Human sensorimotor resting state beta events and aperiodic activity show good test–retest reliability

K. Amande M. Pauls a,b,⇑, Pietari Nurmi c, Heidi Ala-Salomäki c, Hanna Renvall a,c, Jan Kujala d, Mia Liljeström a,c,e

⇑ Corresponding author.
E-mail address: amande.pauls@hus.fi (K. Amande M. Pauls).

Abstract

Objective: Diseases affecting sensorimotor function impair physical independence. Reliable functional clinical biomarkers allowing early diagnosis or targeting treatment and rehabilitation could reduce this burden. Magnetoencephalography (MEG) non-invasively measures brain rhythms such as the somatomotor ‘rolandic’ rhythm which shows intermittent high-amplitude beta (14–30 Hz) ‘events’ that predict behavior across tasks and species and are altered by sensorimotor neurological diseases.

Methods: We assessed test–retest stability, a prerequisite for biomarkers, of spontaneous sensorimotor aperiodic (1/f) signal and beta events in 50 healthy human controls across two MEG sessions using the intraclass correlation coefficient (ICC). Beta events were determined using an amplitude-thresholding approach on a narrow-band filtered amplitude envelope obtained using Morlet wavelet decomposition.

Results: Resting sensorimotor characteristics showed good to excellent test–retest stability. Aperiodic component (ICC 0.77–0.88) and beta event amplitude (ICC 0.74–0.82) were very stable, whereas beta event duration was more variable (ICC 0.55–0.7). 2–3 minute recordings were sufficient to obtain stable results. Analysis automatization was successful in 86%.

Conclusions: Sensorimotor beta phenotype is a stable feature of an individual’s resting brain activity even for short recordings easily measured in patients.

Significance: Spontaneous sensorimotor beta phenotype has potential as a clinical biomarker of sensorimotor system integrity.

Keywords: Magnetoencephalography Resting state Sensorimotor Beta oscillatory activity Aperiodic (1/f) activity Test–retest reliability

1. Introduction

Some neurologic diseases affecting the motor system, such as Parkinson’s disease, are difficult to diagnose at their early stages due to lack of easily observable brain structural changes. Furthermore, disease trajectories and rehabilitation outcomes are variable and often unpredictable. Currently, biomarkers for estimating individual disease courses are lacking. Functional biomarkers reflecting the processes underlying motor dysfunction might help in the differential diagnostics, or in estimating the rate of disease development or the recovery potential in individual patients. Such markers could also improve targeting of treatment and rehabilitation.

Non-invasive electrophysiological recordings, such as electroencephalography (EEG) and magnetoencephalography (MEG),
measure brain activity resulting from the spatial and temporal summation of cellular neural activity of the brain's underlying cortical areas (Buzsáki et al., 2012). The measured activity depends on factors such as neuronal density, size and shape, the anatomy of neural network connections, and their relative activity at any given point (Buzsáki et al., 2012). Thus, MEG and EEG measures reflect the effect of different structural and functional changes in cortical activity and have considerable potential as functional biomarkers.

One promising candidate functional biomarker is the sensorimotor, or ‘rolandic’, 20-Hz beta rhythm which is observed consistently in humans (Hari and Salmelin, 1997) and across other species (Feingold et al., 2015; Haegens et al., 2011; Sherman et al. 2016). Cortical beta activity plays an integral role in several perceptual and cognitive functions and is modulated in a variety of tasks including tactile processing (Haegens et al., 2011; Pfurtscheller et al., 2001), movement (Feingold et al., 2015; Salmelin and Hari 1994), action perception (Babiloni et al., 2002; Hari et al., 1998) and attention (Sacchet et al., 2015; Ede et al., 2011).

Cortical beta band activity displays a characteristic pattern of bursting over time, occurring in intermittent high amplitude ‘beta events’ alternating with lower amplitude periods (Feingold et al., 2015; Jones, 2016). Beta activity is particularly patterned in the sensorimotor cortex (Seedat et al., 2020), where beta event rate predicts behavioral outcome in humans and rodents across tasks (Shin et al., 2017). In humans, spontaneous resting EEG beta band power (Smit et al., 2005; Van Beijsterveldt et al., 1996), as well as beta event parameters (Pauls et al., 2024) have been shown to be heritable.

Neurological conditions with related motor dysfunction, such as stroke (Bartur et al., 2019; Laaksonen et al., 2013, 2012; Parkkonen et al., 2018; Schulz et al., 2021; Rossiter et al., 2014), Parkinson’s disease (Pauls et al., 2022; Vinding et al., 2020) and amyotrophic lateral sclerosis (ALS) (Dukic et al., 2022; Proudfoot et al., 2016) are associated with changes in the sensorimotor cortical beta band signal. Furthermore, sensorimotor beta characteristics correlate with symptom severity (Bartur et al., 2019; Laaksonen et al., 2012; Parkkonen et al., 2018; Pauls et al., 2022; Rossiter et al., 2014) and clinical recovery (Laaksonen et al., 2012, 2013; Parkkonen et al., 2018).

In addition to rhythmic, or ‘periodic’, components, spontaneous cortical activity also contains prominent aperiodic (1/f) components which show exponential decay characteristics. The exponent parameter of the 1/f signal is postulated to reflect excitation-inhibition balance (Gao et al., 2017), arousal and wakefulness (Lendner et al., 2020) and it is modulated, e.g., by attention (Waschke et al., 2021), brain maturation (Hill et al., 2022; McSweeney et al., 2021; Tröndle et al., 2022) and aging (Voytek et al., 2015; Wilson et al., 2022). 1/f behaviour is highlyheritable (Pauls et al., 2024) and appears to be altered in several neurological and neuropsychiatric conditions, such as Parkinson’s disease (Helson et al., 2023), dystonia (Semenova et al., 2021) and ADHD (Ostlund et al., 2021).

Takent together, these studies have shown that sensorimotor beta activity and aperiodic fluctuations are prominent, spontaneously occurring and heritable characteristics of ongoing brain activity that are closely linked to sensorimotor functions, are preserved across mammalian species and show changes associated with sensorimotor symptoms in different neurological disease conditions. Given their salience and ease of acquisition, they are thus potential cortical biomarkers of sensorimotor disease state and its reactivity to treatment. However, a prerequisite for using them as biomarkers is their intraindividual signal stability. In general, test–retest reliability for various MEG responses of clinical interest is good, as suggested, e.g., for measures of somatosensory processing (Illman et al., 2022; Pitolulainen et al., 2018), picture naming (Ala-Salomäki et al., 2021), as well as whole-brain spontaneous oscillatory power (Martin-Buro et al., 2016) and resting state functional connectivity (Garcés et al, 2016). Earlier EEG studies have also demonstrated good test–retest reliability of global beta band power at rest (Fingelkurts et al., 2006; Pollock et al., 1991). Decomposing cortical sensorimotor activity into its different dynamic components, or ‘beta events’, adds detail compared to the assessment of mere beta power globally, and the different beta event parameters’ test–retest reliability has not been assessed before. Therefore, we determined whether and to what extent different spontaneous sensorimotor beta event parameters and aperiodic activity are reliable across sessions.

2. Methods

2.1. Subjects

50 healthy subjects (age mean +/- STD 45 +/- 20 years, range 21–70 years, one ambidextrous, all other right-handed, handedness assessed using the Edinburgh Handedness Inventory) screened to exclude pre-existing neurological disorders, learning disabilities, and language disorders were included in the study after giving written informed consent. The study was approved by the Aalto University ethics committee and carried out in accordance with ethical guidelines set out in the Declaration of Helsinki.

2.2. MEG recordings

Measurements were performed in a magnetically shielded room (Imedco AG, Hägendorf, Switzerland) with a 306-channel Vectorview neuromagnetometer (MEGIN Oy, Helsinki, Finland) consisting of 204 planar gradiometers and 102 magnetometers. The data were measured at MEG Core, Aalto NeuroImaging, Aalto University School of Science. Spontaneous cortical activity was recorded with a 1 kHz sampling rate, continuous head position monitoring (cHPI) and band-pass filtering at 0.03–330 Hz in two separate sessions one-two weeks apart, for five minutes each, while participants were resting with their eyes open. Vigilance was assessed via video during measurements, via behavioural control (eye movements, blinking), and via monitoring of any progressive slowing of the resting-state activity before further analysis.

2.3. MEG signal processing and parameter extraction

Subjects’ data were assessed visually for prominent vigilance effects (e.g. slowing) and to exclude artifacts, and periods with significant artifacts were excluded from further analysis. Overall, data quality was good and there was only minor data loss due to artefact removal (remaining mean recording length 330 seconds, range 306–515 seconds). For suppressing external artifacts, MEG data were preprocessed using the temporally extended signal space separation method (tSSS, (Taulu and Simola 2006) implemented in the MaxFilter software (MEGIN Oy, Helsinki, Finland, version 2.2.15)). Subject’s head movements were compensated based on the cHPI recordings, and individual MEG recordings were transferred to a common head space using MaxFilter’s signal space separation-based head transformation algorithm. One subject was excluded because head transformation to the common space introduced considerable noise, compromising data quality. Further signal processing was done using MNE-python version 1.3 (Gramfort et al. 2013). After band-pass filtering the data to 2–48 Hz with a one-pass, zero-phase, non-causal FIR filter (MNE firwin filter using a Hamming window), power spectral densities (PSD) were calculated using Welch’s method with a non-
overlapping Hamming window and 2048-point Fast Fourier transformation (frequency resolution ~0.5 Hz).

2.4. Channel selection

The subsequent analysis steps are illustrated in Fig. 1 extending the approach previously used (Pauls et al., 2022, 2024). For each hemisphere, we defined a region of interest (ROI) of 15 gradiometer channel pairs per hemisphere centered over the sensorimotor cortices. Pairs of gradiometer channels were combined into one vector PSD. The resulting 15 vector PSDs per hemisphere were then decomposed into a periodic and an aperiodic component using the FOOOF algorithm (Donoghue et al., 2020) in the frequency range from 2–48 Hz and using FOOOF’s default aperiodic fitting mode (‘fixed’). The FOOOF algorithm operates on PSDs in semilog-power space with linearly spaced frequencies (~0.5 Hz frequency resolution in this case), and log-spaced power values. The aperiodic component, \( L \), is modeled using a Lorentzian function:

\[
L = b - \log(k + F^v)
\]

where \( b \) is the broadband offset, \( F \) is the vector of input frequencies, \( v \) is the exponent, and \( k \) is the “knee” parameter, controlling for the bend in the aperiodic component. Here, we did not fit a bend in the aperiodic component, and the knee parameter param-

---

**Fig. 1. Sensorimotor phenotyping.** A. Channel and peak frequency selection. In a predefined region of interest (ROI), channel pairs were combined using vector sum calculation, and the aperiodic (1/f) component was extracted using FOOOF as described previously (Pauls et al. 2022, 2024). Peak channel and frequency within the ROI were selected either (1) manually by selecting the channel with the highest beta spectral peak frequency, (2) in an automated fashion by selecting the channel with the highest periodic beta power, or (3) using automation with visual-manual correction of the selected frequency and channel if necessary. B. Beta event extraction. The raw signal was convolved with a set of complex Morlet wavelets, and the resulting signal averaged either broad-band (14–30 Hz) or narrow-band (+/-2 Hz around the peak beta frequency) to obtain an amplitude envelope which was thresholded at the 75th percentile (red line). Periods exceeding this threshold are defined as ‘beta events’. C. Resting sensorimotor phenotype parameters included PSD peak power and frequency in the 14–30 Hz beta band, total periodic beta power, 1/f exponent (\( \chi \)) and offset, as well as beta event characteristics including event duration and amplitude mean, median, standard deviation, robust maximum, event rate and dispersion (\( CV \)). A short description of all parameters can also be found in Table 2.
eter was hence set to $k = 0$. Note that in FOOOF the decay across frequencies is modeled via the Lorentzian coefficient $\chi$ (chi), whereas when an exponential decay function is used to model $1/f$ behavior the decay is represented by $\lambda$ (lambda).

To flatten the spectrum and reveal remaining periodic components, the aperiodic component exponential decay curve was first calculated using the FOOOF algorithm, and subtracted from the original PSD. The remaining periodic component was then plotted for all 15 vector PSDs. From these flattened vector PSD spectra, the channel pair with the most prominent spectral peak in the beta range was selected per hemisphere (‘the peak channel pair’) and the frequency of the power peak was noted (‘peak beta frequency’) (see Fig. 1A). This choice of channel and frequency was carried out in three different ways: (1) an entirely automated approach, where the ‘peak channel’ was selected based on the area under curve (AUC) of the periodic part of the PSD between 14 and 30 Hz, and the peak was detected automatically as the highest amplitude in this frequency range; (2) a manual peak detection approach, where vector PSD plots were visually inspected, the frequency of the highest peak noted, and the channel with the maximum amplitude at this frequency selected as the ‘peak channel’; and (3) a combined approach where peak channel and peak frequency were selected automatically as described, but all plots underwent visual control and the assignment of peak frequency (and sometimes channel) was re-adjusted if necessary. Typical reasons for correction of the peak frequency (and channel) were cases with strong alpha peaks and weak beta peaks, where the automatically detected beta peak was located on the shoulder of the alpha frequency peak.

### 2.5. Beta event extraction

The channel pair and peak beta frequency corresponding to the chosen peak vector PSD were used for beta event analysis (see Fig. 1B). The chosen peak sensors were identical or adjacent to each other in 46 out of 50 subjects in the left hemisphere, and in 45 (out of 50 subjects) in the right hemisphere. The channel pair’s raw, unfiltered time series data were downsampled to 200 Hz, high pass filtered at 2 Hz and decomposed by convolving the signal with a set of complex Morlet wavelets within the frequency range of 7–47 Hz with 1 Hz resolution and $n_{cycles} = frequency/2$. After this, the amplitude envelope was derived by averaging the signal within a certain beta frequency range. This was done either (1) broadband across the entire beta frequency band (14–30 Hz) or (2) narrow-band, i.e., $\pm 2$ Hz around the individual peak beta frequency chosen in one of the three ways described above. The vector sum over both channels’ beta time series was calculated and rectified, to obtain one beta band amplitude envelope per channel pair. The envelope was smoothed with a 100-ms FWHM kernel and thresholded at the 75th percentile value. Periods exceeding this threshold for 50 ms or longer were defined as beta events. For event amplitude and event duration, the mean, median, robust maximum (defined as mean of the top 5% values) and standard deviation values were calculated (see Fig. 1C for illustration). Furthermore, events per second (event rate) and event dispersion were calculated as described previously (Pauls et al. 2022). Times between beta events were defined as waiting times. To estimate patterning of beta events, the variation of waiting times (henceforth referred to as ‘event dispersion’), we calculated the coefficient of dispersion $C_v$ (Shinomoto, Miura, and Koyama 2005), defined as the waiting times’ standard deviation $\sigma$ divided by their mean $\mu$:

$$C_s = \frac{\sigma}{\mu}$$

For a series of waiting times between events which are independently distributed, $C_v$ takes a value of 1. For an entirely regular event sequence with constant intervals, its value will be 0. If the value is larger than 1, the waiting times are over-dispersed, pointing to the occurrence of clusters of events. All values were calculated for both hemispheres in all subjects to obtain a sensorimotor signature phenotype (Fig. 1C). This ‘resting sensorimotor phenotype’ features included PSD peak power and frequency in the 14–30 Hz beta band, total periodic beta power, 1/f exponent (chi, also referred to as lambda) and offset, as well as beta event characteristics including event duration and amplitude mean, median, standard deviation, robust maximum, event rate and dispersion (see Fig. 1C and Table 2).

Beta event extraction was tested for a range of parameters to investigate their effect on the phenotype results and their stability. The narrow band filter bandwidth was varied (+/- 1, 2, 3, 4, 5 Hz, and broad-band), and the amplitude threshold was tested for different percentiles (50th, 60th, 70th, 75th, 80th, 85th and 90th). The effect of the amount of data (i.e. duration of recording) was also examined by using successively longer data segments (60, 90, 120, 150, 180, 210, 240, 270 and 300 seconds) to estimate the parameters. This was done starting from the beginning of the recording, from the end of the recording, and from the middle of the recording.

### 2.6. Test-retest reliability analysis

Test-retest reliability was assessed using the intraclass correlation coefficient implemented in Pingouin (intraclass_corr function). ICC analyses were conducted on all the phenotype parameters described in Fig. 1C, derived from the same gradiometer channel pair for all parameters per session, by comparing the outcomes between sessions 1 and 2. ICC is defined as follows:

$$ICC(3, 1) = \frac{BMS - EMS}{BMS + (k - 1) + EMS}$$

where $BMS = between-subjects mean square, EMS = error mean square,$ and $k = number of sessions. ICC below 0.4 is considered poor, 0.40–0.59 fair, 0.60–0.74 good, and 0.75–1.00 as excellent consistency (Cicchetti 1994). We used ICC(3,1), with a fixed set of sessions ($n = 2$) during each of which all phenotype parameters were assessed.

### 2.7. Code and data availability

Data cannot be made publicly available due to Finnish data protection law. Data can, however, be shared for research collaboration with an amendment to the research ethics permit and a related data transfer agreement. All analysis code will be made available on GitHub.

### 3. Results

#### 3.1. Test-retest reliability

Test-retest reliability was good or excellent for many of the phenotype features. Fig. 2 shows example scatter plots for three of the beta event parameters and their test–retest reliability in the left hemisphere.

The three approaches using individual narrow-band filtering for determining beta events produced very similar results, whereas using a broad-band filter systematically altered the results (see Fig. 3, discussed in more detail later). Entirely manual peak assignment did not significantly improve results and was very labor-intensive. On the other hand, automated peak assignment missed the PSD beta peak that was by visual evaluation the best one in 14% of the cases. In most of these cases (86%), automatic peak assignment missed the periodic peak, usually in favor of a lower
frequency peak located on the ‘shoulder’ of a large alpha frequency peak. In the remaining 14%, the peak channel changed because the periodic beta peak was visually more distinct in a different channel. We thus chose to work with the approach combining automated peak selection with manual control.

Intraclass correlation coefficient (ICC) values for this approach (‘automated + manual’) are given in Table 1, and for all four approaches in Supplementary Table 1. Table 2 provides a description of all the parameters obtained from the data. Test-retest reliability for most parameters was good or excellent, in particular for the 1/f parameters. Some parameters, notably event dispersion, proved poor (Fig. 2). Interestingly, parameters of event amplitude were more reliable in the left hemisphere, whereas event duration parameters were more stable in the right hemisphere.

3.2. Broad- vs. narrow band event characteristics

Use of broad-band filtering to determine beta event characteristics systematically affected event parameters, shortening event duration, increasing event rate and reducing event amplitudes compared to narrow-band event extraction (see also Fig. 3). Furthermore, broad-band filtering decreased the signal to noise ratio (SNR), as illustrated in Fig. 4.

3.3. Effect of recording length and processing parameters on ICC

Finally, we tested the effect of event extraction parameters (filtering bandwidth & amplitude threshold) as well as recording length on the test–retest reliability (see Fig. 5). Overall, event amplitude showed good or excellent reliability and was relatively invariant to different parameters except for recording duration (Fig. 5, middle column). Event duration was more noise- and parameter-sensitive (Fig. 5, first column). Event rate reliability was fairly stable for longer recording durations. Event dispersion had the lowest reliability of all parameters, also at longer recording durations (Fig. 5, last column).

ICC stabilized reasonably well at 2-min recording length for most beta event parameters, so even short recordings may be enough for obtaining reliable results (Fig. 5, top row). However, event rate estimates benefitted from longer recordings. A bandwidth of 4 Hz was optimal for event duration assessment as well as event rate, after which ICC decreased somewhat (Fig. 5, middle row). Event amplitude appeared invariant to bandwidth. 70–80% percentile thresholds were optimal for both the event duration as well as the event rate parameter, while event amplitude was mostly invariant to this parameter (Fig. 5, bottom row). Only dispersion appeared to benefit from higher percentile thresholds, rais-
Table 1
Test-retest reliability of different resting sensorimotor parameters between two measurement sessions (automated peak selection with manual control). Italics indicate ICC >= 0.75, **bold italics** ICC >= 0.75. CI95 = 95% confidence interval (lower and upper bound); F - F statistic. * The degrees of freedom of the numerator (d1) and the denominator (d2) are 48 for all parameters.

<table>
<thead>
<tr>
<th>PSD parameters</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC</td>
<td>CI95</td>
</tr>
<tr>
<td>peak beta frequency</td>
<td>0.90</td>
<td>[0.83 0.94]</td>
</tr>
<tr>
<td>peak beta amplitude</td>
<td>0.61</td>
<td>[0.39 0.76]</td>
</tr>
<tr>
<td>1/f chi</td>
<td>0.84</td>
<td>[0.74 0.91]</td>
</tr>
<tr>
<td>1/f offset</td>
<td>0.88</td>
<td>[0.8 0.93]</td>
</tr>
<tr>
<td>total periodic beta power</td>
<td>0.82</td>
<td>[0.69 0.89]</td>
</tr>
</tbody>
</table>

Table 2
Description of parameters for which ICC was calculated.

<table>
<thead>
<tr>
<th>PSD parameters</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak beta frequency</td>
<td>Hz</td>
<td>frequency at maximum beta range PSD periodic peak</td>
</tr>
<tr>
<td>peak beta amplitude</td>
<td>(fT/cm)^2</td>
<td>maximum amplitude of the beta range PSD periodic peak</td>
</tr>
<tr>
<td>1/f chi</td>
<td>unitless</td>
<td>exponent of the aperiodic PSD component</td>
</tr>
<tr>
<td>1/f offset</td>
<td>unitless</td>
<td>offset of the aperiodic PSD component</td>
</tr>
<tr>
<td>total periodic beta power</td>
<td>(fT/cm)^2</td>
<td>area under curve (AUC) of the periodic PSD component in the beta range (not including aperiodic PSD part AUC)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Beta event parameters</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>duration mean</td>
<td>milliseconds</td>
<td>mean duration of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>median</td>
<td>milliseconds</td>
<td>median duration of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>standard deviation</td>
<td>milliseconds</td>
<td>standard deviation of duration of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>maximum</td>
<td>milliseconds</td>
<td>maximum duration of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>robust maximum</td>
<td>milliseconds</td>
<td>duration of maximum 5% of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>amplitude mean</td>
<td>fT/cm</td>
<td>mean amplitude of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>median</td>
<td>fT/cm</td>
<td>median amplitude of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>standard deviation</td>
<td>fT/cm</td>
<td>standard deviation of amplitude of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>maximum</td>
<td>fT/cm</td>
<td>maximum amplitude of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>robust maximum</td>
<td>fT/cm</td>
<td>amplitude of maximum 5% of all observed events &gt; 50 ms</td>
</tr>
<tr>
<td>event rate</td>
<td>1/s</td>
<td>number of events/measurement time</td>
</tr>
<tr>
<td>event dispersion (Cv)</td>
<td>unitless</td>
<td>standard deviation/mean for the waiting times, if 0, entirely regular event, if 1, exponential decay behaviour, if &gt;1, clustering of events</td>
</tr>
</tbody>
</table>

The ICC values for the 1/f parameters were independent of the FOOF fitting range (2/3/4–47 Hz).

4. Discussion

We demonstrate that human cortical sensorimotor dynamic cortical beta event parameters and 1/f characteristics as measured with resting-state MEG show good test–retest reliability. The results were robust across a range of analysis parameters including different filtering bandwidths and amplitude thresholds and appear to be stable even for just 2–3 minutes of recording for many parameters. The results suggest that resting sensorimotor beta phenotype is a stable feature of an individual's brain activity with good potential as a clinical biomarker.

4.1. Earlier studies of test–retest reliability of sensorimotor functional measures

Some previous studies have assessed sensorimotor system spontaneous and task-related measures' test–retest reliability. In an MEG study assessing spontaneous resting-state oscillatory beta band power stability, ICCs ranged from 0.74 to 0.86 for frontal and parietal brain areas (Martín-Buro, García, and Maestú 2016), comparable to our results for the left hemisphere beta power. In their study, beta power was separated into low and high beta power (13–20 vs. 20–30 Hz). Furthermore, the total band power included the aperiodic signal component, which was assessed separately.
here. Cortico-kinematic coherence, a measure of proprioception-related brain processing, has been demonstrated to have very good stability, 0.86 for the dominant and 0.97 for the non-dominant hand (Piitulainen et al. 2018). Stimulus-related sensorimotor beta suppression and rebound phenomena in response to sensory (tactile and proprioceptive) stimuli also show good to excellent stability (Illman et al. 2022). However, task-related functional measures rely on some degree of collaboration (and preserved function) from the subject, which can limit applicability in clinical settings. Thus, brief measurements of spontaneous brain activity as used here extend the spectrum of possible applications.

4.2. Are there hemispheric differences in test–retest reliability?

We found a tendency for hemispheric differences in ICC for some of the parameters. The test–retest reliability was slightly higher for the left, dominant hemisphere for the amplitude parameter, probably reflecting the fact that left hemisphere spectra tend to have clearer periodic signal components. Interestingly, the beta event duration parameters as well as event rate were more reliable for the right hemisphere, even though defining the peak was more difficult. A possible explanation for this result is that the dominant left hemisphere has more variable resting activity. On the other hand, the right hemisphere duration parameter may be more stable than the corresponding value in the left hemisphere because of its relatively lower SNR, thus less reflecting the actual periodic beta signal fluctuations and more the general noise level or other, non-periodic activity. This effect can also be seen in the broad vs. narrow band signal extraction: broad-band signal extraction produces good ICC values, especially for the right hemisphere. The ICC difference between ‘broad’ and ‘narrow’ band extraction strategies is bigger for the left hemisphere with a more pronounced PSD beta peak. The difference could also be caused by measurement-related effects: slight natural head tilt to one side could cause the head to be closer to the sensors on one side compared to the other. Here, a head position correction procedure was employed to correct for head position, but differences in signal to noise ratio present before head correction would still remain, making the results there less consistent.

4.3. Potential of resting sensorimotor phenotype as a biomarker

Are the ICC values obtained here sufficient to warrant use as a biomarker? There are no standard values determining acceptable reliability using ICC, but different suggestions have been made for their interpretation (Cicchetti 1994; Koo and Ma, 2016). ICC values above 0.6 (Cicchetti 1994) or 0.75 (Koo and Ma, 2016) have been considered to indicate good test–retest agreement, and values above 0.75 (Cicchetti 1994) or 0.9 (Koo and Ma, 2016) to indicate excellent test–retest agreement. Thus, many of the described event parameters in the present study show good to excellent test–retest reliability. A low ICC can relate to low test–retest agreement but can also relate to lack of variability among subjects (small dynamic range), small number of subjects or a low number of repetitions. The number of subjects in the current study should be sufficient to obtain reasonable ICC values and was the same for all studied parameters. However, the dynamic range was low, e.g., for event dispersion, with most subjects clustering in a very limited range of values, possibly contributing to the low ICC. Overall, the level of test–retest reliability obtained in the current study was good, supporting the use of the described features of interest also in clinical settings. This notion is supported by the confidence intervals that were observed. For the measures that showed good or excellent test–retest agreement, the lower 95% confidence limits suggested fair, good, or even excellent test–retest agreement. However, for the measures showing only fair test–retest agreement (e.g., event dispersion), the lower confidence limits were < 0.2, suggesting that these measures may not serve as viable biomarkers.

Besides technical considerations, biological and pathophysiological considerations are also important for biomarker development. As outlined earlier, sensorimotor beta activity and dynamic beta events are detectable across different mammalian species including humans, non-human primates and rodents (Feingold et al., 2015; Haegens et al., 2011; Hari and Salmelin, 1997;
and they have been found to be heritable (Pauls et al., 2024; Smit et al., 2005; Van Beijsterveldt et al., 1996), suggesting that the brain's sensorimotor signature is quite preserved across the sensorimotor system's evolution. Interestingly, our results show that certain beta event characteristics, e.g., left hemispheric aperiodic behavior and several beta event amplitude measures, showed high test–retest reliability, even though they do not evidence significant heritability (Pauls et al., 2024).

The present results thus further support our previous finding that beta characteristics have variable heritability, possibly related to their underlying generation mechanisms. Furthermore, previous studies have shown disease-related beta changes at the group level (Bartur et al., 2019; Dukic et al., 2022; Laaksonen et al., 2013, 2012; Parkkonen et al., 2018; Pauls et al., 2022; Vinding et al., 2020; Schulz et al., 2021). Finally, the human brain’s sensorimotor system has been extensively studied and is quite well understood.

Fig. 5. Effect of different event extraction parameters on session-to-session event parameter stability as assessed by ICC. Panels depict the following parameters. Top row - effect of recording duration (s); middle row - peak bandwidth (Hz); bottom row - % amplitude threshold, with ICC on the y-axis. Columns. 1st - event duration, 2nd - event amplitude, 3rd - event rate & dispersion. Solid lines - left hemisphere, dashed lines - right hemisphere.
We suggest that these factors, in combination with good test–retest reliability demonstrated here, make sensorimotor beta activity a good candidate electrophysiological biomarker. It is conceivable that test–retest behavior might be different and the signal potentially more heterogeneous or unstable in clinical cohorts, so this will need to be investigated further. Moreover, future studies should also monitor and address in more detail the possible role of cardiac and motor artifacts on the test–retest agreement of the sensorimotor beta activity in different cohorts.

4.4. Limitations and future directions

Although the analysis approach used here was largely automated, it still required some manual preprocessing (adjustment of peaks in some cases with weak beta peak). The need for a human observer is always associated with a certain degree of observer bias, and different observers and experience levels can lead to increased levels of uncertainty. Furthermore, human observers need to be trained to make the process as reliable as possible. Here, only one observer (AP) carried out all manual beta peak selection. For now, human control of results is necessary, and better automation procedures are needed to automate beta event characterization entirely. Alternatively, e.g. more observer-independent methods to assess fluctuations in beta state (e.g. using Hidden Markov Modelling as used in (Seedat et al., 2020)) could reduce the need for human observers. Recently, approaches to determine transient neural events using threshold methods based on estimates of 1/f background behaviour have also been proposed (Seymour et al., 2022; Brady and Bardouille 2022). However, all of these methods will still require some human-observer based choices if looking at specific frequencies as opposed to spectral behaviour across an entire frequency band (e.g. beta band).

Furthermore, sources of variations in signal to noise ratio (SNR) in PSD spectra need to be explored. Some subjects have relatively small periodic components in their PSD spectra for unknown reasons, and poor PSD SNR is not always clearly attributable to measurement noise. In many studies, subjects with poor SNR are excluded before further data analysis. However, after excluding clear problems with measurement quality (external noise), it would be interesting to explore reasons for poor SNR, which might in fact be related to brain processing features or brain state in these subjects. Here, we excluded one subject due to SNR considerations arising from problems related to data collection (poor head positioning during one session). Some subjects had bigger session to session fluctuations than others due to factors not obviously related to measurement factors. Future studies should assess the sources of session-to-session variability, e.g., factors such as vigilance. We assessed vigilance clinically during the measurement (via video) and the eyes open resting condition was used. The raw data was visually assessed for vigilance effects also (exclusion of gradual slowing of activity) to ensure steady vigilance levels. Thus, major fluctuations in vigilance have been excluded, but small changes in alertness or habituation effects across sessions are possible. Quantitative, automated vigilance assessment may also be helpful in the future.

We here show good test–retest reliability for 1/f behavior, an important component of power spectral activity which is still incompletely understood. 1/f subserves several factors, including both subject-related as well as non-subject related factors (such as measurement environment and device). Noise components related to the measurement environment were suppressed using spatiotemporal signal space separation (Taulu and Simola 2006). Noise from the measurement device is expected to be very weak compared to physiological noise from the participants, and is thus unlikely to significantly contribute to the 1/f results. Subjects showed good test–retest reliability (self-sameness) between two measurement sessions carried out on separate days one week apart, suggesting that the 1/f parameters are subject-related and subject-specific. However, as subject-specific physiological signals originating outside the brain (such as eye-movements and cardiac signals) can also show 1/f behavior, it is possible that the 1/f test–retest reliability is also influenced by non-brain signals. For improved interpretability, a more detailed understanding of factors contributing to 1/f signal and their relative weighting in signal generation is needed in the future.

The resting sensorimotor phenotype was assessed at the sensor-level. While source-level approaches add some level of spatial-anatomical resolution, they also introduce more data processing and analysis choices, making the approach less feasible for potential clinical applications and possibly leading to biased estimates. Furthermore, clinically useful source-level analyses would necessitate suitable cortical parcellations to avoid multiplying the amount of data. As MEG is most sensitive to sulcal brain activity, established parcellations taking this into account would be needed. Here, we were specifically interested in testing the reliability of a simple approach using sensor-level data to characterize the resting sensorimotor phenotype for potential clinical applications. If the potential caveats of source-level analyses (amount of data processing, automated analyses, suitable parcellations) can be solved, future studies should address whether more reliable measures can be obtained at the source- compared to the sensor-level.

4.5. Pipeline recommendation

In the current study, recording durations of 2–3 minutes were sufficient to get stable beta event results for most parameters. For most of the examined parameters, short good-quality data segments were preferable to longer data segments with more variable data quality. Only the event rate, and to some extent the event duration parameters, benefited from longer recording times.

Automation of beta peak detection was successful for 86% of the hemispheres: We recommend at least visual control to ascertain correct beta peak assignment. In the future, optimized automatization approaches which work for most subjects would be helpful to eliminate human observer bias. Alternatively, broad-band (13–30 Hz) beta event extraction can be done, but use of a specific beta peak channel and beta band peak for extraction of beta event information increased stability of the beta event duration parameter. A beta amplitude threshold of 75% and 4 Hz bandwidth appears appropriate, giving a good reliability for all beta event parameters. Estimation of event dispersion had low reliability but improved at higher percentage thresholds, so this parameter might require different processing settings.

5. Conclusions

In summary, we demonstrate that a robust resting-state sensorimotor phenotype with good or excellent test–retest stability can be obtained from MEG data in healthy subjects relatively easily even from short, 2–3 minutes long MEG recordings. This resting sensorimotor beta phenotype appears to be a relatively stable feature of an individual’s resting brain activity which can be easily measured also in patient populations, facilitating its use as a potential clinical biomarker.

Conflict of interest statement

None of the authors have potential conflicts of interest to be disclosed.
Acknowledgements

We thank all subjects for participating in the study. We acknowledge the following funding sources: AP received funding from the Academy of Finland (grant number 350242), the Sigrid Juselius Foundation and the Finnish Medical Foundation. Pietari Nurmi received funding from the Finnish Cultural Foundation and the Swedish Cultural Foundation in Finland. Heidi Ala-Salomäki received funding from the Jenny and Antti Wihuri Foundation and the Finnish Cultural Foundation. HR received funding from the Academy of Finland (grant numbers 127401 and 321460). Mia Ljjeström received funding from the Swedish Cultural Foundation in Finland.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinho.2024.03.021.

References


