

Individualized non-invasive deep brain stimulation of the basal ganglia using transcranial ultrasound stimulation

Corresponding Author: Dr Ghazaleh Darmani

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this paper, Darmani and colleagues investigate transcranial ultrasound stimulation (TUS) targeting the basal ganglia as a potential non-invasive deep brain stimulation (DBS) modality. The paper is comprised of two experiments. The first investigates the impact of 2 TUS protocols (theta band, tbTUS, and 10hzTUS) targeting the internal globus pallidus (GPi) on local field potential recordings (LFPs) in 10 participants with implanted DBS in the GPi. The second assesses the behavioral impact of tbTUS-GPi in 15 healthy participants during stop signal task.

Experiment 1 reports a significant increase in total band power (3-30hz) in GPi LFP recordings for both tbTUS and 10hzTUS, compared to both passive and active controls. Analysis of specific frequency changes in 9/10 patients showed an increase in beta-band frequency compared to controls, which was more significant in the 10hzTUS group, and an increase in alpha-band frequency in the tbTUS group. There was additionally a strong correlation with total band power increase and levodopa equivalent daily dose (LEDD)

Experiment 2 reports a significant increase in stop-signal reaction time (SSRT) following tbTUS targeting the GPi, but not the pulvinar. There was also no reported difference in go-trial reaction time with either protocol.

It was a pleasure to review this manuscript, which advances our understanding of TUS as both a potential investigative and therapeutic tool. I have several comments regarding the results presented here however.

Regarding Experiment 1:

1. Whilst participant numbers are small, this is in line with similar work. The participants do however have a marked gender bias, which should be discussed as a source of potential bias.
2. The treatment of hemispherical readings for the same participant as independent, however, may inflate the effect of any changes seen, and should be explicitly discussed.
3. There is some ambiguity in LFP baseline data collection; in figure 1 all LFP baseline data appears to be collected prior to sham trials, while in figure 3 and the discussion, separate LFP baseline data is discussed for each trial. This requires clarification, in particular the time elapsed following sham trial baselines for TUS trials were collected at.
4. The design of experiment 1 potentially significantly impacts the presented results and limits the conclusions that can be drawn. Due to all TUS trials following sham trials, wear-off of medication presents a major confounder, which while discussed, the impact of which is not adequately addressed.
With regards to the total 3-30hz band power increases reported in figure 3, the pooling of all time points may hide the impact of medication wear-off. Were this increase to become more pronounced between t0 and t40, this would support a contribution from beta band increase secondary to medication wear-off.
Indeed, in Figure 5, beta frequency power does appear to have a trend to increase between T0 and T40 in tbTUS vs active and passive sham. That total power change correlating with LEDD also potentially supports a contribution of medication wear-off, and in 10hzTUS-Active Sham, this correlation appears to increase with time.
5. The rationale for exclusion of the dystonia patient from the analysis of frequency bands power change shown in figure 5 requires specific discussion. While figure S2 reports percentage total power increases compared to sham, how this changes

by frequency band and over time is not reported. Where these readings to show a similar pattern of change to other participants, this would strongly support these results being TUS driven.

6. Demonstrating distinct patterns of LFP change with theta and 10hz TUS would be both novel and significant, however the strength of this conclusion is at present limited by the above.

In summary, the work presented here is both interesting and novel. Without revision, however, the results in experiment 1 as presented do not support the stated conclusion to a sufficient degree. Experiment 2, whilst more robust, in isolation unfortunately does not meet the threshold for publication in a high impact journal such as Nature Communications.

Reviewer #2

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)

This manuscript is very interesting and highlights a novel approach to the use of local field potential signals (LFPs) from deep brain stimulating (DBS) electrodes during ultrasound brain stimulation. The results highlight that different ultrasound protocols (ie, varying patterns of applied sonication from theta burst to simple 10 Hz stimulation) may result in power changes in globus pallidus (GP) LFPs in Parkinson patients. The ultrasound stimulation is modeled from individual MRI scans to highlight the focus and range for a spherically focused ultrasound transducer, to show a moderately small spot size (ie, 4.7 mm wide X 27 mm long) at 6 cm within the brain, using a temporal approach to the GP. Similar to transcranial magnetic stimulation (TMS) the location and focus of the ultrasound transducer is controlled by frameless stereotaxy to allow a 2-3 mm reproducibility of localization. A second set of experiments using the same approach to GP highlights a mild disruption of a classic stop signal task in healthy volunteers.

Though the overall quality of the study is high there are multiple clarifications needed which would be helpful to the design and interpretation of the study.

1) DBS details – the DBS electrode in the PD patients is the 4 contact 3387 (or the directional B33015) with 1.5 mm contact and 1.5 mm spacing (as is commonly used in GP DBS). If so, the 4.7 mm width of the 50% max suggests that at least 2 contacts were actively targeted (vs the 1 suggested) with the stimulation, and more considering the aspect outside the 50% max range. Which 2 contacts were most involved with the ultrasound target (Fig. 1 with the DBS reconstruction is a bit misleading)?

2) LFPs – the Percept allows 2 channels out from a “sandwich” configuration – multiple possible contact pairs are mentioned but which was specifically used for each side (ie, 0-3 vs 0-2 bracketing contact 1, etc) for the actual LFPs? It would have been possible to compare (on the single, targeted side) on-target vs off-target pairs for comparison if the focus was on contacts 0 + 1, for example but this did not appear to have been done. Though it is mentioned that analysis was separated for the two sides it is not clear in Fig. 3 if the results (ie, Fig. 3B) are for the stimulated vs unstimulated side? This is confusing since a comparison between the two sides would be the most helpful in addition to the comparison between stimulation paradigms to highlight a focused effect unilaterally. Showing actual spectra in Fig. 3 (ie, similar to the baseline in Fig. 3B for the unilateral, focused stimulation condition) would be very helpful and to discuss frequency bands first rather than overall power. Fig. 4 is less clear and not as helpful.

3) FFT analysis – how was the windowing into 2 sec bins decided? Since there was a longer time of recording 8-10 sec bins may show much better smoothing and more definitive results and better clarification of frequency bands – where the shorter and longer bins compared?

4) Ultrasound modeling – Fig. 2 shows the modeled focusing on multiple structures including the GP, internal capsule and thalamus, almost to the midline (ie, 27 mm). it would be helpful if Fig. 2I included these other structures within the confines of the modeled region of stimulation.

5) On med condition – though the DBS was off for the LFP recordings the patients continued on their (presumed) ordinary med schedule with similar UPDRS, so the LEDD (which is highly variable between patients for optimal med on condition) is not helpful. A more helpful comparison would have been the effect of the ultrasound on med off/stim off vs med on/stim on, to assess if the ultrasound provided any form of treatment (or enhancement) effect on the inherent PD off condition (ie, similar to DBS or meds). This would help considerably show the difference between the theta burst and 10 Hz patterns much more convincingly than as shown in the med on condition and whether the apparent functional neuromodulation might extend to a treatment effect at this dosing.

6) Ultrasound dosing – the resulting W/cm² is very similar to the putaminal levels noted in the Meng et al 2022 (Movement Dis 37: 2133) focused ultrasound BBB study (~ 2 W/cm² across 4 patients) where the localization was confirmed directly using microbubbles, but using the standard HiFu system. Having direct MRI confirmation of the modeled ultrasound localization using microbubbles and BBB opening with contrast verification on MRI would be very helpful in 1-2 patients. Likewise, though the max output from the spherical ultrasound device was used it would be helpful to develop a dose response curve for the putative excitatory (theta burst) and inhibitory (10 Hz) parameters, one of the largely missing but critical aspects of any proposed neuromodulation.

7) The second population of control patients with possible interference on a behavioral task is much less strong of an experiment compared to the actual LFP recordings and adds much less to the overall conclusions. Comparing the theta burst to 10 Hz for differences would have been a helpful differentiation of proposed mechanisms and the results are unclear

as to how the circuits might be affected by parameters.

8) Overall, the study could be significantly improved by comparing presumed on target vs off target effects (on both a single electrode on the stimulated side and comparing the two sides), effects on inherent beta power in the PD off condition vs med on and whether there may be a more apparent neuromodulatory effect on the PD condition. The focus is reasonable but needs to be clarified with respect to actual contacts and the results much better presented. There may considerable off target effects on internal capsule and thalamus vs GP. The actual frequency on meds is presumably mainly alpha and not clear if modulated by eye opening or not and not clear how GP might affect the overall alpha frequency and rhythm with either stimulation paradigm. Though carefully controlled in terms of power and modeling to be safe, many alternative conditions may be much more revealing in this fortuitous population of PD patients with Percept LFP recordings for actual mechanisms and effects in humans.

Reviewer #4

(Remarks to the Author)

Authors Darmani et al presented an intriguing and important approach to understand the effect of transcranial ultrasound stimulation on the basal ganglia. While this is an important question to understand, the following questions arose about the methodology and interpretation of the results presented by the authors:

1. Patient selection: Barring the low sample size, there isn't sufficient clarification of why one patient with dystonia was included in the cohort to study the effects of TUS along with 9 other PD patients. A further point of classification is pertinent in figure 3, where plots C to F: has data across 18 hemispheres, $n = 9$ PD patients – here, the absence 10th data point, i.e., the absence of the dystonia patient is not justified as per knowledge. Clarification on this point would be appreciated.

Statements such as “A similar pattern was also observed for dystonia patients with GPi-DBS” should be toned down drastically given that it is based on only one data point (if at all, since this data point was absent in a plot). It would also be interesting to categorize the patients into a symptom subtype and understand the influence of TUS on each subpopulation, given that one sub population may show different effects to others.

2. Evidence for target engagement: The authors have provided two experimental questions: to employ local field potential recordings in patients using percept Medtronic Percept – this is performed on patients with movement disorders. However, to assess the target engagement, a second experiment using healthy controls is employed. It is not clear if this setup exactly justifies the target engagement since the two tests are performed on different set of subjects.

3. Experimental design and lack of sufficient evidence: While each experiment is setup well, with the second experiment having good design of contrasting GPi sonication with pulvinar nuclei sonication, the statistics seen in Figure 9 are not sufficiently convincing to show that response inhibition changes with stimulation. The difference between GPi pre- and post1 and Pulvinar pre- and post1 stop signal reaction should be compared, in addition to comparing pre- and post. Even a one tailed t-test could provide insight into if this is significant. Additionally, it would be good to show the mean as a highlighted point, rather than median since the t-test also compares the means. This might even show lower difference between the groups.

4. Additional smaller points:

a. To understand the effect of TUS stimulation as a function of LEDD, it would be important to showcase the results of active – passive sham with LEDD.

b. While two protocols of stimulation have been considered here – i.e., tbTUS vs 10 Hz stimulation & their correlation with patient medication profile has been provided to some extent, the effects of each stimulation in relation to clinical improvement is not entirely elucidated clearly. In fact, either of the stimulation protocols have no significant effect on the clinical profile of the patient. This point should be discussed more in detail, and the relation between clinical improvement and target engagement (if this is a positive correlation) should also be specified more clearly.

Reviewer #5

(Remarks to the Author)

This is a really interesting and exciting paper which presents significant results demonstrating changes in signals directly recorded from DBS electrodes following TUS in PD patients, a significant demonstration of target engagement. The paper is overall clearly written and the results well presented. However, it would benefit from a few extra details, clearer explanation and discussion of some concepts and parts of the methods, and some aspects of the results – these minor comments are detailed below.

Page 4: In the description of experiments: ‘The first session obtained MRI images for personalized ultrasound modelling and simulations.’

Suggest wording as e.g. ‘personalized ultrasound sonication planning simulations’. Using both modelling and simulations seems redundant.

‘The subsequent sessions, spaced at least one week apart, were allocated to electrophysiology recordings’

Would be clearer to mention TUS here (and when this description is repeated elsewhere): e.g. ‘... were allocated to TUS and electrophysiology recordings’

Page 5: We used the same ultrasound dose (spatial peak pulse average intensity (ISPPA) × sonication duration (SD) × duty cycle (DC)) in both protocols.

Dose concepts for ultrasound neuromodulation are not yet well understood, what you are describing is really an exposure – an amount of energy per unit area which is incident on the target, it does not account for any specific sensitivity of the tissue or any measure of how much of the energy is absorbed or the resulting mechanical forces for example. Please use the word exposure rather than dose here and elsewhere in the text.

In addition – this was not the same across participants! The in situ Isppa varied between as you show in the figures, the exposure for each person was different. Can you comment on this and why you didn't use the simulations to adjust the exposure?

The pulse timing parameters are a little scattered through the text, it would be useful if they could be collected together here for example (pulse duration added to the figures already given), and reported together with the other parameters in a table in supplementary material.

Figure 1B – it might be more informative to include the spectrum of the signal here.

Page 6: Over what time period did you average to calculate the Ispta?

'However, ISPPA estimates within the target area were similar for the two protocols, ranging from 2.5 to 7 W/cm² (Figure 2G).' – aren't they the same? The set intensity was the same, all that changed was the timing parameters? Or are taking into account the actual position of the transducer in each session for post hoc analysis?

'... guidelines set by the FDA⁴² as well as recommended by the International Transcranial Ultrasonic Stimulation Safety Standards Consortium (Maximal temperature rise anywhere ≤ 2 °C; ISPPA ≤ 190 W/cm²; MI ≤ 1.9)⁴³'

This is slightly misleading – the Isppa limit comes from the FDA guidelines, it is mentioned in the ITRUSST paper but not recommended as a limit.

Page 7: Figure 2 – a single sided colour scale would be easier to see.

Here and elsewhere, e.g. Fig 8. - I don't understand what is meant by the transducer adjustment and adjusted and unadjusted distance to target? – what is it in reference to?

Page 10: Since the intensity varied between participants, was there any correlation between the exposure and the change in recorded signal power?

'A similar pattern was also observed for the dystonia patients with the GPi-DBS lead' – there was only one dystonia patient.

Page 13: 'no correlation was observed between power changes with tbTUS or 10 Hz TUS and the Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) scores' – can you explain what this means and what the reason for looking at these was? Does a higher score mean more severe/advanced disease? Were these used to look for a correlation between response and disease severity? It is not clearly explained either here or in the methods.

Page 16: Figure 8: '(G) Pressure maps derived from ultrasound modelling superimposed on the MRI image of one exemplary subject, highlighting successful targeting of the GPi and pulvinar'
Do you mean that the focal region overlapped with the target region?

Page 17: Discussion – how does the change in LFP after TUS compare with the changes seen due to DBS?

Page 18: We observed that increases in GPi LFP power were stronger with 10 Hz stimulation – greater would be a better description than stronger – or be quantitative – how much larger were the increases?

'... has been demonstrated in animal models¹¹, as well as in studies in healthy subjects' -> healthy human participants.

'We characterized the time course of LFP bands.' – wording – examined the signals at different time points?

This paragraph and other parts of the discussion has many adjectives, please replace them with numbers where possible. 'While our methodology was carefully planned and aligned during the session' – wording – be specific – the transducer was carefully aligned?

Page 19: 'future improvements could be achieved by exploring alternative methods to better stabilize the devices during experiments' – how was the transducer positioned and held? Did you record the actual positions of the transducer during sonication? How close did you get to placing the transducer in the planned position when delivering the sonication?

'that is influenced by the dopaminergic state or disease severity' – the results suggest there is no correlation with disease severity.

STN needs to be defined earlier

'The enhancement of beta band activity, potentially associated with increased motor symptoms in PD, was evident with 10 Hz stimulation' – does this suggest that neuromodulation with this protocol is not suitable for suppressing symptoms? Were

changes in symptoms/behaviour after TUS assessed or observed?

Page 20: '(1) There was no significant difference between baseline powers for the tbTUS and 10Hz TUS compared to the two sham conditions, and the sham condition was always tested before the active TUS conditions' – the sham condition always being tested first would support the suggestion that effects may be related to wearing off of levodopa effects not argue against it.

How often is levodopa usually taken? How long is the study period compared to the usual time between doses?

Page 23/24: 'Both protocols were administered bilaterally (left side first)' – how was this done? How long did it take to reposition the transducer on the other side? What was the time difference?

'To minimize potential auditory confound during TUS67, white noise was played to participants during sonications through earbuds.' – did the participants report that they could hear it or distinguish between active sham, passive sham and sonication?

Page 26: 'Then, a thin layer of ultrasound gel (Wavelength MP Blue, Sabel Med, Oldsmar, FL) was applied to both the participant's head and the transducer' – was the hair not saturated with gel? How did you ensure this resulted in good coupling?

'These coordinates were then fine-tuned through direct anatomical targeting' – what does this mean? Please explain more clearly – you made sure that the coordinates matched what you could see on the participants MRI?

Page 27: 1st paragraph – please report the acoustic and thermal properties used in the simulations and any other settings or parameters used to run simulations – the paper referenced describes the modelling software but not the specific values you used. These could be detailed in supplementary material – see the ITRUSST standardized reporting paper for a guide to what to report when using simulations to make in situ exposure estimates. Please also report the in situ focal dimensions (it would also be useful to know the size of the target brain regions).

Reviewer #6

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

My queries have been answered and I am happy the manuscript is now suitable for publication.

Reviewer #3

(Remarks to the Author)

The comments provided by the authors to respond to the multiple reviewers are very helpful in clarifying the study design and initial, highly preliminary results that the study provides. The authors clearly indicate that this pilot study was significantly abridged by a number of external conditions and overall weak study design, and try to rationalize the existing study in many ways (ie, rather than implement a more robust study design and perform additional experiments), reinforcing the tentative nature of the conclusions:

- 1) the lack of patients off meds + off DBS and the varying duration of on med condition, significantly reducing the apparent beta LFP power and thus preventing a more clear treatment effect;
- 2) performing bilateral GP stimulation vs unilateral and hence having a clear control side that would be less effected (the argument comparing GP homotopic interconnections vs M1 is weak);
- 3) even though this is "focused" the beam and frameless navigation suggest (from the LFP results) that all contacts on the DBS electrode were likely targeted equally so no local control (ie, comparing contacts on the same electrode);
- 4) the effective exposure applied is equal to that commonly used for sonication for BBB opening, and this effect could be an equivalent response - ie, merely BBB opening with local brain irritation and not necessarily a "physiological" neuromodulation, particularly due to the long duration of effects suggested;
- 5) presenting two separate, disparate studies together.

However, the authors have responded well to the multiple reviewers' comments, clearly improving their explanation of the overall weak study design and results and the manuscript. The worry is that these highly preliminary, pilot data (from an excellent group) may be interpreted as stronger data than is actually provided supporting neuromodulation rather than an unknown effect, due to the severe constraints of the study design and results.

Reviewer #4

(Remarks to the Author)

The authors have attempted to address my concerns, and the detailed analysis and responses to my questions are much appreciated. However, some issues continue to remain. The symptom specific analysis is appreciated, however, there is one concern with this analysis: the graph in Figure S5 shows that the significance is for the 10 Hz TUS condition compared to the passive sham condition (not the active sham) – a clarification on the misinterpretation or mislabeling of the graph would be important to draw correct conclusions.

Could the authors also specify whether the significance is based on a two-sample t-test (or in other words an independent sample t-test) or a one sample t-test? This would be important for interpretation. Given the low sample size, it would be very beneficial to see individual data points rather than bar plots to understand variance contributed by each group.

Finally, while it may be correct that the second experiment does not serve as a control experiment, the fact that the second aspect of the target engagement is being performed on a group of healthy population would result in interpretations that cannot be extrapolated to a population of individuals with motor disorders. The underlying point I am trying to raise here is that the experimental design has not been sufficiently justified to draw a common conclusion. This could be added into the limitation section.

Reviewer #5

(Remarks to the Author)

Thank you for addressing my comments, I have nothing further to raise.

Reviewer #6

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this paper, Darmani and colleagues investigate transcranial ultrasound stimulation (TUS) targeting the basal ganglia as a potential non-invasive deep brain stimulation (DBS) modality. The paper is comprised of two experiments. The first investigates the impact of 2 TUS protocols (theta band, tbTUS, and 10hzTUS) targeting the internal globus pallidus (GPi) on local field potential recordings (LFPs) in 10 participants with implanted DBS in the GPi. The second assesses the behavioral impact of tbTUS-GPi in 15 healthy participants during stop signal task.

Experiment 1 reports a significant increase in total band power (3-30hz) in GPi LFP recordings for both tbTUS and 10hzTUS, compared to both passive and active controls. Analysis of specific frequency changes in 9/10 patients showed an increase in beta-band frequency compared to controls, which was more significant in the 10hzTUS group, and an increase in alpha-band frequency in the tbTUS group. There was additionally a strong correlation with total band power increase and levodopa equivalent daily dose (LEDD)

Experiment 2 reports a significant increase in stop-signal reaction time (SSRT) following tbTUS targeting the GPi, but not the pulvinar. There was also no reported difference in go-trial reaction time with either protocol.

It was a pleasure to review this manuscript, which advances our understanding of TUS as both a potential investigative and therapeutic tool. I have several comments regarding the results presented here however.

We sincerely appreciate your positive feedback and the time and effort you dedicated to reviewing our work.

Regarding Experiment 1:

1. Whilst participant numbers are small, this is in line with similar work. The participants do however have a marked gender bias, which should be discussed as a source of potential bias.

We acknowledge the sex imbalance in our DBS study, which included 9 males and 1 female. In contrast, our behavioral study on healthy participants demonstrated a reversed distribution, with a greater representation of females (11 females and 4 males). Severe PD is more prevalent in men than in women. Notably, a documented disparity exists in the proportion of women undergoing DBS surgery, with women accounting for only 30% of all implantations¹. In our study, we ensured inclusivity by reaching out to all eligible patients at our center, without bias based on sex, and recruited all individuals willing to participate,

regardless of their sex. While current research hasn't shown/investigated sex differences in ultrasound stimulation effects, we have noted this limitation in the revised manuscript (page 23) as follows:

"A known sex disparity exists in DBS surgeries for PD, with women accounting for only 30% of all procedures¹. This disparity was also evident in the current study, manifesting as a sex bias. While currently there is no evidence of differential effects of TUS based on sex, future studies might aim for a more balanced sex distribution."

2. The treatment of hemispherical readings for the same participant as independent, however, may inflate the effect of any changes seen, and should be explicitly discussed.

We agree with the reviewer that there is a certain level of dependency between the two GPis, which may indeed vary depending on the level of analysis. For instance, while beta-band oscillations may occur at the same frequencies bilaterally in the basal ganglia, the phase relationship between the two sides might be different. The idea that the two hemispheres can be considered not fully dependent is further supported by the asymmetric motor symptoms observed in Parkinson's disease (PD) patients, related to asymmetric degeneration of dopaminergic neurons in the substantia nigra pars compacta².

Additionally, treating the left and right GPi or STN as independent structures is a common approach in electrophysiological studies of the basal ganglia³⁻⁶. While we acknowledge that this approach may carry a risk of inflating effect sizes, we believe that the assumption of independence is reasonable within the scope of this study. Furthermore, our findings remain consistent with prior research, lending support to the validity of this approach.

We added the following paragraph to the manuscript (page 26) to acknowledge this issue:

"... the data recorded from each hemisphere was considered independently for the analysis. This approach is consistent with common practices in electrophysiological studies of the basal ganglia. However, we note that treating the hemispheres as independent may carry a risk of inflating effect sizes."

3. There is some ambiguity in LFP baseline data collection; in figure 1 all LFP baseline data appears to be collected prior to sham trials, while in figure 3 and the discussion, separate LFP baseline data is discussed for each trial. This requires clarification, in particular the time elapsed following sham trial baselines for TUS trials were collected at.

We recorded LFPs prior to all conditions (pre-sham and pre-TUS). To address this, we have updated the terminology throughout the manuscript to clearly distinguish between "pre-TUS LFP recordings" and "pre-sham LFP recordings." Additionally, we have revised Figures 1 to reflect these changes for consistency and clarity.

4. The design of experiment 1 potentially significantly impacts the presented results and limits the conclusions that can be drawn. Due to all TUS trials following sham trials, wear-off of medication presents a major confounder, which while discussed, the impact of which is not adequately addressed.

Due to our first Research Ethics Board (REB) approval allowing only 2 electrophysiology recordings per patient, our design was structured to test the 2 most replicable offline TUS protocols, including sham conditions to be able to compare their effects. Previous imaging and electrophysiological studies in both animals and humans suggest that the effects of offline TUS protocols can persist up to 1-2 hours post-sonication. That is why, we always began with sham conditions to ensure that pre-TUS LFP recordings, taken 40 minutes post-sham sonication, were neither affected nor contaminated by real TUS effects due to the exploratory nature of the study.

To minimize the potential effects of drug wear-off, we kept experiments brief (maximum 2.5 hours) and scheduled them to begin immediately after the patients' medication intake. Since our group of patients took medication every 4–6 hours, this provided a sufficient buffer to mitigate drug wear-off effects. While we understand the reviewers' concerns, the significant differences observed in LFP modulation across frequency bands and time points following tbTUS and 10 Hz TUS sonications (**Figure 5**) strongly suggest that these modulations are not solely attributable to drug effects. Moreover, we observed a similar pattern in the dystonia patient, who was not on dopaminergic medication (**Figure S3**), further supporting the conclusion that these modulations are not exclusively driven by drug effects.

With regards to the total 3-30 Hz band power increases reported in figure 3, the pooling of all time points may hide the impact of medication wear-off. Were this increase to become more pronounced between t0 and t40, this would support a contribution from beta band increase secondary to medication wear-off. Indeed, in Figure 5, beta frequency power does appear to have a trend to increase between T0 and T40 in tbTUS vs active and passive sham.

Our goal was to provide a comprehensive and holistic view of the dataset. While Figure 3 presents pooled data across all time points, Figure 5 breaks down the data to highlight specific time points and frequency band modulations. If the observed LFP modulations were due to medication wear-off, we would expect a consistent increase or steady modulation of LFP power over time. However, we found no significant difference in beta-band power changes between T0 and T40 in any of the four conditions (Wilcoxon signed-rank test, $p > 0.05$). Additionally, as mentioned in the text, there was no significant difference in pre-TUS and pre-sham power levels across the four conditions, despite tbTUS and 10 Hz TUS being applied later in time compared to the passive and active sham conditions. This further supports the

conclusion that medication dynamics alone cannot fully explain the observed patterns. Nevertheless, future studies could control for the influence of medication wear-off to better isolate stimulation-specific effects.

That total power change correlating with LEDD also potentially supports a contribution of medication wear-off, and in 10HzTUS-Active Sham, this correlation appears to increase with time.

The correlation between percent power change from baseline and LEDD (**Table S1**) shows distinct patterns across conditions. In the '10 Hz TUS – Active Sham' condition, the correlation increases over time (e.g., $r=0.35$ at T0 to $r=0.63$ at T40), potentially reflecting a growing contribution of medication wearing-off. This trend contrasts with the 'tbTUS' conditions, where correlations decrease over time, and the '10 Hz TUS – Passive Sham' condition, where correlations remain relatively stable. Importantly, these inconsistencies, together with the fact that tbTUS and 10 Hz TUS were applied at similar time points with respect to the start of each session, and the differences in the effects observed (e.g., tbTUS showing an increase in theta power during stimulation while 10 Hz TUS exhibits more sustained effects), lead us to conclude that medication wear-off alone is unlikely to be solely responsible for the observed effects in this study (see also our response to your previous comment above). Nevertheless, future studies should control for the influence of medication wear-off using pharmacokinetic measurements to better isolate stimulation-specific effects.

We should also mention that the correlation between tbTUS power changes and LEDD is consistent with findings from other studies. For instance, Grippe et al. demonstrated that tbTUS significantly increased motor evoked potentials (MEPs) amplitudes compared to baseline in both controls and PD patients on medication, but not in PD off medication⁷. Additionally, the bradykinesia subscore of the motor Unified Parkinson Disease Rating Scale (UPDRS) improved in the on-medication state but showed no improvement in the off-medication state⁷. Similarly, Samuel et al. found no UPDRS improvement with tbTUS in PD patients off medication⁸. These findings may suggest that TUS-induced plasticity is dopamine-dependent. This notion is further supported by pharmacological evidence and other existing literature emphasizing the critical role of dopamine in modulating LTP plasticity, a mechanisms potentially behind long lasting effects of TUS⁹.

5. The rationale for exclusion of the dystonia patient from the analysis of frequency bands power change shown in figure 5 requires specific discussion. While figure S2 reports percentage total power increases compared to sham, how this changes by frequency band and over time is not reported. Where these readings to show a similar pattern of change to other participants, this would strongly support these results being TUS driven.

During the initial study design, we planned to recruit two distinct cohorts: PD patients and dystonia patients, both with DBS leads implanted in the GPi, to investigate the effects of TUS based on disease type and characteristics. However, within the study's timeframe, we were only able to recruit one dystonia patient. Given the fundamental differences between these two conditions and to maintain coherence and focus in the main paper, we decided to report the dystonia data in the supplementary information. Here we added the analyses for the dystonia patient the reviewer asked for. Note that **Figure S2** has been renumbered as **Figure S3** in the revised manuscript to reflect the updated order. We also added a new figure (**Figure S4**) to the supplementary information.

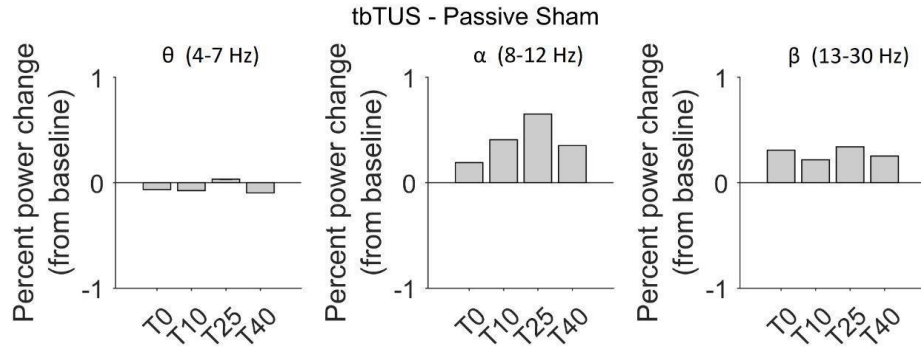
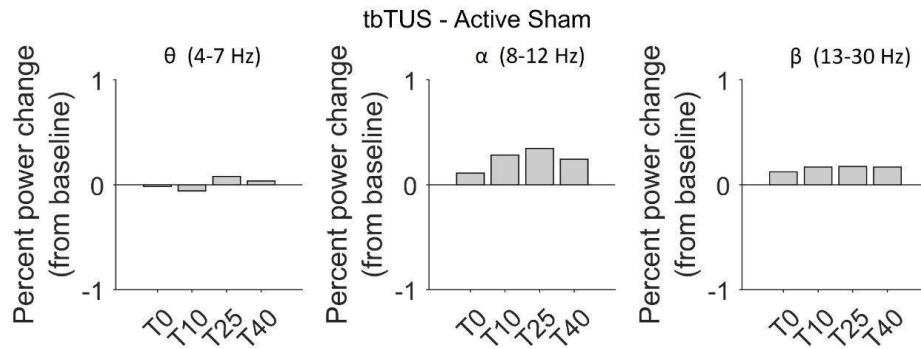
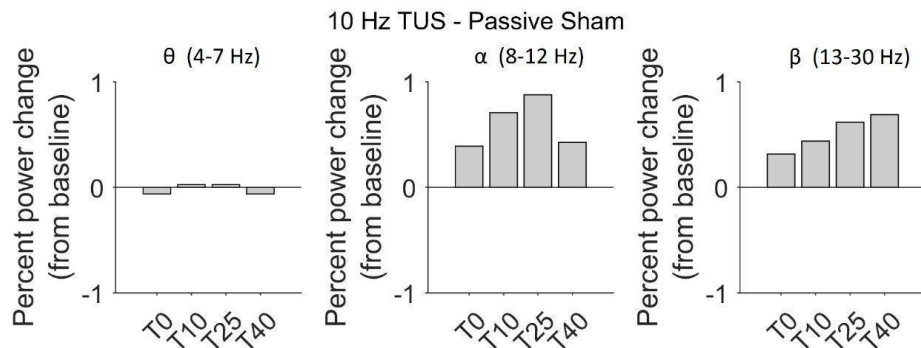
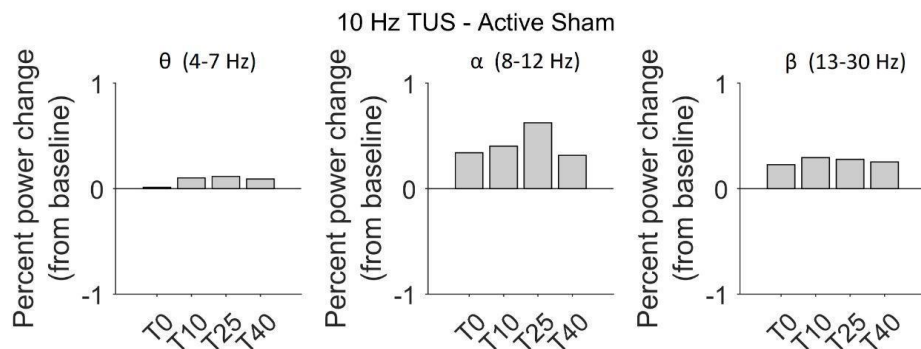
A**B****C****D**

Figure S3. Spectral power changes across different stimulation protocols for n=1 dystonia patient. Similar to the PD patients, both tbTUS and 10 Hz TUS increased the overall spectral power of the GPi LFPs in this dystonia patient.

Since this is single-subject data, we could not perform a statistical analysis. However, based on the observed trends, this data suggests that both tbTUS and 10 Hz TUS increase power in the alpha and beta bands. The increase in beta-band power is consistent with findings in PD patients. In the alpha band, the power increase compared to sham conditions appears to follow an inverted U-shaped curve, peaking around 25 minutes after stimulation. This may suggest that the TUS effects start to diminish after 25 minutes, aligning with our previous works on the motor cortex^{7,10}. In the motor cortex, we observed that the maximum TUS effects on corticospinal excitability, measured by TMS-induced motor-evoked potentials, occurred approximately 30 minutes after stimulation before gradually decreasing back to baseline¹¹. It is important to note that the dystonia patient was taking propranolol (a beta-blocker) only once per day, and therefore, drug wear-off is not a contributing factor in this case.

We would like to emphasize that the lack of theta band modulation in this single dystonia patient, does not contradict the PD results. While dystonia and PD share some symptoms and circuitry, their underlying pathological circuits are largely distinct. Hence, TUS may differentially affect these conditions.

6. Demonstrating distinct patterns of LFP change with theta and 10hz TUS would be both novel and significant, however the strength of this conclusion is at present limited by the above.

We believe that we have convincingly demonstrated at several points in the manuscript that tbTUS and 10 Hz TUS can differentially impact the brain.

1. The thermal rise induced by 10 Hz TUS is significantly larger in the brain (see **Figure 2C, F**).
2. The overall power changes caused by 10 Hz TUS were greater than those induced by tbTUS (**Figure 4C**). Additionally, tbTUS primarily increased LFP power in the theta band during stimulation, while 10 Hz TUS elevated power in the beta band, with effects lasting up to 40 minutes post-stimulation (**Figure 5**).

We also conducted a direct comparison of tbTUS and 10 Hz TUS power changes at different time points across various frequency bands. As shown below (consistent with observations in **Figure 5**), there was a significant difference in power changes between the theta and beta bands. During stimulation, the power increases in the theta band with tbTUS (compared to passive sham) were significantly larger than those observed with 10 Hz TUS (t-test, $p = 0.0082$). In the beta band, however, there was a significant power increase following 10 Hz TUS compared to tbTUS at T2 ($p = 0.0021$, t-test) and a statistical trend at T10 ($p = 0.06$, t-test). For this comparison between tbTUS and 10 Hz TUS, passive sham power values were

subtracted from both tbTUS and 10 Hz TUS measurements. In the revised manuscript, we added a supplementary figure, Figure S2 (see below), that illustrates these results. We also incorporated the following text into the paper (page 11):

“...We also compared tbTUS and 10 Hz TUS directly and found significantly larger power increases in the theta band for tbTUS (Figure S2) compared to 10 Hz TUS (t-test, $p = 0.0082$). Additionally, there was a significant difference in beta power increases at T25 (paired t-test, $p = 0.021$). Altogether, these results indicate that tbTUS and 10 Hz TUS protocols have distinct effects on neural activity within the GPI.”

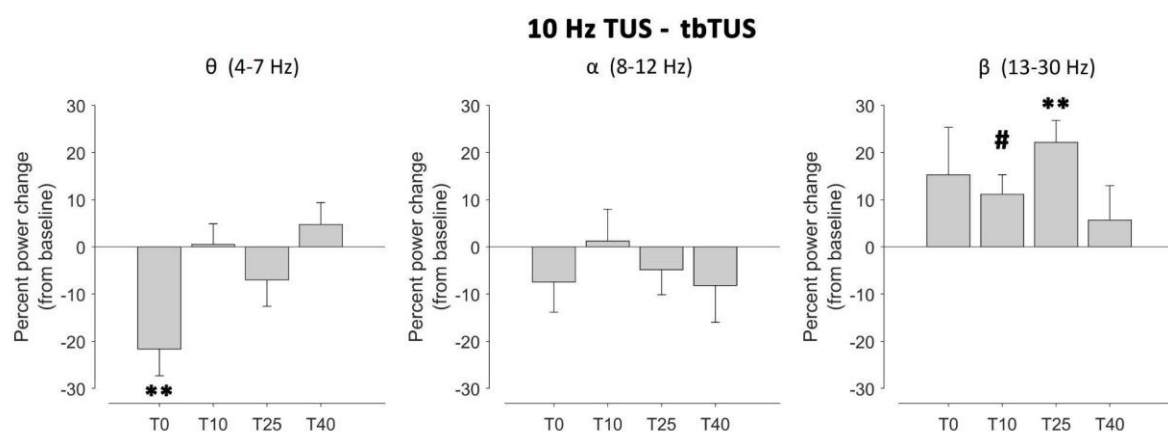


Figure S2. Differential effects of tbTUS and 10 Hz TUS on percent power changes across frequency bands. Bar plots show the percent power changes from baseline in the theta (4–7 Hz), alpha (8–12 Hz), and beta (13–30 Hz) frequency bands at different time points (T0, T10, T25, and T40). Significant differences between conditions are denoted by ** ($p < 0.01$), indicating distinct effects of tbTUS and 10 Hz TUS on neural activity within these frequency bands. Error bars represent the standard error of the mean (SEM). # represents $p = 0.06$.

In summary, the work presented here is both interesting and novel. Without revision, however, the results in experiment 1 as presented do not support the stated conclusion to a sufficient degree. Experiment 2, whilst more robust, in isolation unfortunately does not meet the threshold for publication in a high impact journal such as Nature Communications.

We hope that the additional analyses and clarifications we have provided address the reviewer’s concerns. We acknowledge that there is always room for improvement, and we plan to address some of the remaining concerns and open questions in upcoming studies.

Reviewer #3 (Remarks to the Author):

This manuscript is very interesting and highlights a novel approach to the use of local field potential signals (LFPs) from deep brain stimulating (DBS) electrodes during ultrasound brain stimulation. The results highlight that different ultrasound protocols (ie, varying patterns of applied sonication from theta burst to simple 10 Hz stimulation) may result in power changes in globus pallidus (GP) LFPs in Parkinson patients. The ultrasound stimulation is modeled from individual MRI scans to highlight the focus and range for a spherically focused ultrasound transducer, to show a moderately small spot size (ie, 4.7 mm wide X 27 mm long) at 6 cm within the brain, using a temporal approach to the GP. Similar to transcranial magnetic stimulation (TMS) the location and focus of the ultrasound transducer is controlled by frameless stereotaxy to allow a 2-3 mm reproducibility of localization. A second set of experiments using the same approach to GP highlights a mild disruption of a classic stop signal task in healthy volunteers.

Though the overall quality of the study is high there are multiple clarifications needed which would be helpful to the design and interpretation of the study.

Thank you very much for your precise and insightful comments. We are pleased that you found our study interesting. In the revised manuscript, we have carefully addressed your feedback and incorporated your suggestions.

1) DBS details – the DBS electrode in the PD patients is the 4 contact 3387 (or the directional B33015) with 1.5 mm contact and 1.5 mm spacing (as is commonly used in GP DBS). If so, the 4.7 mm width of the 50% max suggests that at least 2 contacts were actively targeted (vs the 1 suggested) with the stimulation, and more considering the aspect outside the 50% max range. Which 2 contacts were most involved with the ultrasound target (Fig. I with the DBS reconstruction is a bit misleading)?

We targeted contact 1 (considering the beam's width of 4.7 mm). Given that contact 1 has a width of 1.5 mm, the beam theoretically covers 1.6 mm above and below this contact. Thus, the 4.7 mm beam width slightly overlaps with contacts 0 and 2. While we updated **figure 1** to include a revised schematic representation of TUS over DBS leads, the potential targeting errors (2-3 mm) due to neuronavigation inaccuracies, as mentioned in the text and reported in studies involving fiducials and coregistration errors, cannot be excluded.

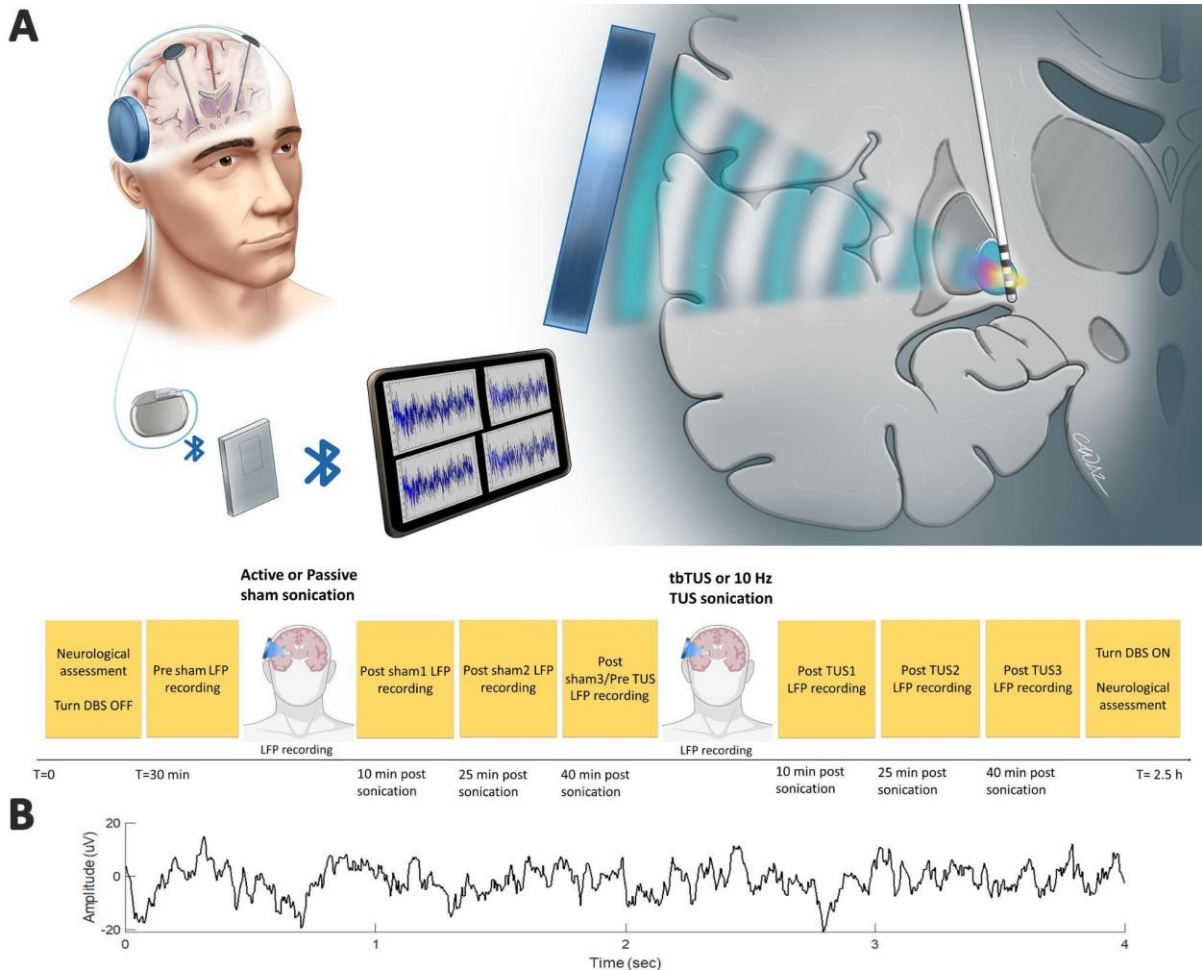


Figure 1. Transcranial ultrasound stimulation (TUS) of the GPI in Parkinson's disease and dystonia patients implanted with deep brain stimulation (DBS) devices.

2) LFPs – the Percept allows 2 channels out from a “sandwich” configuration – multiple possible contact pairs are mentioned but which was specifically used for each side (ie, 0-3 vs 0-2 bracketing contact 1, etc) for the actual LFPs? It would have been possible to compare (on the single, targeted side) on-target vs off-target pairs for comparison if the focus was on contacts 0 + 1, for example but this did not appear to have been done. Though it is mentioned that analysis was separated for the two sides it is not clear in Fig. 3 if the results (ie, Fig. 3B) are for the stimulated vs unstimulated side? This is confusing since a comparison between the two sides would be the most helpful in addition to the comparison between stimulation paradigms to highlight a focused effect unilaterally. Showing actual spectra in Fig. 3 (ie, similar to the baseline in Fig. 3B for for the unilateral, focused stimulation condition) would be very helpful and to discuss frequency bands first rather than overall power.

The reviewer seems to have interpreted that we stimulated only one hemisphere while recording from both hemispheres. However, we indeed stimulated both GPi regions sequentially with only a short delay (roughly 2 minutes)—the time required to reposition the transducer from one side of the head to the other. Thus, both hemispheres were stimulated during the experiment, with LFPs recorded online. Neural responses were also recorded

offline from both sides simultaneously at 10, 25, and 40 minutes. Unfortunately, this approach precludes the possibility of systematically investigating how unilateral stimulation affects contralateral neural responses. This is an important and interesting control experiment, and we plan to explore it in future studies to assess potential cross-hemisphere effects. Nevertheless, our group has previously shown that stimulation of the left M1 with tbTUS not only increases the left M1 corticospinal excitability measured by TMS motor evoked potentials (MEPs), but also decreases the contralateral side, the right M1, MEPs, (Xue et al., 2024). If the same contralateral effects hold for the basal ganglia, our bilateral stimulation may have indeed reduced the overall net effects (due to mutual inhibition). But these are all speculation and need future experiments.

We also analyzed the data from individual contact pairs (e.g., 0-3, 0-2, 1-3), but we did not find any significant differences between them. As an example, we provide an additional figure below showing beta-band comparisons for the tbTUS and 10 Hz TUS protocols versus the sham conditions at various time points for different contact pairs (**Figure R1**). We believe this lack of distinction between contact leads might be due to inherent errors in the navigation system, which operates within an accuracy range of a few millimeters (2-3 mm). Additionally, the focus of the ultrasound beam is sufficiently large (4.7 mm width), that combined with navigation errors, precise targeting of a specific lead contact becomes highly challenging. It is worth noting that LFP signals can be recorded up to 5 mm radially around the lead. That is why we had averaged the contact pairs to reduce the noise effect when reporting the data. This limitation of our method has been discussed in detail in the Discussion section (page 20). It is important to note that averaging LFP signals across contacts in GPi-DBS-implanted PD patients is a common approach⁶.

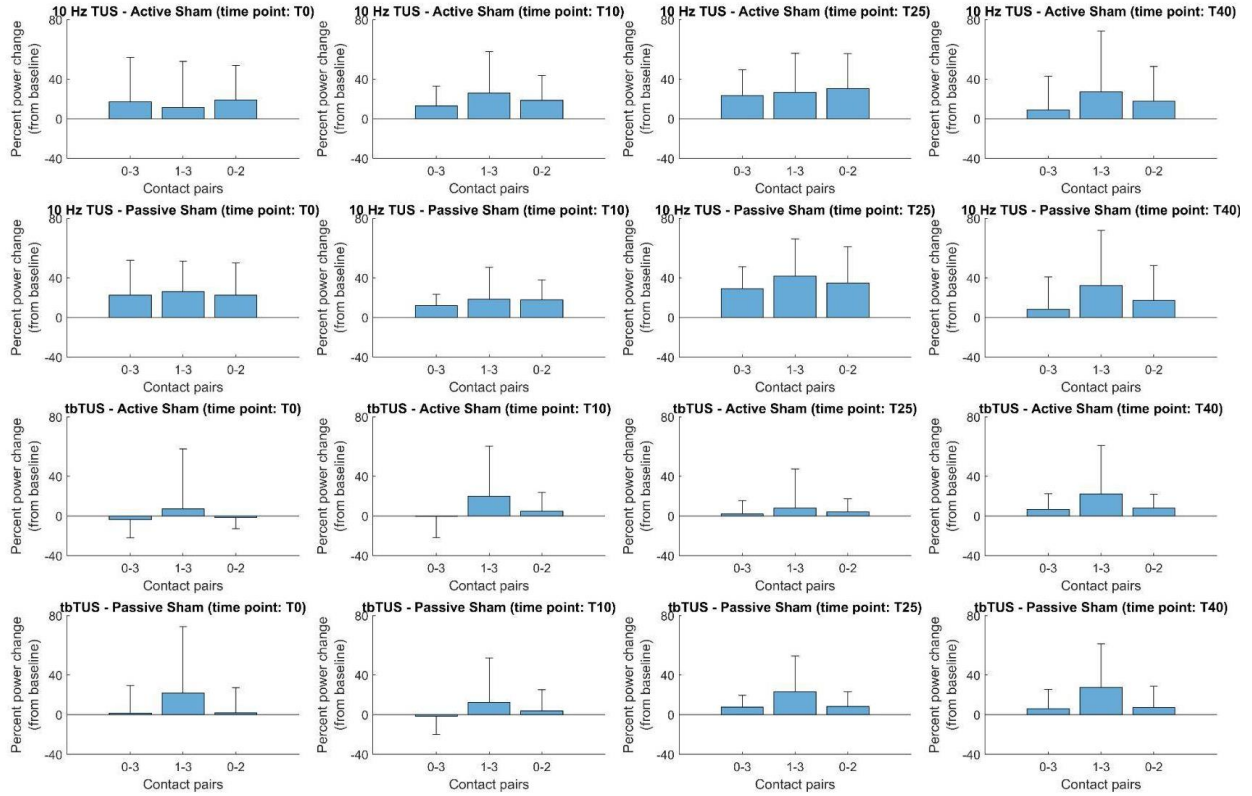


Figure R1. Beta-band comparisons for the tbTUS and 10 Hz TUS protocols versus sham conditions, shown at various time points across different contact pairs. There were no significant variations observed between the different pairs.

In the revised manuscript we acknowledge this limitation in page 20:

“While our methodology was carefully planned and performed during the session, it is important to acknowledge the potential limitations associated with neuronavigation using optical tracking, which may introduce a tracking error of 2-3 mm¹²⁻¹⁴. This level of variability could impact precise engagement with small targets such as a single DBS lead contact. Indeed, neuronavigation error might be one of the reasons that when we analyzed the data from individual contact pairs (e.g., 0-3, 0-2, 1-3), we did not find any significant differences between them. Additionally, the focus of the ultrasound beam is sufficiently large (4.7 mm width) that, when combined with navigation errors, precise targeting of a specific lead contact becomes highly challenging. To address this, we averaged the contact pairs to reduce noise effects when reporting the data. Despite this potential misalignment, our LFP readouts clearly demonstrate TUS-induced modulation in the GPi compared to sham conditions, underscoring the effectiveness of our approach (notably, LFP signals can be recorded up to 5 mm radially around the lead). Future improvements could involve exploring alternative methods to better stabilize devices^{15,16}.”

3) FFT analysis – how was the windowing into 2 sec bins decided? Since there was a longer time of recording 8-10 sec bins may show much better smoothing and more definitive results and better clarification of frequency bands – where the shorter and longer bins compared?

Using a 10-second window can over-smooth the power spectral density (PSD), often attenuating or completely missing pathologic peak activities. We chose a 2-second window for several reasons:

1. Previous studies that analyzed LFPs in PD commonly use window lengths 2 seconds or less to balance temporal and frequency resolution. Our use of a 2-second window aligns with these established practices, ensuring consistency and comparability with prior research^{5,17–20}.
2. While longer windows improve frequency resolution and reduce spectral leakage, the 2-second window strikes a balance between resolution and the need to detect beta-band (13–30 Hz) and theta-band (4–10 Hz) oscillations, which are critical in PD pathophysiology²¹. Additionally, our window length was sufficient to capture ~7 cycles of the lowest frequency of interest (4 Hz in the theta band), which is generally considered necessary for reliable spectral analysis. This ensures the 2-second window is long enough to provide meaningful and accurate estimates of oscillatory activity. Our original analysis clearly identified power peaks within these bands using the 2-second window. In comparison, pathologic peaks were significantly attenuated or missed with a 10-second window, as shown in the example provided. Furthermore, we used a robust spectral analysis approach with the multi-taper method of fast Fourier transform (MTMFFT), incorporating a Hanning taper to reduce spectral leakage and 2 Hz spectral smoothing to enhance the clarity of peaks while reducing noise.
3. During the 10 Hz TUS condition, LFPs were recorded online over a short 40-second stimulation period. Segmenting this data into 10-second windows would yield only 4 trials, which could be very noisy and limit statistical power. By segmenting the data into 2-second trials, we increased the number of trials, improving the reliability of spectral analysis and enabling a more robust evaluation of this condition.

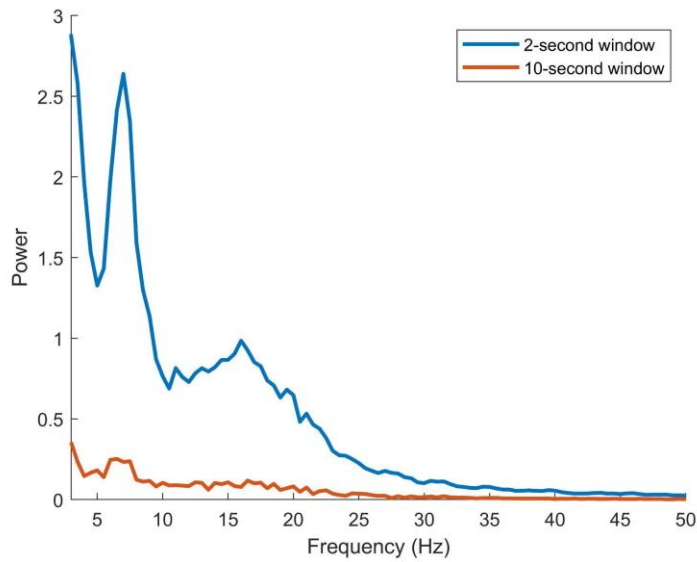


Figure R2. An example of the effects of different window sizes on power spectral density

4) Ultrasound modeling – Fig. 2 shows the modeled focusing on multiple structures including the GP, internal capsule and thalamus, almost to the midline (ie, 27 mm). it would be helpful ti Fig. 2i included these other structures within the confines of the modeled region of stimulation.

We have revised the figures to include additional structures as suggested by the reviewer. The changes can be seen in **Figure 2I-K**.

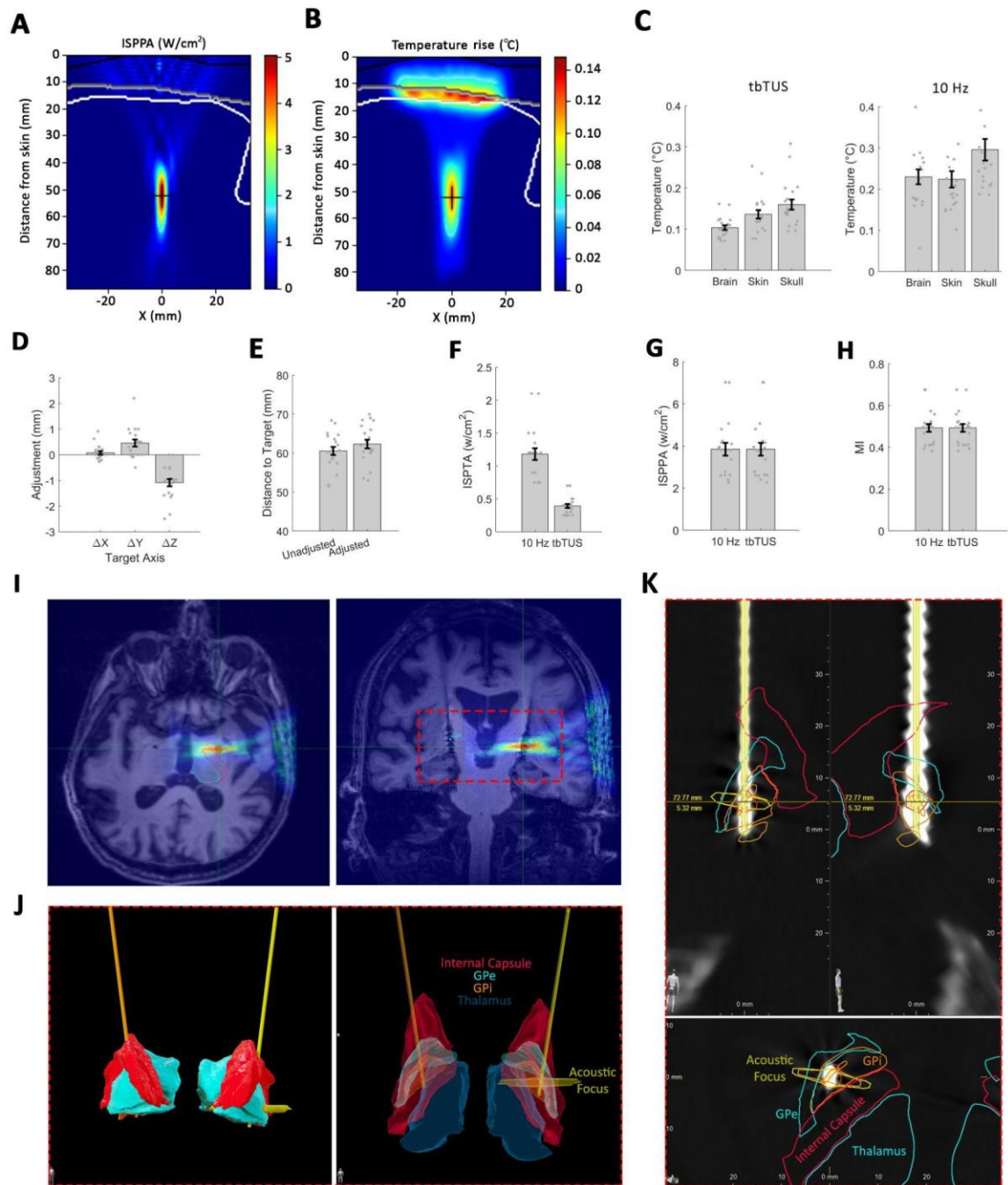


Figure 2. Acoustic simulation estimates of the acoustic pressure distribution and temperature changes in the GPI-DBS patients (A) Spatial intensity distribution for patient S10, with red indicating the highest intensity and blue the lowest. The ultrasound beam's focal point was located within the GPI. (B) Estimated temperature rise in different brain tissues for the same patient S10, with the highest temperature rise observed at the skull due to absorption and heating effects from the ultrasound energy. In both simulations, the transducer was positioned externally at the top, with the skull's outline and ultrasound waves shown penetrating through to the target area. (C-H) Comparative analysis of simulated ultrasound parameters for two sonication protocols, theta burst transcranial ultrasound (tbTUS) and 10 Hz TUS. (C) Estimated temperature in the skin, skull, and brain for both protocols. (D) Transducer adjustments in subject space along the x, y, and z axes in mm to compensate for ultrasound beam deviations based on BabelBrain simulation results. (E) Sonication depth based on BabelBrain simulation results compared to the unadjusted distance from the transducer to the target location. (F-H) In situ

intensity values of Spatial Peak Temporal Average (ISPTA), Spatial Peak Pulse Average (ISPPA), and Mechanical Index (MI) for both protocols. Note that 10 Hz TUS results in higher ISPTA values than tbTUS. Data are presented as mean $\pm$ standard deviation. (I) Pressure maps derived from ultrasound modelling superimposed on the MRI image of patient S04, highlighting the targeting of the DBS lead within the left GPi. (J) Postoperative CT imaging fused with preoperative MRI, featuring automatically segmented 3D visualizations of the GPe (turquoise), GPi (orange), internal capsule (red), thalamus (navy blue), and the DBS lead. These elements are superimposed onto the simulated acoustic focus (yellow) using BrainLab software. (K) The in-line view of the DBS lead artifact for the same patient (S04), along with the simulated acoustic focus (with the epicenter outlined in red, orange, and yellow), is shown. The lower part of Figure K illustrates the size and location of the acoustic focus in comparison to other structures, including the GPe, GPi, internal capsule, and thalamus.

5) On med condition – though the DBS was off for the LFP recordings the patients continued on their (presumed) ordinary med schedule with similar UPDRS, so the LEDD (which is highly variable between patients for optimal med on condition) is not helpful. A more helpful comparison would have been the effect of the ultrasound on med off/stim off vs med on/stim on, to assess if the ultrasound provided any form of treatment (or enhancement) effect on the inherent PD off condition (ie, similar to DBS or meds). This would help considerably show the difference between the theta burst and 10 Hz patterns much more convincingly than as shown in the med on condition and whether the apparent functional neuromodulation might extend to a treatment effect at this dosing.

We agree with the reviewer that the med off/stim off condition can be used to compare the effects of TUS to DBS. However, it is challenging for advanced PD patients who had DBS to be studied in a lengthy study in the med off/stim off condition. based on previous findings, the effects of TUS on PD patients appear to be brain state-dependent, please see our detailed response to reviewer #1 comment #4. Specifically, patients on medication exhibited larger plasticity effects following TUS applications compared to those off medication⁷. This may be attributed to impaired LTP plasticity effects in older PD populations. Moreover, patients on medication showed slight improvements in UPDRS scores after TUS applications, while off-medication PD patients did not exhibit significant effects^{7,8}. Given that the current study served as an exploratory first effort to determine if brain signals could be modulated through TUS, we decided to include the on-med group first and in the case of a reliable response, a follow-up study could incorporate a med off PD group for further clinical investigation (it is important to note that recruitment and conducting such a study would be extremely challenging, as the med-off/DBS-off condition is expected to result in UPDRS scores of 60–80 for patients).

6) Ultrasound dosing – the resulting W/cm² is very similar to the putaminal levels noted in the Meng et al 2022 (Movement Dis 37: 2133) focused ultrasound BBB study (~ 2 W/cm² across 4 patients) where the localization was confirmed directly using microbubbles, but using the standard HiFu system. Having direct MRI confirmation of the modeled ultrasound localization using microbubbles and BBB opening with contrast verification on MRI would be very helpful in 1-2 patients. Likewise, though the max output from the spherical ultrasound

device was used it would be helpful to develop a dose response curve for the putative excitatory (theta burst) and inhibitory (10 Hz) parameters, one of the largely missing but critical aspects of any proposed neuromodulation.

We agree with the reviewer that direct MRI confirmation using microbubbles and BBB opening would be highly beneficial. Unfortunately, there is no Insightec mid-intensity FUS transducer for BBB opening at Toronto Western Hospital. Even if we access the one at Sunnybrook where the study by Meng et al. was conducted, we are currently unable to perform any ultrasound procedures on patients implanted with DBS devices or healthy participants using Insightec machines. This requires comprehensive safety evaluations, as well as ethical approval and Health Canada authorization.

We used the same ultrasound exposure parameters (ISPPA $\times$ SD $\times$ DC) for both protocols to ensure comparable neuromodulatory effects, resulting in an energy delivery of 360 J/cm² for each protocol. Additionally, we conducted a correlation analysis between individual LFP power changes for each protocol and the corresponding ISPPA (W/cm²) at the target calculated by BabelBrian software. However, no significant effects were observed.

We added the following sentences to page 16:

“Finally, we examined the correlations between the ISPPA (as an estimate of exposure) and power changes across various conditions and time points but found no significant correlations. Although no exposure-related changes in the LFPs were observed—possibly due to small individual variations in the ultrasound doses delivered to the target locations—future studies would benefit from developing a exposure-response curve for the tbTUS and 10 Hz TUS protocols.”

7) The second population of control patients with possible interference on a behavioral task is much less strong of an experiment compared to the actual LFP recordings and adds much less to the overall conclusions. Comparing the theta burst to 10 Hz for differences would have been a helpful differentiation of proposed mechanisms and the results are unclear as to how the circuits might be affected by parameters.

The primary goal was to generalize the findings and evaluate the efficacy of individualized TUS at widely used intensity ranges (30 ISPPA W/cm²) on deep brain structures in both health and disease. This was achieved through a combination of electrophysiological and behavioral experiments. We agree that LFP recordings provide stronger and more direct evidence of target engagement. However, the behavioral task complements these findings by demonstrating the behavioral consequences of TUS-induced neuromodulation in a broader population (healthy individuals). Notably, the significant increase in SSRTs following tbTUS

targeting the GPi, but not the pulvinar, highlights the spatial specificity of TUS as a new technique.

8) Overall, the study could be significantly improved by comparing presumed on target vs off target effects (on both a single electrode on the stimulated side and comparing the two sides), effects on inherent beta power in the PD off condition vs med on and whether there may be a more apparent neuromodulatory effect on the PD condition. The focus is reasonable but needs to be clarified with respect to actual contacts and the results much better presented. There may considerable off target effects on internal capsule and thalamus vs GP. The actual frequency on meds is presumably mainly alpha and not clear if modulated by eye opening or not and not clear how GP might affect the overall alpha frequency and rhythm with either stimulation paradigm. Though carefully controlled in terms of power and modeling to be safe, many alternative conditions may be much more revealing in this fortuitous population of PD patients with Percept LFP recordings for actual mechanisms and effects in humans.

We appreciate the reviewer's detailed and constructive feedback. Below, we addressed each aspect of the comment:

1. We acknowledge the importance of explicitly comparing on-target and off-target effects. While our study focused on GPi-targeted TUS, we recognize that off-target effects on structures such as the internal capsule cannot be entirely ruled out. However, our acoustic simulations and modeling indicate that the highest energy of the ultrasound was primarily concentrated within the GPi, with minimal spillover to surrounding structures like the internal capsule. We have included further clarification of this in the revised manuscript and Figure 2. In future studies, incorporating additional electrode contacts, with higher spatial resolution similar to those used in patients with SEEG implants, could provide a deeper understanding of off-target effects.
2. We agree that comparing the effects of TUS in the medication-off versus medication-on states would provide valuable insights into the interaction between dopaminergic states and neuromodulation. However, given previous findings as discussed in details (see response to point 5 above), and the logistical challenges regarding our REB approval, we chose to conduct all sessions in the on-medication state to also prioritize patient comfort (as med-off/DBS-off state would be highly challenging for most PD patients and would significantly reduce our recruitment potential) for this first study.
3. The reviewer raises an important point about alpha activity in the GPi and its modulation by factors such as eye opening. However, we did not observe any significant LFP modulation in the alpha frequency band in our study. We also added to the method section that all the GPi LFP recordings were conducted in a resting state with eyes open across all conditions to maintain consistency.

4. To address contact specific data, we have added to the discussion that when we analyzed the data from individual contact pairs (e.g., 0-3, 0-2, 1-3), we did not find any significant differences between them, most likely due to neuronavigation error (see our response to review #3 comment #2, and **Figure R1**). Additionally, we have updated the **Figure 1** and **Figure 2** to better illustrate the spatial specificity of TUS effects.

Reviewer #4 (Remarks to the Author):

Authors Darmani et al presented an intriguing and important approach to understand the effect of transcranial ultrasound stimulation on the basal ganglia.

Thank you very much for recognizing the importance of the work and your thoughtful feedback. We hope our findings contribute to advancing the understanding of TUS as a neuromodulation tool for deep brain structures.

While this is an important question to understand, the following questions arose about the methodology and interpretation of the results presented by the authors:

1. Patient selection: Barring the low sample size, there isn't sufficient clarification of why one patient with dystonia was included in the cohort to study the effects of TUS along with 9 other PD patients. A further point of classification is pertinent in figure 3, where plots C to F: has data across 18 hemispheres, n = 9 PD patients – here, the absence 10th data point, i.e., the absence of the dystonia patient is not justified as per knowledge. Clarification on this point would be appreciated. Statements such as “A similar pattern was also observed for dystonia patients with GPi-DBS” should be toned down drastically given that it is based on only one data point (if at all, since this data point was absent in a plot). It would also be interesting to categorize the patients into a symptom subtype and understand the influence of TUS on each subpopulation, given that one sub population may show different effects to others.

As in our response to review #1 regarding the dystonia patient, the inclusion of the dystonia patients were intended to provide insights into the effects of TUS in a broader clinical population (both PD and dystonia) with GPi-DBS implants.

However, due to challenges in recruiting additional patients within the study timeframe, the sample size for the dystonia cohort was insufficient to draw conclusions. Consequently, we have decided to present these findings in the supplementary information rather than the main manuscript.

In our analysis, we observed a power increase in this dystonia patient compared to sham conditions, similar to the patterns observed in PD patients (please see our response to the reviewer#1 comment #5). This may suggest that TUS have engaged the target in this patient

as well. However, given the distinct pathophysiology of dystonia versus PD, we presented this patient's data separately. To address the reviewer's suggestion, we revised the statement (page 11) to:

"Notably, TUS of the GPi in the dystonia patient with GPi-DBS implants also increased overall power (Figure S3), similar to the effects observed in PD patients. The power changes were most prominent in the alpha and beta bands (Figure S4). Further research is needed to determine whether these findings can be generalized to a larger population of dystonia patients."

As requested, we also analyzed the potential differential patterns between two subpopulations of PD patients, specifically the akinetic-rigid group (patients 4, 5, 6, 7, 8, 9) and the tremor-dominant group (patients 1, 2, 3). We have added the following figure, which illustrates these results, to the supplementary information as a new figure (Figure S5).

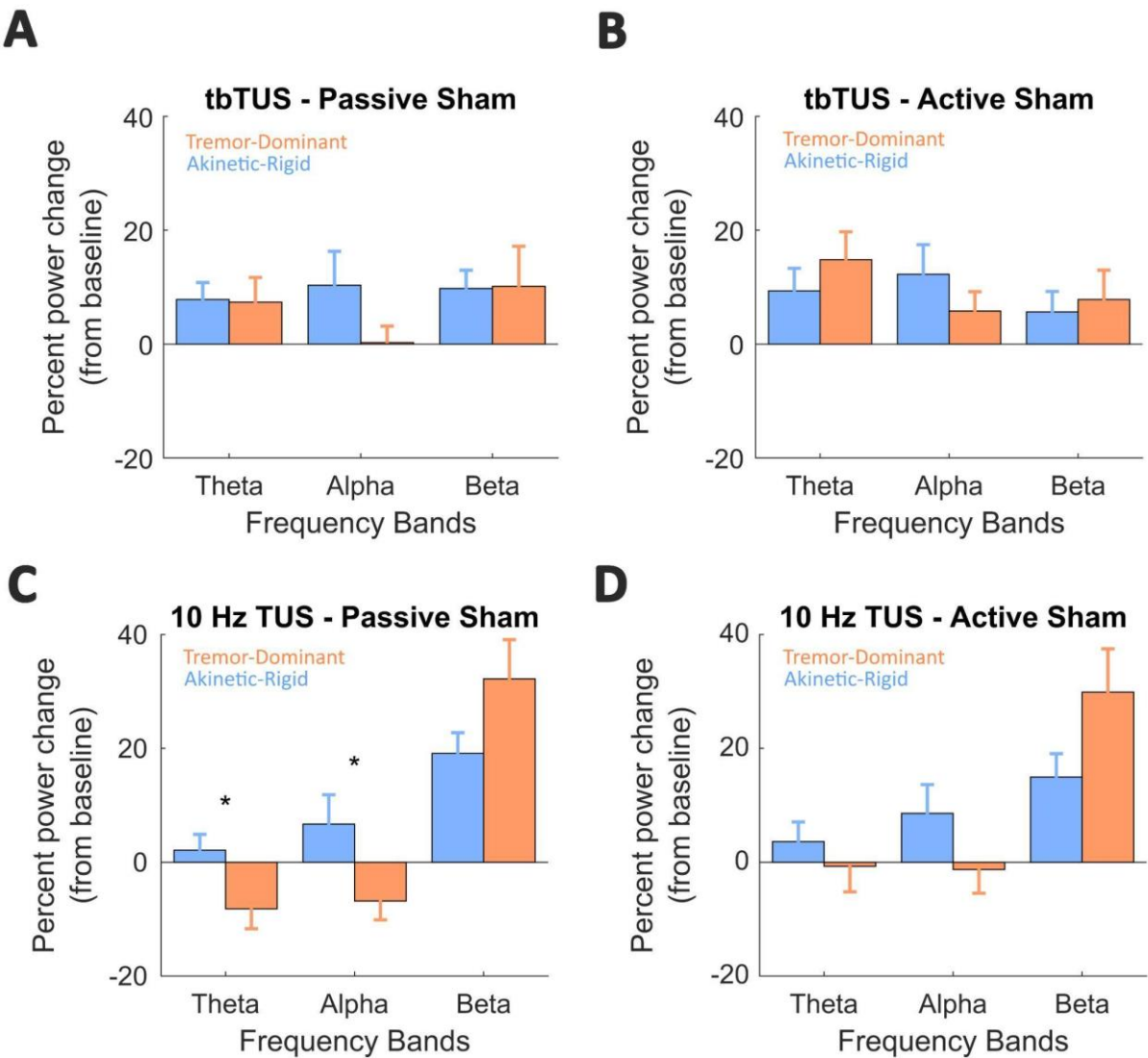


Figure S5. Percent power change from baseline in distinct frequency bands in response to TUS of GPi in two patient subpopulations. Data are presented for two patient populations: tremor-dominant (orange) and akinetic-rigid (blue). Error bars represent the standard error of the mean. Panels compare the effects of different TUS conditions on these populations: (A) tbTUS - Passive Sham, showing no significant changes; (B) tbTUS - Active Sham, also showing no significant changes; (C) 10 Hz TUS - Passive Sham, highlighting significant differences in α and β power (* indicates $p < 0.05$, t-test); and (D) 10 Hz TUS - Active Sham, demonstrating no significant difference in power for the two groups. Statistical significance is denoted by asterisks (* indicates $p < 0.05$, t-test).

Here we found a significant difference in the theta and alpha bands ($p = 0.04$ and $p = 0.04$, t-test) only in response to the 10 Hz TUS condition compared to the active sham condition. This finding is particularly interesting because an increase in 4–10 Hz power in the globus pallidus has been shown to correlate with symptom improvement in PD patients undergoing DBS treatment. Therefore, the benefits of 10 Hz TUS may be more pronounced for akinetic-rigid PD patients compared to tremor-dominant patients. We have included this information in the main text (page 13) as follows:

*“PD patients exhibit diverse symptoms, which may lead to unique effects of TUS for different subpopulations. To explore this, we analyzed power changes across various frequency bands in two distinct subgroups of PD patients: the akinetic-rigid group (patients 4, 5, 6, 7, 8, 9) and the tremor-dominant group (patients 1, 2, 3). As shown in **Figure S5**, we observed a significantly higher theta and alpha power ($p = 0.04$, $p = 0.04$) in the akinetic-rigid compared to the tremor-dominant group in response to the 10 Hz TUS compared to the active sham condition (data pooled across four timepoints). This result is particularly noteworthy, as increased power in the 4–10 Hz range within the globus pallidus has been associated with symptom improvement in PD patients²¹. These findings may suggest that 10 Hz TUS may offer greater therapeutic benefits for akinetic-rigid PD patients compared to tremor-dominant patients in this small cohort.”*

2. Evidence for target engagement: The authors have provided two experimental questions: to employ local field potential recordings in patients using percept Medtronic Percept – this is performed on patients with movement disorders. However, to assess the target engagement, a second experiment using healthy controls is employed. It is not clear if this setup exactly justifies the target engagement since the two tests are performed on different set of subjects.

We thank the reviewer for raising this important point. The experiment on healthy participants was not designed as a control for the experiment on patients with movement disorders. Instead, the two experiments were conducted to address different aspects of target engagement in a broader population and they are complimentary to each other.

In the first experiment on patients, evidence for target engagement was provided by direct recording from deep structures and by the fact that only the active TUS protocols (tbTUS and 10 Hz TUS) applied to the GPi induced significant electrophysiological effects in the target region. Neither the passive sham protocol nor the active sham protocol produced comparable effects. This indicates that the observed effects were specific to the GPi and were not due to non-specific factors or generalized stimulation effects.

In the second experiment on healthy participants, we sought to explore whether TUS-induced modulation of the GPi could produce behaviorally relevant effects. The results showed that only TUS of the GPi (a region well-documented to play a critical role in response inhibition) and not the pulvinar (a region not associated with response inhibition) influenced stop-signal reaction times (SSRTs). These findings reinforce the specificity of TUS for the GPi and further support target engagement by linking it to a functional outcome.

3. Experimental design and lack of sufficient evidence: While each experiment is setup well, with the second experiment having good design of contrasting GPi sonication with pulvinar nuclei sonication, the statistics seen in Figure 9 are not sufficiently convincing to show that response inhibition changes with stimulation. The difference between GPi pre- and post1 and Pulvinar pre- and post1 stop signal reaction should be compared, in addition to comparing pre- and post. Even a one tailed t-test could provide insight into if this is significant. Additionally, it would be good to show the mean as a highlighted point, rather than median since the t-test also compares the means. This might even show lower difference between the groups.

We conducted an independent t-test to compare the changes in SSRT from baseline between the two groups. The analysis revealed a statistically significant difference between GPi and pulvinar SSRTs ($p=0.047$). We did not find any significant difference between the two regions when we compared changes in go trial reaction times from baseline ($p=0.87$, t-test). These results were incorporated into Figure 9E-F, where the comparison between GPi and Pulvinar is visualized.

We also added this information to the revised manuscript and caption of Figure 9 as follows:

“There was also a significant difference between the baseline-normalized SSRT changes (post - pre) for GPi and pulvinar ($p=0.04$, t-test, Figure 9E). We found no significant difference in the go trial reaction times between the two regions ($p=0.87$, t-test, Figure 9F). For these comparisons, data from the first and second post- measurements were pooled.”

We would like to clarify that the data in **Figure 9** actually shows the means, not the medians which was mistakenly mentioned in the original caption. We apologize for this oversight. In

the revised manuscript, this error has been corrected to accurately reflect that means were used.

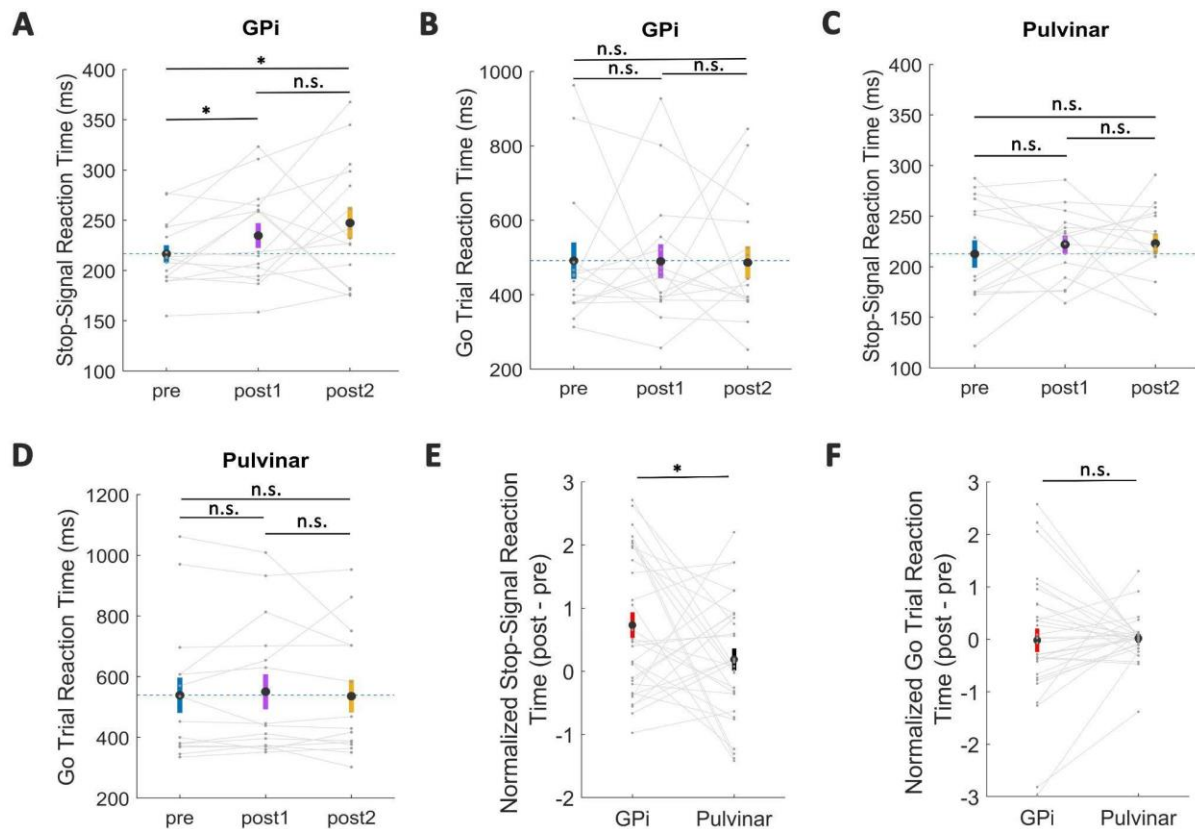


Figure 9. Stop-signal task performance pre- and post-intervention in GPI and pulvinar targets. (A) Stop-signal reaction times (SSRTs) for the GPI target, displayed across three different time points: pre-intervention, post-intervention 1 (post1), and post-intervention 2 (post2). Each data point represents an individual participant's SSRT with group mean marked, demonstrating significantly increased SSRT post-intervention. (B) Go trial reaction times for the GPI target across the same time points as in panel (A), with individual participant data and group mean indicated, showing no significant change. (C) Stop-signal reaction times for the pulvinar target, with individual participant data, displaying no significant change across the time points. (D) Go trial reaction times for the pulvinar target, also showing individual participant data without significant differences across the time points. (E) Comparison of stop-signal reaction times (post - pre) between the two stimulated regions, with individual participant data, displaying significant longer SSRTs when GPI was sonicated. (F) Go trial reaction times did not differ across the two stimulation regions. In E and F to account for differences in baseline variability between the GPI and pulvinar sessions, post-TUS values were normalized by the standard deviation of the pre-TUS values. The asterisk (*) denotes statistically significant differences, and 'n.s.' indicates non-significant differences.

4. Additional smaller points:

- To understand the effect of TUS stimulation as a function of LEDD, it would be important to showcase the results of active – passive sham with LEDD.
- While two protocols of stimulation have been considered here – i.e., tbTUS vs 10 Hz stimulation & their correlation with patient medication profile has been provided to some

extent, the effects of each stimulation in relation to clinical improvement is not entirely elucidated clearly. In fact, either of the stimulation protocols have no significant effect on the clinical profile of the patient. This point should be discussed more in detail, and the relation between clinical improvement and target engagement (if this is a positive correlation) should also be specified more clearly.

In the revised version, we have added the results of the power changes in the active - passive sham correlation with LEDD to the text as well as to the Table S1 in the supplementary information as follows (page 13):

“We did not find any significant correlations between LEDD and power changes in the active sham compared to the passive sham condition during or after cessation of the stimulation (Table S1).”

Table S1. Correlations between percent power change from baseline and LEDD. *r* denotes the Pearson correlation coefficient. *p*-values are uncorrected.

	T0	T10	T25	T40
tbTUS – Passive Sham	<i>r</i> = 0.68 <i>p</i> = 0.001	<i>r</i> = 0.6511 <i>p</i> = 0.0034	<i>r</i> = 0.6146 <i>p</i> = 0.0066	<i>r</i> = 0.5432 <i>p</i> = 0.0198
tbTUS – Active Sham	<i>r</i> = 0.66 <i>p</i> = 0.002	<i>r</i> = 0.6224 <i>p</i> = 0.0058	<i>r</i> = 0.5371 <i>p</i> = 0.0215	<i>r</i> = 0.5612 <i>p</i> = 0.0154
10 Hz TUS – Passive Sham	<i>r</i> = 0.63 <i>p</i> = 0.004	<i>r</i> = 0.6958 <i>p</i> = 0.001	<i>r</i> = 0.6259 <i>p</i> = 0.0055	<i>r</i> = 0.6139 <i>p</i> = 0.0067
10 Hz TUS – Active Sham	<i>r</i> = 0.35 <i>p</i> = 0.14	<i>r</i> = 0.5543 <i>p</i> = 0.0170	<i>r</i> = 0.5745 <i>p</i> = 0.0126	<i>r</i> = 0.6340 <i>p</i> = 0.0047
Active Sham – Passive Sham	<i>r</i> = 0.1495 <i>p</i> = 0.55	<i>r</i> = -0.10 <i>p</i> = 0.63	<i>r</i> = 0.10 <i>p</i> = 0.669	<i>r</i> = -0.12 <i>p</i> = 0.6268

Regarding the lack of correlation between the power changes and the MDS-UPDRS, we have mentioned in the paper that the latest MDS-UPDRS scores recorded in the clinical chart which we used, may not accurately reflect the patients' current clinical status, as some of these values were measured couple of months prior to the experiments.

Please see page 14: "It should be noted that the MDS-UPDRS scores were in the ON-medication state and were the latest evaluations from routine clinical assessments conducted up to a few months before this study."

Reviewer #5 (Remarks to the Author):

This is a really interesting and exciting paper which presents significant results demonstrating changes in signals directly recorded from DBS electrodes following TUS in PD patients, a significant demonstration of target engagement. The paper is overall clearly written and the results well presented. However, it would benefit from a few extra details, clearer explanation and discussion of some concepts and parts of the methods, and some aspects of the results – these minor comments are detailed below.

Thank you very much for your positive feedback and for taking the time to review our paper. We found your detailed comments to be highly constructive for enhancing the clarity and impact of our work for readers.

Page 4: In the description of experiments: 'The first session obtained MRI images for personalized ultrasound modelling and simulations.'

Suggest wording as e.g. 'personalized ultrasound sonication planning simulations'. Using both modelling and simulations seems redundant.

We have revised the wording to: *"The first session obtained MRI images for personalized ultrasound simulations."*

The revised text now appears on page 4 of the manuscript.

'The subsequent sessions, spaced at least one week apart, were allocated to electrophysiology recordings'

Would be clearer to mention TUS here (and when this description is repeated elsewhere): e.g. '... were allocated to TUS and electrophysiology recordings'

We have revised the text to read: *"The subsequent sessions, spaced at least one week apart, were allocated to TUS and electrophysiology recordings."*

The revised wording has also been updated in other instances where this description appears (page 4).

Page 5: We used the same ultrasound dose (spatial peak pulse average intensity (ISPPA) × sonication duration (SD) × duty cycle (DC)) in both protocols.

Dose concepts for ultrasound neuromodulation are not yet well understood, what you are describing is really an exposure – an amount of energy per unit area which is incident on the target, it does not account for any specific sensitivity of the tissue or any measure of how much of the energy is absorbed or the resulting mechanical forces for example. Please use the word exposure rather than dose here and elsewhere in the text.

We agree that the term “exposure” is more accurate in this context. We have replaced “dose” with “exposure” throughout the manuscript where applicable, including the sentence in question, which now reads (page 5):

"We used the same ultrasound exposure (spatial peak pulse average intensity (ISPPA) × sonication duration (SD) × duty cycle (DC)) in both protocols."

In addition – this was not the same across participants! The in situ Isppa varied between as you show in the figures, the exposure for each person was different. Can you comment on this and why you didn't use the simulations to adjust the exposure?

Given our goal of evaluating the efficacy of TUS in deep structures, we utilized the maximum possible output power (30 W/cm²) of our NeuroFUS device to increase the likelihood of observing modulatory effects in each individual (as some studies have reported using significantly higher ISPPA values). This approach was based on the hypothesis that a subsequent correlation analysis could be performed between individualized ISPPA at the target and LFP power modulation to potentially establish a dose-response relationship. However, our correlation analysis revealed no significant effects. This lack of significance may be attributed to the relatively small differences in ISPPA at the target within our study population.

The pulse timing parameters are a little scattered through the text, it would be useful if they could be collected together here for example (pulse duration added to the figures already given), and reported together with the other parameters in a table in supplementary material.

We added the following table as suggested to the supplementary information:

Table S2. Ultrasound protocol parameters. PRI: pulse repetition interval, DC: Duty Cycle, PRF: pulse repetition frequency, FF: Fundamental frequency, ISPPA: spatial-peak pulse-average intensity, ISPTA: spatial-peak temporal average intensity

	Pulse duration	PRI	DC	PRF	Stimulus Duration	FF	ISPPA in water	ISPTA in water
tbTUS	20 ms	200 ms	10%	5Hz	120 s	500 kHz	30 W/cm ²	3 W/cm ²
10 Hz TUS	30 ms	100 ms	30%	10Hz	40 s	500 kHz	30 W/cm ²	9 W/cm ²

Figure 1B – it might be more informative to include the spectrum of the signal here.

We agree with the reviewer that including the spectrum of the signal would be informative. However, we have presented it in Figure 3B as it is part of the results section, whereas Figure 1 is intended to provide introductory information before presenting any results. Additionally, placing it in Figure 1B would disrupt the visual symmetry of the figure. This arrangement ensures clarity and maintains a balanced visual presentation.

Page 6: Over what time period did you average to calculate the Ispta?

The calculation of ISPTA at the target in the BabelBrain involves a simulation process where the temporal averaging is integrated over the sonication duration. In our study, the thermal simulations were performed for a sonication duration of **40 seconds** with a duty cycle of 30% for 10 Hz TUS and **120 seconds** with a duty cycle of 10% for tbTUS. It is worth noting that, because this was a single exposure, the ISPTA is just the product of ISPPA and DC. Please see page 28.

‘However, ISPPA estimates within the target area were similar for the two protocols, ranging from 2.5 to 7 W/cm² (Figure 2G).’ – aren’t they the same? The set intensity was the same, all that changed was the timing parameters? Or are taking into account the actual position of the transducer in each session for post hoc analysis?

As shown in Figure 2G, the ISPPA for both protocols is equal because the same initial pressure and fundamental frequency were used for both. The variation from 2.5 to 7 W/cm² represents

individual ISPPA at the target, calculated based on individual brain characteristics, which is consistent across both protocols.

We edited the text accordingly (page 6):

“However, ISPPA estimates within the target area based on individual brain characteristics were similar for the two protocols, ranging from 2.5 to 7 W/cm² (Figure 2G).”

‘... guidelines set by the FDA⁴² as well as recommended by the International Transcranial Ultrasonic Stimulation Safety Standards Consortium (Maximal temperature rise anywhere $\leq 2^{\circ}\text{C}$; ISPPA $\leq 190\text{ W/cm}^2$; MI ≤ 1.9)⁴³’

This is slightly misleading – the Isppa limit comes from the FDA guidelines, it is mentioned in the ITRUSST paper but not recommended as a limit.

Revised Text (page 6): “The acoustic simulations predicted that the maximum temperature rise, ISPPA, and MI for both tbTUS and 10 Hz TUS protocols remain within the safety limits established by the FDA guidelines (ISPPA $\leq 190\text{ W/cm}^2$), as well as being consistent with the considerations outlined in the International Transcranial Ultrasonic Stimulation Safety Standards Consortium (ITRUSST) consensus paper (maximal temperature rise anywhere $\leq 2^{\circ}\text{C}$; MI ≤ 1.9).”

Page 7: Figure 2 – a single sided colour scale would be easier to see.

We have revised the entire figure to enhance clarity by incorporating the new structures and their labels, as suggested by Reviewer 3, as well.

Here and elsewhere, e.g. Fig 8. - I don't understand what is meant by the transducer adjustment and adjusted and unadjusted distance to target?– what is it in reference to?

Figure 8B:

In the BabelBrain software, the transducer adjustment parameters based on Δx , Δy , and Δz represent the spatial corrections applied to the transducer's position and/or orientation to ensure the ultrasound focus aligns with the intended target in the brain. These adjustments address factors such as misalignment in the initial placement of the transducer, which might not perfectly align with the desired trajectory or target due to anatomical variability or setup constraints.

Figure 8F:

Unadjusted Distance to Target: This refers to the distance measured from the transducer focus to the intended target without accounting for distortions introduced by the skull. It represents the raw alignment of the transducer's focus to the target.

Adjusted Distance to Target: This is the distance measured after accounting for the skull's effects on the ultrasound beam, such as refraction and attenuation which is calculated by BabelBrain. Adjusted distances include corrections applied to ensure the actual focal point aligns as closely as possible with the intended target. For more details on the simulation and modeling steps with BabelBrain, please see Pichardo., 2023.

We have revised the text accordingly to include more information about the transducer adjustment parameters based on Δx , Δy , and Δz and unadjusted and adjusted distance to targets. See page 6.

“The transducer adjustment parameters (Δx , Δy , and Δz) represent spatial corrections made to the transducer's position and orientation to align the ultrasound focus with the intended brain target. These adjustments address initial misalignments due to anatomical variability or setup constraints. **Unadjusted Distance to Target** is the distance from the transducer focus to the intended target without considering skull-induced distortions, representing the raw alignment of the focus to the target. **Adjusted Distance to Target** is the corrected distance after accounting for skull effects, such as refraction and attenuation, calculated using BabelBrain.”

Page 10: Since the intensity varied between participants, was there any correlation between the exposure and the change in recorded signal power?

We did examined the correlations between the ISPPA (as an estimate of the exposure) and power changes in various conditions and time points (see our response to review #3 comment #6). There was no significant correlation between the ISPPA at target and the power changes. As an example, shown below is the correlation plot for power changes during the tbTUS condition (compared to the active sham) and the estimated ISPPA values. In the revised version, we added a sentence about this lack of correlation (Page 16).

“Finally, we examined the correlations between the ISPPA (as an estimate of exposure) and power changes across various conditions and time points but found no significant correlations. Although no exposure-related changes in the LFPs were observed—possibly due to small individual variations in the ultrasound exposure—future studies would benefit from developing an exposure-response curve for *the tbTUS and 10 Hz TUS protocols*.”

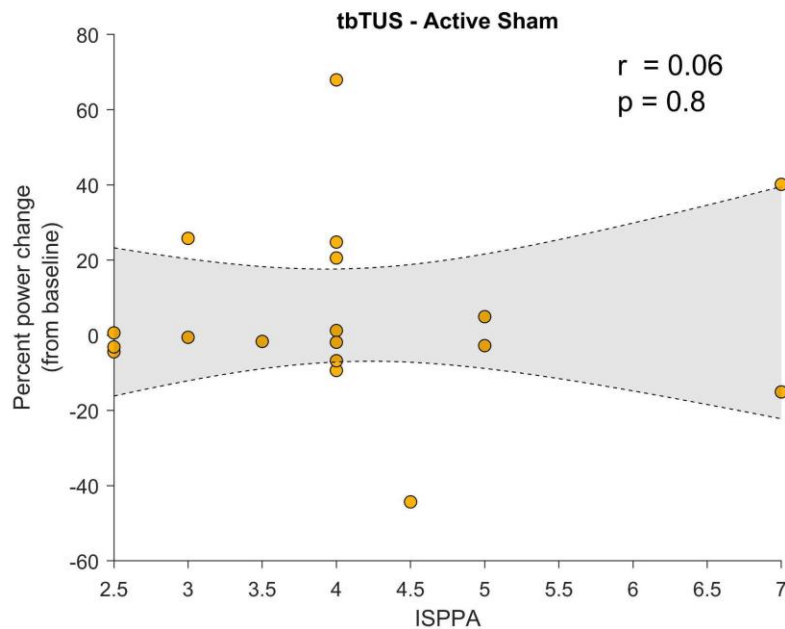


Figure R3. Correlation plot for power changes during the tbTUS condition (compared to the active sham) and the estimated ISPPA values.

‘A similar pattern was also observed for the dystonia patients with the GPi-DBS lead’ – there was only one dystonia patient.

The text has been corrected to accurately reflect that there was only one dystonia patient as this was mentioned by the other two reviewers.

Page 13: ‘no correlation was observed between power changes with tbTUS or 10 Hz TUS and the Movement Disorder Society Unified Parkinson’s Disease Rating Scale (MDS-UPDRS) scores’ – can you explain what this means and what the reason for looking at these was? Does a higher score mean more severe/advanced disease? Were these used to look for a correlation between response and disease severity? It is not clearly explained either here or in the methods.

We added the following paragraph to the discussion to address the reviewers’s comment (page 21):

“... Additionally, no correlation was found between power changes and MDS-UPDRS scores, which were taken from routine clinical assessments before the study. As higher MDS-UPDRS scores reflect more severe or advanced Parkinson’s disease, the lack of correlation suggests that TUS effects may be independent of overall disease severity. Future investigations should directly measure clinical and behavioral outcomes immediately before and after TUS applications to clarify these effects.”

Page 16: Figure 8: '(G) Pressure maps derived from ultrasound modelling superimposed on the MRI image of one exemplary subject, highlighting successful targeting of the GPi and pulvinar'

Do you mean that the focal region overlapped with the target region?

Yes, the statement refers to the focal region of the ultrasound beam overlapping with the intended target region (GPi or pulvinar) as determined by our acoustic modeling.

We have revised the text (page 18) to improve clarity:
"Pressure maps derived from ultrasound modeling were superimposed on the MRI image of one exemplary subject, demonstrating that the focal region of the ultrasound beam successfully overlapped with the intended target regions (GPi and pulvinar)."

Page 17: Discussion – how does the change in LFP after TUS compare with the changes seen due to DBS?

LFP modulations due to DBS in GPi patients are typically characterized by a reduction in beta power, which is associated with pathological activity, and a normalization or an increase in theta-alpha power. In the current study, while 10 Hz TUS increased beta power, an effect that may be undesirable, tbTUS led to an increase in theta power. However, the effects of tbTUS were short-lived, occurring only during the sonication period.

These findings highlight the need for further development of TUS protocols to achieve more sustained and desirable outcomes. Specifically, longer-lasting interventions are required to effectively reduce beta power, the main pathological frequency band in movement disorder patients, while simultaneously normalizing or enhancing theta-alpha power to better mimic the therapeutic effects seen with DBS.

Page 18: We observed that increases in GPi LFP power were stronger with 10 Hz stimulation – greater would be a better description than stronger – or be quantitative – how much larger were the increases?

We changed that sentence to the following (page 19):

"We observed that increases in GPi LFP power were larger with 10 Hz stimulation compared to tbTUS ($10\text{ Hz TUS}_{\text{mean}} - \text{tbTUS}_{\text{mean}} = 5.56$, Figure 4C)."

'... has been demonstrated in animal models, as well as in studies in healthy subjects' -> healthy human participants.

The text has been revised to read (page 19):

"... has been demonstrated in animal models, as well as in studies in healthy human participants."

‘We characterized the time course of LFP bands.’ – wording – examined the signals at different time points?

The text has been revised to read (page 20):

"...examined the LFP power changes at different time points in different frequency bands."

‘While our methodology was carefully planned and aligned during the session’ – wording – be specific – the transducer was carefully aligned?

The transducer was positioned and oriented with high precision, incorporating adjustments informed by BabelBrain's simulations, to ensure effective and accurate ultrasound delivery to the intended target.

The text has been revised (page 20):

"While our methodology was carefully planned and the transducer was positioned and oriented with high precision during the session"

Page 19: ‘future improvements could be achieved by exploring alternative methods to better stabilize the devices during experiments’ – how was the transducer positioned and held? Did you record the actual positions of the transducer during sonication? How close did you get to placing the transducer in the planned position when delivering the sonication?

The transducer was positioned using a neuronavigation system (Brainsight, Rogue Research) to precisely align it with the planned sonication trajectory. Once the target was fully aligned, the transducer was manually held against the head to maintain stability during sonication, with visual feedback provided by Brainsight. While this approach proved effective, we acknowledge that future improvements, such as the use of custom fixation devices or robotic arms, could further enhance the precision and stability of transducer placement, particularly for single or multiple-channel transducers.

‘The enhancement of beta band activity, potentially associated with increased motor symptoms in PD, was evident with 10 Hz stimulation’ – does this suggest that

neuromodulation with this protocol is not suitable for suppressing symptoms? Were changes in symptoms/behaviour after TUS assessed or observed?

We thank the reviewer for raising this important point. Increased beta activity is often associated with worsened motor symptoms in Parkinson's disease. However, clinical scores or behavioral changes following TUS were not assessed in this study, as the primary objective was to evaluate target engagement and the efficacy of TUS using electrophysiological data, rather than its clinical outcomes at this stage.

That said, general neurological safety evaluations conducted before and after each session revealed no adverse effects or worsening of symptoms in patients. This suggests that while beta band activity was enhanced, it did not result in significant observable changes in clinical symptoms within the timeframe of the experiment. Alternatively, the observed changes may have been subtle to be detected through the general neurological assessments performed.

The potential utility of 10 Hz and theta burst TUS for symptom suppression in PD requires further investigation, including studies designed to directly measure clinical scores changes before and after TUS applications. We have clarified this in the revised manuscript (page 21) and highlighted the need for future studies to assess the clinical implications of TUS-induced changes.

“Increased beta activity is often associated with worsened motor symptoms in Parkinson's disease, raising questions about the suitability of 10 Hz TUS for symptom suppression. But Clinical scores or behavioral changes were not directly assessed in this study, as the focus was on target engagement and electrophysiological effects. General neurological safety evaluations revealed no adverse effects or worsening of symptoms, but subtle changes may have been undetectable within the experiment's timeframe. Additionally, no correlation was found between power changes and MDS-UPDRS scores, which were taken from routine assessments months before the study. As higher MDS-UPDRS scores reflect more severe or advanced Parkinson's disease, the lack of correlation suggests that TUS effects may be independent of overall disease severity. Future investigations should directly measure clinical and behavioral outcomes immediately before and after TUS applications to clarify these effects.”

Page 20: '(1) There was no significant difference between baseline powers for the tbTUS and 10Hz TUS compared to the two sham conditions, and the sham condition was always tested before the active TUS conditions' – the sham condition always being tested first would support the suggestion that effects may be related to wearing off of levodopa effects not argue against it.

We would like to clarify our reasoning regarding the lack of baseline power differences and its implications, as well as why we believe that levodopa wear-off does not explain our findings.

The absence of significant differences in baseline power between the sham and TUS conditions suggests that any systematic effect of levodopa wear-off was minimal or did not vary substantially between conditions. This is further supported by the fact that baseline LFP power recordings were conducted approximately one hour apart for sham and TUS conditions, reducing the likelihood of differential wear-off effects.

While we acknowledge that testing the sham condition first could have influenced the observed effects, the differences in LFP power during active TUS conditions (tbTUS and 10 Hz) were protocol-specific. These included distinct frequency band modulations (e.g., beta-band enhancement with 10 Hz and theta-band modulation with tbTUS) and differing durations of effects (e.g., long-lasting effects for 10 Hz versus effects only during sonication for tbTUS). Such protocol-specific effects are unlikely to result from non-specific levodopa wear-off, which would uniformly affect all conditions rather than exhibiting targeted frequency-specific and duration-dependent changes. Moreover, we observed a similar pattern in the dystonia patient, who was not on dopaminergic medication (**Figure S3**), further supporting the conclusion that these modulations are not exclusively driven by drug effects.

How often is levodopa usually taken? How long is the study period compared to the usual time between doses?

Levodopa is typically administered 3 to 4 times per day, with dosing intervals generally ranging from 4 to 6 hours, depending on the patient's symptom severity, response to the medication, and the formulation used (immediate-release vs. extended-release). For our group of patients, this dosing schedule was consistent (4-6 hours). To ensure stable drug levels during our experiments, we minimized the study duration to a maximum of 2.5 hours and scheduled sessions to begin immediately after the patients took their medication. This approach is recommended to help reduce fluctuations in drug efficacy and ensure consistency throughout the on-medication studies.

Page 23/24: 'Both protocols were administered bilaterally (left side first)' – how was this done? How long did it take to reposition the transducer on the other side? What was the time difference?

The transducer was first positioned on the left side using a neuronavigation system to ensure accurate alignment with the planned target (GPi). After completing the stimulation protocol

on the left side, the transducer was repositioned to the right side using the same neuronavigation procedure.

Repositioning the transducer, including the neuronavigation steps to ensure accurate targeting, typically took approximately 2 minutes.

This detail has been added to the revised manuscript (page 25) to ensure clarity in the protocol description.

“Both protocols were administered bilaterally with an intensity of 30 W/cm² ISPPA (measured in water). The transducer was first positioned on the left side using a neuronavigation system to ensure accurate alignment with the planned target (GPi). After completing the stimulation protocol on the left side, the transducer was repositioned to the right side using the same neuronavigation procedure. Repositioning the transducer, including the neuronavigation steps to ensure accurate targeting, typically took approximately 2 minutes.”

‘To minimize potential auditory confound during TUS67, white noise was played to participants during sonications through earbuds.’ – did the participants report that they could hear it or distinguish between active sham, passive sham and sonication?

We added the following information to the method section (page 25):

“None of the patients were able to distinguish between the sham and real conditions, nor did they report hearing the TUS during the experiments. To minimize biases and placebo effects, we ensured that patients were blinded by providing them with only minimal information about the procedure and steps involved in the study.”

Page 26: ‘Then, a thin layer of ultrasound gel (Wavelength MP Blue, Sabel Med, Oldsmar, FL) was applied to both the participant's head and the transducer’ – was the hair not saturated with gel? How did you ensure this resulted in good coupling?

We thank the reviewer for raising this important point, as effective acoustic transmission is critical for successful ultrasound stimulation.

To ensure optimal coupling, we covered the target scalp area with ultrasound gel, paying particular attention to regions with hair. In areas where hair density could impede acoustic transmission, the hair was carefully parted, and additional gel was applied to create a continuous and effective coupling layer. Given that the sonication targeted the frontotemporal region, excessive gel application risked the gel sliding off the head. Therefore, we used an adequate amount of gel to fully cover the area while preventing it from slipping, ensuring effective coupling throughout the procedure.

These details have been added to the revised manuscript (Methods, page 27) to clarify the steps taken to ensure reliable coupling.

“Then, adequate ultrasound gel (Wavelength MP Blue, Sabel Med, Oldsmar, FL) was applied to both the participant's head and the transducer, ensuring the absence of air bubbles between the transducer and the scalp, and sonication was delivered for two minutes. In areas where hair density could impede acoustic transmission, the hair was carefully parted, and additional gel was applied to create a continuous and effective coupling layer.”

‘These coordinates were then fine-tuned through direct anatomical targeting’ – what does this mean? Please explain more clearly – you made sure that the coordinates matched what you could see on the participants MRI?

Yes, it refers to the process of precisely aligning the initial coordinates from the neuronavigation system with the participant’s individual MRI anatomy, ensuring accurate targeting of the intended brain region.

The revised text now reads, (Methods, page 28):
“These coordinates were then fine-tuned to ensure precise alignment with the target region based on the participant's individual MRI anatomy. The fine-tuned coordinates were then transferred to the neuronavigation software (Brainsight, Rogue Research, Canada).”

Page 27: 1st paragraph – please report the acoustic and thermal properties used in the simulations and any other settings or parameters used to run simulations – the paper referenced describes the modelling software but not the specific values you used. These could be detailed in supplementary material – see the ITRUSST standardized reporting paper for a guide to what to report when using simulations to make in situ exposure estimates. Please also report the in situ focal dimensions (it would also be useful to know the size of the target brain regions).

Acoustic parameters are used as described in Table I of Pichardo., 2023. The paper, however, does not report the thermal modelling parameters. While thermal parameters are consultable in the open-source repository, we acknowledge reporting it in the manuscript makes more sense. We added the following table:

“Table S3. Physical parameters used for acoustic and thermal modelling in BabelBrain simulations. $\rho, c_L, c_S, \alpha_L, \alpha_S, K, H$ and ω are, respectively, the density, longitudinal speed of sound, shear speed of sound, longitudinal attenuation, shear attenuation, thermal conductivity, specific heat and perfusion. Acoustic properties ρ, c_L, c_S, α_L and α_S are

calculated from Table I in Pichardo, 2023. Thermal properties K , H and ω are as reported in the IT'IS Foundation database V4.1²²."

Property Material	ρ (kg/m ³)	c_L (m/s)	c_S (m/s)	α_L (Np/m)	α_S (Np/m)	k (W/m/°C)	H (J/kg/°C)	ω (ml/min/kg)
Water	1000	1500	-	-	-	0.6	4178	-
Skin	1116	1537	-	2.3	-	0.37	3391	106
Brain	1041	1562	-	3.45	-	0.51	3630	559
Cortical bone	1896.5	2475	1542	81.5	114.5	0.32	1313	10
Trabecular bone	1738.0	2205	1312	81	114.5	0.31	2274	30

We measured the focal spot dimensions in situ and we can read now in the text (page 29):

"The -6dB length and diameter of the elliptical focal spot in situ were 24.33±3.63 mm (mean ± std) and 4.3±0.5 mm (mean ± std), respectively."

In addition, we added the following info to the text (page 28):

"**Table S3** summarizes physical parameters used for acoustic and thermal modelling. For thermal simulations, the 10 Hz TUS protocol included 40 seconds of sonication with a 30% duty cycle, while the tbTUS protocol included 120 seconds with a 10% duty cycle as input parameters. It is worth noting that, because this was a single exposure, the ISPTA is just the product of ISPPA and DC."

Globus Pallidus Internus (GPi) size: **541.4 ± 81.9** mm³ and Pulvinar size: **1600 ± 260** mm³; These dimensions are approximate and can differ based on factors such as individual anatomy, sex, imaging resolution, and delineation methods.

References

1. Sarica, C. *et al.* Trends and disparities in deep brain stimulation utilization in the United States: a Nationwide Inpatient Sample analysis from 1993 to 2017. *Lancet Reg Health Am* **26**, 100599 (2023).
2. Li, P. *et al.* Hemispheric asymmetry in the human brain and in Parkinson's disease is linked to divergent epigenetic patterns in neurons. *Genome Biology* **21**, 61 (2020).
3. Strelow, J. N. *et al.* Local Field Potential-Guided Contact Selection Using Chronically Implanted Sensing Devices for Deep Brain Stimulation in Parkinson's Disease. *Brain Sciences* **12**, 1726 (2022).
4. Johnson, K. A. *et al.* Globus pallidus internus deep brain stimulation evokes resonant neural activity in Parkinson's disease. *Brain Communications* **5**, fcad025 (2023).
5. Trenado, C. *et al.* Local field potential oscillations of the globus pallidus in cervical and tardive dystonia. *J Neurol Sci* **366**, 68–73 (2016).
6. Swinnen, B. E. K. S. *et al.* Clinically Implemented Sensing-based Initial Programming of Deep Brain Stimulation for Parkinson's Disease: A Retrospective Study. *Neuromodulation* **0**, (2024).
7. Grippe, T. *et al.* Plasticity-Induced Effects of Theta Burst Transcranial Ultrasound Stimulation in Parkinson's Disease. *Mov Disord* (2024) doi:10.1002/mds.29836.
8. Samuel, N. *et al.* Accelerated Transcranial Ultrasound Neuromodulation in Parkinson's Disease: A Pilot Study. *Mov Disord* **38**, 2209–2216 (2023).
9. Shamli Oghli, Y. *et al.* Mechanisms of theta burst transcranial ultrasound induced plasticity in the human motor cortex. *Brain Stimul* **16**, 1135–1143 (2023).
10. Zeng, K. *et al.* Effects of different sonication parameters of theta burst transcranial ultrasound stimulation on human motor cortex. *Brain Stimul* **17**, 258–268 (2024).

11. Zeng, K. *et al.* Induction of Human Motor Cortex Plasticity by Theta Burst Transcranial Ultrasound Stimulation. *Ann Neurol* **91**, 238–252 (2022).
12. Wu, S.-Y. *et al.* Efficient Blood-Brain Barrier Opening in Primates with Neuronavigation-Guided Ultrasound and Real-Time Acoustic Mapping. *Sci Rep* **8**, 7978 (2018).
13. Chaplin, V. *et al.* On the accuracy of optically tracked transducers for image-guided transcranial ultrasound. *Int J Comput Assist Radiol Surg* **14**, 1317–1327 (2019).
14. Xu, L. *et al.* Characterization of the Targeting Accuracy of a Neuronavigation-Guided Transcranial FUS System In Vitro, In Vivo, and In Silico. *IEEE Trans Biomed Eng* **70**, 1528–1538 (2023).
15. Riis, T. S., Losser, A. J., Kassavetis, P., Moretti, P. & Kubanek, J. Noninvasive modulation of essential tremor with focused ultrasonic waves. *J Neural Eng* **21**, (2024).
16. Kusunose, J. *et al.* Design and validation of a patient-specific stereotactic frame for transcranial ultrasound therapy. *IEEE Trans Ultrason Ferroelectr Freq Control* **PP**, (2024).
17. Telkes, I. *et al.* Local field potentials of subthalamic nucleus contain electrophysiological footprints of motor subtypes of Parkinson’s disease. *Proceedings of the National Academy of Sciences* **115**, E8567–E8576 (2018).
18. Bocci, T. *et al.* Asymmetries of the subthalamic activity in Parkinson’s disease: phase-amplitude coupling among local field potentials. *Brain Communications* **6**, fcae201 (2024).
19. Liu, X. *et al.* Subthalamic nucleus input-output dynamics are correlated with Parkinson’s burden and treatment efficacy. *npj Parkinsons Dis.* **10**, 1–15 (2024).
20. Moran, A., Bergman, H., Israel, Z. & Bar-Gad, I. Subthalamic nucleus functional organization revealed by parkinsonian neuronal oscillations and synchrony. *Brain* **131**, 3395–3409 (2008).

21. Silberstein, P. *et al.* Patterning of globus pallidus local field potentials differs between Parkinson's disease and dystonia. *Brain* **126**, 2597–2608 (2003).
22. Baumgartner, C *et al.* IT'IS Database for Thermal and Electromagnetic Parameters of Biological Tissues, Version 4.2. IT'IS Foundation (2024).

Reviewer #1 (Remarks to the Author):

My queries have been answered and I am happy the manuscript is now suitable for publication.

Thank you very much for your positive assessment and support for publication.

Reviewer #3 (Remarks to the Author):

The comments provided by the authors to respond to the multiple reviewers are very helpful in clarifying the study design and initial, highly preliminary results that the study provides. The authors clearly indicate that this pilot study was significantly abridged by a number of external conditions and overall weak study design, and try to rationalize the existing study in many ways (ie, rather than implement a more robust study design and perform additional experiments), reinforcing the tentative nature of the conclusions:

- 1) the lack of patients off meds + off DBS and the varying duration of on med condition, significantly reducing the apparent beta LFP power and thus preventing a more clear treatment effect;
- 2) performing bilateral GP stimulation vs unilateral and hence having a clear control side that would be less effected (the argument comparing GP homotopic interconnections vs M1 is weak);
- 3) even though this is "focused" the beam and frameless navigation suggest (from the LFP results) that all contacts on the DBS electrode were likely targeted equally so no local control (ie, comparing contacts on the same electrode);
- 4) the effective exposure applied is equal to that commonly used for sonication for BBB opening, and this effect could be an equivalent response - ie, merely BBB opening with local brain irritation and not necessarily a "physiological" neuromodulation, particularly due to the long duration of effects suggested;
- 5) presenting two separate, disparate studies together.

However, the authors have responded well to the multiple reviewers' comments, clearly improving their explanation of the overall weak study design and results and the manuscript. The worry is that these highly preliminary, pilot data (from an excellent group) may be interpreted as stronger data than is actually provided supporting neuromodulation rather than an unknown effect, due to the severe constraints of the study design and results.

We appreciate the reviewers' thoughtful comments and have done our best to address their comments. However, the additional experiments the reviewer proposed were not feasible for the following reasons:

I. Direct MRI confirmation using microbubbles and BBB opening

Opening the BBB with microbubbles requires specialized major equipment. We do not have access to the required Insightec 220 kHz apparatus at our hospital so cannot perform the experiments.

II. DBS off–Drug off condition

While this is an important and interesting area of study, this was not part of the original study design and we do not have institutional ethics approval. In addition, the necessary recruitment and experiments would require an entirely new investigation, potentially taking 1–2 years, which is beyond the scope of the current revision.

Regarding point 3

We performed the suggested analyses between DBS lead contacts, but the results were negative as reported in the revised manuscript. This outcome reflects the inherent nature of the technique, which we have discussed in detail in both the manuscript and the rebuttal letter.

Regarding point 4

The hypothesized cavitation effects and BBB opening are highly unlikely. Our reported mechanical index (MI)—a key measure for cavitation risk—remains well below recommended thresholds, and we did not use any microbubbles, which are essential for BBB opening to occur.

Regarding points 2 and 5

We acknowledge these suggestions as minor design variations reflecting personal preference. Although many alternatives exist, our choices were constrained by practical limitations and reflect the best course given our capacity.

At the end, we would like to thank the reviewer for the several constructive comments they provided and the time they devoted to helping us refine our manuscript.

Reviewer #4 (Remarks to the Author):

The authors have attempted to address my concerns, and the detailed analysis and responses to my questions are much appreciated.

Thank you very much for your positive feedback.

However, some issues continue to remain. The symptom specific analysis is appreciated, however, there is one concern with this analysis: the graph in Figure S5 shows that the significance is for the 10 Hz TUS condition compared to the passive sham condition (not the active sham) – a clarification on the misinterpretation or mislabeling of the graph would be important to draw correct conclusions.

The significant differences we found were between the 10 Hz TUS and Passive Sham conditions, but this was mistakenly referred to as 'Active Sham' in the text. We apologize for this oversight. In the revised manuscript, this error has been corrected.

Could the authors also specify whether the significance is based on a two-sample t-test (or in other words an independent sample t-test) or a one sample t-test? This would be important for interpretation. Given the low sample size, it would be very beneficial to see individual data points rather than bar plots to understand variance contributed by each group.

We used a two-sample t-test to assess whether there were significant differences between the two groups of patients. In the revised manuscript, we have added this information and included individual data points (Figure S5).

Finally, while it may be correct that the second experiment does not serve as a control experiment, the fact that the second aspect of the target engagement is being performed on a group of healthy population would result in interpretations that cannot be extrapolated to a population of individuals with motor disorders. The underlying point I am trying to raise here is that the experimental design has not been sufficiently justified to draw a common conclusion. This could be added into the limitation section.

We agree with the reviewer and revised the following paragraph to acknowledge this limitation:

"Altogether, these results provide evidence of the target specificity of TUS neuromodulation. The ability to efficiently modulate deep brain areas not only has therapeutic implications but could also enable the exploration of the causal roles of deep brain regions in sensory, cognitive, and motor functions, which has not been possible with other non-invasive brain stimulation techniques. However, we acknowledge that the behavioral findings from this experiment in healthy participants may not necessarily generalize to individuals with movement disorders."

Reviewer #5 (Remarks to the Author):

Thank you for addressing my comments, I have nothing further to raise.

Thank you very much for your positive assessment.

Reviewer #6 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Thank you very much for your support.