NeuroImage: Clinical 17 (2018) 10–15

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Automated diagnosis of temporal lobe epilepsy in the absence of interictal spikes

Thibault Verhoeven[a,⁎,1], Ana Coito[b,1], Gijs Plomp[c], Aljoscha Thomschewski[d], Francesca Pittau[g], Eugen Trinka[d], Roland Wiest[e], Karl Schaller[f], Christoph Michel[b], Margitta Seeck[g], Joni Dambre[a], Serge Vulliemoz[g], Pieter van Mierlo[a,b]

[a] Department of Electronics and Information Systems, Ghent University, Ghent, Belgium
[b] Functional Brain Mapping Lab, Department of Basic Neuroscience, University of Geneva, Geneva, Switzerland
[c] Perceptual Networks Group, Department of Psychology, University of Fribourg, Fribourg, Switzerland
[d] Department of Neurology, Paracelsus Medical University and Center for Cognitive Neuroscience, Salzburg, Austria
[e] Institute for Diagnostic and Interventional Neuroradiology, University of Bern, Bern, Switzerland
[f] Neurosurgery Clinic, University Hospital Geneva, Geneva, Switzerland
[g] Epilepsy Unit, Neurology Clinic, University Hospital of Geneva, Geneva, Switzerland

A R T I C L E   I N F O

Keywords:
Temporal lobe epilepsy
Diagnosis
Lateralization
EEG
Machine learning

A B S T R A C T

Objective: To diagnose and lateralise temporal lobe epilepsy (TLE) by building a classification system that uses directed functional connectivity patterns estimated during EEG periods without visible pathological activity.
Methods: Resting-state high-density EEG recording data from 20 left TLE patients, 20 right TLE patients and 35 healthy controls was used. Epochs without interictal spikes were selected. The cortical source activity was obtained for 82 regions of interest and whole-brain directed functional connectivity was estimated in the theta, alpha and beta frequency bands. These connectivity values were then used to build a classification system based on two two-class Random Forests classifiers: TLE vs healthy controls and left vs right TLE. Feature selection and classifier training were done in a leave-one-out procedure to compute the mean classification accuracy.
Results: The diagnosis and lateralization classifiers achieved a high accuracy (90.7% and 90.0% respectively), sensitivity (95.0% and 90.0% respectively) and specificity (85.7% and 90.0% respectively). The most important features for diagnosis were the outflows from left and right medial temporal lobe, and for lateralization the right anterior cingulate cortex. The interaction between features was important to achieve correct classification.
Significance: This is the first study to automatically diagnose and lateralise TLE based on EEG. The high accuracy achieved demonstrates the potential of directed functional connectivity estimated from EEG periods without visible pathological activity for helping in the diagnosis and lateralization of TLE.

1. Introduction

Mesial temporal lobe epilepsy (TLE) is the most common type of pharmaco-resistant epilepsy in adults. In order to estimate the localisation of the epileptogenic zone, Electroencephalography (EEG) is recorded to identify pathological activity such as seizures or Interictal Epileptiform Discharges (IEDs). However, in some patients, these are infrequent or completely absent in the EEG recording.

Epilepsy is increasingly recognized as a network disease (Laufs, 2012) and measures of functional relationships between activities of different brain regions could help better understand epileptic networks. Directed functional connectivity estimates the directionality of the functional connections between different brain regions. Several studies have shown that directed functional connectivity measures based on intracranial EEG can help to identify the irritative zone and the seizure onset zone (van Mierlo et al., 2013; Wilke et al., 2009; van Mierlo et al., 2014). Furthermore, directed functional connectivity applied to brain sources estimated from high-density scalp EEG revealed interictal network patterns concordant with cognitive deficits in TLE (Coito et al., 2015) and significant connectivity differences in TLE compared to healthy controls in the absence of interictal spikes (Coito et al., 2016).

Machine learning algorithms have been used for automatic detection and localization of the epileptogenic zone in TLE using a multitude of imaging modalities (Focke et al., 2012; Kamiya et al., 2016; Cantor-

⁎ Corresponding author at: Ghent University, Department of Electronics and Information Systems, Technologiepark 15, 9052 Gent, Belgium.
E-mail address: thibault.verhoeven@ugent.be (T. Verhoeven).
[1] These authors contributed equally to the manuscript.

http://dx.doi.org/10.1016/j.nicl.2017.09.021
Received 25 April 2017; Received in revised form 15 July 2017; Accepted 26 September 2017
Available online 28 September 2017

Rivera et al., 2015; Chiang et al., 2015; Yang et al., 2015; Kerr et al., 2013). However, to the best of our knowledge, no study has attempted to automatically diagnose and lateralise TLE using scalp EEG.

Here, we used EEG-based directed functional connectivity values to build a diagnostic and lateralization classification system for TLE in the absence of visible epileptic activity. Moreover, we compared our results with previous classification studies using other imaging modalities.

2. Materials and methods

2.1. Subjects

Twenty LTLE patients, 20 RTLE patients and 35 healthy subjects were included. Patients were retrospectively selected from the high-density EEG database of the University Hospital of Geneva, University Hospital of Bern and Paracelsus Medical University in Salzburg according to the following inclusion criteria: drug-resistant TLE, unilateral anteromedial localization of the epileptogenic zone confirmed by good surgical outcome (Engel's class I or II, after at least 12 months post-operative follow-up), intracranial EEG or concordant presurgical evaluation methods and the existence of at least a 10–15 min resting-state eyes-closed high-density EEG recording (96–256 channels). All patients had interictal activity on long-term EEG concordant with the diagnosis of unilateral TLE. Most of them had extensive presurgical evaluation including ictal video-EEG, PET, SPECT and electric source imaging. The patients' dataset used in this study was the same as reported in our previous work (Coito et al., 2016). The clinical details can be found in the Supplementary material of the present manuscript.

2.2. Standard protocol approvals, registrations, and patient consents

All patients were evaluated in the epilepsy units of Geneva University Hospital, Switzerland, Bern University Hospital, Switzerland, and Paracelsus Medical University in Salzburg, Austria. The three local ethics committees approved this study. Written informed consent was obtained from all participants in the study.

2.3. EEG, electrical source imaging and directed functional connectivity

Subjects underwent a resting-state eyes-closed recording during presurgical evaluation. The sampling frequency of the recorded EEG ranged between 250 and 1000 Hz. All signals were filtered offline between 1 and 100 Hz and then downsampled to 250 Hz. Sixty epochs of 1 s, free of artefacts and IEDs, during wakefulness were selected per subject. The activity of brain sources during the selected EEG epochs was obtained using Electrical Source Imaging (ESI): an individual head model and a linear distributed inverse solution with biophysical constraints were used (Grave de Peralta Menendez et al., 2004). The grey matter was parcelled in 82 Regions Of Interest (ROI) and the solution point closest to the centroid of each ROI was considered as representative of the source activity in this ROI. This procedure resulted in 82 time-series representing the activity of each individual ROI during the 60 selected epochs.

For each subject and epoch, directed functional connectivity between the 82 source ROIs was estimated using the weighted Partial Directed Coherence (wPDC) (Baccala & Sameshima, 2001; Astolfi et al., 2006; Plomp et al., 2014), and the mean wPDC across epochs was taken.

For each subject, we obtained a 3D connectivity matrix (82 regions × 82 regions × frequency), which represented the outflow from one region to another for each frequency. For further analysis, we reduced the connectivity matrix to 3 frequency bands: theta (4–8 Hz), alpha (8–12 Hz) and beta (12–30 Hz), by calculating the mean connectivity value in each band.

The detailed procedures for EEG recording, electrical source imaging and directed functional connectivity have been described in (Coito et al., 2016) and are also included in the Supplementary material of this manuscript.

2.4. Classification

2.4.1. Feature selection

The calculation of the connectivity between every pair of regions in the three frequency bands resulted in 20.172 features for each individual. An optimal subset of these features was selected to avoid creating false decision rules when training the classifier on the example data. As an example, consider the case where a certain connection is slightly stronger for RTLE compared to LTLE patients in the majority of our patients, but not for the whole population of TLE patients. A classifier taking this contingency as a general rule for lateralization can perform poorly on new subjects. This issue of overfitting to example data becomes more likely with decreasing number of subjects and increasing number of feature values per subject (Mwangi et al., 2014; Guyon & Elisseeff, 2003). To avoid overfitting, we allowed a maximum of one feature per ten subjects, resulting in a maximum of 7 features for diagnosis and 4 features for lateralization.

First, the 82 regions were reduced to a set of 14 regions that showed differences between groups in our previous study (Coito et al., 2016) and are known to be involved in TLE: left and right Hippocampus (Hipp), Amygdala (Amyg), Parahippocampus (PHipp), Anterior Cingulate Cortex (ACC), Posterior Cingulate Cortex (PCC), Olfactory cortex (Olf) and Medial Temporal Pole (TPMid). This left us with 588 features that were used to build the first RF classifier. The importance of each feature f in this classifier was calculated as the decrease in classification performance when the values of f are randomly permuted. As random permutation breaks the link between the feature f and the class labels, this permutation importance reflects how much classification power is lost when this feature is taken out of the design of the system. Following the feature selection method by Genuer et al. (2010), features with a non-significant importance were considered irrelevant and thus removed from the set.

Further reduction was obtained by removing redundant information. For that purpose Genuer et al. (2010) selects the minimal subset of features that contains the maximum amount of discriminant information. The method considers the interaction between features during this selection, which is important as the relevance of an outflow may depend on which other outflows were considered as features. For interpretation of the feature selection result, we calculated the actual interaction effect of a feature f_1 on another feature f_2 as the change in permutation importance of f_2 when f_1 is removed from the design (again by permuting its values). A negative interaction (a decrease in importance) indicates that the discriminative information in f_2 is only relevant when f_1 is included in the design. Higher order interactions (e.g. between three features) can also have an impact. However, with increasing order, more data is required to obtain a reasonably accurate measure of interaction. The first order interaction is given here to illustrate the impact of feature interaction in general.

2.4.2. Random Forests

Random Forest (RF) (Breiman, 2001) is a machine learning technique in which an ensemble of elementary classifiers is trained and its outputs aggregated to classify a new input sample. In RF, the ensemble is composed of many classification and regression trees (Loh, 2011), each trained on a different bootstrap subset of the available samples. When a new input is to be classified, each tree in the ensemble makes the classification and the sample is assigned to the class that was chosen by the majority of the trees.

The scikit-learn library (http://scikit-learn.org/stable/) was used to implement a Balanced RF classifier. This classifier differs from standard RF in the way that subsets containing an equal number of subjects from both classes are used to train the decision trees. Every forest contained 1000 trees. The size of the random set of features from which splits