PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([http://bmjopen.bmj.com/site/about/resources/checklist.pdf](http://bmjopen.bmj.com/site/about/resources/checklist.pdf)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Alterations in the Brain’s Connectome during recovery from severe Traumatic Brain Injury: protocol for a longitudinal prospective study</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Conde, Virginia; Andreasen, Sara; Petersen, Tue; Larsen, Karen; Madsen, Karine; Andersen, Kasper; Akopian, Irina; Madsen, Kristoffer; Hansen, Christian; Poulsen, Ingrid; Kammersgaard, Lars; Siebner, Hartwig</td>
</tr>
</tbody>
</table>

VERSION 1 - REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Damian Cruse</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Birmingham, England</td>
<td></td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

This is a potentially fascinating study that will produce multi-modal brain imaging measures from a challenging group of patients. I have two major comments:

1. The analysis methods are somewhat vague. For example, what measures will go into the graph-theory models or the Granger Causality models - raw EEG data, time-frequency decompositions, ROIs? With EEG data, how can you be sure that a graph theory output reflects cortico-subcortical connectivity? A brief description of the methods for DFA would also be helpful, rather than a reference to a separate study. The Abstract should also contain more detail of the types of data and analyses involved.

2. With regard to the classification models (page 9/10), it seems that a large number of potential features may be included with a relatively small set of observations (30 patients?), which has a danger of overfitting. This will be especially difficult to interpret when published due to the lack of specifics for the type of model and resampling etc. Can the authors provide specifics of the features, or how features will be selected, and the classification methods that will be used to control for overgeneralisation?

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Andrei Irimia</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Southern California, USA</td>
<td></td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

This is a description of an unusually comprehensive and ambitious study of severe TBI. It is likely that the data obtained from such a study will greatly benefit science. Making the data publicly available will likely be crucial for maximizing the scientific impact of this excellent and challenging work. The manuscript is generally well written and I only have a few minor suggestions:
1. Abstract: "...most relevant causes..." --> "...most pervasive causes..." sounds more appropriate. The authors may consider making this change in the introduction as well.

2. "...has offered..." --> "offers" seems more appropriate here.

3. "...will be made comprising of a Magnetic..." --> A. Please omit "of" to render the sentence grammatically correct.

4. The first two sentences in the "Ethics and dissemination" section do not seem to cover content related to either ethics or dissemination.

5. "DOC clinical states" seems to be a bit of a misnomer. "Clinical states associated with DOC" seems more proper.

6. "long-range cortico-cortical neurons" --> I wonder if the authors meant "long-range cortico-cortical neuronal connections" -- The last sentence in the paragraph uses the term "connectivity", which also seems appropriate.

7. "TE=26 MS" --> "TE = 26 ms"

8. ch = channel? The abbreviation is undefined.

9. "b value of 1000" --> This is not a dimensionless quantity. Its units are s/mm^2.

10. "18F" --> the superscript notation should be used consistently throughout.

11. It would be useful to deposit the data from this study in the FITBIR database, so that other researchers can benefit from studying such a rich and exciting dataset.

VERSION 1 – AUTHOR RESPONSE

Replies to reviewer 1 (Damian Cruse) comments

Comment

The analysis methods are somewhat vague. For example, what measures will go into the graph-theory models or the Granger Causality models - raw EEG data, time-frequency decompositions, ROIs? With EEG data, how can you be sure that a graph theory output reflects cortico-subcortical connectivity?

Reply

We thank the reviewer for pointing this out. We made some changes in the paper to make this section of the paper more clear. We changed the sentence “For the resting-state fMRI and EEG data...” to “For the resting-state fMRI data...” to reflect that we didn’t intend to use the EEG data for network analysis. Rather, our intention was to use the resting-state fMRI data (where we have simultaneous EEG) to perform network analysis. We deleted the sentence about the Granger Causality as we don’t have any specific hypothesis about this.

Comment

A brief description of the methods for DFA would also be helpful, rather than a reference to a separate study.
Reply
We added a sentence about the DFA method, thus it now reads:
“For the EEG analyses, de-trended fluctuation analysis (DFA) for the resting state will be performed in accordance to the workflow set out by [38]. DFA measures the degree of long-range temporal correlations in the amplitude envelope in different frequency bands by estimating the so-called “scaling exponent”, which is expected to increase with increasing level of consciousness due to the hypothesized restoration of cortico-cortical connectivity alongside recovery (Rosanova et al. 2012)”

Comment
The Abstract should also contain more detail of the types of data and analyses involved.

Reply
Unfortunately, the word limitations do not allow us to add a lot of detail in the abstract. However, we added this sentence to reflect the overall strategy with respect to the data analyses:
“Data analysis techniques will focus on whole-brain, both data-driven and hypothesis-driven, connectivity measures that will be specific to the imaging modality.”

Comment
With regard to the classification models (page 9/10), it seems that a large number of potential features may be included with a relatively small set of observations (30 patients?), which has a danger of overfitting. This will be especially difficult to interpret when published due to the lack of specifics for the type of model and resampling etc. Can the authors provide specifics of the features, or how features will be selected, and the classification methods that will be used to control for overgeneralisation?

Reply
The reviewer is correct that the sample size here (30 patients) potentially makes it challenging for machine learning methods. We will here rely on Gaussian processes and perform leave-one-out cross-validation to perform unbiased regression analysis. We will include a limited number of features per modality, which reflect long-range connectivity measures. We thereby added more detail to the manuscript, which now reads:
“Finally, in order to assess the advantages of obtaining data from multiple modalities in characterization and prediction of outcomes, we will form multivariate classification and regression models based on multiple modalities. This approach will be based on machine learning and resampling techniques for assessing prediction performance in estimating the clinical evaluation of DOC and recovery from DOC. We will rely on Gaussian process regression, which have as an advantage that regularization parameters to avoid over-fitting can be identified by a computationally efficient Bayesian learning procedure. This avoids the need for an additional cross-validation procedure to identify regularization parameters. Due to the relatively low number of data-samples, we will perform leave-one-out cross-validation to estimate the unbiased prediction performance. The data dimensionality will be kept low and we will include features from the different modalities, which reflect our prior hypothesis (see below) that the structural and functional long-range connectivity is predictive of level of consciousness.”

***********************
Replies to reviewer 2 (Andrei Irimia) comments

Comment
Abstract: "...most relevant causes..." --> "...most pervasive causes..." sounds more appropriate. The authors may consider making this change in the introduction as well.
Reply
We thank the reviewer for the suggestion. Changes have been implemented accordingly.

Comment
"...has offered..." --> "offers" seems more appropriate here.

Reply
Change has been implemented accordingly.

Comment
"...will be made comprising of a Magnetic..." --> A. Please omit "of" to render the sentence grammatically correct.

Reply
Change has been implemented accordingly.

Comment
The first two sentences in the "Ethics and dissemination" section do not seem to cover content related to either ethics or dissemination.

Reply
We thank the reviewer for pointing this out. We have now changed the first paragraph of “Ethics and dissemination”, both in the abstract and in the specific subsection of the manuscript.

"The project has received ethical approval by the local ethics committee of the Capital region of Denmark no. H-4-2013-186 and by the Danish Data Protection no GLO-2014-49 / I-Suite number 03389 (valid until 31/10/2020). A specialized nurse will be present for all the measurements performed on the patients, and a physician will be either present or on call for all the sessions. Any adverse events will be reported to the ethics committee as required by standard protocol. In the early stages of the study, informed consent will be obtained from both the proxy and the patient's general practitioner, a judicial procedure which is required by Danish law. For similar reasons, it is mandatory to obtain informed consent from the included patients as soon as they are capable to reliably perceive and understand information themselves about the purpose of the project, side effects, and discomfort associated with the procedures. However, none of the measurements are known to have any adverse impact on either healthy participants or patients with neurological disorders."

(Please observe that from “In the early stages of the study (...)” to “(...) patients with neurological disorders.” the text has not been changed, but simply moved.)

Changes in abstract are as follows:
“Ethics and dissemination. The project has received ethical approval by the local ethics committee of the Capital region of Denmark and by the Danish Data Protection. Results will be published as original research articles in peer-reviewed journals and disseminated in international conferences. None of the measurements will have any direct clinical impact on the patients included in the study, but may benefit future patients through a better understanding of the mechanisms underlying the recovery process after TBI.”

Comment
"DOC clinical states" seems to be a bit of a misnomer. "Clinical states associated with DOC" seems more proper.

Reply
Change has been implemented accordingly.

Comment
"long-range cortico-cortical neurons" --> I wonder if the authors meant "long-range cortico-cortical neuronal connections" -- The last sentence in the paragraph uses the term "connectivity", which also seems appropriate.

Reply
We have now changed the sentence to "According to the Global Neuronal Workspace theory of consciousness, conscious access occurs when incoming information is made globally available to multiple brain systems through long-range cortico-cortical neuronal connections between prefrontal, parieto-temporal, and cingulate cortices."

Comment
"TE=26 MS" --> "TE = 26 ms"

Reply
Change has been implemented accordingly.

Comment
ch = channel? The abbreviation is undefined.

Reply
We thank the reviewer for pointing this out. "ch" has now been changed to "channel".

Comment
"b value of 1000" --> This is not a dimensionless quantity. Its units are s/mm^2.

Reply
Changes have been implemented accordingly.

Comment
"18F" --> the superscript notation should be used consistently throughout.

Reply
Change has been implemented accordingly.

Comment
It would be useful to deposit the data from this study in the FITBIR database, so that other researchers can benefit from studying such a rich and exciting dataset.

Reply
We thank the reviewer for the suggestion. Unfortunately, the Danish Data Protection agreement does not allow to share individual data with parties not directly involved in the study and, moreover, the informed consent form for relatives and patients does not include this option either (thus, patients have not agreed to their data being shared). We however plan to make the group statistical maps available in the platform NeuroVault, and the following paragraph has been added to the "Ethics and dissemination" subsection:

"Publications will follow open access policies in order to increase the extent of the results outreach. Moreover, group data of brain activation maps will be made available at the public repository NeuroVault (http://www.neurovault.org/). Due to the data protection agreement, as well as ethical consent restrictions, individual data will not be made publicly accessible."
<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Damian Cruse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>University of Birmingham, England</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>01-May-2017</td>
</tr>
<tr>
<td>GENERAL COMMENTS</td>
<td>No further comments.</td>
</tr>
<tr>
<td>REVIEWER</td>
<td>Andrei Irimia</td>
</tr>
<tr>
<td></td>
<td>University of Southern California</td>
</tr>
<tr>
<td></td>
<td>United States of America</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>27-Apr-2017</td>
</tr>
<tr>
<td>GENERAL COMMENTS</td>
<td>Thank you for addressing my suggestions.</td>
</tr>
</tbody>
</table>
|                   | I recommend acceptance of this manuscript.