Validation of “Laboratory-Supported” Criteria for Functional (Psychogenic) Tremor

Petra Schwingenschuh, MD,1,2* Tabish A. Saifee, MRCP,3 Petra Katschnig-Winter, MD, MSc,1 Antonella Macerollo, MD,3 Mariella Koegl-Wallner, MD,1 Valeriu Culea, MD,1 Christine Ghadery, MD, PhD,1 Edith Hofer, PhD,1,4 Tamara Pendl, MD,1 Stephan Seiler, MD,1 Ulrike Werner, PhD,1 Sebastian Franthal, MD,1 Natasha M. Maurits, PhD,5 Marina A. Tijssen, MD, PhD,5 Reinhold Schmidt, MD,1 John C. Rothwell, PhD,3 Kailash P. Bhatia, MD, FRCP,3 and Mark J. Edwards, FRCP, PhD3,6

1Department of Neurology, Medical University of Graz, Graz, Austria
2Division of Neuroradiology, Department of Radiology, Medical University of Graz, Graz, Austria
3Sobell Department, UCL Institute of Neurology, London, United Kingdom
4Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria
5Department of Neurology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
6Department for Cardiovascular and Cell Sciences, St. George’s University of London, London, United Kingdom

ABSTRACT: Background: In a small group of patients, we have previously shown that a combination of electrophysiological tests was able to distinguish functional (psychogenic) tremor and organic tremor with excellent sensitivity and specificity.

Objectives: This study aims to validate an electrophysiological test battery as a tool to diagnose patients with functional tremor with a “laboratory-supported” level of certainty.

Methods: For this prospective data collection study, we recruited 38 new patients with functional tremor (mean age 37.9 ± 24.5 years; mean disease duration 5.9 ± 9.0 years) and 73 new patients with organic tremor (mean age 55.4 ± 25.4 years; mean disease duration 15.8 ± 17.7 years). Tremor was recorded at rest, posture (with and without loading), action, while performing tapping tasks (1, 3, and 5 Hz), and while performing ballistic movements with the less-affected hand. Electrophysiological tests were performed by raters blinded to the clinical diagnosis. We calculated a sum score for all performed tests (maximum of 10 points) and used a previously suggested cut-off score of 3 points for a diagnosis of laboratory-supported functional tremor.

Results: We demonstrated good interrater reliability and test-retest reliability. Patients with functional tremor had a higher average score on the test battery when compared with patients with organic tremor (3.6 ± 1.4 points vs 1.0 ± 0.8 points; P < .001), and the predefined cut-off score for laboratory-supported functional tremor yielded a test sensitivity of 89.5% and a specificity of 95.9%.

Conclusion: We now propose this test battery as the basis of laboratory-supported criteria for the diagnosis of functional tremor, and we encourage its use in clinical and research practice.

Key Words: Psychogenic; functional; diagnostic criteria; electrophysiological criteria

Introduction

A confident clinical diagnosis of functional (psychogenic) tremor (FT) is often possible, but in some cases uncertainty remains and a laboratory-supported level of certainty could aid early positive diagnosis.1 Various electrophysiological tests have been suggested to identify patients with FT,2,3 including a tonic discharge of antagonist muscles approximately 300 milliseconds before tremor onset,4 an increase of tremor amplitudes in response to weighting a limb,4,5 entrainment or...
increase in variability and change of tremor frequency while tapping with the contralateral hand,\textsuperscript{5-8} less accurate tapping performance at requested frequencies,\textsuperscript{5} significant interlimb coherence in bilateral tremors,\textsuperscript{6} and transient arrest of tremor during ballistic movements of the other hand.\textsuperscript{10}

In a pilot study, we performed a head-to-head comparison of the above-mentioned tests in 13 patients with FT and 25 patients with organic tremors (OT). Test sensitivity and specificity of all separate tests varied between 33\% to 77\% and 84\% to 100\%, respectively. A combination of these electrophysiological tests distinguished FT and OT with excellent sensitivity and specificity.\textsuperscript{11} The aims of the current study were to validate this electrophysiological test battery as a tool to diagnose patients with FT with a laboratory-supported level of certainty and to provide data for its potential use in clinical and research practice.

Methods

Patients

We prospectively recruited 38 consecutive patients with FT and 73 patients with OT who presented with unilateral or bilateral arm tremor and who attended the movement disorders clinics of 2 specialized centers (London, Graz). All patients were different from those taking part in the pilot study.\textsuperscript{11} Patients were clinically classified at time of recording (baseline) and reevaluated after 1-year follow-up according to published criteria by 1 of 4 movement disorder experts onsite who were blinded to the test battery results (MJE, KPB, PS, TAS) (see Fig. 1).\textsuperscript{12,13} FT patients were classified as possible, probable, clinically established, and documented.\textsuperscript{12} History was taken, medical records were reviewed, and a standardized clinical examination including the evaluation of entrainment and tremor suppressability was performed. Patients were determined as having a coactivation, alternating, or a mixed tremor, and a psychiatric diagnosis was made according to the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision.\textsuperscript{14} Finally, the electrophysiological test battery results were unblinded, and in patients with FT, a retrospective diagnosis taking these results into account was reached. Before inclusion in the study, written informed consent was obtained from all participants. This study was approved by the local research ethics committee.

Recordings

Patients were seated in a chair. A triaxial accelerometer transducer (Biometrics ACL300, Sensitivity $\pm$ 100 mV/G, Biometrics Ltd, UK) was attached to the dorsal surface of the middle phalanx of the index fingers, which were chosen to optimally pick up signals during the tapping tasks. Surface electromyography (EMG) was recorded from wrist flexors (WF) and wrist extensors (WE). EMG signals were amplified (Digitimer, Welwyn Garden City, UK) and analogue filtered (low pass at 1000 Hz and high pass with 3-millisecond time constant) and sampled at 2000 Hz.

Recordings were performed with arms relaxed and hands hanging freely from the arm rest, with arms/wrists outstretched at shoulder level without and with a 500-gram mass attached to the wrists (loading) during tapping tasks and while performing ballistic movements. For the tapping tasks, participants were instructed to use the index finger of the less-affected hand to tap in time with a metronome at rates of 1, 3, and 5 Hz while ignoring the more symptomatic arm. For the ballistic movement task, participants were instructed to point with the index finger of an outstretched arm to the examiner’s index finger and follow it as fast as possible when the examiner abruptly changed position. During the tapping and ballistic movement tasks, EMG and accelerometry were recorded continuously from the more-affected contralateral arm in the position where tremor was maximal.
The tremor was recorded and analyzed for 30 seconds in each condition. The recording time was 210 seconds, and the overall time needed for the recordings (including setup and pauses) was approximately 15 minutes.

Data Analysis
Analyses were performed as previously described\(^{11}\) by raters blinded to the clinical diagnosis. The first 10 recordings were analyzed independently by 3 blinded raters to test for interrater variability, and the first 10 patients included in the study were tested twice on different days (within 1 month) to test for test-retest variability. The clinical diagnosis was OT in 6 cases and FT in 4 cases.

A fast Fourier analysis of the signals derived from EMG and accelerometry was performed to define peak tremor frequency and tapping frequency, and the total power (TP) of the spectra between 1 and 30 Hz was used as a surrogate for tremor amplitude. All parameters were calculated for each accelerometer axis and then averaged.

Loading Test
We compared TP (accelerometry; more-affected hand) of a 30-second epoch before and after loading. According to our previous study,\(^{11}\) an increase of TP after loading by more than 130% was considered an abnormal response (1 point in test battery).

Response to Ballistic Movements
EMG recordings were used offline to define the onset of each ballistic movement. Tremor was considered sensitive to the execution of a contralateral ballistic movement (1 point in test battery) if a tremor pause or at least 50% decrease in period or amplitude occurred in 7 or more out of 10 trials.\(^{10,11}\)

Coherence Test
In participants with bilateral hand tremor (30 patients with FT; 52 patients with OT), we assessed coherence between EMG of right and left WE. Significant coherence was defined as 2 contiguous bins on the coherence plot that rose above the 99% confidence limit for random coherence at a frequency where there were corresponding peaks in the power spectra (1 point in test battery).\(^{7,11}\)

Tonic Coactivation
We inspected the WE and WF EMG of the tremor-dominant arm for the presence of tonic coactivation at tremor onset (start of rhythmic oscillations on accelerometry). Tonic coactivation phase was defined as tonic discharge of antagonist muscles approximately 300 milliseconds before the onset of tremor bursts (1 point in test battery).\(^{4,11}\)

Tapping Performance
Correct tapping performance for 1, 3, and 5 Hz was predefined as 0.5–1.5 Hz, 2.5–3.5 Hz, and 4.5–5.5 Hz, respectively.\(^{11}\) Tapping performances beyond these ranges were considered as abnormal (1 point in test battery for each tapping frequency).

Tapping Response
Tremor recordings of the nontapping hand were assessed for entrainment, tremor suppression, or pathological frequency shift. The latter was predefined as a frequency shift of at least 19.0%, 26.9%, and 25.7% during contralateral tapping at 1, 3, and 5 Hz, respectively (1 point in test battery for each tapping frequency).\(^{11}\)

We calculated a sum score for all performed tests (maximum 10 points) and used a previously suggested cut-off score of 3 points for a diagnosis of laboratory-supported FT.\(^{11}\)

Statistics
The sample size was calculated based on confidence intervals of sensitivity and specificity as measures of precision. According to the pilot study, sensitivity and specificity were both estimated to be 90%.\(^{11}\) To obtain a 95% confidence interval for the estimated sensitivity and specificity with a width of ±10%, at least 38 patients with FT and at least 38 patients with OT were required. The sample size calculation was performed using the R statistics software and the binomSamSize package.\(^{15}\) As both sensitivity and specificity were close to 100%, the exact method of Fosgate was chosen.\(^{16}\) The recruitment of OT patients was continued until the minimum number of FT patients was reached.

The interrater reliability between the 3 raters and test-retest reliability for the sum scores between the 2 examinations were determined using the intraclass correlation coefficient (ICC). A 2-way, mixed-model ICC with measures of consistency was computed with IBM SPSS Statistics 21, New York, USA.

Sum scores of OT and FT patients were compared by independent-sample \(t\) tests. We performed a 1-way analysis of variance to compare the sum scores of FT subgroups. To compare abnormal and normal test battery results, group analysis was performed with \(2 \times 2\) cross-tabs. We compared the frequency of abnormal test results by chi-square and Fisher’s exact (2-sided) tests and obtained sensitivity and specificity. \(P\) values below .05 were considered statistically significant.

Results
Demographics
Patients with FT (\(n = 38\)) had a mean age of 37.9 ± 24.5 years and a mean disease duration of 3.9 ± 9.0 years. At baseline, 3 patients were clinically diagnosed as possible,
There was a strong agreement between the 3 different raters regarding the sum score with an ICC of 0.984 \((P < .001)\). The test-retest reliability between the sum scores of the 2 examinations was also very high (ICC = 0.918; \(P < .001\)).

Patients with FT had a higher average score on the test battery when compared with the OT patients \((3.6 \pm 1.4 \text{ points vs } 1.0 \pm 0.8 \text{ points}; P < .001)\).

The predefined cut-off score for a diagnosis of laboratory-supported FT with 3 of 10 points yielded a test sensitivity of 89.5% and a specificity of 95.9% \((P < .001)\).

Changing the cut-off score to 4 of 10 points increased test sensitivity (100%), but decreased its specificity (57.9%; \(P < .001\)).

The mean sum scores in patients with possible, probable, clinically established, and documented FT were \(3.6 \pm 0.6, 3.2 \pm 1.2, 3.4 \pm 1.7, \) and \(4.2 \pm 1.3\), respectively. The sum scores did not differ significantly between FT subgroups \((F = 1.058; P = .380)\).

Of 73 OT patients, 70 had either 0, 1, or 2 points; hence only 3 OT patients scored abnormal with a sum score of 3 each. Of 38 FT patients, 34 had at least 3 points (see Fig. 2). The 4 FT patients who scored normal had a sum score of 0 \((n = 1)\) or 1 point \((n = 3; 1\) patient each received 1 point on the Incorrect Tapping Performance at 5 Hz, Tonic Coactivation, and Increase of Amplitude With Loading subtests [for clinical characteristics, see Table 1]).

The frequency of normal and abnormal results of each test in the FT and OT groups is provided in Table 2.

### TABLE 1. Clinical characteristics of the 4 functional (psychogenic) tremor patients with normal test battery results (sum score < 3 points)

<table>
<thead>
<tr>
<th></th>
<th>Patient 1</th>
<th>Patient 2</th>
<th>Patient 3</th>
<th>Patient 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease duration (years)</td>
<td>10</td>
<td>16</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>Sum score</td>
<td>10</td>
<td>16</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>Abnormal subtests</td>
<td>Tonic coactivation</td>
<td>Incorrect tapping performance at 5 Hz</td>
<td>Increase of amplitude with loading</td>
<td>Cocontraction</td>
</tr>
<tr>
<td>Tremor type</td>
<td>Cocontraction</td>
<td>Conversion disorder</td>
<td>Conversion disorder</td>
<td>Alternating</td>
</tr>
<tr>
<td>Psychiatric diagnoses</td>
<td>Conversion disorder</td>
<td>Probable</td>
<td>Probable</td>
<td>Somatization disorder</td>
</tr>
<tr>
<td>Clinical diagnostic level (baseline)</td>
<td>Sudden tremor onset after stressful life event, static disease course, tremor variability during examination, functional gait disorder</td>
<td>Sudden tremor onset after stressful life event, spontaneous remissions, tremor variability over time, false (give-away) weakness</td>
<td>Major depressive disorder, long history of multiple somatizations, sudden tremor onset, tremor variability during examination, attention increases tremor, effortful production, and deliberate slowness of movement</td>
<td>Generalized anxiety disorder, long history of multiple somatizations, functional disability out of proportion to objective findings, attention increases tremor, fluctuation of tremulous body parts during examination, nonanatomical sensory loss</td>
</tr>
<tr>
<td>Explanation of baseline classification</td>
<td>Sudden tremor onset after stressful life event, static disease course, tremor variability during examination, functional gait disorder</td>
<td>Sudden tremor onset after stressful life event, spontaneous remissions, tremor variability over time, false (give-away) weakness</td>
<td>Major depressive disorder, long history of multiple somatizations, sudden tremor onset, tremor variability during examination, attention increases tremor, effortful production, and deliberate slowness of movement</td>
<td>Generalized anxiety disorder, long history of multiple somatizations, functional disability out of proportion to objective findings, attention increases tremor, fluctuation of tremulous body parts during examination, nonanatomical sensory loss</td>
</tr>
<tr>
<td>Clinical diagnostic level (follow-up)</td>
<td>Lost to follow-up</td>
<td>Possible</td>
<td>Lost to follow-up</td>
<td>Lost to follow-up</td>
</tr>
</tbody>
</table>
Patients with a unilateral arm tremor. Continuous tremor, therefore no tremor onset could be determined.

Data missing or artifacts on electromyography recordings of nontapping hand. Abbreviation: NA, not applicable.

<table>
<thead>
<tr>
<th></th>
<th>FT</th>
<th></th>
<th>OT</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal (%)</td>
<td>Abnormal (%)</td>
<td>NA (%)</td>
<td>Normal (%)</td>
</tr>
<tr>
<td>Tapping performance at 1 Hz</td>
<td>60.5</td>
<td>39.5</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Tapping performance at 3 Hz</td>
<td>52.6</td>
<td>47.4</td>
<td>0</td>
<td>86.3</td>
</tr>
<tr>
<td>Tapping performance at 5 Hz</td>
<td>60.5</td>
<td>39.5</td>
<td>0</td>
<td>76.7</td>
</tr>
<tr>
<td>Tapping response at 1 Hz</td>
<td>68.4</td>
<td>28.9</td>
<td>2.6\textsuperscript{a}</td>
<td>80.8</td>
</tr>
<tr>
<td>Tapping response at 3 Hz</td>
<td>50.0</td>
<td>47.4</td>
<td>2.6\textsuperscript{a}</td>
<td>84.9</td>
</tr>
<tr>
<td>Tapping response at 5 Hz</td>
<td>73.7</td>
<td>23.7</td>
<td>2.6\textsuperscript{a}</td>
<td>83.6</td>
</tr>
<tr>
<td>Ballistic movement response</td>
<td>44.7</td>
<td>52.6</td>
<td>2.6\textsuperscript{a}</td>
<td>86.3</td>
</tr>
<tr>
<td>Tonic coactivation</td>
<td>26.3</td>
<td>39.5</td>
<td>34.2\textsuperscript{b}</td>
<td>56.2</td>
</tr>
<tr>
<td>Coherence test</td>
<td>60.5</td>
<td>18.4</td>
<td>21.1\textsuperscript{c}</td>
<td>65.8</td>
</tr>
<tr>
<td>Loading test</td>
<td>76.3</td>
<td>23.7</td>
<td>0</td>
<td>64.4</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.
\textsuperscript{a}Data missing or artifacts on electromyography recordings of nontapping hand.
\textsuperscript{b}Continuous tremor, therefore no tremor onset could be determined.
\textsuperscript{c}Patients with a unilateral arm tremor.

One-Year Follow-Up and Practical Implication of Test Battery Results

Of 38 FT patients, 33 returned for a 1-year follow-up examination, where 2 patients were diagnosed as possible, 1 as probable, 10 as clinically established, and 20 as documented FT.\textsuperscript{1,2} Only 1 of 33 FT patients received a less certain clinical diagnosis at follow-up (probable FT at baseline, possible FT at follow-up). This was 1 of 4 FT patients who scored normal on the test battery. The remaining 3 FT patients (1 probable, 2 clinically established) with a normal test battery result were lost to follow-up (see Fig. 1).

At baseline, 13 patients were clinically diagnosed as FT based on a low level of certainty (possible or probable FT). Of those, 2 were lost to follow-up. Of the remaining 11 patients, 8 moved up to clinically established or documented FT at 1-year follow-up. All of these patients were electrophysiologically identified as FT (sum score \( \geq 3 \) points) (see Fig. 1).

All OT patients were available for a 1-year follow-up, and there were no changes in diagnoses. Considering the baseline clinical diagnosis as gold standard, the test battery identified all but 3 OT patients correctly (sum score \( < 3 \) points). One of these patients was diagnosed with SCA12 and 2 with Parkinson’s disease.

Tremor Type and Test Battery Results

Of 38 FT patients, 9 had an alternating (mean sum score \( 4.3 \pm 1.8 \)), 25 a cocontraction (mean sum score \( 3.3 \pm 1.1 \)), and 4 a mixed tremor type (mean sum score \( 4.0 \pm 1.4 \)). Sum scores did not differ between these groups (\( P = .122 \)). Of the 4 FT patients who scored normal on the test battery, 3 had a cocontraction tremor and 1 had an alternating tremor. The number of normal and abnormal test results in FT patients did not differ between tremor types (\( P = .767 \)).

Tremor Entrainment/Suppression on Clinical Examination and on Test Battery Subtests

In 19 patients, tremor suppressability and/or entrainment were clinically overt at baseline. The test battery showed suppressability and/or entrainment (response at least to 1 tapping task or the ballistic task) in 26 patients (including the 19 mentioned previously).

Of the 12 patients without suppressability and/or entrainment on the test battery, all but 3 had incorrect performances on 1 (\( n = 4 \)), 2 (\( n = 3 \)), or 3 (\( n = 2 \)) tapping tasks. Incorrect task performance during clinical examination was only overt in 5 patients.

Disease Duration and Test Battery Results

Because disease duration was significantly longer in OT than in FT, further tests were performed comparing test battery results in subgroups with shorter versus longer standing disease.

The sum scores differed between FT patients with a disease duration of \(<10\) years (\( n = 23 \)) when compared with those with a disease duration of \( \geq 10 \) years (\( n = 15; \ 4.1 \pm 1.1 \) points vs \( 2.9 \pm 1.5 \) points; \( P = .006 \)). All FT patients who scored normal on the test battery had a disease duration of \( \geq 10 \) years.

In the OT group, there was no difference regarding sum scores between patients with a disease duration of \(<10\) years (\( n = 34 \)) when compared with those with a disease duration of \( \geq 10 \) years (\( n = 39; \ 0.9 \pm 0.7 \) points vs \( 1.1 \pm 0.9 \) points; \( P = .255 \)). The 3 OT patients with an abnormal test battery had a disease duration of \( \geq 10 \) years.

Looking at the entire group of patients with a disease duration of \( \geq 10 \) years (FT, \( n = 15 \); OT, \( n = 39 \)), patients with FT still had a higher average score on the test battery when compared with OT patients (\( 2.9 \pm 1.5 \) points vs \( 1.1 \pm 0.9 \) points; \( P < .001 \)). The predefined cut-off score for a diagnosis of laboratory-
Supported FT of 3 points yielded a test sensitivity of 73.3% and a specificity of 92.3% (P < .001).

In the subgroup of patients with a disease duration of <10 years (FT, n = 23; OT, n = 34), patients with FT had a higher average test battery score when compared with OT patients (4.1 ± 1.1 points vs 0.9 ± 0.7 points; P < .001). The cut-off score of 3 points had a test sensitivity and specificity of 100% (P < .001).

### Modifications to Test Battery

We evaluated modifications to the test battery to assess whether sensitivity and specificity might be further improved. However, the previously proposed test battery remained the one with the best test accuracy (see Table 3).

### Discussion

A diagnosis of FT is based on positive criteria by history and, more important, by examination and can be supported by electrophysiological tests. In this study, we showed that an electrophysiological test battery can discriminate patients with FT and OT with high sensitivity and specificity. This study further showed high interrater and test-retest reliability, and we now propose its use in the work-up of patients with presumed FT. This provides confidence in the use of this tool in clinical and research practice.

In our cohort of FT patients, the initial clinical diagnosis had a low level of certainty (probable or possible) in 13 of 38 patients. This group of patients with initial clinical uncertainty may be the most important one for the implementation of an ancillary objective test in routine practice. In most patients, the clinical diagnosis of FT rose to a higher level of certainty after 1-year follow-up. Our data show that the application of the electrophysiological test battery would have allowed a higher level of certainty 1 year earlier. This may be of clinical relevance because a short delay between symptom onset and diagnosis is regarded as a positive prognostic factor. We therefore propose the test battery to be used in establishing a laboratory-supported level of certainty in patients in whom uncertainty remains after clinical assessment.

In 25 of 38 patients, a rather confident clinical baseline diagnosis (clinically established or documented) was achieved. Although the clinician may not need ancillary tests to gain further diagnostic certainty in this scenario, the test battery can still provide objective evidence and help convey this new and unfamiliar diagnosis to the patient. The initial explanation of the diagnosis is regarded as the most important step in initiating treatment. Sharing diagnostic clinical signs such as entrainment or suppressibility with patients illustrates how the diagnosis is made and is regarded as a powerful way of persuading them about the nature of their illness. If these features are clinically overt, there is no need for electrophysiological tools. However, in addition to those who already showed entrainment or tremor suppressibility during clinical examination, the test battery identified another 7 patients who showed these characteristics and another 9 patients who did not perform the task correctly. In this regard, visualized electrophysiological data can be used as an objective demonstration tool. As recently proposed, diagnostic features such as entrainment may even be useful as a treatment strategy, which would be expected to facilitate acceptance of the diagnosis.

In addition to this clinical utility, we propose that the test battery could be useful in a research setting. Research studies can be aided by having specific objective criteria for diagnosis to reduce diagnostic error and to include patients in experimental studies on the

### Table 3: Sensitivity, specificity, and accuracy of the original and modified versions of the test battery

<table>
<thead>
<tr>
<th>Sum score (mean ± SD)</th>
<th>FT</th>
<th>OT</th>
<th>P&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Cut Off&lt;sup&gt;a&lt;/sup&gt;</th>
<th>False negative</th>
<th>False positive</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>P&lt;sup&gt;c&lt;/sup&gt;</th>
<th>ACC, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original TB</td>
<td>3.6 ± 1.4</td>
<td>1.0 ± 0.8</td>
<td>&lt;.001</td>
<td>≥3</td>
<td>4</td>
<td>3</td>
<td>89.5</td>
<td>95.9</td>
<td>&lt;.001</td>
<td>93.7</td>
</tr>
<tr>
<td>Modified TB&lt;sup&gt;1&lt;/sup&gt;</td>
<td>3.4 ± 1.5</td>
<td>0.6 ± 0.7</td>
<td>&lt;.001</td>
<td>≥3</td>
<td>7</td>
<td>1</td>
<td>81.6</td>
<td>98.6</td>
<td>&lt;.001</td>
<td>92.8</td>
</tr>
<tr>
<td>Modified TB&lt;sup&gt;2&lt;/sup&gt;</td>
<td>3.2 ± 1.3</td>
<td>0.7 ± 0.7</td>
<td>&lt;.001</td>
<td>≥3</td>
<td>4</td>
<td>10</td>
<td>89.5</td>
<td>86.3</td>
<td>&lt;.001</td>
<td>92.4</td>
</tr>
<tr>
<td>Modified TB&lt;sup&gt;3&lt;/sup&gt;</td>
<td>3.0 ± 1.2</td>
<td>0.7 ± 0.6</td>
<td>&lt;.001</td>
<td>≥3</td>
<td>4</td>
<td>7</td>
<td>89.5</td>
<td>90.4</td>
<td>&lt;.001</td>
<td>91.9</td>
</tr>
<tr>
<td>Modified TB&lt;sup&gt;4&lt;/sup&gt;</td>
<td>2.7 ± 1.2</td>
<td>0.6 ± 0.6</td>
<td>&lt;.001</td>
<td>≥3</td>
<td>5</td>
<td>6</td>
<td>86.8</td>
<td>91.8</td>
<td>&lt;.001</td>
<td>90.1</td>
</tr>
</tbody>
</table>

Abbreviations: SD, standard deviation; FT, functional tremor; OT, organic tremor; ACC, accuracy; TB, test battery; Modified TB<sup>1</sup>, test battery without loading test; Modified TB<sup>2</sup>, test battery without 5 Hz tapping performance; Modified TB<sup>3</sup>, test battery without both 5 Hz tapping tasks (performance and response to tapping); Modified TB<sup>4</sup>, test battery without loading test without both 5 Hz tapping tasks.

<sup>a</sup>Cut-off sum score for a diagnosis of functional tremor.

<sup>b</sup>False negative (t test).

<sup>c</sup>False positive (chi-squared test).
basis of uniform standards. This might be particularly useful in multicenter studies where different clinicians are including patients. We suggest that this test battery could be a useful objective test for FT in future research studies.

The equipment required for these tests is simple (2 accelerometers, EMG recording facilities, metronome, 500-gram weight), and we estimate that the time needed for recordings, including set up, would be about 15 minutes. With automated tremor analysis (using custom software developed by the user), analysis should take approximately 15 to 30 minutes per patient.

The test battery yielded high sensitivity and specificity. Nevertheless, there were a few false positive and false negative cases. Of 73 OT patients, 3 scored abnormal results. Clinically, there was no doubt regarding the diagnosis in all 3 patients. Two patients suffered from advanced Parkinson’s disease, and 1 patient was diagnosed with SCA12, thus suffering not only from tremor but also ataxia. Only 1 other OT patient had a combination of tremor and ataxia (FXTAS) and scored normal on the test battery; hence we do not know if the test can discriminate patients with FT from OT in the presence of coexisting ataxia.

Of 38 FT patients, 4 had normal test battery results. Of them, 2 had a low degree of clinical diagnostic certainty (probable FT), and 2 were diagnosed as clinically established FT. Unfortunately, 3 of these patients (1 probable and 2 clinically established FT) were lost to follow-up; hence we do not know for sure if these were truly false negatives or if the clinical diagnoses might have changed over time from FT to OT or even to OT with functional overlay. One FT patient with a negative result was clinically still considered as FT after 1 year, although the clinical level of certainty had further decreased from probable to possible.

All 3 false positive and all 4 false negative cases had a disease duration of at least 10 years. Although test sensitivity and specificity were excellent in the group of patients with a disease duration of <10 years (100% and 100%), they were worse in those with a disease duration of ≥10 years (73.3% and 92.3%). We therefore propose that in patients with long disease durations and unexpected test battery results, the clinical diagnosis should remain the working diagnosis and follow-up of the patients is mandatory. On the other hand, in those with short disease durations, the test battery reliably allows a high level of diagnostic certainty and adequate therapy early in the disease course.

A number of limitations of our study must be highlighted. The main difficulty is that a true test of diagnostic accuracy is unknown because the gold standard is clinical diagnosis by a movement disorders neurologist. Therefore, the sensitivity of the test cannot exceed that of a clinical diagnosis. A potential limitation of this study is that the clinical diagnosis was made by one movement disorder neurologist and not by consensus opinion reached by all 4 experts. However, this would necessitate video assessments, and a previous study had shown divergence of diagnostic opinions provided by clinical assessments of examiners reviewing videotapes. We aimed to strengthen the certainty of the gold standard diagnosis by reassessing patients clinically after 1 year and showed that there was no change in diagnosis in those available for follow-up (106 of 111 patients). As mentioned previously, 3 FT patients with a normal test battery result unfortunately were unavailable for a follow-up examination. If these patients would actually have turned out to be documented FT, test sensitivity would be worse, which clearly is a limitation of this study.

Sensitivity and specificity of the test battery might be further improved by certain modifications. For example, performing multiple trials of each subtest might enhance the chance to detect abnormality in coherence, coactivation, tremor suppression, and so on. However, this would result in prolonging the time needed for the test battery and increase subjectivity. We have shown that alterations to the test battery such as omitting the 5 Hz tapping or the loading tasks worsened test accuracy, which supports the results from our pilot study in this new cohort. The threshold for abnormality in each test could be adjusted to potentially exclude false positives, which was not performed because this was designed as a validation study. The next step, however, may be to further optimize certain subtests and revalidate the modified test battery in a new cohort.

We do not know if the same test battery is useful in FT of other body parts or in patients with a combination of organic and functional diseases. A recently published case report suggested that the use of head accelerometry with entrainment and ballistic movements may be useful in physiological studies of head tremor. However, it may be advisable to use distinct distraction tasks, for example, tapping to a metronome with the less-affected leg or the tongue in leg tremor or head tremor, respectively. We also do not know about the utility of loading in leg tremors. Further studies are needed to confirm the usefulness of electrophysiological tests in these nonarm tremors.

We conclude that clinical assessment remains the most important aspect in evaluating patients with FT, especially in those with longstanding disease. However, the proposed test battery can help strengthen the diagnostic certainty by providing a laboratory-supported level of certainty. This is of particular importance because early recognition and management are critical for a suitable long-term outcome.

References