
Multiprotocol MR Image Segmentation in Multiple Sclerosis: Experience with over 1,000 Studies¹

Jayaram K. Udupa, PhD, László G. Nyúl, MSc, Yulin Ge, MD, Robert I. Grossman, MD

Rationale and Objectives. Multiple sclerosis (MS) is an acquired disease of the central nervous system. Several clinical measures are commonly used to express the severity of the disease, including the Expanded Disability Status Scale and the ambulation index. These measures are subjective and may be difficult to reproduce. The aim of this research is to investigate the possibility of developing more objective measures derived from MR imaging.

Materials and Methods. Various magnetic resonance (MR) imaging protocols are being investigated for the study of MS. Seeking to replace the Expanded Disability Status Scale and ambulation index with an objective means to assess the natural course of the disease and its response to therapy, the authors have developed multiprotocol MR image segmentation methods based on fuzzy connectedness to quantify both macroscopic features of the disease (lesions, gray matter, white matter, cerebrospinal fluid, and brain parenchyma) and the microscopic appearance of diseased white matter. Over 1,000 studies have been processed to date.

Results. By far the strongest correlations with the clinical measures were demonstrated by the magnetization transfer ratio histogram parameters obtained for the various segmented tissue regions. These findings emphasize the importance of considering the microscopic and diffuse nature of the disease in the individual tissue regions. Brain parenchymal volume also demonstrated a strong correlation with clinical measures, which suggests that brain atrophy is an important disease indicator.

Conclusion. Fuzzy connectedness is a viable, highly reproducible segmentation method for studying MS.

Key Words. Brain, MR; magnetic resonance (MR), image processing; sclerosis, multiple MR.

Multiple sclerosis (MS) is an acquired disease of the central nervous system characterized primarily by multifocal inflammation and destruction of myelin. Inflammation and edema are accompanied by different degrees of demyelination and destruction of oligodendrocytes and may be followed by remyelination, axonal loss, and/or gliosis. The disease was first characterized by Charcot (1) and has

since been investigated extensively. The highest frequency of MS occurs in northern and central Europe, Canada, the United States, New Zealand, and South Australia (2). In the United States, it affects approximately 350,000 adults and stands out as the most frequent cause of nontraumatic neurologic disability in young and middle-aged adults (3). In its advanced state, the disease may severely impair the ambulatory ability and may even cause death.

MS is usually classified into three subtypes: relapsing-remitting (RR), secondary-progressive (SP), and primary-progressive. RR disease is characterized by clearly defined disease relapses with full recovery or with sequelae and residual deficit after recovery but no disease progression between relapses. In SP MS, an initial RR disease course is followed by gradual progression, with or without occasional superimposed relapses. Primary-progres-

Acad Radiol 2001; 8:1116-1126

¹From the Medical Image Processing Group (J.K.U., Y.G.) and the Neuroradiology Division (R.I.G.), Department of Radiology, University of Pennsylvania, 423 Guardian Dr, Fourth Floor, Blockley Hall, Philadelphia, PA 19104-6021; and the Department of Applied Informatics, University of Szeged, Hungary (L.G.N.). Received May 21, 2001; revision requested June 5; revision received and accepted July 10. Supported in part by grants NS37172 and NS29029 from the National Institutes of Health. **Address correspondence to J.K.U.**

© AUR, 2001

sive MS is defined by gradual progression from onset with occasional plateaus and temporary minor improvements. The most commonly used clinical measure of disease progression is the Expanded Disability Status Scale (EDSS), introduced by Kurtzke (4). This scale extends from 0 to 10 and is based on the functional systems (visual, brain stem, pyramidal, cerebella, sensory, bowel and bladder, and cerebral) in the lower scores and on ambulation in the higher scores. The clinical quantification of disease severity is subjective and sometimes equivocal. The development of new treatments demands more precise and accurate outcome measures for relatively short trials. Therefore, magnetic resonance (MR) imaging has become one of the most important paraclinical tests for diagnosing MS and monitoring disease progression in MS. A variety of MR imaging protocols have been used to help understand this disease.

T2-weighted Imaging

The sensitivity of T2-weighted imaging in the detection of MS has long been recognized (5,6). MS lesions appear hyperintense on T2-weighted images. The clinical course of MS is mainly defined by the baseline disability of the patient, also called the *disease burden*. The extent of lesions on T2-weighted images, usually expressed as the total volume of lesions (T2LV), is currently regarded as the MR imaging measure of this disease burden. Dual-echo T2-weighted and proton-density (PD)-weighted images are usually acquired, and the lesions are segmented according to combined information from the two images. (The fluid-attenuated inversion-recovery [FLAIR] MR imaging sequence is commonly used when only visual diagnosis is to be performed, without image segmentation. Cystic lesions on PD-weighted or FLAIR images do not appear at the same signal intensity as more traditional plaques.) The positive effects of three currently available injectable medications (Betaseron, Avonex, Copaxone) on T2 lesion volume (T2LV) have been demonstrated (7–10).

An important limitation of T2-weighted imaging is its low pathologic specificity (11). All histopathologic features of MS, such as inflammation, edema, demyelination, axonal loss, gliosis, and remyelination, are seen as hyperintense lesions on T2-weighted images, but not all of these processes are clinically relevant. Reports from several research groups have shown only weak correlations between clinical measures of disability and the traditional T2LV (12–14). Results are sometimes contradictory; for example, some report that T2LV increases consistently

and markedly in RR MS (13,15), and some that it does not (12,16). Reasons for such disagreements include methodologic differences in imaging and processing, the subjectivity of the EDSS, and the inability to account for the microscopic disease. In short, there is increasing evidence that T2LV is not a robust reliable measure of disease or disability despite its use in clinical trials.

T1-weighted Imaging with Gadolinium Enhancement

In MS, gadolinium (gadopentetate dimeglumine) enhancement on MR images is characterized by transient blood-brain barrier abnormality and inflammation that represents the acute stage of the evolution of MS (17,18). On enhanced T1-weighted images, gadolinium enhancement appears as a homogeneous, strongly hyperintense lesion or as a ring-shaped area of hyperintensity at the edge of chronic reactivated lesions. Enhancement is sensitive in demonstrating disease activity and may help distinguish clinical groups (19). Delayed imaging (50–60 minutes after injection), triple doses of gadopentetate dimeglumine, and a single dose with magnetization transfer (MT) saturation to suppress normal brain have all been shown to increase the number of detectable enhancing lesions (20). (Disadvantages of high-dose contrast material-enhanced studies may include more false-positive lesions, increased imaging time, and increased expense.) These results suggest that enhancement is not an “all or none” phenomenon. Enhancement may precede the development of T2 hyperintensity and clinical symptoms, which suggest that the blood-brain barrier abnormality may be the crucial event in the inflammatory cascade (15,21). Further, we know that the disease has microscopic manifestations in normal-appearing white matter (WM), which are not apparent at visual examination of the enhanced images.

MT Imaging

MS is now believed to be a diffuse process that extensively involves the WM and is not restricted to the focal regions of disease activity visible as “lesions” on conventional T2-weighted, PD, and gadolinium-enhanced T1-weighted images. This notion came about mainly from MT image analysis (22) but is also supported by postmortem study (23) revealing microscopic disease characterized by edema, cellular infiltration, and demyelination in macroscopically normal-appearing WM. In MT imaging (24), two consecutive sets of images are required, one with off-resonance presaturation of the relatively immo-

MR Imaging Protocols Currently Used in the Authors' Center

Pulse Sequence	Plane	TR (msec)	TE (msec)	Section Thickness (mm)	Matrix	NSA	FOV (cm)
Fast SE	Axial	2,500	18, 90	3	192	1	22
3D MT Vas TOF	Axial	106	5	5	128	1	22
SE	Axial	600	27	3	192	1	22
SE (gadolinium enhanced)	Axial	600	27	3	192	1	22

Note.—FOV = field of view, NSA = number of signals acquired, SE = spin echo, TE = echo time, 3D = three-dimensional, TOF = time of flight, TR = repetition time.

bile macromolecular protons and one without. An MT ratio (MTR) image is then computed from the two images according to $(M_o - M_s)/M_o$, where M_s and M_o are pixel intensities with and without presaturation. (Since there is a time lag between the acquisitions corresponding to the “on” and “off” sequences, some artifacts due to patient movement may be introduced between the two images.) Authors of published studies (22–27) suggest that the MTR values for WM in normal subjects are highly reproducible within each institution, with a variation of less than 2%. Several studies have demonstrated that the microscopic aspect of the disease can be characterized by the MTR values measured in different regions (WM, lesions, the periphery of lesions) much more specifically than with the T2LV.

T1-weighted Imaging

Another standard MR imaging technique routinely used in MS is T1-weighted imaging (spin echo), which often shows hypointense areas that have been suggested to represent areas of axonal loss and gliosis (28). In recent quantitative studies (29,30), a significant correlation was reported between increase in disability and increase in the volume of these lesions. If hypointense lesions represent areas of severe demyelination, axonal loss, or gliosis (27,28), then the higher rate of new lesions on T2-weighted images accompanied by signal intensity loss on T1-weighted images could reflect a deficit in available repair mechanism in patients with SP disease. Greater “inflammatory” activity may exhaust repair mechanisms, leading to persistent structural loss without much remyelination, which in turn leads to persistent functional loss (31). The quantitative assessment of hypointense lesions on T1-weighted images may resolve part of the clinical-radiologic paradox encountered when one correlates MR imaging lesion load with clinical disability.

We have been developing MR image analysis methods for MS since 1993. Our approaches are guided by two

premises. First, tissue regions in any anatomic object have a heterogeneous composition, with or without disease. Treating them in a binary fashion by considering a voxel to contain either 100% or 0% tissue material is unrealistic and inaccurate. Our segmentation methods and the quantitative measures derived from them take into account the heterogeneity and the fuzziness inherent in tissue regions. Second, as we have described, different MR imaging protocols capture different aspects of the MS disease puzzle. We have therefore taken a multiprotocol approach by using all major protocols current in MS imaging to obtain a composite understanding of the most specific way to characterize MS with image-derived parameters.

MATERIALS AND METHODS

Data Acquisition

The MR imaging protocols we have used are listed in the Table. Every patient recruited in our trial undergoes all acquisitions listed. Our image database currently consists of the following number of patient studies and three-dimensional volume images that have been processed by the methods described in this section: 690 T2-weighted imaging studies; 660 T1-weighted studies, each with and without contrast enhancement; and 670 MT studies. Altogether, the study included 100 patients and over 4,000 three-dimensional scenes. The T2-weighted and PD images in these acquisitions are in registration, since they are acquired simultaneously. Similarly, the two MT images are also in registration. Between these two sets and among other image data, however, registration cannot be guaranteed.

Fuzzy Connectedness Image Segmentation

The method of fuzzy connectedness forms the essential underpinning of all segmentation algorithms used in our MS image analysis efforts. Therefore, we will first outline its principles. These principles are applicable to n -dimen-

sional vector-valued images, but our description will be confined to the three-dimensional case and to scalar-valued images.

We represent a volume image, called a three-dimensional scene (*scene* for short) C by a pair $C = (C, f)$ where C is a rectangular array of voxels, called the *scene domain*, and f is a function that assigns to every voxel $c \in C$ an *intensity* value $f(c)$ from a range $[L, H]$ where L and H are integers. Objects such as WM in the brain are manifest in scenes with a heterogeneity of property values (such as PD) because of the heterogeneity of material composition inherent in the object and the noise, blurring, and background variations introduced by the imaging device. Despite this *graded composition* of intensity values within object regions, human readers perceive regions in the scene belonging to the same object as a whole (*gestalt*). This *hanging-togetherness* of image elements and their graded composition are both fuzzy phenomena, which should be addressed properly for effective image segmentation (32).

We think of any nearby voxels c and d in C as having a *fuzzy adjacency*. The strength $\mu_\alpha(c, d)$ of this fuzzy relation α is smaller the farther c and d are. The idea behind α is to capture the blurring property of imaging devices. We define another fuzzy relation κ in C that assigns to every pair of voxels (c, d) an *affinity*. The strength $\mu_\kappa(c, d)$ of this relation depends on the spatial proximity of c and d (ie, on $\mu_\alpha(c, d)$) and on the similarities between the scene intensities $f(c)$ and $f(d)$ and between the intensity-based properties computed at c and d . The intent here is that κ is "local"; that is, if c and d are far apart (spatially), then their affinity is 0. To define *fuzzy connectedness* K as another fuzzy relation in C , we consider all possible "paths" between all possible pairs of voxels in C . A path between any voxels c and d is simply a sequence of nearby voxels starting from c and ending in d . To every path from c to d we assign a *strength of connectedness*, which is simply the smallest affinity of pairwise voxels along the path (weakest link). Finally, the strength of connectedness $\mu_K(c, d)$ between c and d is the largest of the strengths of connectedness of all paths between c and d . In determining a fuzzy connected object in C , $\mu_K(c, d)$ should be determined for all possible pairs (c, d) of voxels in C . This computationally explosive problem is considerably simplified through some key theoretical results, and finding fuzzy connected objects essentially reduces to dynamic programming (32).

The definition of affinity (33) is a key to the effectiveness of segmentation in this method. The strength of af-

finity $\mu_\kappa(c, d)$ between two voxels c and d depends on the intensities of all voxels within the largest spherical homogeneous region around c and d . The radius of this sphere represents local object size at c and d . Since this radius changes over the scene domain, the actual functional form of the affinity varies over the scene domain, adapting to the local object size and noise characteristics. We have shown that slow background variations do not affect fuzzy connected object segmentations (32–34).

In the rest of this section, we summarize our approach to analyzing the scenes for the different protocols. We think of segmentation as consisting of two related tasks: recognition and delineation. *Recognition* is the high-level process of roughly determining the whereabouts of the object in the scene. *Delineation* is the low-level process of determining the precise spatial extent and voxel-by-voxel material percentage content of the object. Knowledgeable humans can outperform computer algorithms in most recognition tasks, such as those encountered in our application. Conversely, computer algorithms can be devised to delineate objects in scenes more precisely, accurately, and efficiently than manual delineation. With manual delineation it is clearly impossible to indicate the graded material composition within an object. The system we have developed exploits this synergy between human and computer abilities in devising practical solutions that can be used routinely in clinical trials.

Analysis of Fast Spin-Echo T2- and PD-weighted Images

The approach here consists of the following four steps (35). We consider the scene to be vector valued with two values per voxel (the T2- and PD-weighted values).

1. On one section, roughly centrally situated in the brain, an operator specifies a few voxels (seeds). This is a recognition step that uses the superior human knowledge. Points (voxels) are specified for the cerebrospinal fluid (CSF), gray matter (GM), and WM regions but not for lesions.

2. This is a delineation step. The seed voxels are used to determine the fuzzy connected objects that contain them. This results first in a segmentation of WM, GM, and CSF. This knowledge is subsequently used to determine automatically a set of points (determined as holes in the union of GM and WM fuzzy objects) in each three-dimensional lesion object that are subsequently used to delineate the lesions, each as a three-dimensional fuzzy connected object (Fig 1).

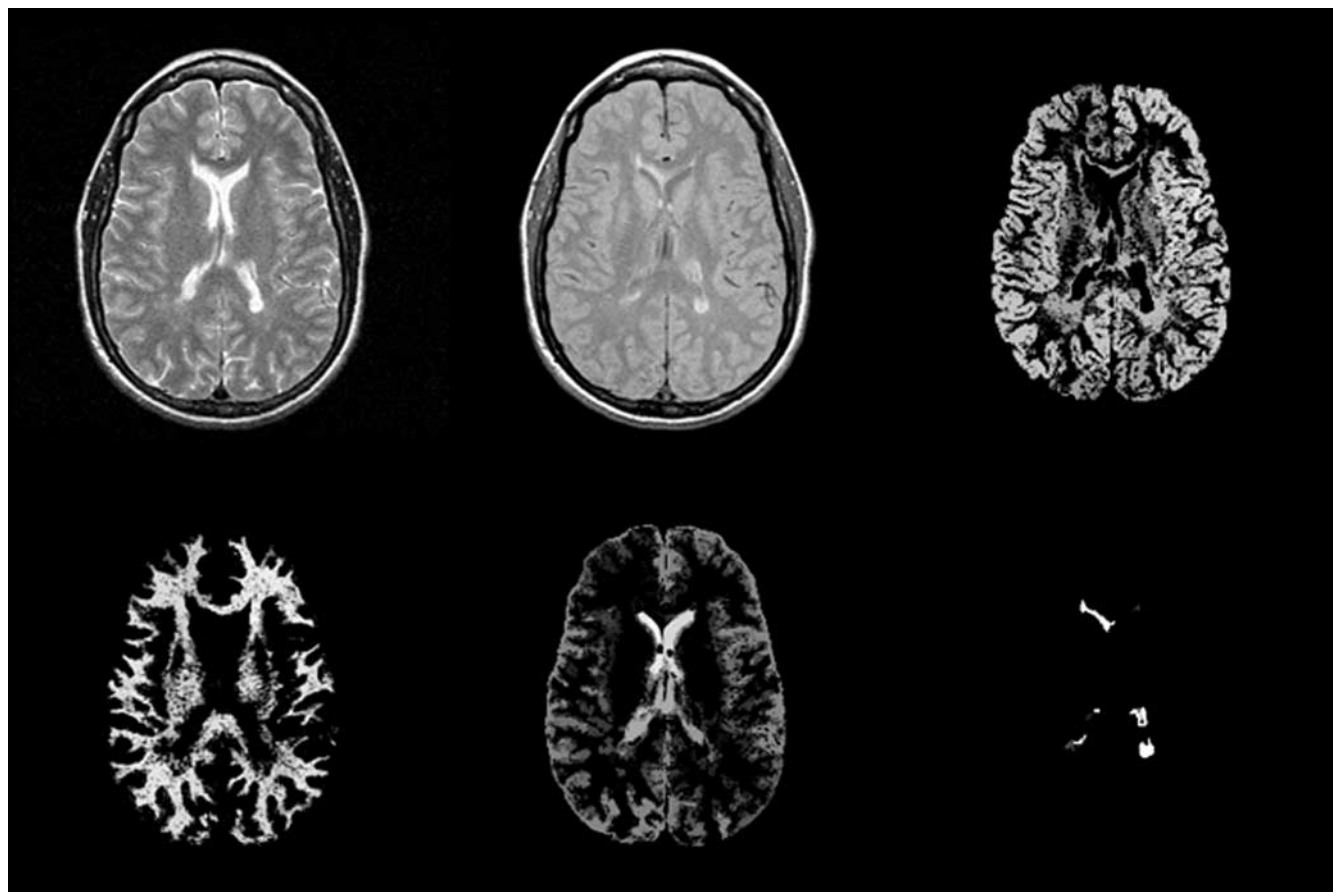


Figure 1. One section of the T2- and PD-weighted images from a patient with MS and the segmented GM, WM, CSF, and lesion objects. The segmentation is done in three dimensions, although the results are shown for one section.

3. This is another recognition step requiring help from a human operator. The operator accepts true lesions and rejects false ones with the click of a mouse button. Each three-dimensional lesion is displayed on one section image passing close to the centroid of the lesion. The operator may override this mode of display and examine the lesions on all or any selected sections. All false-positive findings—usually artifacts and choroid plexus—are eliminated in this way. Although this step is subjective, the number of false-negative findings is very low in our system (see next section for details on validation). Nevertheless, in this step the system allows new seed points to be selected for missed lesions. These are subsequently used in delineating these lesions by repeating part of step 2 above.

4. The final step is the computation of quantitative measures from the segmented objects. These include the number of three-dimensional lesions and their total volume (T2LV), the volume of CSF (CSFV), the volume of the brain parenchyma (BPV) which is the volume of the

union of GM and WM, and normalized BPV, $nBPV = BPV / (BPV + CSFV)$. The purpose of the last measure is to express the BPV independent of subject-to-subject variations in brain size. Thus, $nBPV$ is also computed from the dual-echo T2- and PD-weighted images.

Analysis of Gadolinium-enhanced T1-weighted Images

This analysis (36) includes four steps:

1. A conservative threshold is determined automatically from the histogram of the three-dimensional scene representing gadolinium-enhanced T1-weighted images (36). The purpose of this threshold is to select seed points automatically within the enhancing lesions. This is a recognition step.

2. The seed points are used to determine the fuzzy connected objects that contain them. This delineation step often results in the delineation of vessels. With the volume enclosed by the object used as a criterion, large objects (large vessels) are automatically discarded.

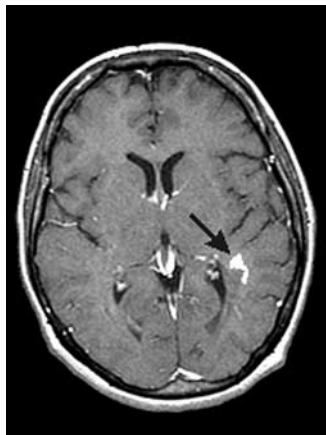


Figure 2. One section of the gadolinium-enhanced T1-weighted images from a patient with MS and the segmented enhancing lesion (arrow).

3. This step is identical to step 3 of the previous section.
4. For the accepted lesions, their total volume and number are computed (Fig 2).

Analysis of MT Images

The steps in this process (37) are as follows:

1. The brain is first segmented from the two MT scenes (with the off-resonance pulse on and off). The brain mask is then used to compute an MTR scene $C_{MTR}^{BP} = (C_{MTR}^{BP}, f_{MTR}^{BP})$, which is such that $f_{MTR}^{BP}(c) = 0$ if c is not in the brain parenchymal region, otherwise $f_{MTR}^{BP}(c)$ will have an MTR value computed from the two MT scenes, as described in the above section on MT imaging.
2. The PD scene is registered with the MT scene obtained without the off-resonance pulse (38). Subsequently, the PD and T2 scenes are redigitized to match with the MT scenes section to section. The redigitized scenes are segmented for WM and GM, as described in the section titled "Analysis of Fast Spin-Echo T2- and PD-weighted Images." The MTR scenes C_{MTR}^{GM} and C_{MTR}^{WM} are computed subsequently.
3. For each of the MTR scenes C_{MTR}^{BP} , C_{MTR}^{WM} , and C_{MTR}^{GM} , a histogram of the whole scene normalized to BPV is computed. From each of these histograms, the following parameters are computed: peak height normalized by BPV (nPH_x), the mode M_x , the 25th and 50th percentiles ($P25_x$, $P50_x$), and the mean m_x , where x stands for BP, WM, or GM (Fig 3).

RESULTS

Validation

Any segmentation effort consists of (a) a theoretical or algorithmic framework, (b) considerable engineering ef-

fort to make the framework work in the application at hand, and (c) evaluation to establish the precision, accuracy, and efficiency of the method for the particular application. *Precision* here refers to the reproducibility of the segmentation results with all subjective actions taken into consideration, including how the patient is positioned in the imager and any operator input required by the algorithms. *Accuracy* denotes the degree of agreement of the result with truth. *Efficiency* indicates the degree of operator help required (the degree of automation), which may be expressed as the operator time required for each study. A compromise in one factor is usually required to achieve an improvement in another. Below we summarize the precision, accuracy, and efficiency of the various analyses described in the previous section (35–37,39). We have evaluated our system extensively, taking into account the subjectivity that appears in steps that require operator input and in the placement of the patient in the imager.

For T2LV, the inter- and intraoperator coefficient of variation was less than 0.9%, the variations among repeated images was approximately 1.5%, the 95% confidence interval for false-negative T2LV was 0%–2.8%, and the mean operator time was 10 minutes per study (Sparc 20 Sun workstation, Mountain View, Calif, reduced to about 7 minutes with a Sun Ultra Sparc workstation). For total volume of lesions on enhanced T1-weighted images (T1ELV), there were no inter- and intraoperator variations, the false-negative rate was 1.3%, and the mean operator time was 2 minutes. For BPV, the inter- and intraoperator coefficient of variation was less than 0.5%, including repeated images, the false-negative rate was less than 1%, and the mean operator time was 5 minutes. For MTR, there was more than 99% correlation for MTR histogram parameters in repeated experiments, and the mean operator time was 10 minutes.

Any segmentation method can and will go wrong in practice. In completely automatic methods, there is the question of how to recognize and delineate false-positive and false-negative findings. By design, our system is not automatic, to facilitate taking recognition help from the user and to minimize false-positive and false-negative findings. This help is obtained as efficaciously as possible, to make the system efficient and practical. We continue to improve the efficiency of our system without sacrificing precision and accuracy.

Clinical Correlation

We have obtained the following results by analyzing the images in our database and correlating those findings

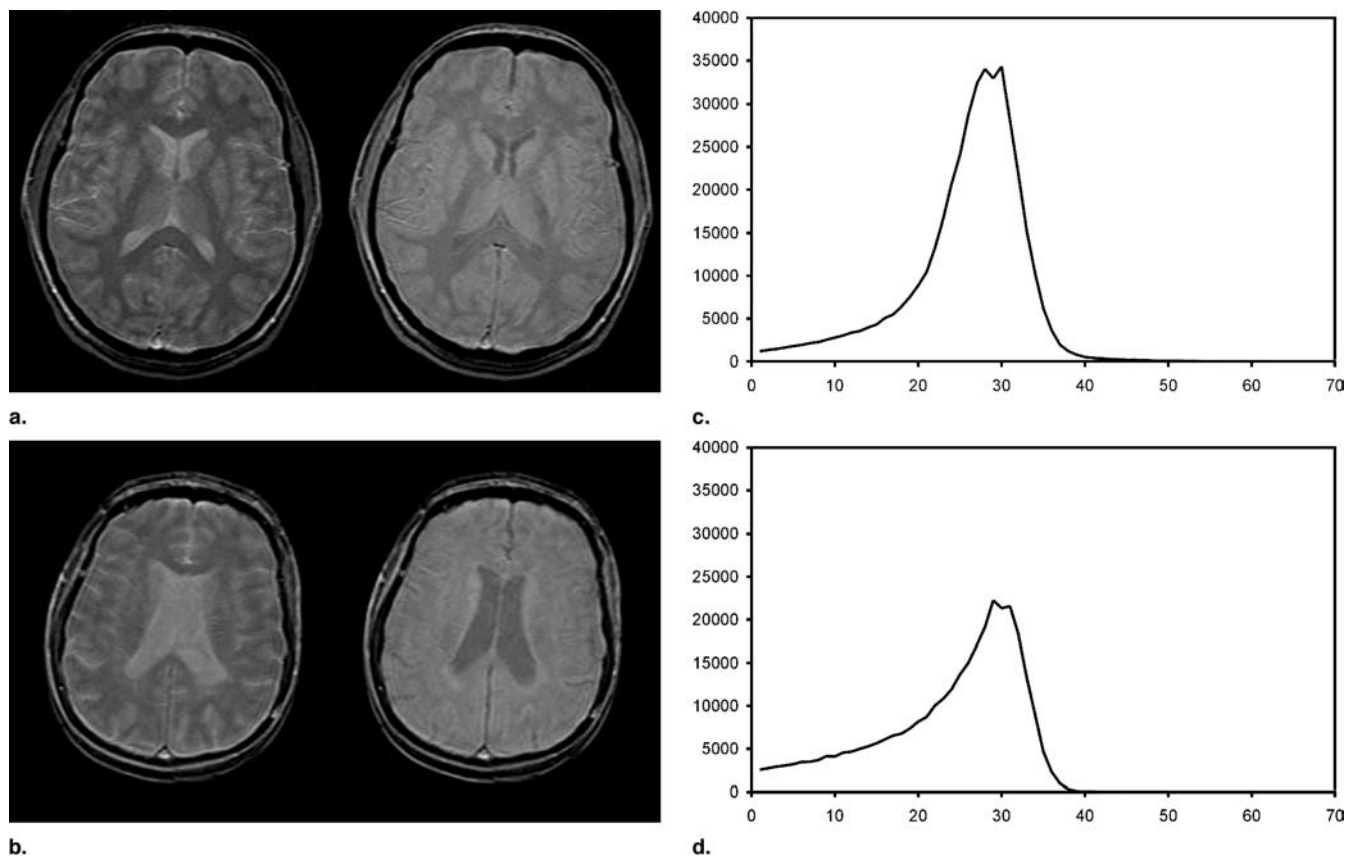


Figure 3. (a, b) One section of the original MT images off-resonance pulse on (left) and off (right) and (c, d) the MTR distributions for a normal subject (a, c) and a patient with MS (b, d).

with the EDSS scores. For T2-weighted imaging (39–45), the strongest correlations were seen in the RR group, for T2LV with nBPV (-0.73 , $P < .001$), nPH_{BPV} (-0.76 , $P < .001$), and T1ELV (0.66 , $P < .001$) both in the RR group. T2LV also showed a good correlation with nBPV (-0.66 , $P < .02$) in the patients with SP disease. In both groups, the correlation of T2LV with EDSS scores was not significant. Generally, we found CSFV to be significantly larger in MS patients and nBPV to be smaller than in age-matched normal control subjects ($P < .005$). A good correlation was seen between T2LV and CSFV (0.73 , $P < .001$).

In comparing patients with and without U-fiber lesions, we found a significant ($P \leq .05$) difference in three of the 11 neuropsychological test scores, which suggests that these lesions may lead to neuropsychological impairment. By analyzing the lesions in GM, we found that cortical GM lesions comprised about 5.7% of the total T2LV, while deep GM lesions comprised about 4.6%. No significant correlation was found between these volumes and clinical disability. The parameter computed from T2- and

PD-weighted scenes that had by far the strongest correlation with disability was nBPV (-0.69 , $P < .004$) for the patients with SP disease, which suggests that brain atrophy may be important for characterizing disease with T2-weighted imaging.

We also analyzed clinical correlation with gadolinium-enhanced T1-weighted imaging (39,41,42,46,47). We generally found that T1ELV was greater in RR than in SP disease ($P < .05$). In the RR group only, T1ELV showed good correlation with nBPV (-0.51 , $P < .009$) and with T2LV. Compared with RR patients receiving placebo, a group of RR patients receiving Copaxone had a significant decrease in the number ($P = .05$) and volume ($P = .01$) of enhancing lesions at the end of 24 months (47). There were no differences in T2LV between the two groups. The loss of brain tissue was significantly less ($P = .02$) in the group receiving Copaxone than in the placebo group.

With MT imaging (41–43,48,49), the nPH_{BP} for the MS study group (both RR and SP disease) was significantly lower than in the control group ($P < .0005$), as

were M_{BP} ($P < .05$), $P25_{BP}$ ($P < .05$), and $P50_{BP}$ ($P < .05$). In particular, nPH_{BP} showed the best correlation with EDSS score for the RR group (-0.44 , $P < .02$). This parameter correlated strongly with $nBPV$ in both the RR (0.69 , $P < .001$) and SP (0.87 , $P < .001$) groups and with CSFV (-0.828 , $P < .0001$) when the two disease groups were combined. MTR parameters also correlated with disease duration ($P < .01$).

Individual neuropsychological test scores correlated with MTR measures ($P < .001$). The unnormalized histogram peak height differed ($P < .05$) among severely impaired, moderately impaired, and normal patients. In a serial MR imaging study of the RR group, nPH_{BP} demonstrated a subtle but significant ($P < .05$) decline over time, but no significant changes in EDSS were noted. M_{WM} and m_{WM} were significantly lower ($P < .05$) in patients with RR MS than in normal controls (50). In longitudinal studies, both M_{WM} and m_{WM} shifted in the direction of normal with increasing disease duration. Up to 44% of new lesions identified on later studies were demonstrated to have originated in the WM region, which had been identified as abnormal on the basis of MTR criteria. As to the GM, M_{GM} and m_{GM} were significantly lower ($P < .01$) in patients with RR MS than in normal controls (51), and nPH_{GM} was inversely correlated with EDSS ($r = 0.65$, $P < .01$).

In summary, in our clinical trials the strongest correlations with disease status at EDSS were observed with $nBPV$ and the MTR parameters for the various tissue components. These findings emphasize the need to consider $nBPV$ and microscopic MS disease in devising objective measures for disease assessment.

DISCUSSION

Automation, Failure, and User Assistance

Any segmentation method can go wrong, and therefore will, if enough studies are analyzed in a routine clinical trial. Therefore, for quality assurance, it is important to have a knowledgeable human operator within the processing loop. Complete automation may therefore be an elusive goal, perhaps reachable after only a great deal of experience within the same imaging modality, protocol, and application setting. As researchers and developers, our aim should be to consider human interaction within the design framework of the segmentation algorithms, with a research goal of minimizing this interaction as much as possible. The system that we presented here may not yet be optimal in terms of the degree of automation,

but it is certainly practical. We continue to tighten this slack in efficiency without sacrificing precision and accuracy.

Standardization of MR Intensity Scale

A major difficulty with the MR imaging techniques for most protocols has been that image intensities do not have a fixed meaning, even within the same protocol, for the same body region, for images obtained with the same imager, and for the same patient (Fig 4). This poses problems in finding the proper window setting for image display and, more important, in image segmentation and analysis. Any segmentation method has parameters. Without the same MR protocol-specific tissue indication for intensity values, it becomes very difficult to set values for the parameters in a manner independent of patient studies, and precision, accuracy, and the efficiency of segmentation may be compromised. For example, it may be necessary to adjust parameters for each study to handle imager-dependent variations in intensity, which will affect efficiency and may also lower precision.

We have recently developed a method (52,53) to standardize the MR image intensity scale. It is a postacquisition processing method that maps nonlinearly the scene intensities in any given scene acquired according to a given protocol for a given body region, standardizing them to a scale so that the same intensities in the transformed scene will have the same tissue-specific meaning. The method is based on deforming histograms so that they are as similar as possible for all scenes of the same body region and protocol (Fig 4). We are using this method in modifying our system, to improve the efficiency of the various quantification processes without sacrificing precision and accuracy.

Standard Intensity-based Analysis

All objects—normal and pathologic—have a heterogeneous tissue composition. Combined with the blurring and noise introduced by the imaging device, this composition makes tissue regions heterogeneous in intensity. We believe that this heterogeneity has tissue-specific information and is useful in characterizing disease stage and severity. Such a characterization becomes impossible when there is imager-dependent intensity variation. As illustrated with MTR analysis (22–27,48) because of the tissue-specific meaning of MTR values, standardization may permit us to treat other protocols (T2-, PD-, and T1-weighted imaging, with and without gadolinium enhancement) in the same tissue-specific way as MT imaging. More important, flat measures such as volume of tissue

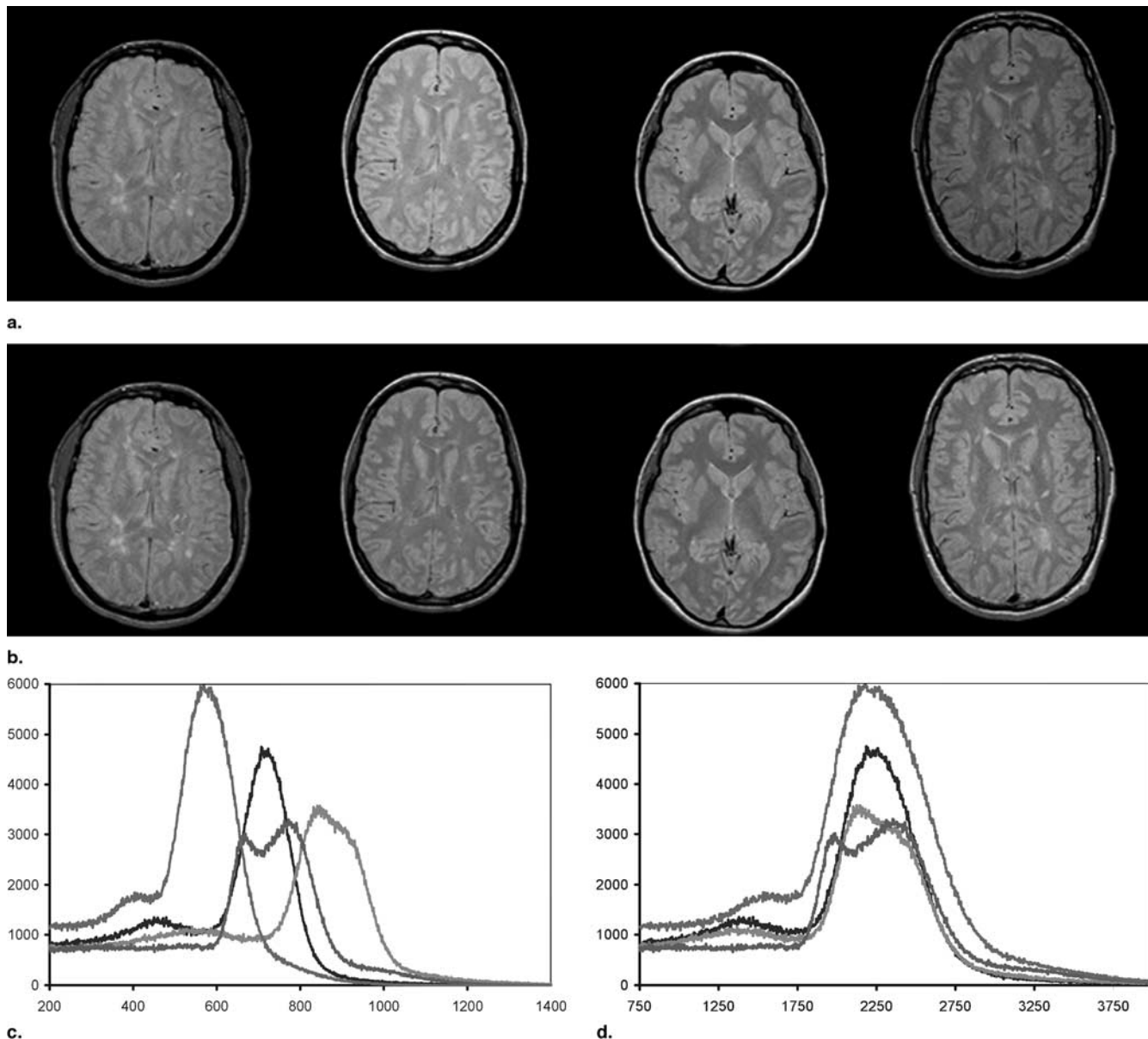


Figure 4. (a) Four sections from four PD-weighted data sets in four MS patients and (c) the corresponding histograms computed over each of these entire three-dimensional scenes. (b) Corresponding sections are shown after standardization, with (d) their histograms.

regions ignore heterogeneity and may lose important disease-specific information. Volume distributions—that is, intensity histograms within segmented tissue regions, as demonstrated by MTR analysis—should be considered as a means for understanding subtle disease processes.

MS Segmentation Workshop

In MS (and other neurologic applications), a variety of MR imaging protocols are used. The imaging parameters used in these protocols vary among institutions. Despite

the numerous brain MR image segmentation methods developed during the past 15 years, none of them is capable of handling variations within the same protocol, much less the variations among protocols. We need a segmentation “workshop” wherein a protocol-specific segmentation method can be quickly fabricated. For the MS application, we believe that the fuzzy connectedness framework combined with MR imaging intensity scale standardization can be used to build such a workshop, and we are working toward this goal.

ACKNOWLEDGMENT

The authors are grateful to Mary A. Blue for typing the manuscript.

REFERENCES

1. Charcot JM. Lectures on the diseases of the nervous system. London, England: New Sydenham Society, 1977.
2. Kurtzke JF. Geographic distribution of multiple sclerosis: an update with special reference to Europe and the Mediterranean region. *Acta Neurol Scand* 1980; 62:65-80.
3. Hauser SL. Multiple sclerosis and other demyelinating disease. In: Is-selbacher KJ, Braunwald E, Wilson JD, Martin JB, Fauci AS, Kasper DL, eds. *Harrison's principle of internal medicine*. New York, NY: McGraw-Hill, 1994; 2287-2295.
4. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983; 33:1444-1452.
5. Lukes SA, Crooks LE, Aminoff MJ, et al. Nuclear magnetic resonance imaging in multiple sclerosis. *Ann Neurol* 1983; 13:592-601.
6. Sheldon JJ, Siddharthan R, Tobias J, et al. MR imaging of multiple sclerosis: comparison with clinical and CT examinations in 74 patients. *AJR Am J Roentgenol* 1985; 145:957-964.
7. Jacobs LD, Cookfair DL, Rudick RA, et al. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Ann Neurol* 1996; 39:285-294.
8. Rudick RA, Goodkin DE, Jacobs LD, et al. Impact of interferon beta-1a on neurologic disability in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG). *Neurology* 1997; 49:358-363.
9. Johnson KP, Brooks BR, Cohen JA, et al. Extended use of glatiramer acetate (Copaxone) is well tolerated and maintains its clinical effect on multiple sclerosis relapse rate and degree of disability. Copolymer 1 Multiple Sclerosis Study Group. *Neurology* 1998; 50:701-708.
10. Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. The IFNB Multiple Sclerosis Study Group and the University of British Columbia MS/MRI Analysis Group. *Neurology* 1995; 45:1277-1285.
11. Miller DH, Grossman RI, Reingold SC, et al. The role of magnetic resonance techniques in understanding and managing multiple sclerosis. *Brain* 1998; 121:3-24.
12. Paty DW, Li DK. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. II. MRI analysis results of a multicenter, randomized, double-blind, placebo-controlled trial. UBC MS/MRI Study Group and the IFNB Multiple Sclerosis Study Group. *Neurology* 1993; 43: 662-667.
13. Miki Y, Grossman RI, Udupa JK, et al. Relapsing-remitting multiple sclerosis: longitudinal analysis of MR images—lack of correlation between changes in T2 lesion volume and clinical findings. *Radiology* 1999; 213:395-399.
14. Thompson AJ, Kermod AG, MacManus DG, et al. Patterns of disease activity in multiple sclerosis: clinical and magnetic resonance imaging study. *Br Med J* 1990; 300:631-634.
15. Kermod AG, Thompson AJ, Tofts P, et al. Breakdown of the blood-brain barrier precedes symptoms and other MRI signs of new lesions in multiple sclerosis: pathogenetic and clinical implications. *Brain* 1990; 113:1477-1489.
16. Filippi M, Horsfield MA, Morrissey SP, et al. Quantitative brain MRI lesion load predicts the course of clinically isolated syndromes suggestive of multiple sclerosis. *Neurology* 1994; 44:635-641.
17. Grossman RI, Gonzalez-Scarano F, Atlas SW, et al. Multiple sclerosis: gadolinium enhancement in MR imaging. *Radiology* 1986; 161:721-725.
18. Kermod AG, Tofts PS, Thompson AJ, et al. Heterogeneity of blood-brain barrier changes in multiple sclerosis: an MRI study with gadolinium-DTPA enhancement. *Neurology* 1990; 40:229-235.
19. Miller DH, Barkhof F, Nauta JJ. Gadolinium enhancement increases the sensitivity of MRI in detecting disease activity in multiple sclerosis. *Brain* 1993; 116:1077-1094.
20. Filippi M, Yousry T, Campi A, et al. Comparison of triple dose versus standard dose gadolinium-DTPA for detection of MRI enhancing lesions in patients with MS. *Neurology* 1996; 46:379-384.
21. Barratt HJ, Miller D, Rudge P. The site of the lesion causing deafness in multiple sclerosis. *Scand Audiol* 1988; 17:67-71.
22. Dousset V, Grossman RI, Ramer KN, et al. Experimental allergic encephalomyelitis and multiple sclerosis: lesion characterization with magnetization transfer imaging. *Radiology* 1992; 182:483-491.
23. Allen IV, McKeown SR. A histological, histochemical and biochemical study of the macroscopically normal white matter in multiple sclerosis. *J Neurol Sci* 1979; 41:81-91.
24. Wolff SD, Balaban RS. Magnetization transfer contrast (MTC) and tissue water proton relaxation in vivo. *Magn Reson Med* 1989; 10:135-144.
25. Filippi M, Campi A, Dousset V, et al. A magnetization transfer imaging study of normal-appearing white matter in multiple sclerosis. *Neurology* 1995; 45:478-482.
26. Loevner LA, Grossman RI, Cohen JA, et al. Microscopic disease in normal-appearing white matter on conventional MR images in patients with multiple sclerosis: assessment with magnetization-transfer measurements. *Radiology* 1995; 196:511-515.
27. Gass A, Barker GJ, Kidd D, et al. Correlation of magnetization transfer ratio with clinical disability in multiple sclerosis. *Ann Neurol* 1994; 36: 62-67.
28. Uhlenbrock D, Sehlen S. The value of T1-weighted images in the differentiation between MS, white matter lesions, and subcortical arterio-sclerotic encephalopathy (SAE). *Neuroradiology* 1989; 31:203-212.
29. van Walderveen MA, Barkhof F, Hommes OR, et al. Correlating MRI and clinical disease activity in multiple sclerosis: relevance of hypointense lesions on short-TR/short-TE (T1-weighted) spin-echo images. *Neurology* 1995; 45:1684-1690.
30. Truyen L, van Waesberghe JH, van Walderveen MA, et al. Accumulation of hypointense lesions ("black holes") on T1 spin-echo MRI correlates with disease progression in multiple sclerosis. *Neurology* 1996; 47:1469-1476.
31. Prineas JW, Barnard RO, Revesz T, et al. Multiple sclerosis: pathology of recurrent lesions. *Brain* 1993; 116:681-693.
32. Udupa JK, Samarasekera S. Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. *Graph Models Image Proc* 1996; 58:246-261.
33. Saha PK, Udupa JK. Scale-based fuzzy connectivity: a novel image segmentation methodology and its validation. In: Hanson KM, ed. *Medical imaging 1999: image processing*. Proc SPIE 1999; 3661:249-257.
34. Saha PK, Udupa JK, Odhner D. Scale-based fuzzy connected image segmentation: theory, algorithms, and validation. *Comput Vis Image Understanding* 2000; 77:145-174.
35. Udupa JK, Wei L, Samarasekera S, et al. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Trans Med Imaging* 1997; 16:598-609.
36. Samarasekera S, Udupa JK, Miki Y, et al. A new computer-assisted method for the quantification of enhancing lesions in multiple sclerosis. *J Comput Assist Tomogr* 1997; 21:145-151.
37. van Buchem MA, Udupa JK, McGowan JC, et al. Global volumetric estimation of disease burden in multiple sclerosis based on magnetization transfer imaging. *AJNR Am J Neuroradiol* 1997; 18:1287-1290.
38. Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 1993; 17:536-546.
39. Ge Y, Grossman RI, Udupa JK, et al. Brain atrophy in relapsing-remitting multiple sclerosis and secondary progressive multiple sclerosis: longitudinal quantitative analysis. *Radiology* 2000; 214:665-670.
40. Miki Y, Grossman RI, Udupa JK, et al. Isolated U-fiber involvement in MS: preliminary observations. *Neurology* 1998; 50:1301-1306.
41. Miki Y, Grossman RI, Udupa JK, et al. Differences between relapsing-remitting and chronic progressive multiple sclerosis as determined with quantitative MR imaging. *Radiology* 1999; 210:769-774.
42. Miki Y, Grossman RI, Udupa JK, et al. Relapsing-remitting multiple sclerosis: longitudinal analysis of MR images—lack of correlation be-

- tween changes in T2 lesion volume and clinical findings. *Radiology* 1999; 213:395-399.
43. Phillips MD, Grossman RI, Miki Y, et al. Comparison of T2 lesion volume and magnetization transfer ratio histogram analysis and of atrophy and measures of lesion burden in patients with multiple sclerosis. *AJNR Am J Neuroradiol* 1998; 19:1055-1060.
 44. Catalaa I, Fulton JC, Zhang X, et al. MR imaging quantitation of gray matter involvement in multiple sclerosis and its correlation with disability measures and neurocognitive testing. *AJNR Am J Neuroradiol* 1999; 20:1613-1618.
 45. Fulton JC, Grossman RI, Udupa JK, et al. MR lesion load and cognitive function in patients with relapsing-remitting multiple sclerosis. *AJNR Am J Neuroradiol* 1999; 20:1951-1955.
 46. Miki Y, Grossman RI, Udupa JK, et al. Computer-assisted quantitation of enhancing lesions in multiple sclerosis: correlation with clinical classification. *AJNR Am J Neuroradiol* 1997; 18:705-710.
 47. Ge Y, Grossman RI, Udupa JK, et al. Glatiramer acetate (Copaxone) treatment in relapsing-remitting MS: quantitative MR assessment. *Neurology* 2000; 54:813-817.
 48. van Buchem MA, Grossman RI, Armstrong C, et al. Correlation of volumetric magnetization transfer imaging with clinical data in MS. *Neurology* 1998; 50:1609-1617.
 49. Patel UJ, Grossman RI, Phillips MD, et al. Serial analysis of magnetization-transfer histograms and Expanded Disability Status Scale scores in patients with relapsing-remitting multiple sclerosis. *AJNR Am J Neuroradiol* 1999; 20:1946-1950.
 50. Catalaa I, Grossman RI, Kolson DL, et al. Multiple sclerosis: magnetization transfer histogram analysis of segmented normal-appearing white matter. *Radiology* 2000; 216:351-355.
 51. Ge Y, Grossman RI, Udupa JK, et al. Magnetization transfer ratio histogram analysis of gray matter in relapsing-remitting multiple sclerosis. *AJNR Am J Neuroradiol* 2001; 22:470-475.
 52. Nyul LG, Udupa JK. On standardizing the MR image intensity scale. *Magn Reson Med* 1999; 42:1072-1081.
 53. Nyul LG, Udupa JK, Zhang X. New variants of a method of MRI scale standardization. *IEEE Trans Med Imaging* 2000; 19:143-150.

ANNOUNCEMENT

An NIH State-of-the-Science Conference, **The Management of the Clinically Inapparent Adrenal Mass (“Incidentaloma”)**, will be held February 4-6, 2002, at the Natcher Conference Center, National Institutes of Health, Bethesda, Md. The conference is presented by the National Institute of Child Health and Human Development, the National Cancer Institute, and the NIH Office of Medical Applications of Research.

This NIH State-of-the-Science Conference has been convened to examine the current state of knowledge regarding the management of the clinically inapparent adrenal mass. Commonly known as adrenal incidentalomas, these are tumors that have been discovered inadvertently, in the course of diagnostic testing or treatment for other conditions. The conference will explore the following key questions: What are the causes, prevalence, and natural history of clinically inapparent adrenal masses? On the basis of available scientific evidence, what is the appropriate evaluation of a clinically inapparent adrenal mass? What criteria should guide the decision on surgical versus nonsurgical management of these masses? If surgery is indicated, what is the appropriate procedure? What is the appropriate follow-up for patients for each management approach? What additional research is needed to guide practice?

During the first day and a half of the conference, experts will present the adrenal incidentalomas findings to an independent non-federal panel. The panel will weigh all of the scientific evidence, then present its draft statement to the audience on the final day of the conference. Melvin Grumbach, MD, Edward B. Shaw Professor of Pediatrics at the University of California, San Francisco, will chair the panel.

To register for this conference or to obtain further information, please visit the NIH Consensus Development Program Web site, <http://consensus.nih.gov>, or contact Channet Williams, Prospect Associates, 10720 Columbia Pike, Silver Spring, MD 20901, phone: 301-592-2130, fax: 301-593-5791, e-mail: adrenalmass@prospectassoc.com.