Movement Disorders
Vol. 14, No. 3, 1999, pp. 502–506

Brief Report

Advantages of a Modified Scoring Method for the Rush Video-Based Tic Rating Scale

*Christopher G. Goetz, MD, *Eric J. Pappert, MD, ‡Elan D. Louis, MD, MS, †Rema Raman, MS, and †Sue Leurgans, PhD

**Departments of Neurological Sciences and †Preventive Medicine, Rush University, Rush Presbyterian St. Luke's Medical Center, Chicago, Illinois; and ‡Department of Neurology and Gertrude H. Sergievsky Center, College of Physicians and Surgeons, Columbia University, New York, NY, U.S.A.*

Summary: Previously, we published a video-based objective rating scale of tics that met reliability and validity criteria for measurement of five domains of tic disability. In the original form, the scale's metric properties did not permit internal comparison of each of the five domains of impairment and did not provide a total score for use as a primary outcome measure. In this study, we retained the original scale and videotape protocol but tested whether a modified scoring system corrected these limitations. The new scoring method rated assigned tic data to ratings of 0–4 on five disability categories: number of body areas, frequency of motor tics, frequency of phonic tics, severity of motor tics, and severity of phonic tics. The sums of these ratings yielded a total score of overall tic disability (0–20). In a series of 31 patients with Gilles de la Tourette syndrome, we assessed Spearman correlation coefficients for the old and new scoring systems as well as the correlation of the new ratings with the objectively derived sections of the Yale Global Tic Severity Scale (YGTSS), another valid and reliable scale used in clinical practice and research. For each domain, the rank order for the scores on the original scale was well retained in the new scores. Likewise, for each domain, ranking with the new scoring system correlated well with scores on the comparable objective item from the YGTSS. The new total score accurately captured the rank order of the combined five domains from the original scale and correlated well with the total objective motor plus phonic tic score from the YGTSS and the YGTSS Tourette Syndrome Overall Impairment Rating. These data demonstrate that the modified videotape-based scoring system retains the essential information gathered in the original Rush scale. The modification provides comparisons among the five assessed domains and a total objectively based disability score that can be used as a single outcome measure for assessing tic disability. **Key Words:** Gilles de la Tourette syndrome—Tourette's syndrome—Tic disorders—Rating scales.

To compliment existing tic rating scales,[1,2] we previously developed a videotape-based objective rating scale for tic assessment.[3,4] This scale successfully fulfills tests for interrater reliability and temporal stability, and correlates well with scales used to assess global changes over prolonged periods. Despite its current use, the scale has two significant limitations. First, although it assesses five domains of importance to tic disorders, the scale's metric properties do not permit internal comparison of one domain with another. Most standard categorical scales currently used in movement disorders rely on 0–4 ranges for each domain to assess and internally compare the gamut of clinical dysfunction in a given disorder.[5–7] With the original tic rating system, it is not possible to compare a treatment's effect on a given domain relative to another because the number value in one domain has no implicit relationship with the same value in another. Second, the original rating system did not permit a composite or global tic score derived from the objectively derived video-based data limiting the scale's use as a primary measure of overall treatment effect. With the aim of addressing these analytical issues without sacrificing the quality of clinical assessments, we retained the

Received July 14, 1998; revisions received October 6, 1998, and January 4, 1999. Accepted January 22, 1999.

Address correspondence and reprint requests to Christopher G. Goetz, MD, Department of Neurological Sciences, Rush University, Rush Presbyterian St. Luke's Medical Center, 1725 W. Harrison St., Chicago, IL 60612, U.S.A.

video protocol and existing rating scale but restructured the scoring methods. We based the scoring modification on the dataset of the original video scale report and tested a new population of patients using both new and old rating scales.

METHODS

Rush Video-Based Filming Protocol and Rating Scale

As previously developed,[3,4] the videotape protocol involves a 10-minute film of patients placed in front of a video camera in a quiet room. Two body views are recorded, full frontal body (far) and head and shoulders only (near) under two conditions, relaxed with the examiner in the room and relaxed with the patient alone in the room. Each segment lasts 2.5 minutes. Only recordings with no examiner present are scored (5 minutes). Five domains are rated: number of body areas involved with tics, motor tic intensity, phonic tic intensity, frequency of motor tics, and frequency of phonic tics. The original scale has met psychometric criteria for interrater reliability, temporal stability, and validity.[3]

New Scoring Method Construction and Description

The original report on the reliability and validity of our video-based scale was based on data from 51 patients.[3] For the new scoring method, we maintained the five original domains of tic disability, but because most clinical scales used in movement disorders use a 0–4 range, we adopted this practice for each domain. We specifically did not rewrite the scale with new descriptive language because the original scale had already passed reliability and validity criteria. Rather, we reorganized the categories of data assignment so all data obtained during the video assessment could be assigned to a simpler 0–4 scoring format. For all domains, we retained 0 as representing normal function without evidence of tic disability. For the severity scales, because the original rating "5" was infrequently represented, we combined 4 and 5 into 4. For the two frequency domains and the body area domain, the statistician (SL) examined the mean and standard deviation identified from the original 51 patients,[3] as well as 0.5, 1.0, and 1.5 standard deviations off the mean. These numbers were delivered to the clinical neurologists (CGG and EJP) who selected cut-off numbers that were clinically reasonable and close to the statistician's calculation (Table 1). A total tic disability score was calculated by summing the values from the five domains.

Test Population

Thirty-one patients volunteered to be videotaped according to the Rush video protocol[3,8] and to be scored on the Yale Global Tic Severity Scale (YGTSS)[9] on the same day. The YGTSS is a scale used in practice and research trials and has fulfilled reliability and validity criteria. The YGTSS evaluator (EJP) had no access to the videotape data. The videotape was reviewed and scored with the original scoring system for the video-based scale by a second investigator (EDL) who had no access to the YGTSS scores. All subjects met DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition) criteria for GTS. The 28 males and three females had a tic onset, defined as the age the parent or patient first observed tics, of age 7.2 ± 2.4 years (mean ± standard deviation) and the current age was 23.4 ± 6.4 years.

Data Analysis

Original data were assigned to the new scoring method by an investigator blind to the videotapes and the YGTSS scores (CGG). Original videotape scores, the new rating scores, and the YGTSS scores were then delivered to the statistician (SL). Summary data were analyzed to present means and standard deviations or frequencies as appropriate. To assess the degree of association of various pairs of scales, Spearman rank correlation coefficients were computed.

RESULTS

Individual Domains

For each domain, the rank order for the scores on the original scale was well retained in the new scoring system. Spearman's correlation coefficient for number of body areas was 0.94, severity of motor tics 0.99, severity of phonic tics 1.0, frequency of motor tics 0.88, and frequency of phonic tics 0.98 (all $p < 0.001$). Based on Spearman's correlation coefficients for each domain, ranking on the new rating method correlated well with the scores on the comparable objective item from the YGTSS: number of body areas versus YGTSS motor tic distribution (item A5) = 0.90 ($p < 0.0001$), severity of motor tics versus YGTSS motor tic intensity (item A2) = 0.62 ($p = 0.0002$), severity of phonic tics versus YGTSS phonic tic intensity (item B2) = 0.56 ($p < 0.0012$), frequency of motor tics versus YGTSS motor tic number (item A1) = .49 ($p < 0.0057$), and frequency phonic tics versus YGTSS phonic tic number (item B1) = 0.48 ($p = 0.0066$).

In addition to high correlations with the objective measures derived by the YGTSS, Spearman's correlation coefficients for each domain correlated well with the historic (subjective) assessment category for the four domains applicable from the YGTSS: motor tic frequency (video) versus YGTSS historic motor tic frequency = 0.54 ($p = 0.0014$), phonic tic frequency (video) versus

TABLE 1. *Comparisons of original Rush Videotape Rating Scale and modified Rush Videotape Rating Scale (1998)*

Original Rush Videotape Rating Scale	Modified Rush Videotape Rating Scale (1998)
No. of Body Areas	No. of Body Areas
____ Eyes ____ Nose ____ Mouth ____ Neck ____ Shoulders ____ Arms ____ Hands ____ Trunk ____ Pelvis ____ Legs ____ Feet	0 = no body area 1 = 1 or 2 body areas 2 = 3 or 4 body areas 3 = 5 or 6 body areas 4 = 7 or more body areas
Motor Tic Frequency	Motor Tic Frequency (tics/min)
Number of tics counted and recorded as raw number	0 = no tics 1 = 1–20 tics/min 2 = 21–40 tics/min 3 = 41–60 tics/min 4 = greater than 60 tics/min
Phonic Tic Frequency	Phonic Tic Frequency
Number of tics counted and recorded as raw number	0 = no tics 1 = 1–5 tics/min 2 = 6–10 tics/min 3 = 11–15 tics/min 4 = greater than 15 tics/min
Severity of Motor Tics	Severity of Motor Tics
0 = absent tics 1 = minimal: could be normal 2 = mild: limited to a single muscle group 3 = moderate: limited to a single body part 4 = severe: involve more than one body part 5 = extreme: complex behaviors	0 = absent tics 1 = minimal: could be normal 2 = mild: limited to a single muscle group 3 = moderate: limited to a single body part 4 = severe: involve more than one body part or complex
Severity of Phonic Tics	Severity of Phonic Tics
0 = absent tics 1 = minimal: could be normal 2 = mild: single words or sounds, separated by at least one breath or 4 sec 3 = moderate: words or sounds repeated 2 or 3 times in series or single obscenities separated by at least 1 breath or 4 sec 4 = severe: words or sounds repeated four or more times in series or obscenities repeated 2–3 times in series 5 = extreme: obscenities repeated ≥4 times in series	0 = absent tics 1 = minimal: could be normal 2 = mild: single words or sounds, separated by at least one breath or 4 sec 3 = moderate: words or sounds repeated 2 or 3 times in series or single obscenities separated by at least 1 breath or 4 sec 4 = severe: words or sounds repeated four or more times in series or obscenities repeated at least 2–3 times in series

YGTSS historic phonic tic frequency = 0.38 ($p = 0.03$), motor tic severity (video) versus YGTSS historic motor intensity = 0.64 ($p < 0.0001$), and phonic tic severity (video) versus YGTSS historic phonic intensity = 0.39 ($p = 0.03$).

Total Score

The new total score accurately captured the rank order of the combined five domains from the original scale (Spearman correlation coefficient = 0.48, $p = 0.006$). Furthermore, the new total score correlated well with the total objective phonic plus motor tic score from the YGTSS (items C1 + C2) (Spearman correlation coefficient = 0.77, $p < 0.0001$). The total score also significantly correlated with the YGTSS Tourette Syndrome Overall Impairment Rating (item E) (Spearman correlation coefficient = 0.53, $p = 0.0024$).

DISCUSSION

Tics are probably the most difficult of all movement disorders to rate with validity and reliability. Clinically, they are remarkably variable[1] and encompass simple eye blinks and sniffing as well as exotic, even bizarre gestures and coprolalic tirades. Second, tics wax and wane spontaneously irrespective of treatments or interventions.[2] Third, tics are under at least partial volitional

control, and many patients with significant problems by report reveal no or minimal tics in the presence of the treating physician.[3] Our original videotape protocol and objective rating scale proved useful for dealing with these problems[3] and has been incorporated into drug trials and government-sponsored research.[4] It has not been tested for reliability and validity in multiple centers.

Our scoring modification maintains the existing scale, video recording protocol, and methods of data collection. It corrects two functional limitations of the existing rating method and expands its potential use to treatment protocols. First, we now can rate frequency, intensity, and anatomic distribution of tics on subscales that all have a comparable and standardized 0–4 range. The anchor points for each cut-off relate directly to our original patient sample, and the gamut of dysfunction for each domain captured ranges from no disability (0) to severe impairment (4). Changes induced by drugs or other treatments that affect one domain more than another can henceforth be detected because each domain's 0–4 anchor points have been prepared in a parallel manner. Our strong Spearman correlation coefficients between the new and old scales as well as the domains of the YGTSS demonstrate that we have captured the essential clinical data in the modification. Second, we can now compute a total score of overall tic severity based on the sum of the objectively based subscores (0–20). This method is used by most other scales in movement disorders[5–7] and, with the modification, allows investigators to use a single composite score of objectively derived data as the primary outcome measure of a GTS study. In developing a total score, however, we do not intend to oversimplify or obscure the implicit complexity of tic evaluations. We recognize these domains as likely distinct and not necessarily equivalent features of tic impairment, but we also face the reality that most research evaluations favor a single primary outcome that encompasses the essential items of interest. A total composite score provides this outcome measure and permits further factor analysis to identify if the five domains are independent of one another or reflective of fewer key items.

We emphasize that this report does not present a new rating scale for tic disorders. In parallel with scales for other movement disorders, there is already a movement to construct a new comprehensive rating scale for tics (Unified Tic Rating Scale).[1,10] Once the domains and anchor points of a new clinical scale are firmly established, we envision a video-based assessment of objective tic function that will be constructed in exact parallel. A similarly constructed office-based clinical assessment and the video-based objective scale will permit a clear definition of each component's use. After over a decade of experience with our video-based scale, we are aware of weaknesses that will need to be corrected in a future scale. For instance, even with the instructions and definitions provided for data collection, a single complex tic that involves a complicated sequence of involuntary movements could still be counted as several individual tics or as one. Furthermore, the intensity rating combines issues of tic magnitude, complexity, and interference which could be separated for greater precision. Factor analysis techniques and multicenter reliability studies will be needed to test the use, reliability, and validity of each item incorporated into a future scale.

Objectively derived videotape data offer several distinct advantages to the study of GTS. First, because tics are fleeting and occur simultaneously in different body regions, they are often difficult to assess completely with a single observation. Videotapes capture the tic repertoire and can be replayed to permit a complete assessment of complex tic material. Second, because tics are often suppressed in the presence of the observer, the videotape maximizes the collection of data in a relaxed setting without direct clinical scrutiny. In our prior study, the number of tics occurring in the presence of an observer was only 27% of the number documented on videotape when no observer was present during filming.[3] Third, videotape protocols have been validated, incorporated into working versions of standardized scales like the Unified Tic Rating Scale,[1,10] and are currently used in multicenter trials of GTS (Tourette Study Group, NIH #5-RO1 NS33654-03). Concerns have been expressed over the short timeframe of filming and the questionable correlation of videotape findings with overall clinical function. In recognition of these concerns and based on the study by Chappell et al.,[11] we previously modified our protocol to adopt our current 5-minute total data acquisition period.[1,4] Chappell et al. examined the temporal stability of tic counts and found that a 5-minute videotape segment provided a reliable index of the frequency of both motor and phonic tics over a 30-minute period, and highly correlated with overall tic severity evaluated with the YGTSS and the Clinical Global Impression Scale for Tourette Syndrome. Based on these several considerations, we view videotape data as a reliable and valid means of assessing tic dysfunction with the advantages of objectivity and permanent documentation. This added modification to the analytical method of handling video-based data offers a new means to use the data to provide a single overall assessment of tic disability in addition to a clear delineation of the specific impact of anatomic distribution of tics, motor tic frequency and severity, and phonic tic frequency and severity.

REFERENCES

1. Kompoliti K, Goetz CG. Clinical rating and quantitative assessment of tics. *Neurol Clin North Am* 1997;15:239–253.
2. Shapiro AK, Shapiro ES, Young JG, Feinberg TE. Measurement in tic disorders. In: Shapiro AK, Shapiro ES, Young JG, Feinberg TE, eds. *Gilles de la Tourette Syndrome,* 2nd ed. New York, NY: Raven Press, 1988:451–480.
3. Goetz CG, Tanner CM, Wilson RS, Shannon KM. A rating scale for Gilles de la Tourette's syndrome: description, reliability, and validity data. *Neurology* 1987;37:1542–1544.
4. Goetz CG, Stebbins GT, Thelen JA. Talipexole and adult Gilles de la Tourette's syndrome: double-blind, placebo-controlled clinical trial. *Mov Disord* 1994;9:315–317.
5. Fahn S, Elton RL, members of the UPDRS development committee. Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Calne DB, Goldstein M, eds. *Recent Developments in Parkinson's Disease,* vol 2. Florham Park, NJ: Macmillan Health Care Information, 1987:153–164.
6. Goetz CG, Stebbins GT, Shale HM, et al. Utility of an objective dyskinesia rating scale for Parkinson's disease: inter- and intrarater reliability assessment. *Mov Disord* 1994;9:390–394.
7. Kurlan R, Riddle M, Como P. *Tourette Syndrome Unified Rating Scale.* Bayside, NY: Tourette Syndrome Association Inc, 1988.
8. Pappert EJ, Goetz CG, Louis ED, Blasucci L, Leurgans S. Longitudinal objective assessment of Gilles de la Tourette syndrome. *Neurology.* In press.
9. Leckman JF, Riddle MA, Hardin MT, et al. The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. *J Am Acad Child Adolesc Psychiatry* 1989;28:566–573.
10. Goetz CG, Tanner CM, Stebbins GT, Leipsig G, Carr WC. Adult tics in Gilles de la Tourette syndrome. *Neurology* 1992;42:784–788.
11. Chappell PB, McSwiggan-Hardin MT, Scahill L, et al. Videotape tic counts in the assessment of Tourette's syndrome: stability, reliability, and validity. *J Am Acad Child Adolesc Psychiatry* 1994; 33:386–393.