A comparison of two sleep spindle detection methods based on all night averages

Citation for published version:

Digital Object Identifier (DOI):
10.3389/fnhum.2015.00052

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Frontiers in Human Neuroscience

Publisher Rights Statement:
Copyright © 2015 Ujma, Gombos, Genzel, Konrad, Simor, Steiger, Dresler and Bódizs. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
A comparison of two sleep spindle detection methods based on all night averages: individually adjusted vs. fixed frequencies

Péter Przemyslaw Ujma1, Ferenc Gombos2, Lisa Genzel3, Boris Nikolai Konrad4, Péter Simor5,6, Axel Steiger2, Martin Dresler4,7* and Róbert Bódizs1,2

1 Institute of Behavioral Science, Semmelweis University, Budapest, Hungary
2 Department of General Psychology, Pázmány Péter Catholic University, Budapest, Hungary
3 Centre for Cognitive and Neural Systems, University of Edinburgh, Edinburgh, UK
4 Department of Clinical Research, Max Planck Institute of Psychiatry, Munich, Germany
5 Department of Cognitive Sciences, Budapest University of Technology and Economics, Budapest, Hungary
6 Nyíregyháza Teaching Hospital, National Institute of Psychiatry and Addictions, Budapest, Hungary
7 Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, Netherlands

*Correspondence:
Martin Dresler, Department of Clinical Research, Max Planck Institute of Psychiatry, Kneipelinsstraße 2-10, 80904 München, Germany
e-mail: dresler@mpipsykl.mpg.de

Sleep spindles are frequently studied for their relationship with state and trait cognitive variables, and they are thought to play an important role in sleep-related memory consolidation. Due to their frequent occurrence in NREM sleep, the detection of sleep spindles is only feasible using automatic algorithms, of which a large number is available. We compared subject averages of the spindle parameters computed by a fixed frequency (FixF) (11–13 Hz for slow spindles, 13–15 Hz for fast spindles) automatic detection algorithm and the individual adjustment method (IAM), which uses individual frequency bands for sleep spindle detection. Fast spindle duration and amplitude are strongly correlated in the two algorithms, but there is little overlap in fast spindle density and slow spindle parameters in general. The agreement between fixed and manually determined sleep spindle frequencies is limited, especially in case of slow spindles. This is the most likely reason for the poor agreement between the two detection methods in case of slow spindle parameters. Our results suggest that while various algorithms may reliably detect fast spindles, a more sophisticated algorithm primed to individual spindle frequencies is necessary for the detection of slow spindles as well as individual variations in the number of spindles in general.

Keywords: EEG, sleep spindles, sigma waves, automatic detections, fixed frequency method, IAM, comparison

INTRODUCTION

Sleep spindles are oscillations emerging from interacting thalamocortical, corticothalamic, and reticular networks in NREM sleep (Steriade and Deschenes, 1984; Amzica and Steriade, 2000; Steriade, 2000; Fogel and Smith, 2011), which are thought to play an important role in sleep-related brain plasticity (Genzel et al., 2014). Due to their trait-like nature and relationship to plasticity, sleep spindles are frequently studied as candidate indexes of individual variations in cognitive performance. Sleep spindles are remarkably individual features: sleep spindle parameters are characterized by high intra-individual stability and inter-individual variability (De Gennaro et al., 2005), a strong genetic background (De Gennaro et al., 2008), and a correlation with anatomical properties of the brain (Piantoni et al., 2013; Saletin et al., 2013).

Due to their high prevalence and specific signal properties automatic detection methods have proven to be viable and preferable alternatives to visual detection. Some of the earliest studies (Broughton et al., 1978; Campbell et al., 1980) used phase-locked loop devices for automatic sleep spindle detection and already reported an adequate agreement with visual detection. An early combined software-hardware system (Ferri et al., 1989) also reliably replicated visual spindle detection results. Software solutions for automatic spindle detection were introduced somewhat later (Schimicek et al., 1994) and reported relatively high (approx. 70%) specificity for 90% sensitivity, while an improved method (Devuyst et al., 2006) could increase this to almost 76% in a clinical sample. More recently, sophisticated automatic sleep spindle detection methods using artificial neural networks (Acır and Güzelış, 2004; Ventouras et al., 2005) and decision trees (Duman et al., 2009) reached even higher performance, with correct classification frequently exceeding 90%.

Automatic sleep spindle recognition was further refined by adapting algorithms that take into account the inter-individual differences in sleep spindle activity, which vastly exceed intra-individual variation (De Gennaro et al., 2005) and emerge—among others—as a function of age and sex (Driver et al., 1996; Carrier et al., 2001; Huupponen et al., 2002; Genzel et al., 2012). Sleep spindle detection methods have been developed to operate with individually adjusted amplitude limits (Huupponen et al., 2000, 2007; Ray et al., 2010). A novel algorithm (Bódizs et al.,
2009; Ujma et al., 2014) based on the electrophysiological fingerprint theory of human sleep (De Gennaro et al., 2005, 2008) is the Individual Adjustment Method (IAM), which takes into account inter-individual variations not only in the amplitude, but also in the frequency of sleep spindles. In the IAM, sleep spindles are therefore not only detected based on individual amplitude thresholds, but also within the exact frequency bands where they are present in a given individual. A similarly adaptive detection method (based on a probabilistic model) is reported in Nonclercq et al. (2013).

A comparison of four different spindle detection methods (Huupponen et al., 2007) reported acceptable, but not overwhelming concordance. A recent study (Warby et al., 2014) investigated the agreement in spindle detection between expert human raters, non-experts recruited in an internet crowdsourcing effort, and automatic detection algorithms. Concordance was strongest among human experts, followed by non-experts operating in a crowdsourcing scheme, and weakest among automatic algorithms.

While the progress in automatic sleep spindle detection methods is impressive, there are numerous concerns which must be addressed in this field. A practical criticism may arise from the fact that automatic sleep spindle detections are frequently validated against visual detections: however, agreement in the visual scoring of spindles is not perfect (Campbell et al., 1980; Warby et al., 2014), the visual detection of spindles is often considered as a consensus from several raters which may bias results (Ray et al., 2010), and—despite stronger agreement among human raters than algorithms (Warby et al., 2014)—the use of human expert opinion as an absolute gold standard is philosophically questionable in itself (Bódizs et al., 2009).

Further criticism must be given to the fact of the use of standard signal detection terminology (such as sensitivity and specificity) in case of sleep spindle detection algorithms. Sleep spindles are frequent phenomena, but even so the vast majority of a sleep EEG recording does not consist of sleep spindles. Therefore, correct negative classifications are by far the most common result produced by any sleep spindle detector, which might drastically inflate specificity. The ratio of correct hits and false detections—including misses and false positives—would be a much more conservative, but also more informative measure of detection performance.

Sleep spindles are not only biological signals, but important markers of individual traits (De Gennaro et al., 2005, 2008) as well as powerful correlates of human cognition (among others: Bódizs et al., 2005; Schabus et al., 2006; Fogel et al., 2007; Ujma et al., 2014). Therefore, an alternative option in order to assess detection algorithms would be to investigate how much they can reproduce trait-like individual averages (instead of comparing individual spindle detections).

To our knowledge, it has never been investigated how strongly spindle measures of different detection methods are correlated if not individual spindle detections, but subject averages are considered. This can evidently not predicted from the signal detection characteristics of the comparison of individual spindle detections of various methods—albeit the literature usually reports moderate agreement between the individual spindle detections of different algorithms, it is unknown whether the different spindle samples obtained by different methods approximate the same individual averages. Therefore, the aim of our study was to reveal the correlation between individual sleep spindle parameters calculated with two different detection methods.

**MATERIALS AND METHODS**

**SUBJECTS**

We examined polysomnographic data of 161 healthy volunteers (88 males, 73 females, age between 17 years and 69 years, mean age 29.4 years, SD 10.7 years) recorded on the second night spent in a sleep laboratory. All procedures were approved by the responsible institution’s ethical board and subjects gave informed consent. A semi-structured interview excluded any history of neurologic or psychiatric disease, but six subjects suffered from frequent nightmares. Subjects were free of drugs and prescription medication (except for contraceptives, all data self-reported). Alcohol and excessive caffeine consumption (over two cups of coffee before noon) was not allowed. Eight subjects were smokers, while the rest were non-smokers (self-reported). This dataset used for analysis was the same as in Ujma et al. (2014), except for the inclusion of one female subject who was excluded from the previous study due to her unavailable IQ score.

**SLEEP RECORDINGS**

All subjects spent two nights in a sleep laboratory and polysomnographic data from the 2nd night was used for analysis. Since the study was performed in cooperation between multiple sleep laboratories, recordings were performed in four slightly different designs.

For 31 subjects, recordings were performed with 18 EEG electrodes using a Flat Style SLEEP La Mont Headbox device with a HBX32-SLP preamplifier (La Mont Medical Inc. USA), with a sampling rate of 249 Hz, hardware prefiltering 0.5–70 Hz and a precision of 12 bit.

For 16 subjects signals were collected, prefiltered (0.33–1500 Hz, 40 dB/decade anti-aliasing hardware input filter), amplified and digitized with 4096 Hz/channel sampling rate (synchronous) and 12 bit resolution by using the 32 channel EEG/poly-system (Brain-Quick BQ 132S, Micromed, Italy). A further 40 dB/decade anti-aliasing digital filter was applied by digital signal processing which low-pass filtered the data at 450 Hz. Finally, the digitized and filtered EEG was undersampled at 1024 Hz.

For 114 subjects, recordings were performed with a Comlab 32 Digital Sleep Lab device (Schwarzer, Germany) with a sampling rate of 250 Hz, hardware prefiltering 0.53–70 Hz and a precision of 8 bit. In 94 of these subjects, 22 EEG electrode sites were used, while in the others 20 subjects 10 EEG electrodes were used. Common recording sites in all subjects which were used in the analysis were Fp1, Fp2, F3, F4, Fz, F7, F8, C3, C4, Cz, P3, P4, T3, T4, T5, T6, O1, and O2, all referred to the mathematically linked mastoids. For the 20 subjects with only 10 electrodes, data from Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1 and O2 was available and for the other electrodes these subjects were treated as missing data.

In order to correct for potentially different baseline amplitudes depending on the recording device (Vasko et al., 1997),
the analog-digital conversion and filtering characteristics of all recording devices were measured and sleep spindle amplitudes were corrected for the measured differences as follows (Ujma et al., 2014). We determined the amplitude reduction rate of each recording system by calculating the proportion between digital (measured) and analog (generated) amplitudes of sinusoid signals at typical sleep spindle frequencies (10, 11, 12, 13, 14, and 15 Hz) for both inducing (40 and 355 μV amplitude) signals. Machine-specific amplitude reduction rates were given as the mean amplitude ratio between digital and analog values at the two amplitudes and six measured frequencies. Sleep spindle amplitudes were corrected by dividing their calculated values by the amplitude reduction rate of the recording system. Given the individual- and derivation-specific adjustment inherent to both the Fixed frequency method (FixF) and the IAM, sleep spindle densities and durations are amplitude-insensitive measures. Thus, there is no need for the compensation of the different recording systems in these values.

Sleep recordings of the second nights were scored according to standard criteria (Iber et al., 2007) on a 20 s basis and artifacts were removed by visual inspection on a 4 s basis. Sleep spindle analysis was performed on artifact-free segments of NREM sleep.

ANALYSES

Fixed frequency method of sleep spindle analysis

For the FixF method we determined the 11–13 Hz range as a slow spindle frequency band and the 13–15 Hz window as a fast spindle frequency band. These frequencies were selected to ensure consistency with previous studies (Schabus et al., 2006, 2007, 2008; Chattburn et al., 2013), which used a similar approach for the separation of slow and fast spindles.

Sleep spindles were automatically detected within artifact-free NREM sleep periods on every EEG derivation. For slow spindle detection, data were bandpass-filtered between 11 Hz and 13 Hz. The root mean squares of the filtered signals were determined for 0.25 s length time windows. Next a threshold was calculated at the 95th percentile of the root mean square values for every EEG derivation. A spindle was identified when at least two consecutive root mean square time points exceeded the threshold, and the duration criterion (≥0.5 s) was met. Four spindle characteristics were calculated; these were density (number of spindles/min); amplitude (peak-to-peak difference in voltage, expressed in μV); duration (s), and frequency (number of cycles/s, in Hz). The same procedure was followed for detecting fast spindles, using a band pass filter of 13–15 Hz (Schabus et al., 2007; Gruber et al., 2013).

Sleep spindle analysis according to the IAM

The second sleep spindle detection algorithm was the IAM (Bódizs et al., 2009). This sleep spindle detection method takes into account both inter-individual variations and intra-individual consistency in sleep spindle frequency (De Gennaro et al., 2005, 2008), analyzing sleep spindles at the individual peak frequency for all subjects.

The IAM procedure (Bódizs et al., 2009) consisted of several steps as described below (illustrated on Figure 1).

i. Average amplitude spectra. Non-overlapping 4 s artifact-free NREM sleep EEG segments are Hanning-tapered (50%), then zero-padded to 16 s. Average amplitude spectra of all-night NREM sleep EEG derivations is computed between 9 Hz and 16 Hz by using an FFT routine (frequency resolution: 0.0625 Hz).

ii. Individually adjusted frequency limits of slow and fast sleep spindles. Determination of the individual slow and fast sleep spindle frequencies is based on second order derivatives of the 9–16 Hz amplitude spectra. In order to avoid small fluctuations in convex and concave segments average amplitude spectra of 0.0625 Hz resolution (i) is decimated (resulting in a resolution of 0.25 Hz) before calculating the derivation-specific second-order derivatives in this frequency range. Derivation-specific second order derivatives of the amplitude spectra are then averaged over all EEG derivations resulting in a whole-scalp second order derivative for each subject. Individual-specific frequency limits of sleep spindles are defined as pairs of zero crossing points encompassing a negative peak in the whole-scalp second order derivatives. These zero-crossing points are rounded to the closest bins within the high-resolution (0.0625 Hz) amplitude spectra obtained in step i. Two pairs of individual-specific frequency limits and corresponding ranges are defined (one for slow and one for fast spindles). In cases of uncertainty (lack of zero crossing points indicating slow spindles or partial overlap between slow and fast sleep spindles in some cases), frequencies with predominance of power in averaged frontal (Fp1, Fp2, F3, F4, Fz, F7, F8) over averaged centro-parietal (C3, C4, Cz, P3, P4) amplitude spectra were considered as slow spindle frequencies (N = 18). There was no case of uncertainty related to the individual-specific frequency boundaries of fast sleep spindles.

iii. Individual-specific spindle middle frequencies. Slow spindle middle frequency of a given subject was quantified as the arithmetic mean of the individual-specific lower and upper limits for slow spindling as obtained above (ii). In case of fast sleep spindling the arithmetic mean of the lower and the upper frequency limits of fast sleep spindles were considered. Low spindle middle frequency of a given subject was quantified as the arithmetic mean of the pair of slow spindle frequency limits (ii) in the average amplitude spectra (i).

a. The number of high resolution (0.0625 Hz) frequency bins (i) falling in the individual-specific slow- and fast sleep spindle frequency ranges (ii) is determined.

b. The amplitude spectral values (i) at the individually adjusted frequency limits for slow and fast sleep spindles (ii) are determined. This is performed in a derivation-specific manner.

c. Number of bins for slow and fast sleep spindling (iv/a) are multiplied by the arithmetic mean of the pairs of derivation-specific amplitude spectral values for slow and
FIGURE 1 | The Individual Adjustment Method (IAM) of sleep spindle analysis. (A) Four-second EEG epoch Hanning-tapered and zero padded to 16 s. (B) Fast Fourier Transformation (FFT) is used to calculate 9–16 Hz average amplitude spectra of all night NREM sleep EEG from Hanning-tapered and zero-padded segments (derivations: Fp1, Fp2, F3, F4, Fz, F7, F8, T3, T4, T5, T6, C3, C4, Cz, P3, P4, O1, O2 referred to the mathematically-linked mastoids). (C) Amplitude spectra are decimated (down-sampled) by a factor of 4. (D) Second order derivatives of the decimated amplitude spectra. (E) Calculating the whole-scalp second order derivatives by averaging all series. The resulting average series is overplotted with the averaged frontal (Fp1, Fp2, F3, F4, Fz, F7, F8) and centro-parietal (C3, C4, Cz, P3, P4) amplitude spectra (the left-side Y axis is for average second-order derivatives, while the second Y axis on the right is for average amplitude spectral). Appropriate zero-crossing points encompassing individual-specific slow and fast sleep spindle bands are selected on the 9–16 Hz frequency scale. (F) Derivation-specific amplitude criteria are calculated. (G) Thresholding of the envelopes of the slow and fast-spindle filtered signal.
fast sleep spindle frequency limits (iv/b), respectively. Outcomes are individual- and derivation specific amplitude criteria for slow and fast sleep spindle detections.

v. Envelopes of sleep spindling. EEG data is band-pass filtered for the slow and fast spindle frequency ranges by using an FFT-based Gaussian filter with 16 s windows: \( f(x) = e^{-\left(\frac{(x-x_m)}{w/2}\right)} \), where \( x \) varies between zero and the Nyquist frequency according to the spectral resolution, \( x_m \) is the middle frequency of the spindle range (iii), and \( w \) is the width of the spindle range (ii) (ii and iii). Filtered signal is rectified and smoothed by a moving average weighted with a Hanning window of 0.1 s length and multiplied with \( \pi/2 \) (the latter is the inverse of the mean of a rectified sine wave).

vi. Detection and characterization of sleep spindles. If envelopes of this band-pass filtered and rectified data (v) exceed the individual and derivation-specific threshold as defined above (iv) for at least 0.5 s, a sleep spindle is detected. Sleep spindles detected this way are analyzed and average sleep spindle density (number of spindles per minute), sleep spindle duration (s), as well as median and maximum amplitude (expressed as all-night means of intra-spindle envelopes in \( \mu \)V at the middle of the detected spindles and at the maxima of the spindles, respectively) is calculated for the subject.

STATISTICS
FixF and IAM spindle parameters were compared using paired-sample \( t \)-tests (\( \alpha = 0.05 \)). The Benjamini-Hochberg method of false detection rate correction was performed in order to correct for multiple comparisons.

We computed Pearson’s point-moment correlation coefficients between comparable sleep spindle measures (that is, sleep spindle parameters computed from the same electrode) produced by IAM, and the FixF method.

RESULTS
IAM FREQUENCY BANDS
For the IAM method, individual slow spindle lower frequency limits ranged from 8.98 Hz to 12.95 Hz (mean: 10.96 Hz), while higher frequency limits ranged from 10.14 Hz to 13.7 Hz (mean: 11.9 Hz). Slow spindle middle frequencies ranged from 9.59 to 13.28 Hz (mean: 11.43 Hz). Fast spindle lower frequency limits ranged from 11.82 Hz to 14.77 Hz (mean: 13.06 Hz), while higher frequency limits ranged from 13.04 Hz to 16.03 Hz (mean: 14.36 Hz). Fast spindle middle frequencies ranged from 12.49 Hz to 15.38 Hz (mean: 13.71 Hz).

Individual slow spindle frequency bands were on average 0.94 Hz wide (range: 0.34–2.2 Hz). Individual fast spindle frequency bands were on average 1.3 Hz wide (range: 0.84–1.89 Hz).

Figure 2 shows the distribution of individual sleep spindle frequencies.

FixF vs. IAM SPINDLE PARAMETERS
IAM provides an approximately twice higher sleep spindle density than the FixF method in case of both slow and fast spindles as well as 1.5–2 times longer sleep spindle durations. Standard deviations of the individual averages of the FixF parameters are much smaller than in case of IAM parameters, even proportionally to the lower mean values.

Sleep spindle parameters are shown in Table 1. It must be noted that while FixF and IAM amplitude measures are displayed and compared, they are not expected to be on the same scale due to the narrower frequency band of IAM and the fact that in the FixF method amplitude was expressed as the mean maximum peak-to-peak voltage difference within a spindle, while in IAM amplitude was defined as the mean maximum of intra-spindle envelopes of the individually band-passed EEG.

The difference between comparable FixF and IAM spindle parameters is significant in all cases at \( p < 0.0001 \), and all comparisons remain significant after correction for multiple comparisons.

CORRELATIONS BETWEEN FixF AND IAM SPINDLE PARAMETERS
Despite the differences in the results, individual spindle parameters obtained with the FixF and IAM methods are strongly correlated in case of the amplitude and duration of fast spindles. These correlations are always over 0.5 for amplitude and over 0.4 for duration and they are highest (>0.8 for amplitude, >0.7 for
Table 1 | Sleep spindle parameters calculated by IAM and the fixed frequency method (FixF).

<table>
<thead>
<tr>
<th>C3</th>
<th>C4</th>
<th>F3</th>
<th>F4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>Density</td>
<td>Density</td>
<td>Density</td>
</tr>
<tr>
<td>FixF</td>
<td>IAM</td>
<td>FixF</td>
<td>IAM</td>
</tr>
<tr>
<td>Fz</td>
<td>O1</td>
<td>P3</td>
<td>P4</td>
</tr>
<tr>
<td>Slow spindles</td>
<td>F8</td>
<td>Slow spindles</td>
<td>T4</td>
</tr>
<tr>
<td>Duration</td>
<td>Duration</td>
<td>Duration</td>
<td>Duration</td>
</tr>
<tr>
<td>Amplitude</td>
<td>Amplitude</td>
<td>Amplitude</td>
<td>Amplitude</td>
</tr>
<tr>
<td>Cz</td>
<td>O2</td>
<td>F8</td>
<td>T4</td>
</tr>
<tr>
<td>Slow spindles</td>
<td>Density</td>
<td>Slow spindles</td>
<td>Density</td>
</tr>
<tr>
<td>Duration</td>
<td>Duration</td>
<td>Duration</td>
<td>Duration</td>
</tr>
<tr>
<td>Amplitude</td>
<td>Amplitude</td>
<td>Amplitude</td>
<td>Amplitude</td>
</tr>
<tr>
<td>F4</td>
<td>P4</td>
<td>F8</td>
<td>T4</td>
</tr>
<tr>
<td>Slow spindles</td>
<td>Density</td>
<td>Slow spindles</td>
<td>Density</td>
</tr>
<tr>
<td>Duration</td>
<td>Duration</td>
<td>Duration</td>
<td>Duration</td>
</tr>
<tr>
<td>Amplitude</td>
<td>Amplitude</td>
<td>Amplitude</td>
<td>Amplitude</td>
</tr>
</tbody>
</table>

(Continued)
duration) in derivations where fast spindles are most prominent (central and parietal electrodes) as well as in occipital derivations. There is, surprisingly, a negative correlation between fast spindle density calculated by the IAM and the FixF method.

There is only a week concordance between FixF and IAM slow spindle parameters. There is no significant FixF-IAM correlation in case of slow spindle density and duration, and only a modest estimation in case of slow spindle amplitude ($r < 0.5$ except for F3).

Table 2 presents the Pearson correlation coefficients depicting the linear relationship between corresponding IAM and FixF spindle parameters on all electrodes.

Given that 1) our sample consisted of several datasets recorded on various EEG devices and 2) the FixF ranges we analyzed—while based on previous literature—did not correspond well to the frequency ranges computed by IAM, we reanalyzed our sample divided in subsamples as well as with different FixF ranges set with slow spindles between 10 Hz and 12.5 Hz and fast spindles between 12.5 Hz and 15 Hz. In both re-analyses, we attempted to replicate our most prominent results, and investigated fast spindle parameters on P4 and slow spindle parameters on F3. F3 was selected over Fz because of the higher availability of this electrode in the sample.

Results are similar across subsamples: that is, fast spindle density is negatively correlated; slow spindle density and duration are not correlated, slow spindle amplitude is moderately and positively correlated while fast spindle duration and amplitude are strongly and positively correlated. FixF-IAM correlations for slow spindles on F3 are as follows for density ($r_{\text{Budapest1}} = 0.427, p = 0.016; r_{\text{Budapest2}} = -0.032, p = 0.908; r_{\text{Munich}} = 0.129, p = 0.086$), duration ($r_{\text{Budapest1}} = 0.086, p = 0.647; r_{\text{Budapest2}} = -0.143, p = 0.597; r_{\text{Munich}} = -0.072, p = 0.448$) and amplitude ($r_{\text{Budapest1}} = 0.353, p = 0.052; r_{\text{Budapest2}} = 0.498, p = 0.049; r_{\text{Munich}} = 0.519, p < 0.001$). FixF-IAM correlations for fast spindles on P4 are as follows for density ($r_{\text{Budapest1}} = -0.28, p = 0.127; r_{\text{Budapest2}} = -0.282, p = 0.291; r_{\text{Munich}} = -0.359, p < 0.001$), duration ($r_{\text{Budapest1}} = 0.844, p < 0.001; r_{\text{Budapest2}} = 0.661, p = 0.005; r_{\text{Munich}} = 0.805, p < 0.001$) and amplitude ($r_{\text{Budapest1}} = 0.75, p < 0.001; r_{\text{Budapest2}} = 0.798, p < 0.001; r_{\text{Munich}} = 0.861, p < 0.001$).

Application of the new frequency bands also did not change the pattern of consistency of our methods significantly. With the 10–12.5 Hz FixF windows, FixF-IAM correlations for slow spindles on F3 are the following: $r_{\text{density}} = 0.083, p = 0.292$; $r_{\text{duration}} = -0.069, p = 0.39; r_{\text{amplitude}} = 0.419, p < 0.001$. With the 12.5–15 Hz FixF windows, FixF-IAM correlations for fast spindles on P4 are the following: $r_{\text{density}} = -0.149, p = 0.06; r_{\text{duration}} = 0.802, p < 0.001; r_{\text{amplitude}} = 0.66, p < 0.001$.

**DISCUSSION**

While previous studies compared sleep spindle detections between various manual and automatic methods (Huupponen et al., 2007; Warby et al., 2014), to our knowledge no previous study compared individual averages of sleep spindle parameters...
Table 2 | Correlation coefficients and p-values between compatible sleep spindle parameters calculated by IAM and the fixed frequency method.

<table>
<thead>
<tr>
<th></th>
<th>C3 Slow</th>
<th>Fast</th>
<th>Fz Slow</th>
<th>Fast</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>−0.072</td>
<td>0.095</td>
<td>0.379*</td>
<td>−0.292*</td>
</tr>
<tr>
<td>p</td>
<td>0.364</td>
<td>0.230</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>C4 Slow</td>
<td>Fast</td>
<td>O1 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>−0.090</td>
<td>0.049</td>
<td>0.297*</td>
<td>−0.341*</td>
</tr>
<tr>
<td>p</td>
<td>0.255</td>
<td>0.541</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Cz Slow</td>
<td>Fast</td>
<td>O2 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>−0.067</td>
<td>0.056</td>
<td>0.320*</td>
<td>−0.296*</td>
</tr>
<tr>
<td>p</td>
<td>0.430</td>
<td>0.511</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>F3 Slow</td>
<td>Fast</td>
<td>P3 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>0.159</td>
<td>−0.050</td>
<td>0.623*</td>
<td>−0.296*</td>
</tr>
<tr>
<td>p</td>
<td>0.044</td>
<td>0.529</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>F4 Slow</td>
<td>Fast</td>
<td>P4 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>0.162</td>
<td>−0.078</td>
<td>0.365*</td>
<td>−0.235*</td>
</tr>
<tr>
<td>p</td>
<td>0.041</td>
<td>0.328</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>F7 Slow</td>
<td>Fast</td>
<td>T3 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>0.085</td>
<td>−0.086</td>
<td>0.429*</td>
<td>−0.121</td>
</tr>
<tr>
<td>p</td>
<td>0.259</td>
<td>0.310</td>
<td>0.000</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>F8 Slow</td>
<td>Fast</td>
<td>T4 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>0.071</td>
<td>−0.074</td>
<td>0.421*</td>
<td>−0.040</td>
</tr>
<tr>
<td>p</td>
<td>0.404</td>
<td>0.383</td>
<td>0.000</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>Fp1 Slow</td>
<td>Fast</td>
<td>T5 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>0.010</td>
<td>0.104</td>
<td>0.277*</td>
<td>0.052</td>
</tr>
<tr>
<td>p</td>
<td>0.901</td>
<td>0.191</td>
<td>0.000</td>
<td>0.512</td>
</tr>
<tr>
<td></td>
<td>Fp2 Slow</td>
<td>Fast</td>
<td>T6 Slow</td>
<td>Fast</td>
</tr>
<tr>
<td></td>
<td>Density</td>
<td>Duration</td>
<td>Amplitude</td>
<td>Density</td>
</tr>
<tr>
<td>r</td>
<td>0.237*</td>
<td>−0.158*</td>
<td>0.297*</td>
<td>−0.025</td>
</tr>
<tr>
<td>p</td>
<td>0.002</td>
<td>0.046</td>
<td>0.000</td>
<td>0.753</td>
</tr>
</tbody>
</table>

Significant correlations (after multiple comparison correction) are marked with an asterisk.

calculated by various methods. Moreover, comparisons of individual detections were usually performed on many spindles from a small number of subjects. We investigated the convergent validity of two well-known algorithms by correlating all-night averages of individual sleep spindle parameters in a large database of subjects. In this approach, the agreement between individual detections is admittedly less important than agreement between individual averages. Overall, our results highlight both similarities and differences in the two sleep spindle detection methods we compared, and they do not provide overwhelming evidence for the convergence of the two methods.
IAM is tuned to individual spindle frequencies as well as individual and derivation-specific amplitude limits, making it inherently more sensitive as evidenced by higher spindle density and longer duration. FixF, on the other hands, focuses on the upper 5% of the amplitude distribution of filtered EEG signals. While FixF appears to detect “the tips of the icebergs” with this approach, the fast spindles detected by FixF are able to realistically approximate the same fast spindle durations and amplitudes as the IAM. Concordance is much weaker, however, in case of slow spindle amplitude, completely absent in case of slow spindle density and duration, while a very surprising negative correlation between fast spindle densities were found. To explain these findings, some empirical tendencies must be considered.

First, while the 13–15 Hz FixF window for fast spindles was similar to the empirically determined individual frequencies of IAM fast spindles, this was not the case for the 11–13 Hz slow spindle window. Fast spindle middle frequencies were below 13 Hz in only 11.24% of all cases and over 15 Hz in 1.87% of cases, while slow spindle middle frequencies were below 11 Hz in 27.5% of all cases and over 13 Hz in 1.25% of all cases. This poor demarcation of slow spindles in the FixF method might explain why FixF slow spindle parameters correlate more strongly with IAM fast spindle parameters than IAM slow spindle parameters (FixF slow vs. IAM fast correlations on Cz: \( r_{\text{density}} = -0.092, p = 0.275 \); \( r_{\text{duration}} = 0.547, p < 0.001 \); \( r_{\text{amplitude}} = 0.603, p < 0.001 \), with similar tendencies on all electrodes, see Table 2 for correlations with IAM slow spindle parameters). This finding, together with poor agreement on density measures suggests that some FixF slow spindles may actually be classified as fast spindles by the IAM procedure and vice versa, explaining the confusion in both density measures and slow spindle parameters in general. This phenomenon is exemplified by some dissimilar findings in the field. That is, both slow and fast sleep spindle measures correlate with cognitive abilities in cases when the FixF method is used (Schabus et al., 2006, 2008), while in case of IAM fast spindles are much more stable correlates of cognitive performance (Bódizs et al., 2005, 2008; Ujma et al., 2014). It must be noted that sleep spindles are not stationary sinusoidal processes: they are known to exist around a single maximal spectral peak relative to the background EEG (Gottselig et al., 2002; Geiger et al., 2011) or by using an adaptive, probabilistic model (Nonclercq et al., 2013).

In sum, our results show that in case of fast spindles, duration and amplitude can be estimated reliably with both fixed and individual frequency methods. Much less consistency can be reached in case of slow spindles, and fixed cutoff frequencies may also lead to a poor separation of slow and fast spindles. Our results suggest that the cutoff frequencies and bandwidths for slow and fast spindles must be selected carefully and individually determined frequency bands should be considered.

It is notable that the concordance between the two methods is generally highest on typical spindle locations (frontal electrodes for slow spindles and centro-parietal electrodes for fast spindles). Concordance is usually lowest on temporal leads, but remains relatively high in occipital leads, in line with the relatively high spindle amplitude on these electrodes reported in the same dataset (Ujma et al., 2014). Lead-specific findings suggest that the lack of concordance between different spindle detection algorithms is especially problematic when non-prominent (e.g., temporal) leads are investigated.

There are limitations of our study that must be mentioned. First, the technical standards of the American Academy of Sleep Medicine.
Comparison of spindle detection methods

Ujma et al.

Results of different spindle detection methods

FIGURE 3 | Visual and automatic sleep spindle detections in a 20 s N2 sleep segment recorded from a healthy adult male. Unfiltered EEG with the visual detections of an expert rater are shown in the top data line. Appropriately filtered EEG with automatic (IAM and FixF) slow/fast spindle detections are shown in the bottom 4 data lines.

Medicine (2007) are not met in several subsamples of our study. That is, the analog-to-digital conversion rate is low (8 bits) in the largest subsample (N = 114), while the sampling rates are close to the minimally required values (249 and 250 Hz) in two subsamples. Second, while the study compared methodically well-established methods with previous practical applications in science, it must be acknowledged that IAM and the FixF algorithm operate with different philosophical underpinnings, they are designed to detect different features: the FixF method considers the background-relative amplitude of the filtered signal as the key feature of a spindle event, while IAM looks for an amplitude threshold based on the inflection points of the individual EEG.

REFERENCES


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.