scientific reports

Check for updates

OPEN

# Focused ultrasound thalamotomy improves voice tremor in essential tremor: objective insight from artificial intelligence

Francesca Pistoia[1,2], Patrizia Sucapane[2,7], Francesco Asci[3,7], Gennaro Saporito[1], Giovanni Costantini[4], Valerio Cesarini[4], Giovanni Saggio[4], Simona Sacco[1] & Antonio Suppa[5,6]✉

**The impact of Magnetic Resonance Imaging-guided Focused Ultrasound (MRgFUS) on vocal tremor is largely underexplored. We used artificial intelligence to investigate changes in vocal tremor in ET patients undergoing MRgFUS. Eighty-three controls and ET patients (mean age-SEM 70.0–1.0 years), including 39 patients without and 44 patients with clinically overt voice tremor, were assessed before and 24 h after MRgFUS targeting the ventral intermediate nucleus (Vim). Voice recordings were collected to calculate the receiver operating characteristic (ROC) curves, the likelihood ratios (LRs) for clinical-instrumental correlations, and the oscillatory activity peak by means of artificial intelligence algorithms. We found that before and after Vim-MRgFUS, artificial intelligence discriminated voices in ET and controls objectively with high accuracy. Also, we verified that Vim-MRgFUS improved voice tremor in ET. Moreover, the analysis showed positive correlations between LRs and clinical scores of voice tremor. Lastly, Vim-MRgFUS reduced the 4–6 Hz oscillatory activity peak in ET patients. Therefore, our observations pave the way for objective voice evaluations in ET patients following Vim-MRgFUS, for telemedicine purposes.**

**Keywords** Essential Tremor, Magnetic Resonance Imaging-guided Focused Ultrasound, Voice Tremor, Tremor Frequency, Artificial Intelligence

Magnetic Resonant-guided Focused UltraSound (MRgFUS), targeting the thalamic ventralis intermedius nucleus (Vim), has been recently approved by the U.S. Food and Drug Administration (FDA) for the surgical treatment of medically resistant upper limb tremor in ET[1,2]. MRgFUS is a non-invasive, highly technological technique able to deliver focused beams of ultrasound energy to specific deep brain structures, leading to coagulative necrosis with focal irreversible neuronal death[3–5]. Although the Vim-MRgFUS can improve upper limb tremors effectively, the impact of this procedure on tremors affecting other body locations, including vocal tremors, has only been partially investigated[6–8]. Preliminary studies have shown contrasting results, with some research indicating that Vim-MRgFUS may improve vocal tremor[6,7,9], while others showed voice worsening following the procedure[10,11]. Gaining advances in the field and clarifying Vim-MRgFUS-related voice changes in ET would require innovative technological approaches.

Over the last years, the analysis of voice impairments in neurologic disorders through artificial intelligence (AI) has been identified as a promising non-invasive and reliable instrumental tool for supporting clinicians in recognizing the disease, evaluating disease severity and monitoring therapeutic response[12–17]. In a previous observation, we objectively and automatically detected voice impairments in a large cohort of ET patients, with and without a clinically overt voice tremor, through AI, reporting a prominent oscillatory peak activity at 4–6 Hz band[16]. We also objectively detected the symptomatic effect of pharmacologic treatment on voice impairment in ET, also showing a significant reduction of the oscillatory peak activity in treated patients[16]. Hence, owing to the

[1]Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy. [2]Department of Neurology, San Salvatore Hospital, L'Aquila, Italy. [3]Neurology Unit, Department of Neurosciences and Sensory Organs, AO San Giovanni – Addolorata, Rome, Italy. [4]Department of Electronic Engineering, University of Rome Tor Vergata, Rome, Italy. [5]Department of Human Neurosciences, Sapienza University of Rome, Viale dell'Università, 30, 00185 Rome, Italy. [6]IRCCS Neuromed Institute, Pozzilli, IS, Italy. [7]These authors contributed equally: Patrizia Sucapane and Francesco Asci. ✉email: antonio.suppa@uniroma1.it

high complexity and multidimensionality of the human voice, a machine-learning- (ML)-based approach, able to combine thousands of changing features dynamically, would be the ideal tool to objectively disclose possible benefits and side effects on voice in patients with ET following Vim-MRgFUS.

To the best of our knowledge, no study has yet investigated the impact of Vim-MRgFUS on voice production in a large cohort of ET patients using ML-based voice analysis. Also, none has assessed clinical-instrumental correlations in ET patients following MRgFUS to clarify the clinical impact of neuroablative procedures on voice tremor.

This study aimed to investigate voice tremor changes in a large, consecutive cohort of ET patients undergoing MRgFUS, using spectral analysis and ML-based methods. The support vector machine (SVM), a specific machine-learning tool, was used for calculating the receiver operating characteristic (ROC) curves, as diagnostic outputs for all the comparisons included in the analysis, thus providing statistical measures of sensitivity, specificity, positive and negative predictive values, accuracy and area under the curve (AUC). Also, we assessed clinical-instrumental correlations by calculating for each patient the likelihood ratios (LRs), a machine-learning measure of voice impairment severity. Lastly, we used a spectral analysis for extracting and comparing the oscillatory activity peak (i.e., the dominant tremor frequency index) in ET before and 24 h after MRgFUS.

## Results

### Participants

Demographic and anthropometric parameters (i.e., age, weight, height and BMI) were normally distributed in HS and ET, with and without voice tremor ($p < 0.05$) and were comparable between HS and ET as well as between $ET_{VT+}$ and $ET_{VT-}$ ($p > 0.05$). The chi-square test showed a higher prevalence of males than females in all cohorts of patients: ET (75.9%), $ET_{VT-}$ (82.0%) and $ET_{VT+}$ (70.4%)($p > 0.05$), in line with previous research[18]. Our cohort of ET patients showed a clinically overt voice tremor in 53% of cases. Clinical scores (i.e., MMSE, FAB and BDI) were comparable among HS and ET ($p > 0.05$ for all comparisons). At baseline, $ET_{VT+}$ patients showed higher scores on the FTM ($U = 532.5$; $z = 2.71$; $p < 0.01$), FTM-v ($U = 0.0$; $z = 7.79$; $p < 0.01$), and VHI ($U = 390.50$; $z = 4.05$; $p < 0.01$) scales than $ET_{VT-}$ patients (Table 1). Lastly, patients with $ET_{VT-}$ and $ET_{VT+}$ undergoing right or left Vim surgery showed comparable clinical scores (FTM), before treatment ($p > 0.05$).

### MRgFUS and voice tremor: clinically-based assessment

Vim-MRgFUS induced a significant clinical improvement in overall tremor in ET, as demonstrated by reduced FTM scores ($ET_{before}$: 47.2–1.4; $ET_{after}$: 26.0–1.0; z=−7.0; W = 249; $p < 0.01$). Similarly, FTM scores significantly decreased in $ET_{VT+}$ ($ET_{VT+before}$: 51.2–1.8; $ET_{VT+after}$: 29.3–1.7; z=−4.7; W = 132; $p < 0.01$) and $ET_{VT-}$ ($ET_{VT-before}$: 42.9–2.1; $ET_{VT-after}$: 24.0-1.1.0.1; z=−5.4; W = 0; $p < 0.01$), following surgical procedures, showing a comparable magnitude in $ET_{VT-}$ and $ET_{VT+}$ undergoing right or left hemispheric side procedure. Also, as reported by the FTM-v, Vim-MRgFUS significantly improved voice tremor in $ET_{VT+}$ (patients with a clinically overt voice tremor) after surgical procedures when compared to before ($ET_{VT+before}$: 1.5 – 0.1; $ET_{VT+after}$: 1.2 – 0.1; z=−2.9; W = 79; $p < 0.01$), showing no difference with respect to hemispheric side surgery. Interestingly, voice tremor was also significantly reduced in the overall group of ET patients, following MRgFUS, as suggested by the decrease in the FTM-v score ($ET_{before}$: 0.8 – 0.1; $ET_{after}$: 0.5 – 0.1; z=−3.3; W = 81; $p < 0.01$). Lastly, when considering results from self-assessed voice evaluation, only a trend of VHI decrease following MRgFUS was achieved, without reaching statistical significance.

### MRgFUS and voice tremor: ML-based assessment

Before MRgFUS, artificial intelligence achieved high diagnostic performance when comparing voices in HS and ET patients, showing sensitivity = 98.8%, specificity = 98.7%, PPV = 98.8%, NPV = 98.8%, accuracy = 98.2%, and AUC = 0.998, when applying an optimal diagnostic threshold value of 0.33 and 10-folds cross-validation (Youden index = 0.99) (Fig. 1A; Table 2).

Also, ML-based procedures achieved high diagnostic performance when comparing voice recordings in ET patients before and after MRgFUS: associated criterion = 0.19, Youden index = 0.72, sensitivity = 78.5%, specificity = 79.7%, PPV = 79.2%, NPV = 81.6%, accuracy = 80.4%, and AUC = 0.865 (Fig. 1B; Table 2). The diagnostic accuracy became even higher when we compared voice recordings $ET_{VT+}$ and $ET_{VT-}$ before and after MRgFUS. Indeed, when voice recordings were compared in $ET_{VT+}$ before vs. after MRgFUS, we achieved high diagnostic performances, as shown by sensitivity = 95.5%, specificity = 93.3%, PPV = 95.8%, NPV = 94.2%, accuracy = 94.8%, and AUC = 0.955, using an optimal diagnostic threshold value of 0.82 and 10-folds cross-validation (Youden index = 0.92) (Fig. 1C; Table 2). Similarly, significantly high diagnostic accuracy was reported for $ET_{VT-}$ before vs. after MRgFUS comparison, as shown by sensitivity = 90.0%, specificity = 96.0%, PPV = 95.5%, NPV = 93.0%, accuracy = 96.7%, and AUC = 0.965 (Table 2), possibly pointing to an effect of MRgFUS on subtle voice alterations in ET other than voice tremor.

A final observation is that although MRgFUS improved voice emission in ET, it was not able to restore a normal pattern of voice in ET as indicated by the significant accuracy of our last comparison between controls and ET after the neuroablative procedure. Indeed, we achieved sensitivity = 70.0%, specificity = 65.0%, PPV = 49.9%, NPV = 54.1%, accuracy = 95.2%, and AUC = 1, when applying discretization and 10-folds cross-validation (Youden index = 1), using an optimal diagnostic threshold value of 0.05 (Fig. 1D; Table 2).

The detailed output of SVM-based voice classification is shown in Table 2.

### Correlation analyses

In the whole group of ET patients, the Spearman test disclosed that FTM positively correlated with VHI and FTM-v both collected before ($r = 0.26$, $p < 0.05$; $r = 0.33$, $p < 0.01$) and after MRgFUS ($r = 0.29$, $p < 0.05$; $r = 0.49$,

| | N° | Age (years) | Weight (kg) | Height (cm) | BMI | DD (years) | MMSE | FAB | BDI | FTM | | | FTM-v | | | VHI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | Before | After | b/a | Before | After | b/a | Before | After | b/a |
| HS | 83 | 68.9–1.3 | 75.5–1.6 | 171.7–1.4 | | - | 29.4–0.4 | 15.0-0.3.0.3 | 1.2–0.6 | - | - | | - | - | | - | - | |
| ET | 83 | 70.0–1.0 | 78.0-1.8.0.8 | 169.3–1.2 | 28.0-1.6.0.6 | 13.3–1.2 | 27.8 – 0.2 | 14.6 – 0.3 | 1.6 – 0.3 | 47.2–1.4 | 26.0–1.0 | **P < 0.01*** | 0.8 – 0.1 | 0.5 – 0.1 | **P < 0.01*** | 2.2–0.7 | 2.2–0.6 | p > 0.05 |
| $ET_{VT-}$ | 39 | 70.6–1.6 | 76.2–1.9 | 171.0-1.4.0.4 | 26.0-0.5.0.5 | 14.4–1.9 | 28.2–0.3 | 14.6 – 0.5 | 1.2–0.3 | 42.9**–2.1 | 24.0-1.1.0.1 | **P < 0.01*** | 0.0**–0.0 | 0.0–0.0 | *p* > 0.05 | 0.6**–0.2 | 0.4 – 0.2 | *p* > 0.05 |
| $ET_{VT+}$ | 44 | 69.5–1.2 | 79.5–2.9 | 167.8–1.9 | 29.8–2.9 | 12.3–1.4 | 27.3–0.4 | 14.6 – 0.4 | 1.9 – 0.4 | 51.2**–1.8 | 29.3–1.7 | **P < 0.01*** | 1.5**–0.1 | 1.2 – 0.1 | **P < 0.01*** | 4.8**–1.5 | 3.8–1.2 | p > 0.05 |

**Table 1**. Demographic and clinical features of healthy subjects (HS) and patients with Essential Tremor (ET), with ($ET_{VT+}$) and without ($ET_{VT-}$) a clinically overt voice tremor. *BDI* Beck Depression Inventory, *BMI* body mass index, *DD* disease duration, *ET* essential tremor, $ET_{VT-}$ ET patients without clinically overt voice tremor, $ET_{VT+}$ ET patients with clinically overt voice tremor, *FAB* Frontal Assessment Battery, *FTM* Fahn Tolosa Marin rating scale for essential tremor, before and after thalamotomy, *FTM-v* Fahn Tolosa Marin rating scale for essential tremor, voice impairment sub-item, before and after thalamotomy, *HS* healthy subjects, *MMSE* Mini-Mental State Evaluation, *VHI* Voice Handicap Index. Results are expressed as average - standard error of the mean (SEM).

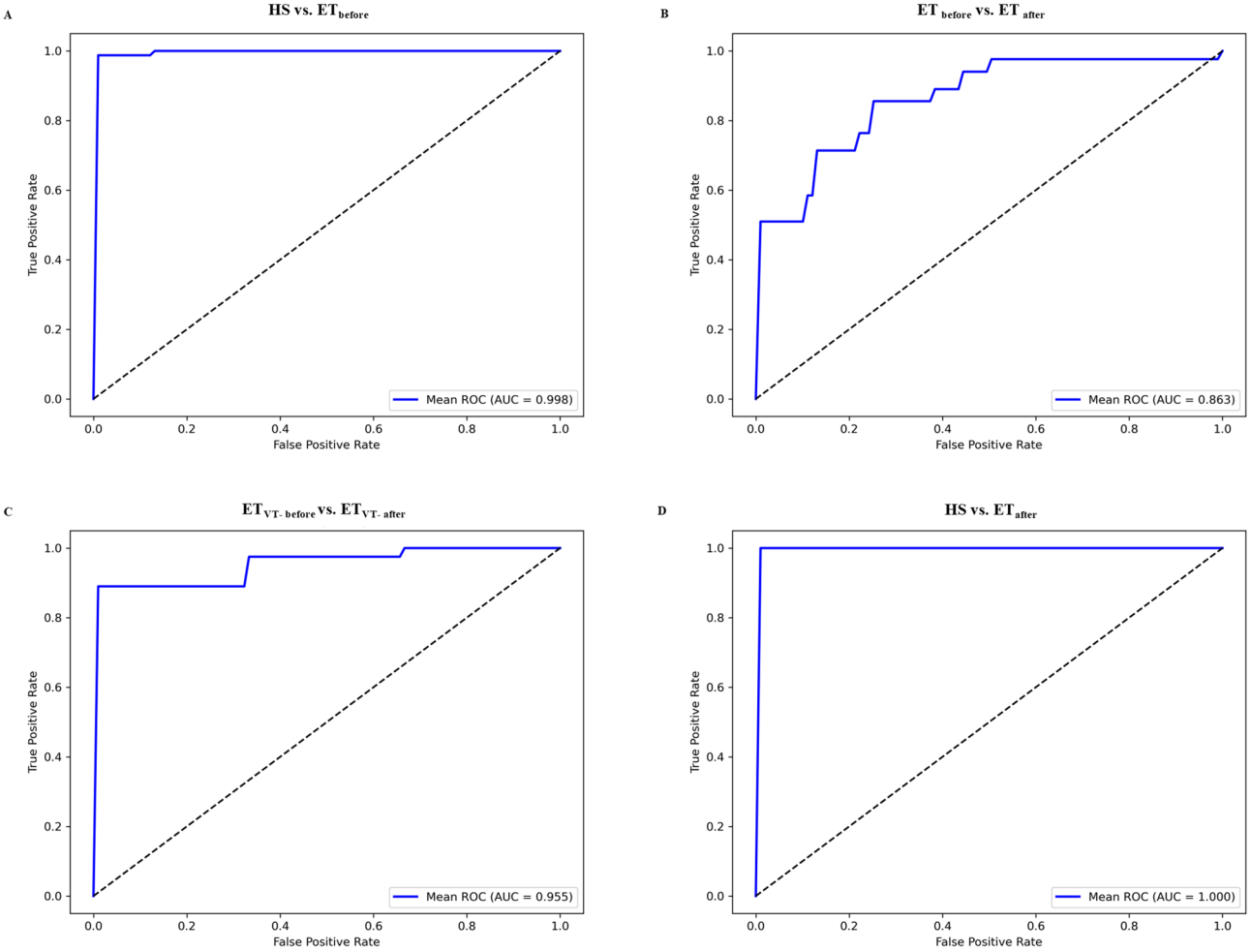


**Fig. 1**. The receiver operating characteristic (ROC) curves analysis. **A** HS vs. $ET_{before}$; **B** $ET_{before}$ vs. $ET_{after}$; **C** $ET_{VT+before}$ vs. $ET_{VT+after}$; **D** HS vs. $ET_{after}$ patients. Note that ROC cureves were calculated through the support vector machine (SVM) classifier in healthy subjects (HS) and patients with essential tremor (ET), with ($ET_{VT+}$) and without ($ET_{VT-}$) a clinically overt voice tremor, before and after MRgFUS. AUC: area under the curve.

| Comparison | Associated criterion | Youden Index | Se (%) | Sp (%) | PPV (%) | NPV (%) | Acc (%) | AUC |
|---|---|---|---|---|---|---|---|---|
| HS vs. $ET_{before}$ | 0.33 | 0.99 | 98.8 | 98.7 | 98.8 | 98.8 | 98.2 | 0.998 |
| $ET_{before}$ vs. $ET_{after}$ | 0.19 | 0.72 | 78.5 | 79.7 | 79.2 | 81.6 | 80.4 | 0.865 |
| $ET_{VT+\ before}$ vs. $ET_{VT+\ after}$ | 0.82 | 0.92 | 95.5 | 93.3 | 95.8 | 94.2 | 94.8 | 0.955 |
| $ET_{VT-\ before}$ vs. $ET_{VT-\ after}$ | 0.37 | 0.91 | 90.0 | 96.0 | 95.5 | 93.0 | 96.7 | 0.965 |
| HS vs. $ET_{after}$ | 0.05 | 1.00 | 70.0 | 65.0 | 49.9 | 54.1 | 95.2 | 1.000 |

**Table 2**. Diagnostic Performance of the Artificial Intelligence-based Voice Analysis. *ET* essential tremor, $ET_{VT+}$ ET patients with clinically overt voice tremor, $ET_{VT-}$ ET patients without clinically overt voice tremor, $ET_{before}$ ET patients before undergoing MRgFUS, $ET_{after}$ ET patients following MRgFUS, *HS* healthy subjects, *Se* sensitivity, *Sp* specificity, *PPV* positive predictive value, *NPV* negative predictive value, *Acc* accuracy, *AUC* area under the curve.

$p < 0.01$). Concerning clinical-instrumental correlations, LRs positively correlated with FTMv in ET patients after MrgFUS ($r = 0.35$, $p < 0.01$).

### MRgFUS and voice tremor: spectral analysis

Spectral analysis showed an oscillatory tremor activity at 4–6 Hz in ET patients before and after thalamotomy. The paired Student t-test showed a significant reduction in the tremor frequency index in ET patients after Vim-MRgFUS than before ($t = 4.17$, $p < 0.01$) (Fig. 2).

## Discussion

In this study, by using a non-invasive artificial intelligence tool, we demonstrated that MRgFUS targeting the Vim effectively improved voice tremor in patients with ET, as confirmed by the 4–6 Hz oscillatory peak reduction. Our analysis also suggested a possible role of Vim-MRgFUS in improving, but not restoring, additional subtle voice abnormalities in ET, possibly not perceived by self-assessed voice analysis. The present findings would pave the way to the objective assessment of voice changes in patients with ET, before and after Vim-MRgFUS.

We applied a rigorous protocol, thus excluding several potential biases. First, patients and controls were comparable in demographic, anthropometric and cognitive features. Also, linguistic and cultural confounding were avoided since all participants were native Italian speakers and the speech task consisted of sustained vowel/e/ emission[13,16,19,20]. Lastly, all voice samples were recorded using a high-definition audio recorder embedded in a commercially available smartphone, to exclude technical artefacts, according to standardized procedures[13].

Our clinically- and ML-based analysis applied to a large dataset of voice features first confirmed our previous report in ET patients[16]. The high diagnostic accuracy achieved in classifying ET and healthy controls pointed to machine learning as a new potentially helpful tool for supporting clinicians in the objective evaluation of voice tremor in ET, in a culture-free and ecological scenario, for telemedicine purposes[16]. The first new finding in this study concerned the objective recognition of symptomatic improvement of voice tremor in patients with ET following Vim-MRgFUS. Clinically, we also found a significant decrease in FTM-v in ET patients, reflecting a notable effect of thalamotomy on voice impairment, in parallel to the expected reduction of upper limb tremor, as shown by decreased scores at FTM[21]. Our findings agreed with previous clinical studies that pointed out a remarkable improvement in voice tremor following MRgFUS thalamotomy in ET[7,22,23].

Spectral analysis first confirmed that voice tremor in ET patients manifested a mean dominant frequency of 4–6 Hz, possibly reflecting an abnormal activation of the physiologic neuronal oscillator underlying the integrated activity of vocal fold vibration and resonant structures, as suggested by our previous observations[16,19,24]. Interestingly, we demonstrated for the first time that Vim-MRgFUS significantly decreased the peak of 4–6 Hz frequency voice intensity in ET patients following the surgical procedure as an objective instrumental measure of the perceptual voice improvement scored by the FTM-v.

Our ML-based analysis objectively showed that Vim-MRgFUS improved voice tremor as demonstrated by the high diagnostic accuracy in distinguishing ET patients before and after thalamotomy. The diagnostic accuracy of our ML-based analysis became even higher when comparing only the subgroup of patients with clinically overt tremor ($ET_{VT+}$) before and after Vim-MRgFUS. To the best of our knowledge, only a few studies in the field have instrumentally assessed the effect of Vim-MRgFUS on voice tremor in ET patients through standardized spectral voice analysis. In the first study[6], the authors observed no significant changes in the frequency and amplitude of voice tremor, following bilateral FUS thalamotomy, thus reflecting a lack of effect of Vim-MRgFUS on speech duration, pause events or duration and oro-motor control during single-syllable repetition[25]. Other

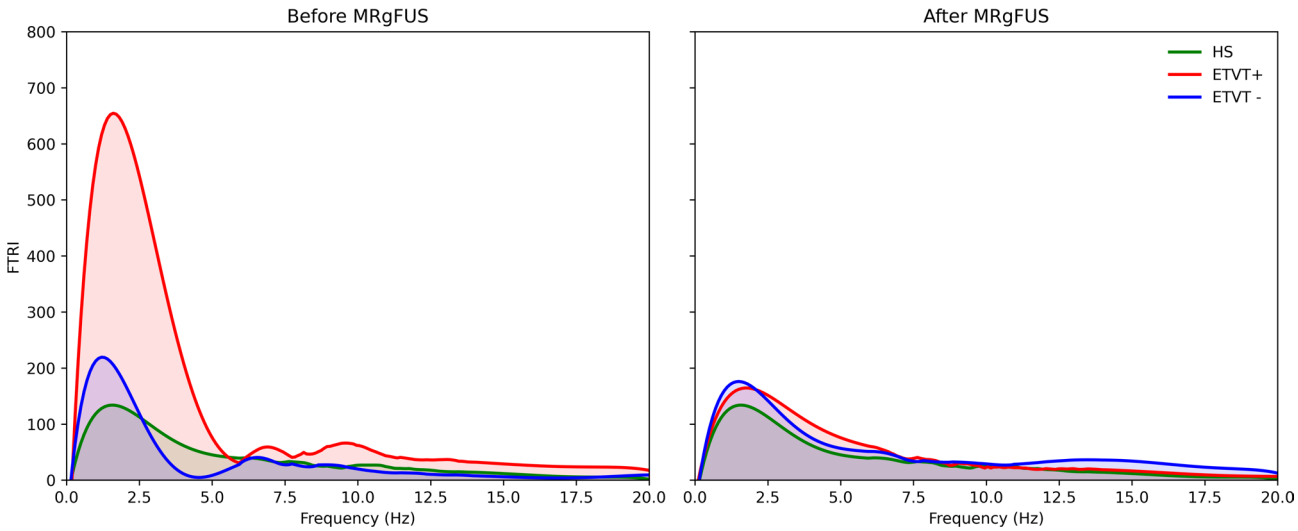


**Fig. 2**. Frequency tremor intensity index (FTRI) in a healthy subject and in two representative ET patients, with ($ET_{VT+}$) and without ($ET_{VT-}$) vocal tremor, before (panel A) and after (panel B) Vim-MRgFUS. Note that the oscillatory activity of voice tremor peaks in the 4–6 Hz frequency band and decreases after MRgFUS.

researchers, however, have reported a significant reduction in the fundamental frequency variability extracted from the sustained emission of the vowel/a/, in ET patients following unilateral Vim-MRgFUS[6]. We argued that the contrasting results achieved by spectral analysis would depend on the specific methodology used since the standardized acoustic analysis focuses only on limited vocal parameters that cannot express the complex and multidimensional phenomenon of human phonation. Rather, AI is an ideal tool under this scenario for objective recognition of voice abnormalities in ET. Indeed, when critically reviewing the voice features more relevant for AI-based analysis, we observed that those extracted by the spectral analysis, including the slope, centroid and harmonicity, ranked in the first positions. Since these features are known to reflect tremulous voice activity, we argued that the high accuracy reached by our AI-based methodology in detecting voice improvement in ET following Vim-MRgFUS, largely depended on the automatic recognition of voice tremor. These findings are fully in line with results provided by the standard spectral analysis by means of the tremor frequency index, which showed a significant reduction in the 4–6 Hz oscillatory tremor activity in ET patients after thalamotomy.

Another relevant finding provided by our study was that Vim-MRgFUS did not restore voice in ET, as demonstrated by the high accuracy when comparing voice features in controls and ET following MRgFUS, despite the significant symptomatic improvement. As previously observed following pharmacological treatments[16], Vim-MRgFUS cannot normalize voices in ET. Accordingly, also following Vim-MRgFUS voices can be easily classified and distinguished from those recorded in controls using AI-based analysis. The overall conclusion is that voice changes in ET would include abnormal features partly resistant to pharmacological and neuroablative strategies[16].

It would be rather tough to clarify the neural underpinning of the significant Vim-MRgFUS-induced voice tremor improvement in patients with ET. So far, a large amount of data from neurophysiological and neuroimaging studies have suggested that the pathophysiologic basis of tremor in ET would be related to abnormalities of the cerebello-thalamo-cortical network[26,27]. Furthermore, recent investigations based on functional imaging have reported large-scale changes in the cerebello-thalamo-cortical network following Vim-MRgFUS, which would be related to the significant improvement of tremor in patients with ET after the surgical procedure[28]. In line with this evidence, it is worth noting that the cerebello-thalamo-cortical network represents a crucial circuit involved in the human phonological loop[29–31], that consists of a complex and integrated neural network of cortico-subcortical brain regions, involved in speech planning, programming, and articulation[29–31]. We hypothesized that the MRgFUS targeting the Vim would exert an overall beneficial symptomatic effect on voice tremor in ET owing to a more physiological engagement of the cerebello-thalamo-cortical network as part of the human phonological loop.

When considering the overall reproducibility of our findings in the general population of ET patients undergoing MRgFUS, several limitations should be addressed. In our cohort, the prevalence of a clinically overt voice tremor in ET patients was about 50%, thus significantly outlasting the estimated prevalence of tremors affecting the vocal tract in the previous series (i.e., about 25%)[32]. This would likely reflect the clinical features of ET cohorts requiring access to advanced therapies, including MRgFUS. Similarly, the higher prevalence of males than females in our cohort of ET patients suggests that males are more likely to access advanced therapies, including MRgFUS, than females confirming previous epidemiologic studies[18]. Lastly, in our patients with ET, we cannot exclude that, besides vocal fold oscillations, voice tremor reflected also tremulous components generated by other muscles contributing to phonatory, articulatory and respiratory functions (i.e., laryngeal, tongue and diaphragm muscles, respectively).

In conclusion, this study demonstrated for the first time that artificial intelligence represents a non-invasive reliable tool for detecting a significant improvement in voice tremor in ET following therapeutic MRgFUS targeting the Vim, in an echological scenario, for telemedicine purposes. Voice emission improvement in ET patients following Vim-MRgFUS may result from positive effects on functional connectivity in specific neural networks involved in physiological human phonation, including the cerebello-thalamo-cortical circuit. Future research in the field will further explore and clarify the possible persistent symptomatic effects of MRgFUS outlasting our short-term assessment in patients with ET.

## Methods

### Study design and participants

We recruited 83 patients affected by ET (20 females, 63 males, mean age-SEM 70.0–1.0 years, range 31–86 years) and 83 age- and gender-matched healthy subjects (36 females and 47 males; mean age-SEM 68.9–1.3 years, range 29–84 years). Participants were enrolled over two years at the Department of Applied Clinical Sciences and Biotechnologies, University of L'Aquila, Italy. All subjects gave written informed consent, which was approved by the institutional ethics committee of the University of L'Aquila (n. 01/2021), according to the Declaration of Helsinki. The diagnosis of ET was established according to updated clinical criteria[33–35]. Inclusion criteria were a diagnosis of ET with clinical recommendation for undergoing MRgFUS and the ability to comply with the study protocol. Patients with voice alterations due to conditions other than essential tremor were excluded.

### Clinical assessment

Patients and controls underwent clinical otolaryngologic and phoniatric evaluation to exclude bilateral/unilateral hearing loss, respiratory disorders and additional non-neurologic disorders impacting voice emission[17]. Clinical symptoms and signs of ET were scored using the Fahn Tolosa Marin (FTM) rating scale for essential tremor[32,36,37]. The perceptual (i.e., clinical) evaluation of voice abnormalities in ET was achieved by an independent rater using the specific item for speech evaluation included in the FTM scale (FTM-v), during the overall motor assessment[37]. A standardised patient-based 30-item scale, the Voice Handicap Index (VHI), Italian version, originally developed by speech-language pathologists (SLPs) in patients with laryngeal dysphonia, was also used for the self-assessment of the functional, physical, and emotional aspects of speech production, thus achieving

a further score of voice impairment[38,39]. In all participants, cognitive functions were scored through the Mini-Mental State Evaluation (MMSE)[40] and the Frontal Assessment Battery (FAB)[41], whereas mood was assessed using the Beck's Depression Inventory (BDI) scale[42].

According to the presence of clinically overt voice tremor, the overall cohort of 83 patients was divided into a subgroup of ET patients without a clinically overt voice tremor ($ET_{VT-}$) (39, 7 females, 32 males, mean age-SEM 70.6 ± 1.6 years, range 31–86 years) and a subgroup of ET patients manifesting voice tremor ($ET_{VT+}$) (44, 13 females, 31 males, mean age-SEM 69.5–1.2 years, range 51–86 years)[33]. At baseline, patients with ET were included in the cohort of $ET_{VT+}$ when they scored ≥ 1 at FTM-v, in line with our previous study[16]. In ET patients, the clinical and instrumental assessment of motor and voice abnormalities was performed at baseline (i.e., before the Vim-MRgFUS) and 24 h after the Vim-MRgFUS. ET patients chronically treated with pharmacological compounds commonly used for improving symptoms (i.e., beta-blockers (BB) or benzodiazepines (BZD)[43] underwent the experimental procedures following an appropriate tapering withdrawal period of at least one week. Participant demographic and clinical features are summarized in Table 1.

## MRgFUS treatment

Before MRgFUS treatment, MRI scans were acquired in all patients using a 3 T scanner (MR750w, GE Healthcare), including FLAIR (slice 3–0.3, TR 11000, Freq FOV 24, Phase FoV 0.8), GRE (slice 3–0.3, TR 960, Freq FoV 26, Phase FoV 0.75), SWI (Slice 2 mm, freq FOV 24, phase FOV 0.85) and DWI (slice 3–0.3, TR 10550, freq FOV 26, phase FOV 0.8) sequences on axial planes, T2 (slice 3.0–0.3, TR 7854, freq. FOV 26, phase FOV 0.8) sequences on coronal and axial planes, and a volumetric T1 3D IR FSPGR (BRAVO) sequence (slice 1 mm, TR 8.5, freq FOV 25.6, phase FOV 0.8). Then, unilateral ablations of the Vim (contralateral to the most symptomatic or disabling side) were performed using the Neuro ExAblate machine (NeuroAblate 4000, InSightec Ltd, Israel). More in detail, the right Vim was ablated in 5 $ET_{VT-}$ and in 9 $ET_{VT+}$, whereas the left Vim was considered for procedure in 34 $ET_{VT-}$ and in 35 $ET_{VT+}$. We recorded several procedural parameters, including SDR values and sonication parameters (namely, Skull Area (mm2), number of active elements Target coordinates (mean, mm), Accumulated Thermal Dose (ATD) Area (mm2), Accumulated Thermal Dose (ATD) Temperature (◦C), number of total Sonications (n), mean Power (W), mean Energy (J), number of Target Movements, meanx Temperature (◦C), and Sonication Duration (s)). Twenty-four hours after the Vim-MRgFUS, a post-procedural MRI was acquired to measure the thalamotomy lesion size.

## Voice recordings

The recording session started by asking participants to sit on a chair in the middle of a silent room, in order to mitigate background noise. As voice tasks for our investigation, we used the sustained emission of a close-mid front unrounded vowel/e/for at least 5 s, produced with usual voice intensity, pitch, and quality. We selected this voice task since sustained vowel emission is a reliable language- and culture-free task for detecting voice abnormalities including voice tremor in patients with ET, according to our previous report in the field[16]. The audio tracks of speech tasks were recorded through a smartphone, which was held by the examiner in front of the participants' faces, at about 30 cm from the mouth, during the task recordings. The placement and input gain of the recording device allowed us for eliminating clipping during technical recordings, with graphic alerts on the recording applications notifying eventual clipping to operators. All voice samples collected in this study from controls and patients were recorded by expert neurologists participating in the study, using a smartphone, equipped with a high-definition microphone and a dedicated application. Recordings were achieved by adopting a linear pulse-code modulation (PCM) format (.wav) at a sampling rate of 44.1 kHz, 16-bit depth, without compressions or filtering[13,16].

## Audio processing and machine-learning analysis

Raw audio tracks were preprocessed according to data curation procedures, by means of the REAPER (Rapid Environment for Audio Production, Engineering, and Recording) (version 7.30, Cockos Inc., Rosendale, NY, USA[44], an advanced software of Digital Audio Workstation (DAW). More in detail, the REAPER software was applied to lossless.wav files for a first check of the audio spectral waveforms, for audio segmentation, for peak normalization and lastly to leverage a common maximum, thus avoiding audio clipping. Then, the Izotope RX7 (iZotope, n.d.) toolbox, which provides the fine-tuning of the algorithm, was used for removing the background noise and artefacts from each audio track using spectral subtraction techniques, according to multi-band spectral subtraction techniques[45,46]. Moreover, audio tracks underwent feature extraction through the OpenSMILE toolbox[47], using the INTERSPEECH 2016 feature dataset[48], which consists of 6373 digitalized acoustic features, as a result of a smooth moving average over temporal windows on the audio file. The feature dataset entails the most relevant domains in voice analysis, including energy, frequency, cepstrum and Mel-frequency Cepstral Coefficients (MFCC)[49], RASTA-style filtering[50], loudness measurements, and voicing probability[51].

Then, all the relevant extracted features underwent the feature selection pre-processing, using the Correlation-based Feature Selector (CFS), which is a supervised heuristic based on a maximum relevance/minimum redundancy principle. The optimal subset was selected using a forward Greedy Stepwise search method[52]. This method allowed us to select the optimal subset via progressive enlargement of the cardinality while evaluating the factor of merit defined by the CFS, which is proportional to the correlation of the features in a given subset to the class label, and inversely proportional to the correlation among features. Then, all the relevant features selected by the CSF were ranked by relevance using the Information Gain Attribute Evaluation (IGAE) algorithm, that measures the information gained concerning the class.

After the pre-processing, the audio features underwent classification procedures. The classification was limited to the 20 most relevant features ranked by the CFS[17], and streamlined the data needed for machine-learning purposes, in order to minimize data needed for machine-learning purposes[16]. A list of the first 20

features which represent functionals applied to audio low-level descriptors (LLDs), extracted from the vowel for the comparison between HS vs. $ET_{before}$, $ET_{before}$ vs. $ET_{after}$, $ETVT_{+before}$ vs. $ETVT_{+after}$ and HS vs. $ET_{after}$ is reported in Supplementary Material 1. Given the relatively small dataset here considered, a Support Vector Machine (SVM) with a linear kernel and soft margins was used as a classifier. The SVM classifier is suitable for small datasets and noisy data since it allows for reducing the likelihood of "overfitting". More in detail, we applied the Platt's Sequential Minimal Optimization method, based on iteratively solving analytically small sub-problems of minimization, to perform the hyperparameter optimization and to train SVMs and solve their quadratic programming problem. The SVM was also calibrated using a logistic regressor to convert its score-like output into probabilistic values suitable for producing ROC curves (see also Suppa et al., 2023[14] for further methodological details).

Lastly, we performed a further machine-learning analysis for clinical-instrumental correlation purposes, after achieving feature extraction and selection, in parallel to the SVM classification procedures. We used a feed-forward artificial neural network (ANN) with ReLu activation functions in the hidden layers, using an architecture consisting of a 20-neurons input layer corresponding to the number of features, a 10-neurons hidden layer and a two-neuron output layer for classification employing a softmax function, trained with the aid of an Adam optimizer. The ANN was trained for classification, but only used to compute a continuous numerical value called the "Likelihood Ratio" (LR), used to assess possible correlations between Machine Learning mechanics and clinical indicators, reflecting the degree of voice impairment in each patient with the FTM-v. The LRs were extracted from the last layer as the raw neuron values before softmax, normalized to a unitary peak leading to a vector containing each training example (patient). Lastly, a Spearman correlation test was applied against the FTM-v associated with each patient, providing p-values (see Fig. 3 for further details about the experimental paradigm).

## Spectral analysis

Praat software was used to perform voice spectral analysis in ET patients before and after Vim-MRgFUS, according to standardized procedures[16,24]. In line with our previous research on voice tremor in ET, the tremor intensity frequency was extracted as a measure of the dominant 4–6 Hz frequency band, whereas the tremor intensity index (power), reported as the "frequency tremor intensity index" (FTRI) was calculated as a measure of the amplitude of the oscillatory peak activity, for each patient with ET before and after MRgFUS[16,24].

## Statistical analysis

The normality of the demographic and anthropometric parameters of HS and ET patients (age, height, and weight) was assessed using the Kolmogorov-Smirnov test. The Mann-Whitney U test was used to compare demographic and anthropometric parameters in HS and ET patients, as well as clinical scores (i.e., FTM, FTM-v, VHI, MMSE, FAB and BDI scores) in $ET_{VT+}$ and $ET_{VT-}$, by also considering the hemispheric side lesion (i.e., right or left hemispheric surgery). The Wilcoxon signed-rank test was used to compare FTM, FTM-v, and VHI scores in $ET_{VT+}$ and $ET_{VT-}$ before and after Vim-MRgFUS, considering also the hemispheric side lesion. ROC analyses were performed to identify the optimal diagnostic cut-off values to discriminate between HS and ET before and after Vim-MRgFUS as well as between $ET_{VT+}$ before and after Vim-MRgFUS and $ET_{VT-}$ before and after Vim-MRgFUS, according to standardized procedures[13,16]. In more detail, the cut-off values were calculated as the point of the curves with the highest Youden index (sensitivity + specificity – 1) to maximize the sensitivity and specificity of the diagnostic tests. Therefore we calculated the Youden Index (YI) and its optimal criterion value, the associated criterion (Ass. Crit.). We also applied ROC curve discretization, by means of Maximum Area Under the ROC Curve-based Discretization (MAD), a supervised technique implemented to maximize the classification performance, by selecting optimal cut-points. Lastly, we applied a 10-fold cross-validation for training the SVM in order to exploit the full dataset, by creating a 90 – 10% training-test split. The final accuracy results were averaged across the ten folds. The sensitivity, specificity, positive and negative predictive values as well as accuracy and area under the curve were also calculated. Spearman's rank correlation coefficient was used to assess correlations between clinical scores (FTM, FTM-v, VHI) and LR values before and after MRgFUS. The paired Student t-test was used to compare the frequency tremor intensity in ET before and after Vim-MRgFUS. Unless otherwise stated, all values were presented as mean – standard error of the mean (SEM). A p-value $< 0.05$ was considered statistically significant. Statistical analyses were performed using Statistica version 10 (StatSoft, Inc).

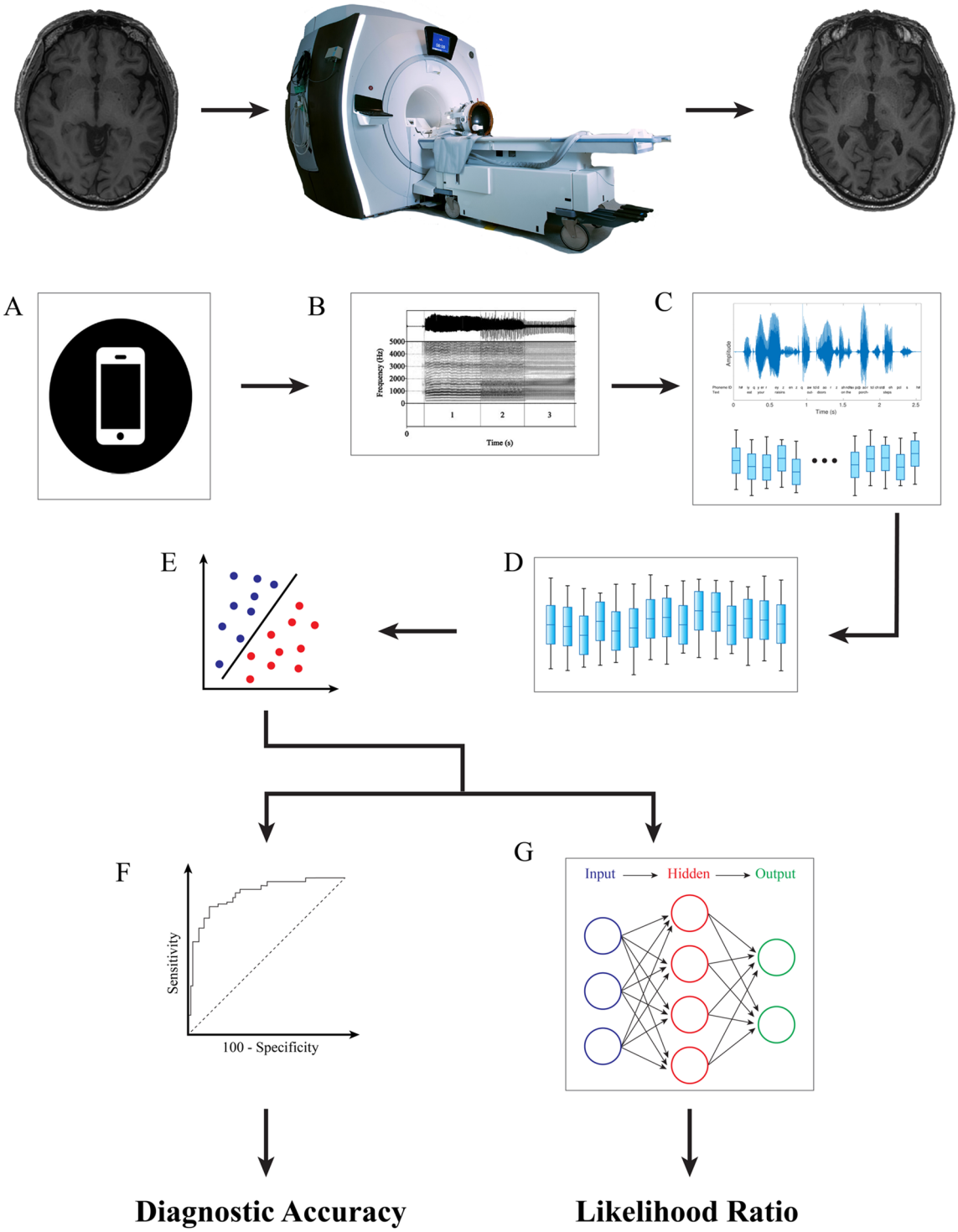


**Fig. 3**. Experimental design. **A** Recording of voice samples through a commercially-available smartphone; **B** narrow-band spectrogram of the acoustic voice signal; **C** feature extraction; **D** feature selection; **E** feature classification; **F** ROC curve analysis; **G** LR values calculated through ANN. Note that all the procedures of artificial intelligence analysis summarized in panels A-G have been performed in ET patients before and after MRgFUS (see the top panel).

## Data availability

The corresponding author has full access to the data used in the analyses in the manuscript. All clinical and instrumental data are stored offline and are available on reasonable request to the corresponding author.

Received: 9 December 2025; Accepted: 23 March 2026
Published online: 04 April 2026

## References


1. Elias, W. J. et al. A randomized trial of focused ultrasound thalamotomy for essential tremor. *N Engl. J. Med.* **375**, 730–739 (2016).
2. Elias, W. J. et al. A pilot study of focused ultrasound thalamotomy for essential tremor. *N. Engl. J. Med.* **369**, 640–648 (2013).
3. Binder, D. K., Shah, B. B. & Elias, W. J. Focused ultrasound and other lesioning in the treatment of tremor. *J. Neurol. Sci.* **435**, 120193 (2022).
4. Kobus, T. & McDannold, N. Update on clinical magnetic resonance–guided focused ultrasound applications. *Magn. Reson. Imaging Clin. N. Am.* **23**, 657–667 (2015).
5. Shetty, N. Essential tremor-Do we have better therapeutics? A review of recent advances and future directions. *Curr. Neurol. Neurosci. Rep.* **22**, 197–208 (2022).
6. Larner, P. et al. Voice improvement after essential tremor treatment via focused ultrasound and deep brain stimulation. *Laryngoscope* **134**, 367–373 (2024).
7. Kaplitt, M. G. et al. Safety and efficacy of staged, bilateral focused ultrasound thalamotomy in essential tremor: An open-label clinical trial. *JAMA Neurol.* **81**, 939 (2024).
8. Pearce, J. J., Thoma, J., Vinson, K. & Sani, S. Treatment of vocal tremor with bilateral magnetic resonance imaging-guided focused ultrasound of the ventral intermediate thalamic nucleus: Illustrative case. *J. Neurosurg. Case Lessons* **5**, CASE2351 (2023).
9. Arcadi, A. et al. Magnetic resonance-guided focused ultrasound (MRgFUS)-thalamotomy for Essential Tremor: Lesion location and clinical outcomes. *Mov. Disord.* **39**, 1015–1025 (2024).
10. Alshaikh, J. & Fishman, P. S. Revisiting bilateral thalamotomy for tremor. *Clin. Neurol. Neurosurg.* **158**, 103–107 (2017).
11. Krack, P., Martinez-Fernandez, R., del Alamo, M. & Obeso, J. A. Current applications and limitations of surgical treatments for movement disorders: Surgical treatments for movement disorders. *Mov. Disord.* **32**, 36–52 (2017).
12. Asci, F. et al. Acoustic analysis in stuttering: A machine-learning study. *Front. Neurol.* **14**, 1169707 (2023).
13. Asci, F. et al. Machine-learning analysis of voice samples recorded through smartphones: The combined effect of ageing and gender. *Sensors* **20**, E5022 (2020).
14. Suppa, A. et al. Effects of deep brain stimulation of the subthalamic nucleus on patients with Parkinson's disease: A machine-learning voice analysis. *Front. Neurol.* **14**, 1267360 (2023).
15. Suppa, A. et al. Voice in Parkinson's Disease: A machine learning study. *Front. Neurol.* **13**, 831428 (2022).
16. Suppa, A. et al. Voice analysis with machine learning: One step closer to an objective diagnosis of Essential Tremor. *Mov. Disord.* https://doi.org/10.1002/mds.28508 (2021).
17. Suppa, A. et al. Voice analysis in adductor spasmodic dysphonia: Objective diagnosis and response to botulinum toxin. *Parkinsonism Relat. Disord.* **73**, 23–30 (2020).
18. Song, P. et al. The global prevalence of essential tremor, with emphasis on age and sex: A meta-analysis. *J. Glob Health.* **11**, 04028 (2021).
19. Hlavnička, J. et al. Characterizing vocal tremor in progressive neurological diseases via automated acoustic analyses. *Clin. Neurophysiol.* **131**, 1155–1165 (2020).
20. Tripoliti, E. Voice tremor and acoustic analysis: Finding harmony through the waves. *Clin. Neurophysiol.* **131**, 1144–1145 (2020).
21. Peters, J. et al. Outcomes of focused ultrasound thalamotomy in tremor syndromes. *Mov. Disord.* https://doi.org/10.1002/mds.29658 (2023).
22. Iacopino, D. G. et al. Preliminary experience with a transcranial magnetic resonance–guided focused ultrasound surgery system integrated with a 1.5-T MRI unit in a series of patients with essential tremor and Parkinson's disease. *NeuroSurg. Focus.* **44**, E7 (2018).
23. Martínez-Fernández, R. et al. Bilateral staged magnetic resonance-guided focused ultrasound thalamotomy for the treatment of essential tremor: a case series study. *J. Neurol. Neurosurg. Psychiatry.* **92**, 927–931 (2021).
24. Brückl, M., Ghio, A. & Viallet, F. Measurement of Tremor in the Voices of Speakers with Parkinson's Disease. *Procedia Comput. Sci.* **128**, 47–54 (2018).
25. Scantlebury, N. et al. Safety of bilateral staged magnetic resonance-guided focused ultrasound thalamotomy for essential tremor. *Mov. Disord. Clin. Pract.* **10**, 1559–1561 (2023).
26. Muthuraman, M., Heute, U., Deuschl, G. & Raethjen, J. The central oscillatory network of essential tremor. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2010**, 154–157 (2010).
27. Raethjen, J. & Deuschl, G. The oscillating central network of essential tremor. *Clin. Neurophysiol.* **123**, 61–64 (2012).
28. Dahmani, L. et al. Focused ultrasound thalamotomy for tremor treatment impacts the cerebello-thalamo-cortical network. *npj Parkinsons Dis* **9**, 90 (2023).
29. Baddeley, A. D. & Hitch, G. J. The phonological loop as a buffer store: An update. *Cortex* **112**, 91–106 (2019).
30. Fuertinger, S., Horwitz, B. & Simonyan, K. The functional connectome of speech control. *PLoS Biol* **13**, e1002209 (2015).
31. Valeriani, D. & Simonyan, K. The dynamic connectome of speech control. *Philos Trans R Soc Lond B Biol Sci* **376**, 20200256 (2021).
32. Louis, E. D. & Gerbin, M. Voice handicap in essential tremor: A comparison with normal controls and Parkinson's disease.. *Tremor Other Hyperkinet Mov (N Y)* https://doi.org/10.7916/D8KD1WN3 (2013).
33. Bhatia, K. P. et al. Consensus Statement on the classification of tremors. from the task force on tremor of the International Parkinson and Movement Disorder Society: IPMDS Task Force on Tremor Consensus Statement. *Mov. Disord.* **33**, 75–87 (2018).
34. Deuschl, G., Bain, P. & Brin, M. Consensus statement of the Movement Disorder Society on tremor. *Mov. Disord.* **13**, 2–23 (2008).
35. Elble, R. J. Diagnostic criteria for essential tremor and differential diagnosis. *Neurology* **54**, S2-6 (2000).
36. Elble, R. et al. Task force report: Scales for screening and evaluating tremor: Critique and recommendations. *Mov. Disord.* **28**, 1793–1800 (2013).
37. Ondo, W. et al. Comparison of the Fahn-Tolosa-Marin Clinical Rating Scale and the Essential Tremor Rating Assessment Scale. *Mov Disord Clin Pract* **5**, 60–65 (2018).
38. Jacobson, B. H. et al. The Voice Handicap Index (VHI): Development and validation. *Am. J. Speech Lang. Pathol.* **6**, 66–70 (1997).
39. Schindler, A. et al. Cross-cultural adaptation and validation of the Voice Handicap Index into Italian. *J Voice* **24**, 708–714 (2010).
40. Folstein, M. F., Folstein, S. E. & McHugh, P. R. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975).
41. Dubois, B., Slachevsky, A., Litvan, I. & Pillon, B. The FAB: A Frontal Assessment Battery at bedside. *Neurology* **55**, 1621–1626 (2000).
42. Beck, A. T. An inventory for measuring depression. *Arch. Gen. Psychiatry* **4**, 561 (1961).
43. Ferreira, J. J. et al. MDS evidence-based review of treatments for essential tremor. *Mov. Disord.* **34**, 950–958 (2019).
44. REAPER | Audio Production Without Limits. https://www.reaper.fm/
45. Er, M. B., Isik, E. & Isik, I. Parkinson's detection based on combined CNN and LSTM using enhanced speech signals with variational mode decomposition. *Biomed. Signal Process. Control* **70**, 103006 (2021).
46. Upadhyay, N. & Karmakar, A. Speech enhancement using spectral subtraction-type algorithms: A comparison and simulation study. *Procedia Comput. Sci.* **54**, 574–584 (2015).

47. Eyben, F. & Schuller, B. openSMILE:): The Munich open-source large-scale multimedia feature extractor. *SIGMultimedia Rec* **6**, 4–13 (2015).
48. Schuller, B. The INTERSPEECH 2016 computational paralinguistics challenge: 17th Annual Conference of the International Speech Communication Association, INTERSPEECH 2016. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* 08-12-September-2016, 2001–2005 (2016).
49. Bogert, B. P. The quefrency alanysis of time series for echoes; Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking. *Time Series Analysis,* 209–243 (1963).
50. Hermansky, H. & Morgan, N. RASTA processing of speech. *IEEE Trans. Speech Audio Process.* **2**, 578–589 (1994).
51. Hess, W. *Pitch and Voicing Determination of Speech with an Extension Toward Music Signals* 181–212 (2008). https://doi.org/10.1007/978-3-540-49127-9_10.
52. Mostafa, S. A. et al. Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease. *Cogn. Syst. Res.* **54**, 90–99 (2019).

## Author contributions

1. Research project: (A) Conception (FP, FA, SS, AS), (B) Organization (FP, PS, FA, GSap), (C) Execution (PS, FA, GSap, VC, GC); 2. Statistical Analysis: (A) Design (FA, VC, GC, AS), (B) Execution (FA, VC, GC), (C) Review and Critique (FP, GSag, AS); 3. Manuscript Preparation: (A) Writing of the first draft (FP, PS, FA), (B) Review and Critique (GSag, SS, AS).

## Funding

The author(s) declared that financial support was received for this work and/or its publication. The publication of this articlewas supported by "Current Research – Ministry of Health to IRCCS I.N.M. Neuromed.

## Declarations

## Competing interests

Authors VC, GC, and GSag are advisory members of VoiceWise S.r.l., spin-off company of University of Rome Tor Vergata (Rome, Italy) developing voice analysis solutions for diagnostic purposes. The other authors declare they received no funding from any institution, including personal relationships, interests, grants, employment, affiliations, patents, inventions, honoraria, consultancies, royalties, stock options/ownership, or expert testimony for the last 12 months.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-45935-x.

**Correspondence** and requests for materials should be addressed to A.S.