NEUROLOGY 1989;39:291-302

The Disability Status Scale for multiple sclerosis:

Apologia pro DSS sua

John F. Kurtzke, MD

The Disability Status Scale (DSS),[1] with its associated Functional Systems (FS)[2,3] and the latter's final revisions with the expanded DSS (EDSS),[4] has provided a system for encoding the neurologic deficits in MS which has had considerable use. It may be of value, then, to review the evolution of this system and its bases in the standard neurologic examination and the neuropathology of this disease.

In 1953, we were attempting to evaluate the efficacy of a new treatment in MS, the use of isoniazid (INH).[5] It was obvious that we needed, first, a group of patients to compare the subjects with, and, second, a way to measure the effects of therapy. As to the latter, the only method of overall assessment that was then available and that might have been applicable in a treatment trial was that of Alexander.[6] This comprised a complex schema of 30 neurologic signs and disabilities, some duplicated, with the items given arbitrary and differing numbers. The sum of these scores then defined the patient's neurologic status. This proved both unworkable and undesirable.

It was necessary, therefore, to devise a new method. The first step was to define what indeed are the deficits found at neurologic examination in this disorder. This was considered the primary question, since the entire basis of clinical neurology is that signs and symptoms indicate the localization of lesion(s) in the neuraxis. As a corollary, the more signs the more numerous or extensive should be the lesions, and the fewer signs, the fewer or smaller. Because symptoms can be unreliable, the goal was a system that was limited to objectively verifiable deficits. But it had to be one that would categorize the deficits seen in MS within a manageable number of groupings, each one mutually exclusive of all the others, and yet that together would define all the abnormalities of the nervous system clinically manifest in this disorder. The exclusivity was necessary to avoid repeatedly counting the same phenomena.

Available for this purpose were the records of nearly 250 men, all veterans of World War II, who had been hospitalized for MS at the VA Hospital, Bronx, New York between 1944 and 1953. What was unique to this material was its completeness. At that time, neurology at the Bronx VA was an affiliate of Cornell, and H.G. Wolff of Cornell was director of training. At Cornell, the medical students then received, in the sophomore year, a course in neurologic diagnosis distinct from physical diagnosis. There was a 54-page typed neurologic examination format that had to be followed, necessitating some 20 pages of handwritten material per patient. As part of this, all negative findings had to be listed as well as all the positives. Further, the examination was organized by body parts (head, neck, upper extremities, trunk, lower extremities, pelvis). Included, of course, was a rather detailed assessment of mental status. I still provide this examination outline for the use of our own medical students and residents. While then, however, not quite so extensive, this methodology carried over to the neurologic units at New York Hospital and the VA. Anyone who knew Wolff will recall his insistence on precision, accuracy, and thoroughness. Therefore, the VA records provided for findings present *and* absent, with measures of severity, and organized by the part(s) of the body affected.

After recording all the individual findings among these patients, it was possible to consolidate them into eight mutually exclusive types of impairment. Table 1, first published in 1956,[7] shows an overview of these 1953 data for that subset of the MS patients that we had used as controls for the INH study.[5] The distributions by step within each type were not provided in reference 7, and I no longer have these data. Since the unit of study is the patient and not one or several types of impairment, it was still necessary to provide a single figure that described the patient as a whole. However, when scores were added for each type of impairment to give a single summed number for each patient, it was clear that the sum rapidly plateaued, far below the theoretical maximum score, and it then bore little resemblance to the obvious impairments the severely involved patients had. Further, when the hospital admission and discharge examinations (the latter also quite complete)

From the Neurology Service and Neuroepidemiology Research Program, Veterans Administration Medical Center, and the Departments of Neurology and of Community and Family Medicine, Georgetown University School of Medicine, Washington, DC.

Received August 2, 1988. Accepted for publication in final form August 8, 1988.

Address correspondence and reprint requests to Dr. Kurtzke, Neurology Service (127), VA Medical Center, Washington, DC 20422.

Table 1. Neurologic involvement by type of impairment* in 172 patients with MS admitted to VA Hospital, Bronx, NY, 1944-1953 for exacerbations of 2 years' duration or less (mean = 6.0 months) and with illness of up to 300 months' total duration (mean = 59.7, median = 30 months). Modified from Kurtzke 1956.[7]

Type of impairment*	N	*p*N
Pyramidal	152	0.8837
Cerebellar	135	0.7849
Brainstem	144	0.8372
Sensory	110	0.6395
Sphincter	83	0.4826
Others	61	0.3547

* "... 0 to 5 scale for each of ... pyramidal, cerebellar tract, brainstem, sensory, bowel and bladder, and others (miscellaneous plus cerebral, including optic)."[7]

Table 2. Frequency of involvement by Functional System (FS) at diagnostic bout admission examination, US Army WW II hospital series; 335 definite MS with complete information on all eight FS. Data of Kurtzke et al 1972.[13]

Functional System	Involvement			
	Yes (1)	No (0)	*p* (1)	*p* (0)
Major				
Pyramidal (P)	276	59	0.82388	0.17612
Cerebellar (Cll)	257	78	0.76716	0.23284
Brainstem (BS)	253	82	0.75522	0.24478
Sensory (S)	168	167	0.50149	0.49851
Minor				
Bowel & Bladder (BB)	66	269	0.19701	0.80299
Visual (V)*	116	219	0.34627	0.65373
Cerebral (Cb)†	71	264	0.21194	0.78806
Other (O)	43	292	0.12836	0.87164

* Neuropathic signs.[11]
† Includes mood change only (grade 1).

were compared, changes in the overall scores did not reflect adequately the changes in the patients. In addition, more than a few patients improved in one area while worsening in another. Also, as I was slow to realize, these are *not* true numbers that can properly be added, but rather separate rank-order scales with no assurance whatsoever that one scale is equivalent to another.

It was for all these reasons that the Disability Status Scale[1] was devised as an overall measure of neurologic involvement in this disease, which, however, was based on these separate types of impairment—soon to be called functional groups[2] and then Functional Systems.[3] At that time, the types of impairment were looked on as merely a way to consolidate the findings in order to grade the DSS. It was the DSS that was then used as the primary measure in our trial of INH, and the results, when compared with the group of equivalent MS patients previously hospitalized, were strikingly positive. Our paper, submitted in August 1954, was published in October 1954.[5] It aroused considerable interest, and it was in fact the precipitant for the first randomized, placebo-controlled, double-blind, multicentered therapeutic trial ever conducted in MS.[8] Here too the DSS was used to measure change (quite successfully), with several other methods as well. This study proved to be a flat negative as to INH, no matter how one looked at the results. And indeed our own later findings led us also to conclude that, regardless of dose, INH has no demonstrable effect in MS.[9]

Functional Systems. While the DSS has proved useful, not only in therapeutic trials but also in measures of disease severity, course, and prognosis, its validity as a means of quantitating the amount of lesioned tissue present has always rested upon the FS.

To explore further the FS versus MS pathophysiology, we were fortunate to have available another unique resource.[10] There were 762 men whose final diagnosis was MS admitted to US Army hospitals during the World War II period. During the war, almost all the neurologic talent in this country (at least those below age 45 or 50) were in the service. The named General (Army) Hospitals were, for the most part, transplanted University faculties. A diagnosis of MS was the basis for medical discharge from service—a step not taken lightly during the war. Most soldiers in whom the diagnosis of MS was entertained were therefore transferred to the General Hospitals or other major units for evaluation, and virtually all were examined by neurologists or neuropsychiatrists. In addition, because of the administrative requirements for a medical discharge, the documentation of findings was unusually complete, so much so that there was little difficulty in assigning FS and DSS scores for each examination, with the sole exception of the Visual FS. This last requires corrected acuities *and* mapped-out visual fields for its employment (standard in the Cornell axis but not in the Army). We therefore had to rely on a measurement of neuropathic optic signs[11] for the Visual Functions instead of the scale.

Based also on special follow-up neurologic examinations and extensive Veterans Administration records, we reassessed the diagnosis in these 762 men in the early 1960s and classified 476 of them as definite MS and another 51 as probable MS,[10] using, for the former, the criteria later adopted by the Schumacher panel.[12]

Signs, symptoms, and laboratory features at the time of the original Army diagnosis were explored in detail.[13] The bout for which the patients were then hospitalized, called the (Army) diagnostic bout, occurred in all instances in soldiers on active duty and free of neurologic complaints before that episode, even though previous (undiagnosed) bouts had taken place in 56% of the 527 MS patients. For those with prior bouts, median duration from onset to diagnosis was 39 months, though for their diagnostic bout the median was 2 months. The median to diagnosis for those whose onset bout was the diagnostic bout was 3 months. Thus, the diagnostic bout data provided a cross-sectional view of the clinical phenomena in a nationwide series of early MS, which were uncontaminated by preexisting deficits.

Table 3. Specific patterns of involvement (1) or no involvement (0) within each of eight Functional Systems (FS) at diagnostic bout admission examination, US Army WW II hospital series; 335 definite MS with complete information on all FS. Data of Kurtzke et al 1972.[13]

FS patterns: Major \ Minor		(BB)	(V)	(Cb)	(O)	FS patterns											
	0000	1000	0100	0010	0001	1100	1010	1001	0101	0110	0011	1110	1101	1011	0111	1111	Total
	0000	—†	1	3	—	—	—	—	—	—	1	—	—	—	—	—	—	5
(P)	1000	3	3	1	—	1	—	—	—	—	—	—	—	—	—	—	—	8
(Cll)	0100	1	—	1	1	—	1	—	—	1	—	—	—	—	—	—	—	5
(BS)	0010	6	—	9	—	1	—	—	—	—	—	—	—	—	—	—	—	16
(S)	0001	—	—	1	—	—	—	1	—	—	—	—	—	—	—	—	—	2
	1100	8	2	2	1	1	—	—	1	—	2	—	—	—	—	—	—	17
	1010	7	—	5	1	1	1	—	—	—	—	1	—	—	—	—	—	16
	1001	5	2	3	2	2	2	—	1	—	—	—	—	—	—	—	—	17
	0101	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	1
	0110	9	1	5	1	1	1	—	—	—	3	1	—	—	—	—	—	22
	0011	1	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	2
	1110	29	4	16	8	3	4	2	—	2	6	—	1	—	3	—	—	78
	1101	14	3	2	2	2	1	—	2	—	—	—	1	—	—	—	—	27
	1011	6	—	3	—	1	1	—	—	—	1	—	—	—	—	—	—	12
	0111	1	—	1	1	—	1	—	—	1	—	—	—	—	—	1	—	6
	1111	31	11	12	15	7	5	2	2	2	6	1	4	1	1	1	—	101
Total		121	27	65	32	21	17	5	6	6	19	3	6	1	4	2	—	335

Major FS: Pyramidal (P), Cerebellar (Cll), Brainstem (BS), Sensory (S).
Minor FS: Bowel & Bladder (BB), Visual (V), Cerebral (Cb)*, Other (O).

* Includes grade 1 (mood).
† Number of patients.

Table 2 describes the frequency of involvement within each FS among all 335 definite MS with complete information as to presence (and severity) or absence on all eight FS. The first four FS (Pyramidal, Cerebellar, Brainstem, Sensory) are considered the major FS because of their frequencies, in each instance clearly in excess of those for the other four, which are then called the minor FS. If we denote involvement on any FS by a "1" and no involvement by a "0," then each patient's neurologic status can be defined with an 8-digit binary number. For example, 1101 0000 would describe a patient with Pyramidal + Cerebellar + Sensory involvement, all other FS being normal.

Table 3 shows the observed frequencies of all FS patterns in this series at the diagnostic bout examination. There are 256 (2^8) possible patterns of such involvement into which a patient may fall. Obviously, these patterns are not equally distributed among the patients. It is clear that patterns with multiple (three or four) major FS involved are the most common, whereas, even among those, the minor FS are far more often singly affected; indeed that holds true also for patients with only one or two major FS being abnormal. Single-system involvement was rare: only 14 such occurred, with ten of these major FS, even in this early bout of illness.

Is there any biologic significance to the specific patterns observed? Is involvement in one FS dependent on involvement in another? To answer the latter, the null hypothesis of independence (that each FS is independent of the others) must be rejected. Anatomically, they certainly are separate and mutually exclusive, but that does not imply that their coincidence would also be so. In fact, *a priori* judgment would favor the opposite view, that certain FS should "go together" with certain others. If each FS is independent, then the product of each of the observed frequencies for each of the eight will define the proportion of cases expected to have that specific pattern. For example, the pattern of Pyramidal + Cerebellar + Brainstem + Sensory, all others normal, is depicted as 1111 0000. Under the hypothesis of independence, the expected frequency for that specific pattern is, from table 2, 0.82388 (P+) × 0.76716 (Cll+) × 0.75522 (BS+) × 0.50149 (S+) × 0.80299 (BB−) × 0.65373 (V−) × 0.78806 (Cb−) × 0.87164 (O−) = 0.086316. For the 335 cases, the expected number of cases with this specific pattern is thus 28.916 cases (335 × 0.086316); there were 31 observed. For Pyramidal alone, the expectation is 2.828 cases (335 × 0.008441), and three were observed. Expectations for each of the 256 possibilities were similarly calculated.

Table 4 lists all the patterns for which the expected number is 5.0 or more, ranked according to the expected frequencies. When compared with the frequencies actually observed for these same patterns, then, with a Chi-square goodness-of-fit test, the hypothesis of independence is clearly not rejected ($\chi^2_{14} = 20.58$, $p > 0.10$). Figure 1 shows the continued agreement between observed and expected numbers for each specific pattern through the first 86 of the 256 patterns, which 86 include 90% of all the cases: 90.31% of expected, 90.45% of observed. The close relation of O versus E persists even with expectations well under 1.0 case. The other data underlying this figure are in table 5. Note the persistence of major FS patterns.

Only 14 patterns, observed or expected, describe the

Table 4. Patterns* of involvement by Functional System (FS) in US Army WW II hospital series of MS at diagnostic bout, with observed (O) and expected (E) numbers of cases for all patterns with E ≥ 5.0. Modified from Kurtzke 1983.[4]

Rank†	Pattern*	O	E	Cumulative p‡ O	Cumulative p‡ E
1	1111 0000	31	28.916	0.09	0.09
2	1110 0000	29	28.774	0.18	0.17
3	1111 0100	12	15.316	0.21	0.22
4	1110 0100	16	15.225	0.26	0.26
5	1101 0000	14	9.372	0.30	0.29
6	1100 0000	8	9.316	0.33	0.32
7	1011 0000	6	8.776	0.35	0.35
8	1010 0000	7	8.724	0.37	0.37
9	1111 0010	15	7.777	0.41	0.40
10	1110 0010	8	7.730	0.44	0.42
11	1111 1000	11	7.094	0.47	0.44
12	1110 1000	4	7.052	0.48	0.46
13	0111 0000	1	6.181	0.48	0.48
14	0110 0000	9	6.145	0.51	0.50
15+	All other	164	168.610	1.00	1.00
256	Total	335	335.008	1.00	1.00

* Specific pattern for involvement (1) or no involvement (0) in P, Cll, BS, S, BB, V, Cb, O functional systems in cited order, all cases with complete information on all eight FS.
† Rank order of expected frequency (E) for the specific pattern, with expectations based upon product of individual FS frequencies observed under the hypothesis of independence for involvement in each FS versus the others.
‡ Cumulative proportion (p) of total series for observed (O) and expected (E) patterns.

neurologic involvement in half of all the cases (table 4). One-third of all cases are defined by only six patterns, and one-fourth by four. Pyramidal + Cerebellar + Brainstem involvement, with or without Sensory *and* with or without a single minor FS, comprise eight of these 14 patterns. For only two of the 14 is either Cerebellar (ranks 7, 8) or Pyramidal (ranks 13, 14) involvement not present. In none of these 14 were there abnormalities for Other FS, attesting (as did its total frequency in table 2) to the appropriateness of this miscellaneous category. Only six of the 14 had any minor FS involved, each one single, and each one included: Visual (ranks 3, 4), Cerebral (9, 10), Bowel & Bladder (11, 12).

The Functional Systems scales. As stated, the scales were first devised back in 1953 with grades of 0 (normal) to 5 (maximal impairment) for each, except that the Other FS was a "yes-no" category (1,0) with the specific findings recorded but not rated. As mentioned above, there has been some tinkering with the steps and their definitions over the years, and the final set to be used should be the one provided in ref 4.

However, the basis for the scales remains as originally conceived, with increasing numbers given to increasing impairment as measured by both severity and extent of involvement for the deficits *as they occur* in this disease. For each FS, the observed deficits in the Bronx VA control series were detailed. For Brainstem, for example, nystagmus was by far the most common sign, and dysphagia occurred only when multiple and severe brainstem signs were otherwise present. Bowel & Bladder grades were based on observation of function in this hospitalized series. For dysfunction below the neck, the signs were allotted to the parts of the limbs or trunk affected, with measures of severity, and then consolidated by limb. In this manner, the FS grades were defined. As noted, I no longer have these data, but other material is illustrative of the process.

In the Army WW II series there were nearly 2,000 neurologic examinations among the 527 MS cases (476 definite, 51 probable) over the first 15 to 20 years of illness, some 40% of which were performed either for VA administrative purposes unrelated to exacerbations requiring hospitalization, or as part of our 1961-1963 follow-up effort.[10,15,16] In table 6 is the distribution of grades on the Pyramidal FS versus the pattern of limbs that were recorded at that examination as symptomatic for motor involvement (stiffness, weakness, etc). There were many more multiply symptomatic limbs for the higher grades of Pyramidal involvement than seen for the lower grades, but throughout there was an impressive tendency for (1) lower-extremity involvement and (2) symmetry of involvement. These trends increase regularly with increasing FS grades. The two-limbs-both pattern was found for ipsilateral as well as crossed involvement. Virtually all three-limb complaints included both lower extremities. This also holds true for coordination. Sensory FS did have some bilateral upper extremities in the three-limb group. This pattern for motor symptoms was quite similar to that from the Bronx VA series for motor *signs* by limb and severity, data which, as stated, no longer exist. As with any count of historical information, there is clearly some "noise" in the system (note grade 0 complaints), but the overall patterning seems reliable.

Surprisingly, the findings are very similar for Cere-

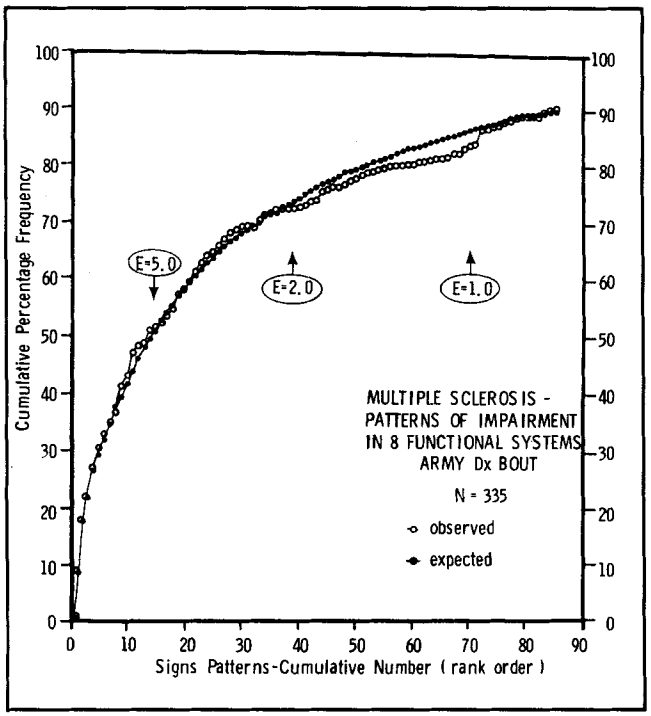

Figure 1. Cumulative percentage frequency for patterns of neurologic involvement within eight Functional Systems (FS) in an early bout of MS for the first 86 of the possible 256 patterns, with each specific pattern ranked according to its expectation under the hypothesis of independence of involvement in any FS versus the others. Observed (open circles) and expected (closed circles) frequencies are drawn for the same patterns at each rank. Expected numbers of 5.0, 2.0, and 1.0 cases for specific patterns are indicated on the graph, with patterns to the left having higher and patterns to the right lower expected frequencies than indicated. Data are based on admission diagnostic bout examinations in 335 definite MS with complete information on all eight FS, US Army WW II series. From Kurtzke 1972.[13]

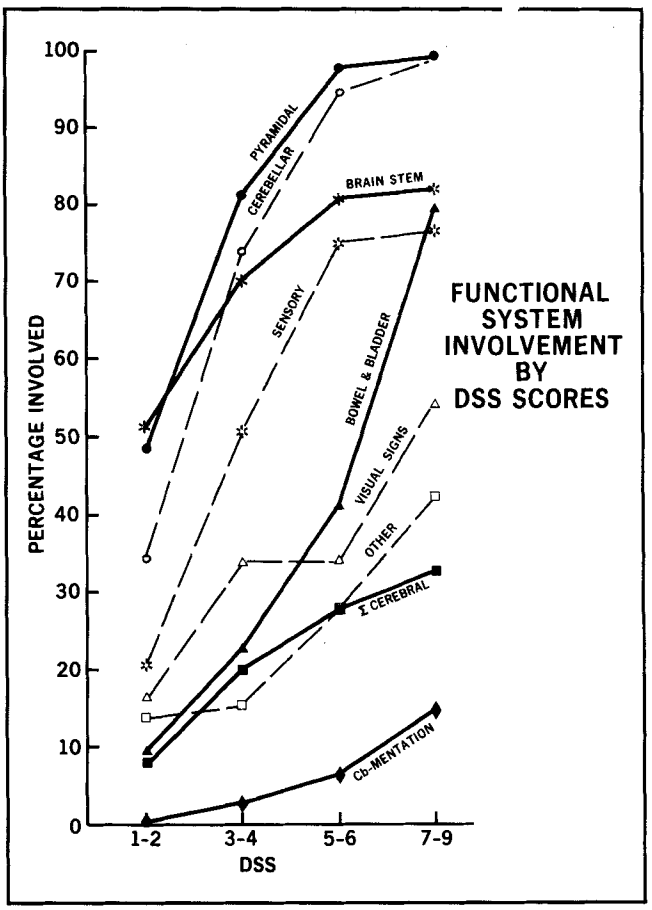

Figure 2. Percentage frequency involvement for each Functional System according to Disability Status Scale (DSS) score groups in some 2,000 neurologic examinations among 527 definite and probable MS during the first 15 to 20 years of illness, Army WW II series overview. From Kurtzke 1984.[18]

bellar FS grades versus coordination symptoms by limb (table 7). There is once again the marked preponderance of lower limb impairment and an even more pronounced tendency to symmetry. Again, these findings are more marked with higher FS grades.

The picture is quite different for Sensory signs versus symptoms (table 8). There is an almost equal frequency of involvement by limb, and there is a general trend to increasing symptoms for *all* the limb patterns with higher sensory FS grades, with very little lower limb or bilateral predilection.

These comparisons, though, do not provide the direct information that the neurologic examinations afforded in constructing the scales—all of which were much more strongly correlated with the scales than were the symptoms. For example, the frequency of motor or coordination *signs* by limb was, of course, 0 for FS grade 0. In his monograph on MS, Ragnar Müller[17] did provide a quite complete description of the neurologic involvement in nearly 600 Swedish patients with MS whom he personally reexamined. His data are contrasted with the motor, coordination, and sensory symptoms in table 9. There is a striking agreement between his motor signs and our symptoms as to involvement by limb. Unfortunately, neither his material nor any others of which I am aware permit a similar assessment for the Cerebellar system involvement. The Bronx VA data showed the pattern for signs recorded above for the Army symptoms for Pyramidal and Cerebellar, and the findings for Pyramidal were almost identical to Müller's.

Müller's data at sensory examination also are very similar to those from the Bronx VA material and are in striking contrast with the Army symptoms. Here too, just as with coordination and motor symptoms, we see the tendency to symmetry and the predilection for the lower extremities, more marked (as to frequency, at least) for the posterior column functions than for superficial sensation.

The FS and the DSS. Figure 2 shows the frequency of involvement in each FS for the Army WW II series overview according to DSS score groupings.[18] The DSS ranges from 0 (normal) to 10 (death due to MS), and, in my experience at least, tends to show a reasonably normal (Gaussian) distribution with maximal frequen-

Table 5. Patterns of involvement by Functional System (FS) in US Army WW II hospital series at diagnostic bout, with observed (O) and expected (E) numbers of cases for patterns 15 to 86*

Rank	Pattern	O	E	Rank	Pattern	O	E
15	1101 0100	2	4.964	51	1101 0001	2	1.380
16	1100 0100	2	4.935	52	1100 0001	1	1.372
17	1011 0100	3	4.648	53	1101 0110	0	1.335
18	1010 0100	5	4.621	54	1100 0110	2	1.327
19	1111 0001	7	4.258	55	1011 0001	1	1.292
20	1110 0001	3	4.233	56	1010 0001	1	1.285
21	1111 0110	6	4.119	57	1011 0110	1	1.250
22	1110 0110	6	4.095	58	1010 0110	0	1.243
23	1111 1100	5	3.758	59	1101 1100	1	1.218
24	1110 1100	4	3.736	60	1100 1100	0	1.211
25	0111 0100	1	3.274	61	1111 0011	1	1.145
26	0110 0100	5	3.255	62	1011 1100	1	1.141
27	1001 0000	5	2.844	63	1110 0011	0	1.138
28	1000 0000	3	2.828	64	1010 1100	1	1.134
29	1101 0010	2	2.521	65	0101 0100	0	1.061
30	1100 0010	1	2.505	66	0100 0100	1	1.055
31	1011 0010	0	2.360	67	1111 1001	2	1.045
32	1010 0010	1	2.346	68	1110 1001	0	1.039
33	1101 1000	3	2.299	69	1111 1110	4	1.011
34	1100 1000	2	2.286	70	1110 1110	1	1.005
35	1111 0101	2	2.256	71	0011 0100	1	0.994
36	1110 0101	2	2.242	72	0010 0100	9	0.988
37	1011 1000	0	2.153	73	0111 0001	0	0.910
38	1010 1000	0	2.140	74	0110 0001	1	0.905
39	0101 0000	0	2.004	75	0111 0110	0	0.880
40	0100 0000	1	1.992	76	0110 0110	3	0.875
41	1111 1010	2	1.908	77	0111 1100	1	0.803
42	1110 1010	2	1.897	78	0110 1100	1	0.799
43	0011 0000	1	1.876	79	1001 0010	2	0.765
44	0010 0000	6	1.865	80	1000 0010	0	0.760
45	0111 0010	1	1.662	81	1101 0101	0	0.731
46	0110 0010	1	1.653	82	1100 0101	0	0.727
47	0111 1000	0	1.517	83	1001 1000	2	0.698
48	0110 1000	1	1.508	84	1000 1000	3	0.694
49	1001 0100	3	1.507	85	1011 0101	0	0.684
50	1000 0100	1	1.498	86	1010 0101	0	0.680
					Ranks 1-86 total:	303	302.541

* Footnotes as in table 4.

cies in the mid-range and low tails at either end.[1-4,7,14-16,18]

The impressive point here is that there is a clearly increasing frequency of involvement in every one of the eight FS with increasing DSS scores, even for functions that obviously have nothing to do with ambulation.

Not only does the frequency increase with higher DSS scores, this is also true of the severity, as is readily seen for both Pyramidal and Cerebellar FS (figure 3). There is a regular progression of a shift to the right in both these curves, while the overall distribution is either flat or irregular below the highest grades. Data underlying these figures are published.[15,16] Brainstem and Sensory findings are similar, particularly for the former (figure 4). Shifts in Sensory FS are modest at the higher DSS, and severe sensory loss is really not a very common feature in this disease. The trend, though, is still apparent. Bowel & Bladder as well as Visual FS do tend to worsen with increasing DSS, but this is obscured by the lower frequency of any involvement within these minor FS (figure 5). Cerebral and the "yes-no" Other FS behave the same, but, here also, frequencies involved overall also demonstrate their appropriate listing as "minor" (figure 6).

The FS and pathophysiologic inferences. To my mind, the DSS is a good reflection of the FS, and the FS accurately reflect the neurologic examination. For pragmatic use as a tool in measuring the course of illness, whether in therapeutic trials or otherwise, this might be considered sufficient. But I think we can make inferences as to the anatomic distribution of lesions from the FS findings.

The first conclusion is that MS is, first and foremost, a disease of the white matter of the spinal cord and brainstem. This is indicated by the high frequency of the long tract findings (Pyramidal, Sensory). Secondly,

Table 6. Pyramidal FS versus motor symptoms by limb, US Army WW II series overview. Data of Kurtzke 1970.[15,16]

Motor symptoms by limb	Pyramidal FS grades 0	1	2	3	4	5 + 6	Total
None	76.2*	73.2	19.3	10.0	3.6	0.7	31.8
1-UE	3.6	0.9	12.4	2.7	1.2	0	3.5
1-LE	5.1	5.2	18.5	9.6	2.4	0	7.3
2-UE	0.3	0.4	1.7	2.1	0.4	0	1.0
2-LE	8.4	12.6	29.2	34.0	40.7	18.5	24.7
2-both	2.7	3.9	11.2	24.9	10.3	1.3	11.1
3(2 LE)	2.1	1.3	2.6	5.7	11.5	4.6	4.7
4	1.5	2.6	5.2	11.0	30.0	74.8	15.9
Total	99.9	100.0	100.1	100.0	100.1	99.9	100.0
(N)	(332)	(231)	(233)	(438)	(253)	(151)	(1,638)
[%]	[20.3]	[14.1]	[14.2]	[26.7]	[15.4]	[9.2]	[99.9]

UE/LE Upper/lower extremity (limb).
* Percentages.

Table 7. Cerebellar FS versus coordination symptoms by limb, US Army WW II series overview. Data of Kurtzke 1970.[15,16]

Coordination symptoms by limb	Cerebellar FS grades 0	1	2	3	4 + 5	Total
None	76.6*	69.1	33.5	14.1	7.1	41.1
1-UE	1.7	0.9	4.2	2.2	0	2.4
1-LE	1.2	0.9	4.0	1.2	2.0	2.1
2-UE	1.2	0.9	5.6	7.7	8.2	4.8
2-LE	15.5	20.9	37.7	40.7	14.3	29.5
2-both	1.0	0.9	3.3	1.5	3.1	1.9
3(2 LE)	0.7	1.8	4.2	4.9	2.0	3.1
4	2.0	4.5	7.5	27.7	63.3	15.1
Total	99.9	99.9	100.0	100.0	100.0	100.0
(N)	(406)	(110)	(427)	(405)	(98)	(1,446)
[%]	[28.1]	[7.6]	[29.5]	[28.0]	[6.8]	[100.0]

UE/LE Upper/lower extremity (limb).
* Percentages.

Table 8. Sensory FS versus sensory symptoms by limb, US Army WW II series overview. Data of Kurtzke 1970.[15,16]

Sensory symptoms by limb	Sensory FS grades 0	1	2	3-5*	Total
None	75.0†	64.5	42.2	37.6	58.9
1-UE	2.8	5.4	8.4	8.1	5.4
1-LE	3.3	5.4	5.7	8.1	4.9
2-UE	3.3	3.2	6.9	5.9	4.8
2-LE	5.7	9.7	13.6	10.0	8.9
2-both	3.6	2.2	9.1	15.4	7.0
3	1.5	2.2	5.0	5.0	3.1
4	4.8	7.5	9.1	10.0	7.0
Total	100.0	100.1	100.1	100.1	100.0
(N)	(689)	(93)	(419)	(221)	(1,422)
[%]	[48.5]	[6.5]	[29.5]	[15.5]	[100.0]

* Only 11 patients grade 4 or 5.
UE/LE Upper/lower extremity (limb).
† Percentages.

the lesions of cord and stem that are responsible are multiple and bilateral. This is indicated by the independence for FS patterns and the distributions of limb involvement. Bilaterality is obvious, but multiplicity must also be necessary because of the striking preponderance of lower limb involvement over the uppers, compatible with the thesis that the longer is a given neural pathway, the more likely is it to be attacked, and multiply so. And since there is a much greater extent of these tracts in the cord than in the brainstem, the lesions will probably be more numerous in the former. The ratios of the upper: lower involvement further support this. Also, because of the multiplicity of lesions required for impairment implicit in the bilaterality, the lower limb predilections, and the independence of FS patterns, we would not expect tissue-destructive lesions in this disease; there should not be wallerian degeneration.

What is unexpected in this context is that the Cerebellar FS findings *also* reflect predominantly spinal cord disease. We are taught that, aside from a system degeneration like Friedreich's ataxia, we do not see cerebellar ataxia from spinal cord lesions, the usual explanation given being the bilateral nature of at least the ventral spinocerebellar tracts. To involve both, one lesion would also involve both pyramidal tracts, and a paralyzed patient cannot be ataxic. Here the limb involvement shows the same preponderance of lower over upper, and an even greater tendency to bilaterality than with the other long tracts. Bowel & Bladder involvement is also evidence of spinal cord disease. The Sensory involvement favors the large posterior columns in like manner. The sensory *symptoms*, though, equally distributed by body-part, might indicate the dorsal root

Table 9. Symptoms and signs in MS by type of impairment and limbs involved: A. symptoms by limb in US Army WW II series overview; B. signs by limb in prevalent MS cases of Sweden (Müller 1949).[17] Data of Kurtzke 1970.[15,16]

Limbs involved	Type of impairment						
	Motor		Coordination	Sensory	Sensory signs (B)		
	Symptoms (A)	Signs (B)	symptoms (A)	symptoms (A)	Touch-pain	Position	Vibration
None	31.8†	36.1	41.1	58.9	62.5	49.9	36.8
1-UE	3.5	1.2	2.4	5.4	?*	0	0
1-LE	7.3	5.2	2.1	4.9	?	0	0
2-UE	1.0	0.2	4.8	4.8	2.0	3.0	4.1
2-LE	24.7	30.2	29.5	8.9	24.9	30.9	37.4
2-both	11.1	3.6	1.9	7.0	?	0	0
3(2 LE)	4.7	11.0	3.1	3.1	?	0	0
4	15.9	12.5	15.1	7.0	3.5	16.3	21.8
Total	100.0	100.0	100.0	100.0	100.0	100.1	100.1
(N)	(1,638)	(582)	(1,446)	(1,422)	(541)	(541)	(541)

* Incomplete definition for the other 7.1 percent.
UE/LE Upper/lower extremity (limb).
† Percentages.

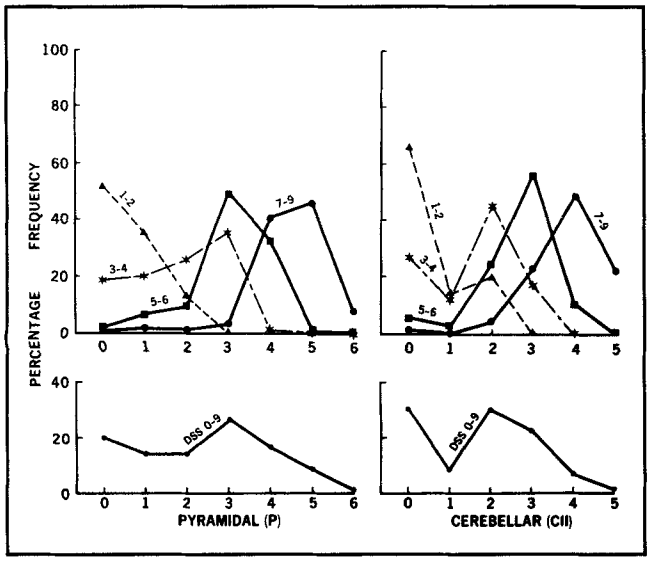

Figure 3. Percentage frequency involvement by grade for Pyramidal and Cerebellar FS according to DSS score groups from the Army WW II series overview. The lower graphs show FS distributions for the entire series.

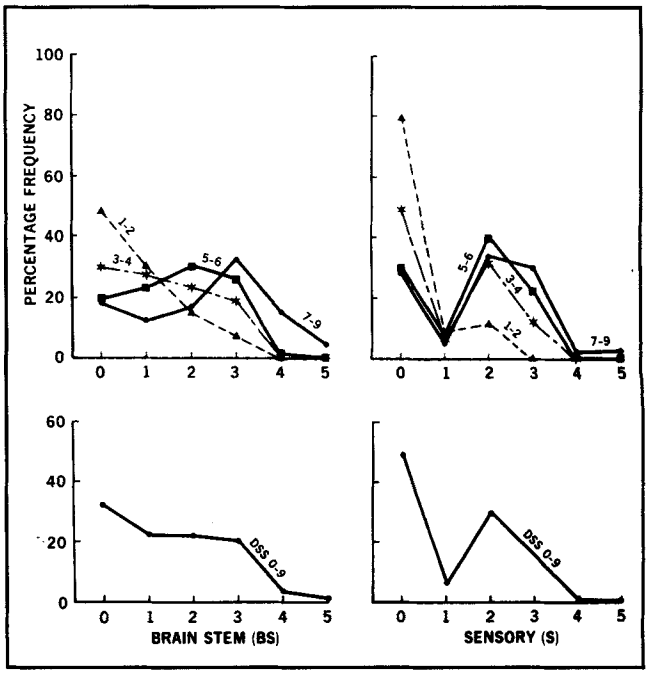

Figure 4. Percentage frequency involvement by grade for Brainstem and Sensory FS according to DSS score groups, as in figure 3.

entry zone as the likely locus for the undefinable "numbness" that is so common in this disease.

Thus, the greatest part of the clinical involvement in MS reflects multiplicity of lesions in the spinal cord, with Pyramidal, Cerebellar, Sensory, and Bowel & Bladder systems affected. Only the Brainstem signs are left as a major FS, and this then reflects the second most common site of lesions inferred in this disorder.

Therefore, we expect that the predominant neuropathology will be that of multiple axon-sparing lesions of the spinal cord and brainstem. Thus, MS is not only a disease of the *myelin*, it is above all a disease of the *myelon* (μυελός, μυελόν), the (white matter of the) neural tube. That this is indeed the case was illustrated years ago by Charcot.[19] Figures 7 and 8 show the multiple and irregular plaques in the spinal cord, some occupying the entire cross section of the cord, and clearly demonstrating the absence of wallerian degeneration. A large proportion of spinal cord white matter is affected. In contrast, the proportion of cerebral white with grossly visible lesions is much less (figure 9), and the major ones are concentrated about the lateral ventricles.

MS pathologically is almost the reverse of CNS vascular disease: the cord is universally affected, and most cerebral lesions are around the ventricles. In all instances the MS lesions would seem to reflect the occurrence, spread, and confluence of perivascular (perivenular) lesions.

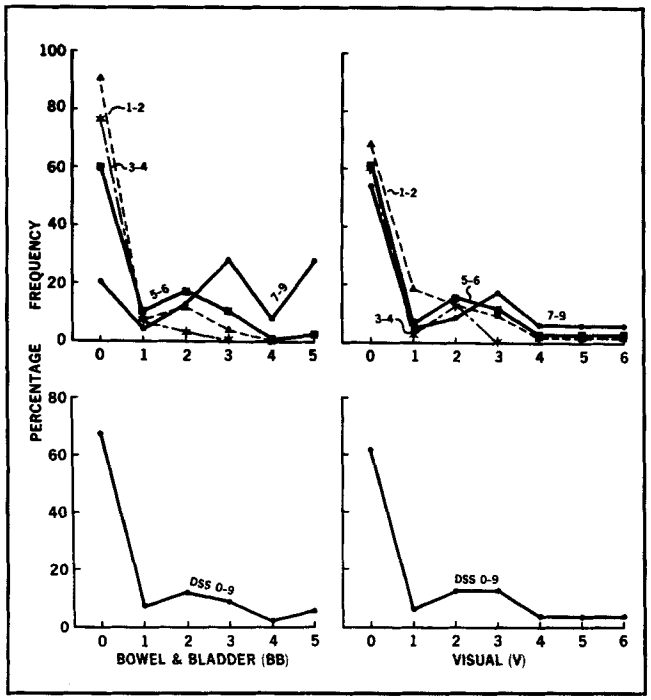

Figure 5. Percentage frequency involvement by grade for Bowel & Bladder and Visual FS according to DSS score groups, as in figure 3. Visual functions are those for 392 VA Hospital patients.

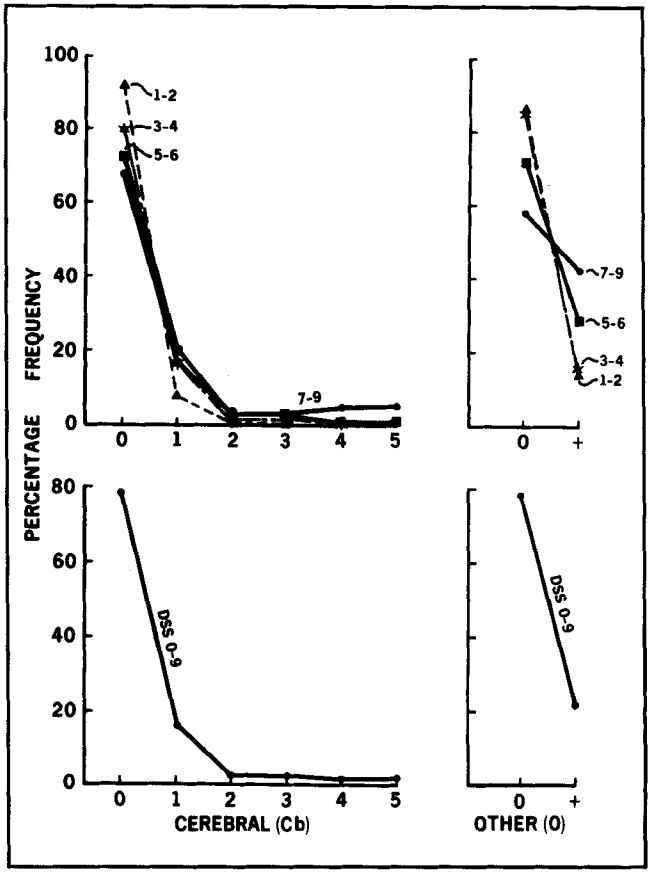

Figure 6. Percentage frequency by grade for Cerebral FS and frequency for Other FS according to DSS score groups, as in figure 3.

Therefore, the FS do reflect the qualitative and quantitative aspects of the pathophysiology of MS, and the DSS reflects an overview of the entire patient insofar as this disease affects his neurologic behavior as summarized by the FS. It should be clear, also, why the FS per se can properly be subordinated in the more severely affected patient by measures of the ability to walk, and, in the most severe, by the bedridden state.

The EDSS was devised in response to objections that the DSS did not have enough steps—it was not "sensitive" enough to measure change in treatment trials. Based on the data cited here, the only feasible method to enlarge the scale was to split each step (1 to 9) into two (0.5 to 9.5), since in our material no one DSS step was discrepant enough to believe it should really be two or more. In order to accomplish this expansion, very arbitrary definitions of both FS scores and ambulation abilities were required.

The (E)DSS + FS—Critiques. This bifid system of DSS + FS has, in my view, proved to be both useful and reliable, the latter attested to by Kuzma et al[20] for DSS + FS, and by Amato et al[21] for EDSS + FS. Amato et al[22] were much less impressed with the reliability, however, when different examiners independently examined the patients. This, though, is precisely what Kuzma et al[20] had done. In the Italian work, basic data to assess the conclusions were not provided, but it seems that the authors had found that the interobserver differences lay more in the neurologic examinations per se. It may be then that "The fault, dear Brutus, is not in our scales, But in ourselves," to misquote Cassius (*Julius Caesar*, I.ii, 146-7). No scale can be more accurate than its basic data. I have previously implied that one should require a full step change on the EDSS, not the 0.5 step that is possible, before inferring a change for better or worse in the patient.[4]

A more extensive critique has been that of Willoughby and Paty.[23] They conclude that "the EDSS has important flaws . . . the title is inappropriate. More substantial problems include inadequate precision in defining the degree of impairment in some functional categories . . . and the use of a mixture of neurologic signs elicited on examination and subjective information obtained from the patient in defining the overall scale."

Willoughby and Paty provide an excellent recapitulation of the varied attempts to assess neurologic impairment in MS. I agree that they are etymologically correct as to the (E)DSS, but the name came before the WHO categories they describe, and, personally, I think that observation is really not important.

The objection that neurologic involvement does not measure the pathologic extent of MS is, to me, in large measure incorrect. "Relatively small lesions in the spinal cord or brainstem may produce marked impairment while more extensive pathologic changes in the cerebral hemispheres may have few neurologic signs"[23] is a totally appropriate statement when we are looking at tumors or infarcts or hemorrhages. But, as detailed

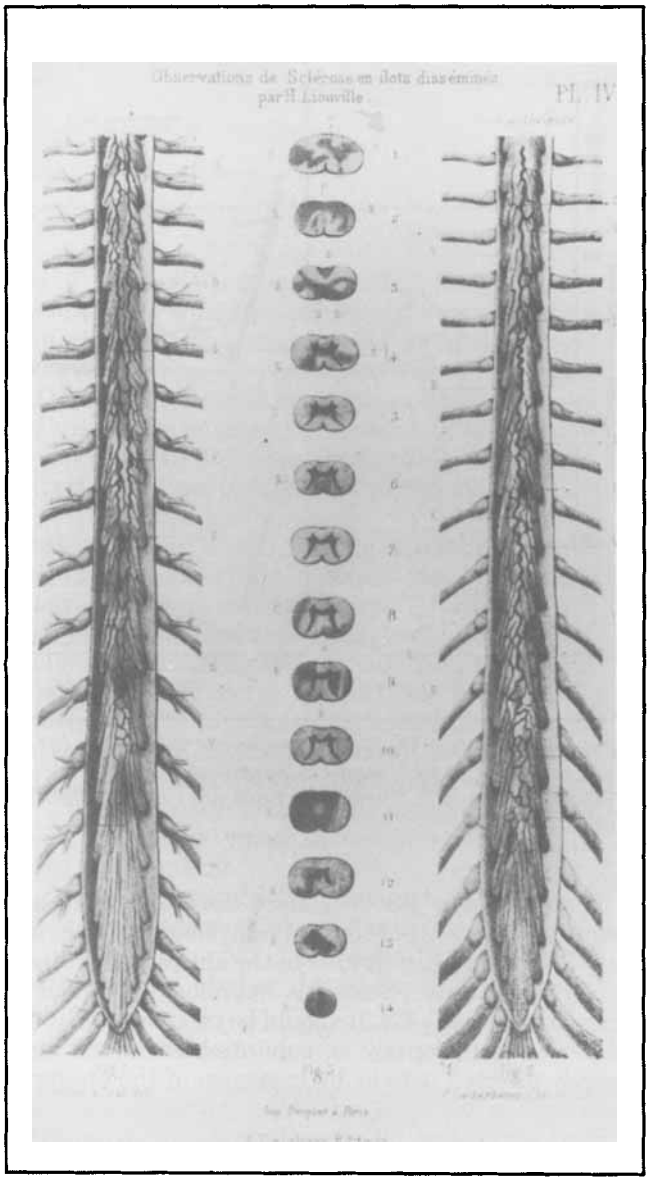

Figure 7. Plate III. MS spinal cord viewed from posterior (Fig. 1) and anterior (Fig. 2) aspects, with cross sections at the levels specified (Fig. 3). From Charcot 1877.[19]

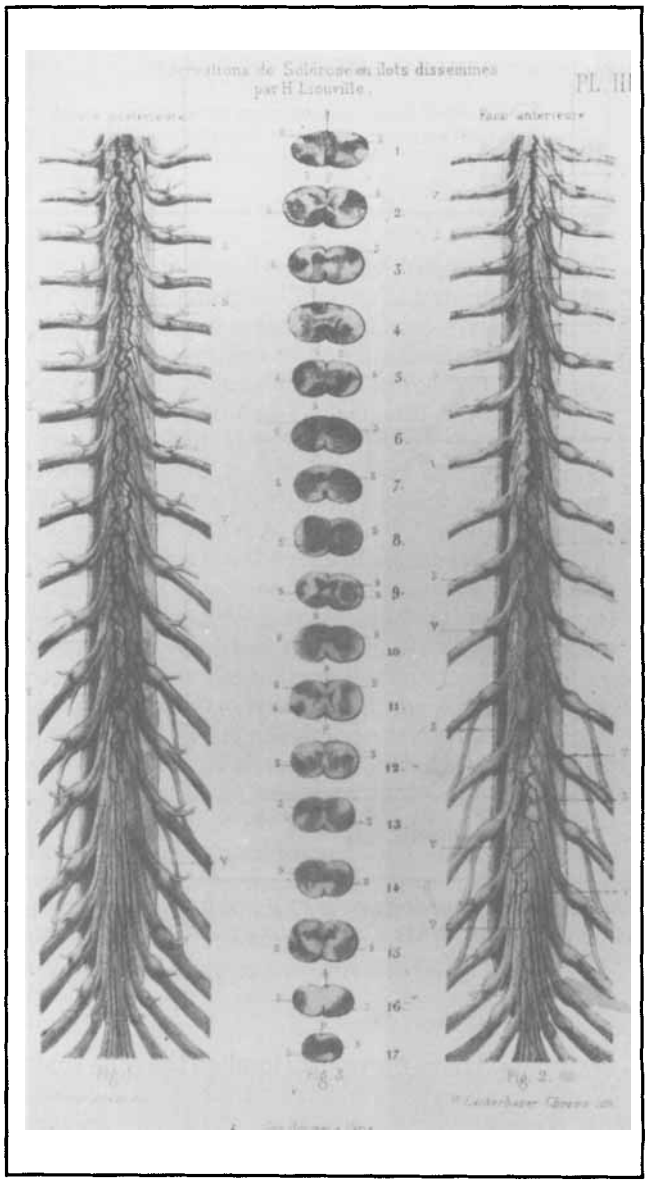

Figure 8. Plate IV. MS spinal cord as in figure 7. From Charcot 1877.[19]

above, it does *not* appear to me to reflect the pathophysiology of MS. To oversimplify, any MS dysfunction is likely to be the result of multiple lesions within a given neural pathway, and the longer the pathway, the more frequent (and the more severe) will be the involvement. Thus my definition of MS as a myelon disease. (What the UBOs [unidentified bright objects] in MRI denote, is, as I have said,[24] anyone's guess.)

Their reference that the Minimal Record of Disability published by the National MS Society[25] added a separate scale for spasticity is correct, as is their observation that it "further confuses the original concept." It was done over my objections; spasticity has always been included as part of Pyramidal Function.

I will refer to their "Problems with the DSS" in the order cited. (1) The name and "a component related to the patient's ability to work and carry out activities of daily living." To the latter, these were *explanations* and not definitions for certain steps. The problem here may be that the DSS appeared before publication and explanation of the FS, from which it was derived.

(2) The FS *are* mutually exclusive; Pyramidal, Cerebellar, Sensory *always* referred to phenomena below the neck. Again, the infarct model is raised here, and, for MS, I believe it is incorrect.

(3) Originally *and* to date, the FS (and DSS) *are* limited to "objectively verifiable defects." As noted, the FS were first defined from hospitalized patients. I am sure that Willoughby and Paty will agree that, in that setting, incontinence is "objectively verifiable." Similarly, gait performance was part of the examination in the Cornell axis. For convenience, we *do* often accept statements as to sphincters and gait, but they *are* provable, and this is what the scales were based on. Especially if one wishes to use the EDSS versus the DSS, then I think it is his responsibility to observe whether the patient *can* walk 100 or 200 or 300 meters, or whatever.

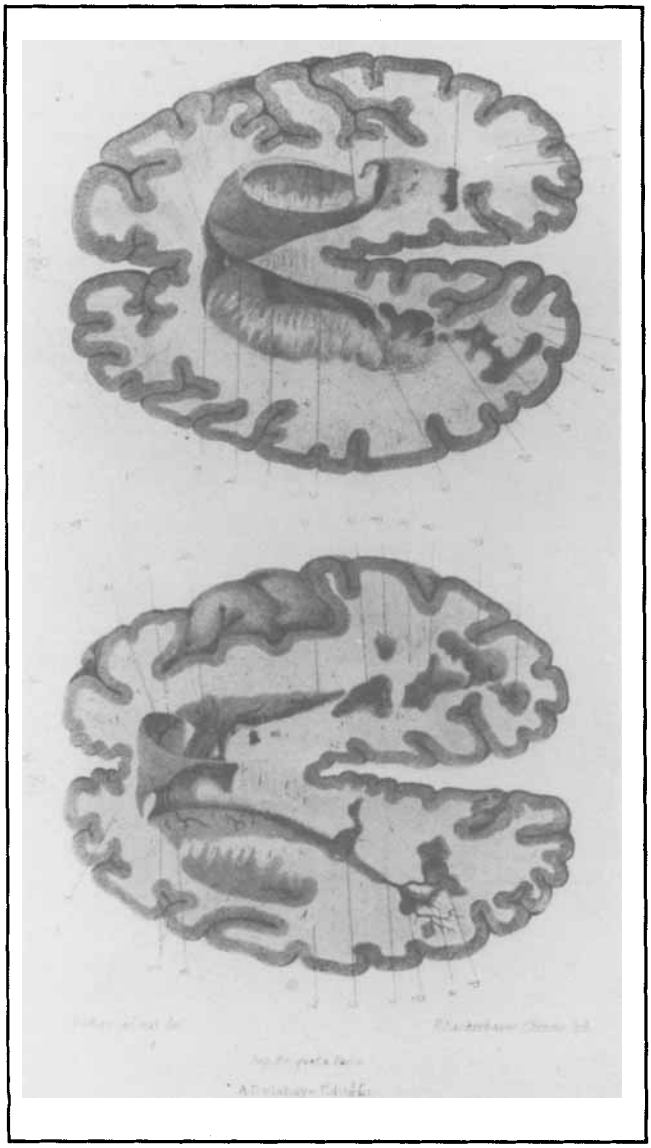

Figure 9. Plate II. MS brain. Horizontal cross section of brain showing the islets of sclerosis in different regions, white and gray (Fig. 1); another section of the same brain (Fig. 2). From Charcot 1877.[19]

(4) Assignment of grades can be difficult. It is much less so if one follows the prescriptions given from the beginning of the scales, ie, to "bracket" the relevant grade. A patient will clearly be, let us say, neither Cerebellar 0 or 1 nor 4 or 5. Thus, he is either 2 or 3. To which one is he closer? That is his grade. I really don't think Brainstem offers that much difficulty either.

(5) The emphasis on ambulation in the (E)DSS is, as detailed above, based upon the neurologic impairments seen in this disease. A patient with *only* loss of vision in one eye, as they cite, does not *have* much of his neuraxis clinically affected.

(6) The Gaussian distribution of the DSS in the series I reviewed has not been found in their experience. I think this is interesting, but not really vital to the argument. It would be of interest, perhaps, were I to recode their series to see if we are really differing in interpretations or whether their cases do differ as to the findings. I find it difficult to believe that all those series that we coded—including the nearly 2,000 exams done by multiple examiners in multiple settings in the Army series—are really artifactual. No patient was ever coded by us with the view that we needed more or fewer subjects with DSS 4 or 5 or whatever. One might observe that the *use* of a cane is not necessarily equated with the *need* for same. If I scored their patients as they did, my action would perhaps then be to redefine steps 4 to 6 on the (E)DSS, but not to throw the baby out with the bath water.

As to the recommendations of Willoughby and Paty,[23] to each of their points I will respond here as to whether the (E)DSS + FS does, in my opinion, meet their goals (criteria) or not: (1) subscales on standard exam: yes; (2) clear description of grades: yes; (3) selection on ease and reliability: yes; separate limbs: no—but see prior discussions above; (4) bigger score, higher impairment: yes; (5) combination of subscales: yes; addition of subscales: no—see above.

In summary, I do wish Willoughby and Paty well in defining a better scale. Personally, I would change my system only, perhaps, by returning to the DSS vice the EDSS. We have more steps, but I really wonder if the gain is worth the possibly overly rigid definitions entailed. However, if it is performed properly, the EDSS is readily collapsed into the DSS. Further, I wonder if perhaps some of the objections to the (E)DSS by workers other than our Canadian brethren are really a "slay the messenger" response when, based upon such scales, a treatment trial is concluded to be negative or a laboratory finding unrelated to severity or activity in this disease. It may well be that the (E)DSS + FS rating scheme could be likened to democracy, which has been called the worst form of government—except for all others.

References

1. Kurtzke JF. A new scale for evaluating disability in multiple sclerosis. Neurology 1955;5:580-583.
2. Kurtzke JF. On the evaluation of disability in multiple sclerosis. Neurology 1961;11:686-694.
3. Kurtzke JF. Further notes on disability evaluation in multiple sclerosis, with scale modifications. Neurology 1965;15:654-661.
4. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444-1452.
5. Kurtzke JF, Berlin L. The effects of isoniazid on patients with multiple sclerosis: preliminary report. Am Rev Tuberc 1954;70:577-592.
6. Alexander L. New concept of critical steps in course of chronic debilitating neurologic disease in evaluation of therapeutic response: longitudinal study of multiple sclerosis by quantitative evaluation of neurologic involvement and disability. Arch Neurol Psychiat 1951;66:253-271.
7. Kurtzke JF. Course of exacerbations of multiple sclerosis in hospitalized patients. Arch Neurol Psychiat 1956;76:175-184.
8. Veterans Administration Multiple Sclerosis Study Group. Isoniazid in treatment of multiple sclerosis: report on Veterans Administration cooperative study. JAMA 1957;163:168-172.
9. Kurtzke JF, Berlin L. Isoniazid in treatment of multiple sclerosis. JAMA 1957;163:172-174.
10. Nagler B, Beebe GW, Kurtzke JF, Kurland LT, Auth TL, Nefzger MD. Studies on the natural history of multiple sclerosis. 1. Design

and diagnosis. Acta Neurol Scand 1966;42(suppl 19):141-156.
11. Kurland LT, Beebe GW, Kurtzke JF, et al. Studies on the natural history of multiple sclerosis. 2. The progression of optic neuritis to multiple sclerosis. Acta Neurol Scand 1966;42(suppl 19):157-176.
12. Schumacher GA, Beebe G, Kibler RF, et al. Problems of experimental trials of therapy in multiple sclerosis: report by the Panel on the Evaluation of Experimental Trials of Therapy in Multiple Sclerosis. Ann NY Acad Sci 1965;122:552-568.
13. Kurtzke JF, Beebe GW, Nagler B, Auth TL, Kurland LT, Nefzger MD. Studies on the natural history of multiple sclerosis. 6. Clinical and laboratory findings at first diagnosis. Acta Neurol Scand 1972;48:19-46.
14. Kurtzke JF. Migration and latency in multiple sclerosis. In: Field EJ, Bell TM, Carnegie PR, eds. Multiple sclerosis: progress in research. Amsterdam: North-Holland, 1972:208-228.
15. Kurtzke JF. Clinical manifestations of multiple sclerosis. In: Vinken PJ, Bruyn GW, eds. Handbook of clinical neurology, vol 9. Multiple sclerosis and other demyelinating diseases. Amsterdam: North-Holland, 1970:161-216.
16. Kurtzke JF. Neurologic impairment in multiple sclerosis and the Disability Status Scale. Acta Neurol Scand 1970;46:493-512.
17. Müller R. Studies on disseminated sclerosis with special reference to symptomatology, course and prognosis. Acta Med Scand 1949;133(suppl 222):1-214.
18. Kurtzke JF. Disability rating scales in multiple sclerosis. Ann NY Acad Sci 1984;436:347-360.
19. Charcot J-M. Leçons sur les Maladies du Système Nerveux Faites a la Salpêtrière. Recueillies et Publiées par Bourneville. Tome Premier. Troisième édition. Paris: V. Adrien Delahaye et Cie, 1877.
20. Kuzma JW, Namerow NS, Tourtellotte WW, et al. An assessment of the reliability of three methods used in evaluating the status of multiple sclerosis patients. J Chronic Dis 1969;21:803-814.
21. Amato MP, Groppi C, Siracusa GF, et al. Inter- and intra-observer reliability in the employment of Kurtzke scoring system in multiple sclerosis. Ital J Neurol Sci 1987;3(suppl 6):129-131.
22. Amato MP, Fratiglioni L, Groppi C, Siracusa G, Amaducci L. Interrater reliability in assessing functional systems and disability on the Kurtzke scale in multiple sclerosis. Arch Neurol 1988;45:746-748.
23. Willoughby EW, Paty DW. Scales for rating impairment in multiple sclerosis: a critique. Neurology 1988;38:1793-1798.
24. Kurtzke JF. Multiple sclerosis: what's in a name? Neurology 1988;38:309-316.
25. National Multiple Sclerosis Society. Minimal record of disability for multiple sclerosis. New York: NMSS, 1985.

The Disability Status Scale for multiple sclerosis: Apologia pro DSS sua

John F. Kurtzke

Neurology 1989;39;291
DOI 10.1212/WNL.39.2.291