

Brief-Report

Screening for Action Tremor in Epidemiological Field Surveys: Assessing the Reliability of a Semi-Quantitative, Visual, Template-Based Scale for Rating Hand-Drawn Spirals

Elan D. Louis^{1,2,3,4*}, Qianhua Zhao⁵, Haijiao Meng⁵ & Ding Ding^{5*}

¹ G.H. Sergievsky Center, College of Physicians and Surgeons, Columbia University, New York, United States of America, ² Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, United States of America, ³ Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia University, New York, United States of America, ⁴ Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, United States of America, ⁵ Institute of Neurology, Huashan Hospital, Fudan University, WHO Collaborating Center for Research and Training in Neurosciences, Shanghai, People's Republic of China

Abstract

Background: Population-based epidemiological studies of essential and other tremors have need of a rapid yet accurate means to assess tremor, especially mild tremors. Handwriting is often affected by tremor, and a hand-drawn spiral can provide investigators with objective rather than self-reported data. We present a semi-quantitative, ordinal scale to rate hand-drawn spirals. The scale, which includes values for mild tremor, is accompanied by photographic examples of spirals of each rating, providing a visual template for guidance.

Methods: This study, conducted within the framework of a population-based epidemiological study of 5,000 individuals aged 60 and older in Shanghai, asked enrollees to draw an Archimedes spiral with each hand. Spirals were rated using an ordinal scale (0, 0.5, 1, 1.5, 2, and 3). Three raters rated an initial set of 548 spirals. Four raters rated a subsequent set of 200 spirals using a visual template for guidance.

Results: Initial agreement (548 spirals) was good (r values ranged from 0.49 to 0.62, all $p < 0.001$). Subsequent agreement (200 spirals and using visual template) improved (r values ranged from 0.67 to 0.91, all $p < 0.001$).

Discussion: This tool will be useful to researchers who are attempting to rapidly assess action tremor in their field surveys.

Keywords: Essential tremor, epidemiology, rating scale, spiral, reliability

Citation: Louis ED, Zhao Q, Meng H, et al. Screening for action tremor in epidemiological field surveys: assessing the reliability of a semi-quantitative, visual, template-based scale for rating hand-drawn spirals. *Tremor Other Hyperkinet Mov* 2012;2: <http://tremorjournal.org/article/view/46>

*To whom correspondence should be addressed. E-mail: EDL2@columbia.edu; dding99@yahoo.com

Editor: Elan D. Louis, Columbia University, United States of America

Received: June 15, 2011 **Accepted:** July 7, 2011 **Published:** April 16, 2012

Copyright: © 2012 Louis 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 author(s) and source are credited; that no commercial use is made of the work; and that the work is not altered or transformed.

Funding: Dr. Louis was supported by R01 NS039422 from the National Institutes of Health. Dr. Ding was supported by 1R21AG028182-01A1 from the National Institutes of Health and 09DZ1950400 from the Shanghai Science and Technology Committee, China.

Competing Interests: The authors report no conflict of interest.

Introduction

Population-based epidemiological studies of the prevalence, incidence, and determinants of a variety of types of tremor, including essential tremor (ET),^{1–3} toxin-induced tremors,^{4–8} and enhanced physiological tremor,⁹ require a rapid yet accurate means to assess tremor. The assessment of action tremor in such settings poses practical challenges. For example, although a neurologist specializing in movement disorders is best positioned to provide the most valid measure of tremor severity, it is often not feasible for such a person to examine thousands of participants in field settings.

Portable tools for the assessment of tremor, such as the digitizing tablet,¹⁰ are available, but these have not been validated in such field studies.

Handwriting is often affected by tremor,¹¹ and a hand-drawn spiral can provide investigators with objective rather than subjective (i.e., self-reported) data. Spirals can be collected in the field and rapidly rated at a later point in time.

A critical issue is that population-based studies often involve the evaluation of individuals whose tremor is on the mild end of the spectrum.¹² Hence, standard rating scales (e.g., 0, 1, 2, 3), used in

clinical settings to assess more severe tremor, may not be of value. We have devised a semi-quantitative scale for rating spirals: the scale more precisely distinguishes between tremors in the mild end of the spectrum (e.g., ratings of 0, 0.5, 1, 1.5), which is in marked contrast to previous scales we have used that did not do so.

In this report, we present the scale and describe our methods to improve upon the application of the scale so as to maximize its reliability. The issue of reliability is important, that is, whether different raters are able to use the scale to derive comparable ratings. Indeed, demonstrating reliability is an important initial step in epidemiological research as reliability may be the only measure of data quality in situations in which validity is difficult to assess.¹³ Importantly, the scale we describe is accompanied by photographic examples of spirals of each rating, providing a visual template for guidance, and, as we show in this report, the use of the visual template improves upon reliability. We hope this tool will be of use to researchers who are attempting to rapidly assess tremor in their field surveys and populations.

Methods

Study setting

This study was conducted as part of a population-based epidemiological study, the Prevalence Study of Alzheimer's Disease and Mild Cognitive Impairment in Shanghai, China, which began in October 2009. The study, which sampled 5,000 individuals aged 60 and older in an urban community, focuses on cognitive impairment and dementia, although several other neurological disorders, including ET, will be assessed. At the time of enrollment, written informed consent was obtained from all participants and/or their legally acceptable representative. The study was approved by the ethics committee of Huashan Hospital, Fudan University, Shanghai.

Spiral drawings

Enrolled participants were asked to draw an Archimedes spiral with each hand. Spirals were drawn on a standard 8.5 × 11 inch sheet of paper using a pen while the participant was seated at a table. The paper was centered at right angles directly in front of them and held down by their other hand. The drawing hand was not allowed to rest or be supported when the spiral was being drawn. Participants started at the center of the page, without lifting their pen.¹⁴

Spiral rating: In-person training session

Two investigators (a neurologist, Q.Z., and a neuroepidemiologist, D.D.) underwent a 2-hour, in-person training session in New York with a senior movement disorder neurologist with expertise in tremor evaluation (E.D.L.). During the training session, the three individuals co-reviewed 300 hand-drawn spirals using the rating scale described below. The 300 spirals ranged in severity from normal (no tremor) to marked tremor.

Spiral rating: Initial independent ratings

Two months later, the three investigators, blinded to clinical information, independently rated spirals that were available from the first 548 enrollees. The rating scale was as follows:

0 = Absolutely no oscillations anywhere.

0.5 = Subtle, low-amplitude oscillations are present in a few spots. Oscillations are not consistently present throughout the spiral.

1.0 = Low-amplitude oscillations are present in multiple places. Examining at least one of the spiral's quadrants reveals the presence of these low-amplitude oscillations that occur in each larger and larger line of the spiral within that quadrant.

1.5 = Low-amplitude oscillations are present in multiple places and oscillations can at times reach moderate amplitude.

2 = Moderate-amplitude oscillations that are present in many areas of the spiral.

3 = Oscillations reach large amplitude in one or more places. Lines may overlap. Pen may lift off the paper.

On each spiral, the raters were careful to distinguish clear, regular oscillations from sloppiness, spatial errors, and other irregularities or movement disfluencies that were not strictly oscillatory.

Spiral rating: Second independent ratings using visual templates

To further improve on the inter-rater reliability, the senior movement disorder neurologist carefully assembled visual examples of spirals that he had rated, including two examples each of the spirals rated as 0.5, 1.0, 1.5, 2, and 3 (Figures 1–5). These spirals were electronically scanned to produce a visual template to be used as a reference guide during the rating process. The investigative team in China was asked to assemble 200 spirals with a range of scores (0–3) so that they could be independently rated by the movement disorder neurologist, the two previous investigators on the Chinese team, and a new investigator (a second neurologist) on the Chinese team (H.M.). The new investigator did not have the benefit of the initial in-person training session in New York or experience with the prior 548 rated spirals. These 200 spirals were independently rated by the four investigators.

Statistical analyses

Statistical analyses were performed using SPSS (Version 18.0). Agreement between raters in the initial set of 548 spirals and in the subsequent set of 200 spirals was assessed using Spearman's correlation coefficient (r) and intraclass correlation coefficients (ICCs). For ICCs, absolute agreement was assessed rather than consistency.

Results

In the initial set of 548 spirals, the US neurologist assigned the following ratings: 0 ($n = 5$), 0.5 ($n = 280$), 1 ($n = 208$), 1.5 ($n = 35$), 2 ($n = 17$) and 3 ($n = 3$). The agreement between raters ranged from $r = 0.49$ ($p < 0.001$) to $r = 0.62$ ($p < 0.001$), indicating good agreement (Table 1). ICC values tended to be slightly higher than r values (Table 1).

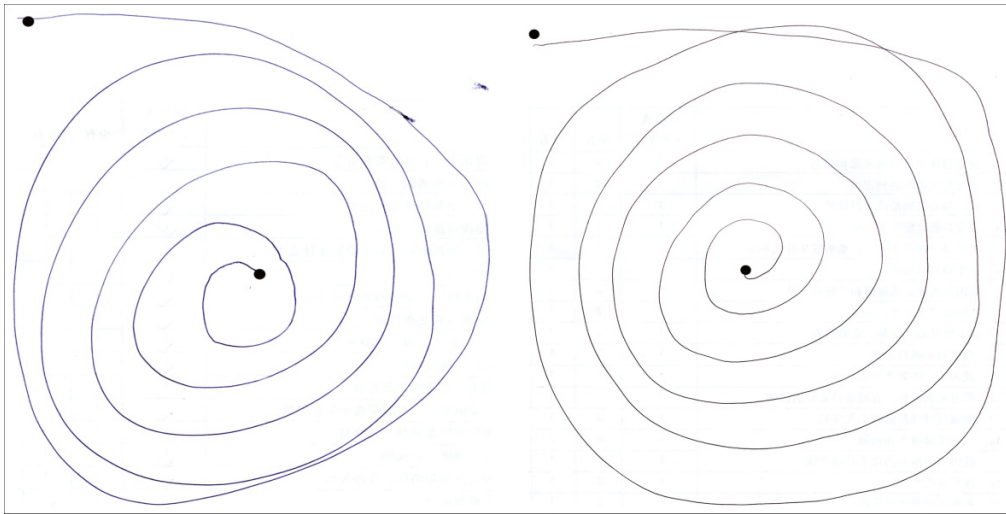


Figure 1. (A, B) Two Spirals that Received Ratings of 0.5.

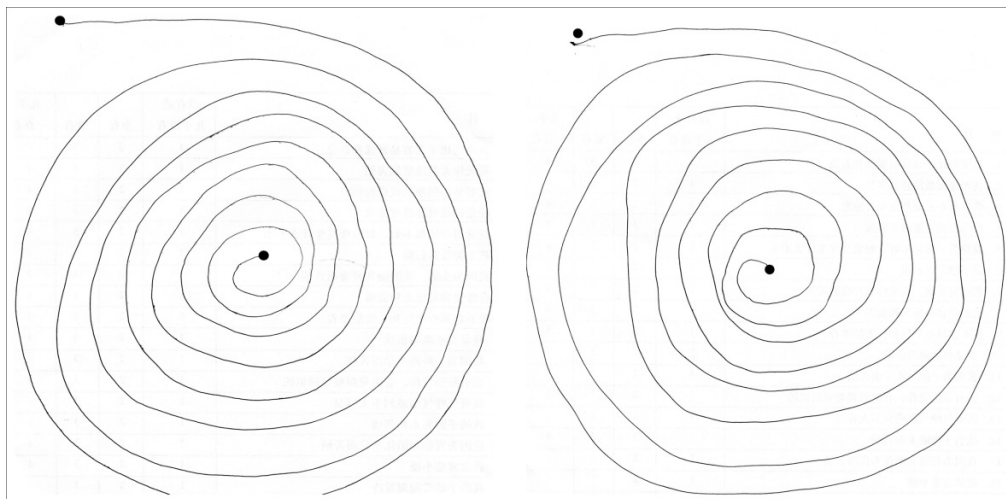


Figure 2. (A, B) Two Spirals that Received Ratings of 1.0.

In the subsequent set of 200 spirals, the US neurologist assigned the following ratings: 0 ($n = 4$), 0.5 ($n = 20$), 1 ($n = 84$), 1.5 ($n = 48$), 2 ($n = 38$) and 3 ($n = 6$). The agreement between raters ranged from $r = 0.67$ ($p < 0.001$) to $r = 0.91$ ($p < 0.001$), indicating good agreement (Table 2). ICC values tended to be slightly higher than r values (Table 2). The agreement between the raters improved when compared with the initial agreement (548 spirals). Thus, for the agreement between the US neurologist and the Chinese neurologist 1, the r value increased from 0.56 to 0.74; for the agreement between the US neurologist and the Chinese neuroepidemiologist, the r value increased from 0.49 to 0.73; and for the agreement between the Chinese neurologist 1 and the Chinese neuroepidemiologist, the r value increased from 0.62 to 0.91 (Tables 1 and 2). Chinese neurologist 2, who had received no in-person training and who only used the visual templates, had high agreement with the other

raters ($r = 0.67$ with the US neurologist; $r = 0.78$ with Chinese neurologist 1; and $r = 0.87$ with the Chinese neuroepidemiologist).

We assessed whether the agreement between raters differed with respect to the rating of spirals that were drawn with the right or left hand. Agreement was slightly higher for the left hand; for example, for the agreement between the US neurologist and Chinese neurologist 1, r (right) = 0.70 and r (left) = 0.78; for the agreement between the US neurologist and Chinese neurologist 2, r (right) = 0.60 and r (left) = 0.78; and between the two Chinese neurologists, r (right) = 0.74 and r (left) = 0.87. We also stratified ratings into low (i.e., the US neurologist assigned ratings of 0 or 0.5) vs. high (the US neurologist assigned ratings of 2 or 3) to see whether the severity of tremor affected the level of agreement. We did not find that agreement differed across these two strata.

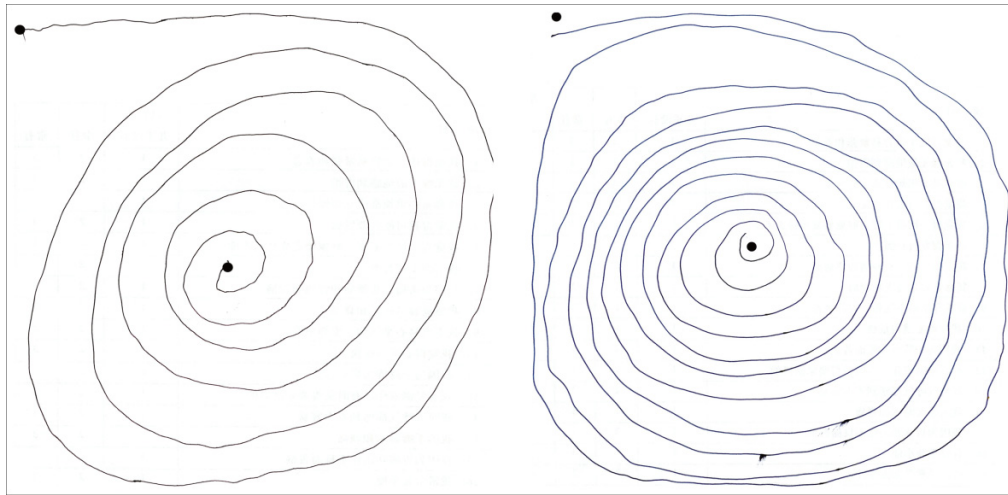


Figure 3. (A, B) Two Spirals that Received Ratings of 1.5.

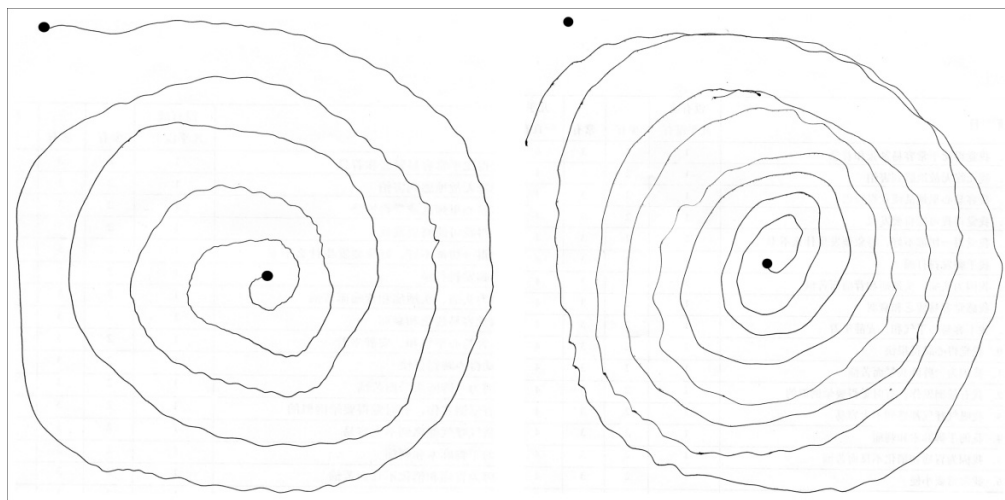


Figure 4. (A, B) Two Spirals that Received Ratings of 2.

Discussion

Movement disorder neurologists cannot practically be sent into the field to personally examine thousands of study subjects nor is it practicable to videotape thousands of neurological examinations in the field for later viewing. Hence, a screening procedure is necessary. The problem is that screening questionnaires for ET lack sensitivity, particularly for mild tremor.¹⁵ As an alternative, handwriting samples allow for the rapid collection of objective rather than self-reported data. Furthermore, for other types of tremor (e.g., toxin-induced tremors) the value of screening questionnaires is not known; empiric data on the presence and severity of tremor are of greater value. We are aware of one other example of a rating scale for action tremor that provides visual examples; however, a problem with that scale is that ratings are from 0 to 10, and this large number of scale steps has the potential to produce discrepancies in rater agreement.¹⁶

In the current study, we demonstrated that the semi-quantitative rating scale we use is reliable. Indeed, the second Chinese neurologist, who had had no in-person training and who only used the visual templates, demonstrated a high agreement with the other raters ($r = 0.67$ with the US neurologist; $r = 0.78$ with Chinese neurologist 1; and $r = 0.87$ with the Chinese neuroepidemiologist). The photographic examples of spirals of each rating provide an easy to use visual template for guidance, and this improves reliability.

This visual template-based method can be used in a variety of field settings. For example, it is often important to decide in the field whether to incorporate a more detailed assessment after an initial screening evaluation, and that decision needs to be made on the spot. Thus a screening spiral that receives a rating above a certain predetermined threshold (e.g., 1.5) might be an entrée to a second and more detailed diagnostic neurological examination on the same day of

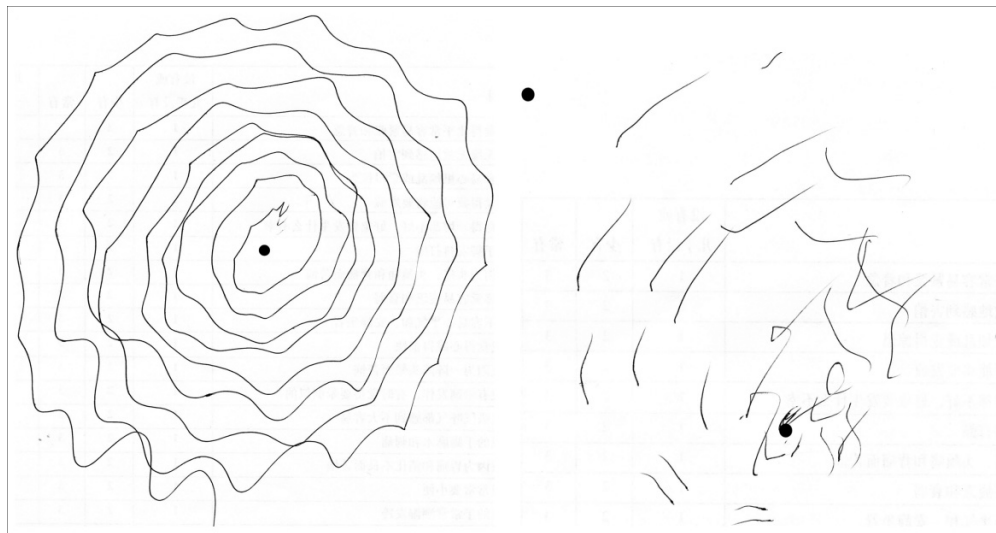


Figure 5. (A, B) Two Spirals that Received Ratings of 3.

Table 1. Inter-rater Agreement in Ratings of Tremor in the First Set of 548 Spirals

	US Rater	Neurologist 1 (China)	Neuroepidemiologist (China)
US Rater		$r = 0.56, p < 0.001$ $ICC = 0.68, p < 0.001$	$r = 0.49, p < 0.001$ $ICC = 0.49, p < 0.001$
Neurologist 1 (China)			$r = 0.62, p < 0.001$ $ICC = 0.67, p < 0.001$
Neuroepidemiologist (China)			

Abbreviations: ICC, intraclass correlation coefficient; r, Spearman's correlation coefficient.

Table 2. Inter-rater Agreement in Ratings of Tremor in Second Set of 200 Spirals

	US Rater	Neurologist 1 (China)	Neurologist 2 (China)	Neuroepidemiologist (China)
US rater		$r = 0.74, p < 0.001$ $ICC = 0.79, p < 0.001$	$r = 0.67, p < 0.001$ $ICC = 0.70, p < 0.001$	$r = 0.73, p < 0.001$ $ICC = 0.77, p < 0.001$
Neurologist 1 (China)			$r = 0.78, p < 0.001$ $ICC = 0.81, p < 0.001$	$r = 0.91, p < 0.001$ $ICC = 0.92, p < 0.001$
Neurologist 2 (China)				$r = 0.87, p < 0.001$ $ICC = 0.89, p < 0.001$
Neuroepidemiologist (China)				

Abbreviations: ICC, intraclass correlation coefficient; r, Spearman's correlation coefficient.

testing. The other important issue is that the use of these templates will ensure that field workers have calibrated their ratings against the visual examples of scores assigned by a senior movement disorder specialist with expertise in tremor. The initial screen is critical; if inaccurate, cases will either be under-ascertained or alternatively too many cases will be referred for a second, more-detailed evaluation, placing an undue burden on study resources.

We recognize that this study had limitations. First, it is important to recognize that these spirals assess action tremor but they do not allow one to definitively distinguish parkinsonian or dystonic action tremor from ET. This being said, the presence of micrographia or the absence of a consistent spiral axis¹⁷ would argue in favor of parkinsonism and dystonia. Also, if handwriting is not affected by tremor, then this spirography method will under-ascertain tremor cases. As the focus of this paper was on inter-rater agreement, we did not present data using Cronbach's alpha, which is a measure of the internal consistency of data. The current analyses used pen and paper to capture spiral data. Whether similar field data could be captured on a computer or a digitizing tablet and then rated reliably remains to be determined.

It is our hope that the tool we describe here will be of use to researchers who are attempting to rapidly screen for tremor in their field surveys and population-based studies, and will pave the way for population-based studies of tremor that utilize an objective but not burdensome initial screening process.

References

1. Benito-Leon J, Bermejo-Pareja F, Louis ED. Incidence of essential tremor in three elderly populations of central Spain. *Neurology* 2005;64:1721–1725, <http://dx.doi.org/10.1212/01.WNL.0000161852.70374.01>.
2. Dogu O, Sevim S, Camdeviren H, et al. Prevalence of essential tremor: Door-to-door neurologic exams in Mersin Province, Turkey. *Neurology* 2003;61:1804–1806.
3. Louis ED, Benito-Leon J, Bermejo-Pareja F. Population-based prospective study of cigarette smoking and risk of incident essential tremor. *Neurology* 2008;70:1682–1687, <http://dx.doi.org/10.1212/01.wnl.0000311271.42596.32>.
4. Wastesson G, Lamoureux D, Sallsten G, Beuter A, Barregard L. Quantitative assessment of neuromotor function in workers with current low exposure to mercury vapor. *Neurotoxicology* 2008;29:596–604, <http://dx.doi.org/10.1016/j.neuro.2008.03.005>.
5. Louis ED. Kinetic tremor: Differences between smokers and non-smokers. *Neurotoxicology* 2007;28:569–575, <http://dx.doi.org/10.1016/j.neuro.2006.12.006>.
6. Louis ED, Jiang W, Pellegrino KM, et al. Elevated blood harmaline (1-methyl-9H-pyrido[3,4-b]indole) concentrations in essential tremor. *Neurotoxicology* 2008;29:294–300, <http://dx.doi.org/10.1016/j.neuro.2007.12.001>.
7. Bose-O'Reilly S, Drasch G, Beinhoff C, et al. Health assessment of artisanal gold miners in Tanzania. *Sci Total Environ* 2010;408:796–805, <http://dx.doi.org/10.1016/j.scitotenv.2009.10.051>.
8. Blond M, Netterstrom B. Neuromotor function in a cohort of Danish steel workers. *Neurotoxicology* 2007;28:336–344, <http://dx.doi.org/10.1016/j.neuro.2006.07.010>.
9. Louis ED, Ford B, Pullman S, Baron K. How normal is 'normal'? Mild tremor in a multiethnic cohort of normal subjects. *Arch Neurol* 1998;55:222–227, <http://dx.doi.org/10.1001/archneur.55.2.222>.
10. Elble RJ, Sinha R, Higgins C. Quantification of tremor with a digitizing tablet. *J Neurosci Methods* 1990;32:193–198, [http://dx.doi.org/10.1016/0165-0270\(90\)90140-B](http://dx.doi.org/10.1016/0165-0270(90)90140-B).
11. Louis ED, Ford B, Wendt KJ, Lee H, Andrews H. A comparison of different bedside tests for essential tremor. *Mov Disord* 1999;14:462–467, [http://dx.doi.org/10.1002/1531-8257\(199905\)14:3<462::AID-MDS1012>3.0.CO;2-V](http://dx.doi.org/10.1002/1531-8257(199905)14:3<462::AID-MDS1012>3.0.CO;2-V).
12. Louis ED, Ford B, Wendt KJ, Cameron G. Clinical characteristics of essential tremor: Data from a community-based study. *Mov Disord* 1998;13:803–808, <http://dx.doi.org/10.1002/mds.870130508>.
13. Louis ED, Ford B, Bismuth B. Reliability between two observers using a protocol for diagnosing essential tremor. *Mov Disord* 1998;13:287–293, <http://dx.doi.org/10.1002/mds.870130215>.
14. Hafeman D, Ahsan H, Islam T, Louis E. Betel quid: Its tremor-producing effects in residents of Araihaazar, Bangladesh. *Mov Disord* 2006;21:567–571, <http://dx.doi.org/10.1002/mds.20754>.
15. Louis ED, Ford B, Lee H, Andrews H. Does a screening questionnaire for essential tremor agree with the physician's examination? *Neurology* 1998;50:1351–1357.
16. Bain P, Findley LJ. Assessing Tremor Severity. London: Smith-Gordon; 1993.
17. Louis ED, Yu Q, Floyd AG, Moskowitz C, Pullman SL. Axis is a feature of handwritten spirals in essential tremor. *Mov Disord* 2006;21:1294–1295, <http://dx.doi.org/10.1002/mds.20915>.