

## Brief Reports

## Digitizing Tablet and Fahn–Tolosa–Marín Ratings of Archimedes Spirals have Comparable Minimum Detectable Change in Essential Tremor

Rodger J Elble<sup>1\*</sup> & Aaron Ellenbogen<sup>2</sup>

<sup>1</sup> Department of Neurology, Southern Illinois University School of Medicine, Springfield, IL, USA, <sup>2</sup> Michigan Institute for Neurological Disorders, Farmington Hills, MI, USA

### Abstract

**Background:** Drawing Archimedes spirals is a popular and valid method of assessing action tremor in the upper limbs. We performed the first blinded comparison of Fahn–Tolosa–Marín (FTM) ratings and tablet measures of essential tremor to determine if a digitizing tablet is better than 0–4 ratings in detecting changes in essential tremor that exceed random variability in tremor amplitude.

**Methods:** The large and small spirals of FTM were drawn with each hand on two consecutive days by 14 men and four women (age  $60 \pm 8.7$  years [mean  $\pm$  SD]) with mild to severe essential tremor. The drawings were simultaneously digitized with a digitizing tablet. Tremor in each digitized drawing was computed with spectral analysis in an independent laboratory, blinded to the clinical ratings. The mean peak-to-peak tremor displacement (cm) in the four spirals and mean FTM ratings were compared statistically.

**Results:** Test–retest intraclass correlations (ICCs) (two-way random single measures, absolute agreement) were excellent for the FTM ratings (ICC 0.90, 95% CI 0.76–0.96) and tablet (ICC 0.97, 95% CI 0.91–0.99).  $\log_{10}$  tremor amplitude ( $T$ ) and FTM were strongly correlated ( $\log T = \alpha \text{FTM} + \beta$ ,  $\alpha \approx 0.6$ ,  $\beta \approx -1.27$ ,  $r = 0.94$ ). The minimum detectable change for the tablet and FTM were 51% and 67% of the initial assessment.

**Discussion:** Digitizing tablets are much more precise than clinical ratings, but this advantage is mitigated by the natural variability in tremor. Nevertheless, the digitizing tablet is a robust method of quantifying tremor that can be used in lieu of or in combination with clinical ratings.

**Keywords:** Essential tremor, rating scale, spirometry, minimum detectable change

**Citation:** Elble RJ, Ellenbogen A. Digitizing tablet and Fahn–Tolosa–Marín ratings of Archimedes spirals have comparable minimum detectable change in essential tremor. *Tremor Other Hyperkinet Mov.* 2017; 7. doi: 10.7916/D89S20H7

\*To whom correspondence should be addressed. E-mail: relble@siu.edu

**Editor:** Elan D. Louis, Yale University, USA

**Received:** May 20, 2017 **Accepted:** June 20, 2017 **Published:** July 7, 2017

**Copyright:** © 2017 Elble et al. This is an open-access article distributed under the terms of the Creative Commons Attribution–Noncommercial–No Derivatives License, which permits the user to copy, distribute, and transmit the work provided that the original authors and source are credited; that no commercial use is made of the work; and that the work is not altered or transformed.

**Funding:** This work was supported by a grant from the Spastic Paralysis Research Foundation of Kiwanis International, Illinois–Eastern Iowa District. The clinical trial from which these data were derived was funded by Jazz Pharmaceuticals.

**Financial Disclosures:** Rodger Elble received grant support from the Spastic Paralysis Research Foundation of Kiwanis International, Illinois–Eastern Iowa District. He received consulting fees from Sage Therapeutics and is a paid video rater for InSightec. Aaron Ellenbogen has received consulting fees from Sage Therapeutics, Sunovion, and Allergan. He has received honoraria for speaking from Allergan, Arbor, Ipsen, Lundbeck, and U.S. World Meds.

**Conflicts of Interest:** Rodger Elble was a paid consultant in the clinical trial that produced the data in this study and was responsible for analyzing the tablet data. Aaron Ellenbogen was a paid clinical investigator in the clinical trial and performed all of the clinical ratings.

**Ethics Statement:** This study was performed in accordance with the ethical standards detailed in the Declaration of Helsinki. The authors' institutional ethics committee has approved this study, and all patients have provided written informed consent.

### Introduction

Tremor rating scales provide crude, nonlinear, subjective assessments of tremor severity.<sup>1</sup> The Fahn–Tolosa–Marín (FTM) tremor rating scale<sup>2</sup> uses 0–4 anchors to assess tremor in drawings of Archimedes spirals. The Bain and Findley scale uses 0–10 anchors.<sup>3</sup> Both scales have a strong logarithmic relationship with tremor amplitude measured with a digitizing tablet, consistent with the Weber–Fechner law of psychophysics.<sup>4–7</sup>

Digitizing tablets are capable of providing linear objective measures of tremor in writing and drawings.<sup>6,8–13</sup> The Wacom Intuos 3 digitizing tablet (www.wacom.com) has been used most commonly and has an accuracy of  $\pm 0.25$  mm and a sampling frequency of 100 samples/s, which are adequate for measuring the amplitude and frequency of a tremor that is visible to the unaided eye.<sup>8,9</sup> Digitizing tablets are unable to detect pen motion when the pen tip is greater than 1 cm above the tablet surface and lack sufficient sensitivity to measure physiologic

tremor. Thus, tablets, like clinical ratings, have ceiling and floor effects at the extremes of tremor amplitude.

The greater precision of tablets, relative to rating scales, enables one to detect much smaller changes in tremor amplitude. However, this advantage of tablets is diminished when random variability in tremor is large. Tablets measure random variability precisely, but a change in tremor must exceed random variability to be recognizable as a statistically significant change resulting from treatment or disease progression (minimum detectable change).<sup>4,9</sup> Therefore, we sought to determine if a digitizing tablet is better than FTM part B spiral ratings in detecting changes in essential tremor that exceed random variability in tremor amplitude.

**Methods**

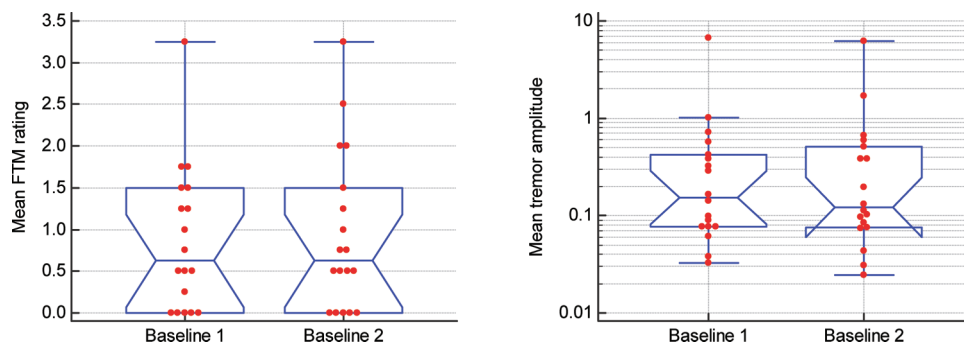
Twenty patients were enrolled in an unpublished open-label pharmacokinetic–pharmacodynamic study of sodium oxybate for the treatment of essential tremor, conducted by Jazz Pharmaceuticals. Details of the study design can be found on ClinicalTrials.gov (<https://clinicaltrials.gov/ct2/show/study/NCT00598078>). All patients participated after giving their informed written consent, approved by a local human subjects committee. The patients stopped all drugs for tremor at least five half-lives before the study. They also abstained from alcohol and caffeine for 48 hours. Fourteen men and four women (age  $60 \pm 8.7$  years [mean  $\pm$  SD]) with mild to severe essential tremor completed the study in which placebo or sodium oxybate was administered orally at 8 a.m. on three consecutive days. Baseline assessments of tremor were performed each day between 7 and 8 a.m. Tremor was quantified with the FTM rating scale and a digitizing tablet. All patients were examined by the same neurologist (A.L.E.). The paper with the large and small FTM spiral templates was mounted on a Wacom Intuos 3 digitizing tablet so that the same drawings were rated and digitized. Tremor amplitude in each digitized drawing was computed in an independent central laboratory using spectral analysis. The software used is available online.<sup>9</sup> The technician performing the tablet analyses was blinded to the tremor ratings and study design. The grand average of mean peak-to-peak tremor displacement (cm) in the four spirals (large and small spirals drawn with each hand) was compared with the grand average of the four FTM spiral ratings.

A paired t test analysis of the baseline FTM spiral ratings and tablet measures on days 1 and 2 revealed a statistically significant practice effect or carryover effect from day 1 to day 2. The mean FTM spiral rating decreased slightly (1.21 to 0.88,  $t=-3.011$ ,  $p=0.008$ ), as did the log-transformed tablet measure (geometric mean 0.28 to 0.20,  $t=-2.431$ ,  $p=0.026$ ). By contrast, the baseline FTM and tablet means were statistically identical on days 2 and 3 (mean FTM spiral ratings, 0.88 and 0.94,  $t=0.719$ ,  $p=0.48$ ; geometric mean tablet measures, 0.20 to 0.19,  $t=-0.457$ ,  $p=0.65$ ). We therefore used the data from days 2 and 3 in this study to estimate test–retest reliability and MDC. In this, study, baseline 1 refers to the baseline data from study day 2, and baseline 2 refers to the data from study day 3. Baseline assessments from these two days were used to compute test–retest reliability (two-way random single measures intraclass correlations [ICCs], absolute agreement) and minimum detectable change (MDC) for the FTM spiral ratings and digitizing tablet measurements.<sup>14</sup>

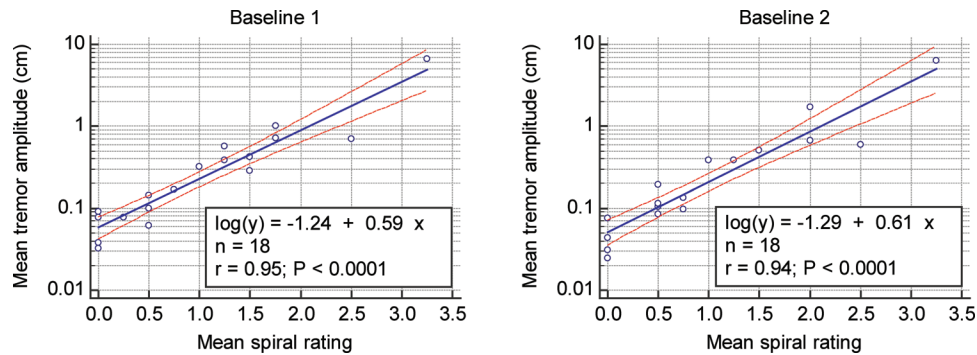
MDC was computed using the formula  $MDC=SDd \cdot 1.96$ , where SDd is the standard deviation of the differences for the two measurements.<sup>14</sup> For the grand average of the four FTM spiral ratings, MDC was expressed as a percentage of the baseline 1 mean (MDC%). The tablet data were positively skewed, so  $\log_{10}$  transformation was performed to normalize these data. Note that SDd of log-transformed data is a ratio, and the MDC is therefore also a ratio; they are not log SDd and log MDC of the non-transformed data.<sup>15</sup> MDC% of the log-transformed data is expressed as a percentage of the baseline geometric mean, using the equation  $MDC\%=(1-10^{-MDC}) \cdot 100$ .<sup>15</sup> All statistical analyses were performed with MedCalc<sup>®</sup> statistical software ([www.medcalc.org](http://www.medcalc.org)).

**Results**

The mean spiral ratings did not differ statistically from a normal distribution (D’Agostino–Pearson test:  $p=0.16$  for baseline 1 data and  $p=0.13$  for baseline 2 data). The tablet data were positively skewed and deviated significantly from a normal distribution (D’Agostino–Pearson test:  $p<0.0001$  for baseline 1 and 2 data), so  $\log_{10}$  transformation was performed to normalize these data, producing data that did not deviate significantly from a normal distribution (D’Agostino–Pearson test:  $p=0.25$  for baseline 1 data and  $p=0.18$  for baseline 2 data). The FTM ratings exhibited a floor effect in this patient population (Figure 1).



**Figure 1. Distributions of Mean FTM Spiral Ratings and Log Tremor Amplitudes.** Notched box and whisker plots of mean FTM spiral ratings and tablet measures are shown for the two baseline assessments. The baseline medians did not differ significantly for either measure of tremor severity. The FTM data exhibit a floor effect at 0. FTM, Fahn–Tolosa–Marin.



**Figure 2. Linear Regression Equations for Log<sub>10</sub> Tremor Amplitude vs. FTM.** Regression lines (blue) and 95% confidence intervals (red broken lines) are shown for log mean tremor amplitude vs. mean FTM spiral rating for baseline 1 and 2 assessments. FTM, Fahn–Tolosa–Marin.

**Table 1. Minimum Detectable Change Results for FTM and Tablet Measures of Tremor Amplitude**

	FTM Mean (cm)	SDd FTM Baseline 1–2	Tablet Geometric Mean (cm)	SDd Tablet Baseline 1–2	MDC%	
					FTM	Tablet
Baseline 1	0.88	0.41	0.20	0.16 <sup>4</sup>	90% <sup>1</sup>	51% <sup>3</sup>
Baseline 2	0.94		0.19		67% <sup>2</sup>	

Abbreviations: FTM, Fahn–Tolosa–Marin; MDC, Minimum Detectable Change (SDd·1.96); SDd: Standard Deviation of the Differences.

<sup>1</sup>MDC%: percentage of baseline 1 mean FTM.

<sup>2</sup>MDC%: percentage of baseline 1 mean FTM, computed with the Weber–Fechner equations in Figure 2.

<sup>3</sup>MDC%: percentage of baseline 1 geometric mean.

<sup>4</sup>SDd of log-transformed data.

Regression analysis revealed a very strong linear Weber–Fechner relationship ( $\log T = \alpha \cdot \text{FTM} + \beta$ ) between mean FTM spiral ratings and log mean tablet tremor amplitudes  $T$  (cm) for baseline 1 and baseline 2 measurements (Figure 2). Test–retest ICC was excellent for the FTM ratings (ICC 0.90, 95% CI 0.76 to 0.96) and log-transformed tablet measures (ICC 0.97, 95% CI 0.91 to 0.99).

The MDC for the digitizing tablet was 51% of the baseline geometric mean tremor amplitude (Table 1). The MDC for FTM was 90% of mean baseline spiral rating. However, FTM is a non-linear ordinal scale, so computing % change is not valid.<sup>5,16</sup> Therefore, we converted FTM to actual tremor amplitude using the average slope and intercept in Figure 2 for the two regression equations relating FTM and log tremor amplitude (average slope  $\alpha=0.6$  and intercept  $\beta=-1.26$ ), and we computed the MDC% using the following equations derived by Eble and colleagues.<sup>5</sup>

$$\log T = \alpha \cdot \text{FTM} + \beta$$

$$\left(\frac{T_2 - T_1}{T_1}\right) 100 = \left[10^{\alpha(\text{FTM}_2 - \text{FTM}_1)} - 1\right] 100$$

$$\alpha = 0.6 \tag{1}$$

$$\text{FTM}_2 - \text{FTM}_1 = \text{MDC} = \text{SDd} \cdot 1.96 = -0.41 \cdot 1.96 = -0.80$$

$$\left(\frac{T_2 - T_1}{T_1}\right) 100 = \text{MDC}\% = (10^{0.6 \cdot \text{MDC}} - 1) 100 = -67\%$$

In the above calculations, SDd (–0.41) is the standard deviation of the differences between the baseline 1 and baseline 2 FTM scores. This estimate of MDC% (67%) is similar to that found for the tablet.

Our estimates of MDC% appear to be very robust and not dependent on normalization of the data. We computed the MDC% of the tablet data without log transformation, using the baseline 1 mean (0.62 cm) and the SDd of the two baselines (0.20 cm). Using these values, the MDC% is as follows:

$$\text{MDC}\% = 0.20 \cdot 1.96 \cdot 100 / 0.62 = 63\% \tag{2}$$

### Discussion

This is the first blinded study demonstrating a strong correlation between tablet and FTM spiral ratings, and this study provides much-needed estimates of test–retest reliability and MDC% for tablet and FTM spiral ratings. We have shown that tablet measures are highly correlated with FTM tremor ratings. The test–retest ICC for the tablet was only marginally better than the FTM ICC. However, the FTM ICC probably would have been lower if different raters had been used to assess the two baselines because intra-rater reliability is much better than inter-rater reliability for tremor rating scales.<sup>1</sup> Also, we compared the average ratings and amplitudes of 4 spirals, and this is known to reduce test–retest variability.<sup>17</sup>

Haubenberger and colleagues<sup>6</sup> found a strong ( $r > 0.9$ ) logarithmic relationship between tablet measures of tremor amplitude and the Bain and Findley 0–10 ratings of tremor in Archimedes spirals, and the slope of this relationship was 0.2436. From this relationship for 0–10 ratings, the slope for a 0–4 scale can be estimated as  $0.2436 \cdot (10/4) = 0.601$ ,<sup>16</sup> which is what we found in this study. Thus, the logarithmic relationship between tablet measures and tremor ratings is robust, regardless of the scale that is used.

There is no published evidence that the Bain and Findley 0–10 ratings are more sensitive to change than FTM 0–4 ratings.<sup>1</sup> Hopfner and coworkers<sup>18</sup> estimated the minimum detectable change of the Bain and Findley scale to be 2 points, or 20% of the maximum rating 10. We found the MDC of the mean FTM spiral rating to be 0.8 points, which is 20% of the maximum rating 4.

Detectable change in essential tremor is limited by the considerable natural variability of tremor amplitude over time. The variability in tremor amplitude is so great that the MDC (the smallest detectable change exceeding random variability) of the digitizing tablet is similar to the MDC of the FTM 0–4 ratings and the Bain and Findley 0–10 ratings. Digitizing tablets are much more precise than clinical ratings, but this advantage is mitigated by the natural variability in tremor.

Digitizing tablets have potential floor and ceiling effects. They cannot measure tremor that is not visible because their accuracy is roughly  $\pm 0.25$  mm. They also cannot record tremor that is so severe that the pen tip does not remain within 1 cm of the tablet surface. However, FTM ratings had an obvious floor effect in our patient population, but the tablet exhibited no floor effect for these patients. Tremor severity was not great enough in our patient cohort to examine a ceiling effect for the tablet vs. FTM.

Nevertheless, the digitizing tablet is clearly a valid and robust method of quantifying tremor. It can be used in lieu of, or in combination with, clinical ratings of tremor in Archimedes spirals. The tablet provides an accurate, clinically meaningful assessment of tremor amplitude. These devices cost a few hundred dollars, and free software for tremor analysis is available on the internet.<sup>6,9</sup>

Our study has limitations. Our estimates of test–retest reliability and MDC% were computed using two baseline assessments at the same time of the day on two consecutive days, while controlling for tremor medications, caffeine, and alcohol. Random test–retest variability might be greater if the interval between assessments was longer and if the other controls were less stringent. Our results need to be confirmed using baseline assessments at intervals of 1 week and 1 month, which are common intervals of assessment in clinical trials.

## References

1. Elble R, Bain P, Forjaz MJ, et al. Task force report: scales for screening and evaluating tremor: critique and recommendations. *Mov Disord* 2013;28:1793–1800. doi: 10.1002/mds.25648
2. Fahn S, Tolosa E, Marin C. Clinical rating scale for tremor. In: Jankovic J, Tolosa E, editors. *Parkinson's disease and movement disorders*. 2nd ed. Baltimore: Williams & Wilkins, 1993. p 225–234.
3. Bain PG, Findley LJ. *Assessing tremor severity: a clinical Handbook*. London: Smith-Gordon, 1993.
4. Elble R, Hellriegel H, Raethjen J, Deuschl G. Logarithmic relationship between head tremor and 5-point tremor rating. *Mov Disord* 2011;26:S375–376. doi: 10.1002/mds.23579
5. Elble RJ, Pullman SL, Matsumoto JY, Raethjen J, Deuschl G, Tintner R. Tremor amplitude is logarithmically related to 4- and 5-point tremor rating scales. *Brain* 2006;129:2660–2666. doi: 10.1093/brain/awl190
6. Haubenberger D, Kalowitz D, Nahab FB, et al. Validation of digital spiral analysis as outcome parameter for clinical trials in essential tremor. *Mov Disord* 2011;26:2073–2080. doi: 10.1002/mds.23808
7. Gescheider GA. *Psychophysics: the fundamentals*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 1997.
8. Haubenberger D, Abbruzzese G, Bain PG, et al. Transducer-based evaluation of tremor. *Mov Disord* 2016;31:1327–1336. doi: 10.1002/mds.26671
9. Elble RJ, McNames J. Using Portable Transducers to Measure Tremor Severity. *Tremor Other Hyperkinet Mov* 2016;6. doi: 10.7916/D8DR2VCC
10. Rudzinska M, Izowski A, Banaszkiwicz K, Bukowczan S, Marona M, Szczudlik A. Quantitative tremor measurement with the computerized analysis of spiral drawing. *Neurol Neurochir Pol* 2007;41:510–516.
11. Pullman SL. Spiral analysis: a new technique for measuring tremor with a digitizing tablet. *Mov Disord* 1998;13:85–89. doi: 10.1002/mds.870131315
12. Elble RJ, Sinha R, Higgins C. Quantification of tremor with a digitizing tablet. *J Neurosci Methods* 1990;32:193–198. doi: 10.1016/0165-0270(90)90140-B
13. Elble RJ, Brilliant M, Leffler K, Higgins C. Quantification of essential tremor in writing and drawing. *Mov Disord* 1996;11:70–78. doi: 10.1002/mds.870110113
14. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005;19:231–240.
15. Spooner J, Dressing SA, Meals DW. Minimum detectable change analysis. *Tech Notes* 7, December 2011. [http://www.bae.ncsu.edu/programs/extension/wqg/319monitoring/tech\\_notes.htm](http://www.bae.ncsu.edu/programs/extension/wqg/319monitoring/tech_notes.htm): U.S. Environmental Protection Agency; 2011.
16. Deuschl G, Raethjen J, Hellriegel H, Elble R. Treatment of patients with essential tremor. *Lancet Neurol* 2011;10:148–161. doi: 10.1016/S1474-4422(10)70322-7
17. Elble RJ, Lyons KE, Pahwa R. Leveteracetam is not effective for essential tremor. *Clin Neuropharmacol* 2007;30:350–356. doi: 10.1097/WNF.0b013e31807a32c6
18. Hopfner F, Erhart T, Knudsen K, et al. Testing for alcohol sensitivity of tremor amplitude in a large cohort with essential tremor. *Parkinsonism Relat Disord* 2015;21:848–851. doi: 10.1016/j.parkreldis.2015.05.005