# Evaluating Factors That Can Influence Spirography Ratings in Patients with Essential Tremor

William G. Ondo, MD, Anchi Wang, MD, Madhavi Thomas, MD, Kevin Dat Vuong, MA

Parkinson's Disease Center and Movement Disorders Clinic, Department of Neurology, Baylor College of Medicine, Houston, Texas

## Abstract

**Objective:** To evaluate the safety and efficacy of tetrabenazine (TBZ) in the control of hyperkinetic movement disorders and to examine its age-related tolerability.
**Background:** Spirograph drawings are used in most comprehensive assessments of essential tremor (ET). Nevertheless, several different paradigms are used and no effort has been made to compare these.
**Methods:** We used two different cohorts to assess different aspects of spiral rating. In the first, we had subjects simulate different levels of effort by writing 1. "normally", 2. "slowly and carefully", 3. "softly", and 4. "rapidly" using both their dominant and non-dominant hands. In the second subjects wrote 1. freehand, 2. traced a previously drawn spirograph, and 3. drew in between the lines of two drawn spirals. (Figure 1) Subjects drew each with both "supported" (regular writing) and "unsupported" writing. The spirals were coded, randomized and blindly rated on a 0-9 scale.
**Results:** Unsupported drawings were consistently rated as worse than supported spirals, and the dominant hand was generally better than the non-dominant hand. Spirals traced on top of previously drawn spirals tending to have the highest test scores, whereas freehand spirals were consistently rated with the lowest scores. Inter-rater and intra-rater reliability was also best for the "on the line" traced spirographs, especially unsupported. "Effort" had little effect on ratings.
**Conclusion:** Based on our results, we recommend that assessment of ET include unsupported, "on the line" tracing spirals.

## Introduction

Essential tremor (ET) [1] is arguably the most common movement disorder, effecting between 1-2% of the population. [2] No single assessment measure of ET has gained universal acceptance, and different ET measures often do not correlate well with each other. [3-5] Archimedes spirograph drawings are employed in most comprehensive tremor rating scales. [4,6-8] Nevertheless, these different scales use different variations of spirograph drawing, which have never been compared against each other. Furthermore the consequences of subject writing style and effort have never been evaluated. In order to define which method of spirograph drawing best captures meaningful and reproducible data, we have systematically compared different spiral writing techniques in ET patients.

## Methods

Two separate patient cohorts were recruited from the Baylor College of Medicine Parkinson's Disease Center and Movement Disorders Clinic. All patients met the criteria for probable essential tremor [9] and were recruited over 9-month periods, in 1997-1998 and 2001-2002. All demographic data was obtained from direct interview.

In the first group, we were interested in how subject effort may influence performance on tremor spirographs. Subjects drew spirographs freehand on a blank paper with a standard size ballpoint pen, and were allowed full arm support. In order to simulate different effort patterns that we witness in clinical practice, subjects were instructed to write the spirographs 1. "normally", 2. "slowly and carefully", 3. "softly", and 4. "rapidly" using both their dominant and non-dominant hands. Subsequently, all spirograph drawings were coded, randomized, and blindly rating (A.W.) using the Bain and Finley scale (0-9). [4] We used this rating system exclusively because it has published examples by which raters could be trained, and we felt it would be more sensitive to minor differences than a (0-4) scale. Since this group contained a high number of ET patients with concurrent dystonia (DT), we also compared pure ET to DT.

In the second group, we were interested in how three different previously used spirograph drawing paradigms affected results. Subjects drew 1. freehand, 2. traced a previously drawn spirograph, and 3. drew in between the lines of two drawn spirals. (Figure 1) Subjects drew each both "supported" (regular writing) and "unsupported" (allowing only the pen to touch the writing surface) with both their dominant and non-dominant hands (total of 12 spirographs). The order with which they drew these was computer randomized to avoid any fatiguing or learning effect. The spirals were then copied, coded, and blindly read by two different raters (A.W. and M.T.) on two different occasions approximately three months apart. The means of the initial ratings were used for analysis comparing the different methods. We then calculated Spearman rho inter-rater and Kendall's W intra-rater reliability measures based on ordinal data. Finally, we compared various spiral paradigms to standard (0-4) visual assessment of postural outstretched hand tremor.

All analyses were performed via SPSS v10. Parametric statistical procedures included independent samples t-tests while nonparametric procedures entailed Friedman, Mann-Whitney U, and Wilcoxon signed rank tests. Exact *p*-values are reported whenever appropriate.

## Results

In the first group of 83 subjects (47 male, 73 right hand dominant), the mean age was 61.4±16.6 years and the mean duration of tremor was 21.5±16.6 years. Thirty-seven (44.5%) had some form of concurrent dystonia (DT): cervical 19, upper extremity 16 (6 with concurrent cervical), cranial 4 (1 with concurrent cervical), and lower extremity 1.

Spirograph scores while patients were writing "normal", "fast", and "careful" were similar, but "soft" scores were worse in both dominant ($p<0.001$) and non-dominant ($p<0.001$) hands. (Table 1) This same pattern was seen in both pure ET ($p<0.001$) and DT ($p<0.001$). Dominant hand scores were better than non-dominant scores in "normal" ($p<0.001$), "fast" ($p<0.001$) and "slow" ($p<0.001$) spirals but were similar in the "soft" group. This difference, however, was only seen in the pure ET group, "normal" ($p<0.001$), "fast" ($p<0.001$) and "slow" ($p<0.001$). The DT subjects had similar spiral scores in both hands. Overall, ET scores were worse than DT scores ($p<0.0001$).

In the second group of 54 subjects (28 male), the mean age was 62.8±15.1 years and mean duration of tremor was 20.1±16.5 years. Only eight subjects (14.8%) had any dystonia, and 48 (88.9%) were right hand dominant.

There were several significant differences in spirals as a function of writing paradigms. First, overall, dominant hand spirals were worse than non-dominant hand spirals ($p<0.05$). This difference, however was exclusively powered by supported writing in all three paradigms "on line" ($p<0.001$), "in between lines" ($p<0.005$), and "freehand" ($p<0.001$). Unsupported spirals did not differ between hands. Second, unsupported spirals were worse that supported spirals in all three paradigms "on line" ($p<0.05$), "in between lines" ($p<0.001$), and "freehand" ($p<0.001$).

The three groups "on line", "in between lines", "freehand" also differed. Overall, "freehand" scores were rated as less severe than "in between lines" ($p<0.001$) and "on the line" ($p<0.001$), which tended to have the highest score. The same pattern was seen in all combined supported and all combined unsupported spirals. The results, however, varied as a combined function of both hand (D vs. ND) and whether support was allowed. In the supported dominant hand, "freehand" was less severe than "on the line" ($p<0.01$) but only tended to be less severe than "in between lines" ($p=0.08$). In the unsupported dominant hand, "freehand" was only less severe than "on the line" ($p<0.005$). Also in the unsupported dominant hand "in between lines" tended to be less severe than "on the line" ($p=0.06$). In the supported non-dominant hand, "freehand" was less severe than "on the line" ($p<0.05$), whereas in the unsupported non-dominant hand, "freehand" was less severe than "in between lines" ($p<0.005$).

Inter-rater reliability was excellent for "on the line", good for "in between lines", and poor for "freehand". (Table 3) It was best for the unsupported "on the line" spirals. Intra-rater reliability was excellent in all three paradigms, but best for "on the line" spirals. (Table 4)

Overall, spirals correlated with outstretched visual assessment of postural tremor (Spearman rho=0.54, $p<0.001$). All three were similar, however correlation of all unsupported spirals was modestly higher than all supported spirals (r=0.57 vs. r=0.50).

## Discussion

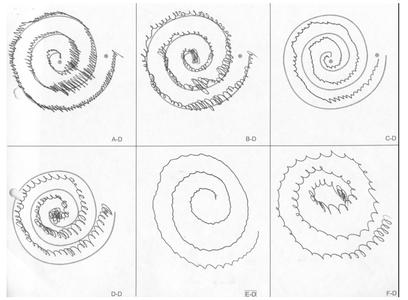
Our study demonstrated significant differences among spirographs drawn by the same subject that were mostly a function of writing paradigms. Unsupported drawings were consistently rated as worse than supported spirals, and dominant hand was generally better than the non-dominant hand. Spirals traced on top of previously drawn spirals tending to have the highest test scores, whereas freehand spirals were consistently rated as the lowest scores. Inter-rater and intra-rater reliability was also best for the "on the line" traced spirographs. Patients who applied minimal pressure drew worse spirographs, but our simulation of "patient effort" did not otherwise significantly alter ratings. Drawing speed (rapid vs. slow) often resulted in different appearing spirographs but did not affect final scoring.

We would consider the ideal spirograph drawing for therapeutic studies to be 1. sensitive to score modest tremor, 2. sensitive to change and grades of severity, represented by a larger intra-rater standard deviation, 3. possess high inter-rater reliability, and 4. possess high intra-rater reliability. Our data suggests that unsupported "on the line" spirals should be used in future trials.

It is not surprising that unsupported spirals were worse than supported spirals. We can, however, only speculate on why "on the line" spirals showed the highest scores, greatest variance, and had the best inter-rater and intra-rater reliability. Keeping in mind that the final score depends on both the subject and the rater, we suspect that this may be mostly due to the rater having a constant fixed point (the traced spiral) from which to judge deviation. Conversely, a single line may place more pressure or impose a more specific task onto the subject, which could in turn worsen tremor. We do not know why supported "on the line" spirals had such better inter-rater reliability than unsupported spirals.

Potential weaknesses of our study include those associated with a tertiary referral center. However, there was actually a bias within our center toward less severe tremor because most severe ET patients have undergone VIM deep brain stimulation, and were thus excluded. This may have "normalized" the cohort to that seen in the community. We included tremor patients with dystonia to be intentionally inclusive. In our experience it is difficult to clearly segregate these two groups in the absence of overt dystonia involving a distant body part, [10] and we suspect that many patients without overt dystonia, but with irregular, jerky tremor may have some dystonic physiology. There were to few DT patients in the second cohort to make any meaningful comparisons. Despite these potential weaknesses, we feel that out data supports the use of unsupported line tracing spirals in future evaluations of ET.

Figure 1.
Sample of Different Spiral Drawing Paradigms with Dominant Hand




Legend to Figure 1.

On the line
- Supported: A-D
- Unsupported: B-D

Between the lines
- Supported: C-D
- Unsupported: D-D

Freehand
- Supported: E-D *
- Unsupported: F-D

* (start here)

Table 1.
Comparison of Tremor Spirographs "Effort"

| | Total score | Spiral Rating (ET) | Spiral Rating (DT) | *p* [a] (ET vs. DT) |
|---|---|---|---|---|
| **Dominant hand** | | | | |
| Normal | 3.76 ± 2.29 | 4.49 ± 1.85 | 2.94 ± 1.47 | 0.005 |
| Slow | 3.61 ± 2.19 | 4.11 ± 1.78 [c] | 3.05 ± 1.43 | 0.05 |
| Soft | 5.24 ± 2.23 | 5.84 ± 1.71 | 4.54 ± 1.70 | 0.01 |
| Fast | 3.73 ± 2.78 | 4.37 ± 2.31 | 3.00 ± 2.05 | 0.05 |
| **Non-dominant hand** | | | | |
| Normal | 4.31 ± 2.37 | 5.53 ± 1.70 [b] | 2.83 ± 1.15 | 0.0001 |
| Slow | 4.31 ± 1.95 | 5.20 ± 1.59 [d] | 3.24 ± 0.96 | 0.0001 |
| Soft | 5.20 ± 2.16 | 6.16 ± 1.56 | 4.08 ± 1.29 | 0.0001 |
| Fast | 4.20 ± 2.64 | 5.60 ± 1.96 | 2.54 ± 1.39 | 0.0001 |

[a] Wilcoxon Signed Rank Test
[b] Worsening with the non-dominant hand compared to dominant hand, $p<0.005$
[c] Improvement with slow writing when compared to normal writing, $p<0.05$
[d] Improvement with slow writing when compared to normal writing, $p<0.005$

Table 2.
Comparison of Different Spiral Drawing Paradigms by Handedness (Mean ± SD, *N* = 54)

| Spiral Drawing Paradigms | Dominant hand | Non-dominant hand |
|---|---|---|
| **On the line** | | |
| Supported | 2.89 ± 1.56 | 3.46 ± 1.65 |
| Unsupported | 3.39 ± 2.10 | 3.67 ± 2.19 |
| **Between the lines** | | |
| Supported | 2.83 ± 1.28 | 3.22 ± 1.53 |
| Unsupported | 3.67 ± 2.04 | 3.85 ± 2.12 |
| **Freehand** | | |
| Supported | 2.57 ± 1.59 | 3.22 ± 1.98 |
| Unsupported | 3.13 ± 1.93 | 3.35 ± 2.30 |

Table 3.
Inter-rater Reliability of Different Spiral Drawing Paradigms (Mean ± SD, *N* = 54) [a]

| Spiral Drawing Paradigms | Rater 1 | Rater 2 | Kendall's W |
|---|---|---|---|
| **On the line** | 3.35 ± 1.58 | 4.28 ± 1.59 | 0.75 |
| Supported | 3.18 ± 1.46 | 3.92 ± 1.48 | 0.40 |
| Unsupported | 3.53 ± 1.90 | 4.64 ± 1.86 | 0.80 |
| **Between the lines** | 3.39 ± 1.47 | 4.11 ± 1.66 | 0.50 |
| Supported | 3.03 ± 1.24 | 3.83 ± 1.50 | 0.41 |
| Unsupported | 3.76 ± 1.90 | 4.38 ± 1.96 | 0.25 |
| **Freehand** | 3.07 ± 1.61 | 3.48 ± 1.70 | 0.17 [b] |
| Supported | 2.90 ± 1.54 | 3.29 ± 1.61 | 0.15 [b] |
| Unsupported | 3.24 ± 1.86 | 3.67 ± 1.92 | 0.10 [c] |

[a] All Kendall's concordance coefficients, $p<0.0001$
[b] $p<0.01$
[c] $p<0.05$

Table 4.
Intra-rater Reliability of Different Spiral Drawing Paradigms (3-month, *N* = 28) [a]

| Spiral Drawing Paradigms | Both raters [b] | Rater 1 | Rater 2 |
|---|---|---|---|
| **On the line** | 0.93 | 0.84 | 0.87 |
| Supported | 0.87 | 0.67 | 0.79 |
| Unsupported | 0.88 | 0.78 | 0.89 |
| **Between the lines** | 0.90 | 0.72 | 0.93 |
| Supported | 0.74 [c] | 0.50 [c] | 0.79 |
| Unsupported | 0.91 | 0.75 | 0.91 |
| **Freehand** | 0.90 | 0.75 | 0.93 |
| Supported | 0.84 | 0.59 [d] | 0.85 |
| Unsupported | 0.90 | 0.81 | 0.88 |

[a] All Spearman rho correlations, $p<0.0001$
[b] Mean scores from both raters
[c] $p<0.01$
[d] $p<0.001$

## Key References

1. Jankovic J. Tardive syndromes and other drug-induced movement disorders. Clin Neuropharmacol 1995;18:197-214.
2. Goetz CG. Tardive dyskinesia, in Movement Disorders. Edited by Watts R, Koller W. New York, McGraw-Hill, 1997,519-526.
3. Bain P, et al., Criteria for the diagnosis of essential tremor. Neurology, 2000. 54(11 Suppl 4):S7.
4. Findley LJ, Epidemiology and genetics of essential tremor. Neurology, 2000. 54(11 Suppl 4):S8-S13.
5. Bain PG, et al., Assessing the impact of essential tremor on upper limb function. J Neurol, 1993;241(1):54-61.
6. Bain PG, et al., Assessing tremor severity. J Neurol Neurosurg Psychiatry, 1993;56(8):868-873.
7. Ondo W, et al., Gabapentin for essential tremor: a multiple-dose, double-blind, placebo-controlled trial. Mov Disord, 2000;15(4):678-682.
8. Louis ED, et al., A teaching videotape for the assessment of essential tremor. Mov Disord, 2001;16(1):89-93.
9. Fahn S, Tolosa E, Conceppcion M, Clinical rating scale for tremor, in Parkinson's Disease and Movement Disorders, J. Jankovic and E. Tolosa, Editors. 1993, Williams and Wilkins: Baltimore. 271-280.
10. Elble RJ, et al., Quantification of essential tremor in writing and drawing. Mov Disord, 1996;11:70-78.
9. Elble RJ, Diagnostic criteria for essential tremor and differential diagnosis. Neurology, 2000;54(11 Suppl 4):S2-S6.
10. Jedynak, C.P., A.M. Bonnet, and Y. Agid, Tremor and idiopathic dystonia. Mov Disord, 1991;6(3):230-236.

Poster designed by Kevin Dat Vuong, MA