechtschaffen A: Current perspectives on the function of sleep. Perspect Biol Med 1998;41:359–390.
265 Borbély AA: A two process model of sleep regulation. Hum Neurobiol 1982;1:195–204.
266 Daan S, Beersma D, Borbély A: Timing of human sleep: recovery process gated by a circadian pacemaker. Am J Physiol 1984;246:R161–R183.
267 Cajochen C, Dijk DJ: Electroencephalographic activity during wakefulness, rapid eye movement and non-rapid eye movement sleep in humans: comparison of their circadian and homeostatic modulation. Sleep Biol Rhythms 2003;1:85–95.
268 Dijk DJ, Duffy JF, Riel E, Shanahan TL, Czeisler CA: Ageing and the circadian and homeostatic regulation of human sleep during forced desynchrony of rest, melatonin and temperature rhythms. J Physiol 1999;516:611–627.
269 Achermann P, Borbély AA: Simulation of daytime vigilance by the additive interaction of a homeostatic and a circadian process. Biol Cybern 1994;71:115–121.
270 Van Dongen HPA, Dinges DF: Investigating the interaction between the homeostatic and circadian processes of sleep-wake regulation for the prediction of waking neurobehavioural performance. J Sleep Res 2003;12:181–187.
271 Åkerstedt T, Folkard S, Portin C: Predictions from the three process model of alertness. Aviat Space Environ Med 2004;75(suppl):A75–A83.
272 Belyavin AJ, Spencer MB: Modeling performance and alertness: the QinetiQ approach. Aviat Space Environ Med 2004;75(suppl):A93–A103.
273 Mistlberger RE: Circadian regulation of sleep in mammals: role of the suprachiasmatic nucleus. Brain Res Rev 2005;49:429–454.
274 Dijk DJ, Jewett ME, Czeisler CA, Kronauer RE: Reply to technical note: nonlinear interactions between circadian and homeostatic processes: models or metrics? J Biol Rhythms 1999;14:604–605.
275 Lavie P: Ultrashort sleep-waking schedule. III. 'Gates' and 'forbidden zones' for sleep. Electroencephalogr Clin Neurophysiol 1986;63:414–425.
276 Bes F, Jobert M, Schulz H: Modeling napping, post-lunch dip, and other variations in human sleep propensity. Sleep 2009;32:392–398.
277 Hoever P, Dorffner G, Beneš H, Penzel T, Danker-Hopfe H, Barbanoj MJ, Pillar G, Saletu B, Polo O, Kunz D, Zeitlhofer J, Berg S, Partinen M, Bassetti CL, Högl B, Ebrahim IO, Holsboer-Trachsler E, Bengtsson H, Peker Y, Hemmeter UM, Chiossi E, Hajak G, Dingemanse J: Orexin receptor antagonism, a new sleep-enabling paradigm: a proof-of-concept clinical trial. Clin Pharmacol Ther 2012;91:975–985.
278 Kräuchi K, Cajochen C, Möri D, Graw P, Wirz-Justice A: Early evening melatonin and S-20098 advance circadian phase and nocturnal regulation of core body temperature. Am J Physiol 1997;272:R1178–R1188.
279 Cajochen C, Kräuchi K, Möri D, Graw P, Wirz-Justice A: Melatonin and S-20098 increase REM sleep and wake-up propensity without modifying NREM sleep homeostasis. Am J Physiol 1997;272:R1189–R1196.
280 Sadeh A: The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev 2011;15:259–267.
281 Sadeh A, Hauri PJ, Kripke DF, Lavie P: The role of actigraphy in the evaluation of sleep disorders. Sleep 1995;18:288–302.
282 Sadeh A, Acebo C: The role of actigraphy in sleep medicine. Sleep Med Rev 2002;6:113–124.
283 Tryon WW: Issues of validity in actigraphic sleep assessment. Sleep 2004;27:158–165.
284 Acebo C, LeBourgeois MK: Actigraphy. Respir Care Clin North Am 2006;12:23–30.
285 Tonetti L, Pasquini F, Fabbri M, Belluzzi M, Natale V: Comparison of two different actigraphs with polysomnography in healthy young subjects. Chronobiol Int 2008;25:145–153.
286 Vallières A, Morin CM: Actigraphy in the assessment of insomnia. Sleep 2003;26:902–906.
287 Lichstein KL, Stone KC, Donaldson J, Nau SD, Soeffing JP, Murray D, Lester KW, Aguillard RN: Actigraphy validation with insomnia. Sleep 2006;29:232–239.
288 Van Someren EJ: Actigraphic monitoring of movement and rest-activity rhythms in aging, Alzheimer's disease and Parkinson's disease. IEEE Trans Rehabil Eng 1997;5:394–398.
289 Wilson SJ, Rich AS, Rich NC, Potokar J, Nutt DJ: Evaluation of actigraphy and automated telephoned questionnaires to assess hypnotic effects in insomnia. Int Clin Psychopharmacol 2004;19:77–84.
290 Natale V, Plazzi G, Martoni M: Actigraphy in the assessment of insomnia: a quantitative approach. Sleep 2009;32:767–771.
291 Paquet J, Kawinska A, Carrier J: Wake detection capacity of actigraphy during sleep. Sleep 2007;30:1362–1369.
292 Martin JL, Hakim AD: Wrist actigraphy. Chest 2011;139:1514–1527.
293 Saksvik IB, Bjorvatn B, Harvey AG, Waage S, Harris A, Pallesen S: Adaptation and readaptation to different shift work schedules measured with sleep diary and actigraphy. J Occup Health Psychol 2011;16:331–344.
294 Berger AM, Wielgus KK, Young-McCaughan S, Fischer P, Farr L, Lee KA: Methodological challenges when using actigraphy in research. J Pain Symptom Manage 2008;36:191–199.
295 Chae KY, Kripke DF, Poceta JS, Shadan F, Jamil SM, Cronin JW, Kline LE: Evaluation of immobility time for sleep latency in actigraphy. Sleep Med 2009;10:621–625.
296 Denise P, Bocca ML: Effects of zolpidem 10 mg, zopiclone 7.5 mg and flunitrazepam 1 mg on night-time motor activity. Eur Neuropsychopharmacol 2003;13:111–115.
297 Peterson BT, Chiao P, Pickering E, Freeman J, Zammit GK, Ding Y, Badura LL: Comparison of actigraphy and polysomnography to assess effects of zolpidem in a clinical research unit. Sleep Med 2012;13:419–424.
298 Stanley N: Actigraphy in human psychopharmacology: a review. Hum Psychopharmacol 2003;18:39–49.
299 Borbély AA: Ambulatory motor activity monitoring to study the time course of hypnotic action. Br J Clin Pharmacol 1984;18(suppl 1):83S–86S.
300 Klösch G, Gruber G, Anderer P, Saletu B: Activity monitoring in sleep research, medicine and psychopharmacology. Wien Klin Wochenschr 2001;113:288–295.
301 Stanley N, Hindmarch I: Actigraphy can measure antidepressant-induced daytime sedation in healthy volunteers. Hum Psychopharmacol 1997;12:437–443.

302 Wichniak A, Skowerska A, Chojnacka-Wójtowicz J, Tafliński T, Wierzbicka A, Jernajczyk W, Jarema M: Actigraphic monitoring of activity and rest in schizophrenic patients treated with olanzapine or risperidone. J Psychiatr Res 2011;45:1381–1386.
303 Allen R, Chen C, Soaita A, Wohlberg C, Knapp L, Peterson BT, García-Borreguero D, Miceli J: A randomized, double-blind, 6-week, dose-ranging study of pregabalin in patients with restless legs syndrome. Sleep Med 2010;11:512–519.
304 Paul MA, Gray G, Sardana TM, Pigeau RA: Melatonin and zopiclone as facilitators of early circadian sleep in operational air transport crews. Aviat Space Environ Med 2004;75:439–443.
305 Littner M, Kushida CA, Anderson WM, Bailey D, Berry RB, Davila DG, Hirshkowitz M, Kapen S, Kramer M, Loube D, Wise M, Johnson SF, Standards of Practice Committee of the American Academy of Sleep Medicine: Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002. Sleep 2003:26:337–341.
306 Morgenthaler T, Alessi C, Friedman L, Owens J, Kapur V, Boehlecke B, Brown T, Chesson A Jr, Coleman J, Lee-Chiong T, Pancer J, Swick TJ, Standards of Practice Committee of the American Academy of Sleep Medicine: Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. Sleep 2007;30:519–529.
307 Atkinson AJ Jr, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, et al, Biomarkers Definitions Working Group: Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69:89–95.
308 Ritsner MS, Gottesman I: Where do we stand in the quest for neuropsychiatric biomarkers and endophenotypes and what next?; in Ritsner MS (ed): The Handbook of Neuropsychiatric Biomarkers, Endophenotypes and Genes. New York, Springer, 2009, vol 1: Neuropsychological endophenotypes and biomarkers, pp 3–21.
309 Paterson LM, Nutt DJ, Wilson SJ: Sleep and its disorders in translational medicine. J Psychopharmacol 2011;25:1226–1234.
310 Ivarsson M: Sleep changes as translational pharmacodynamic biomarkers. Eur Pharm Rev 2009;5:44–50.
311 Ruigt GS, Engelen S, Gerrits A, Verbon F: Computer-based prediction of psychotropic drug classes based on a discriminant analysis of drug effects on rat sleep. Neuropsychobiology 1993;28:138–153.
312 Wouters K, Abrahantes JC, Molenberghs G, Ahnaou A, Drinkenburg WH, Bijnens L: A comparison of doubly hierarchical discriminant analyses for multiple class longitudinal data from EEG experiments. J Biopharm Stat 2008;18:1120–1135.
313 Steiger A, Kimura M: Wake and sleep EEG provide biomarkers in depression. J Psychiatr Res 2010;44:242–252.
314 Kimura M, Steiger A: Sleep EEG provides biomarker in depression; in Turck CW (ed): Biomarkers for Psychiatric Disorders. New York, Springer, 2008, pp 355–403.
315 Pillai V, Kalmbach DA, Ciesla JA: A meta-analysis of electroencephalographic sleep in depression: evidence for genetic biomarkers. Biol Psychiatry 2011;70:912–919.
316 Modell S, Ising M, Holsboer F, Lauer CJ: The Munich vulnerability study on affective disorders: stability of polysomnographic findings over time. Biol Psychiatry 2002;52:430–437.
317 Modell S, Ising M, Holsboer F, Lauer CJ: The Munich vulnerability study on affective disorders: premorbid polysomnographic profile of affected high-risk probands. Biol Psychiatry 2005;58:694–699.
318 Lauer CJ, Modell S, Schreiber W, Krieg JC, Holsboer F: Prediction of the development of a first major depressive episode with a rapid eye movement sleep induction test using the cholinergic agonist RS 86. J Clin Psychopharmacol 2004;24:356–357.
319 Murck H, Nickel T, Künzel H, Antonijevic IA, Schill J, Zobel A, Steiger A, Sonntag A, Holsboer F: State markers of depression in sleep EEG: dependency on drug and gender in patients treated with tianeptine or paroxetine. Neuropsychopharmacology 2003;28:348–358.
320 Macher JP, Luthringer R, Staner L: Spectral EEG sleep profiles as a tool for prediction of clinical response to antidepressant treatment. Dialogues Clin Neurosci 2004;6:78–81.
321 Dumont GJ, de Visser SJ, Cohen AF, van Gerven JM, Biomarker Working Group of the German Association for Applied Human Pharmacology: Biomarkers for the effects of selective serotonin reuptake inhibitors (SSRIs) in healthy subjects. Br J Clin Pharmacol 2005;59:495–510.
322 Wilson S, Argyropoulos S: Antidepressants and sleep: a qualitative review of the literature. Drugs 2005;65:927–947.
323 Carney RM, Shelton RC: Agomelatine for the treatment of major depressive disorder. Expert Opin Pharmacother 2011;12:2411–2419.
324 Geuze E, Vermetten E: Disordered sleep in posttraumatic stress disorder: clinical presentation, research findings, and implications for treatment; in Golbin AZ, Kravitz HM, Keith LG (eds): Sleep Psychiatry. London, Taylor & Francis, 2004, pp 277–304.
325 Breslau N, Roth T, Burduvali E, Kapke A, Schultz L, Roehrs T: Sleep in lifetime posttraumatic stress disorder: a community-based polysomnographic study. Arch Gen Psychiatry 2004;61:508–516.
326 Lavie P: Sleep disturbances in the wake of traumatic events. N Engl J Med 2001;345:1825–1832.
327 Monti JM, Monti D: Sleep in schizophrenia patients and the effects of antipsychotic drugs. Sleep Med Rev 2004;8:133–148.
328 Cohrs S: Sleep disturbances in patients with schizophrenia: impact and effect of antipsychotics. CNS Drugs 2008;22:939–962.
329 Petit D, Gagnon JF, Fantini ML, Ferini-Strambi L, Montplaisir J: Sleep and quantitative EEG in neurodegenerative disorders. J Psychosom Res 2004;56:487–496.
330 Abad VC, Guilleminault C: Sleep and psychiatry. Dialogues Clin Neurosci 2005;7:291–303.
331 Roehrs T, Roth T: Drug-related sleep stage changes: functional significance and clinical relevance. Sleep Med Clin 2010;5:559–570.
332 Viola AU, Archer SN, James LM, Groeger JA, Lo JC, Skene DJ, von Schantz M, Dijk DJ: PER3 polymorphism predicts sleep structure and waking performance. Curr Biol 2007;17:613–618.
333 Goel N, Banks S, Mignot E, Dinges DF: DQB1*0602 predicts interindividual differences in physiologic sleep, sleepiness, and fatigue. Neurology 2010;75:1509–1519.
334 Schierenbeck T, Riemann D, Berger M, Hornyak M: Effect of illicit recreational drugs upon sleep: cocaine, ecstasy and marijuana. Sleep Med Rev 2008;12:381–389.
335 Hatzinger M, Hemmeter UM, Brand S, Ising M, Holsboer-Trachsler E: Electroencephalographic sleep profiles in treatment course and long-term outcome of major depression: association with DEX/CRH-test response. J Psychiatr Res 2004;38:453–465.
336 Doran SM, Wessel T, Kilduff TS, Turek F, Renger JJ: Translational models of sleep and sleep disorders; in McArthur R, Borsini F (eds): Animal and Translational Models for CNS Drug Discovery. New York, Elsevier, 2008, vol 1: Psychiatric disorders, pp 395–456.
337 Hajós M: Targeting information-processing deficit in schizophrenia: a novel approach to psychotherapeutic drug discovery. Trends Pharmacol Sci 2006;27:391–398.
338 Thaker GK: Schizophrenia endophenotypes as treatment targets. Expert Opin Ther Targets 2007;11:1189–1206.
339 Javitt DC, Spencer KM, Thaker GK, Winterer G, Hajós M: Neurophysiological biomarkers for drug development in schizophrenia. Nat Rev Drug Discov 2008;7:68–83.

340 Hajós M, Rogers BN: Targeting alpha7 nicotinic acetylcholine receptors in the treatment of schizophrenia. Curr Pharm Des 2010;16:538–554.
341 Cirelli C: The genetic and molecular regulation of sleep: from fruit flies to humans. Nat Rev Neurosci 2009;10:549–560.
342 Revel FG, Gottowik J, Gatti S, Wettstein JG, Moreau JL: Rodent models of insomnia: a review of experimental procedures that induce sleep disturbances. Neurosc Biobehav Rev 2009;33:874–899.
343 Seugnet L, Suzuki Y, Thimgan M, Donlea J, Gimbel SI, Gottschalk L, Duntley SP, Shaw PJ: Identifying sleep regulatory genes using a *Drosophila* model of insomnia. J Neurosci 2009;29:7148–7157.
344 Veasey S: Insight from animal models into the cognitive consequences of adult sleep-disordered breathing. ILAR J 2009;50:307–311.
345 Qu S, Le W, Zhang X, Xie W, Zhang A, Ondo WG: Locomotion is increased in a11-lesioned mice with iron deprivation: a possible animal model for restless legs syndrome. J Neuropathol Exp Neurol 2007;66:383–388.
346 Chen L, Brown RE, McKenna JT, McCarley RW: Animal models of narcolepsy. CNS Neurol Disord Drug Targets 2009;8:296–308.
347 Bode FJ, Stephan M, Wiehager S, Nguyen HP, Björkqvist M, von Hörsten S, Bauer A, Petersén A: Increased numbers of motor activity peaks during light cycle are associated with reductions in adrenergic alpha(2)-receptor levels in a transgenic Huntington's disease rat model. Behav Brain Res 2009;205:175–182.
348 Garcia-Garcia F, Ponce S, Brown R, Cussen V, Krueger JM: Sleep disturbances in the rotenone animal model of Parkinson disease. Brain Res 2005;1042:160–168.
349 McDowell KA, Hadjimarkou MM, Viechweg S, Rose AE, Clark SM, Yarowsky PJ, Mong JA: Sleep alterations in an environmental neurotoxin-induced model of parkinsonism. Exp Neurol 2010;226:84–89.
350 Jyoti A, Plano A, Riedel G, Platt B: EEG, activity, and sleep architecture in a transgenic AβPPswe/PSEN1A246E Alzheimer's disease mouse. J Alzheimers Dis 2010;22:873–887.
351 Platt B, Welch A, Riedel G: FDG-PET imaging, EEG and sleep phenotypes as translational biomarkers for research in Alzheimer's disease. Biochem Soc Trans 2011;39:874–880.
352 Machado-Vieira R, Kapczinski F, Soares JC: Perspectives for the development of animal models of bipolar disorder. Prog Neuropsychopharmacol Biol Psychiatry 2004;28:209–224.
353 Roybal K, Theobold D, Graham A, DiNieri JA, Russo SJ, Krishnan V, Chakravarty S, Peevey J, Oehrlein N, Birnbaum S, Vitaterna MH, Orsulak P, Takahashi JS, Nestler EJ, Carlezon WA Jr, McClung CA: Mania-like behavior induced by disruption of CLOCK. Proc Natl Acad Sci USA 2007;104:6406–6411.
354 Matos G, Tufik S, Scorza FA, Cavalheiro EA, Andersen ML: Sleep, epilepsy and translational research: what can we learn from the laboratory bench? Prog Neurobiol 2011;95:396–405.
355 Antonijevic I, Artymyshyn R, Forray C, Rabacchi S, Smith K, Swanson C, Tamm J, Mazin W, Gerald C: Perspectives for an integrated biomarker approach to drug discovery and development; in Turck CW (ed): Biomarkers for Psychiatric Disorders. New York, Springer, 2008, pp 355–403.