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*

Wallace W. Tourtellotte, Armin F. Haerer,† John F. Simpson, J. W. Kuzma,‡ J. Sikorski

Department of Neurology and the Department of Biostatistics, University of Michigan, Ann Arbor, Mich.

The long term goal of this study is to design a battery of objective clinical neurological tests which are appropriate for patients with multiple sclerosis, so that the results of therapeutic trials in this disease might be evaluated with less bias and hence be more valid. In this study some of the conventional neurologic tests used by clinicians[1, 2] to evaluate functional capacity of multiple sclerosis patients have been replaced in part by methods which lend themselves to more objective measurements.

Since no previous studies along this line have been reported, it was necessary to design experiments on normal individuals and patients with multiple sclerosis to determine some of the limitations of the battery of tests before undertaking a therapeutic trial.

METHODS

The battery of clinical quantitative neurological tests used in this study is shown in TABLE 1. The test battery consisted of three parts, visual acuity and functions of the upper and lower extremities. The details for the administration of this battery of neurologic tests were modified somewhat during the stages of experimentation with normal subjects as new information was obtained. The final form of instructions for the examiner used on all the multiple sclerosis patients was as follows:

Description of the Quantitative Clinical Tests

Visual acuity with glasses. The Standard Snellen visual acuity test was used; the corrected vision was measured in each eye in turn. The visual acuity is established by finding a line on the chart for which the patient makes no more than three errors and the next larger print is read without error. Deductions were made for incorrectly identified letters. A conversion table was prepared for each possible response; for example, 20/25-3 corrected vision was equivalent to 0.75 visual acuity.

When the patient's visual acuity was so impaired that the largest letters on the Snellen reading chart could not be identified, the Lebensohn reading card was used. The visual acuity was estimated by permitting the patient to use the card *ad lib;* that is, he could use peripheral vision at any distance. A conversion table was prepared for each letter size similar to the Snellen visual acuity test conversion table except that deductions were not made for incorrectly identified letters and that no consideration was given for distance between eye and letters.

*Supported in part by a grant from The Upjohn Company, Kalamazoo, Mich. and The Kenneth H. Campbell Foundation for Neurological Research, Grand Rapids, Mich., as well as a grant from the University of Michigan Computor center.

†Present address: 5803A, Billhymer, Fort Knox, Ky.

‡Present address: Health Sciences Computing Facility, U.C.L.A., Los Angeles, Calif.

TABLE 1
LIST OF CLINICAL QUANTITATIVE NEUROLOGICAL FUNCTIONS TESTED IN THIS STUDY

I. Visual acuity
II. Functions of the upper extremity
A. Hand speed
B. Hand coordination
C. Strength (grip, wrist, deltoid)
D. Fatigue
1. Hand speed
2. Hand coordination
3. Grip strength
E. Sensation
1. Finger vibration
2. Finger two-point discrimination
III. Functions of the lower extremity
A. Station stability
1. Eyes open
2. Eyes closed
B. Foot speed
C. Strength (foot and hip flexor)
D. Fatigue
1. Foot speed
2. Hip flexor strength
E. Vibratory sense (toe)

On subsequent testing the subject was instructed to read the line which was his previous best score.

Speed and fatigue of hand. For this an Adam-Clay 9 key laboratory counter was used. The subject seated in a standard chair and table arrangement was asked to strike the same button with his index finger for 5, 10, 20 and 30 consecutive seconds, as fast as possible. The right hand was tested before the left hand. (In all subsequent tests the right side was tested first.) The rest period between trials was the time required to record the result. The 10-second interval (second trial) is used in subsequent analyses as the speed of the hand tests, and the 30-second interval (fourth trial) divided by three is referred to subsequently as hand speed fatigue test.

Coordination and fatigue of hand. An Adam-Clay 9 key laboratory counter was used for these tests. The subject was instructed to strike the white keys only, from left to right with the index finger, for 5, 10, 20 and 30 consecutive seconds as fast and as accurately as possible. All the counts (red and white keys) registered on the counter are recorded; the rest period between trials was the time to record these data.

Frequently a subject will strike a red key by mistake. In order to get a score which will reflect coordination more appropriately, an adjusted score is obtained by subtracting twice the score recorded for the red keys from the score recorded for the white keys. Occasionally a negative net score will result from this correction; such a score was arbitrarily set to zero.

The adjusted 10 second test score (second trial) is referred to in subsequent analyses as the test for coordination of the hand, and the adjusted 30 second test

score (fourth trial) divided by three is referred to subsequently as the test for fatigue of hand coordination.

Grip and fatigue strength. The Jamar hand dynamometer was used for this; the variable grip handle position was maintained in the second position. The subject was instructed to maintain his grip as high on the handle as possible for five seconds. The rest period between the five trials consists of the time to record the reading and reset the dynamometer dial. The first trial is referred to as grip strength and the fifth trial, as grip strength fatigue.

Wrist strength. A calibrated Newman myometer (60 pound gauge) with attached handle pushers was used for all subsequent strength tests. The wrist was positioned by putting the subject's arm on the arm of the standard chair so that the wrist in maximum dorsiflexion was at the end of the arm of the chair; the hand was in the form of a fist. The instrument pad was centered over the third metacarpal; force was then applied at right angles to the dorsum of the hand. The subject generally gives away suddenly. The rest period between the two trials consisted of the time to reset the dial and record the datum. The first trial was used as the wrist strength unless the patient did not perform the test properly; under these circumstances the second trial result was considered.

Deltoid strength. A Newman myometer as described was used. The deltoids were positioned for the test by flexing the arm at the elbow and abducting the upper arm to a right angle. The instrument pad was placed at right angles over the lateral aspect of the elbow joint. The rest period between the two trials consisted of the time to reset the dial. The first trial was used as the deltoid strength unless the patient did not perform the test properly; under these circumstances the second trial result was considered.

Finger vibratory sensibility. A Biothesiometer (Model PVD) which vibrates at 120 cycles per second was modified to double its range (0 to 100 microns). The subject was instructed to place the pad of second finger down onto the vibrator, so that the fingerprint whorl was over the center of the vibrator. The subject was given a recognizable stimulus, then beginning with zero voltage, the amplitude of vibration was increased gradually by turning the voltage knob evenly and slowly, with the examiner watching the voltage dial, until the subject reported his first perception (threshold) of vibration. This procedure was repeated three times and the threshold values converted to microns and averaged to arrive at a score.

If a patient was found to have a very high threshold, his responses appeared more consistent if the voltage was advanced at five volt steps rather than smoothly.

Finger two-point discrimination. Sweet's two point compass calibrated in millimeters (mm.) was used. The subject was given a recognizable stimulus of $>$ 10 mm. to the index finger across the pad at the center of the fingerprint whorl. The stimulus was decreased by about one mm. per trial, until the subject gave three consecutive responses of perceiving one point at the same measurement. If this was the threshold value, the subject correctly recognized the next increment of one mm.

Station stability. In a corner with grab supports to prevent falling, the subject was asked to stand on one foot without touching the walls with eyes open. The test ended when the patient touched to catch his balance or when he performed the test for 30 seconds. The best of three tests was taken as the station stability with eyes open.

The test was carried out in a similar fashion with the eyes closed.

Speed and fatigue of foot. A foot pedal similar to a gas pedal on a car was constructed to measure the speed of foot tapping. The subject was instructed to tap the foot pedal, with sufficient force to have an attached counter register it, for 5, 10, 20 and 30 consecutive seconds, as fast as possible. For proper positioning of

the foot on the pedal platform the heel should be directly over the pivot point. The test was demonstrated to the subject and a trial of several seconds was permitted. The rest period between trials was the time required to record the results. The 10 second interval (second trial) is used in subsequent analyses as the speed of the foot, and the 30 second interval (fourth trial) divided by three is referred to as foot speed fatigue.

Foot strength. The Newman myometer described above was used for this test. The foot was positioned by placing the foot in maximum dorsiflexion with the heel in a groove in a two inch wooden foot platform. The foot platform was notched to fit around the chair legs, so as to prevent it from sliding during the test. The instrument pad was placed over the second, third and fourth metatarsals; force was applied at a right angle to the dorsum of the foot. The rest period between two trials consisted of the time to reset the dial and record the reading. The majority of normal males used in this study could resist a force of 84 pounds which is the maximum reading on the myometer. The first trial was used as the strength of the dorsiflexion of the foot unless considered unreliable due to lack of patient cooperation; under these circumstances the second result was considered.

Hip flexor strength and fatigue. The Newman myometer described above was used for these tests. The hip flexors were positioned by having the subject slide forward in the chair and then the extended leg was raised to the horizontal just before testing. The instrument pad was placed over the ankle and force was applied at right angles until the heel contacted the floor; the reading was noted just before the heel touched the floor, since the gauge is jarred by this contact. The rest periods between the five consecutive trials consisted of the time to record the reading and to reset the dial. Rarely can an individual resist 84 pounds, which is the maximum value on the myometer. The first trial was referred to as hip flexor strength and the fifth trial as the hip flexor strength fatigue.

TABLE 2
GENERAL CHARACTERISTICS OF THE NORMAL COLLEGE STUDENTS USED IN THIS STUDY TO DETERMINE SOME OF THE LIMITATIONS OF THE PROPOSED TEST BATTERY

Characteristics	Females	Males		
		Group 1	Group 2	Normal male right handed medical students W.W.T. data
Total students	20	17	10	8
per cent right handed	95	88	90	100
Average age in years	21	24	31	26
(range)	(19–22)	(18–37)	(22–38)	(22–34)
Weight in pounds	125	165	177	173
(range)	(110–150)	(135–202)	(140–230)	(140–205)
Height in inches	64	69	69	71
(range)	(60–68)	(67–74)	(68–73)	(68–76)

TABLE 3
GENERAL CHARACTERISTICS OF THE AMBULATORY MULTIPLE SCLEROSIS PATIENTS USED IN THIS STUDY TO EVALUATE THE APPLICATION OF THE PROPOSED TEST BATTERY

Characteristics	Group 1	Group 2
	Outpatients (no therapy)	Inpatients (no therapy)
Total patients	10	40
Per cent females	70	64
Per cent right-handed	89	100
Average age (years)		
Female : Mean	37	38
Range	28–51	16–51
Male : Mean	40	39
Range	23–52	28–51
Weight (pounds)		
Female : Mean	117	131
Range	89–156	92–183
Male : Mean	174	167
Range	150–199	136–199
Height (inches)		
Female : Mean	56	64
Range	60–67	60–70
Male : Mean	70	68
Range	69–71	61–73
Duration of disease (at time of examination) (years)		
Mean	5	5
Range	1–17	1–20
Per cent of patients who required appliances to walk at time of examination	40	47
Average Kurtzke total disability score at time of examination	4	5
Diagnostic grouping* (per cent)		
I. (Relapses $\leq$8 weeks)	33 (in remission)	37 (in relapse)
II.	56	52
III.	11	11

The 40 patients reported in this TABLE are the same as those used in the therapeutic trial; only the data obtained at the initial examination, that is, before the drugs were administered are shown in this TABLE.

*Group I: *Multiple sclerosis manifested by bouts.* The patient may be either in a relapse or a remission at the time of the examination. Group II: *Multiple sclerosis manifested by bouts followed by a chronic phase.* The patient may be stationary or progressively worsening; relapses are rare. Group III: *Multiple sclerosis manifested by a chronic course.* Patient may be stationary or progressively worsening; relapses are rare.

Toe vibratory sensibility. The Biothesiometer and the technique of using the equipment was the same as described above for finger vibratory sensibility. The subject's foot was placed flat on the floor and then the vibrator was positioned at the base of the first toe on the joint interspace. In patients with very impaired vibratory sense the sensation was referred; for example, "in my leg." Only sensation perceived in the finger or at the base of the toe were recorded.

As well as carrying out all the above tests in a given order on patients, a routine neurological examination was done in order to score the patient's neurological function according to Kurtzke's system.[2] The entire examination took about 45 minutes.

Description of Normal Subjects

In TABLE 2 are presented the general characteristics of the normal college students used in this study to determine the limitations of the proposed test battery. The individuals were paid volunteers from the student body of the medical school or the school of nursing (females); they were judged free of disease in the past or at present by history.

Description of Multiple Sclerosis Patients

The definition of multiple sclerosis beyond reasonable clinical doubt was based upon the criteria presented in another report in this monograph.[3] Also the patients used in this study were selected after consideration of the criteria for exclusion of a patient from a therapeutic trial presented in Schumacher's report.[3] In TABLE 3 are presented the general characteristics of the ambulatory multiple sclerosis patients used in this study.

Design of the Therapeutic Trial

The design of the therapeutic trial also followed closely that laid down by Schumacher's committee.[3] We divided the medical staff into two teams. The therapists were Armin F. Haerer and John F. Simpson. They selected the patients by a preprinted set of rules[3] and also prescribed the drugs in a predetermined random fashion; the patients were unaware of what drug they were receiving. The therapists were also responsible for patient care. The only responsibility the examiner (Wallace W. Tourtellotte) had in this study was to carry out the quantitative neurological examination, as well as score the neurological function based on a classical neurological examination according to Kurtzke;[2] he only saw the patients at the time of the examination, and he had no knowledge of the kind of medications the patients were receiving. Effort to eliminate bias at the patient and examiner level was taken.

In TABLE 4 are given the general characteristics of the ambulatory multiple sclerosis patients used in the therapeutic trial.

TABLE 5 shows the time when the examinations were carried out and the dosages of the five treatments used. Immediately after the initial neurological examination the treatment was started. For example, Solu-Medrol,® 160 mg. was given every day intramuscularly for seven days, then the second quantitative neurological examination was carried out. The Solu-Medrol® was then tapered at the rate of 25 mg. every day for six days. At this point the third neurological examination was carried out. These large dosages of steroids were recommended to us by Robert F. Kibler; his experiences are also reported in this monograph.[4]

In addition to the above medications the patients were all placed on a 900 mg. sodium diet which contained approximately 1,800 calories. Potassium chloride was given one gram three times a day and Maalox,® Gelusil® or Amphojel,® 30 ml. in between meals and at bedtime.

All patients were given appropriate physical therapy in the morning and afternoon under the direction of professional therapists. Other than this the patients were up *ad lib* and encouraged to be at full activity.

TABLE 4
GENERAL CHARACTERISTICS OF THE AMBULATORY MULTIPLE SCLEROSIS PATIENTS USED IN THE THERAPEUTIC TRIAL

Characteristics	All patients in trial	Subgroups				
		Vitamin B_{12}	Depo-acth®	Solu-Medrol®	Medrol®	Alpha-drol®
Total patients	40	7	8	8	8	9
Per cent females	64	57	75	88	50	67
Per cent right-handed	100	100	100	100	100	100
Average age (years)						
Female : Mean	38	40	37	36	38	38
Range	16–51	35–44	29–47	16–51	31–46	17–46
Male : Mean	39	43	30	33	43	39
Range	28–51	41–46	32–40	—	29–51	28–53
Weight (pounds)						
Female : Mean	131	118	135	130	127	140
Range	92–183	105–150	92–183	103–178	102–140	109–180
Male : Mean	167	161	180	144	175	155
Range	136–199	157–166	174–187	144	136–199	144–170
Height (inches)						
Female : Mean	64	65	64	63	65	64
Range	60–70	62–67	61–67	60–66	63–66	61–70
Male : Mean	68	71	72	66	69	68
Range	61–73	69–73	70–73	66	66–71	67–69
Duration of disease at time of examination (years)						
Mean	5	7	5	4	6	5
Range	1–20	3–15	2–18	1–10	2–14	1–20
Per cent of patients who required appliances to walk at time of admission	47	43	43	50	50	50
Average Kurtzke total disability score at time of examination	5	5	6	5	5	6
Diagnostic grouping (per cent)						
I. (Relapses $\leqq$ 8 weeks)	37	29	38	50	37	30
II.	52	57	62	38	50	55
III.	11	14	0	12	12	15

TABLE 5
TIME OF THE QUANTITATIVE NEUROLOGICAL EXAMINATION AND THE FIVE TREATMENTS USED

Initial Exam	Treatment for 1st week	Second Exam	Treatment for 2nd week	Third Exam
	Vitamin B_{12} : 1000 μgm., Q.D., I.M.		1000 μgm., Q.D., I.M.	
	Depo-acth® : 120 mg., Q.D., I.M.		Decrease 20 mg., Q.D.	
	Solu-Medrol® : 160 mg., Q.D., I.M.		Decrease 25 mg., Q.D.	
	Medrol®* : 160 mg., Q.D., Oral		Decrease 25 mg., Q.D.	
	Alphadrol®† : 60 mg., Q.D., Oral		Decrease 10 mg., Q.D.	

*Medrol® : 6 α Methyl Prednisolone.
†Alphadrol® : 6 α Fluoro Prednisolone.

On the morning of the quantitative neurological examination physical therapy was postponed until after the examination. As a further effort to standardize conditions, patients were brought to the examining room in a wheel chair.

Statistical Methods

The statistical methods used in this study have been presented in Kuzma's doctoral thesis,[5] and will also be presented elsewhere.[6] A 4 × 4 Graeco-Latin square design with two observations per cell was used to estimate the variability of learning and fatigue of the subjects and to determine if there was a difference between observers.

Kuzma[5] also discusses the problems related to correlated observations such as that which resulted from repetitive testing in an attempt to produce fatigue. In the present study fatigue is presented as the last trial of a series, whereas Kuzma has utilized a more powerful statistic (time-response curves).

RESULTS AND DISCUSSION

The battery of clinical quantitative neurological tests listed in TABLE 1 has been applied to normal individuals, to untreated patients with multiple sclerosis and to patients participating in a therapeutic trial.

Normal Individuals

Since no one has used a similar battery of clinical quantitative neurological tests, we designed experiments on normal college students to determine the limitations of the battery of tests and to redefine the tests, if necessary.

Some comments on this battery of tests which will probably be extended when more normal individuals are examined will be presented first. It was found that with the speed tests, such as hand speed, hand coordination, and foot speed, a five-second trial served as a warm-up period. Along the same line, when station stability was tested, especially with the eyes closed, a bad start could result in a low score; therefore, it was decided to choose the best score out of three trials.

The Newman myometer (60 pound gauge) which can only measure 84 pounds did not permit evaluation of strength of dorsiflexors of the foot in the majority of normal male subjects and occasionally the hip flexor strength as defined for this study. However, we continued to use the myometer since this was not the case for normal females or any of the multiple sclerosis patients.

TABLES 6, 7 and 8 present the results of the test battery done on normal individuals; the general characteristics of the normal individuals were shown in TABLE 2. The data in columns 2, 3, and 4 in TABLES 6, 7 and 8 were some of the first we obtained; they were obtained by John Sikorski under the direction of Wallace W. Tourtellotte. The data in column 2 of TABLES 5, 6 and 7 are the only normal female data we have. The data in columns 3 and 4 were composed of 27 different normal male medical students (17 in group 1 and 10 in group 2); they were examined four months apart. The agreement between the two samples was considered good.

It can be gleaned from these TABLES (7 and 8) that performance of normal females differs in certain respects from normal males; the most marked differences appear in the strength tests.

The data in TABLES 6, 7 and 8 also suggest that certain tests of the battery are sensitive enough to detect handedness in both females and males. Speed of hand and grip strength appear to differ most when the right hand is compared to the left. We have also analyzed the data separating the few left-handers from the data presented in TABLES 7 and 8 and the differences become even more obvious.

In TABLES 6, 7 and 8 are also presented some data obtained by Wallace W. Tourtellotte (data in last columns). These data are presented separately for several reasons, even though the sample size is small and hence subject to errors inherent in testing small samples. First, these data were obtained by the latest techniques; second, the examiner (Wallace W. Tourtellotte) who made these observations is the same examiner used in the therapeutic trial to be reported below; third, in analyzing the data in the therapeutic trial we discovered that if normal results for a given test (within mean of WWT data minus two standard deviations) were analyzed separately from abnormal results (outside mean of WWT data minus two standard deviations) a number of significant differences was disclosed; and fourth, by inspection most of the results obtained by WWT do not differ too much from those obtained by John Sikorski (columns 3 and 4, TABLES 6, 7 and 8) except when the technique of carrying out the test was changed drastically.

Also shown in TABLES 6, 7 and 8 are coefficients of variation for most of the mean test results. The coefficient of variation (standard deviation times 100 divided by the mean) is a convenient expression, since one can compare the absolute

TABLE 6
VISUAL ACUITY (CORRECTED WITH GLASSES) OF NORMAL COLLEGE STUDENTS

	20 Female students		Male Students				
			Group 1 (17 students)		Group 2 (10 students)		8 Normal male right handed medical students W.W.T. data
	Mean (C.V.* %)		Mean (C.V. %)		Mean (C.V. %)		Mean (C.V. %)
	Right	Left	Right	Left	Right	Left	Right
Visual acuity (Snellen)	1.02 (26)	1.10 (24)	1.12 (30)	1.06 (30)	—	—	1.31 (39)

*C. V. means coefficient of variation; standard deviation × 100/mean.

TABLE 7
FUNCTION OF THE UPPER EXTREMITIES OF NORMAL COLLEGE STUDENTS USED IN THIS STUDY TO DETERMINE SOME OF THE LIMITATIONS OF THE PROPOSED TEST BATTERY

Tests	20 Females		Male Students				
			Group 1 (17 Students)		Group 2 (10 Students)		8 Normal male right handed medical students W.W.T. data
	Mean (C.V. %)		Mean (C.V. %)		Mean (C.V. %)		Mean (C.V. %)
	Right	Left	Right	Left	Right	Left	Right
Speed of hand (strokes per 10 seconds)							
2nd trial	65 (10)	61 (13)	70 (11)	65 (10)	74 (16)	68 (11)	77 (7)
Fatigue (4th trial)	—	—	—	—	62 (10)	60 (8)	65 (9)
Coordination of hand (strokes per 10 seconds)							
2nd trial	34 (23)	32 (32)	39 (23)	34 (16)	41 (15)	40 (16)	46 (22)
Fatigue (4th trial)	—	—	—	—	39 (13)	35 (13)	40 (19)
Strength (pounds)							
Grip							
1st trial	62 (16)	53 (20)	104 (20)	92 (18)	103 (17)	99 (17)	100 (26)
Fatigue (5th trial)	53 (19)*	46 (20)*	91 (23)*	78 (19)*	76 (27)	71 (25)	77 (22)
Wrist	34 (17)	31 (17)	57 (15)	52 (20)	53 (14)	53 (21)	58 (16)
Deltoid	28 (12)	28 (14)	52 (21)	52 (25)	—	—	58 (13)
Sensation							
Vibration (microns)	0.1 (21)	0.1 (28)	0.1 (44)	0.1 (52)	0.1 (57)	0.1 (69)	0.1 (41)
Two point (mm.)	2.8 (15)	2.8 (14)	2.8 (9)	2.8 (9)	—	—	3.2 (10)

*Grip fatigue test used on these individuals differed from that defined in the METHODS; grip was only maintained two seconds, instead of five, for each trial.

value from one test to another because it is calculated as a percentage of the mean. For example, in the row entitled "speed of hand second trial" note that the coefficient of variation for females was 10 per cent for the right hand and for data obtained by WWT on normal males it was seven per cent; this was considered good agreement.

The highest coefficients of variation were obtained when vibratory sense was measured; 41 per cent for the finger tip and 70 per cent for the base of the great toe. Greater attention to details of conducting the test has not decreased this variation much. It is our impression that the decision of each subject as to what constitutes vibration threshold must be different; that is, it is not due to faulty technique of conducting the test. Furthermore, since the vibratory stimulus perceived is very small, on the average 0.1 for the index finger tip, a subject has to

TABLE 8
FUNCTION OF THE LOWER EXTREMITIES OF NORMAL COLLEGE STUDENTS USED IN THIS STUDY TO DETERMINE SOME OF THE LIMITATIONS OF THE PROPOSED TEST BATTERY

Tests	20 Female Students		Male Students				
			Group 1 (17 students)		Group 2 (10 students)		8 Normal male Right handed Medical students W.W.T. data
	Mean (C. V. %)		Mean (C. V. %)		Mean (C. V. %)		Mean (C. V. %)
	Right	Left	Right	Left	Right	Left	Right
Station stability (seconds)							
Eyes open	—	—	—	—	—	—	> 30
Eyes closed	—	—	—	—	—	—	> 30
Speed of foot (strokes per seconds)							
2nd trial	42(19)	34(28)	45(22)	42(16)	40(21)	43(21)	53(11)
Fatigue (4th trial)					34(16)	34(15)	39(9)
Strength (pounds)							
Foot	59(18)	58(19)	77(13)	75(15)	81(9)	81(7)	> 84
Hip flexor							
1st trial	28(24)*	29(26)*	38(30)*	34(31)*	—	—	61(30)
Fatigue (5th trial)	28(23)*	27(20)*	38(32)*	36(36)*	—	—	57(30)
Sensation							
Vibration (microns)	0.3(52)	0.3(41)	0.4(60)	0.4(64)	0.8(72)	0.6(107)	0.5(70)

*Hip flexor strength test used on these individuals differed from that defined in the METHODS; in the sitting position the hip and knee were flexed and then pressure through the myometer was applied at right angles directly over the knee.

be very accurate or else it will appear statistically that his perception is erratic. For example, when on duplicate testing scores of 0.1 and 0.15 microns were obtained it would appear statistically that his performance was highly variable, but clinically, since the stimulus was small, this was good agreement.

In TABLE 6 are shown the coefficients of variation for visual acuity; it is our impression that the high coefficient of variation of 39 per cent obtained by WWT may not be due to faulty technique of conducting the examination, but may be an expression of the variability of this sample which differed from the other three samples.

It might be of some interest to present coefficient of variations on two cerebrospinal fluid constituents obtained in our laboratory for comparison to those obtained by the proposed test battery. A modified biuret total protein procedure has a coefficient of variation of 5.1 per cent on replicate determinations of the same sample; when the coefficient of variation for 137 normal college students was

calculated it was found to be 27 per cent (10.2 standard deviation × 100/38.2 mg./100 ml.). Moreover, the coefficient of variation for counting normal white cells in cerebrospinal fluid (< 4.7 white cells / cu. mm.) was 20 per cent on replicate determinations of the same sample and 61 per cent (1.0 × 100/1.7) for the same 137 normal college students used to determine the normal total protein value. It would appear on comparison with these laboratory tests that the coefficients of variation presented in TABLES 6, 7 and 8 for the proposed battery of neurological tests are reasonable.

The next experiment done was to determine if normal subjects on repetitive testing fatigued, that is, obtained significantly lower scores or learned how to do the tests better, that is, obtained significantly higher scores. The first data we obtained along this line were done on the females and males (TABLES 6, 7 and 8 group 1 and 2) by John Sikorski; only the initial data are shown. For females and group 1 males he repeated the examination within two days. For group 2 males he repeated the examination for the next two days and again after three days of rest.

The equality of variances[5, 6] between the first and second examinations for females was tested with a double-tail F-test. The variances for grip strength for the right hand and hip flexor strength were significantly different ($p < 0.05$). Since the mean differences between the first and second examination were less than five per cent, these changes were considered only technically significant.

The equality of variance for group 1 males was tested statistically in the same manner as mentioned above. Only finger vibration sense for the right hand showed significantly different variances. This again was thought to be a technically significant difference.

For the group 2 males equality of variances between the four examinations was also tested. The variance for the left foot strength was significant at the five per cent level. As mentioned earlier, this may be explained by the constant reading recorded when the limit of the instrument was exceeded.

These preliminary data on repetitive testing done one to four days apart on the same normal individuals utilizing the paired Student test[5] indicated that learning and fatigability in the normal individuals was not sufficient to be statistically significant.

A more powerful design (4 x 4 Graeco-Latin square) to test if learning and fatigability on repetitive testing are significant factors was suggested by Kuzma.[5, 6] Eight normal right-handed medical students were tested every day for four days, then after one week of rest they were reexamined every day for another four days. The first week's data were averaged for a given test and compared to the second week's data. TABLE 9 presents the only neurologic test which had significantly

TABLE 9

ONLY NEUROLOGIC TEST WITH SIGNIFICANTLY DIFFERENT MEANS FROM WEEK 1 TO WEEK 2: 8 NORMAL MALE RIGHT-HANDED MEDICAL STUDENTS

Test	Week 1	Week 2
	Mean (C. V. %)	Mean (C. V. %)
Hip Flexor Strength (pounds)	48 (16)	58 (13)

different means between weeks; the hip flexor strength increased 120*per cent from week 1 to week 2. It is our impression that this improvement was due to organized daily exercise (five trials per day for eight days) on a muscle group sensitive to training; therefore therapeutic trials should be so designed as to avoid "training" of the hip flexors; that is, perhaps, weekly examinations instead of daily examinations. We have concluded from this study that the proposed test battery is not influenced by repetition of examinations on normal individuals, except hip flexor strength which increases with repetitive testing.

The next question asked was as follows: Can different examiners get the same results with the proposed test battery? Two neurologists and two physical therapists were selected as examiners and the same 4 x 4 Graeco-Latin design employing eight normal medical students mentioned above was used to answer this question. The examiners initially were instructed by Wallace W. Tourtellotte until he felt that they could carry out the tests adequately. Each of the four examiners examined two of the subjects on four subsequent days and after one week the entire program was repeated. The data presented in TABLE 10 show the neurologic tests having significantly different examiner means. For example, in the row entitled wrist strength note that physical therapist 2 obtained a value of 47 pounds, whereas the other three examiners ranged from 58 to 60. It was apparent that this examiner was not performing the test in the same way as the other examiners. On the other hand, what she was finding was consistent since the coefficient of variation was similar to the others, 12 and 9 per cent, respectively.

In view of these data we reviewed the techniques of performing the examination

TABLE 10
LIST OF NEUROLOGICAL TESTS HAVING SIGNIFICANTLY DIFFERENT EXAMINER MEANS
RIGHT-SIDED DATA: 8 NORMAL MALE RIGHT-HANDED MEDICAL STUDENTS

Tests	Examiners			
	W. W. T.; Senior Neurologist Mean (C. V. %)	A. F. H.; Junior Neurologist Mean (C. V. %)	Physical Therapist 1 Mean (C. V. %)	Physical Therapist 2 Mean (C. V. %)
Upper Extremity				
Wrist strength (pounds)	58 (9)	59 (9)	60 (9)	47 (12)*
Deltoid strength (pounds)	58 (13)	57 (14)	63 (12)	53 (15)
Two-point discrimination (MM.)	2.6 (13)	2.8 (12)	2.3 (15)	2.6 (13)
Vibration sense (microns)	0.19 (64)	0.16 (76)	0.51 (24)	0.16 (76)
Lower Extremity				
Speed of foot (Strokes per 10 seconds)	53 (9)	48 (10)	53 (9)	50 (10)
Hip strength (pounds)	60 (12)	53 (14)	56 (13)	51 (14)
Hip strength fatigue (5th Trial; pounds)	62 (12)	53 (14)	57 (14)	49 (16)
Vibration (microns)	0.5 (37)	0.6 (32)	1.1 (17)	0.6 (31)

*Underlined data were found to be responsible for significant observer means.

and retrained all the examiners. This time the examiners met every day for one hour and tested each other; at the end of five days we were all getting similar results.

After we were all retrained we examined another group of 12 normal right-handed medical students; only those tests shown in TABLE 10 were performed. TABLE 11 shows that only three neurologic tests remained significantly different among the four examiners. Examination of wrist strength apparently is "tricky" and requires a skill which may be difficult to teach. On the other hand, it is our impression that the significant differences discovered on retesting for finger two-point discrimination and finger vibration sensibility are technical dfferences, that is, the dfferences between the means should be disregarded, since they are so small. It should also be noted that the coefficients of variation decreased markedly for vibration sense for three of the four examiners on retesting. This was probably due to the fact that we decided to carry out this test on the center of the index fingerprint whorl. It is our impression that the majority of ladies who are strong enough to be physical therapists who carry out muscle testing are strong enough to carry out our proposed muscle tests. We have concluded from this portion of the study that the proposed battery of neurological tests can be performed by neurologists and by physical therapists reliably providing they are properly trained.

TABLE 11
LIST OF NEUROLOGICAL TESTS HAVING SIGNIFICANTLY DIFFERENT EXAMINER MEANS IN RETESTING EXPERIMENT
RIGHT-SIDED DATA: 8 NORMAL MALE RIGHT-HANDERS

Tests	Examiners Means			
	W. W. T. Mean (C. V. %)	A. F. H. Mean (C. V. %)	Physical Therapist 1 Mean (C. V. %)	Physical Therapist 2 Mean (C. V. %)
Wrist strength (pounds)				
Initial	58 (9)	59 (9)	60 (9)	47 (12)
Retest	51 (9)	58 (8)	42 (12)	43 (11)
Two-point discrimination on Index finger tip (mm.)				
Initial	2.6 (13)	2.8 (12)	2.3 (15)	2.6 (13)
Retest	3.2 (11)	2.8 (13)	3.1 (12)	3.0 (12)
Vibration sense on index Finger tip (microns)				
Initial	0.19 (64)	0.16 (76)	0.51 (24)	0.16 (76)
Retest	0.09 (34)	0.08 (41)	0.10 (32)	0.13 (24)

Ambulatory Multiple Sclerosis Patients

The next series of experiments were done to provide information on the applicability of this test battery to patients with multiple sclerosis. Patients with Kurtzke

total disability score[2] of 6* or less (at the time of admission) were used in this study, with the exception of patients who had relapsed from a score of 6 or less to a score of 7 or more, eight weeks before admission to the study.

The general characteristics of the two groups of patients, 10 outpatients and 40 inpatients, studied are shown in TABLE 3. In general the neurological health of the outpatients was better than the inpatients; note the mean Kurtzke total disability of 4 and 5 for outpatients and inpatients, respectively. This may be due in part to the fact that outpatients were purposely selected because their neurological status by history had not changed for eight or more weeks prior to the examination. On the other hand, the 40 inpatients reported here are the same patients as those used in the therapeutic trial; 37 per cent of these patients had had a worsening of their neurological status by history eight or less weeks prior to the examination.

We encountered some problems when the proposed test battery was applied to patients with multiple sclerosis. The manner of quantitating the visual acuity when dense central scotoma existed was mentioned under METHODS. Only the hand coordination test requires visual acuity of 0.1 (20/200) or better in one eye; only two patients out of 50 could not perform this test because their visual acuity was < 0.1.

One patient on testing right-hand speed had noticeably variable speed during each of the four trials especially the 20 and 30 second trial; "He seemed to break into a sudden burst of speed."

In station stability testing occasionally the patient was so weak he was unable to raise his leg to test one-legged station stability; for such patients we arbitrarily assigned a score of 0. In future experiments we will also include station stability in both feet with eyes open and closed.

In some patients muscle strength testing posed problems. Frequently, zero strength was recorded. In two patients out of 50 involuntary leg jerks and cramps were induced with hip flexor strength (also fatigue) testing. In another patient ankle clonus interfered with the foot strength test.

In one patient (same patient mentioned above) in this series clonus interfered with the foot-speed test. It was possible to continue these tests in spite of these involuntary movements.

When a patient had an increased sensory threshold either for two-point discrimination or vibration, it was difficult for him to respond reliably; it appeared as if he had broadened his threshold, that is, there was a wide scatter of responses. Also in the lower extremities toe vibration sense was not elicited in eight patients out of 50 when they were stimulated with 100 microns; they were assigned values of 100 arbitrarily.

It should be mentioned that patient cooperation and motivation to perform the proposed test battery to the best of their ability appeared good.

*Definition of Kurtzke[2] total disability score 4 named by him *Relatively severe dysfunctions in an ambulatory patient* is as follows: The dysfunction should not be so severe as to prevent the ability to work and/or carry on normal activities of living and to use public transportation.

Definition of Kurtzke total disability score 5 named by him *Severe dysfunction in an ambulatory patient* is as follows: Usually patient cannot work, but he should be up and about all day; he can walk unaided several blocks; also severely decreased visual acuity was included in this category.

Definition of Kurtzke total disability score 6 named by him *Assistance required for walking* is as follows: Usually the patient cannot work; but he should be up and about most of the day; he can walk short distances with aid of canes, crutches, and short double upright day; he can walk short distances with aid of canes, crutches, and short double upright braces; this category did not include patients requiring fulltime use of a walker or paraplegics.

The data dealing with the untreated ambulatory multiple sclerosis patients are presented in TABLES 12 through 15. Since all the examinations were carried out by Wallace W. Tourtellotte, it was decided to use data obtained by this examiner on eight normal volunteers as a basis for comparison; this is recorded in the second columns of TABLES 12, 13 and 14.

Also in TABLES 12, 13 and 14 the data on multiple sclerosis patients have been stratified into those patients who fell within the range (mean $\pm$ 2 standard deviations) and those who fell outside this range. In this study patients only fell on the subnormal (or abnormal) side of the range. The outer limit of the range for a given test was determined by subtracting two standard deviations from the mean of WWT data, except for two-point discrimination and vibratory sensibility in which the threshold was raised, so that a mean plus two standard deviations was used. For example, in TABLE 13 the hand speed was 77 strokes per 10 seconds; the lower limit of the range was 66 (77- 77 $\times$ 0.14). Under the assumption that these observations are normally distributed it is expected that about 2.5 per cent of the subjects will fall below 66. When this cut off point of 66 was applied to out- and inpatients, 90 and 87 per cent were abnormal, respectively.

This type of analysis has several justifications. It is reasonable to expect in a therapeutic trial that patients who are normal will not be affected by the therapy; on the other hand, it is reasonable to assume that if a medication is effective it should improve the function of those tests which are subnormal.

This technique of separating the normal functions from the subnormal has another advantage; it can decrease the variance and hence significant differences can be discovered with fewer observations.

TABLE 12
VISUAL ACUITY (CORRECTED WITH GLASSES) OF RIGHT EYE OF AMBULATORY UNTREATED MULTIPLE SCLEROSIS PATIENTS USED IN THIS STUDY TO EVALUATE THE APPLICABILITY OF THE PROPOSED TEST BATTERY

	8 Normal male right-handed medical students W.W.T. data mean (C.V. %)	Ambulatory Multiple Sclerosis Patients							
		Group 1				Group 2			
		10 Outpatients				40 Inpatients			
		Mean (C.V. %)			% of patients who were abnormal	Mean (C.V. %)			% of patients who were abnormal
		Overall	Within normal range	Abnormal		Overall	Within normal range	Abnormal	
Visual Acuity (Snellen)	1.31 (39)	1.10 (18)	1.10 (18)	—	0	0.82 (39)	0.9 (29)	0.12 (52)	10

Prior to now we were unable to stratify mathematically a group of patients into those with normal and those with subnormal function. It is realized that the dividing point between normal and subnormal function at present may not be a valid one, since we do not think that we have a large enough "normal" group of subjects to compare to the multiple sclerosis patients. This cut off point will

TABLE 13
FUNCTION OF THE UPPER RIGHT EXTREMITIES OF AMBULATORY UNTREATED MULTIPLE SCLEROSIS PATIENTS USED IN THIS STUDY TO EVALUATE THE APPLICABILITY OF THE PROPOSED TEST BATTERY

Tests	8 Normal male right-handed medical students W.W.T. data Mean (C.V. %)	Ambulatory Multiple Sclerosis Patients							
		Group 1				Group 2			
		10 Outpatients				40 Inpatients			
		Mean (C.V. %)			% of patients who were abnormal	Mean (C.V. %)			% of patients who were abnormal
		Overall	Within normal range	Abnormal		Overall	Within normal range	Abnormal	
Speed of hands (strokes per 10 seconds)									
2nd trial	77 (7)	57 (12)	68	57 (11)	90	52 (19)	69 (5)	48 (27)	87
Fatigue (4th trial)	65 (9)	49 (11)	58	47 (8)	80	45 (30)	58 (8)	45 (22)	80
Coordination of hand (strokes per 10 seconds)									
2nd trial	46 (22)	28 (34)	32 (12)	20 (37)	40	24 (33)	32 (14)	19 (24)	63
Fatigue (4th trial)	40 (19)	27 (31)	32 (13)	21 (28)	40	22 (28)	28 (8)	19 (24)	58
Strength (pounds)									
Grip									
1st trial	100 (26)	63 (42)	73 (29)	33 (37)	25	64 (53)	73 (36)	40 (29)	25
Fatigue (5th trial)	77 (22)	47 (44)	57 (26)	19 (50)	25	49 (54)	58 (34)	28 (35)	30
Wrist	58 (16)	38 (26)	49 (21)	34 (19)	70	33 (33)	47 (13)	29 (24)	80
Deltoid	58 (13)	30 (34)	56 (30)	28 (29)	80	29 (46)	55 (14)	24 (31)	85
Sensation									
Vibration (microns)	0.1 (41)	1.7 (224)	0.1 (73)	2.8 (73)	70	2.8 (410)	0.2 (19)	4.8 (264)	57
Two point (mm.)	3.2 (10)	6.7 (43)	3.8 (13)	8 (27)	60	5.8 (44)	3.8 (17)	6.5 (30)	75

TABLE 14
FUNCTION OF THE RIGHT LOWER EXTREMITIES OF AMBULATORY UNTREATED MULTIPLE SCLEROSIS PATIENTS USED IN THIS STUDY TO EVALUATE THE APPLICABILITY OF THE PROPOSED TEST BATTERY

Tests	8 Normal male right-handed medical students W.W.T. data Mean (C.V. %)	Ambulatory Multiple Sclerosis Patients							
		Sample 1				Sample 2			
		10 Outpatients				40 Inpatients			
		Mean (C.V. %)			% of patients who were abnormal	Mean (C.V. %)			% of patients who were abnormal
		Overall	Within normal range	Abnormal		Overall	Within normal range	Abnormal	
Station stability (seconds)									
Eyes open	> 30	13 (91)	> 30	7 (149)	80	5 (140)	> 30	4 (146)	97
Eyes closed	> 30	4 (151)	—	4 (154)	100	1 (125)	—	1 (125)	100
Speed of foot (strokes per 10 seconds)									
2nd trial	53 (11)	24 (57)	47	21 (65)	90	17 (57)	42	16 (48)	97
Fatigue (5th trial)	39 (9)	18 (60)	32	16 (61)	90	14 (55)	32	13 (50)	97
Strength (pounds)									
Foot	> 84	49 (61)	—	49 (61)	100	35 (56)	84	34 (53)	97
Hip flexor									
1st trial	61 (30)	24 (42)	30 (21)	15 (32)	60	20 (66)	35 (47)	11 (51)	62
Fatigue (5th trial)	57 (30)	18 (57)	29 (4)	12 (66)	70	16 (79)	33 (30)	10 (52)	75
Sensation									
Vibration (microns)	0.5 (70)	19 (127)	1.0	23 (115)	90	30 (77)	0.7 (43)	34 (69)	87

TABLE 15
PER CENT OF UNTREATED AMBULATORY MULTIPLE SCLEROSIS PATIENTS WITH ABNORMAL RESULTS FOR THE VARIOUS TESTS (RIGHT SIDE)

Per cent of patients with abnormal results	Test	
	Inpatients	Outpatients
90–100	Station stability (eyes open and closed) Foot speed (also fatigue) Foot strength	Station stability (eyes closed) Foot speed (also fatigue) Foot strength Toe vibration sense Hand speed
76–89	Toe vibration sense Hand speed (also fatigue) Deltoid strength Wrist strength	Station (eyes open) Hand speed fatigue Deltoid strength
51–75	Hip strength (also fatigue) Finger two point discrimination Finger vibration sense Hand coordination (also fatigue)	Wrist strength Hip strength (also fatigue) Finger two point discrimination Finger vibration sense
31–50		Hand coordination (also fatigue)
0–30	Grip strength (also fatigue) Visual acuity	Grip strength (also fatigue) Visual acuity

become more valid when we are better able to define what is normal function. It is reasonable to use the data obtained by WWT, however, for the following reasons: he examined all the multiple sclerosis patients, the data are based on our latest definition of the neurological tests, and they are in reasonable agreement with other data on college students (see TABLES 6, 7 and 8). The objections to the WWT data as normal for the multiple sclerosis patients are as follows: The sample size is small (eight patients), the individuals were all right-handed, they were all males, and since they were all college students they probably represent a select stratum of the population, different from the multiple sclerosis patients.

The visual acuity data in TABLE 12 show no significant abnormalities in the outpatient group. On the other hand, the inpatient group had 10 per cent of patients who were abnormal.

The remainder of the data for the test battery are presented in TABLES 13 and 14. It can be seen that all functions tested are unusually abnormal when compared to the normal results. Inspection of the coefficients of variation for the two multiple sclerosis groups showed a marked increase over the values which we found in normal individuals. For example, note speed of foot: the coefficient of variation was 11 per cent for normal individuals and 57 per cent for both in- and outpatients. This is an expected result, since some of the patients have practically normal foot speed and others may have practically none.

It was of great interest to note that the fatigue tests we have designed (hand speed, hand coordination, grip strength, foot speed and hip flexor strength) did

not show as much fall-off as expected when compared to the normal data. No experiments have been designed to explain for this unexpected finding.

The data in TABLE 15 summarizes some of the data presented in TABLES 12, 13 and 14. It should be noted that 90-100 per cent of patients had difficulty with their lower extremities (station stability [eyes open and closed], foot speed [also fatigue] and strength). On the other hand, visual acuity and grip strength (also fatigue) was least affected (0-30 per cent). We are concluding from these data that the high incidence of abnormality of some tests in the proposed test battery should suggest that they are appropriate to evaluate quantitatively the neurological status of multiple sclerosis patients.

In view of the fact that abnormalities can be discovered with the proposed test battery the following questions arise: Are they important functional abnormalities in relation to daily tasks of living? Are we quantitating important functions or are we evaluating "bits" that do not indicate integrated neurological function? Since we cannot answer these questions now with our limited experience, we have carried out the classical neurological examination on all patients and summarized this in terms of Kurtzke's scores.

An experiment was designed to provide information on the extent of learning and fatigue associated with repeated testing of patients with multiple sclerosis. The 10 multiple sclerosis outpatients were examined once a week for three weeks. The time-response curve analysis[5, 6] was used to test the following hypothesis: there were no significant linear regression (learning or fatigue) trends over the period of the three examinations. The following tests were found to show a significant learning effect from the first to the third examination: hand coordination fatigue (119 per cent of original); grip strength fatigue (110 per cent); finger two-point discrimination (110 per cent); and in the lower extremity only hip flexor strength and fatigue (121 and 137 per cent of original, respectively). No tests had a significant fatigue trend from the first to the third examination.

If we would decide arbitrarily that a change to below 80 or above 120 per cent of original values were a technically significant difference, then only the hip flexor strength and fatigue have increased significantly from the first to the third examination; this result is similar to that found in normal individuals where the hip flexor strength test increased to 121 per cent of the original value after more vigorous repetitive testing (see TABLE 9).

We have concluded from this limited study on 10 multiple sclerosis outpatients that the proposed test battery is not affected by weekly repetition of examinations, except hip flexor strength and hip flexor strength fatigue which improve with repetitive testing; these results are similar to more extensive studies on normal individuals.

It would appear from these studies on ambulatory multiple sclerosis patients that we can use the proposed battery of tests in a short term therapeutic trial (up to three weeks) and probably can use it to follow the course of the disease on a long term basis (years). It has also occurred to us that it might be possible to determine the depth of a relapse if pre-relapse quantitative testing had been carried out; we have had one patient (M.G.) which illustrated this usage of the test battery.

Therapeutic Trial

A study which used the proposed battery of quantitative neurological tests in a therapeutic trial in multiple sclerosis was carried out for one year, July 1962 to June 1963; 40 patients were selected by preprinted criteria for the therapeutic trial;[3] the design of the experiment was presented in the section entitled METHODS.

When the data obtained from the five treatment groups were analyzed statistically by the paired Student t test no differences ($p > 0.05$) between the neurologic tests obtained on the first examination (no drug) and the second examination (high dosage for one week) and/or third examination (dosage tapered to zero over one week period) were discovered.

The data obtained from the five treatment groups also were analyzed statistically by the analysis of variance techniques. It was found that the patients assigned at random to the five treatment groups were not significantly different in their mean performance for each of the neurologic tests before the administration of the drugs (examination one) except for a technically significant difference for foot strength. Furthermore, average hand speed was significantly increased in the group treated with Medrol® after the second examination, but this difference did not hold up after the medication was decreased (third examination).

Based on the above results the bio-statistician in this study, Jan W. Kuzma, recommended the following: "It appears that there are trends and that real differences may be discovered, if they exist, if the sample size for each group would be about 25. This would be another 60 patients assigned at random to the five groups. Then this type of analysis should be performed again." At the rate we admitted patients to this study it would take another 18 months. Perhaps a cooperative study among several neurological centers should be considered in order to expedite this type of study.

Since this recommendation could not be carried out, the data were reanalyzed in a different way. First, patients treated with Solu-medrol,® Medrol® and Alphadrol® were combined and relabeled steroid treatment group; the three treatment groups were now vitamin B_{12}, seven patients; DEPO-ACTH,® eight patients; and steroids, 25. This is a reasonable manipulation in view of the fact that the three synthetic adreno-cortoids probably have an identical mechanism of action. Second, the right and left sided data were combined; rather than compare right to right and left to left. This may be a reasonable procedure, since multiple sclerosis is an asymmetrical (patchy) affliction and therefore one could make the assumption that right and left-sided may not be dependent. Third, the right and left-sided data for a given neurological test at the initial examination for a given treatment were then stratified into results which were within normal range (mean of W.W.T. $\pm$ 2 standard deviations) or subnormal. These two groups of data were then compared by the paired Student t test to results obtained for examination 2, as well as examination 3. The separation of the data into normal function versus subnormal function has some merit; it is reasonable to assume that a patient who has normal hand speed will not improve further, whereas diseased tracts which affect hand speed and hence result in a subnormal hand speed are more likely to have a favorable treatment response. In other words, if the data which are least likely to be affected are removed prior to statistical analysis, the statistical tool becomes more powerful and hence is likely to discover smaller but real changes.

The results of combining the data and the statistical analyses for each test and each treatment group are shown in TABLE 16; 55 statistical differences were discovered; however, we arbitrarily decided that any change to a value between 80 to 120 per cent of the initial examination was only a technically statistical difference. The underlined data in the table were "neurologically," as well as statistically significant; there were 32 differences. There may be a pattern to the data. All changes were in the direction of improvement except toe vibration in the vitamin B_{12} group for the second examination. Also, more significant improvements were discovered in the lower extremities than in the upper extremities; in the upper extremities none of the "neurological," as well as statistical improvements were found in both examination 2 and 3, whereas in the lower extremities this was the

rule. In the lower extremity vitamin B_{12} improved two functions on both examinations, namely, station stability (eyes open) and speed of foot; ATCH improved four different functions, namely, foot speed fatigue, foot strength, hip flexor strength and fatigue; and steroids improved the same two as vitamin B_{12}, as well as those improved by ACTH. If these are significant neurological improvements, it would appear that steroids and probably ACTH do more than vitamin B_{12}; other than the improvements in station stability (eyes open) and foot speed which may be the placebo effect; foot speed fatigue, hip flexor strength and hip flexor strength fatigue may improve to 124 to 196 per cent of the first examination when the patient is on ACTH or steroids.

We also analyzed the test results which fell into the normal range for the three treatments in a manner similar to that mentioned above; there were only three values which were considered "neurologically" and statistically significant. They occurred only in the steroid group. They were as follows: hip flexor strength 126 per cent exam 1 versus exam 2 and hip flexor strength fatigue, 121 and 124 per cent, exam 1 versus 2, and exam 1 versus 3, respectively. These data confirm our previously stated impression that test results which fall into the normal range are unlikely to be improved by treatment.

Because of the above results and those reported by Kibler[4] in this symposium, we reworked the data again. We went through the clinical records of the patients used in this therapeutic trial and chose those patients who stated they had a significant weakness in the lower extremities less than two months before the initial examination in the therapy study. Furthermore, the patient had to have normal to subnormal function prior to the relapse; he had to be able to walk at least one block without the aid of appliances. Three patients in the Depo-ACTH,® and nine patients in the steroid group (Solu-Medrol,® Medrol,® or Alphadrol®) were admitted to this study; unfortunately no patients in the vitamin B_{12} group were accepted.

The data shown in TABLE 17 included only those improvements which were neurologically significant, i.e. the paired Student t test from one examination to another had to indicate a significant difference ($p < 0.05$) and then the change had to fall outside of 80-120 per cent of the original values. The improvements fit about the same pattern as gleaned from the data presented in TABLE 16; i.e. most of the improvement was in the function of the lower extremities; however, some of the tests do show much greater improvement. No test results significantly worsened during the course of this therapeutic trial.

The patients were also rated neurologically according to the system recommended by Kurtzke.[2] In most cases the patients did not change their total disability score or systems score from one examination to another, even though it was obvious by clinical neurological examination and quantitative neurological examination that the patient was better.

It was our impression that this therapeutic trial warrants continuation; more patients should be obtained to determine whether the improvements produced by ACTH and steroids are better than a vitamin treatment (vitamin B_{12}).

In TABLE 18 are recorded the special conditions occurring in the patients during the study period. Nothing serious was observed, even though the steroid dosages were very large.

There are many objections to the therapeutic trial which we have conducted so far. It is obvious that more patients must be studied; perhaps the battery of tests proposed is irrelevant; and maybe a different dosage schedule should be used.

If a significant neurological improvement is established, it will then be necessary to determine whether this improvement is related to the pathogenesis of multiple sclerosis or whether it is symptomatic.

TABLE 16

SUMMARY OF THE RESULTS AND THE STATISTICAL ANALYSES OF THE BATTERY OF TESTS CARRIED OUT IN THE THERAPEUTIC TRIAL; RIGHT AND LEFT SIDED DATA WERE COMBINED;[1] ONLY THE ABNORMAL TEST RESULTS ARE PRESENTED; AND THE PATIENTS TREATED WITH MEDROL,® SOLU-MEDROL,® AND ALPHADROL® WERE COMBINED AND LABELLED STEROID GROUP. UNDERLINED VALUES (PER CENT CHANGE IN FUNCTION OF ORIGINAL VALUES $\geqq 120$ TO $\leqq 80$) ARE ARBITRARILY CONSIDERED CLINICALLY SIGNIFICANT

Function	Treatment Group	Exam 1 Per cent of patients abnormal (mean value; C.V.[2])	Exam 1 vs 2[3] Per cent of original function if significant ($p < 0.05$)	Exam 1 vs 3[3] Per cent of original function if significant ($p < 0.05$)
Visual acuity (Snellen)	Vit. B_{12} (7 patients)	7 (0.23;—)	—	—
	ACTH (8 patients)	0	—	—
	Steroids (25 patients)	12 (0.1;48)	N.S.[4]	N.S.
Hand speed (strokes per 10 seconds)	Vit. B_{12}	100 (46;26)	104	N.S.
	ACTH	100 (44;27)	N.S.	116
	Steroids	92 (49;23)	109	109
Hand speed fatigue (strokes per 10 seconds)	Vit. B_{12}	93 (38;23)	N.S.	N.S.
	ACTH	100 (38;22)	N.S.	110
	Steroids	82 (40;24)	110	113
Hand speed coordination (strokes per 10 seconds)	Vit. B_{12}	93 (21;16)	N.S.	N.S.
	ACTH	69 (19;17)	N.S.	N.S.
	Steroids	56 (19;24)	122	119
Hand speed coordination fatigue (strokes per 10 seconds)	Vit. B_{12}	93 (19;19)	109	112
	ACTH	62 (20;13)	N.S.	N.S.
	Steroids	52 (17;26)	N.S.	116
Grip strength (pounds)	Vit. B_{12}	29 (36;21)	N.S.	N.S.
	ACTH	38 (42;11)	122	117
	Steroids	22 (39;29)	117	114
Grip strength fatigue (pounds)	Vit. B_{12}	50 (29;19)	N.S.	N.S.
	ACTH	38 (30;10)	132	118
	Steroids	56 (32;25)	118	115
Wrist strength (pounds)	Vit. B_{12}	71 (29;32)	N.S.	N.S.
	ACTH	94 (29;18)	N.S.	N.S.
	Steroids	82 (27;25)	N.S.	106
Deltoid strength (pounds)	Vit. B_{12}	79 (26;38)	N.S.	N.S.
	ACTH	100 (26;28)	109	112
	Steroids	84 (23;29)	127	116

Finger vibration sensibility (microns)	Vit. B_{12}	43 (1.4;56)	N.S.	N.S.
	ACTH	69 (0.7;42)	N.S.	N.S.
	Steroids	52 (3;180)	N.S.	N.S.
Finger two point discrimination (mm.)	Vit. B_{12}	86 (8.5;29)	N.S.	80
	ACTH	93 (7; 9)	N.S.	N.S.
	Steroids	76 (6;22)	N.S.	N.S.
Station stability; eyes open (seconds)	Vit. B_{12}	100 (5;56)	167	156
	ACTH	100 (2;76)	N.S.	N.S.
	Steroids	96 (6;110)	121	136
Station stability; eyes closed (seconds)	Vit. B_{12}	100 (2;126)	N.S.	N.S.
	ACTH	100 (1; 96)	N.S.	N.S.
	Steroids	100 (1;108)	N.S.	N.S.
Hip flexor strength (pounds)	Vit. B_{12}	64 (17;16)	N.S.	N.S.
	ACTH	78 (10;74)	173	196
	Steroids	56 (12;46)	144	144
Hip flexor strength fatigue (pounds)	Vit. B_{12}	78 (12;29)	N.S.	N.S.
	ACTH	81 (8;66)	158	153
	Steroids	66 (10;52)	141	143
Foot strength (pounds)	Vit. B_{12}	100 (41;56)	115	N.S.
	ACTH	100 (29;70)	122	122
	Steroids	98 (34;51)	113	111
Foot speed (strokes per second)	Vit. B_{12}	100 (16;62)	121	123
	ACTH	100 (11;60)	N.S.	N.S.
	Steroids	98 (17;49)	122	127
Foot speed fatigue (strokes per 10 second)	Vit. B_{12}	100 (14;68)	N.S.	N.S.
	ACTH	100 (8;72)	135	142
	Steroids	100 (11;55)	134	134
Toe vibration sensibility (microns)	Vit. B_{12}	86 (37;78)	129	N.S.
	ACTH	87 (32;88)	50	N.S.
	Steroids	82 (30;76)	N.S.	87

[1]Since multiple sclerosis is a patchy affliction, the data were combined under the assumption that right and left-sided observations were not dependent. These data should be interpreted with this assumption in mind.

[2]C.V. means coefficient of variation; standard deviation × 100/mean.

[3]Used paired student t test.

[4]N.S. means not significant, $p > 0.05$.

TABLE 17
SUMMARY OF THE RESULTS AND THE STATISTICAL ANALYSES OF THE BATTERY OF TESTS CARRIED OUT IN THE THERAPEUTIC TRIAL OF THOSE PATIENTS JUDGED TO HAVE A RELAPSE OF WEAKNESS IN THE LOWER EXTREMITIES

Tests*	Treatment Group	Exam 1 Mean (C.V. %)	Exam 1 vs. 2 per cent of original function, if significant	Exam 1 vs. 3 per cent of original function, if significant
Hand coordination	Steroid	18 (27)	135	135
Grip strength fatigue	ACTH	29 (135)	147	123
Station stability (Eyes open)	Steroid	7 (48)	137	172
Hip flexor strength	ACTH Steroid	6 (98) 19 (9)	210 125	250 129
Hip flexor strength fatigue	Steroid	13 (57)	184	193
Foot strength	ACTH	21 (82)	166	195
Foot speed	ACTH Steroid	21 (48) 6 (130)	126 250	133 280
Foot speed fatigue	ACTH Steroid	3 (171) 16 (38)	300 131	370 137
Toe vibration	Steroid	32 (120)	73	N.S.

*Units for the test are the same as shown in TABLE 16.

SUMMARY

Evidence has been presented that certain neurological functions in normals and ambulatory patients with multiple sclerosis are measurable quantities (visual acuity, 10 upper extremity functions and eight lower extremity functions). There was no significant learning or fatigue effect on repetitive testing except hip flexor strength. The battery of tests were teachable to para-medical personnel.

The therapeutic trial which was conducted has not shown any significant differences between vitamin B_{12}, Depo-Acth,® Solu-Medrol,® Medrol® and Alphadrol;® however, if only patients with subnormal function were analyzed, a number of significant differences especially in the lower extremities, were discovered in the ACTH and steroid groups, and fewer in the vitamin B_{12} group. More patients will have to be studied to determine if the significant differences discovered are not the result of sampling errors.

ACKNOWLEDGMENTS

Some aspects of this study were considered by J. W. Kuzma in his dissertation as a partial fulfillment of the requirements for doctor of philosophy in the Department of Biostatics, University of Michigan.

TABLE 18
SPECIAL CONDITIONS OBSERVED IN PATIENTS DURING STUDY PERIOD

Special Conditions	Vitamin B_{12}	Depo-ACTH®	Solu-Medrol®	Medrol®	Alphadrol®
Weight change (average pounds)	0	−1	−5	−5	−4
Edema	12*	0	0	0	0
Moon-face	0	38	12	12	22
Acne	0	0	0	12	22
Blood pressure rise (diastolic $\geqq$ 90)	0	25	12	25	11
Blood sugar rise ($\geqq$ 102)	0	25	12	25	0
Mood changes	43	38	38	12	44
Infection	0	0	0	**	0
Diarrhea	0	25	12	12	11
Postlumbar puncture Headache	14	50	38	25	22
Number of patients	7	8	8	8	9

*% of patients in group.
**1 Abscess injection site; 13th day.

The authors would like to acknowledge the technical assistance of Mrs. Elza Rencis Bryan and Mr. Irwin Oleinick who gave assistance in the collection and reduction of the data, and Mrs. Louise Dapprich for preparing the program and data checks. The authors are also grateful to Dr. R. D. Remington for the valuable guidance regarding many statistical aspects of this study.

REFERENCES

1. ALEXANDER, L. 1951. New concept of critical steps in course of chronic debilitating neurologic disease in evaluation of therapeutic response: A longitudinal study of multiple sclerosis by quantitative evaluation of neurologic involvement and disability. A.M.A. Arch. Neurol. Psychiat. **66:** 253-271.
2. KURTZKE, J. F. 1961. On the evaluation of disability in multiple sclerosis. Neurology **11:** 686-694.
3. SCHUMACHER, G. 1965. Whither research in demyelinating disease? This Annal.
4. KIBLER, ROBERT F. 1965. Large dose cortico steroid therapy of experimental and human demyelinating diseases. This Annal.
5. KUZMA, J. W. 1963. A statistical study of various aspects of a battery of clinical neurological tests. Ph.D. Thesis. Univ. of Michigan. Ann. Arbor, Mich.
6. KUZMA, J. W., TOURTELLOTTE, W. W. & REMINGTON, R. D. 1964. Quantitative clinical neurological testing. II. Some statistical aspects. J. Chronic Dis. In press.