OPEN PEER REVIEW REPORT 2

Name of journal: Neural Regeneration Research
Manuscript NO: NRR-D-21-00335
Title: Efficacy and safety of transcutaneous auricular vagus nerve stimulation paired with conventional rehabilitation training in acute stroke patients
Reviewer's Name: Robert Andrew Morrison
Reviewer's country: USA

COMMENTS TO AUTHORS

This study examined the safety and efficacy of transcutaneous auricular vagus nerve stimulation (ta-VNS) on motor, sensory, and emotional metrics of recovery after acute stroke rehabilitation. For 20 days, acute stroke patients underwent 20 minute sessions of either active or control stimulation immediately preceding motor and sensory rehabilitation. Follow up assessments indicated active ta-VNS enhanced motor, sensory, and emotional recovery compared to control stimulation, as indicated by Fugl Meyer assessment, Wolf Motor Function Test, Stroke Impact Scale, and the Hospital Anxiety Depression Scale. These improvements persisted after one year. There were no obvious treatment-related adverse health incidents.

These findings are promising, but there are several severe limitations that are not addressed in the manuscript's current form that greatly limit its utility and interpretability.

Major Concerns

1) The authors state that the study is double blinded, and that a method of randomization was used to ensure that patients were not aware if they were in the active or control stimulation groups. The intensity of ta-VNS was "individually adjusted by the patients according to their tolerance". This is a common methodology for determining intensity, but suggests that patients would easily be able to perceive stimulation, so they would not be blinded as to which group they were in as the authors suggest. This substantially weakens the impact of this study, as all the effects attributed to ta-VNS could be due to the placebo effect from perception of stimulation, and not actually due to vagal nerve activation. Common strategies to account for this are the use of control groups who receive transcutaneous stimulation on areas of the ear that are not innervated by the auricular vagus nerve, or even surveying subjects as to what group they believed they were in, but it appears as though none of these mitigating strategies were used.

2) Page 5, line 21, the authors state that "the same (stimulation) procedure was performed on the control ta-VNS group without stimulation". Does this mean that the control group did receive stimulation in order to determine their tolerance to current intensity, and then it was shut off during their stimulation sessions, or did they not receive ta-VNS at any point? This should be clarified.

3) Was the procedure for determining current intensity of stimulation based on patient tolerance performed only on the first day of ta-VNS, or was it ongoing throughout the 20 days of ta-VNS delivery? This should be clarified in the manuscript.

4) The authors should make it more apparent that stimulation is being delivered immediately before rehabilitation sessions, and not actually "paired" during, as is commonly done in implantable VNS therapy. For example, the title of the manuscript states that VNS is "paired" with conventional

rehabilitation, and given the contrast that the authors make between ta-VNS and implantable VNS in the manuscript this seems like an inaccurate description of the experimental design.

5) Many studies have demonstrated that activation of the vagus nerve and its targets in the brain are heavily dependent on stimulation parameters used, especially current intensity. If the current intensity of stimulation was variable from patient to patient, what was the mean and standard deviation of stimulation? It is becoming increasingly common to report these stimulation metrics.

6) The results section is generally lacking specific statistical tests used for each comparison. Section 3.5 states there are significant differences in HADS scores between groups, but does not provide details to back the claim up.

7) All table and figure captions should state the statistical test used.

8) Were the statistical tests, group sizes, outcome measures, or other aspects of this study preregistered anywhere before data collection began?

Minor Concerns

1) In Figure 3, the motor portion of the FMA-L exam has a maximum score of 34, but the y axis on panel 3d extends to 40, which is an unobtainable score. The same goes for panel 3e, as the FMA-S exam has a max score of 24, but the y axis extends to 30. Errors such as this should be checked for the other metrics as well.

2) Discussion Page 8 Line 44, "The potential of VNS to treat epilepsy[15],traumatic brain injury[16],Parkinson's disease[17], stroke and other neurological disorders has been proposed."
VNS has been successfully used to treat epilepsy for over two decades, so I wouldn't characterize it as a potential or proposed treatment.

3) Discussion Page 8 Line 46, the authors state "The serious side effects of VNS therapy such as vocal cord palsy and dysphagia were reported by the Dawson et al after device implantation[13].The damage is caused by surgical operation or VN direct stimulation."
This seems like a mischaracterization of these complications. The Dawson study specifically states there were no serious adverse events stemming from treatment or implantation. One patient had left vocal cord palsy (which subsided by 9 months follow-up) and dysphagia (subsided after 3 weeks). Another patient had dysphagia (subsided in 2 weeks). Attributing these two patients' complications directly to damage via surgical operation or direct stimulation of the vagus nerve seems disingenuous, as the root causes of these problems were never specifically identified in the Dawson study. I would suggest the authors report a more realistic portrayal of the safety of implantable VNS. There are certainly possibilities for adverse effects with any surgical intervention, and the authors are right to point out that ta-VNS is an attractive alternative to implantable VNS for this reason, but this overstates the risk of adverse events stemming from surgery. For example, adverse events stemming from surgical complications only accounts for 3-6% of all patients with a VNS implant.
(https://doi.org/10.1111/ene.12629)

4) Discussion Page 8 Line 52, "Repair after vagal nerve injury is performed in approximately 50 percent of cases[18]".
This is incorrect, and massively overstates the risk of complications after implantable VNS surgery. The citation on this statement states that ~46% of patients with a VNS implant will undergo a battery replacement or surgical revision, but does not state that these surgeries are due to vagal nerve injury. In

fact, the article specifically says, "there was no obvious manifestation of vagus nerve injury associated with (device) removal" in the study.
Using the information in this cited article, of the 46% of patients needing a second surgery, the most common reason was replacement of the battery (which has a finite lifespan, and is expected to be replaced) (27%), followed by removal for poor efficacy (9%), and revision associated with lead malfunction (8%).

5) Discussion Page 10 Line 15, the authors propose that the stagnation of sensory recovery after active ta-VNS from one month onwards "could be due to a ceiling effect, or that somatosensory functions were fully recovered". It may be helpful to include that these patients' FMA-S scores were very much nearing the max score of 24, indicating full sensory function.

6) The ta-VNS device used is cited as (Bohua, Weihai, China). I am unable to find any further information about this device. A device model or similar identifying number should be included if available.

7) The study was performed in a very wide age range (18-80 years old). Did the authors observe any effect of age on efficacy?

OPEN PEER REVIEW REPORT 3

Name of journal: Neural Regeneration Research
Manuscript NO: NRR-D-21-00335
Title: Efficacy and safety of transcutaneous auricular vagus nerve stimulation paired with conventional rehabilitation training in acute stroke patients
Reviewer's Name: Jonathan M. Borkum
Reviewer's country: USA

COMMENTS TO AUTHORS

I think the authors are to be commended for having successfully carried out a logistically challenging study, combining intensive rehabilitation with (in the experimental group) transcutaneous auricular vagus nerve stimulation, and following the subjects for one year with virtually no attrition. The sample size is nicely justified by the power calculation. The finding that 20 days of stimulation just prior to rehabilitation was still having positive effects at 1-year follow-up is, needless to say, incredibly useful.

There were a few areas that I felt should be addressed in the write-up:

1. Although there were extensive, very good efforts to conceal group assignment to the participants, rehabilitation therapists, and the researchers, it sounds like the control group had the TENS-type device "without stimulation" (section 2.4, line 24), which I am assuming means that the unit was inserted but not switched on. In contrast, the experimental group was adjusting the parameters of stimulation to tolerance. This seems rather limited blinding (as opposed to, say, stimulation of a sham location, or using some other stimulus not expected to be effective. I feel this limitation of blinding should be mentioned in the "limitations" section at the end of the Discussion.

2. (Section 2.7, line 25): The various dependent measures were each analyzed with a 2-way ANOVA. Am I correct in assuming that for each measure there was a separate ANOVA at each post-baseline time point? If so, I feel this should be stated explicitly.

3. (Section 3.3, first paragraph): So, the two groups differed throughout the study on heart rate and systolic and diastolic blood pressure? Figure 2 is very helpful. I think this difference at baseline and throughout the study should be discussed a bit. I assume there is no reason to believe it could have affected outcome, but this, too, should be mentioned.

4. (Section 3.4). As you no doubt know, when conducting a statistical test to determine that two groups are comparable at baseline, the p value is often set to .10 rather than .05, because the concern is for Type II error (not seeing a difference that is there) rather than Type I error (seeing a difference that is not there). Even without this adjustment, there is a strong trend ($p = .06$) for the experimental group to have higher SIS quality of life and lower anxiety at baseline than the control group. I assume the difference is too small to have likely affected the rehabilitation outcome, but I feel the issue should be treated in the Discussion section.

5. (Section 3.5): I feel the results for the HADS should be described in a bit more detail.

6. (Third page of Discussion section, line 28): The authors contend that the improvement in depression

and anxiety, and improvement in function, confirms that depression and anxiety have a negative effect on treatment outcome. However, the reverse causality is also possible - depression and anxiety might have improved because functioning was getting better. I feel the Discussion should mention this possibility also.

7. Forgive me for not understanding but, Section 2.1, line 9: What is a "pragmatic" trial? Do you mean parallel groups?

8. Conflict of interest: To be explicit, then, you have no connection with the manufacturer of the unit, yes?