Title
Methods for Measuring Brain Morphologic Features on Magnetic Resonance Images: Validation and Normal Aging

Permalink
https://escholarship.org/uc/item/1ww159vp

Journal
Archives of Neurology, 47(1)

ISSN
0003-9942

Authors
Jernigan, TL
Press, GA
Hesselink, JR

Publication Date
1990

DOI
10.1001/archneur.1990.00530010035015

Peer reviewed
Methods for Measuring Brain Morphologic Features on Magnetic Resonance Images

Validation and Normal Aging

Terry L. Jernigan, PhD; Gary A. Press, MD; John R. Hesselink, MD

In this article, methods are described for measuring brain morphologic features, using computer-analyzed magnetic resonance (MR) images. The aim in developing these methods was to provide a means by which degeneration in brain structures could be quantified during life in patients with dementia, dysphasia, psychiatric disorders, and the acquired immunodeficiency syndrome.

The protocol used produces a set of images within which each pixel has been classified as resembling gray matter, white matter, cerebrospinal fluid (CSF), or signal hyperintensities. From these images, estimates are computed of the volumes of the different pixel classes within certain cerebral zones of interest. The procedures require considerable subjective input from the operator, who is "blind" to patient identification, but the interoperator reliability has been estimated to be high.

Neuropathologic examination of autopsy material from patients whose in vivo MR images have been analyzed is underway. However, too few brain specimens have now become available to permit sound inferences.

In this report we describe morphologic changes across the age range observed with our methods in a group of 58 normal volunteer subjects. These changes will be compared with those noted in earlier neuropathology and brain imaging studies.

AGE-RELATED CHANGES IN BRAIN MORPHOLOGIC FEATURES

Numerous neuropathologic studies have shown an age-related decrease in brain weight and brain volume in nondemented subjects. Age-related decreases in cortical volume and cortical neuronal populations have been suggested, particularly in superior temporal and frontal regions.

Most cross-sectional computed tomography (CT) studies of cerebral atrophy have demonstrated that age-related changes in ventricular and sulcal sizes increase slowly until about age 60 years and increase more rapidly thereafter. In several of these investigations dramatic increases in the variability were observed in older age groups. Thus, perhaps it is not surprising that in some studies, especially those in which the age range of subjects was restricted, no significant age increases were observed in ventricular size. Gado et al., who used volumetric methods for measuring CSF, were able to demonstrate increasing brain atrophy on CT with 1-year longitudinal follow-ups of normal elderly subjects. Bird et al., however, did not observe increasing atrophy on follow-up scanning of elderly subjects, although a subgroup of the sample, who declined on cognitive testing, did show increased ventricular size.

In an early study with CT, periventricular CT values in an area of white matter were found to decline in normal aging. More recently, diffuse periventricular lucency on CT, called leukoaraiosis, has been better defined, identified in nonsymptomatic elderly subjects, and linked to soft neurologic signs and vascular risk factors in this population.

Magnetic resonance imaging studies show that signal hyperintensities occur very frequently in normal elderly people, although they are rare in asymptomatic young people. Magnetic resonance imaging may be more sensitive to the white matter changes occurring in the elderly, and these abnormalities may be related to subclinical functional impairment. This study evaluates the sensitivity of our methods for detecting and measuring the age-related central nervous system changes reported in these earlier investigations.

SUBJECTS, MATERIALS, AND METHODS

The subjects were 58 normal volunteers recruited to participate in several neurobehavioral investigations. Twenty-three were female and 35 were male. All were screened by detailed medical and psychiatric interviews for evidence of significant disease (ie, diabetes, heart disease, substance abuse, developmental intellectual abnormality, or psychiatric illness. All adult subjects were living independently in the community; all were students, employed, or retired from employment. The subject group had a mean age of 45 years (range, 8 to 79 years).

Imaging Protocol

Magnetic resonance imaging was performed with a 1.5-T superconducting magnet (Signa; General Electric, Milwaukee, Wis). A standard protocol was adopted for the acquisition of MR brain images to be analyzed by the Brain Imaging Laboratory in the Department of Psychiatry. Three consecutive spin-echo pulse sequences were used to obtain images in each of three orthogonal planes (Fig 1). Using a T1-weighted sequence (TR = 600 milliseconds, TE = 20 milliseconds) sequence, sagittal images centered at the midsagittal plane were acquired to visualize the corpus callosum, brain stem, and other medial hemisphere surface landmarks. Subsequently, proton-density weighted (PDW) and T2-weighted images were obtained simultaneously for each section, using an asymmetric, multiple-echo sequence (TR = 2000 milliseconds, TE = 25 and 70 milliseconds). The sequence was used twice to obtain images of the entire brain in the axial and coronal planes in each subject (Fig 1, B
through E). Slice thickness was 5 mm with a 2.5-mm gap between successive slices in all instances. A 256 × 256 matrix and 24-cm field of view were used in all examinations.

Image Preprocessing

Signal intensity variations due to inhomogeneity of the magnetic field and imperfect radiofrequency pulses render simple image analysis techniques using signal thresholds unreliable. Improvement in spatial homogeneity of signal values is achieved by digital filtering of those pixel values from the soft tissues and fluid-filled spaces in and surrounding the brain. An interactive computer program is used with which the operator designates a pixel in the center of each isolated brain region in the image. A simple algorithm is used to determine the brain region by defining the designated pixel as a “brain pixel” and iteratively checking for pixels contiguous with brain pixels that are within a predetermined intensity range. Figure 2 shows the regions defined as brain in a typical section near the base of the skull. This step is performed on the T2W images. The high signal intensity CSF surrounding the brain is thus included in the “brain” regions and may subsequently be measured.

Brain pixels are further processed as follows: High-pass filtering is accomplished by applying a low-pass, two-dimensional filter, 32 × 32 pixels in size, with two iterations, and then subtracting the resultant matrix from the original matrix. A similar method was described previously in a report on techniques for analyzing CT images.

Pixel Classification

Each pixel within the “brain” region is classified into one of four categories based on its signal values in the T2W and PDW images. The categories were chosen to correspond to (1) CSF, (2) white matter, (3) gray matter, and (4) abnormal signal hyperintensities. All analyses are performed on images with randomly assigned subject numbers. The operators are blind to subject identification and diagnostic information.

Computation of “CSF” and “Gray and White” Images.—In a preliminary analysis, multiple linear regressions were performed to compute linear combinations of the pixel values (T2W and PDW) that would result in good separation of CSF, white matter, and gray matter. Samples of T2W and PDW values were obtained by interactively designating the locations of fully volumed white matter, gray matter, and CSF. Discriminant function analyses yielded the two optimal linear combinations of T2W and PDW values for CSF/brain and gray matter/white matter separation, respectively.

Using these results two new images are computed from the original images. The CSF image, shown in Fig 3, C, provides good contrast between CSF and brain, while the gray and white image provides improved contrast between gray matter and white matter (Fig 3, D).

Selection of Classification Criteria.—The criterion value for separation of white and gray matter is obtained by manually sampling these areas on the gray and white matrix, computing mean signal values for each sample, and using the midpoint between the two means as the criterion.

Sampling in regions of fully volumed CSF is not possible in many cases. Therefore, average differences between white matter and CSF values obtained in the preliminary analysis were used to set the CSF criterion as 24 U greater than the mean CSF matrix value of the white matter sample.

All pixel values are examined to determine whether they meet criteria for abnormal signal hyperintensities as follows: the pixel must have a higher signal intensity value on the CSF matrix than the mean of the white matter sample, a higher value on the PDW matrix than the mean of the gray matter sample, and a higher value than the estimated value of CSF on the T2W matrix.
Because signal values drift from section to section, especially near the edges of the magnetic field, all criteria above are calibrated separately for each section relative to the mean signal values of a sample of white matter from that section.

Region Definition

Trained operators, blind to subject age or diagnosis, manually separate cerebellar from cerebral areas, left from right hemispheres, and the cortical from subcortical regions, using a stylus-controlled cursor on the displayed image. Thus, separate estimates of the four classes of pixels are made for these areas. The fully processed images are illustrated in Fig. 4. Different pixel classes are color coded as follows: right hemisphere cortical gray matter is yellow and subcortical gray is light blue, left hemisphere cortical gray is white, and subcortical gray is green. The CSF is red, white matter is black, and signal hyperintensities are blue; however, CSF and signal hyperintensity pixels in or overlying the cortex are measured separately from those in deep structures or ventricles. The major measures were obtained by summing pixels in a given class, for each region, over all axial sections. Totals are then expressed as proportions of the total number of supratentorial pixels. An overall measure of cerebral volume loss is computed by expressing the total number of non-CSF cerebral pixels as a proportion of the supratentorial volume. This measure will be referred to as the cerebral proportion. The other measures to be presented are of cortical sulcal and ventricular CSF, cortical gray matter, and signal hyperintensities in cortical and subcortical zones. The hyperintensity proportions are multiplied by 100.

Interoperator Reliability

To estimate the interoperator reliability, 20 full sets of axial images from 20 different individuals were processed independently by two operators. The Table summarizes the interoperator correlations for the four classes, summed over all sections showing the cerebrum, and expressed as proportions of the total number of supratentorial pixels. The correlations range from .84 to .98 and reflect adequate reliability of the method across operators. Means and SDs are also given in the Table and confirm virtually identical classification criteria.

Statistical Analysis

Analyses of age effects were conducted for the major morphologic variables. Three-term polynomial regressions were performed for each morphologic variable and age. In addition, the residuals were examined for evidence of increasing variability in the measures, using similar polynomial regressions. The function presented includes only those terms that significantly increased the prediction based on the lower-order polynomial.

RESULTS

Routine clinical readings of the MR examinations were performed. In 33 cases no abnormalities were noted. A
single hyperintensity was noted in a 32-year-old woman, a 41-year-old man, and a 54-year-old woman. Scattered hyperintensities in deep cerebral white matter, mostly in the centrum semiovale, were noted in one 44-year-old man and seven women and five men older than 60 years of age. Both advanced volume loss for age and similar scattered hyperintensities were noted in a 49-year-old man, a 59-year-old woman (who also had a right cerebellar venous angioma), a 62-year-old man, and a 73-year-old woman. Advanced volume loss alone was noted in a 32-year-old woman and a 35-year-old man. A right cerebellar venous angioma was noted in a 9-year-old boy, and several tiny punctate hyperintensities in deep white matter were noted in a 10-year-old boy. Mild increase of signal intensity in the pons was noted in a 44-year-old man.

For each brain measure the simple linear correlation with age was highly significant. A significant linear decrease occurred over the age range in the cerebral proportion measure ($r = -0.82, P < .001$). The function is illustrated in Fig 5. Significant linear increases were observed in both cortical ($r = 0.58, P < .001$) and ventricular ($r = 0.65, P < .001$) CSF measures, as shown in Figs 6 and 7, respectively. The measure of cortical gray matter showed a curvilinear decrease with age (Fig 8) (multiple $r = .80$ with linear and quadratic age terms, $P < .001$). Both measures of signal hyperintensities were related to age with highly nonlinear functions, as illustrated in Figs 9 and 10 for the cortical condition (multiple $r = .60$ with linear, quadratic, and cubic terms, $P < .001$) and the subcortical condition (multiple $r = .68$ with linear, quadratic, and cubic terms, $P < .001$), respectively.

**COMMENT**

The age function shown in Fig 5 closely resembles that for a similar measure taken from neuropathologic material by Davis and Wright. In their study, as in the present study, brain volume declined from an average of about 93% of cranial volume at age 20 years to about 82% at age 80 years. Similarly, the functions in Figs 6 and 7 show increases in sulcal and ventricular CSF comparable to those demonstrated with CT. The measurements of volume loss presented herein appear to be comparable to those from quantitative CT measures and may approach the accuracy of postmortem measures, with the added advantage that the CSF may be measured locally.

The decrease in cortical gray matter volume also agrees with neuropathologic findings. Hubbard and Anderson reported values for a measure of the cerebral cortical volume for 19 normal controls. Their measure was expressed as a proportion of total cranial capacity and ranged from 0.45 to 0.39 in their 23- to 95-year-old subjects. Our measure is expressed as a proportion of the supratentorial cranium, excluding the posterior fossa; however, adjusting our values to include the volume of the posterior fossa in the denominator of the ratio yields values similar to the neuropathologic data. Our adjusted volumes range from 0.46 to 0.30 in the adult subjects. Thus, the range we observed is somewhat larger than in the earlier study. To some degree this increase in range would be expected with an increase in sample size from 19 to 58. It is possible, however, that the increased variability is also due to reduced sensitivity of this method relative to the postmortem methods. Potential contamination of the gray matter measures when white matter has abnormally lengthened T2 is discussed below. In any event, however, it seems likely that the age decreases in our measures of cortical gray matter reflect to a substantial degree those changes reported in postmortem brains. It is hoped that our correlating neuropathology studies will identify which cortical elements are lost in association with the observed volume changes.

In Fig 4, three sections from a young normal subject are shown in the top row, and three comparable sections from an elderly normal subject are shown in the lower row. Although these subjects were chosen randomly from the two subgroups, most of the age differences are discernible in the images. The CSF in both ventricles and sulci is increased in the elderly control relative to the younger subject. The cortical ribbon also appears somewhat thinner in the elderly subject.

The evidence from clinical MR studies suggests that signal hyperintensities are increasingly prevalent in normal subjects with advancing age. Our findings confirm this observation and suggest that hyperintensities occur in both cortical and subcortical zones. Visual comparison of the processed images with the original images reveals that pixels in small visible hyperintensities within white matter are frequently classified as gray matter pixels with our computer-aided methods. In fact, in most instances, the signal intensity values of these pixels, in both PDW and T2W matrices, are well within the range of normal gray matter, owing to partial voluming of white matter with hyperintensities or subtle degrees of demyelination. These small hyperintensities are therefore “missed.” This limitation is to some extent balanced, however, by an increased sensitivity to hyperintense areas adjacent to gray matter and CSF, where visual identification is often difficult. The measurement of hyperintensities with this method appears to be quite conservative, and possibly significant areas of abnormally increased signal may be missed. Also, normally hyperintense areas, such as those commonly occurring in association with vessels and at the tips of the ventricles, usually ignored by neuroradiologists, are necessarily included in our measures. Further studies are needed to determine the clinical correlates, if any, of our measures of signal hyperintensities.

**LIMITATIONS OF THE METHOD**

Interoperator reliability for the pixel classification step seems to be high. This sets a limit on the reliability of the overall method and ultimately on its validity. However, the reliability computed herein has no bearing on the reliability of image-analysis results across multiple examinations of the same individual since the same set of images was analyzed by both operators.

One weakness limiting the validity of our method is the necessity for obtaining reference samples from within the brain values. If the white matter is abnormal, an inappropriate estimate of the characteristic value for white matter may result, with a concomitant inappropriate adjustment of the classification criteria. Thus, an elevation
Fig 5.—The linear decrease of the cerebral proportion across the age range. The center line is a linear fit to the data ($r = -0.82, P < .001$). No significant change in the variance was observed. The upper and lower lines are 2 SDs above and below the regression line.

Fig 6.—The linear increase in cortical sulcal cerebrospinal fluid (CSF) across the age range. The center line is a linear fit to the data ($r = 0.58, P < .001$). No significant change in the variance was observed. The upper and lower lines are 2 SDs above and below the regression line.

Fig 7.—The linear increase in ventricular cerebrospinal fluid across the age range. The center line is a linear fit to the data ($r = 0.65, P < .001$). No significant change in the variance was observed. The upper and lower lines are 2 SDs above and below the regression line.

Fig 8.—The curvilinear decrease in cortical gray matter across the age range. The center curve is a polynomial fit to the data with significant linear and quadratic terms ($multiple r = 0.80, P < .001$). A linear increase in the variance was observed. The upper and lower curves diverge from the solid curve as predicted from this increase.

Fig 9.—The curvilinear increase in cortical signal hyperintensities across the age range. The center curve is a polynomial fit to the data with significant linear, quadratic, and cubic terms ($multiple r = 0.60, P < .001$). A nonlinear increase in the variance was also observed ($multiple r = 0.50$ with linear, quadratic, and cubic terms $P < .002$). The upper and lower curves represent the age-predicted 2-SD limits.

Fig 10.—The curvilinear increase in subcortical signal hyperintensities across the age range. The center curve is a polynomial fit to the data with significant linear, quadratic, and cubic terms ($multiple r = 0.68, P < .001$). A nonlinear increase in the variance was also observed ($multiple r = 0.83$ with linear, quadratic, and cubic terms, $P < .001$). The upper and lower curves represent the age-predicted 2-SD limits.
of T1 of white matter, a common occurrence in the brains of elderly subjects, could lead to a systematic misclassification of true gray matter pixels as "white matter" pixels. The operators attempt to exclude pathologic white matter from their sample selections. However, subtle variations may not be detectable or avoidable. To examine our data for evidence of significant inaccuracy due to this factor, we computed the correlations between the mean signal values in the white matter samples and the cortical gray matter proportions. The correlations were very small and not statistically significant. Furthermore, a partial correlation removing the effects of variation in white matter signal intensity yielded a slightly higher linear correlation between the cortical gray matter proportion and age (−.81 vs. −.77). Thus, changes in gray matter classification induced by lengthened T1 in white matter appear to account for a negligible degree of variation in our gray matter proportions, and there is no evidence that the age-related decrease in cortical gray matter volume is mediated by white matter changes.

In some white matter regions, such as the forceps major, pixels may have values identical to cortical gray matter on both PDW and T2W images. Since the classification is based on signal intensity values, these pixels will be misclassified as gray matter, leading to a systematic overestimate of gray matter volume, especially in the subcortical region, with a commensurate underestimate of white matter.

Another limitation of the method arises when fully volumed white matter samples are not available on each section processed. This occurs most commonly at the periphery of the brain: at the base and the vertex in the axial series and at the frontal and occipital poles in the coronal series. These are also the regions in which the signal strength tends to be most variable. Frequently, on these sections, an estimate of the white matter value must be taken from a section several centimeters removed, and the classification may at times be inappropriate.

Currently, we use partially volumed samples on the sections or avoid using the data from these sections when possible.

A modification to the method that could substantially improve its validity, reducing both errors due to signal alterations in tissue and errors due to unavailability of tissue samples, would involve the inclusion of a standardization object in the image. The object would be present on all sections, preferably producing MR signals on both images in or near the range of those obtained within brain. Sampled values from such an object could be used in automated adjustments of the criteria. Attempts to include such an object to improve the method are underway.

This investigation was supported in part by grants from the Medical Research Service of the Veterans Administration and by grant NS22943 from the National Institute of Neurological Disorders and Stroke, Multidisciplinary Research Center for the Study of the Neurological Basis of Language, Learning and Behavior Disorders in Children, San Diego, Calif.

Invaluable assistance in the recruitment of elderly subjects has been provided by the staff of the San Diego National Alzheimer's Disease Research Center directed by Robert Katzman, MD. Al Ahumada, PhD, and Renee Dupont, MD, gave assistance during the development of the methods. Sarah Archibald, MA, Katherine Grant, and David Foster assisted in image analysis. Mary Jenkins, MA, helped with preparation of the manuscript.

References

32. Awad IA, Johnson PC, Spetzler RF, Hodak JA. Incidental subcortical lesions identified on magnetic resonance imaging in the elderly, II: postmortem pathological correlations. Stroke. 1986;17:1090-1097.

32 Arch Neurol—Vol 47, January 1990

Brain Morphometric Features—Jernigan et al