N33-10118

14. The Use of a Battery of Tracking Tests in the Quantitative Evaluation of Neurological Function*

B. S. REPA, J. W. ALBERS, A. R. POTVIN, AND W. W. TOURTELLOTTE
The University of Michigan

The potential of tracking tasks for use in clinical applications has been recognized for many years. While investigators have demonstrated the usefulness of these tasks in drug research and in measuring the performance of pathological subjects, (e.g., Stark and Iida (ref. **1)** and Angel et al. (ref. **2)** few have made effective use of tracking measures in clinical trials. At the University of Michigan's Neurology Research Laboratory, a tracking test battery has been applied in a drug trial designed to compare the efficacy of L-DOPA and amantadine to that of L-DOPA and placebo in the treatment of **28** patients with Parkinson's disease. The drug trial provided an ideal opportunity for objectively evaluating the usefulness of tracking tests in assessing changes in neurologic function.

Evaluating changes in patient performance resulting from disease progression and controlled clinical trials is of great importance in establishing effective treatment programs. Clinicians are usually able to classify a given neurologic function of the patient into broad categories such as supernormal, normal, and abnormal (mild, moderate, or severe); but they often have difficulty in detecting small but significant changes in the patient's function over time. One attempt at a more objective and quantitative neurologic examination was initiated at The University of Michigan Medical Center several years ago by Dr. Wallace W. Tourtellotte. Investigators have long known that the total performance capabilities of an individual cannot be specified on the basis of a single performance test. For this reason, Dr. Tourtellotte devised a battery of sensory and motor performance tests which are now collectively referred to as the Clinical Quantitative Neurological Examination (CQNE) (table **1).**

TABLE **1.**—*The Clinical Quantitative Neurological Examination (CQNE) Test Items*

- Vision: Visual acuity
- Upper Extremities—
  - Strength of movements:
    - Grip
    - Wrist dorsiflexion
    - Shoulder abduction
  - Control of movements:
    - Steadiness
      - Hole steadiness, supported and unsupported
      - Force steadiness, supported and unsupported
      - Finger tremor, resting and sustension
    - Simple reaction time
    - Speed of hand
    - Speed-coordination of hand
    - Rotary pursuit
    - Finger dexterity
      - Purdue Pegboard
      - Pencil rotation
  - Fatigue of movements:
    - Grip strength
    - Speed of hand
    - Speed-coordination of hand
  - Sensation:
    - Touch, hand
    - Vibration sense, index finger
    - Position sense
    - Two-point discrimination

Considerable experience has been gained with the CQNE as it has been used in a number of studies using asymptomatic subjects to obtain

* This research effort has been supported in part by NIH training grant 5 to 1 GM01289-07, 1970–71, NASA contract NSr 23–005–364, and an equipment grant from the University of Michigan Institute of Science and Technology.

normative data and in several therapeutic trials involving multiple sclerosis and Parkinson's disease patients (Tourtellotte et al. (ref. **3)** and Kuzma et al. (ref. **4)).**

As part of a continuing effort to improve the clinical testing program, a tracking test battery was studied as a possible source of future tests for inclusion in the CQNE. To be effective in a clinical environment, the battery had to provide measurements that required minimum run lengths and the fewest trials possible to establish stable parameter estimates. Extensive training time was a luxury that simply could not be afforded. Furthermore, all tests and measures had to allow the use of on-line data reduction schemes. The battery that was selected is very similar to the one described by Jex and Allen at last year's meeting (ref. *5);* it includes step tracking, random tracking, and critical tracking (table **2).** The tests were kept as simple, yet as comprehensive and challenging, as possible. Modifications in the display screen and control stick were necessary to accommodate patients with various sensori-motor disabilities.

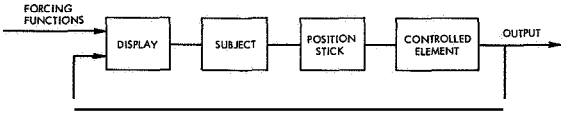
Whenever new quantitative tests for measuring neurologic disorders are developed it is of interest to examine the performance of normal subjects as well as patients on the tests. Reliabilities and learning effects are more effectively measured with normal subjects due to the possibility of large variations in patients' performance which can be justifiablyattributed to their pathological condition. Ten age-matched normals were used in a test-retest study to determine reliability measures for the tracking battery (table **3).** All reliability coefficients were found to be significant at or above the **5** percent level with the exception of movement time for a right to left transition, and the coefficient for this test

TABLE 3.—*Reliability of Tracking Test Battery Involving 10 Matched Normals With a 3 Week Interval Between the First and Second Examinations*

Test	r [1]	$\frac{2r}{1+\lvert r\rvert}$ [2]
Step tracking		
Reaction time, right to left	[3] **0.75**	**0.86**
Reaction time, left to right	[4] **.82**	.90
Movement time, right to left	**.60**	**.75**
Movement time, left to right	[3] **.67**	.80
Random tracking		
Integral of absolute error	[5] **.91**	**.95**
Critical tracking		
Reciprocal of critical root	[5] **.96**	.98

[1] Pearson product moment correlation coefficient.
[2] Spearman-Brown split-half correlation formula.
[3] $p \leqq .05$. [4] $p \leqq .01$. [5] $p \leqq .001$.

TABLE 2.—*General Tracking Task Descriptions*



Type of tracking	Forcing function	Display	Controlled element	Performance measures
Step	**Rectangular pulse with alternating ± 14 centimeter amplitude and pulse width from 2.7 to 5.7 sec.**	**Pursuit**	**K**	**Reaction time, movement time**
Random	**Random noise with cutoff frequency of 1.0 rad/sec.**	**Compensatory**	**K**	**Integral of absolute error**
Critical	**None**	**Compensatory**	$\frac{K}{S-\lambda}$; $\lambda_0 = 1.0$ rad/sec $\lambda = 0.05$ rad/sec	**Reciprocal of critical root**

TABLE 4.—*Learning in Tracking Test Battery Involving 10 Matched Normals With a 3 Week Interval Between the First and Second Examinations*

Test	Exam I Mean	Exam I SD	Exam II Mean	Exam II SD	Difference	% Change [1]	*t*-Difference	*t*-% Change
Step tracking [2]								
Reaction time, right to left	308	44	305	43	−3	−.3	.27	.10
Reaction time, left to right	297	37	297	51	0	−.2	.00	.07
Movement time, right to left	510	80	493	106	−17	−2.7	.62	.53
Movement time, left to right	596	109	530	116	−66	−10.5	2.29	2.41 [4]
Random Tracking [3]								
Integral of absolute error	1.93	.54	1.89	.52	−.04	−1.4	.68	.35
Critical tracking [2]								
Reciprocal of critical root	371	56	361	61	−10	−2.6	1.77	1.79

[1] $\%\ \text{Change} = \frac{1}{10}\sum_{i=1}^{10}\frac{\text{Score 2i} - \text{Score 1}_i}{\text{Score 1}_i} \times 100.$

[2] Units are in milliseconds.

[3] Units are in $\frac{\text{centimeter-seconds}}{\text{second}}$

[4] $= p \leqq .05$.

measure just barely missed the cutoff point. The same group of 10 normals was used to measure learning effects(table **4).** Although all test scores showed an improvement on the aecond exam, none of the improvements were statistically significant.

Another important reason for using normal subjects on new tests is to establish normative performance levels. Since it is the goal of the physician to bring the performance of patients to the predisease level, it is meaningful to express patient data as a percentage of that obtained from matched normal controls. This was done for the drug study, which used a randomized, double-blinded, crossover design. The 28 Parkinson's disease patients were randomly assigned to two groups, the first group receiving L-DOPA +amantadine first and L-DOPA+placebo second, with the second group receiving just the opposite schedule (table *5).* Treatment groups were combined for analysis, and scores were expressed as a percentage of matched normal levels (table **6).** Relative to the normal subjects, patients performed better on step tracking and critical tracking than on random tracking. Improvements were modest when amantadine was taken in addition to L-DOPA.

TABLE 5.—*Experimental Paradigm*

Group	No. of patients	Medication taken during week 1–3	4–6
1	14	*L−D+A* *	*L − D+P*
2	14	*L−D+P*	*L − D+A*

* *L − D=L-DOPA.* *A* =Amantadine. *P* =Placebo.

Based on other tests administered in the drug trial, the effect of adding amantadine to L-DOPA was found to be beneficial but weak. Thus, while the trends in the CQNE scores favored L-DOPA+ amantadine, only grip strength, hand coordination, pencil rotation, 2-point discrimination, and resting tremor showed statistically Significant changes. The tracking test measures all showed improvements favoring the L-DOPA+amantadine treatment group (table **7).** The critical task measure and left to right movement time showed improvements significant at the *5* percent level. While changes in random tracking scores and right to left reaction time scores were 10 percent or more, large variations in scores among patients prevented these changes from being statistically significant. The tracking

TABLE 6.—*Performance of Patients in the Tracking Test Battery Expressed as a Percentage of Matched Adult Normal Function*

Test	Matched adult normal function Mean ± 2SD	Patients on placebo %	Patients on placebo SD	Patients on drugs %	Patients on drugs SD
Step tracking [1]					
Reaction time, right to left	303 ± 78	83	19	90	22
Reaction time, left to right	294 ± 67	83	17	86	18
Movement time, right to left	489 ± 220	78	22	80	20
Movement time, left to right	568 ± 234	76	22	84	23
Random tracking [1]					
Integral of absolute error	1.8951.16	61	23	65	17
Critical tracking [2]					
Reciprocal of critical root	362 ± 128	78	17	81	18

[1] Units are in millisec.

[2] Units are in $\frac{\text{centimeter-sec}}{\text{sec}}$

TABLE 7.—*Results of Tracking Test Battery Involving 28 Parkinson Patients: Comparison Between L-DOPA+Placebo and L-DOPA+Amantadine Treatment Groups*

Test	L-DOPA + amantadine Mean	L-DOPA + amantadine SD	L-DOPA + placebo Mean	L-DOPA + placebo SD	Difference	% Change [1]	*t*-Difference	*t*-% Change
Step tracking [2]								
Reaction time, right to left	359	91	385	102	27	10	1.61	2.32 [4]
Reaction time, left to right	358	84	368	81	10	4	.75	1.32
Movement time, right to left	642	145	679	215	4	7	1.18	1.46
Movement time, left to right	717	191	820	289	10	16	2.32 [4]	2.89 [5]
Random tracking [3]								
Integral of absolute error	3.04	.74	3.36	1.35	.32	11	1.42	1.66
Critical tracking [2]								
Reciprocal of critical root	463	96	486	110	22	5	2.23 [4]	2.58 [4]

[1] $\%\ \text{Change} = \frac{1}{28}\sum_{i=1}^{28}\frac{\text{Score } 2_i - \text{Score } 1_i}{\text{Score } 1_i}$.

[2] Units are in millisec.

[3] Units are in $\frac{\text{centimeter-sec}}{\text{sec}}$

[4] $p \leqq .05$. [5] $p \leqq .01$.

measures still appeared to be at least as sensitive as most measures in the CQNE in detecting changes in performance.

The final selection of a test for inclusion in the CQNE depends upon the test's satisfactory fulfillment of a number of criteria relating to the neurological function being tested, instrumentation, test data, subject requirements, and examiner-requirements (table 8). The direct application of tracking tasks to a clinical trial has shown that they are indeed capable of satisfying these criteria. In closing, it should be mentioned that tracking tasks and other quantitative testing procedures are not meant as a substitute for

TABLE 8.—*Criteria for Test Selection*

Criteria related to the neurological function tested:
The function must relate meaningfully to the status of the subject's nervous system.
Criteria related to the instrument:
The instrument must be small and capable of being used in a small area.
Initial, operating, and maintenance costs of the instrument must; not be prohibitive.
Criteria related to the test data:
The data must be truly quantitative, i.e., at least of interval strength.
The data must be objective, i.e., reliable.
The data must be sensitive enough to detect changes in the neurological function being evaluated.
Criteria related to the subject:
The "supernormal" healthy young adult should be challenged by the test, and yet at the same time the test should not be beyond the ability of the patient.
The subject should be reasonably interested and motivated by the test.
Learning effects should be at a minimum.
The subject must not be *so* fatigued by the test as to prohibit the completion of succeeding tests in the battery.
The idea of the test must be simple enough to be easily communicated to the subject.
Criteria related to the examiner:
A trained physical therapist must be capable of administering the test.

sound clinical judgement, but they do provide the medical investigator with information that is often impossible to obtain from observation alone, particularly in detecting and documenting changes in a patient's condition.

REFERENCES

1. STARK, L.; AND IIDA, M.: Dynamical Response of the Movement Coordination System of Patients with Parkinson Syndrome. Quarterly Progress Rept. No. 63, Research Laboratory of Electronics, MIT, Oct. 15, 1961, pp. 204–213.
2. ANGEL, R. W.; ALSTON, W.; AND HIGGINS, J. R.: Control of Movement in Parkinson's Disease. Brain, vol. XCIII, pp. 1–14.
3. TOURTELLOTTE, W. W.; HAERER, A. F.; SIMPSON, J. F.; KUZMA, J. W.; AND SIKORSKI, J.: Quantitative Clinical Neurological Testing. I. A Study of a Battery of Tests Designed to Evaluate in Part the Neurological Function of Patients With Multiple Sclerosis and Its Use in a Therapeutic Trial. N.Y. Acad. Sci., vol. 122, 1965, p. 480.
4. KUZMA, J. W.: TOURTELLOTTE, W. W.; AND REMINGTON, R. D.: Quantitative Clinical Neurological Testing. 11. Some Statistical Considerations of a Battery of Tests. J. Chron. Dis., vol. 18, 1965, pp. 303–311.
5. JEX, H. R.; AND ALLEN, R. W.: Research on a New Human Dynamic Response Test Battery. Presented at Sixth Annual Conference on Manual Control, Air Force Institute of Technology, Wright-Patterson Air Force Base, Apr. 7–9, 1970.