Reliability for non-invasive somatosensory cortex localization: Implications for pre-surgical mapping

Jack Solomon a,b, Shaun Boe a,b,d, Timothy Bardouille b,c,d,*

a Laboratory for Brain Recovery and Function, Dalhousie University, 5869 University Avenue, Halifax, Nova Scotia, Canada B3H 1X7
b Department of Psychology and Neuroscience, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia, Canada B3H 4R2
c Biomedical Translational Imaging Centre (BIOtic), QEII Health Sciences Centre, 3900-1796 Summer Street, Halifax, Nova Scotia, Canada B3H 3A7
d School of Physiotherapy, Dalhousie University, 5869 University Avenue, Halifax, Nova Scotia, Canada B3H 1X7

A R T I C L E   I N F O

Article history:
Received 12 May 2015
Received in revised form 21 August 2015
Accepted 5 October 2015
Available online 9 October 2015

Keywords:
Magnetoencephalography
Source localization
Pre-surgical mapping
Intersession reliability
Somatosensory cortex
Median nerve stimulation
Neurosurgery

A B S T R A C T

Objectives: In patients with epilepsy or space occupying tumors in cortical regions, surgical resection is often considered as the primary treatment. Pre-surgical neuroimaging can provide a detailed map of pathological and functional cortex, leading to safer surgery. Mapping can be achieved non-invasively using magnetoencephalography (MEG), and is concordant with invasive findings. However, the reliability of MEG mapping between sessions is not well established. The inter-session reliability is an important property in pre-surgical mapping to establish resection margins, but repeated scans are impracticable.

The present study sought to quantify the intersession reliability of MEG localization of somatosensory cortex (S1).

Patients and methods: Eighteen healthy individuals underwent MEG sessions on 3 consecutive days. Five participants were excluded due to technical issues during one of the three days. Each session included clinical-style S1 localization using electrical stimuli to each median nerve at sub-motor thresholds. The 35 ms peak of the somatosensory evoked field was used for localizing S1 in each session using a single equivalent current dipole model. Intersession reliability was quantified using two methods. Average Euclidean Distance (AED) quantified the difference in localization between each session and the intersession mean localization. Session Euclidean Distance (SED) quantified the difference in localization between each pair of sessions.

Results and discussion: Results showed the AED was 4.8 ± 1.9 mm, whereas the SED was 8.3 ± 3.4 mm. While the AED values obtained parallel those reported previously in smaller samples, the SED values were substantially larger.

Conclusion: Clinicians should consider up to an 8 mm confidence interval around the estimated location of S1 based on MEG pre-surgical mapping.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

In patients diagnosed with epilepsy or space occupying tumors, a valid protocol to alleviate symptoms is to resect the damaged cortex or the tissue serving as the focal point for the epileptic seizures. Surgical procedures are considered for those who have disabling impairments due to the nature of the damage, those who are refractory to their medications and those who have medication side effects that alter their quality of life. In order to reduce cognitive and/or motor impairment as a direct outcome of surgery, it is imperative to cause minimal or no damage to the functional cortex surrounding the area that is to be resected. Clinicians elucidate functional cortex, in part, using functional mapping techniques. For use in surgical planning, it is essential that functional mapping is accurate (identified the correct brain area to spare) and reliable (the same result would be found if repeated). The gold standards for mapping are invasive, and include the Wada test [1] and direct cortical stimulation (DCS), both of which are used based on validated accuracy and reliability. As an example, DCS involves implanting electrodes directly onto the patient’s brain during an awake craniootomy. Stimulation via the implanted electrodes inhibits areas of
the cortex while the subject performs specific tasks. If function is affected by the stimulation, then the area is deemed an "area of eloquence", and in-turn is identified as one to spare during surgery.

Non-invasive pre-surgical functional mapping during planning can alleviate some of the challenges associated with invasive options, such as the Wada test and DCS, potentially leading to better patient outcomes. Procedures such as magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and navigated transcranial magnetic stimulation [2] are emerging as viable non-invasive alternatives for pre-surgical mapping [3]. Given the invasive nature of procedures like DCS, there are inherent surgical risks with severe health consequences [4]. The DCS procedure also requires compliance from the patient to complete the required tasks under challenging circumstances. This limits the applicability of the procedure to groups with challenges for compliance, such as young children and patients with neurological or cognitive deficits [5]. Furthermore, DCS is usually performed during the respective surgery, which limits the amount of time and individuals available for consulting about DCS findings prior to resection.

In order to impact positively on decisions related to surgical approach and resection margins, non-invasive techniques such as MEG need to demonstrate both a high level of accuracy and intersession reliability. From a clinical perspective, accuracy and reliability are important to ensure that the MEG localization correctly identifies the targeted neural tissue, and that the coordinates provided by the localization would not change if the scan were repeated. This latter feature is important for characterizing the variability associated with the localization technique, as completing multiple sessions of pre-surgical mapping to establish variability is not a viable option in a clinic setting.

The accuracy of MEG for pre-surgical cortical mapping has been established in the primary somatosensory cortex (S1) via comparison to invasive mapping, which is the gold standard mapping method, and by surgical outcome [6, 7]. Localization of S1 is a good test case for pre-surgical mapping techniques as robust patterns of activation can be determined in a single patient using MEG and DCS. However, the inter-session reliability (i.e., consistency between sessions) of localization using MEG pre-surgical mapping has not been well examined in past. Quantifying intersession reliability will provide clinicians with an estimate of how much the localization provided by a single MEG scan could differ due to the variability associated with the imaging technique itself, toward establishing critical margins of error during surgery.

Accordingly, the primary objective of this study is to provide a robust measure of the intersession reliability of S1 localization with MEG. Measures of intersession reliability have previously been established, ranging between 2.8 and 7 mm, but only for single participants [6, 8]. While both studies established baseline measurements for intersession reliability, there are two issues with this previous work. Firstly, measures of reliability were derived from a single participant, and may not be representative of the larger population. Secondly, the methodology for determining intersession reliability was not described in either report. Owing to these limitations, we have further studied the inter-session reliability of MEG localization to better assess the utility of MEG for pre-surgical mapping.

To achieve our primary objective, the current study seeks to examine intersession reliability of MEG S1 localization in a cohort of individuals using two different approaches to quantify reliability. We anticipate that localization of S1 will be consistent across sessions. Results of the study will increase our understanding of the intersession reliability, and in-turn variability, of localization using MEG, allowing clinicians to make more informed decisions about resection margins during pre-surgical planning.

2. Methods

2.1. Subjects

Eighteen healthy right-handed volunteers participated in the study (10 females; 24.7±3.8 years). All subjects were free of neurological disorder and each provided written, informed consent. Prior to the onset of the study, subjects were screened for compatibility with MEG according to institutional procedure. The Research Ethics Board at the IWK Health Centre approved the study.

2.2. Data acquisition

Neuroimaging data was collected using a 306 channel MEG system (Elekta Neuromag Oy, FL) using standard pre-surgical functional mapping procedures. The vertical and horizontal electrooculogram was also obtained using four electrodes, with one superior and one inferior to the left eye, and one just lateral to the left and right eye, with a ground electrode attached to the collar bones. Additionally, four head position indicator coils were placed on the subject’s head; two on the forehead and one on each mastoid process. The positions of the coils, anatomical landmarks, and a 200-point head model were digitized using a Polhemus digitization device (Polhemus Incorporated, Vermont, USA). During scanning, coils were activated continuously to permit tracking of head movement. All data were acquired continuously at a sampling rate of 1500Hz and a bandwidth of 0.1–500Hz, and recorded to a file for off-line analysis.

To generate robust activation in S1, percutaneous electrical stimulation was applied to the left and right median nerves at the level of the wrist (DST7A Constant Current Stimulator, Digitimer, England). Prior to MEG scanning and for each wrist, stimulator output was increased until a visible twitch of the median innervated thenar muscles was observed. Stimulator output was then reduced until the twitch was no longer visible, but the participant reported a sensory response with each stimulus. Eighty stimuli were then applied to each wrist in random sequence (to a maximum of 4 consecutive stimuli on one side) with an inter-stimulus interval varying between 1 and 2 s. Markers indicating the onset and side (left or right wrist) of each stimulus were acquired with the continuous MEG recording to facilitate event-related analysis. Participants attended three experimental sessions performed at approximately the same time on consecutive days. Data reported was collected as part of a larger study examining neurofeedback and brain activation patterns, the methods and results of which have been previously reported [9].

2.3. Data analysis

Data analysis was completed using standard pre-surgical functional mapping procedures. Temporal signal space separation and head position estimation was completed on all MEG data (Maxfilter, Elekta AB, Stockholm, SE) [10]. Datasets were excluded from further analysis if rotation or translation exceeded 3 degrees or 5 mm, respectively. Data was then down sampled to 250 Hz and a low-pass filter of 70 Hz was applied to attenuate any high-frequency noise. Lastly, an additional artifact removal was done using independent component analysis to remove artifacts, including components that were highly correlated with the electrooculogram time course [11]. The filtered data was then epoched into 600 ms sections (100 ms pre-stimulus to 500 ms post-stimulus) using the event markers for stimulation onset. The resulting 160 epochs were binned according to side (N=80 for the left and right hemisphere, respectively) and averaged. Baseline correction was applied using a pre-stimulus window of −50 to −25 ms. The somatosensory evoked field (SEF) pattern was then manually selected as the field topography at the
latency of peak magnetic field strength in the hemisphere contralateral to the side of stimulation nearest to 35 ms post-stimulus. Source localization of the SEF was achieved using dipole fitting to minimize the least-squares error between the measured and projected field for a single equivalent dipole (xfit software, Elekta Neuromag Oy, FL). Magnetic fields were projected using a single sphere head model, where the sphere was fit to the 200-point head digitization acquired prior to the MEG scan. The source location was recorded in Cartesian co-ordinates, along with the latency of the SEF and the goodness of fit to the measured field. The positive x-, y- and z-axes were directed toward the right, anterior and superior directions, respectively. This procedure was repeated for each participant and for each of the 3 sessions.

2.4. Determining intersession reliability

Intersession reliability was quantified in two ways. In both cases, Euclidean distance (ED) was used as a univariate measure of change in localization. The first method to calculate intersession reliability (“Average Euclidean Distance”) was based on the difference in localization between each session and the inter-session mean localization. The second method (“Session Euclidean Distance”) was based on the difference in localization between each pair of sessions. The mean x-, y- and z-axis values of the three localizations (i.e., one localization for each day) were calculated to create an average co-ordinate. The average coordinate was then used as a reference point to obtain the mean difference as the distance between the average coordinate and the individual co-ordinates for each of the participant’s three sessions. The mean differences were calculated along the x-, y- and z-axes, in order to obtain the average mean difference along each axis. For each method, three measures of Euclidean distance were generated per participant per hemisphere. All mean differences and Euclidean distance measures are accompanied by a standard deviation and a confidence interval (alpha = 0.05).

3. Results

Of the 18 total participants, 5 were removed from the analysis due to technical errors with one of the three consecutive MEG sessions. Three participants were removed due to issues associated with head movement in the MEG, and two participants were removed due to an inability to accurately identify the peak activation of the SEF.

The average SEF in the left hemisphere occurred at 33 ± 2 ms and was localized to 41.5 ± 3.0 mm, 6.2 ± 5.7 mm, and 85.1 ± 7.1 mm (x-axis coordinate ± SD, y-axis coordinate ± SD and z-axis coordinate ± SD) with an average goodness of fit (GOF) of 89 ± 5%. Similarly in the right hemisphere the average SEF occurred at 34 ± 2 ms and was localized to −39.0 ± 4.7 mm, 4.0 ± 6.1 mm and 85.5 ± 6.5 mm with an average GOF of 90 ± 6%.

The mean differences in source localization within the left and right hemispheres are shown in three dimensions in Figs. 1 and 2.

For visualization purposes, the intersession average location is subtracted from each participants' locations in these figures. The mean difference for 89.1% of data points is less than 5 mm along any cardinal axis. The variability of S1 localization in the right hemisphere (Fig. 1) is smaller in the x-axis (lateral/medial) than in the z-axis (superior/inferior) or y-axis (anterior/posterior), which have relatively similar distributions. In the left hemisphere (Fig. 2), the distribution of variability along the lateral/medial dimension is smaller than in the other dimensions.

In the right hemisphere, the average mean difference of localization for S1 was 2.3 ± 1.8 mm (CI = 0.6 mm), 3.4 ± 2.0 mm (CI = 0.6 mm), and 2.5 ± 1.4 mm (CI = 0.4 mm) for the x-, y- and z-axis, respectively. The maximum mean difference was 5.96 mm along the x-axis, 8.77 mm along the y-axis, and 6.57 mm along the z-axis. The average mean difference for each subject was 2.3 ± 3.3 mm (CI = 0.6 mm) in any direction. In the left hemisphere, the average mean difference of localization for S1 was 1.6 ± 1.2 mm (CI = 0.4 mm), 2.5 ± 1.9 mm (CI = 0.6 mm) and 2.2 ± 1.7 mm (CI = 0.5 mm) for the x-, y- and z-axis, respectively. The

---

**Fig. 1.** Variability in localization of primary somatosensory cortex (S1) activation during left median nerve stimulation across sessions and participants. Each point represents the localization of S1 for a single session with respect to the mean localization across three sessions in the (A) coronal, (B) sagittal, and (C) axial planes. Data for all 13 participants are shown.

**Fig. 2.** Variability in localization of primary somatosensory cortex (S1) activation during right median nerve stimulation across sessions and participants. Each point represents the localization of S1 for a single session with respect to the mean localization across three sessions in the (A) coronal, (B) sagittal, and (C) axial planes. Data for all 13 participants are shown.
maximum mean difference was 4.63 mm along the x-axis, 5.37 mm along the y-axis and 6.20 mm along the z-axis. The average mean difference was 2.1 ± 2.7 mm (CI = 0.5 mm) along any axis.

Fig. 3 shows the distribution of mean differences in S1 localization between sessions. The mean difference has a negatively skewed distribution along each axis and in both hemispheres. Furthermore, median values for the mean difference in the left hemisphere occur between 2–3 on the y-axis, and 1–2 mm along the x-axis and z-axis. This indicates that the majority of the S1 localizations are with 3 mm of the mean location. In the right hemisphere, the distribution appears slightly less negatively skewed across axes. The y- and z-axes display a larger range of mean difference values as compared to the x-axis. Furthermore, median mean difference values are between 1 and 2 mm for the x-axis, as compared to 3–4 mm for the y-axis and 2–3 mm for the z-axis. This indicates that the least inter-session variability in localization occurs in the medial–lateral direction. This is true for both hemispheres.

Fig. 4 shows the intersession reliability quantified as both Average and Session EDs. The intersession reliability is greatly reduced in either hemisphere when using the Average Euclidean Distance method. The Average ED in the left hemisphere was 4.3 ± 1.8 mm (CI = 1.0 mm), 5.3 ± 1.9 mm (CI = 1.0 mm) in the right hemisphere, and 4.8 ± 1.9 mm (CI = 0.7 mm) when both hemispheres are considered together. When using the Session ED the obtained values are significantly larger. In the right hemisphere was 9.2 ± 3.5 mm (CI = 1.1 mm), 7.5 ± 3.1 mm (CI = 1.0 mm) in the left hemisphere and 8.3 ± 3.4 mm (CI = 0.8) when both hemispheres are considered. The right hemisphere exhibited poorer intersession reliability and had higher median EDs for both methods of calculations. The Average ED had a median of 3–4 mm in the left hemisphere and 5–6 mm in the left hemisphere. The Session ED had a median in the right hemisphere between 8 and 9 mm, in comparison to the left hemisphere that had the median ED fall between 7 and 8 mm.

4. Discussion

The purpose of this study was to quantify the intersession reliability associated with localization of S1 using a pre-surgical mapping protocol in MEG for the purpose of informing the surgical decision making process. Intersession reliability across three sessions was calculated using two measures. Regardless of whether reliability was quantified using the difference from the intersession mean location or the difference between individual scans, our result showed that intersession reliability is on the high end of values reported in previous literature. While previous reports on single participants have suggested that intersession reliability may be as low as 2 mm, our study indicates a value of 4–8 mm over a larger group of individuals depending on the measure of reliability used. Given that resource limitations make repeat clinical scans infeasible, estimates of the variability that is likely to occur in localization between sessions using MEG have important implications for pre-surgical mapping. Our data suggest that this variability may be larger than previously estimated. This finding should be taken into account when interpreting MEG source localization during surgical planning.

In both hemispheres, the anterior–posterior direction revealed the poorest intersession reliability (left hemisphere: 2.5 mm, right
hemi-

The right hemisphere also demonstrated a wider distribution of inter-session reliability than the left hemi-

The larger variability in the right hemisphere contributed heavily to the poorer inter-session reliabil-

This implies that the source localization of S1 is less reliable in the right hemisphere than in the left hemisphere. Slightly poorer reliability in the right hemisphere may occur because stimu-

olution is occurring to the non-dominant side in this case. However, further work is necessary to investigate this hemispheric differ-

ence. Regardless, our results suggest that the most caution in pre-surgical planning will be necessary when using MEG localization of S1 in the right hemisphere and along the anterior–posterior axis in right-handed individuals.

The ED measure describes the straight-line distance between the localizations of S1 for a given participant. From the current data, ED calculated between each session (i.e., “Session ED”) estimated the magnitude of inter-session reliability as 8.3 ± 3.4 mm. The ED to the intersession mean location (i.e., “Average ED”) estimated reliability as 4.8 ± 1.9 mm. This reduced value does not represent any true change in the variability in the data; it simply highlights the difference between two methods to calculate the reliability of source localization data. While session ED highlights the likely distance between two localizations of the same response on different days, the average ED highlights the likely distance of a given localization from the best estimate available after three sessions. The values obtained for inter-session reliability in MEG S1 localization using the Average Euclidean Distance fit well with previous litera-

ure, while the values obtained from the Session Euclidean Distance are substantially larger [6,8]. Unfortunately, it is unclear which methods are most representative of the processes used in previous studies. Another explanation for the larger values of inter-session reliability reported in our study is that the larger number of partic-

ipants increased the variability in the experiment. This is unlikely, however; calculating ED measurements on each individual’s data eliminated inter-subject variability. One existing challenge to pre-surgical mapping in MEG is the limited ability to assess data quality at the time of the scan. In our cohort, five of eighteen participants were excluded due to data quality issues on one of the three consecutive days, which amounts to an approximately 10% failure rate. Scans were unsuccessful due to excessive head movement or an inability to accurately identify the peak activation of the SEF. Newly developed online head track-

ing algorithms [12,13] will likely improve the ability to ameliorate losses due to head movement. However, the quality of the SEF is usually assessed offline once the patient has left the MEG suite due to the time involved in data analysis. Clinical MEG users would be well served by fast, automated analysis methods that can determine if the evoked field maps are of sufficient quality for source localization within minutes of scan completion.

Inter-session reliability is an important measure to know for clinical applications, as this measure encompasses many forms of variability that are not included when multiple scans are performed in the same session. In MEG, one major source of variability and limit-

ation to MEG source localization accuracy is the co-registration of the MEG scanner to the patient’s head. This co-registration is usu-

ally achieved via the spatial localization of head position indic-

ator coils attached to anatomical landmarks on the patient’s head. Recent data has suggested that new head digitization technologies can improve MEG source localization accuracy from 5 mm to 2 mm, and improve inter-session reliability approximately by a factor of 2, as compared to the commonly used methodology described in the current study [Polhemus] [14,15]. Thus, novel digitization tech-
nologies can have a major impact on the inter-session reliability reported herein. Another potential negative impact on inter-session reliability can occur due to variability in the selection of the peak amplitude of the SEF component occurring 35 ms post-stimulus. Although the 35 ms component of the SEF was always used for localization in our study, inaccuracies in the peak latency selection can add variability in the resulting coordinates of source localiza-

tion. However, the contribution of latency selection to variability is likely minimal, as compared to the variability due to the currently used head digitization method. Importantly, the inter-session reliabil-

ity calculated in this study accounts for these types of source of variability, giving a realistic measure with meaningful implications for clinical MEG users.

When planning a surgical procedure, inter-session reliability in functional neuroimaging is an important measure for clinicians, as the localization reported is subject to variability inherent to the methodology. Knowing the extent of this variability can inform surgical decisions as to the extent of the cortex which may be implicated in a specific function. This information is critical in surgical decision making to reduce the risk of functional deficit fol-

owing resection of cortical matter. This study demonstrates that inter-session reliability of MEG S1 localizations using median nerve stimulation is slightly larger than previously reported. Specifically, based on current best practices, surgical teams should consider up to an 8 mm confidence interval around the estimated source loca-

tion when interpreting MEG S1 localization based on median nerve stimulation.

Conflict of interest

The authors declare that there is no conflict of interest.

Acknowledgements

This research was supported by funding from the Nova Scotia Health Research Foundation (Grant Number: MED-DI-1551). This sponsor did not have any involvement with the study.

References


