EEG spectral fingerprinting as a new method of safeguarding data integrity in multicenter trials

Georg Gruber, Peter Anderer, Silvia Parapatics, Sebastian Schmittner, Georg Dorffner

Methological Question: How to assure data integrity of EEG data in multicenter trials

Introduction: Phase II and III CNS clinical trials involving the polysomnographic measurement of sleep (PSG) and wake EEG are usually multi-centric, involving many (up to 100) different sites. Visual scoring and computerized analysis of the acquired raw data is commonly performed at a central scoring center. This scenario implies that PSG-/EEG data need to be transferred to the central scorer. However, up to now there is no proof that the data sent are indeed matching the correct subject. The dataflow from site to data center involves several (sometimes manual) steps and, thus, is potentially prone to undeliberate or even deliberate confusion or duplication of data. Despite the fact that this kind of error is rare, they do occur and might spoil the statistical results of a trial.

Aims: This paper proposes a new and innovative way to solve this problem, taking advantage of the fact that the human sleep EEG is characteristic for an individual (except in monozygotic twins) comparable to a fingerprint.

Methods: It has been shown that the frequency spectrum of the NREM sleep EEG is substantially different among individuals, highly stable within individuals (Lewandowski et al., 2013). Out of 174 PSG recordings, 155 (89%) could be correctly identified from the spectrum alone.

The method is equally robust to experimental challenges even as massive as total sleep deprivation (Tarokh et al 2015). It is genetically determined and has been shown to be one of the most heritable traits in humans (De Gennaro et al. 2008).

Results: Based on these findings, in our solution the spectral signature of each PSG recording received is matched to recordings from the same subject acquired at other time points. Any mismatch is revealed instantly. Likewise, any duplicate data would be disclosed, even when timestamps are manipulated. This is not possible with any checksum based algorithm. Moreover, the presented solution has the advantage of detecting incorrect EEG settings, mixed-up electrodes or recordings contaminated by abundant artifacts.

Conclusion: The presented method allows to detect discrepancies in data immediately and thus is able to take corrective actions in real time.

References:

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 Oct;64(4):455-60.

Lewandowski A, Rosipal R, Dorffner G. On the Individuality of Sleep EEG Spectra. J Psychophysiol. 2013 Jul 22;27(3):105-112.

Tarokh L, Rusterholz T, Achermann P, Van Dongen HP. The spectrum of the non-rapid eye movement sleep electroencephalogram following total sleep deprivation is trait-like. J Sleep Res. 2015 Aug;24(4):360-3.