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) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

| TITLE (PROVISIONAL) | Connectomic profiling and Vagus nerve stimulation Outcomes Study (CONNECTiVOS): A prospective observational protocol to identify biomarkers of seizure response in children and youth |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------|
| AUTHORS             | Siegel, Lauren; Yan, Han; Warsi, Nebras; Wong, Simeon; Suresh, Hrishikesh; Weil, Alexander G; Ragheb, John; Wang, Shelly; Rozzelle, Curtis; Albert, Gregory W; Raskin, Jeffrey; Abel, Taylor; Hauptman, Jason; Schrader, Dewi V; Bollo, Robert; Smyth, Matthew D; Lew, Sean M; Lopresti, Melissa; Kizek, Dominic J.; Weiner, Howard L; Fallah, Aria; Widjaja, Elysa; Ibrahim, George M. |

VERSION 1 – REVIEW

| REVIEWER             | Papacostas, Savvas |
|----------------------|--------------------|
|                      | Cyprus Institute of Neurology and Genetics |
| REVIEW RETURNED      | 26-Sep-2021 |

GENERAL COMMENTS

None, the proposed study is well described

| REVIEWER             | Garcia, Camilo |
|----------------------|----------------|
|                      | Cleveland Clinic Florida, Neurology |
| REVIEW RETURNED      | 04-Oct-2021 |

GENERAL COMMENTS

Dear Editor, the study entitled "Connectomic profiling and Vagus nerve stimulation Outcomes Study (CONNECTiVOS): A prospective observational protocol to identify biomarkers of seizure response" by Sigel L. et al is a fantastic idea and long awaited study. Vagus nerve stimulation as a therapy for epilepsy has been widely used without a clear target of population in the clinical practice of patients with intractable/pharmacoresistant epilepsy with variable results. The current practice of epilepsy which is clearly contributing in neuroscience requires a better assessment of pathways and networks where VNS has a better role. This study is needed, it is the pivot for future neuromodulation therapies and the inclusion of artificial intelligence in the assessment of patients with intractable epilepsy helps in the "profiling" of better candidates to certain therapies. The only part that I did not see in the methods is the type of stimulation delivered with the VNS system, since different therapies may provide different results. Therefore, clarification of stimulation is advised for comparison among participants.

| REVIEWER             | Vespa, Simone |
|----------------------|---------------|
|                      | UCLouvain |
| REVIEW RETURNED      | 20-Dec-2021 |

GENERAL COMMENTS

This research protocol presents an ambitious yet solid and well-designed prospective clinical trial, addressing the search for predictive network-based biomarkers for Vagus Nerve Stimulation in a pediatric population. The primary endpoint is the assessment of the clinical
effectiveness in a pediatric population. This endpoint is instrumental to the classification needed to develop the predictive algorithms for seizure response, which appears to be the key outcome of impact of this study.

The leading research team has renowned expertise in the methodological aspects of data collection and analysis. The potential risks due to the multicentricity of the study are outlined and, overall, well addressed. Clinical assessment will be carried out in a standardized manner. A challenge might be posed by the diversity of neuroimaging and neurophysiological data received from the different centers.

The protocol is sound and well-built. However, I recommend that the authors provide further details on certain points of the study design which remain not entirely clear.

Main comments:

Title:
The fact that a pediatric population is addressed is an essential point of the study. I suggest adding in the title.

Recruitment (Line 192-194):
The planned sample size could be justified with more detail, in terms of power calculation. How was the reported number deemed a sufficient sample size to test the predictive model? Was this simply based or expected recruitment rates? I recommend that a power calculation based on known variables, present in background studies from the main research group, is presented in order to justify the sample size. The sample size might indeed differ between neuroimaging and electrophysiological analyses, since as mentioned later at Line 254, patients not able to undergo MRI might still be enrolled for neurophysiology. Moreover, a discrepancy is found with the bullet point at Line 69 (where the number of 500 is mentioned), while later only the 200 patient-parent pairs globally is mentioned. Please clarify the timeline of open patient enrolment.

Study design:
A potential confounder when assessing VNS effectiveness, and/or its induced changes on neurophysiological parameters over time, are the AEDs concurrently administered. The authors could explain whether guidelines for AEDs changes from baseline through follow-up are foreseen across the different centers or if physicians will be free to adapt AEDs since baseline. It should be explained whether no-AED-change windows (6 or 12 months after baseline) have been foreseen to counter this possible bias, and if not, justify the choice.

Outcomes (Line 214):
The authors mention that "all outcomes will be measured at baseline, 6 months, 12 months...". However, it is also stated that connectivity within the VagAN will be assessed only once post-operatively, through MRI, at 12 months. The authors should better clarify which outcome parameters will be evaluated post-operatively.

(Line 258);
The authors should state whether all participating centers avail the head T/R coils needed for the 12-months follow up acquisition in implanted patients, which are later reported as required at Line 273.
Informed consent to disclose incidental findings might be required. Please add what measures are taken at level of research ethics.

Data analysis:

The authors should clarify how they will operate the clinical assessment with cases presenting multiple seizure types, whether they will consider overall seizure frequency or focusing on most disabling seizure types.

For the sake of completeness, the authors are encouraged to give some brief details on the range of machine learning approaches that will be tested for the prediction algorithms. Also, the expected division of cohort in training and validation datasets, with respect to the foreseen sample size, might be further discussed.

REVIEWER Foit, Niels
McGill University

REVIEW RETURNED 11-Jan-2022

GENERAL COMMENTS

Siegel and co-workers are proposing a vigorous research plan to evaluate connectomic changes and effectiveness of VNS treatment in children. Despite its growing popularity, our understanding of the biological underpinnings of VNS therapy and potential implications is limited, thus prospective studies are clearly necessary. The particular patient population makes such investigations difficult, which the authors address. The overall concept of the study is excellent, with only a few open questions remaining:

Major Issues:

1) The authors will include children which are subjected to VNS therapy, including those with previous resective surgery. While it is clear to me that VNS therapy is reserved for highly refractory seizure conditions, the authors should clarify which inclusion criteria would trigger eligibility for VNS therapy and study participation.

2) Naturally, a cohort of VNS therapy patients will be quite heterogeneous. However, inclusion of children with previous resections are likely to exhibit heavily altered connectomes. Disentangling disease effects from surgery and VNS is a statistical challenge and I would advise against including these children.

3) The authors state that language abilities are not required. While this is appropriate for small children, neuropsychological evaluation should be standardized in modern epilepsy therapy. The authors should consider dividing the study cohort by age and should routinely collect a basic inventory of such test data. This would also clarify why such investigations are only "occasionally performed" and who might qualify.

4) The authors need to elaborate further which connectome parameters are to be examined and how connectome data will be processed. For example, fractional anisotropy is not a connectome marker - it is a non-specific diffusion imaging marker. The same holds true for "intrinsic bold signal correlations".

REVIEWER Mondal, Banashree
Institute of Neurosciences Kolkata, neurology

REVIEW RETURNED 13-Jan-2022
GENERAL COMMENTS

Very well written proposal with good clarity and challenging outcome

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Savvas Papacostas, Cyprus Institute of Neurology and Genetics
Comments to the Author:
None, the proposed study is well described.

Thank you to Dr. Papacostas for your time in reviewing our manuscript.

Reviewer: 2
Dr. Camilo Garcia, Cleveland Clinic Florida
Comments to the Author:
Dear Editor, the study entitled “Connectomic profiling and Vagus nerve stimulation Outcomes Study (CONNECTiVOS): A prospective observational protocol to identify biomarkers of seizure response” by Sigel L. et al is a fantastic idea and long awaited study. Vagus nerve stimulation as a therapy for epilepsy has been widely used without a clear target of population in the clinical practice of patients with intractable/ pharmacoresistant epilepsy with variable results. The current practice of epilepsy which is clearly contributing in neuroscience requires a better assessment of pathways and networks where VNS has a better role. This study is needed, it is the pivot for future neuromodulation therapies and the inclusion of artificial intelligence in the assessment of patients with intractable epilepsy helps in the “profiling” of better candidates to certain therapies.
The only part that I did not see in the methods is the type of stimulation delivered with the VNS system, since different therapies may provide different results. Therefore, clarification of stimulation is advised for comparison among participants.

Thank you for your comments. All children will undergo optimization of VNS, as deemed appropriate by the treating neurologist. There is no prescribed format for programming of the device, given that these institutions are highly specialized centres with individual stimulation parameters. We are however recording stimulation parameters used at each followup interval. Table 1 already lists the stimulation parameters that will be collected at each visit. However, we have included extra text in the manuscript to highlight the importance of measuring this covariate among participants (lines 237-243).

"VNS stimulation settings
Stimulation settings of the patient’s current parameters will be recorded prior to adjustment or changes.
i. Percentage of time and autostimulation function
ii. Current (mA) normal mode, current (mA) autostimulation, and current (mA) of the magnet
iii. Heart rate sensitivity and heart rate threshold
iv. system resistance”

Reviewer: 3
Mr. Simone Vespa, UCLouvain
Comments to the Author:
This research protocol presents an ambitious yet solid and well-designed prospective clinical trial, addressing the search for predictive network-based biomarkers for Vagus Nerve Stimulation in a pediatric population. The primary endpoint is the assessment of the clinical effectiveness in a pediatric population. This endpoint is instrumental to the classification needed to develop the predictive algorithms for seizure response, which appears to be the key outcome of impact of this study.

The leading research team has renowned expertise in the methodological aspects of data collection and analysis. The potential risks due to the multicentricity of the study are outlined and, overall, well addressed. Clinical assessment will be carried out in a standardized manner. A challenge might be posed by the diversity of neuroimaging and neurophysiological data received from the different centers.

Thank you for your comments and time in reviewing this manuscript.

The protocol is sound and well-built. However, I recommend that the authors provide further details on
certain points of the study design which remain not entirely clear.

Main comments:

Title:
The fact that a pediatric population is addressed is an essential point of the study. I suggest adding in the title.

Thank you for this suggestion. We have revised our title: "Connectomic profiling and Vagus nerve stimulation Outcomes Study (CONNECTiVOS): A prospective observational protocol to identify biomarkers of seizure response in children and youth"

Recruitment (Line 192-194):
The planned sample size could be justified with more detail, in terms of power calculation. How was the reported number deemed a sufficient sample size to test the predictive model? Was this simply based or expected recruitment rates? I recommend that a power calculation based on known variables, present in background studies from the main research group, is presented in order to justify the sample size.
The sample size might indeed differ between neuroimaging and electrophysiological analyses, since as mentioned later at Line 254, patients not able to undergo MRI might still be enrolled for neurophysiology.
Moreover, a discrepancy is found with the bullet point at Line 69 (where the number of 500 is mentioned), while later only the 200 patient-parent pairs globally is mentioned. Please clarify the timeline of open patient enrolment.

We appreciate the questions related to sample size. The discrepancy is explained because 500 is the maximum expected enrollment, and therefore the bullet point lists on line 70, “This will be the largest study, enrolling up to 500 patients…”
We had calculated a sample size of 52 patients based on a conservative effect size of 0.4, a desired power of 0.8 and a level of significance of 0.05. This is based on the clinical effect from several studies, including the two meta-analyses below:
1. Schevernels H, van Bochove ME, De Taeye L, Bombek K, Vonck K, Van Roost D, De Herdt V, Santens P, Raedt R, Boehler CN. The effect of vagus nerve stimulation on response inhibition. Epilepsy & Behavior. 2016 Nov 1;64:171-9.
2. Jain P, Arya R. Vagus Nerve Stimulation and Seizure Outcomes in Pediatric Refractory Epilepsy: Systematic Review and Meta-analysis. Neurology. 2021 Jun 1;96(22):1041-51.
The challenge of adequately determining sample size in this study is that there are several clinical correlates that we hope to measure to find clinical predictors of seizure outcome. Therefore, without the benefit of the final model, accounting for the need to perform hierarchical multiple regression given the multiple sites, and the high likelihood of missing data, we felt it safest (and most realistic) to aim for a sample size of 200 patients. From the biomarker perspective, this is a fairly parsimonious estimate as in neuroimaging studies, previous studies and our own pilot data have demonstrated that only 26 patients are needed differentiate between signal change between two groups, based on
1. Smitha KA, Raja KA, Arun KM, Rajesh PG, Thomas B, Kapilamoorthy TR, Kesavadas C. Resting state fMRI: A review on methods in resting state connectivity analysis and resting state networks. Journal of Neuroradiology. 2017 Aug. 30(4): 305-317.
We have added several sentences to summarize these facts on lines 197-202:
"A sample size calculation with a conservative effect size of 0.4, desired power of 0.8 and significance of 0.05 would require only 52 patients for a paired analysis of seizure frequency before and after VNS. However, given the number of institutions involved, expectation of a hierarchical mixed-effects model, and the high likelihood of missing data across multiple institutions, we believe that 200 patient-parent pairs is a parsimonious target to produce accurate and reliable analyses."

Study design:
A potential confounder when assessing VNS effectiveness, and/or its induced changes on neurophysiological parameters over time, are the AEDs concurrently administered.
The authors could explain whether guidelines for AEDs changes from baseline through follow-up are
foreseen across the different centers or if physicians will be free to adapt AEDs since baseline. It should be explained whether no-AED-change windows (6 or 12 months after baseline) have been foreseen to counter this possible bias, and if not, justify the choice.

This study hopes to understand and analyze current use of VNS in practice and will not add additional clinical parameters to patient care. For results to be generalizable and to demonstrate a clinically minimal difference for patients, the effect of VNS should show a difference in seizure frequency for patients already labelled with drug-refractory epilepsy. This is summarized in this additional sentence, on lines 224-226:

"This study will not enforce a no-drug-change window to mimic realistic clinic practice and understand the effect of VNS for patients with medically refractory epilepsy."

Outcomes
(Line 214):
The authors mention that "all outcomes will be measured at baseline, 6 months, 12 months...". However, it is also stated that connectivity within the VagAN will be assessed only once post-operatively, through MRI, at 12 months. The authors should better clarify which outcome parameters will be evaluated post-operatively.

Thank you for this clarification. We will highlight that clinical seizure outcomes and VNS stimulation settings will be measured at baseline, 6 months, 12 months, and 2 years. The MRI will be acquired before surgery and at 12 months. The original sentence has been modified on line 223:

"All clinical seizure control and VNS stimulation settings will be measured at baseline, 6 months, 12 months and 24 months post-implantation."

(Line 258);
The authors should state whether all participating centers avail the head T/R coils needed for the 12-months follow up acquisition in implanted patients, which are later reported as required at Line 273. All participating centers will have T/R coils and we have amended the manuscript to clarify this detail on line 294-295:

"Additionally, these images will be acquired at 3T with a head transmit/receive coil, which is available at all participating institutions acquiring these images."

(Line 285): Informed consent to disclose incidental findings might be required. Please add what measures are taken at level of research ethics.

This is a very important comment and we have added a statement in our Ethics and Dissemination paragraph on lines 207-208:

"Any incidental findings will be shared with a physician within the patient’s circle of care for disclosure and appropriate follow-up."

Data analysis:
(Line 323):
The authors should clarify how they will operate the clinical assessment with cases presenting multiple seizure types, whether they will consider overall seizure frequency or focusing on most disabling seizure types.

Thank you for this statement. When measuring seizure outcome, we intend to measure the severity of the main (most disabling) seizure type as well as the total seizure frequency (all types). We have clarified this in the manuscript on lines 345-346:

"For each participant, there will be frequency measurements of the main (most disabling) and the total (all types) of seizures."

(Line 359):
For the sake of completeness, the authors are encouraged to give some brief details on the range of machine learning approaches that will be tested for the prediction algorithms. Also, the expected division of cohort in training and validation datasets, with respect to the foreseen sample size, might be further discussed.

We have added some details to describe the supervised machine learning on lines 385-393:
"We will utilize supervised machine learning algorithms, during which is a labelled training dataset is used first to train the underlying algorithm, then applied to an unlabeled test dataset to categorize them into similar groups."

The details regarding sample size were addressed in the above response.

Reviewer: 4
Dr. Niels Foit, McGill University

Comments to the Author:

Siegel and co-workers are proposing a vigorous research plan to evaluate connectomic changes and effectiveness of VNS treatment in children. Despite its growing popularity, our understanding of the biological underpinnings of VNS therapy and potential implications is limited, thus prospective studies are clearly necessary. The particular patient population makes such investigations difficult, which the authors address. The overall concept of the study is excellent, with only a few open questions remaining:

We thank Prof Foit for the very thoughtful comments.

Major Issues:

1) The authors will include children which are subjected to VNS therapy, including those with previous resective surgery. While it is clear to me that VNS therapy is reserved for highly refractory seizure conditions, the authors should clarify which inclusion criteria would trigger eligibility for VNS therapy and study participation.

We wholly agree with the reviewer that these patients do present with various issues related to the effects of surgery. As such, we have amended our protocol to emphasize that this population will undergo subgroup analysis (lines 353-354):

“A subgroup analysis will be performed for children with resective surgery.”

We have not entirely excluded them because of emerging evidence that intrinsic connectivity (including subcortical connectivity) in these patients predicts response to therapy:

1. Taylor PN, Sinha N, Wang Y, Vos SB, Tisi J, Miserocchi A, McEvoy AW, Winston GP, Duncan JS. The impact of epilepsy surgery on the structural connectome and its relation to outcome. Neuroimage: Clinical. 2018: 18, 202-214.

2. He X, Doucet GE, Pustina D, Sperling MR, Sharan AD, Tracy JI. Presurgical thalamoc “hubness” predicts surgical outcome in temporal lobe epilepsy. Neurology. 2017 Jun 13; 88(24:2285-2293.

Importantly, shared circuitry within this “unique subgroup” of participants and other participants in the study, may indeed inform the development of a common biomarker of interest.

2) Naturally, a cohort of VNS therapy patients will be quite heterogeneous. However, inclusion of children with previous resections are likely to exhibit heavily altered connectomes. Disentangling disease effects from surgery and VNS is a statistical challenge and I would advise against including these children

We wholly agree with the reviewer that these patients do present with various issues related to the effects of surgery. As such, we have amended our protocol to emphasize that this population will undergo subgroup analysis (lines 353-354):

“A subgroup analysis will be performed for children with resective surgery.”

We have not entirely excluded them because of emerging evidence that intrinsic connectivity (including subcortical connectivity) in these patients predicts response to therapy:

1. Taylor PN, Sinha N, Wang Y, Vos SB, Tisi J, Miserocchi A, McEvoy AW, Winston GP, Duncan JS. The impact of epilepsy surgery on the structural connectome and its relation to outcome. Neuroimage: Clinical. 2018: 18, 202-214.

2. He X, Doucet GE, Pustina D, Sperling MR, Sharan AD, Tracy JI. Presurgical thalamoc “hubness” predicts surgical outcome in temporal lobe epilepsy. Neurology. 2017 Jun 13; 88(24:2285-2293.

Importantly, shared circuitry within this “unique subgroup” of participants and other participants in the study, may indeed inform the development of a common biomarker of interest.

3) The authors state that language abilities are not required. While this is appropriate for small children, neuropsychological evaluation should be standardized in modern epilepsy therapy. The authors should consider dividing the study cohort by age and should routinely collect a basic inventory of such test data. This would also clarify why such investigations are only "occasionally performed" and who might qualify.

The product of this study will allow us to indeed stratify on the basis of age and we do collect a robust of clinical data. Neuropsychological testing, however, is not routinely performed for patients undergoing VNS at all centres (in contrast to resective epilepsy surgery procedures, where it is universally performed). Therefore, we have included a sentence to highlight that pre-operative and >12 month neuropsychological evaluation will be collected when available on lines 226-228:

"If available, neuropsychologic test data are collected prior to surgery and a year after VNS"
implantation, depending on individual site resources and clinical indications."

4) The authors need to elaborate further which connectome parameters are to be examined and how connectome data will be processed. For example, fractional anisotropy is not a connectome marker - it is a non-specific diffusion imaging marker. The same holds true for "intrinsic bold signal correlations". We thank the reviewer for raising these critical clarifications regarding the biomarkers we wish to measure. These additions are found on lines 358-362:

"Analysis of the neuroimaging data will consist of i) measures of microstructure (fractional anisotropy), ii) inter-regional structural connectivity (on the basis of streamlines fibre-tracking), iii) measures of functional connectivity (bold correlations in fMRI, envelope amplitude correlations in MEG, bandlimited phase synchrony in MEG) and iv) evoked fields in MEG."

Reviewer: 5
Ms. Banashree Mondal, Institute of Neurosciences Kolkata
Comments to the Author:
Very well written proposal with good clarity and challenging outcome

Thank you, Ms. Mondal for your time in assessing our manuscript.

**VERSION 2 – REVIEW**

| REVIEWER          | Vespa, Simone  |
|-------------------|----------------|
| UCLouvain         |                |
| REVIEW RETURNED   | 31-Jan-2022    |

**GENERAL COMMENTS**
The authors have addressed most concerns, however a definition of the ideally suited machine learning algorithm has not been given. I agree with the authors that a supervised method is well-suited, however the statistical method need to be specified. According to their description, this is likely a support vector machine - those might however be too simple for the data complexity. I would like the authors to specifically elaborate on how and why which algorithm will be chosen.

| REVIEWER         | Foit, Niels    |
|------------------|----------------|
| McGill University|                |
| REVIEW RETURNED  | 07-Feb-2022    |

**GENERAL COMMENTS**
I am satisfied with the thorough amendments performed by the authors to the research protocol, which in my opinion is now ready to be published.

**VERSION 2 – AUTHOR RESPONSE**

Reviewer: 3
Mr. Simone Vespa, UCLouvain
Comments to the Author:
I am satisfied with the thorough amendments performed by the authors to the research protocol, which in my opinion is now ready to be published.

Thank you, Mr. Vespa for your time and contribution in reviewing our manuscript.

Reviewer: 4
Dr. Niels Foit, McGill University
Comments to the Author:
The authors have addressed most concerns, however a definition of the ideally suited machine learning algorithm has not been given. I agree with the authors that a supervised method is well-suited, however the statistical method need to specified. according to their description, this is likely a support vector machine - those might however be too simple for the data complexity. I would like the authors to specifically elaborate on how and why which algorithm will be chosen.

Thank you for your comment, Dr. Foit. At this time, we will be considering different supervised ML models to identify the options that would provide the best predictive power (logistic, random forest, decision tree, XGBoost, Neural network, SVM). We have listed these possibilities within our manuscript, on lines 391 – 393:

“Different supervised machine learning models will be tested to identify the method to provide the best predictive power; a non-comprehensive list of model options include logistic, random forest, decision tree, SVM, XGBooster, neural network.”