Individual slow-wave morphology is a marker of aging

Péter P. Ujma\textsuperscript{a,b,*}, Péter Simor\textsuperscript{a,c}, Axel Steiger\textsuperscript{d}, Martin Dresler\textsuperscript{e,**,1}, Róbert Bódizs\textsuperscript{a,b,1}

\textsuperscript{a}Semmelweis University, Institute of Behavioural Science, Budapest, Hungary
\textsuperscript{b}National Institute of Clinical Neuroscience, Budapest, Hungary
\textsuperscript{c}Institute of Psychology, ELTE, Eötvös Loránd University, Budapest, Hungary
\textsuperscript{d}Max Planck Institute of Psychiatry, Munich, Germany
\textsuperscript{e}Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, the Netherlands

**Corresponding author at: Martin Dresler Donders Institute, Kapittelweg 29, 6525 EN Nijmegen, the Netherlands. Tel.: 024-3610984.
*Corresponding author at: Institute of Behavioural Sciences, Semmelweis University, Nagyvárad tér 4, Budapest 1089, Hungary. Tel.: +36 1 459-1500 x56454.
E-mail addresses: ujma.peter@med.semmelweis-univ.hu (P.P. Ujma), martin.dresler@donders.ru.nl (M. Dresler).
1 These authors made equal contribution.

ABSTRACT

Slow-wave activity is a hallmark of deep non-rapid eye movement sleep. Scalp slow wave morphology is stereotypical and it is highly correlated with the synchronized onset and cessation of cortical neuronal firing measured from the surface or depth of the cortex. It is also strongly affected by aging, and these changes are causally associated with age-related cognitive decline. We investigated how normal aging affects the individual morphology of the slow wave and whether these changes are captured by the summary slow wave parameters generally used in the literature. We recorded full-night polysomnography in 176 participants (age 17–69 years) and automatically detected slow waves. We established individual slow morphologies using average amplitude at 501 data points for each participant and also calculated the individual average slow-wave amplitude, average ascending and descending slope steepness, halfwave duration, and the total number of slow waves (gross parameters). Using least absolute shrinkage and selection operator penalized regression, we found that SW gross parameters explain up to 60% of age variance but using fine morphology up to 80% of age variance can be accounted for. This predictive power was greatest when data from multiple channels were averaged, in midline derivations and in the first quarter of the night. Young participants had faster slow-wave polarity reversals, suggesting a more efficient initiation and termination of slow-wave downstate and upstate. Our results demonstrate the superiority of the high-resolution slow wave morphology as a biomarker of aging and highlight downstate-upstate transitions as promising targets of restorative stimulation-based interventions.

© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Electroencephalographic recordings in non-rapid eye movement (NREM) sleep are characterized by slow, large amplitude waves (“slow waves” [SWs]). These SWs are cortical in origin (Amzica and Steriade, 1995; Cercsca et al., 2010; Steriade et al., 1993), with thalamic regulation (Crunelli and Hughes, 2010; David et al., 2013), and they reflect the rhythmic, highly synchronized onset and cessation of cortical neuronal firing (Nir et al., 2011; Steriade et al., 1993). This particular pattern of neural firing has an important role in what is perhaps the most important function of NREM sleep: the normalization of synaptic connections formed during previous wakefulness (Tononi and Cirelli, 2014). In line with this observation, SWs are more frequent and have greater amplitude in participants and experimental conditions where stronger synaptic connections are expected to form before sleep: in younger participants, whose synaptic plasticity is greater in general (Carrier et al., 2001, 2011; Feinberg and Campbell, 2013; Kurth et al., 2012; Pőtári et al., 2017) or in case of sleep deprivation (Borbely et al., 1981) or intensive, experience-rich wakefulness (Huber et al., 2004).

SWs—even when recorded from the scalp—accurately reflect the temporal dynamics of the synchronized onset and cessation of neural firing, as well as its spatial extent over the cortical mantle below the scalp electrodes (Nir et al., 2011). The detailed morphology of SWs (slope steepness, minor structural differences, frequency) reflects the characteristics of neural firing, and it changes as a function of age and previous wakefulness: steeper,
larger, smoother SWs indicate stronger synchronized firing and therefore stronger synaptic connections (Huber et al., 2007; Riedner et al., 2007). As it is expected, this pattern is more common in young participants, early in the night and after extended, or more active wakefulness (Bersaglieri and Achermann, 2010; Colrain et al., 2010; Huber et al., 2013; Vyazovskiy et al., 2008).

The age-related reduction of SW activity is well known (Carrier et al., 2001, 2011). SW density and amplitude is lower in older individuals, and it is especially associated with cognitive decline (Dresler et al., 2014; Yaffe et al., 2014), whereas retained SW activity at a higher age is associated with better cognitive and physical health (Anderson and Horne, 2003; Mazzotti et al., 2014). Our recent study (Pötári et al., 2017) has shown that high IQ is associated with a significantly attenuated age-related reduction in SW spectral power, indicating that retained cortical plasticity may be an important mechanism behind the better cognitive and perhaps somatic health of participants with high general cognitive ability.

In sum, SWs reflect the ability of cortical neuron populations to engage in synchronized activity through synaptic connections, and they are affected by aging, resulting in worse cognitive functioning. This renders SW activity a promising marker of age-related cognitive problems.

SWs have multiple characteristics which reflect different properties of the underlying synchronized neuronal firing. The amplitude of waves reflects the maximum extent of firing synchrony (Nir et al., 2011; Vyazovskiy et al., 2008), the steepness of slopes reflects the rapidness of the buildup and cessation of neuronal firing (Esser et al., 2007; Riedner et al., 2007) and frequency—not independently from the former—reflects the overall synaptic load (Carrier et al., 2011). It has been demonstrated that different SW parameters are differently affected by aging (Carrier et al., 2011), and age-related structural changes in the brain have different effects on density and amplitude (Dube et al., 2015; Latreille et al., 2019).

Although they provide more temporal resolution than power spectral density or the total number of SWs, SW slope and amplitude are crude approximations of the true morphology of SWs and therefore can only approximately track the age-related changes of it. Owing to their highly stereotypical morphology, it is possible to investigate SWs with a methodology similar to what is used in case of event-related potentials (ERPs) (Key et al., 2005). ERP analysis focuses on the average of EEG signals following several functionally identical events (usually the repeated presentation of a stimulus) and it identifies stereotypical patterns in EEG activity which may be obscured in individual cases but appear in the averaged recording. In the sleep EEG, such analyses were previously performed on a specific type of SW, the evoked K-complex (Crowley et al., 2002). K-complexes are morphologically similar to ordinary SWs, which can also appear spontaneously but can also be elicited by stimulation (Halász, 2005). Evoked K-complex components, computed in an identical manner to ERPs in wakefulness, are not only sensitive to the modality of the stimulus with which they were elicited but also to sex, age, and disease (Colrain and Baker, 2011; Colrain et al., 2010, 2011; Crowley et al., 2002). It is reasonable to assume that the shape of spontaneously occurring NREM SWs is also subject to such age-related variation, but this has never been empirically investigated.

An ERP-like analysis on spontaneous SWs is possible even in the absence of a trigger stimulus by aligning the EEG signal to a stereotypical morphological element of SWs (such as the negative peak) and averaging accordingly. This analysis reveals not only the major microstructural parameters of SWs (such as halfwave frequency or ascending/descending slope steepness) but also identifies specific wave components and their precise dynamics at a temporal resolution equal to the sampling frequency of the recording, resulting in substantial additional information about SW dynamics.

The goal of this study was to investigate the age-related changes in the high-resolution temporal morphology of spontaneously occurring NREM SWs using a large existing sleep EEG database of healthy participants of various ages. We hypothesized that the association between age and SW characteristics is stronger if the detailed SW morphology is also considered in multivariate models. As a result of this study, we were able to pinpoint candidate electrophysiological markers of healthy aging beyond the change of major SW elements such as amplitude or density. We found that younger participants are characterized by faster SW phase transitions both preceding and following negative peaks, suggesting a more rapid buildup and cessation of SW downstates.

2. Materials and methods

2.1. Participants

Pre-existing data from 176 healthy participants (mean age 29.9 years, SD 10.62 years, range 17–69 years; 95 males with mean age = 30.35 years, SD = 10.51 years, range 17–69 years; 81 females with mean age = 29.25 years, SD = 10.68 years, range 18–60 years; t(age by sex) = 0.68, p = 0.49) from a multicenter database of the Max Planck Institute of Psychiatry (Munich, Germany), the Psychophysiology and Chronobiology Research Group of Semmelweis University (Budapest, Hungary), and the Budapest University of Technology and Economics (Pötári et al., 2017; Ujma et al., 2014) was used in this retrospective study. The exact breakdown of participant ages by sex is provided in Supplementary table S1. Data from one 19-year-old male participant was excluded because of abnormal average SW morphology. Study procedures were approved by the ethical boards of Semmelweis University, the Medical Faculty of the Ludwig Maximilian University, or the Budapest University of Technology and Economics. All participants were volunteers who gave informed consent in line with the Declaration of Helsinki. According to semistructured interviews with experienced psychiatrists or psychologists, all participants were healthy, had no history of neurologic or psychiatric disease, and were free of any current drug effects, excluding contraceptives in females. Consumption of small habitual doses of caffeine (maximum 2 cups of coffee until noon), but no alcohol, was allowed. Six male and 2 female participants were light-to-moderate smokers (self-reported), and the rest of the participants were nonsmokers. Further details about participant selection criteria and study protocols can be found in the studies referenced previously.

2.2. Electroencephalography

All participants underwent all-night polysomnography recordings for 2 nights, and data from the second night were used for all analyses. Scalp EEG electrodes were applied according to the 10–20 system (Jasper, 1958) and referenced to the mathematically linked earlobes. Impedances were kept at <8 kΩ. EEG was sampled at 250 Hz for 115 participants, 249 Hz for 29 participants, and 1024 Hz for 15 participants, always resampled at 250 Hz (see Section 2.3). Sleep EEG was visually scored on a 20-second basis according to standard criteria (Iber et al., 2007). A visual scoring of artifacts was also performed on a 4-second basis. EEG preprocessing was implemented in Fericio’s EEG (@Ferenc Gombos, Budapest, Hungary). Further details about the technical details of the sample can be found in Table 1 in the study by Ujma et al. (2014).

2.3. Slow-wave detection

We used multiple methods of SW detection to test the robustness of our findings. We ran detections on 2 signal derivations: (1)
the averaged data of frontal channels and (2) individual channel data. For frontal SW detection, the signal of the electrodes F3 and F4 was band-pass–filtered to 0.5–4 Hz (two-way finite impulse response filter implemented in EEGLAB) and averaged. SWs were detected on this averaged signal. An SW was detected if a negative deflection of the signal persisted for 0.2–1 second and exceeded an amplitude threshold based on previous studies. To test the robustness of results across several possible amplitude thresholds, detections were performed using 3 different values and used in separate subsequent analyses:

1. Maximal negative deflection $<-30$ $\mu$V, with the lower 50% of the amplitude distribution discarded as putative false positives (Bersaglieri and Achermann, 2010)—note that, in this case, the exact amplitude and distribution thresholds were not identical to the cited study because of the differences in study goals.
2. Maximal negative deflection $<-37.5$ $\mu$V, with peak-to-peak amplitude to the subsequent positive deflection $>75$ $\mu$V (Piantoni et al., 2013).
3. Maximal negative deflection $<-75$ $\mu$V, with peak-to-peak amplitude to the subsequent positive deflection $>140$ $\mu$V (Massimini et al., 2004).

Putative SWs which were detected even partially outside N2 or SWS and/or coincided with a visually detected artifact epoch were discarded. All SWs detected using each detection threshold were extracted from the EEG signal used for their detection and averaged for each participant, yielding 3 individual average SWs for each participant, one according to each detection criterion. To control for inter-individual differences in baseline EEG voltage because of sex and body size effects, we also calculated a standardized SW for each participant, in which the mean amplitude of the average SW was 0 and its standard deviation 1 within the same participant, regardless of raw voltage. Because this was performed for the average waves detected using all 3 methods, it resulted in a total of 6 average SWs for participant, 3 consisting of EEG voltage time series and another 3 of standardized values of the former.

For all-scalp detections we ran, the above described methods (2) and (3) on the data of 18 EEG channels (Fp1, Fp2, F3, F4, Fz, F7, F8, C3, C4, Cz, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6). Data from F7, F8, T3, T4, T5, T6, Fz, and Cz were not available for 20 participants and treated as missing data in all subsequent analyses. Simultaneous detections on multiple different electrodes (negative peaks separated by no more than 200 ms) were grouped together as a single traveling wave, and SWs only exceeding the 75 $\mu$V amplitude criterion were only considered if they co-occurred by these definitions with at least one 140 $\mu$V SW on another channel (Ujma et al., 2018).

Macroscopic features of SW morphology frequently used in EEG literature (henceforth referred to as “gross parameters”) were represented by the number of SWs, their average amplitude (negative peak voltage in $\mu$V), and their average slopes (descending downstate [DDS], ascending downstate [ADS], ascending upstate [AUS], descending upstate [DUS], in all cases expressed in $\mu$V/s) and their average negative and positive halfwave durations, which were calculated for each participant. In addition to these, for the 30 $\mu$V frontal model which used a dynamic detection threshold corresponding to the 50th percentile of the amplitude distribution of these relatively small SWs, the value of the 50th percentile was also used as a gross parameter. For all-scalp SW detections, we also calculated mean lag (the mean difference of the timing of SW negative peaks on this channel and the channel the first co-occurring SWs occurred on) as a gross parameter. Gross parameters were 2-transformed across all participants of the same sex to eliminate possible sex differences.

The fine morphology of SWs was represented by 501 equally spaced time points including and sampled $\pm 1$ second around the negative peak of the average individual SWs, thus each separated by 4 milliseconds. Amplitude values at these time points were obtained using linear interpolation for the participants whose original EEG sampling frequency was not 250 Hz. Amplitude values at each of the 501 sampling points were averaged across all SWs within the same participants, yielding an average individual SW morphology consisting of 501 variables, each representing the average SW amplitude at one 4 msec sampling point. For average frontal channel data, the fine morphology of SWs was calculated using all 3 detection criteria. For all-scalp detections in case of SWs co-occurring on multiple channels, SW morphology was measured around each channel’s individual SW peak.

SW detection and further statistical analyses were implemented in MATLAB R2017a (The MathWorks, Natick, MA) using EEGLAB (Delorme and Makeig, 2004) and custom scripts.

2.4. Statistical analysis

This study aimed to reveal whether the fine morphology of SWs is affected by aging beyond its effects on gross parameters: formally, whether the SW voltage values included in the measurement of fine morphology account for additional age variance beyond what is accounted for by gross parameters. Standard least square regression modeling of this relationship was precluded by the large number of predictors: although age and EEG data were only available for 176 participants, gross parameters and fine morphology of SWs were represented by 507 predictors (508 in case of 30 $\mu$V models), leading to a large degree of model under-determination. We aimed to solve the problem of under-determination using least absolute shrinkage and selection operator (LASSO) regression (Tibshirani, 1996). LASSO is an iterative learning algorithm which aims to minimize the following function:

$$
\sum_{i=1}^{n}(y_i - \sum_{j} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| 
$$

Where, $y_i$ is the dependent variable with n observations, x_{ij} are predictors (total number: p), $\beta_j$ are regression coefficients assigned to the predictors, and $\lambda$ is a penalty parameter which increases in each iteration of LASSO. It is easy to see that the parts of (1) before $\lambda$ refer to the residual or prediction error of the regression model; however, in each iteration, an additional error term is added for each nonzero regression coefficient $|\beta_j|$, and this additional error is larger for larger values of $\lambda$. In other words, with higher values of $\lambda$, the model is forced to constrain more regression coefficients to 0, leaving only the strongest predictors in the model, until ultimately no predictors remain. LASSO balances the number of predictors against prediction accuracy, finding a value of $\lambda$ at which enough predictors are left in the model to account for maximal variance but not so many that it would result in under-determination. Formally, LASSO seeks a $\lambda$ value for which optimal fit to (1) is achieved in a holdout sample. LASSO is useful when only a subset of the measured predictors are really associated with the dependent variable (Lello et al., 2017) or when a small subset or a single best candidate must be selected from multiple highly correlated predictors (Krapohl et al., 2017), and in the biological sciences, its use is especially common in animal breeding genetics (de los Campos et al., 2013).

LASSO was implemented in MATLAB with 5-fold cross-validation. We used LASSO on 7/8 of our sample (N = 154) for training and 1/8 (N = 22) for validation. Because of 5-fold cross-validation, the model was trained to iteratively find the
of the amplitude of the individual average SW at “time,” expressed in seconds relative to the negative peak.

We recalculated our models by averaging SWs after triggering them to the negative-positive zero crossing instead of the negative peak to calculate fine morphology. Even in this case, the best-fit LASSO model identified average SW amplitude at multiple time points as significant predictors of age beyond the effect of gross parameters, but these models did not provide better fit in the holdout sample than those using gross parameters only ($R^2_{\text{dynamicgross}} = 47.4\%$, $R^2_{\text{750Vgross}} = 65.4\%$, $R^2_{\text{1400Vgross}} = 59.7\%$, $R^2_{\text{dynamicfull}} = 32.3\%$, $R^2_{\text{750Vfull}} = 48.8\%$, $R^2_{\text{1400Vfull}} = 66.1\%$, $R^2_{\text{dynamicfullzscore}} = 34.3\%$, $R^2_{\text{750Vfullzscore}} = 28.9\%$, $R^2_{\text{1400Vfullzscore}} = 50.6\%$, $1/6$ positive signs, $W = 1, p = 0.063$ based on the Wilcoxon test).

3.2. All-scalp SWs

We next investigated potential topographic differences in the predictive potential of SW gross parameters and fine morphology. Gross parameters provided the best fit to age data on T6 ($R^2 = 62\%$), whereas SW morphology provided the most additional explained variance on Fp1 ($\Delta R^2 = 19\%$). Overall, age could best be predicted using gross parameters and SW morphology on Cz ($R^2 = 66\%$). Regression models also using SW morphology provided better fit to data over gross parameter-only models on 14 derivations (14/18 positive signs, $W = 25, z = 2.64, p = 0.008$ based on the Wilcoxon test). Typically, SW morphology was more important for predicting age along the sagittal midline derivations, especially the extremes (frontopolar and occipital), and less important in temporal derivations. On average, SW morphology provided 6.8% of additional explained variance. Notably, all models—including those using frontal single-channel data—underperformed compared with the models using SW detections on the averaged signal of the 2 frontal channels. Fig. 3 illustrates topographic differences in the predictive power of gross parameters and SW morphology.

3.3. Ultradian effects

We next investigated potential ultradian differences in the predictive potential of SW gross parameters and fine morphology. We reran LASSO models on gross parameters and fine morphology data calculated separately from the 4 sleep quartiles of all detections of the same individual. We performed this analysis both on averaged frontal data and all single channels individually. In both models, age could generally be best predicted using SW data from the first sleep quartile and predictive power dropped in the subsequent 4 quartiles. Using the data from averaged frontal channels, full models significantly outperformed gross parameter—only models substantially in the first ($\Delta R^2_{\text{mean}} = 22.05\%$, 6/6 positive signs, $W = 0, p = 0.031$) and second ($\Delta R^2_{\text{mean}} = 25.29\%$, 6/6 positive signs, $W = 0, p = 0.031$) sleep quartiles, but the difference remained a tendency in the third ($\Delta R^2_{\text{mean}} = 5.77\%$, 5/6 positive signs, $W = 1, p = 0.063$) quartile and full models significantly underperformed in the fourth ($\Delta R^2_{\text{mean}} = -16.43\%$, 0/6 positive signs, $W = 21, p = 0.031$). Using all-scalp data, there was a significant difference in the second ($\Delta R^2_{\text{mean}} = 11.28\%$, 14/18 positive signs, $W = 21, z = 2.81, p = 0.005$) quartile. There was no significant difference in the first ($\Delta R^2_{\text{mean}} = 0.4\%$, 8/18 positive signs, $W = 79, z = 0.28, p = 0.77$), the third ($\Delta R^2_{\text{mean}} = 5.9\%$, 13/18 positive signs, $W = 48, z = 1.64, p = 0.1$), and the last ($\Delta R^2_{\text{mean}} = 3.8\%$, 10/18 positive signs, $W = 73, z = 0.54, p = 0.58$) quartiles (Fig. 4).

Clear topographic differences in the relative contribution of SW gross parameters and fine morphology to prediction performance emerged across cycles. In the first cycles, where predictive power was

\[
\text{age} = 90.41 - 0.15 \times \text{Count} + 3.19 \times \text{Amplitude} \\
+ 1.63 \times \text{DUS slope} + 14.21 \times \text{zamp}_{-1} - 2.28 \times \text{zamp}_{-0.424} \\
- 9.04 \times \text{zamp}_{-0.416} + 33.2 \times \text{zamp}_{0.176} + 1.52 \times \text{zamp}_{-0.1} \\
- 10.93 \times \text{zamp}_{-0.048} + 10.85 \times \text{zamp}_{0.036} \\
+ 11.97 \times \text{zamp}_{0.04} - 10.5 \times \text{zamp}_{0.188} - 0.88 \times \text{zamp}_{0.372} \\
- 0.54 \times \text{zamp}_{0.376} - 13.61 \times \text{zamp}_{0.456} + 2.24 \times \text{zamp}_{0.628} \\
+ 5.96 \times \text{zamp}_{0.664} - 18.87 \times \text{zamp}_{0.912} - 6.87 \times \text{zamp}_{0.956} \\
\]

(2)

where count, amplitude, and DUS slope refer to the z-scores (across participants with the same sex) of the corresponding gross parameters using the 140 $\mu$V detection criterion, whereas zamp$_{\text{Prime}}$ refers to the z-score (across the average SW of the same individual)
maximal and the contribution of fine morphology to prediction accuracy the lowest, the predictive power of gross parameters and fine morphology was evenly distributed across the scalp. In cycles 2 and 3, the predictive power of gross parameters was strongest in more posterior channels, but the additional accuracy of fine morphology remained maximal at frontal locations. In the fourth cycle, while prediction accuracy was the lowest and the overall difference between gross parameter—only and full models was not significant, fine morphology on posterior channels contributed substantially to prediction accuracy (ΔR²Q2 = 42%, ΔR²T6 = 33%, ΔR²T4 = 17%) (Fig. 5).

3.4. Contributing elements of SW fine morphology

Which exact elements of SW fine morphology predict age over the effects of gross parameters? To determine this, we took all B-values from best-fit LASSO models, summed them within 125 msec
temporal bins, and plotted them against an average SW (Fig. 6). SWs in younger participants had more negative voltage immediately before but more positive voltage immediately after and 125–250 msec before the negative peak. This effect was most prominent in averaged frontal and frontocentral derivations. Another minor effect was the presence of a later, larger positive peak after the negative peak in younger individuals, which showed a steeper decline thereafter.

In other words, younger participants were characterized by a more rapid initiation and dissipation of the SW downstate, and a later, more prominent upstate, which also declined faster. This effect was present over and beyond what was captured by gross parameters such as SW slope and halfwave duration.

3.5. Sex differences

The modest size of our replication sample did not allow for the rigorous testing of sex differences in age prediction accuracy. Nominally, both the 140 μV average frontal model and the selected Fp1 single-channel model showed higher accuracy in females (N = 9, rfrontal_gross = 0.91, rfrontal_full = 0.96; rFp1_gross = 0.88, rFp1_full = 0.93) than in males (N = 13, rfrontal_gross = 0.65, rfrontal_full = 0.86; rFp1_gross = 0.49, rFp1_full = 0.75), but at this sample size, none of these differences were statistically significant. In the full sample, we observed a stronger correlation between ascending slope steepness and age in females, while early fine morphology amplitude correlated with age positively in males and negatively in females. In addition, immediately before the onset of the downstate, a more negative age-amplitude correlation was seen in males. These findings are illustrated on Fig. 7.

3.6. SWs mediating the relationship between age and cognitive performance

If the relationship between age and cognitive performance is mediated by age-related changes in SW characteristics, then the correlation between age and cognitive performance should be attenuated after controlling for SW characteristics. We assessed this by using different measures of cognitive performance we had in our database: the Zahlen verbindungs test (ZVT, N = 107), a trail-making test, and 2 nonverbal IQ tests as available in each participant, Raven’s Advanced Progressive Matrices (APMs) and the Culture Fair Test (CFT), henceforth referred to as IQ, although we did not norm raw scores to a deviation IQ score (N = 158 with at least one test score). We converted raw test scores to Raven’s APM equivalents (the Raven’s APM raw score corresponding to the same IQ as the other test, given the participant’s age as the actual test score) before averaging. Raven’s APM and CFT measure general

<table>
<thead>
<tr>
<th>Detection threshold</th>
<th>Gross parameters only</th>
<th>SW morphology only</th>
<th>SW morphology only, normalized</th>
<th>Full model</th>
<th>Full model, with normalized SW morphology</th>
</tr>
</thead>
<tbody>
<tr>
<td>75 μV</td>
<td>47.64%</td>
<td>71.16%</td>
<td>75.58%</td>
<td>67.02%</td>
<td>75.24%</td>
</tr>
<tr>
<td>140 μV</td>
<td>58.45%</td>
<td>80.41%</td>
<td>76.49%</td>
<td>72.14%</td>
<td>77.32%</td>
</tr>
</tbody>
</table>

Rows refer to the threshold of the SW detection method, whereas columns refer to differences in model construction (predicting age based on gross parameters only, SW morphology only or both, and whether SW morphology was z-transformed within-participant to eliminate amplitude differences).
Fig. 3. Topographic differences in the predictive power of SW gross parameters and fine morphology. Panel (A): the topographical distribution of explained variance in the holdout sample using regression coefficients from LASSO with gross parameters only (left) or gross parameters and SW morphology (middle) and the difference of explained variance between the 2 models (right). O1 data were manually removed for 1 participant with outlying data (correlation between predicted and actual age with and without this participant: \( r_{\text{gross, with}} = 0.45 \), \( r_{\text{gross, without}} = 0.66 \); \( r_{\text{full, with}} = 0.04 \), \( r_{\text{full, without}} = 0.66 \)). Panel (B): the correlation between predicted age (axis x) and actual age (axis y) in the holdout sample based on gross parameters only (left) or gross parameters and SW morphology (middle). All data are from Fp1 where the difference between the performance of the 2 models was maximal. Panel (C): Pearson’s correlation coefficients between individual average morphology at 501 data points and age after regressing out the effects of gross parameters (left axis, red line) and the regression coefficients from the best-fit LASSO model (right axis, black line). All data are from Fp1 where the difference between the performance of the 2 models was maximal. Note that SW morphology shows substantial correlations with age even after controlling for gross parameters, and LASSO fits the best predictive model to include data points with high residual correlation. Abbreviations: LASSO, least absolute shrinkage and selection operator; SW, slow wave. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 4. The predictive power (proportion of age variance explained in the holdout sample) of SW characteristics across sleep quartiles. Panel (A) shows the average predictive power of the 18 models using SW data from each EEG derivation. Error bars show 95% confidence intervals computed from SD across channels. Panel (B) shows the predictive power of models using SW detections on averaged frontal data. Full model with z-transformation refers to models in which the average SW was z-transformed for each participant. Abbreviation: SW, slow wave.
cognitive ability and test-specific abilities (Gignac, 2015), whereas ZVT is a measurement of general cognitive ability and mental speed, both of which are negatively affected by aging (Salthouse, 2004). Both IQ ($r = -0.22$) and ZVT ($r = 0.49$) scores were significantly correlated with age (note that because ZVT measures reaction time, a positive correlation indicates worse performance at later age).

We calculated Pearson’s partial correlation coefficients between age and ZVT/IQ, correcting for SW characteristics (Fig. 8) and assessed whether the corrected correlation falls outside the confidence intervals of the uncorrected correlation coefficient. We did not find evidence for significant mediation, but a tendency for it emerged in case of both test scores if we controlled for amplitude during early descending downstates and late ascending upstates. In other words, our findings provide some limited evidence that the loss of SW steepness may play a functional role in worse cognitive functioning at higher ages.

4. Discussion

We trained LASSO regression models to predict age from SW gross parameters and fine morphology in 154 participants and validated our models in an independent holdout sample of 22 participants. Although our findings can be viewed as a proof of concept for predicting age (or rather a ‘brain age’ (Sun et al., 2019)) from sleep EEG data, in our view, the most important insights they provide is into how SWs in NREM are affected by normal aging. Our study is by its nature correlational but because the route of causation between chronological age and SW characteristics can only run from the former to the latter, learning which SW characteristics predict age best is an indicator of which ones are the most affected by aging. Our study revealed which age-related changes in SWs are adequately captured by easily calculable gross parameters, which ones are only captured by the more accurate fine

Fig. 5. Topographic differences in the predictive power of SW parameters across sleep quartiles. The left column of topographic plots shows the proportion of age variance explained in the holdout sample by gross parameter-only models using data from each electrode. The middle column illustrates topographic differences of the proportion of variance explained by full models (gross parameters plus SW morphology). The right column shows the difference in variance explained between the 2 models. Abbreviation: SW, slow wave.

Fig. 6. The elements of SW fine morphology which predicted age over the effects of gross parameters. Regression coefficients from the best-fit LASSO models were summed up within 125 msec bins across all models with similar predictors: averaged frontal SWs with or without z-transformation with different amplitude criteria (Panel A) and single-channel SWs from frontocentral (Fp1, Fp2, F3, F4, C3, C4, Fz, Cz), posterior (P3, P4, O1, O2, T5, T6), or lateral (F7, F8, T3, T4) channels (Panel B). All sums of B-values are plotted against the grand average 140 μV frontal SW for illustration. Note the different y-axis limits for frontocentral waves.
morphology, and how these changes and their proportions are different across EEG derivations and sleep cycles.

We found that gross parameters are indeed highly correlated with age, sharing up to 60%–65% of their overall variance with it. This was higher than what was found in previous studies, such as $r = 0.49$ between age and SW density in 82 participants (Carrier et al., 2011), $r = 0.28$ for log delta power in 211 participants (Schwarz et al., 2017), and $r = 0.28$ and $r = 0.62$ for relative delta power in 2 subsamples (N = 87 and N = 72, respectively) of the present study (Póthari et al., 2017), highlighting the superiority of a multivariate approach compared with simply using spectral power in the investigation of age-related sleep EEG changes (ForCarrier et al., 2011, the correlation was calculated from $F = 13.9$ (Fig. 2) for the dichotomous age difference, corresponding to $d = 0.835$ and $r = 0.39$ at N = 82. This correlation was corrected for the downward bias of dichotomization, assuming a cutoff at 52.4% of the sample based on subsample sizes, corresponding to the normal deviate $z = 0.06$ and the normal ordinate 0.4. For Schwarz et al., 2017, the correlation was calculated from the t-values in Table 4.). Recently, Sun et al. (2019) could predict 21.2% of age variance from sleep macrostructure, but 67.2% using sleep microstructure, similar to our gross parameters. Still, our study shows that age-related changes in SWs are far from fully captured by changes even by all the gross parameters. The individual average shape of the SW remains correlated with age after regressing out the variance also captured by gross parameters (Fig. 2, panel B; Fig. 3, panel C), and LASSO regression could be trained to capture this correlation to predict even more of the age variance (up to 80%). In other words, SW characteristics are very strongly indicative of the participants’ age, but approximate measures of the SW shape are not sufficient to fully measure age effects. This requires that we take into account

**Fig. 7.** Sex differences in the relationship between age and SW characteristics. Panel (A) shows Pearson’s correlation coefficients by sex between SW fine morphology using individual average frontal SWs based on the 140 μV detection threshold. Amplitude is measured as a negative deflection, resulting in a positive correlation with age. The shaded areas show the 95% confidence intervals of the correlation coefficient. The grand average SW is shown for guidance. Panel (B) shows Pearson’s correlation coefficient by sex between SW gross parameters based on the 140 μV detection threshold. Error bars show 95% confidence intervals. Abbreviations: SW, slow wave; DDS, descending downstate; ADS, ascending downstate; AUS, ascending upstate; DUS, descending upstate.

**Fig. 8.** The mediation of the relationship between age and cognitive performance by SW characteristics. Panel (A): Pearson’s partial correlation coefficients between age and average IQ test raw scores (left) and ZVT raw scores (right) corrected for each of the gross parameters. Green lines show the confidence intervals of the uncorrected correlation coefficient ($r_{ZVT} = 0.49$; $r_{IQ} = 0.22$). Axis x is placed at the uncorrected correlation coefficient so the size and direction of bars show the difference of corrected and uncorrected correlations. Panel (B): Pearson’s partial correlation coefficients (thick black line) between age and average IQ test raw scores (left) and ZVT raw scores (right) corrected for amplitude at each of the 501 SW fine morphology sampling points. The shaded green area shows the confidence intervals of the uncorrected correlation coefficient ($r_{ZVT} = 0.49$; $r_{IQ} = 0.22$). The grand average SW (gray) and the corresponding B-values from the best-fit 140 μV model (red, multiplied by 100 for visibility) are shown for guidance. Abbreviations: SW, slow wave; DDS, descending downstate; ADS, ascending downstate; AUS, ascending upstate; DUS, descending upstate; ZVT, Zahlen verbindungs test. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
the precise waveform. Further analyses revealed several other important details of this effect.

First, a more stringent definition of SWs enables a better prediction of age, suggesting that the discovery of age effects on SWs is more sensitive to false positives than false negatives. Eliminating false positives by choosing a higher detection threshold or averaging data across channels resulted in SW data with higher accuracy in predicting age. Models with higher SW detection thresholds and consequently fewer detected SWs always had higher accuracy (Table 1) and SW detections based on averaged frontal channel data provided the best fit, better than both individual frontal channels or any other (Table 1, Fig. 3). Our findings are in line with those of Sun et al. (2019) who also found that although EEG data using frontal channels could predict age slightly better than central or occipital channels, all were outperformed by models using data from all electrodes.

Second, although not all of the SW-age covariance captured by the fine morphology is captured by gross parameters, the converse is not true. Fine morphology generally provided better fit to age data than only gross parameters—however, a fine morphology-only model provided a fit at least as good as the one achieved by the full models also using gross parameters. In other words, fine morphology beyond its own independent predictive power possesses all of the predictive power of gross parameters as well.

Third, there are clear topographic differences in the correlation between age and SW characteristics. Age effects on SW gross parameters are not the strongest on frontal derivations where SWs are generally the most prominent but rather on more posterior derivations. Older participants produce fewer, smaller, and shallower SWs in more posterior cerebral areas even if frontal SWs are relatively unaffected. This is, however, mainly because at more frontal locations, the gross parameters fail to measure age-related changes that the fine morphology does capture. Full models mainly predict age better because the fine morphology explains additional variance on frontal and central channels (Fig. 3), which is in line with the study by Sprecher et al. (2016) who also found that age-related changes in SW power are widely spread out across the scalp and the study by Sun et al. (2019) who found only small differences between the predictive power of frontal, central, and occipital channels.

Fourth, the correlation between age and SW characteristics is variable across sleep cycles (in our study approximated by sleep quartiles). SW characteristics are most predictive of age in the first cycle, and in this cycle, only models based on average frontal SWs fine morphology could outperform gross parameter-only models. Fine morphology provided the most value in predicting age in cycle 2; its role was reduced in cycle 3 and disappeared by cycle 4. In other words, age effects on SWs are strongest in early sleep, when SWs are generally the most frequent and prominent. Gross parameters, however, do not capture age effects adequately toward midsleep, and fine morphology provides better accuracy. In late sleep, age effects on SWs are generally weaker and fully captured by gross parameters.

SW morphology was generally associated with age during phase transitions, especially late descending and early ascending downstates and late descending upstates (possibly indicating the initiation of the next downstate). These findings suggest that age-related changes in SW morphology especially affect the speed and degree of the initiation and cessation of SW upstate and downstate. These findings are interpretable in light of previous literature using high-density scalp EEG, coregistered invasive and scalp EEG in humans or computer and rodent models. SWs accurately reflect the onset and cessation of cortical neuronal firing, even in human recordings (Csercsa et al., 2010; Nir et al., 2011). Consequently, the shape of the SW reflects how rapidly and to what extent neuronal firing is synchronized (Esser et al., 2007; Riedner et al., 2007; Vyazovskiy et al., 2007). In line with the view that SWs reflect a homeostatic response to long-term potentiation taking place in the preceding wakefulness (Tononi and Cirelli, 2014), frequent, high-amplitude SWs with steep slopes are seen under conditions of high homeostatic pressure for sleep, such as early in the night (Esser et al., 2007; Riedner et al., 2007) or in case of young participants with a higher degree of neuronal plasticity (Feinberg and Campbell, 2013). Based on these studies, the current results can be interpreted as a more rapid initiation and cessation of synchronized cortical neuronal firing in younger participants, in line with the finding that in older participants, both the homeostatic need for SWs and the ability to respond to this need by SW generation is reduced (Mander et al., 2017). In line with previous findings, age effects were strongest in the first quartile of sleep when sleep pressure is the strongest and approximately equally distributed across the scalp (Sprecher et al., 2016).

Age-related changes in any SW measure are important because alterations in NREM sleep are thought to play a functional role in age-related cognitive decline (Mander et al., 2017). In one study, NREM SW activity reductions statistically mediated the age-related reduction in overnight memory consolidation (Mander et al., 2013). Following up animal experiments demonstrating the effect of sleep deprivation on beta amyloid accumulation (for a summary, see [Ju et al., 2013]), a structural equation modeling study demonstrated that the relationship between beta amyloid accumulation and reduced memory function is mediated by SW activity (Mander et al., 2015). This suggests that an important part of Alzheimer pathology may be related not directly to the beta amyloid burden but to the resulting disruption of NREM sleep (Lucy and Holtzman, 2015). Although it would be tentative to see both the cognitive symptoms and NREM sleep changes in Alzheimer’s disease as an extreme case of normal aging, it must be considered that while normal aging involves a reduction of a wide range of slow frequency activity not limited to delta, very low frequencies are especially affected in Alzheimer’s disease (Mander et al., 2017). The molecular composition of beta amyloid proteins is also different in normal aging and Alzheimer’s disease (Piccini et al., 2005). Therefore, it requires further research to establish the degree to which SW changes in normal aging and in dementia are analogous.

We found that although the mediation of the age-related reduction on cognitive test performance by SW characteristics remained a trend, they were strongest at the SW phase transitions where the SW fine morphology was independently associated with age. Our cognitive tests measured either general cognitive ability and test-specific abilities (Raven APM, CFT) (Gignac, 2015) or a combination of general cognitive ability and mental speed (ZVT) both of which are known to be strongly negatively associated with aging (Salthouse, 2004). In line with this observation, ZVT performance was the most strongly correlated with age. Although this effect did not reach statistical significance, the correlation between age and cognitive performance was attenuated after statistically controlling for some SW characteristics, especially SW amplitude around phase transitions, suggesting that the neurobiological mechanism underlying reduced SW steepness may also play a role in age-related reductions in cognitive performance. Further research with a more detailed neuropsychological assessment battery and preferably with a larger sample size should investigate the role SW characteristics play in mediating age-related reductions in cognitive performance.

This has important translational potential as extant evidence already suggests that cognitive functioning in wakefulness, also in old age, can be improved by artificially increasing SW activity in NREM sleep through heating and auditory or electrical stimulation (Ladenbauer et al., 2017; Papalambros et al., 2017; Wilckens et al.,...
SW characteristics that are particularly affected by aging, especially if they also statistically mediate age-related reductions in cognitive performance, can be especially promising targets of such restorative interventions. Our study provides preliminary evidence that SW phase transition speed and steepness may be one such characteristic.

Our study suffers from a number of limitations. First, due to the cross-sectional nature of our study, we had a limited ability to establish cause-effect relationships. Second, although our participants spent 2 nights in the laboratory, no EEG recording was performed on the first night. Therefore, we are also unable to assess the internight reliability of predicting age from EEG performance and the possible role of first-night effects, and our results are representative for the second night. Third, although we used multiple methods to detect SWs to test the robustness of the prediction of age from SW characteristics, our methodology did not cover the full range of existing SW detection algorithms.

Disclosure

All authors do not declare any actual or potential conflicts of interest.

Acknowledgements

The research was supported by institutional funding of the authors. Participants provided written informed consent according to approval of the local ethics boards in Munich and Budapest. Péter P. Ujma was supported by the ÚNKP-17-4 New National Excellence Program of the Ministry of Human Capacities. Péter Simor was supported by the Hungarian Scientific Research Fund (NKFI FK 128100) of the National Research, Development and Innovation Office and by the ÚNKP-18-4 (Bolyai +) New National Excellence Program of the Ministry of Human Capacities and by the Bolyai János Research Scholarship of the Hungarian Academy of Sciences.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neurobiolaging.2019.04.002.

References

Gignac, G.E., 2015. Raven's is not a pure measure of general intelligence: implications for g factor theory and the brief measurement of g. Intelligence 52, 71–79.


