Predictive Value of Lesions for Relapses in Relapsing-remitting Multiple Sclerosis

James A. Koziol, Simone Wagner, David F. Sobel, Lloyd S. Slivka, John S. Romine, Jack C. Sipe and Hans-Peter Adams

AJNR Am J Neuroradiol 2001, 22 (2) 284-291
http://www.ajnr.org/content/22/2/284

AJNR Am J Neuroradiol 22:284–291, February 2001

Predictive Value of Lesions for Relapses in Relapsing-remitting Multiple Sclerosis

James A. Koziol, Simone Wagner, David F. Sobel, Lloyd S. Slivka, John S. Romine, Jack C. Sipe, and Hans-Peter Adams

BACKGROUND AND PURPOSE: **Recent studies have suggested that enhancing lesions on contrast-enhanced T1-weighted MR images are predictive of impending exacerbations in cases of relapsing-remitting multiple sclerosis. We examined whether enhancing lesions, new enhancing lesions, and new hypointense lesions (''black holes'') could accurately predict exacerbations in a cohort of 50 patients with relapsing-remitting multiple sclerosis within a time frame of up to 6 months.**

METHODS: **Data were obtained from 50 patients with relapsing-remitting disease. All patients underwent monthly MR imaging and clinical examinations for a period of 12 months. Putative predictors of clinical relapse were defined from enhancing lesions, new enhancing lesions, and new black hole outcomes, and their operating characteristics were studied.**

RESULTS: **Overall, the positive predictive values (PV+) of enhancing lesions, new enhancing lesions, or new black holes for an exacerbation did not exceed 0.25 and the negative predictive values (PV−) were all near 0.9. The best predictor for new enhancing lesions was the occurrence of new enhancing lesions in each of the previous 3 months (PV+: 0.79 [95% confidence interval, 0.651–0.900]; PV−: 0.83 [95% confidence interval, 0.751–0.887]). Similarly, new black holes were predicted best by the occurrence of new black holes in each of the previous 2 months (PV+: 0.54 [95% confidence interval: 0.372–0.697]; PV−: 0.85 [95% confidence interval, 0.790–0.896]).**

CONCLUSION: **None of the MR markers could predict an impending relapse with any reasonable degree of precision. Rather, the absence of MR markers is associated with a more favorable clinical course (ie, fewer relapses).**

MR imaging has become the most important paraclinical test for diagnosing multiple sclerosis, for delineating its natural history, and potentially for use as an objective quantitative outcome measure in assessing the response of patients with multiple sclerosis to experimental therapy (1, 2). The images reveal the multiple, primarily periventricular, lesions that grow and shrink at different rates in various regions of the brain. Acute inflammatory lesions associated with multiple sclerosis enhance after the injection of contrast media because of a breakdown of the blood-brain barrier. The appearance of enhancing lesions has been widely accepted as a measure of disease activity in cases of multiple sclerosis (3).

Nevertheless, disease-related activity, as measured from MR images, remains a complex issue. New lesions, enhancing lesions, hypointense or hyperintense lesions, and changes in lesion size have all been cited as potential measures of pathophysiological mechanisms in cases of multiple sclerosis, with varying but generally modest degrees of correlation with clinical assessments (2). The clinical usefulness of MR imaging in the assessment of patients with multiple sclerosis is predicated on the assumptions that tissue types can be accurately and precisely classified by images and that there is a relation to the clinical course of disease (4).

It is this latter assumption that we investigate herein. In particular, we examine whether three MR imaging–derived markers—enhancing lesions, the appearance of new enhancing lesions, and the occurrence of new hypointense lesions (black holes, new black holes) on contrast-enhanced T1-weighted conventional spin-echo images—are predictive of short-term outcomes, such as the occurrence of

Received April 5, 2000; accepted after revision, July 10, 2000.

From the Department of Molecular and Experimental Medicine (J.A.K., S.W., L.S.S., J.C.S., H.-P.A.), The Scripps Research Institute, the Department of Radiology (D.F.S.) and the Division of Neurology (J.S.R., J.C.S.), Scripps Clinic, La Jolla, CA, and the Department of Neurology (S.W.), Ruprecht-Karls University, Heidelberg, Germany.

Address reprint requests to James A. Koziol, PhD, The Scripps Research Institute, Department of Molecular and Experimental Medicine, MEM216, 10550 N. Torrey Pines Road, La Jolla, CA 92037.

exacerbations in relapsing-remitting multiple sclerosis. Our goal was to assess whether individual patient events (exacerbations) can be temporally predicted with any level of confidence based on these MR imaging findings. Such determination would lend further support to the usefulness and value of MR imaging in providing markers of disease activity and progression in cases of multiple sclerosis, with obvious implications relating to therapeutic intervention on the individual patient level.

Methods

Study Design

MR imaging and exacerbation data were gathered monthly during the course of a randomized, double-blind, placebo-controlled clinical trial assessing the efficacy of cladribine for the treatment of relapsing-remitting multiple sclerosis. The primary results of the trial are reported elsewhere (5). We herein focus on short-term clinical outcomes (exacerbations in particular) and attempt to predict the clinical manifestations of relapsing-remitting multiple sclerosis based on MR imaging findings, including the presence of enhancing lesions, the occurrence of new enhancing lesions, and the occurrence of new black holes, on contrast-enhanced T1-weighted conventional spin-echo images. The analyses reported herein are based on the 50 patients (of the 52 who were enrolled) whose conditions were evaluable at 12 months. Demographic data relating to these 50 patients are presented in Table 1.

Clinical neurologic examinations of all patients were performed at study entry and were repeated every month for the 1st year of the trial, as well as within 48 hours of report by a patient of a relapse (exacerbation). A clinical relapse was defined as the appearance of new symptoms or worsening of an existing symptom attributable to multiple sclerosis and accompanied by objective worsening of neurologic findings. To be scored as a relapse, the alterations must have been preceded by disease stability or improvement lasting for at least 30 days and the worsening must have lasted at least 24 hours and have occurred in the absence of fever. All relapses were identified by the attending neurologists (J.S.R., J.C.S.).

MR Imaging Analyses

Following the protocol for this clinical trial, patients underwent MR imaging at baseline (time of entry into the trial) and then at monthly intervals thereafter during the initial year of the trial. All MR imaging was performed on a 1.5-T General Electric Signa imager at the MR imaging facility at Scripps Clinic. T2- and proton density–weighted images were obtained using a conventional spin-echo sequence with 2500/30/90 (TR/TE/TE). Sections were 4 mm thick, with a 1-mm intersection gap. T1-weighted images of 3-mm thickness and 0 intersection gap were obtained approximately 10 minutes after the administration of gadopentotate dimeglumine (Magnevist; Berlex Laboratories, Wayne, NJ) (0.1 mmol/kg) to ensure optimal time for transmigration of the contrast agent across the blood-brain barrier.

Two observers (S.W., D.F.S.) agreed on the definition of hypointense lesions as described by Truyen et al (6) and, in a preliminary study, evaluated a number of images of patients with multiple sclerosis who were not included in the clinical trial. Subsequently, one observer (S.W.) undertook quantification of numbers of enhancing lesions, new enhancing lesions, and new black holes on a monthly basis; clarification and review were provided by the other observer (D.F.S.). Both observers were blinded to the clinical data, including treatment assignment.

TABLE 1: Baseline demographic and clinical characteristics of 50 relapsing-remitting multiple sclerosis patients

	Placebo (n = 24)	Cladribine (n = 26)
Sex		
Male	7	8
Female	17	18
Race		
White	24	23
Other	0	3
Age (yrs)		
Mean	40.1	44.0
25th Percentile	36.5	38.5
50th Percentile	41.0	44.5
75th Percentile	44.0	49.5
Range	31–52	31–52
Years with Symptoms		
Mean	9.1	10.2
25th Percentile	3.5	4.5
50th Percentile	9.0	8.0
75th Percentile	12.5	12.5
Range	1–25	1–29
Number of exacerbations in previous year		
1	12	5
2	5	16
3 or 4	7	5
Baseline EDSS		
Mean	3.8	3.9
25th Percentile	2.5	2.3
50th Percentile	3.5	3.5
75th Percentile	5.3	5.5
Range	2.5–6.5	2–6.5
Baseline SNRS		
Mean	75.8	76.1
25th Percentile	67.0	65.5
50th Percentile	75.5	78.5
75th Percentile	86.0	86.5
Range	54–98	41–93

Reproducibility of counts was assessed by the random selection of 56 images, and reevaluation was performed by one of the observers (S.W.) approximately 2 months after completion of all counting. The kappa statistic was calculated to assess the intrarater level of agreement with the replicate counts (7). Reproducibility was reasonably good regarding counts of enhancing lesions (kappa = 0.67, standard error = 0.07) and regarding counts of black holes (kappa = 0.70, standard error = 0.08).

Statistical Methods

We considered a number of putative predictors of clinical relapse in a particular month: enhancing lesions, new enhancing lesions, new black holes, or combinations thereof in the nearest preceding MR image or in a consecutive sequence of previous monthly MR images. For example, to predict an event in month 5 of the trial, the findings of the four preceding MR images were available and could be used in this endeavor. We used the logical operators AND and OR to combine the putative predictors across months when combining information.

For each patient, we calculated the sensitivity, specificity, positive predictive value, and negative predictive value associated with each of these various MR indices relative to the

TABLE 2: 2 × 2 Contingency table and derived values for an individual patient

	Event		
Marker	Present	Absent	Total
Present	a	b	a + b
Absent	c	d	c + d
Total	a + c	b + d	n

Note.—Operating characteristics of a marker/predictor relative to the occurrence (presence) or nonoccurrence (absence) of a subsequent clinical event (exacerbation) summarized for any individual patient, over the course of n = a + b + c + d months.

Sensitivity (S+) = a/(a + c); specificity (S−) = d/(b + d); positive predictive value (PV+) = a/(a + b); negative predictive value (PV−) = d/(c + d); and hit rate (HR) = (a + d)/n.

clinical event of exacerbation. Note that during the course of the clinical trial, each patient's monthly outcome (occurrence [presence] or nonoccurrence [absence] of a clinical relapse during that month), together with the MR imaging–derived predictor for the clinical outcome that month can be summarized in a 2 × 2 contingency table, which renders straightforward calculation of the operating characteristics of the MR index as a prognostic test (Table 2).

We then combined the sample estimates across patients with a random effects model, as used in metaanalysis (8). The choice of the random effects model rather than the fixed effects model was made under the assumption that patients with relapsing-remitting multiple sclerosis are likely heterogeneous in terms of both MR imaging findings and clinical events and that this heterogeneity should be taken into account for inferential purposes. In this regard, we used the arcsine transformation on the individual estimates of the parameters of sensitivity, specificity, and positive and negative predictive values and previous combinations via the random effects model to ensure stability of the variances of the individual patient estimates prior to combining via the random effects model. For each potential prognostic index that we constructed, we can summarize the operating characteristics by means of point estimates and associated 95% confidence intervals.

Results

Demographics relating to all 50 patients with relapsing-remitting multiple sclerosis randomized into the clinical trial and with conditions that were evaluable at 12 months are presented in Table 1. Exacerbations, enhancing lesions, new enhancing lesions, and new black holes relating to these patients are summarized in Table 3. Based on the information presented in Table 1, baseline characteristics of the two treatment cohorts seem comparable. On the other hand, note the profound effect of cladribine on enhancing lesions, as presented in Table 3.

On a monthly basis, the proportions of patients with enhancing lesions, new enhancing lesions, or new black holes always exceeded the proportions of patients with exacerbations. With few exceptions, however, all these monthly proportions were below 50% (Table 3). We summarized individual patient findings and outcomes, as presented in Table 2, and then combined sensitivity, specificity, positive predictive value, and negative predictive value across all patients by means of a random effects model to allow for patient heterogeneity in these measures. The results up to 6 months with markers/predictors combined by the OR function are presented in Figure 1. In this regard, the OR function denotes that the presence of the marker (enhancing lesions, new enhancing lesions, or new black holes) in any of the indicated number of preceding months (1–6 months) is considered to be a marker that is ''present'' in the Table 2 dichotomization.

Note that the results shown in Figure 1 are presented separately for the placebo group, the cladribine group, and, for positive and negative predictive values, the combined data. Because cladribine reduces or eliminates enhancing lesions, sensitivities of the MR indices derived from them will tend to be higher in the placebo group than in the cladribine group; correspondingly, specificities will tend to be higher in the cladribine group than in the placebo group. In comparison, predictive values remain relatively unchanged across the two treatment groups and are combined for increased precision of summary measures of predictive values.

TABLE 3: Number of patients with exacerbations and MR markers by month in trial

A. Placebo (n = 24)

	Month in Trial											
	1	2	3	4	5	6	7	8	9	10	11	12
Exacerbations		2	3	0	2	2	4	2	1	4	4	5
Enhancing lesions	9	10	9	11	14	14	11	13	7	12	11	15
New enhancing lesions		10	7	11	12	14	11	13	5	11	10	12
New black holes		5	1	2	4	4	4	6	4	6	3	3

B. Cladribine (n = 26)

	Month in Trial											
	1	2	3	4	5	6	7	8	9	10	11	12
Exacerbations		4	0	4	0	5	0	3	1	1	2	3
Enhancing lesions	13	11	7	4	5	3	0	0	0	0	0	0
New enhancing lesions		10	3	2	2	1	0	0	0	0	0	0
New black holes		4	6	4	4	8	4	2	3	5	1	7

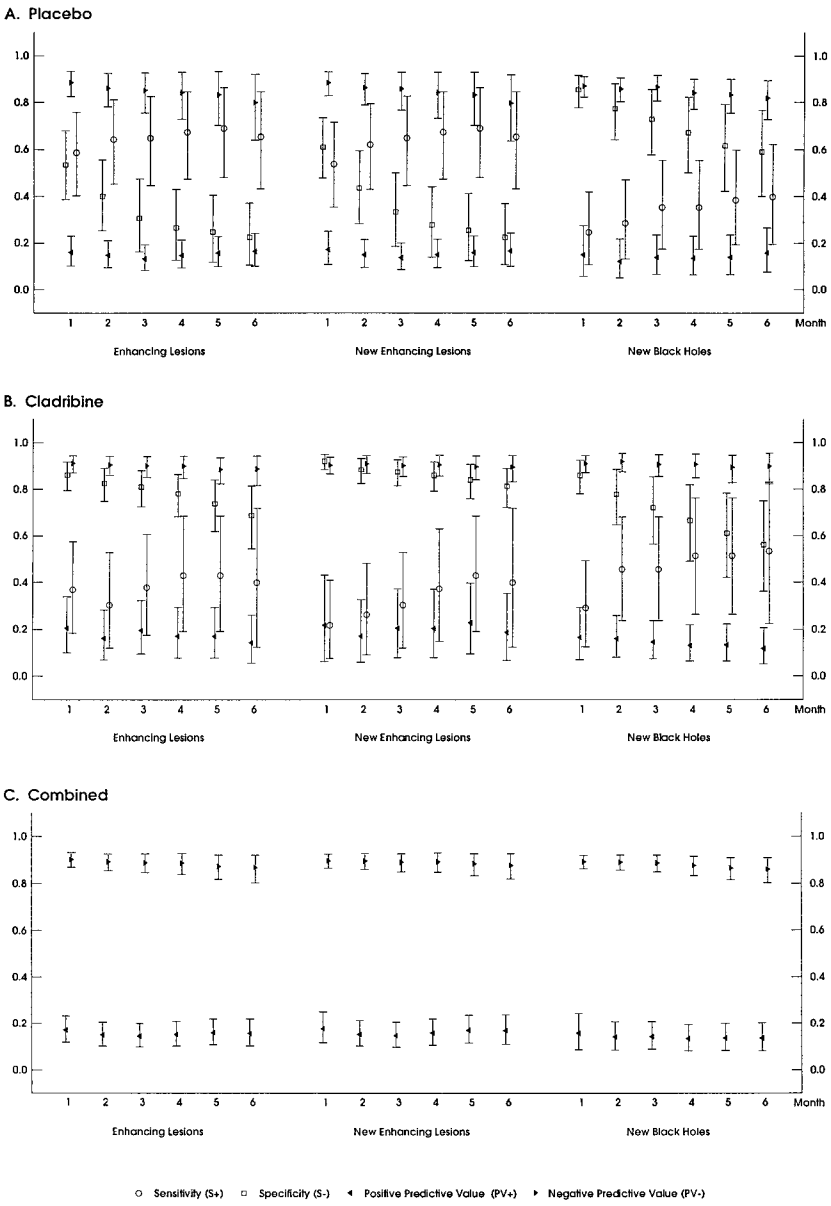

FIG 1. Summary operating characteristics and associated 95% confidence intervals for MR markers (OR combination) relative to the occurrence of subsequent relapse. Month denotes the number of months and MR images preceding the event time point. MR markers are combined across each of these time segments of months using the OR function. If the MR marker is present in any of the preceding months, the marker is set to 1; if absent in all months, the marker is set to 0.

Indices derived from enhancing lesions and new enhancing lesions have sensitivities ranging from 0.54 to 0.69 but lower specificities, from 0.22 to 0.61, relative to the occurrence of exacerbations among the placebo cohort. The corresponding indices from the cladribine group are clearly dissimilar, with sensitivities from 0.21 to 0.43 and specificities from 0.69 to 0.92. Note the trade-off between sensitivity and specificity; as one increases (with number of months combined by OR), the other decreases. In contrast to the enhancing lesion and new enhancing lesion findings, the indices derived from new black holes are not confounded by drug treatment; in the two cohorts, sensitivities range from 0.25 to 0.53 and specificities range from 0.56 to 0.86.

Compared with sensitivities and specificities, positive and negative predictive values are relatively constant in both cohorts, with no evidence of statistically significant differences. Positive predictive values are all low, ranging from 0.12 to 0.23 in the two cohorts, but negative predictive values are all in excess of 0.80, ranging up to 0.92 for enhancing lesions, new enhancing lesions, and new black holes.

Considering the marked heterogeneity in sensitivities and specificities between the two cohorts for enhancing lesions and new enhancing lesions, we chose not to combine these indices into overall summary measures. With no evidence of heterogeneity in any of the indices relating to predictive values, we do provide overall summary measures for them in Figure 1C. In these combined indices, positive predictive values are all in a narrow range from 0.13 to 0.18, and, similarly, negative predictive values are all between 0.86 and 0.90. Note the narrowness of the corresponding 95% confidence intervals; the

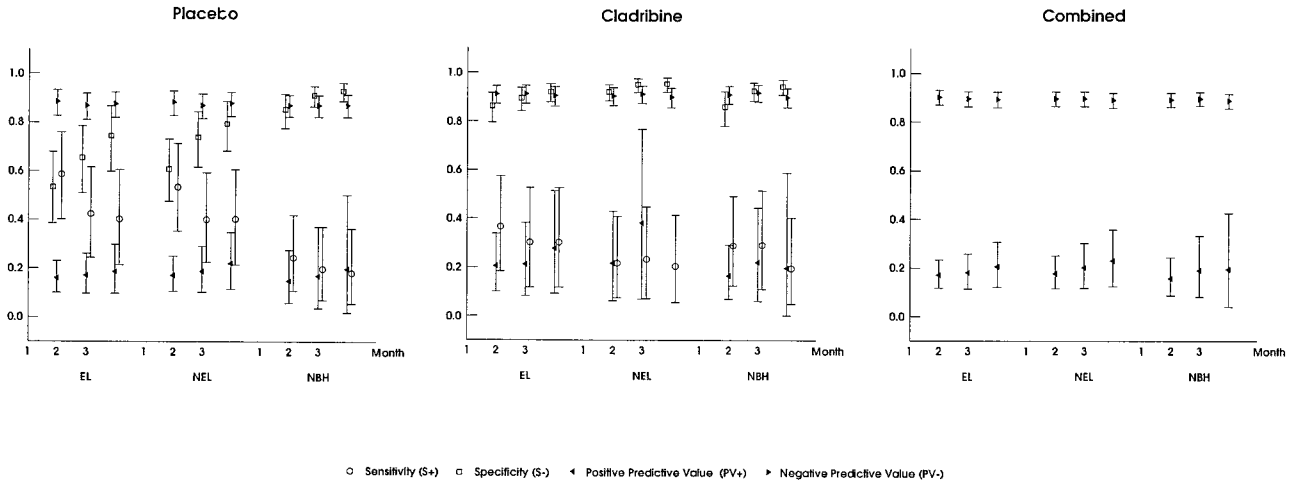

FIG 2. Summary operating characteristics and associated 95% confidence intervals for MR markers (AND combination) relative to the occurrence of a subsequent relapse. Month denotes the number of months and MR images preceding the event. MR markers are combined across each of these time segments of months using the AND function. That is, if the MR marker is present in all of the preceding months, the marker is set to 1; if absent in any of the months, the marker is set to 0.

TABLE 4: Operating characteristics of MR markers as predictors for exacerbations

Estimate (95% CI)	Event Exacerbation		
MR-Marker/Predictor Time to Event (Months)*	Enhancing Lesions Every 3	New Enhancing Lesions Every 3	New Black Holes Every 3
Positive predictive value	0.21 (0.121–0.306)	0.23 (0.124–0.357)	0.20 (0.041–0.426)
Negative predictive value	0.89 (0.859–0.923)	0.89 (0.857–0.920)	0.89 (0.855–0.916)
Sensitivity	0.36 (0.220–0.508)	0.31 (0.180–0.459)	0.19 (0.085–0.321)
Specificity	0.85 (0.778–0.903)	0.89 (0.841–0.929)	0.94 (0.911–0.959)
Hit Rate	0.86 (0.785–0.915)	0.89 (0.831–0.932)	0.93 (0.890–0.961)

Note.—CI, confidence interval; every 3, occurrence of MR marker in the 3 consecutive monthly MR images immediately preceding the clinical event (exacerbation).

upper limit of all the 95% confidence intervals for positive predictive values is <0.25, and the lower limit of all the 95% confidence intervals for negative predictive values is >0.80.

We also used a more stringent criterion, as defined by the AND function, to combine MR predictors. A predictor was said to be present if the MR marker occurred in all of the preceding months, otherwise not. Results to 3 months are depicted in Figure 2. Note that sensitivity decreases as more months are added to the criterion, and specificity increases. In the combined indices, positive predictive values are all <0.24 and negative predictive values are all approximately 0.89.

The criteria that yielded the highest positive predictive values for exacerbations are presented in Table 4. As in the figures, the precision of the summary predictive value estimates, as reflected by the narrowness of the corresponding 95% confidence intervals, is remarkably good, implying the unlikelihood that positive predictive values would be much larger in magnitude than the values reported herein, had our cohort size been substantially larger.

We also examined the operating characteristics of the prediction of various MR indices among themselves. We briefly summarize these findings: enhancing lesions are predicted best by the occurrence of new enhancing lesions during each of the previous 2 months (PV+: 0.82 [95% confidence interval, 0.714–0.898]; PV−: 0.81 [95% confidence interval: 0.739–0.869]), new enhancing lesions are predicted best by new enhancing lesions during each of the previous 3 months (PV+: 0.79 [95% confidence interval, 0.651–0.900]; PV−: 0.83 [95% CI, 0.751–0.887]), and new black holes are best predicted by new black holes during each of the previous 2 months (PV+: 0.54 [95% confidence interval, 0.372–0.697]; PV−: 0.85 [95% CI, 0.790–0.896]). The operating characteristics of these predictions degraded by including other putative predictors such as exacerbations and the other lesion type, thus suggesting autonomous and self-propagating underlying disease processes for these two types of lesions.

Discussion

Exacerbations are a hallmark of relapsing-remitting multiple sclerosis; if remission is incomplete, permanent clinical worsening may ensue (9, 10).

Early prediction of potential exacerbations in individual patients would thus be valuable for clinical management of multiple sclerosis, especially for preemptive or ameliorative therapy, thereby offering the potential of delaying the onset of progressive disease. Moreover, patients likely to encounter a relapse would be the preferred patient group for testing the efficacy of experimental therapy. There are no incontrovertible predictors of exacerbations in individual patients; hence, it is of interest to investigate whether conventional MR imaging would be at all useful for this purpose.

In the current study, we identified a number of putative markers of exacerbations from the monthly MR imaging findings of a cohort of patients with relapsing-remitting multiple sclerosis and examined the operating characteristics of these markers when considered from the point of view of prediction of exacerbations. In this regard, we emphasize that clinical concern relating to these potential MR markers should not be regarding the intrinsic sensitivity and specificity of the markers but the accuracy of positive and negative test results for predicting clinical events. We were primarily interested in short-term predictive values of MR imaging findings and whether presence or absence of enhancing lesions, new enhancing lesions, or new black holes were reliable markers of subsequent clinical exacerbations, with a time frame of up to 6 months. We found that neither the occurrence of enhancing lesions nor that of new black holes was useful for predicting clinical events (small positive predictive values, typically <0.2) but that in the absence of enhancing lesions or new black holes, clinical exacerbations were unlikely to occur (high negative predictive values, typically approximately 0.9). Our estimates of predictive values are very precise, as reflected by the narrowness of the corresponding 95% confidence intervals depicted in Figures 1 and 2. The total number of MR images examined in this study is almost 600. Therefore, it is extremely unlikely that the positive predictive value of any MR marker would be much larger in magnitude than 0.25 (an upper limit of the reported 95% confidence intervals shown in Fig 1), a value that we judge to be of limited clinical usefulness. We conclude that when the pre-MR imaging likelihood of an exacerbation is low, a normal MR image (no enhancing lesions, no new black holes) tends to exclude the possibility of an exacerbation but a positive result is not particularly helpful in inferring that an exacerbation is about to occur.

Factors relating to our MR imaging methodology and evaluations might have influenced our findings (4, 11, 12). However, we relied solely on one experienced assessor (S.W.), with high intrarater reliability, for primary enumeration of enhancing lesions and black holes; we thereby avoided potential pitfalls associated with interrater variability (13). The application of triple doses of contrast agent would have increased the number of enhancing lesions (14), but this would likely have resulted in increased sensitivity and decreased specificity of any of the MR indices without improving the positive predictive value. It is unlikely that our determination of new black holes from contrast-enhanced and not from unenhanced T1-weighted MR images is critical. O'Riordan et al (15) confirmed a very strong correlation of T1 hypointense lesion volumes before and after contrast enhancement (15).

A second technical issue relates to our adoption of a random effects model rather than a fixed effects model for the pooling of operating characteristics across patients (8). This choice was predicated on the underlying assumption that, as in most randomized clinical trials, our patients constitute a random sample from an underlying population of patients with relapsing-remitting multiple sclerosis to which we want to generalize findings. In practice, the random effects model incorporates the degree of observed patient-to-patient variability into the estimates of the variability of the pooled indices of predictive values. This in turn enlarges the confidence intervals for the overall estimates of predictive values, relative to the fixed effects model, and produces a more conservative if more realistic view of the precision of the overall estimates.

Cladribine effectively eliminates all enhancing lesions in cases of relapsing-remitting multiple sclerosis and secondary progressive multiple sclerosis after standard therapeutic doses are administered for a period of $\geq$1 year (5, 16, 17). In this regard, the abrogation of enhancing lesions from cladribine is apparently more complete than with other immunomodulators used therapeutically in cases of multiple sclerosis. However, the mechanisms leading to these losses of enhancement are not precisely known, and whether the pathways of the various immunomodulators differ is speculative. Nonetheless, excacerbations did occur in our cohort despite the absence of enhancing lesions from six consecutive monthly MR images (cladribine induced or not). Conversely, only a minority of patients with enhancing lesions suffered from a relapse within 6 months. Operationally, the virtual elimination of enhancing lesions in the cladribine cohort leads to decreased sensitivities and increased specificities of putative MR markers of exacerbations relative to the placebo cohort, but the more important predictive values remain virtually identical. Within the 6-month time frame examined herein, enhancing lesions are of limited usefulness for predicting relapse.

Koudriavtseva et al (18) reported that the presence of a single enhancing lesion is predictive of subsequent relapse; ie, that enhancing lesions seen on T1-weighted MR images of patients with relapsing-remitting multiple sclerosis was positively correlated with the occurrence of relapses during the next 6 months. We attempted to refine their finding by focusing on a shorter time horizon and on individual patient outcomes rather than an ag-

gregate analysis. From the predictive values for relapse based on enhancing lesions in our longitudinal study, we would preferentially state that the absence of any enhancing lesions is highly predictive of no subsequent relapse but the presence of enhancing lesions is of substantially less value for predicting subsequent relapse.

Similarly, Kappos et al (19) concluded that enhancing lesions seem to help identify patients with a high risk of relapses. They had found that the correlation between enhancing frequency and relapse rate was modest during the 1st year and weakened during the 2nd year of their study. We focused on MR events more proximate to the actual clinical event than did Kappos et al; within this time frame, there is little predictive value for relapse attributable to the presence of blood brain–barrier disruption.

The study presented by van Walderveen et al (20) revealed a correlation between the number of enhancing lesions at study entry and development of new enhancing lesions during the study period. This agrees with our results, and we can quantitate our findings by means of our rather high positive and negative predictive values. However, we cannot confirm another finding of the Amsterdam group, that enhancing lesions predict subsequent new black holes. In our cohort, new black holes were predicted best by their occurrence in previous months. Patients in whom production of black holes is an ongoing process have the highest likelihood of producing new black holes, and patients in whom this process is not ongoing have a lower likelihood of producing new black holes. The positive predictive value for this event was higher than the prediction of new black holes from previous new enhancing lesions or exacerbations. We therefore suggest that accumulation of black holes may be, at least partially, an independent and self-propagating process; once started, further activity is likely. We suggest that the same is true for enhancing lesions.

As black holes may well represent a long-term marker of disease severity in cases of multiple sclerosis, it is perhaps not surprising that there is no strong level of association between new black holes and occurrence of exacerbations in cases of relapsing-remitting multiple sclerosis. Truyen et al (6) found no significant correlation between T1 hypointensity and subsequent relapse rate in cases of either relapsing-remitting multiple sclerosis or secondary progressive multiple sclerosis. We confirm the lack of positive predictive value for subsequent clinical exacerbations considering the occurrence of new black holes in our cohort of patients with relapsing-remitting multiple sclerosis.

We conclude with a general comment concerning our focus and strategy in this study. During the past several years, there has been a broad paradigmatic approach to the analysis of MR imaging findings in multiple sclerosis clinical studies: this approach is correlational in nature, with emphasis on the relatedness (via statistical correlations) of MR imaging findings (eg, burden of disease as determined from T1- or T2-weighted MR images) with clinical outcomes (eg, Kurtzke's Expanded Disability Status Scale scores) (21–23). Generally, the reported correlations between clinical outcomes and radiologic findings are lower than 0.6 and in most instances are approximately 0.2 to 0.4 in absolute magnitude (24, 25). In contrast, our approach represents a paradigmatic shift in that we have more expressly focused on the prediction of clinical outcomes based on MR imaging findings. Our aim was to determine whether actual prediction of individual clinical relapses based on MR imaging findings can be achieved with accuracy and precision. We think that the correlational approach is misplaced in the prognostic setting we have envisioned; rather, accurate prediction of clinical outcomes from covariate information, such as MR imaging findings embodied in the statistical concept of regression, should immediately be more useful and meaningful to clinicians and patients than knowledge that MR imaging findings and the clinical course of disease are (weakly) correlated.

Conclusion

Neither the presence of enhancing lesions, the appearance of new enhancing lesions, nor the occurrence of new hypointense lesions on contrast-enhanced T1-weighted conventional spin-echo MR images is useful for predicting, within a precise time frame of $\leq$6 months, the subsequent occurrence of exacerbations in cases of relapsing-remitting multiple sclerosis. Rather, the absence of these MR imaging–derived markers is associated with a more favorable clinical course.

Acknowledgments

The authors gratefully acknowledge the support of Ernest Beutler, MD, the principal investigator of the cladribine trials at Scripps Clinic. The clinical trials were supported by the Stein Endowment Fund, Grant RR00833 from the National Institutes of Health, Johnson & Johnson Pharmaceutical Research Institute, SFP-1004, and the Garate Memorial Fund for multiple sclerosis research. Dr. Wagner was supported by a grant from the Deutsche Forschungsgemeinschaft, WA 1058/1–2.

References

1. Paty DW, Li DK. **Interferon beta-1b is effective in relapsing-remitting multiple sclerosis: II. MR analysis results of a multicenter, randomized, double-blind, placebo-controlled trial: UBC MS/MR Study Group and the IFNB Multiple Sclerosis Study Group.** *Neurology* 1993;43:662–667
2. Miller DH, Grossman RI, Reingold SC, McFarland HF. **The role of magnetic resonance techniques in understanding and managing multiple sclerosis.** *Brain* 1998;121:3–24
3. Miller DH, Barkhof F, Nauta JJ. **Gadolinium enhancement increases the sensitivity of MR in detecting disease activity in multiple sclerosis.** *Brain* 1993;116:1077–1094
4. Miller DH. **Guidelines for MR monitoring of the treatment of multiple sclerosis: recommendations of the US Multiple Sclerosis Society's task force.** *Mult Scler* 1996;1:335–338

5. Romine JS, Sipe JC, Koziol JA, Zyroff J, Beutler E. **A double-blind, placebo-controlled, randomized trial of cladribine in relapsing-remitting multiple sclerosis.** *Proc Assoc Am Physicians* 1999;111:35–44
6. Truyen L, van Waesberghe JH, van Walderveen MA, et al. **Accumulation of hypointense lesions (“black holes”) on T1 spin-echo MR correlates with disease progression in multiple sclerosis.** *Neurology* 1996;47:1469–1476
7. Cohen J. **A coefficient of agreement for nominal scales.** *Educ Psychol Meas* 1960;20:37–45
8. Fleiss JL. **The statistical basis of meta-analysis.** *Stat Methods Med Res* 1993;2:121–145
9. Poser CM. **Exacerbations, activity, and progression in multiple sclerosis.** *Arch Neurol* 1980;37:471–474
10. Poser CM, Paty DW, Scheinberg L, et al. **New diagnostic criteria for multiple sclerosis: guidelines for research protocols.** *Ann Neurol* 1983;13:227–231
11. Filippi M, Horsfield MA, Ader HJ, et al. **Guidelines for using quantitative measures of brain magnetic resonance imaging abnormalities in monitoring the treatment of multiple sclerosis.** *Ann Neurol* 1998;43:499–506
12. Tofts PS. **Standardisation and optimisation of magnetic resonance techniques for multicentre studies.** *J Neurol Neurosurg Psychiatry* 1998;64:S37–S43
13. Barkhof F, Filippi M, van Waesberghe JH, Campi A, Miller DH, Ader HJ. **Interobserver agreement for diagnostic MR criteria in suspected multiple sclerosis.** *Neuroradiology* 1999;41:347–350
14. Filippi M, Rovaris M, Capra R, et al. **A multi-centre longitudinal study comparing the sensitivity of monthly MRI after standard and triple dose gadolinium-DTPA for monitoring disease activity in multiple sclerosis: implications for phase II clinical trials.** *Brain* 1998;121:2011–2020
15. O’Riordan J, Gawne Cain M, Coles A, et al. **T1 hypointense lesion load in secondary progressive multiple sclerosis: a comparison of pre versus post contrast loads and of manual versus semi automated threshold techniques for lesion segmentation.** *Mult Scler* 1998;4:408–412
16. 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
17. 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 U S A* 1996;93:1716–1720
18. Koudriavtseva T, Thompson AJ, Fiorelli M, et al. **Gadolinium enhanced MR predicts clinical and MR disease activity in relapsing-remitting multiple sclerosis.** *J Neurol Neurosurg Psychiatry* 1997;62:285–287
19. Kappos L, Moeri D, Radue EW, et al. **Predictive value of gadolinium-enhanced magnetic resonance imaging of relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis.** *Lancet* 1999;353:964–969
20. van Walderveen MA, Truyen L, van Oosten BW, et al. **Development of hypointense lesions on T1-weighted spin-echo magnetic resonance images in multiple sclerosis: relation to inflammatory activity.** *Arch Neurol* 1999;56:345–351
21. Prentice RL. **Surrogate endpoints in clinical trials: definition and operational criteria.** *Stat Med* 1989;8:431–440
22. Miller DH. **Multiple sclerosis: use of MR in evaluating new therapies.** *Semin Neurol* 1998;18:317–325
23. Molyneux PD, Filippi M, Barkhof F, et al. **Correlations between monthly enhanced MR lesion rate and changes in T2 lesion volume in multiple sclerosis.** *Ann Neurol* 1998;43:332–339
24. Koziol JA, Wagner S, Adams HP. **Assessing information in T2-weighted MR scans from secondary progressive MS patients.** *Neurology* 1998;51:228–233
25. Adams HP, Wagner S, Sobel DF, et al. **Hypointense and hyperintense lesions on magnetic resonance imaging in secondary-progressive MS patients.** *Eur Neurol* 1999;42:52–63