

Responsiveness of the Scripps Neurologic Rating Scale During a Multiple Sclerosis Clinical Trial

James A. Koziol, Adriana Lucero, Jack C. Sipe, John S. Romine, Ernest Beutler

ABSTRACT: *Objective:* The Scripps neurologic rating scale (SNRS) is a summary measure of individual components comprising a neurological examination, designed for use in multiple sclerosis (MS). Our objective is to evaluate the responsiveness of the SNRS, within the context of a 2-year, randomized, double-blind crossover study of the efficacy of cladribine for treatment of secondary progressive MS. *Methods:* Effect sizes were determined for the SNRS and its components, separately for each treatment group (initial placebo, and initial cladribine) over both years of the clinical trial, using a standard random effects model. *Results:* Individual components tended to show positive effect sizes (improvement) during periods of active therapy in both treatment groups, and negative effect sizes (deterioration) during periods of no active therapy. Summation indices derived from the individual components of the SNRS seemed somewhat more stable than the individual components. The two components mentation and mood, and bladder, bowel, or sexual dysfunction, were rather unresponsive in our clinical trial. *Conclusions:* Changes in the components of the SNRS over the course of our clinical trial were consistent between the two treatment groups. Most components were moderately responsive; and, the summary SNRS score appropriately summarized the moderate magnitudes of change evinced in the individual components.

RÉSUMÉ: *Sensibilité de l'échelle neurologique de Scripps lors d'un essai thérapeutique dans la sclérose en plaques.* *Objectif:* L'échelle neurologique de Scripps (ÉNS), conçue pour être utilisée dans la sclérose en plaques (SEP), est une mesure sommaire de composantes individuelles incluant un examen neurologique. L'objectif de l'étude était d'évaluer la sensibilité de l'ÉNS dans le contexte d'une étude randomisée, en double insu avec chassé-croisé, de l'efficacité de la cladribine dans le traitement de la SEP secondaire progressive. *Méthodes:* L'ampleur des effets a été déterminée pour l'ÉNS et ses composantes séparément pour chaque groupe de traitement (traitement initial: placebo ou cladribine) au cours des deux ans de l'étude clinique au moyen d'un modèle standard à effets aléatoires. *Résultats:* Les composantes individuelles tendaient à démontrer des effets positifs (amélioration) pendant les périodes de traitement actif dans les deux groupes de traitement et des effets négatifs (détérioration) pendant les périodes sans traitement actif. Les indices de sommation dérivés des composantes individuelles de l'ÉNS semblaient plus stables que les composantes individuelles. Les deux composantes état mental et humeur, et dysfonction vésicale, intestinale ou sexuelle n'étaient pas sensibles lors de cet essai thérapeutique. *Conclusions:* Les changements dans les composantes de l'ÉNS au cours de cet essai thérapeutique étaient concordants entre les deux groupes de traitement. La plupart des composantes étaient modérément sensibles et le score sommaire résumait de façon appropriée l'ampleur modérée des changements démontrés par les composantes individuelles.

Can. J. Neurol. Sci. 1999; 26: 283-289

We have recently undertaken a two-year, randomized, placebo-controlled, double-blind crossover study examining the efficacy of cladribine in the treatment of secondary progressive multiple sclerosis.^{1,2} In the first year of this trial, patients were given either four monthly courses of cladribine or placebo. There was an eight-month washout period; then, in the second year patients who had been treated previously with cladribine were crossed over to placebo and conversely. Two neurologists (JR and JS) with no knowledge of which medication each patient was receiving, examined the patients monthly and provided scores for two neurological instruments: the Kurtzke expanded disability status scale (EDSS) and the Scripps neurologic rating scale (SNRS). Each patient was examined by the same neurologist over the course of the study. All neurological examinations in this clinical trial were conducted in the Scripps General Clinical

Research Center (GCRC); in addition, the neurologists were blinded to prior scores until after their current examinations were completed. We reported^{1,2} that cladribine therapy was associated with positive outcomes, namely, the elimination of lesions on T1-weighted MRI scans under gadolinium enhancement and a favorable influence on the monthly SNRS and EDSS scores in patients treated with cladribine.

A prerequisite for adoption of neurologic impairment scales as outcome measures in clinical studies is evaluation of their

From the Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California (JAK, EB), and the Division of Neurology, Scripps Clinic, La Jolla, California (AL, JCS, JSR)

RECEIVED JULY 28, 1998. ACCEPTED IN FINAL FORM JULY 13, 1999

Reprint requests to: James A. Koziol, Department of Molecular and Experimental Medicine, 311N, The Scripps Research Institute, 15550 North Torrey Pines Road, La Jolla, CA 90237

scientific properties of reliability, validity, and responsiveness.³ We have previously reported on certain psychometric properties of the SNRS instrument, in particular construct validity^{4,5} and reliability.¹ We now address the issue of responsiveness of the SNRS, that is, the ability of the instrument to detect significant changes in health status over time, within the context of our clinical trial.

It is common to use the EDSS, the SNRS, or some other evaluative instrument as an outcome measure in MS therapeutic trials, both to estimate a patient's degree of disability or impairment at any point during the trial, and to assess changes that might be attributable to treatment effects.⁶ Whereas the reliability and validity of an instrument bear upon its usefulness for descriptive or predictive purposes, its responsiveness is the essential attribute necessary for its adoption as an appropriate outcome measure in a clinical trial.⁷ Our investigation of the responsiveness of the SNRS and its components reveals that most of the components, as well as the summary SNRS, change appreciably and consistently in the time course of our clinical trial, hence are useful as outcome measures.

METHODS

We have previously reported results from a two-year, placebo-controlled, randomized double-blind crossover study to evaluate cladribine in the treatment of secondary progressive multiple sclerosis.^{1,2} In the first year, patients were given intravenous (i.v.) infusions of cladribine on a schedule of 0.10 mg/kg/day for seven days as four monthly courses for a total of 2.8 mg/kg, or placebo. During the second year, patients treated with placebo during the first year were given i.v. infusions of 0.10 mg, 0.05 mg, and 0.05 mg of cladribine per kilogram of body weight per day for seven consecutive days in three consecutive monthly courses, for a total dose of 1.4 mg/kg; patients who had been previously treated with cladribine were crossed over to placebo.

Magnetic resonance imaging (MRI) was performed at six month intervals on all patients, using a 1.5 T General Electric Signa scanner. T2 and proton density weighted images were obtained using a conventional spin-echo sequence with repetition times of 2500 ms and echo delay times of 30 and 90 ms. Sections were 4 mm thick with a 1 mm interslice gap. Special attention was given to careful repositioning of patients to ensure reproducible slice sections. Regions of demyelination on T2-weighted scans were outlined on filmed images; a volume rendering software program (CEMAX, Santa Clara, California) was used to convert pixel counts for each slice into volumes. All scans were interpreted and marked by one neuroradiologist who had no knowledge of treatment assignment; and, all volumetric analyses were done by the same technologist. T1-weighted scans were also recorded, but are not considered here.

Patients were evaluated monthly by two neurologists (JS and JR), with the Kurtzke expanded disability status scale (EDSS) and the Scripps neurologic scoring scale (SNRS). We have

Scripps Neurological Rating Scale				
System Examined	Degree of Impairment			
	Normal	Mild	Moderate	Severe
Mentation and Mood	10	7	4	0
Cranial Nerves				
Visual Acuity	5	3	1	0
Fields, Discs, Pupils	6	4	2	0
Eye Movements	5	3	1	0
Nystagmus	5	3	1	0
Lower Cranial Nerves	5	3	1	0
Motor				
RU	5	3	1	0
LU	5	3	1	0
RL5	3	1	0	
LL	5	3	1	0
DTRS				
UE	4	3	1	0
LE	4	3	1	0
Babinski				
R	2	0	0	0
L	2	0	0	0
Sensory				
RU	3	2	1	0
LU	3	2	1	0
RL	3	2	1	0
LL	3	2	1	0
Cerebellar				
UE	5	3	1	0
LE	5	3	1	0
Gait; Trunk and Balance	10	7	4	0
Special Category				
Bladder/Bowel/ Sexual Dysfunction	0	-3	-7	-10
<hr/> OVERALL SNRS SCORE (Maximum = 100) <hr/>				

Figure 1: Worksheet for the Scripps Neurological Rating Scale. Assessments are made on each item during the neurological examination. Points assigned for each component are subtotaled, and points for autonomic dysfunction are subtracted, to arrive at the overall SNRS score.

discussed reliability of the EDSS and SNRS in Sipe et al.¹ and Beutler et al.² and have commented on issues relating to validity of the scales in Sipe et al.⁴ and Koziol et al.⁵

The SNRS was specifically devised for use in multiple sclerosis, and is based on assessment of each component of a standard neurological examination protocol.⁴ The worksheet associated with the SNRS is shown in Figure 1. The assignment

Table 1: Effect sizes of individual components of the SNRS for the initial cladribine treatment group

Component of SNRS	Period 1	Period 2
Mentation and Mood	0.12	-0.44
Cranial Nerves		
Visual Acuity	0.24	-0.07
Fields, Discs, Pupils	0.01	0.01
Eye Movements	0.04	-0.39
Nystagmus	0.09	-0.20
Lower Cranial Nerves	0.21	-0.19
Motor		
RU	0.19	-0.23
LU	0.54	-0.23
RL	-0.08	-0.18
LL	0.32	-0.41
DTRS		
UE	0.36	-0.14
LE	0.55	0.21
Babinski		
R	0.18	-0.33
L	0.24	-0.07
Sensory		
RU	0.15	-0.02
LU	0.19	-0.16
RL	0.24	-0.05
LL	0.15	0.02
Cerebellar		
UE	0.36	-0.04
LE	0.71	-0.14
Gait; Trunk and Balance	0.19	-0.59
Special Category	0.22	0.12
Bladder/Bowel/Sexual Dysfunction		

Note: Patients randomized to initial cladribine received active therapy during Period 1, Months 1 to 12, and were then crossed over to placebo in Period 2, Months 13 to 24. See text for further details.

of points in the SNRS directly reflects the examiner's clinical assessment of each component in the neurological examination; points assigned to each component are summed to arrive at the overall SNRS score.

Our aim here is to examine the responsiveness of the SNRS to change, in the context of the clinical trial we undertook. Data consist of the monthly overall SNRS scores, as well as the individual component scores, for the 21 (out of 24) patients initially randomized to cladribine and 22 (out of 24) patients initially randomized to placebo, who completed the 24-month study. (Reasons for withdrawal from the study are discussed in Beutler et al.⁴)

We use a standard statistical procedure, the random effects

model,⁸ for the analysis of the longitudinal data comprising the scores on each component per patient over the 24 months of the trial. The random effects model allows within-subject correlations in scores across the repeated observations, and yields estimates of random person-specific effects, which represent a patient's trend across the time frame of the study. Results here are based on a linear model in which scores vary linearly over each one-year period, since no significant improvement in fit could be found with higher order models. Thus each individual has a random intercept, and a random slope over the first period and the second period. Hence, for each patient, an intercept and linear slope for each 12-month interval are estimated in addition to overall average intercept and slopes in each group, with the differences between an individual trend

Table 2: Effect sizes of individual components of the SNRS for the initial placebo treatment group

Component of SNRS	Period 1	Period 2
Mentation and Mood	0.37	-1.16
Cranial Nerves		
Visual Acuity	-0.03	0.32
Fields, Discs, Pupils	0.29	0.68
Eye Movements	-0.78	0.02
Nystagmus	-0.97	0.06
Lower Cranial Nerves	-0.00	-0.24
Motor		
RU	-0.06	-0.02
LU	-0.17	0.90
RL	-0.22	0.40
LL	-0.43	0.52
DTRS		
UE	0.30	0.71
LE	-0.17	1.29
Babinski		
R	0.04	0.46
L	-0.19	0.54
Sensory		
RU	0.13	0.44
LU	0.10	0.81
RL	-0.22	0.25
LL	-0.15	0.31
Cerebellar		
UE	-0.33	0.70
LE	-0.12	0.47
Gait; Trunk and Balance	-0.49	-0.04
Special Category	-0.37	0.33
Bladder/Bowel/Sexual Dysfunction		

Note: Patients randomized to initial placebo received placebo during Period 1, Months 1 to 12, and then were crossed over to cladribine therapy in Period 2, Months 13 to 24. See text for further details.

Table 3: Effect sizes of summary indices of the SNRS, and of the EDSS, for the initial cladribine treatment group

Component of SNRS	Period 1	Period 2
Cranial Nerves	0.15	-0.08
Motor		
Right	-0.03	-0.23
Left	0.41	-0.42
Upper	0.36	-0.27
Lower	0.12	-0.31
Total	0.19	-0.36
DTRS	0.61	0.11
Babinski	0.23	-0.24
Sensory		
Right	0.24	-0.04
Left	0.18	-0.04
Upper	0.18	-0.15
Lower	0.20	-0.00
Total	0.21	-0.03
Cerebellar	0.59	-0.09
Overall SNRS	0.65	-0.34
EDSS	-0.06	0.25
T2 Lesions (MRI)	-0.04	0.26

estimate and that subject's group estimate indicating how an individual subject's trend differs from the subgroup as a whole. In this regression context, we implicitly assume that deviations of observed scores from expected values are approximately normally distributed. This assumption is likely to be valid with summary scales such as the SNRS; and, the technique should be robust to moderate departures from normality.

We quantify responsiveness by means of effect sizes, for both the overall SNRS and its components. The basic principle of an effect size is to take the mean change in a variable and divide it by the baseline standard deviation of that variable.^{9,10} Typically, it is calculated as the difference between the mean score at time 2 and the mean score at time 1, divided by the standard deviation of the mean score at time 1.^{11,12} (Any measure of responsiveness will typically relate clinically meaningful change scores in a health status instrument to the variability of scores in the target population¹³). We determined the effect sizes for the overall SNRS and its components using the fitted values of each patient from the random effects model. There is some nonspecific variability in SNRS scores;⁵ hence to be responsive, the instrument must detect change over and above this nonspecific degree of change.¹³ The random effects model smooths out the within-subject variability over time, so effect size estimates from this model should be more stable than determinations from the raw data.

All computations were performed in SPlus,¹⁴ using the mixed effects model subroutine written by Pinheiro and Bates,¹⁵ as implemented on a Sun Sparc 2 workstation.

RESULTS

We have reported previously² that most patients originally randomized to cladribine therapy tend to demonstrate improvement in SNRS scores over the first 12 months of the trial. Analogously, most patients originally randomized to placebo tend to show improvement in SNRS scores over the second 12 months of the trial, when they had been crossed over to cladribine therapy. The analogous pattern with EDSS scores also obtains; see Figure 2.

In Tables 1 and 2, we give the average effect sizes of the various components of the Scripps Neurologic Rating Scale (SNRS) over the course of the clinical trial. In Tables 3 and 4, we present analogous results from similar analyses based on summations of the individual indices within patients. For example, we introduced four different summations of the individual components comprising the motor and sensory evaluations: Right = RU + RL, Left = LU + LL, Upper = RU + LU, and Lower = RL + LL. A positive effect size indicates an improvement in that component, and a negative effect size indicates a deterioration in that component.

Table 4: Effect sizes of summary indices of the SNRS, and of the EDSS, for the initial placebo treatment group

Component of SNRS	Period 1	Period 2
Cranial Nerves	-0.37	0.39
Motor		
Right	-0.19	0.27
Left	-0.39	0.73
Upper	-0.13	0.33
Lower	-0.32	0.52
Total	-0.29	0.51
DTRS	0.04	1.40
Babinski	-0.09	0.52
Sensory		
Right	-0.09	0.40
Left	-0.05	0.52
Upper	0.13	0.71
Lower	-0.19	0.28
Total	-0.07	0.47
Cerebellar	-0.25	0.66
Overall SNRS	-0.50	0.75
EDSS	0.53	-0.10
T2 Lesions (MRI)	0.29	0.15

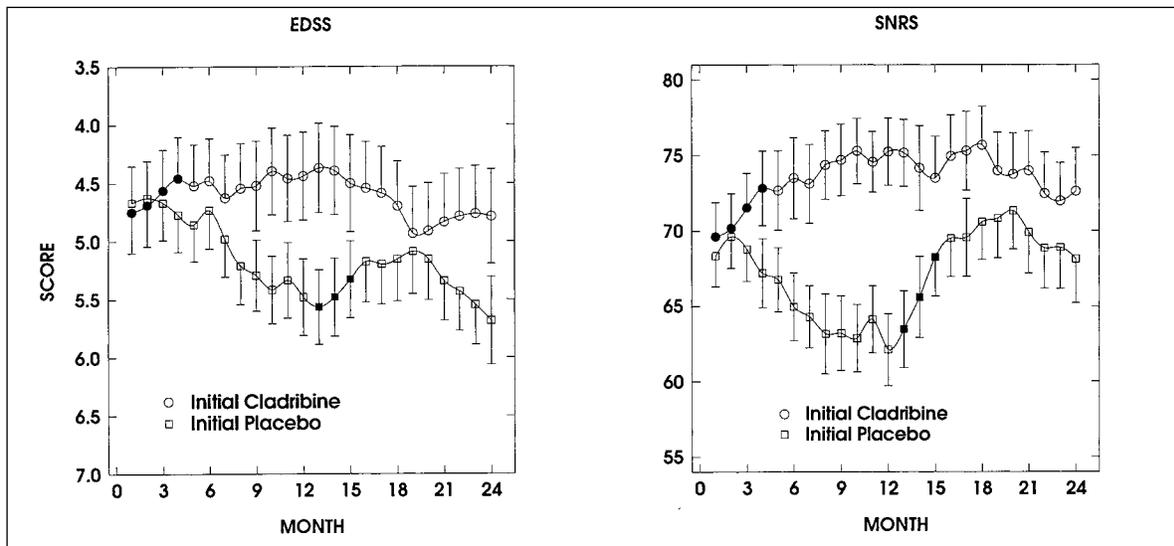


Figure 2: Overall EDSS and SNRS scores during the course of the clinical trial. The solid symbols indicate months when drug was administered. The group of patients denoted by circles was given 0.7 mg of cladribine per kg of body weight, monthly for four months. The group of patients denoted by squares was given 0.7 mg of cladribine per kg of body weight for one month, followed by two months of drug at 0.35 mg per kg. Bars represent 1 standard error of the mean.

Given the crossover design of the clinical trial, one might anticipate that effect sizes in the initial cladribine group would tend to be positive in the Month 1 to Month 12 period (Period 1), with some deterioration in the Month 13 to Month 24 period (Period 2). This pattern seems to hold for most of the components, though there are some exceptions: mentation and mood seems to decline rather severely in Period 2 compared to the improvement in Period 1, and DTRS (summation) tends to show continued improvement in Period 2. Conversely, one might expect some deterioration in effect sizes during Period 1 for the initial placebo group, but improvement during Period 2. Again, this pattern seems borne out, with the notable exception of mentation and mood, and lower cranial nerves. These expected patterns are more readily apparent with the summation indices, Tables 3 and 4, than with the individual components, Tables 1 and 2.

The initial placebo group received active therapy during Period 2, hence effect sizes during this period might be compared to those of the initial cladribine group during Period 1. In this comparison, the effect sizes for the initial placebo group are somewhat higher than those of the initial cladribine group, a reflection of the fact that scores for the initial placebo group declined rather dramatically during Period 1 (so that baseline values for the initial placebo group during Period 2 are smaller than the corresponding baseline values for the initial cladribine group during Period 1). Note also that the direction of change is generally consistent across the individual components within each period for both treatment groups, and that the effect sizes for the overall SNRS reflect this, tending to exceed those of the individual components in absolute magnitude.

For comparison purposes, we also give in Tables 3 and 4 the effect sizes for the EDSS, obtained using the same analytic technique as with the SNRS and its components. (Recall that the

directions of shift for the EDSS and the SNRS are reversed: clinical worsening would be related to an increase in the EDSS and a positive effect size, but a decrease in the SNRS and a negative effect size.) Effect sizes for the EDSS and SNRS are congruent, in that they both connote improvement in Period 1 for the initial cladribine group, and in Period 2 for the initial placebo group (the respective periods of active drug therapy); and, there is deterioration during Period 2 in the initial cladribine group, and in Period 1 in the initial placebo group, the respective periods without active drug therapy. Effect sizes are generally smaller for the EDSS than for the SNRS.

Lastly, we also give in Tables 3 and 4 the effect sizes from the burden of disease measurements, that is, lesion volumes, obtained from the T2-weighted MRI scans obtained at six month intervals over the course of the trial. Again, the random effects model was used for these determinations. In the original cladribine group, T2-weighted lesion volumes remained relatively constant overall during Period 1. The negative sign of the effect size connotes some diminution in the burden of disease, but the summary effect size is nonetheless near zero. During Period 2, the initial cladribine group is off therapy, and there is a corresponding marked rise in T2-weighted lesion volumes. Here, the magnitude of the effect size, .257, is nearly identical to that evinced by the initial placebo group during Period 1, namely .294. On the other hand, the initial placebo group during Period 2 continues to show a rise in T2-weighted tumor volumes, albeit at a much reduced rate compared to Period 1.

DISCUSSION

Few of the many instruments available for measuring neurological outcomes have undergone comprehensive

evaluation.³ Our goal here has been to provide further information about the scientific properties of the Scripps neurologic rating scale (SNRS), in particular, its responsiveness to change within the context of a clinical trial of cladribine for treatment of secondary progressive multiple sclerosis. Indeed, responsiveness together with reliability and validity are necessary requirements for instruments designed to measure changes in health status over time in clinical trials, so as to infer treatment differences from group differences.¹³

We have used effect sizes, that is, standardized estimates of the magnitudes of statistically determined experimental effects, to quantify responsiveness of the SNRS and its components in our clinical trial. The effect size provides a standardized unit of measurement for comparing the sizes of changes in different clinical measures within a study.¹⁶ Hence the effect size approach provides translation of clinical findings into a numerical quantity with utility from a number of different perspectives.¹¹ In particular, effect sizes can be used as benchmarks for understanding changes in health status measures, and can be compared across studies to provide a readily interpretable evaluation of treatment effects. Cohen⁹ presented a detailed discussion of effect size measures and their interpretation.

Cohen suggested that an effect size of around 0.20 is considered small, one of 0.50 indicates moderate differences, and one of 0.80 or above, large differences. From this perspective, one might characterize the responsiveness of most of the components of the SNRS to be in the small to moderate range; only mentation and mood and DTRS have effect sizes exceeding 1.0 in absolute value, with the mentation and mood value during Period 2 for the initial placebo group being an anomaly. The effect sizes for the overall SNRS tend to exceed those of the individual components in magnitude, as might be expected with the direction of change being generally consistent across the components.

Examination of the individual components of the SNRS relating to motor and sensory function and other “directional” items is obviously useful for tracking changes within patients over time. However, since location of impairment is not constant across patients, the summary indices (Tables 3 and 4) provide more appropriate population estimates of change.

All patients in our clinical trial were secondary progressive. Nevertheless, it might be inferred from Figure 2 that these secondary progressive MS patients were somewhat atypical, in that patients off therapy (initial placebo patients in Period 1, and initial cladribine patients in Period 2) underwent marked deterioration in terms of EDSS and SNRS scores. A larger confirmatory trial of cladribine in chronic progressive multiple sclerosis has subsequently been undertaken: a beneficial effect was seen in presence and volume of gadolinium-enhanced lesions on T1-weighted MRI scans, but cladribine had no demonstrable effect on disability. This latter outcome might be attributable to a different case mix of chronic progressive patients than the Scripps cohort. (In particular, there was little deterioration in clinical scores among placebo patients in the

confirmatory trial. Corresponding effect sizes would likely cluster around zero.)

As mentioned previously, effect sizes provide useful summary measures for comparison of treatment groups, both within individual trials and in separate clinical trials. This point is illustrated with the effect sizes from the MRI findings, which indicate that cladribine therapy is associated with a diminution of the growth rate of lesions found on T2-weighted MRI scans in this cohort of secondary progressive MS patients. A concern regarding the use of MRI as a marker of disease activity in MS has been the apparent lack of association between MRI findings and disease activity seen clinically (as assessed, perhaps, by neurologic rating scales). In this study, however, the effect sizes from the MRI findings are qualitatively similar to those from the scales and components; the treatment groups can indeed be distinguished on the basis of the effect sizes.

A potential limitation of our study is the crossover design of the clinical trial. Drug therapy in each stage (one-year period) lasted a maximum of four months; so in particular there was an eight-month washout period (relevant especially for the initial cladribine group) prior to crossover in the second year of the trial. Double-blinding was maintained throughout the two-year course of the trial, so that attending neurologists and patients were blinded to treatment assignment. Moreover, a valid, if inefficient, method of analysis in this study would be comparison of the two treatment groups solely on the basis of the first year (Period 1) experience. In this regard, our conclusions would remain unchanged had we examined responsiveness of the SNRS solely over the first year, with no question of carryover. Moreover, it is reassuring that findings related to responsiveness in Period 2 are congruent with those from Period 1, and therefore lend further credence to our conclusions.

Although there is no consensus concerning the identification of a gold standard, many investigators regard the EDSS as a preferred measure of treatment outcome in clinical trials of experimental therapies in multiple sclerosis.^{17,18} In the context of this clinical trial, we previously reported⁵ that the EDSS and SNRS were not strongly correlated within individual patients, and that the scales were differentially sensitive to clinical changes over time, with the EDSS indicating a more abrupt, and the SNRS a more gradual, change in the clinical course of disease. The effect sizes given here suggest that the SNRS might be more sensitive to detecting clinical changes than the EDSS. This should be interpreted with some caution, however, as the rather weak association between EDSS and SNRS scores strongly suggests that the underlying constructs that they are intended to measure are not coincident.

Let us remark that we have used the SNRS to quantify the severity of relapses in a different clinical trial of cladribine in relapsing-remitting MS. Our motivation for this was the belief that the severity of relapses can be quantitated with more sensitivity using the SNRS rather than the EDSS, but the present study provides only indirect evidence in support of this concept.

The EDSS and the SNRS represent operationally a reduction

of multivariate data – that is, assessment of function in different neurologic systems and assessment of patients' physical abilities – to a univariate quantity. In general, there are advantages accruing to the use of a summary score, including the avoidance of multiplicity, clinical validity (so long as the individual components comprising the pooled index are each clinically important disease manifestations that have face validity), and improved sensitivity due to expected reduction in measurement noise.¹³ Nevertheless, the fundamental question arises as to whether the reduction of multivariate information to an overall univariate score effects an adequate representation of the information available with the original assessments. In particular, it is not at all obvious how to determine the optimal weights or more generally the best method of combining the individual measures into a single index. We have previously advocated a multivariate statistical technique, principal components analysis, for this problem:¹⁹ this method identifies a linear combination of the individual measures that provides maximal spread (variation) among patients so as to maximize the information content of that combination.

The SNRS converts the neurological examination into an ordinal impairment scale with an arbitrary weighting system. In particular, several potentially prognostic variables are combined into an overall risk score. It is therefore not at all surprising that various components of the SNRS appear more responsive than the overall summary measure, and others, less responsive. In particular, the various items comprising the SNRS are not expected to be homogeneous, because they measure different attributes of a complex clinical syndrome rather than different facets of the same attribute. If one wishes to develop a new assessment system, then the results presented here suggest items which are clearly unresponsive, marking them as candidates for deletion: an excess of irrelevant items might well obscure the ability of a multiple-item index such as the SNRS to detect a change.^{7,20} In particular, the mentation and mood component evinces a pattern somewhat inconsistent with those of the other components; and, the special category of bladder, bowel, or sexual dysfunction seems rather insensitive to change. (It is noteworthy that these are the only items on the SNRS which are not "objective", in the sense that, rather than being scored directly from neurological examination, the attending neurologist queries the patient.) Hence these components are not likely to be useful as part of a multivariate or composite outcome measure; more detailed neuropsychological assessment might be appropriate. Moreover, if examination of the individual components is considered desirable, then the summation indices introduced in Tables 3 and 4 seem rather more stable than the individual components in Tables 1 and 2.

ACKNOWLEDGEMENTS

This research was supported in part from NIH grants RR00833 and NS30218, the Sam and Rose Stein Charitable Trust Fund, and a grant from the R. W. Johnson Pharmaceutical Research Institute. This is manuscript number 11173-MEM from The Scripps Research Institute.

REFERENCES

1. Sipe JC, Romine JS, Koziol JA, McMillan R, Zyroff J, Beutler E. Cladribine in treatment of chronic progressive multiple sclerosis. *Lancet* 1994;344:9-13.
2. Beutler E, Sipe JC, Romine JS, Koziol JA, McMillan R, Zyroff J. The treatment of chronic progressive multiple sclerosis with cladribine. *Proc Natl Acad Sci USA* 1996;93:1716-1720.
3. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. *J Neurol Neurosurg Psychiatry* 1996;60:127-130.
4. Sipe JC, Knobler RL, Braheny SL, Rice GP, Panitch HS, Oldstone MB. A neurologic rating scale (NRS) for use in multiple sclerosis. *Neurology* 1984;34:1368-1372.
5. Koziol JA, Frutos A, Sipe JC, Romine JS, Beutler E. A comparison of two neurologic scoring instruments for multiple sclerosis. *J Neurol* 1996;243:209-213.
6. Whitaker JN, McFarland HF, Rudge R, Reingold SC. Outcomes assessment in multiple sclerosis clinical trials: a critical analysis. *Multiple Sclerosis* 1995;1:37-47.
7. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chron Dis* 1985;38:27-36.
8. Laird NM, Ware JH. Random effects models for longitudinal data. *Biometrics* 1982;38:963-974.
9. Cohen J. *Statistical Power Analysis for the Behavioral Sciences* (Rev. Ed.). New York: Academic Press, 1977.
10. Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A, Mowat A. Importance of sensitivity to change as a criterion for selecting health status measures. *Qual Health Care* 1992;1:89-93.
11. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health studies. *Med Care* 1989;27:S178-S189.
12. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. *Control Clin Trials* 1991;12:142S-158S.
13. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chron Dis* 1987;40:171-178.
14. Statistical Sciences Inc. *SPlus User's Manual*, Version 3.0. Seattle: StatSci, 1991.
15. Pinheiro JC, Bates DM. *lme and nlme: mixed effects models methods and classes for S and Splus*. Technical report. In: University of Wisconsin: Department of Biostatistics, 1995.
16. Rudick R, Antel J, Confavreux C, et al. Clinical outcomes assessment in multiple sclerosis. *Ann Neurol* 1996;40:469-479.
17. Noseworthy JH, Vandervoort MK, Hopkins M, Ebers GC. A referendum on clinical research in multiple sclerosis: the opinion of the participants in the Jekyll Island workshop. *Neurology* 1989;39:977-981.
18. Ellison GW, Myers LW, Leake BD, et al. Design strategies in multiple sclerosis clinical trials. *Ann Neurol* 1994;36:S108-S112.
19. Koziol JA, Hacke W. Multivariate data reduction by principal components, with application to neurological scoring instruments. *J Neurol* 1990;237:461-464.
20. Wright JG, Feinstein AR. A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *J Clin Epidemiol* 1992;45:1201-1218.