Commentary

Quantitative evaluation of neuropharmacological trials

A. R. Potvin, Ph.D., J. W. Albers, M.D., Ph.D., B. S. Repa, Ph.D., W. G. Henderson, Ph.D., J. E. Walker, M.D., R. F. Stribley, M.S., R. W. Pew, Ph.D., and W. W. Tourtellotte, M.D., Ph.D. *Los Angeles, Calif., Arlington, Tex., Ames, Iowa, and Ann Arbor, Mich.*

Department of Neurology at Veterans Administration Wadsworth Hospital and University of California at Los Angeles, Biomedical Engineering Program at University of Texas at Arlington, Department of Biostatistics at University of Iowa, Ames, and Bioengineering Program at University of Michigan, Ann Arbor

The clinician's ability to classify a patient's neurophysiological deficits into such categories as mild, moderate, and severe does not readily allow the detection of small but significant changes in function over time. Clinical neurologists, for example, are discovering that the standard neurological examination is not sufficient to critically evaluate new drugs and surgical techniques that purport to improve the performance of patients in the activities of daily life.[1] As a result, the quantitative evaluation of the functional capacity of patients with

nervous system disorders has been receiving increased attention during the past few years. Clinicians are attempting to augment the standard neurological examination by assembling batteries of quantitative tests that measure the sensory, psycho-

---

Special abbreviations used

QENF: Quantitative Examination of Neurological Function

CQNE: Clinical Quantitative Neurological Examination

SADLE: Simulated Activities of Daily Living Examination

NPE: Neuro-Psychological Examination

---

motor, and cognitive performance of patients in clinical trials.

Most test batteries in use today suffer serious deficiencies. The tests, their measures, and their administration lack uniformity. The reliability and validity of the

test measures are often not evaluated. Control groups, when used, lack uniformity and are often poorly matched to patient groups. Methods for motivating individuals to perform in all tests to the best of their ability are often not reported. The contribution of learning improvements in clinical trials as a result of repeated testing is usually unknown. For these reasons, critical evaluation and comparisons of the results of most of these studies are qualified.

One attempt to develop an objective quantitative battery of tests capable of measuring small changes in the neurological function of patients involved in neuropharmacological trials was initiated more than a decade ago by Tourtellotte and his associates. It involved a team of neurologists, pharmacologists, biomedical engineers, psychologists, biostatisticians, programmers, and paramedical personnel. The purpose of this report is to describe the long-term development, evaluation, and use of this large-scale test battery, called the Quantitative Examination of Neurological Function (QENF). It is our belief that evaluations of the QENF, summarized herein, represent the minimal criteria that a quantitative test battery should meet before being used to determine the efficacy of a neuropharmacological trial.

The QENF evolved in several stages during the past decade. Each stage was marked by test development, test evaluation, and test use in a neuropharmacological trial. In this report, each stage (early development, introduction of steadiness and tracking tasks, and expansion of the present QENF) is presented sequentially with temporal overlaps noted.

Early development

Tourtellotte was among the first to quantitate the functional capacity of patients by assembling a battery of sensory and motor performance tests, called the Clinical Quantitative Neurological Examination (CQNE).[15, 36, 37] The purpose of the CQNE was to obtain, whenever possible, quantitative measures of the neurological func-

tions that are evaluated qualitatively in the standard neurological examination. The CQNE was specifically designed to quantitatively measure many of the neurological functions that comprise the complex symptoms of multiple sclerosis.

In a study conducted by Ivers and Goldstein,[12] the most frequent symptoms of multiple sclerosis were found to be retrobulbar neuritis, diplopia, paresthesia, weakness or incoordination of the upper and lower extremities, and gait difficulty. Impairment of the upper or lower extremities is often the result of disabilities such as contractures, weakness, excessive or premature fatigue, incoordination, or paresthesia. These disabilities may act alone or in combination. Prehensibility (the grasping function), representing the most important function of the upper extremities, requires adequate strength of the grip, wrist, and shoulder; stamina; speed and coordination of movements; and proprioceptive and protective sensation. Ambulation, representing the most important function of the lower extremities, requires adequate strength of hip and foot, stamina, speed and coordination of movements, and proprioceptive and protective sensation.

These functions were emphasized in the selection of tests and the development of the CQNE. The initial battery of the CQNE tests measured the following neurological functions: visual acuity; two-point discrimination of the index finger; vibration sense of the finger and toe; station with eyes open and closed; speed of movement of the hand and foot; coordination of the hand; strength of various muscle groups; and fatigue measures of the speed, coordination, and strength tests.

Kuzma and associates[12, 14-16, 37] designed a series of experiments to examine this test battery statistically. Much of their work involved evaluating the reproducibility of the test battery using normal young adult subjects under various testing conditions (different examiners, different times of the day, different days, and different weeks). The results indicated good inter- and intra-examiner reproducibility

once the examiners (two clinicians and two physical therapists) were carefully trained. These results, supported by a later study,[8] indicate that paramedical personnel (in this case, physical therapists) are as capable as clinicians in administering the tests.

The results of this evaluation of the test battery had important implications. The clinician's time with his patients is not significantly increased since paramedical personnel administer the tests and analyze the data. Administration of tests by paramedical personnel ensures a more perfect "blinding" of a study since paramedical personnel are less likely to detect telltale effects of a treatment than a clinician.

In addition to the reproducibility studies, Kuzma examined learning effects and fatigue effects using young normal adults and multiple sclerosis patients. From the results of the learning studies for each group, it was concluded that, except for hip strength, there were no significant improvements when the tests were administered on consecutive days. From the results of the fatigue studies, it was found that the measures for grip strength fatigue, hip strength fatigue, hand speed fatigue, and foot speed fatigue were meaningful measures of fatigue effects and that the measures were reproducible.

This initial test battery, with minor modification, was used in part to obtain data from multiple sclerosis patients for evaluating two therapeutic trials[36, 37] and the National Cooperative Study.[10, 31, 38] The results for each trial were in agreement with subjective methods for assessing changes in neurological function. The major conclusion of these early studies is that the CQNE test battery, when administered by trained paramedical personnel, especially registered physical therapists, provides a quantitative and reliable assessment of neurological function.

Development of steadiness and tracking tasks

During the past six years, Biomedical Engineers[21] have developed sub-batteries of tests measuring steadiness and tracking behavior for inclusion in the CQNE. In these tests, the subject is required to perform continuous limb movements graded along various spatial and temporal dimensions. Other tests in the CQNE do not adequately capture many of the basic features of the spatial-temporal organization of movements characteristic of many neurological disorders. The movements required in a steadiness or tracking task are a compromise between highly restricted movements that make response processes difficult to study and totally unrestricted ones that result in recording and scoring difficulties. Furthermore, the standard steadiness and tracking apparatus, described below, provides complete access to the stimulus and response records for both on-line and off-line analysis and quantitative as well as qualitative interpretations.

Steadiness tasks. Albers and associates[1-5] developed a sub-battery of tests for the CQNE to measure one component of neurological function, steadiness, or tremor. They demonstrated the applicability of a human operator control system, consisting of a constant force tracking task using a force stick with a visual display and auditory feedback. The subject was required to maintain a constant muscle tension on the force stick using only the index finger. The studies conducted by Albers and his associates were fundamental in later developing both steadiness and tracking tasks for clinical use.

Clinical steadiness tests were first developed in 1970 for use by paramedical personnel.[18] The clinical instrumentation was also used with an accelerometer to obtain quantitative measures of resting, sustention, and static intention tremor. To complete the sub-battery of steadiness tests, a hole steadiness test was modified for clinical use. A study involving 60 normal adult subjects, 10 patients with multiple sclerosis, and 10 patients with Parkinson's disease, was designed to evaluate the steadiness tests. The test measures were found to be reliable. No significant differences in performance were found on the tests

between the dominant and nondominant body sides. Males, however, were shown to be less steady than females. Steadiness was found to decrease significantly with increasing age in tests performed with the arm supported and to remain invariant with increasing age in tests performed with the arm unsupported.

The clinical instrumentation was also used in part to evaluate the efficacy of amantadine (Symmetrel) in parkinsonian patients who were taking levodopa. Results indicate that with amantadine amplitude measures of steadiness improve from 65% to 89% of the normal function of age-matched controls.

To measure a different type of steadiness, Stribley and associates[34, 35] developed a test battery of postural stance using a force platform. Twenty normal young adult subjects were recruited to evaluate measures of narrow stance, wide stance, and one-leg stance with and without support. No significant differences were found between male and female subjects or between dominant and nondominant sides in the one-leg stances. Significant differences were found between eyes-open and eyes-closed trials. As a result of these early studies, the postural stance tasks are presently being modified for use with CQNE.

Tracking tasks. More recently Repa and associates[27-30] developed a battery of tracking tasks for inclusion with the CQNE. The battery is similar to the one described by Jex and Allen[13]; it includes step tracking, random tracking, and critical tracking. Ten normal subjects, age-matched to parkinsonian patients involved in a therapeutic trial, were used in a test-retest study (one month test-retest interval) to determine reliability measures and learning effects. All test measures were found to be reliable. Although all test scores showed an improvement on the second examination, these improvements were not statistically significant. This tracking task battery was also used as a part of a neuropharmacological trial to evaluate the

efficacy of levodopa plus amantadine. The test measures, all showing improvements favoring levodopa plus amantadine, appear to be at least as sensitive as most measures in the QENF for detecting changes in performance. Potvin and associates[22] are presently modifying this tracking task battery for clinical use.

D-*Amphetamine trial.* The sub-batteries of upper extremity steadiness and tracking tasks were used to determine the effects of 10 mg of d-amphetamine and a placebo on motor performance in 6 normal subjects.[4, 6] Medication was administered in a random double-blind crossover design on two occasions, one week apart. No significant differences between the effects of the medication on resting tremor, sustention tremor, or precision hole steadiness were found. However, several compensatory tracking tasks that required sustained concentration and motor coordination were significantly improved with d-amphetamine. The results suggest that the tracking tasks were sufficiently sensitive to detect changes in performance attributable to d-amphetamine that were so small that the subjects themselves could not tell whether they had received d-amphetamine or placebo.

Extension of CQNE and development of QENF

Expansion of the CQNE owes much to Fleishman and his associates,[7, 8] who, from the 1940's onward, administered more than 200 different tests to thousands of basic trainee airmen. The methodology that Fleishman established and the results that he obtained together have served as a guide in modifying and extending the test battery.

The present test battery includes psychomotor tests that are intended to measure many of the traits found by Fleishman. These tests include: rotary pursuit; simple reaction time, Purdue pegboard; large peg, small peg, and pencil rotation; arm supported and arm unsupported steadiness

using (1) a hole steadiness apparatus and (2) a force control stick; resting, sustention, and static intention tremor using an accelerometer; hand speed; foot speed; hand coordination; and foot coordination.

Psychomotor tests represent relatively unpracticed tasks that attempt to measure abilities or traits that determine an individual's limit of performance in a skilled task. An individual's abilities are thought to be dependent upon long-term learning and genetic factors and not likely to change once he reaches adulthood.[9] Abilities differ from skills in the sense that skills are highly practiced tasks. A skill is thought to be dependent upon abilities.

The underlying assumptions relating skills to abilities measured by psychomotor tests remain unproved. However, if a patient involved in a clinical trial claims that he is now better able to button buttons, or to pick up coins, or to put on a shirt, it is important that his claim be verifiable in an objective manner. Consequently, a series of timed tests that simulate simple skills of daily living has been included in the test battery. The Simulated Activities of Daily Living Examination (SADLE)[20] presently includes the following tests: putting on a shirt, managing visible buttons, zipping a garment, putting on gloves, dialing a telephone, tying a bow, manipulating safety pins, handwriting, picking up coins, unwrapping a Band-Aid, squeezing toothpaste, cutting with a knife, using a fork, and arising from a chair.

The SADLE makes no attempt to separately measure any basic components or abilities, but focuses on simple motor skills. For example, one timed test, which requires the subject to unbutton and button a garment as rapidly as possible, reflects a compound measure of the subject's reaction time, his speed and coordination in moving to the button, his steadiness, and his fine finger manipulative abilities. The measures from the SADLE provide little or no information as to the nature of improvement in a clinical trial. From the results of the examination, it cannot be determined if improvements are due to changes in strength, reaction time, coordination, mental state, sensation, or other basic components. Nonetheless, the SADLE measures specifically those functions that are of great importance to the patient. The CQNE and SADLE complement each other. The SADLE measures a patient's capability to carry out his activities of daily living. If this capability is altered in a clinical trial, the CQNE describes which basic components have caused the change.

In addition to new psychomotor and SADLE tests, a battery of tests called the Neuro-Psychological Examination (NPE) has been assembled by Smith[33] for use with the CQNE and SADLE. These tests attempt to assess specific higher and lower cerebral functions. They include the Peabody picture vocabulary test, a color naming and recognition subtest from the Eisenson test for aphasia, a double simultaneous (face-hand) simulation test, a discrimination test of right-left and body parts, the Raven Colored Progressive Matrices test, a written and oral digit substitution test, and selected Wechsler Adult Intelligence Scale subtests (digit span, digit symbol, similarities, picture completion, picture arrangement, and object assembly).

Since the CQNE, SADLE, and NPE all purport to evaluate neurological function quantitatively, the term *Quantitative Examination of Neurological Function* (QENF) is used to describe the three examinations collectively. The QENF presently includes tests for many of the functions of the intact nervous system. With the wide variety of tests in the QENF, it is possible to select certain tests for a particular neuropharmacological trial while eliminating others that are not relevant.

Recent evaluations of QENF

The recent expansion of the QENF required extensive evaluation. A series of studies was designed to examine the effects of motivation, short-term and long-term

learning, sex, and age on performance. In addition, reliability of the test measures, various types of performance measures, and data reduction techniques were examined. A review of these studies is found below.

Motivation and short-term learning. A study of the QENF was designed to: (1) investigate the success that clinicians and paramedical personnel achieve in uniformly and adequately motivating subjects, (2) compare two different kinds of normal control groups that are often used for evaluating therapeutic trials, and (3) assess the effects of short-term learning (one-half hour test–retest interval) among patients and normal subjects on successive examinations.[19, 23, 24]

Twenty normal young adult subjects, 20 normal older adult subjects, and 20 patients (10 with Parkinson's disease and 10 with multiple sclerosis) participated in the study. Each subject group was divided into two subgroups, a control subgroup and an incentive subgroup. Substantial monetary incentives and additional verbal encouragement (in addition to that ordinarily given by the examiner) given to the incentive subgroups but not to the control subgroups were found insufficient to differentially improve the performance of the matched subgroups on a subset of the tests in the battery. There were no important differences in motivation between older adult normal subjects related to patients and young adult normal subjects unrelated to patients and paid to volunteer their services. These results indicate that clinicians and paramedical personnel are capable of motivating subjects uniformly and adequately.

Analysis of the control subgroup data for short-term learning effects revealed that young adult normal subjects showed the largest improvement in performance; older adult normal subjects showed less improvement; and patients with Parkinson's disease and multiple sclerosis showed essentially no improvement. From these results, it is doubtful that appreciable improvements in the performance of patients

in clinical trials can be attributed to learning.

Reliability, long-term, and short-term learning effects. A study[25] was designed to evaluate the reliability of the test measures in the QENF in a manner consistent with the administration of the tests in clinical trials. A one-month test-retest interval was chosen. Normal young adult subjects were used since large variations in patients' performance could justifiably be attributed to disease progression and to exacerbations. Alternatively, the large variation in patient performance could artificially inflate reliability coefficients. Twenty normal young adult subjects were evaluated once in all tests of the QENF and one month later were contacted for re-evaluation. At the completion of the second examination, the subjects repeated a subset of tests to evaluate differences between long-term (one month test-retest interval) and short-term learning (one-half hour test-retest interval).

The results of the study indicate that: (1) most test measures in the test battery are reliable, and (2) there are no important differences between short-term and long-term learning effects. Of 72 tests evaluated, the tests for fatigue, error, vibration, using a fork, and zipping a garment were found to be unreliable. With respect to learning effects, the largest improvements were found for rotary pursuit followed by many of the tests of coordination, particularly those requiring manual or finger dexterity. Only slight changes were found for tests of vision, strength, and speed of hand and foot.

When considered with previous studies, the present results lend support to the hypothesis that appreciable improvements in the performance of patients in neuropharmacological trials cannot be attributed to learning effects.

Performance measures. Potvin[19] reviewed methods for expressing the performance of patients in neuropharmacological trials and also methods of data reduction and concluded that the performance of pa-

tients should be expressed as a percentage of age-sex–matched normal function. In analyzing the results of a neuropharmacological trial, the clinician is more interested in knowing how close to normal the treatment has brought the patient than in knowing that the functional level of the patient is double or triple his functional level on entry to the trial. It is therefore reasonable to express the performance of patients in all tasks in the QENF as a percentage of normal function. Many tests in the QENF show significant differences in performance of males and females[17, 19] and therefore require that performance be matched by sex. As discussed below, many tests in the QENF show significant effects of age on performance and therefore require that performance be matched by age. Once the performance of patients in neuropharmacological trials is expressed as a percentage of age-sex–matched normal function, the data may be reduced by combining individual tests into meaningful indices of neurological function.[26, 39, 40]

Age effects. A study[26] was designed to examine the effects of age on performance in the QENF and the importance of selecting proper control groups for assessing clinical trials. Forty normal young adult subjects, 10 patients with Parkinson's disease and their 10 normal spouses, and 10 patients with multiple sclerosis and their 10 normal spouses participated in this study. Each subject group had equal numbers of males and females and differences due to sex were taken into account prior to analysis for age effects.

Significant decreases in performance with increasing age were found for: (1) tests requiring fine coordinated movements of the dominant hand, (2) tests of steadiness performed with the arm in a supported position, and (3) tests of reaction time, tandem gait, and sensation. Nonsignificant decreases in performance with increasing age were found for: (1) most motor tests performed by the nondominant hand, (2) tests of steadiness performed with the arm in an unsupported position,

and (3) tests of vision, strength, and speed. Where there were significant differences among the three normal subject groups, it was the oldest normal subject group that differed from the two younger subject groups.

Normal young adults did not perform significantly better than normal subjects in the age range of patients with multiple sclerosis; but, normal young adults performed significantly better than normal subjects in the age range of patients with Parkinson's disease, especially on tasks requiring fine skilled movements of the dominant hand and coordinated activities of the lower extremities. These results indicated that the performance of patients with multiple sclerosis could be expressed as a percentage of the function of either age-matched normal controls or normal young adult controls. The performance of patients with Parkinson's disease could be expressed only as a percentage of the function of age-matched normal controls.

Recent neuropharmacological trials. Forty-two ambulatory patients with Parkinson's disease participated in a double-blind crossover trial of amantadine and placebo.[39] A comprehensive battery of both quantitative and qualitative tests was carried out on each patient after he entered the study, after previous standard treatment was discontinued or reduced to a minimal tolerable dose, while on placebo, while on amantadine, and at three-week intervals. Quantitative measures revealed significant improvement in 10 of 19 SADLE tests, in several tests of strength and station, and in all tests of coordination and gait. When the amantadine scores were compared to placebo scores, an average improvement of 29% occurred in the SADLE, 14% in tests of coordination, 11% for gait, and 3% for strength. Sensation and neuropsychological performance were unaffected, and side effects were minimal. Comparison of amantadine scores with entry scores (first examination) suggested that amantadine may be superior to classical medications.

The quantitative results were in general agreement with qualitative assessments. The neurologists judged amantadine superior to placebo in 74% of the patients and found that all relevant symptoms and physical signs improved. Sixty-four per cent of the patients on amantadine experienced subjective improvement compared to 21% on placebo. Twenty-eight of the 42 patients with Parkinson's disease participating in the amantadine trial continued their participation in the therapeutic trial. Each patient began levodopa therapy as amantadine was gradually withdrawn and reached a stable, maximal tolerable dose of levodopa within 4 to 5 months. After evaluation on levodopa alone, patients were randomly allocated into two treatment groups, using a double-blind crossover design to compare levodopa with amantadine to levodopa with placebo.

Quantitative measures showed levodopa superior to amantadine in grip strength, foot speed, foot coordination and gait, and in 6 of the 19 tests in the SADLE.[40] Levodopa was shown to be superior to amantadine in the relief of functional disability (dressing, hygiene, feeding, speech) and in improving neurological signs (hand tremor, finger coordination, and associated movements), in general agreement with quantitative measures.

The combination of levodopa and amantadine proved more effective than levodopa and placebo on the quantitative measures of grip strength, pencil rotation, step tracking, critical tracking, foot speed, foot coordination, and gait in 3 of 19 tests in the SADLE, and in 1 NPE test. Once again, quantitative and qualitative assessments of changes in neurological function were in general agreement. The results of the therapeutic trial indicate that amantadine provides a weak synergistic effect when given with levodopa.

Conclusions

The development, evaluation, and use of a large-scale battery of tests called the

Quantitative Examination of Neurological Function (QENF) have been described. For more than a decade, this research has involved the joint participation of neurologists, pharmacologists, biomedical engineers, psychologists, biostatisticians, programmers, and paramedical personnel.

The test battery has been critically evaluated with respect to age, sex, motivation, and learning effects. In addition, the reliability of the test measures has been assessed, the consensual validity of some steadiness tests has been reported, the importance of adequately training paramedical personnel has been determined, and the importance of selecting proper control groups for assessing clinical trials has been established. Quantitative and qualitative assessments of changes in neurological function obtained from several neuropharmacological trials have been found to be in general agreement.

The results of evaluations of the QENF to date indicate that quantitative test batteries should not be used to assess a neuropharmacological trial prior to extensive evaluation to assure a correct assessment. A comprehensive evaluation of a quantitative test battery is required in order to critically compare the results with those of others. Evaluations of the QENF to date encourage its continued development and the eventual realization of an automated examination of neurological function to objectively assess changes in the performance of patients in neuropharmacological trials.

References

1. Albers, J. W., Tourtellotte, W. W., Pew, R. W., et al.: Quantification of motor performance in the clinical neurological examination. Abstracts, Proceedings of the Eighth International Conference on Medicine and Biological Engineering, pp. 4-12, 1969.
2. Albers, J. W., Potvin, A. R., Tourtellotte, W. W., et al.: Investigation of physiological and pathological tremors. Abstract, Twenty-third Annual Conference on Engineering in Medicine and Biology, p. 235, 1970.
3. Albers, J. W.: A quantitative investigation of

tremor in asymptomatic and pathological subjects, Doctoral Dissertation, University of Michigan, 1970.

4. Albers, J. W., Repa, B. S., Domino, E. F., et al.: Effects of d-amphetamine on quantitative measures of motor performance, Proceedings, Eighth Annual NASA–University Conference on Manual Control, 1972.

5. Albers, J. W., Potvin, A. R., Tourtellotte, W. W., et al.: Quantification of hand tremor in the clinical neurological examination, IEEE Trans. on Biomed. Engr., BME 20:27-37, 1973.

6. Domino, E. F., Albers, J. W., Potvin, A. R., et al.: Effects of d-amphetamine on quantitative measures of motor performance, CLIN. PHARMACOL. THER. 13:251-257, 1972.

7. Fleishman, E. A.: Dimensional analysis of psychomotor abilities, J. Exp. Psychol. 48:437-454, 1954.

8. Fleishman, E. A.: Development of a behavior taxonomy for describing human tasks: A correlational-experimental approach, J. Appl. Psychol. 51:1-10, 1967.

9. Gagne, R. M., and Fleishman, E. A.: Psychology and human performance: An introduction to psychology, New York, 1959, Henry Holt & Co.

10. Henderson, W. G.: Application of some statistical methods including factor analysis to a battery of clinical quantitative neurological tests for evaluating disability in multiple sclerosis, Doctoral Dissertation, University of Michigan, 1970.

11. Henderson, W. G., Tourtellotte, W. W., and Potvin, A. R.: Training physical therapists to administer a quantitative clinical neurological examination. Manuscript in preparation.

12. Ivers, R. R., and Goldstein, N. P.: Multiple sclerosis: A current appraisal of symptoms and signs, Mayo Clin. Proc. 38:457-466, 1963.

13. Jex, H. R., and Allen, R. W.: Research on a new human dynamic response test battery, Proceedings of Sixth Annual NASA University Conference on Manual Control, 1970.

14. Kuzma, J. W.: A statistical study of various aspects of a battery of clinical neurological tests, Doctoral Dissertation, University of Michigan, 1963.

15. Kuzma, J. W., Tourtellotte, W. W., and Remington, R. D.: Quantitative clinical neurological testing. II. Some statistical considerations of a battery of tests, J. Chronic Dis. 18:303-311, 1965.

16. Kuzma, J. W., Namerow, N. S., Tourtellotte, W. W., et al.: An assessment of the reliability of three methods used in evaluating the status of multiple sclerosis patients, J. Chronic Dis. 21:803-814, 1969.

17. Potvin, A. R.: An analysis of asymptomatic

performance in the clinical quantitative neurological examination, MS thesis in psychology, University of Michigan, 1970.

18. Potvin, A. R., Albers, J. W., Tourtellotte, W. W., et al.: The development of clinical instruments for measuring steadiness. Abstracts, Twenty-fourth Annual Conference on Engineering in Medicine and Biology, p. 104, 1971.

19. Potvin, A. R.: The effects of age, motivation and learning on performance in the quantitative examination of neurological function, Doctoral Dissertation, University of Michigan, 1971.

20. Potvin, A. R., Tourtellotte, W. W., Dailey, J. A., et al.: Simulated activities of daily living examination, Arch. Phys. Med. Rehabil. 53: 476-487, 1972.

21. Potvin, A. R., Tourtellotte, W. W., Albers, J. W., et al.: Neurobioengineering and clinical trials, Southwestern IEEE Conference Record, pp. 155-159, 1972.

22. Potvin, A. R., Tourtellotte, W. W., Albers, J. W., et al.: A portable tracking task battery for evaluating clinical trials, Abstracts, Twenty-sixth Annual Conference on Engineering in Medicine and Biology, p. 162, 1973.

23. Potvin, A. R., Tourtellotte, W. W., Pew, R. W., et al.: The effects of motivation and short-term learning on performance in the quantitative examination of neurological function, J. Med. Instrum., 1973. (Abst.)

24. Potvin, A. R., Tourtellotte, W. W., Pew, R. W., et al.: Motivation and learning in the quantitative examination of neurological function, Arch. Phys. Med. Rehabil. 7:75, 1973.

25. Potvin, A. R., Tourtellotte, W. W., Albers, J. W., et al.: Reliability, long and short-term learning in the quantitative examination of neurological function. Manuscript in preparation.

26. Potvin, A. R., Tourtellotte, W. W., Albers, J. W., et al.: The importance of age effects in the assessment of clinical trials. In press, J. Chronic Dis.

27. Repa, B. S., Albers, J. W., Potvin, A. R., et al.: The use of a battery of tracking tasks in the quantitative evaluation of neurological function, Proceedings of the Seventh Annual NASA–University Conference on Manual Control, 1971.

28. Repa, B. S., Albers, J. W., and Tourtellotte, W. W.: Applications of tracking to the evaluation of motor disabilities, Proceedings of the Joint NASA–United Cerebral Palsy Research Foundation Conference on Technology and the Neurologically Handicapped, 1971.

29. Repa, B. S.: The use of a tracking task battery in the quantitative evaluation of neurological function, Doctoral Dissertation, University of Michigan, 1972.

30. Repa, B. S., Albers, J. W., Pew, P. W., et al.: Clinical applications of tracking, Proceedings of the Eighth Annual NASA–University Conference on Manual Control, 1972.

31. Rose, A. S., Kuzma, J. W., Kurtzke, J. F., et al.: Cooperative study in the evaluation of therapy in multiple sclerosis: ACTH vs. placebo, Final Rept. Neurol. 20:1-59, 1970.

32. Schumacher, G. A., Beebe, G., Kibler, R. F., et al.: Problems of experimental trials of therapy in multiple sclerosis: Report to the panel on the evaluation of experimental trials of therapy in multiple sclerosis, N. Y. Acad. Sci. 122: 552-568, 1965.

33. Smith, A.: Objective indices of severity of chronic aphasia in stroke patients, J. Speech Hear Disord. 36:167-207, 1971.

34. Stribley, R. F., Albers, J. W., Tourtellotte, W. W., et al.: A measure of station in the quantitative examination of neurological function, Abstracts, 25th Annual Conference on Engineering in Medicine and Biology, p. 233, 1972.

35. Stribley, R. F., Albers, J. W., Tourtellotte, W. W., et al.: A quantitative study of stance in normal subjects. In press, Arch. Phys. Med. Rehabil.

36. Tourtellotte, W. W., and Haerer, A. F.: The use of oral corticosteroid in the treatment of multiple sclerosis: a double-blind study, Arch. Neurol. 12:536-545, 1965.

37. Tourtellotte, W. W., Haerer, A. F., Simpson, J. F., et al.: Quantitative clinical neurological testing, I. A study of 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. 122:480-505, 1965.

38. Tourtellotte, W. W.: Addendum to the National Cooperative ACTH Study, *in* Wolfgram, F., Allison, G. W., Stevens, J. G., and Andrews, J. M., editors: Multiple sclerosis immunology, biology and ultrastructure, UCLA Forum in Med. Sci., No. 16, New York, 1972, Academic Press, Inc., pp. 527-528.

39. Walker, J. E., Albers, J. W., Tourtellotte, W. W., et al.: A qualitative and quantitative evaluation of amantadine in the treatment of Parkinson's disease, J. Chronic Dis. 25:149-183, 1972.

40. Walker, J. E., Potvin, A. R., Tourtellotte, W. W., et al.: Amantadine and levodopa in the treatment of Parkinson's disease, CLIN. PHARMACOL. THER. 13:28-36, 1972.