
R N Bryan, T A Manolio, L D Schertz, C Jungreis, V C Poirier, A D Elster and R A Kronmal

http://www.ajnr.org/content/15/9/1625
A Method for Using MR to Evaluate the Effects of Cardiovascular Disease on the Brain: The Cardiovascular Health Study


PURPOSE: To do a pilot study for the Cardiovascular Health Study (a population-based, longitudinal study of coronary heart disease and stroke in adults 65 years of age and older designed to identify risk factors related to cerebrovascular disease, particularly stroke): (a) to determine the feasibility of adding brain MR to the full-scale study; (b) to evaluate the reliability of standardized MR image interpretation in a multicenter study; and (c) to compare the prevalence of stroke determined by MR with that by clinical history. METHODS: Protocol-defined MR studies were performed in 100 subjects with clinical histories of stroke and 203 subjects without reported histories of stroke. MR scans were independently evaluated by two trained neuroradiologists for the presence of small (≤3 mm) and large (>3 mm) "infarctlike" lesions. The sizes of the cerebral sulci and lateral ventricles and the extent of white matter disease were graded on a scale of 0 to 9.

RESULTS: Eighty percent of the Cardiovascular Health Study participants who were invited to undergo MR studies agreed to do so; 95% of those agreeing to the procedure successfully completed the exams. Intrareader and interreader reliability of infarctlike lesion identification was high for large lesions (κ, 0.71 and 0.78, respectively) but not for small lesions (κ, 0.71 and 0.32, respectively). Relaxed intrareader and interreader κ scores for sulcal and ventricular sizes and extent of white matter disease were greater than 0.8. MR evidence of infarctlike lesions was present in 77% of the participants with histories of stroke but was also present in 23% of the participants without clinical histories of stroke. Seventy-nine percent of the infarctlike lesions were larger than 3 mm.

CONCLUSIONS: This preliminary study indicates that a large, prospective, epidemiologic study of elderly subjects using MR scans of the brain for identification of cerebrovascular disease is feasible and that the interpretative results are reproducible, and suggests that MR evidence of stroke is more prevalent than reported clinical history of stroke.

Index terms: Heart; Brain diseases; Brain, infarction; Brain, magnetic resonance; Magnetic resonance, comparative studies


The Cardiovascular Health Study is a longitudinal, multicenter, epidemiologic study of cardiovascular disease in 5201 men and women 65 years of age and older (1). The study is designed to determine risk factors for cerebrovascular disease based on the association of possible risk factors with the prevalence and incidence of clinical and subclinical cerebrovascular disease in this cohort.

Magnetic resonance (MR) imaging was added to the Cardiovascular Health Study to assess the extent and severity of cerebrovascular disease in this population, because MR is noninvasive and very sensitive to brain disease and carries minimal risk (2). MR findings related to cerebrovascular disease include focal parenchymal lesions with abnormal MR signals such as infarcts and hematomas (3–5), diffuse parenchymal lesions with abnormal MR
Materials and Methods

Cardiovascular Health Study Cohort

The Cardiovascular Health Study cohort was recruited from four US communities: Forsyth County, NC; Sacramento County, Calif; Washington County, Md; and Pittsburgh (Allegheny County), Pa (1).

Extensive physical and laboratory evaluations were obtained at baseline to identify the presence and severity of cardiovascular disease and known or putative risk factors. Details of the overall study design have been described (1).

Classification of Cardiovascular Events

Criteria for stroke and transient ischemic attack are similar to those used in the Systolic Hypertension in the Elderly Program (14). We use the following definitions:

- Stroke: abrupt onset of new neurologic deficit lasting at least 24 hours, with specific location confirmed by unequivocal physical examination or laboratory data and without evidence of an underlying nonvascular cause.
- Transient ischemic attack: rapid onset of focal neurologic deficit lasting less than 24 hours, assessed to be caused by ischemia, without evidence of underlying nonvascular cause.

MR Feasibility Trial

This study of cerebral MR examinations of 303 Cardiovascular Health Study participants 65 to 95 years of age was undertaken to evaluate the feasibility of performing MR on the entire Cardiovascular Health Study cohort. A Cardiovascular Health Study MR Imaging Reading Center was established to assist in the design of the MR protocol, evaluate its feasibility, and examine the reproducibility of interpretative results. The study included 100 randomly selected participants who had reported prior strokes and 203 persons who had not reported strokes and had not had them during follow-up. The MR exams were prospectively performed with standardized spin density-, T1-, and T2-weighted pulse sequences on 1.5-T MR scanners at three field centers and a 0.35-T scanner at one field center (2). Axial images were angled to be parallel to the anterior commissure–posterior commissure line (17).

Scan data were stored on 0.5-inch magnetic tapes that were mailed to the MR Imaging Reading Center, which archived and analyzed the image data and then transmitted the results to the Cardiovascular Health Study Coordinating Center. This was performed in a completely digital fashion, without any hard copies of images or interpretative results. The image data transmission and archival format was ACR-NEMA 2.0. The ACR-NEMA image files of the MR instrument manufacturer were converted to the JHU standard ACR-NEMA 2.0 version, which is the common image format in the Image Archival and Retrieval System used for a long-term (10 year) storage. The Image Archival and Retrieval System hardware includes an optical disk "jukebox" (Kodak, Rochester, NY) managed by an Image File Server (Vortech, Dallas, Tex). The Image Archival and Retrieval System stores the images on optical media for retrieval at a later date. Images from optical storage were transmitted to a Vortech Personal Display System 4 work station used for interpretation.

The Personal Display System 4 monitors measure 16 inches diagonally with 1024 × 1024 pixel elements and
256 gray-scale intensities. A mouse or track ball is used to select and manipulate images. A number of image-manipulation features are available on the display workstation, including window/level, zoom, measure, select a scout line, annotate, and print. The system software is Macintosh (Apple Computer, Cupertino, Calif) based, providing a user-friendly, menu-driven working environment.

Based on visual workstation image analysis, supplemented by simple cursor measurements, the MR image data results were reported on the Cardiovascular Health Study MR Scan Coding form, a menu-driven form based on File Maker Pro, which is Macintosh based. The results were entered directly into the coding form database, which is located in a second Macintosh computer adjacent to the image work station. The complete form was designed to be easily completed within 15 minutes by a trained reader viewing the images on the work station. The results categories were carefully designed to be as comparable as possible with other Cardiovascular Health Study variables, particularly those of the Cardiovascular Health Study Neurologic Evaluation for Stroke/Transient Ischemic Attack form. The files of the results, formatted in the American Standard Code for Information Interchange, were sent to the Coordinating Center, where the complete Cardiovascular Health Study database resides, via electronic mail.

Classification of “infarctlike” lesions was based on size, location, and signal intensity. Suspected infarcts were reported as small (≤3 mm) or large (>3 mm) infarctlike lesions (Figs 1 and 2). The locations of small infarcts was reported, whereas the sizes and locations of large infarcts were reported. Size was reported as maximum anterior-to-posterior and right-to-left dimensions. Location was reported by selection of one or more choices from a “pop-up” menu containing 16 anatomic locations.

Infarcts were defined as lesions without mass effect but with abnormal signal intensity in a vascular distribution. Nonhemorrhagic infarcts involving cortical gray matter had to be bright on spin density- and T2-weighted images relative to normal gray matter. They could be hypointense or isointense on T1-weighted images. Similarly, nonhemorrhagic infarcts involving the deep nuclear region had to be bright on spin density-weighted images and bright on T2-weighted images. They also could have been isointense or dark on T1-weighted images. The requirement for hyperintensity on spin density-weighted images was intended to distinguish small deep nuclear region infarcts from dilated perivascular spaces. For the purposes of this study, the deep nuclear region was defined to include the caudate nucleus, lentiform nucleus, internal capsule, external capsule, extreme capsule, and thalamus.

Nonhemorrhagic infarcts in the white matter likewise had to be bright on spin density-weighted images and bright on T2-weighted images, but they also had to be dark on T1-weighted images, approaching the T1 hypointensity of cerebrospinal fluid. This T1 hypointensity requirement was intended to distinguish actual white matter infarction from the more prevalent, nonnecrotic, white matter disease.

Any hemorrhagic nonmass lesion in a vascular distribution was recorded as an infarct with hemorrhage. Signal criteria included heterogeneously increased signal on T1-weighted images and heterogeneously decreased signal on T2-weighted images.

In addition, measures of atrophy were reported, including ventricular and sulcal size (on a semiquantitative scale of 0 to 9) as well as extent of white matter disease (also on a semiquantitative scale of 0 to 9) (2). Image interpretation was based on “pattern matching” of individual subject scans to a library of 40 example studies retained at the MR Imaging Reading Center.

All MR exams were interpreted by two board-certified radiologists (second-year neuroradiology fellows) who had been trained by the MR Imaging Reading Center principal investigator. Training used a detailed reading manual containing anatomic diagrams and definitions of terms. A library of example cases, which included large and small infarcts, all grades of ventricular and sulcal size, and extent of white matter disease, was used for training and subsequent reference.

Twenty percent of the readings were duplicated blind to assess reproducibility—10% within readers and 10% between readers. Analysis included determination of intrareader and interreader k scores of reliability for reporting major variables from 60 randomly selected studies.

### Results

A total of 334 participants agreed to undergo MR, but matching of cases and control subjects, no-show and ineligible patients, and unreadable tapes led to only 303 actually having scans completed and interpreted. An additional 85 eligible participants refused to undergo imaging, for an acceptance rate of 334 of 419 or 79.7%.

Reasons for refusal follow.

<table>
<thead>
<tr>
<th>Reason for Refusal</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not interested</td>
<td>20</td>
</tr>
<tr>
<td>Sick</td>
<td>21</td>
</tr>
<tr>
<td>Caring for relative</td>
<td>4</td>
</tr>
<tr>
<td>Claustrophobia</td>
<td>15</td>
</tr>
<tr>
<td>Scheduling constraints</td>
<td>25</td>
</tr>
</tbody>
</table>

Among the “scheduling constraints” group, 15 were out of town or had other temporary scheduling constraints that prevented their having the exams during the relatively short duration of the feasibility study (120 days).

Based on a questionnaire answered 1 month after their MR scans, 96% of the participants who completed the examinations stated that they would be willing to have follow-up MR scans.

The individual scans were evaluated at the MR Imaging Reading Center for protocol ob-
Fig 1. Deep-nuclear-region lesions. A, D, and G, Spin density–weighted images (3000/30 [repetition time/echo time]); B, E, and H, T2-weighted images (3000/10); C, F, and I, T1-weighted images (500/20).

A–C, Small (<3-mm) infarct, left internal capsule (arrow). Perivascular space, right caudate nucleus (curved arrow). The latter has no signal increase on the spin density–weighted image. Neither lesion category has high interobserver reproducibility.

D–F, Small infarct, left lentiform nucleus (arrow). Large (6 × 3-mm) infarct, right lentiform nucleus (curved arrow).

G–I, Multiple large infarcts, right and left lentiform/capsular regions and thalami (arrows).
Fig 2. Large cortical and subcortical white matter infarcts. A, D, and G, Spin density-weighted images; B, E, and H, T2-weighted images; C, F, and I, T1-weighted images.

A–C, Cortical/subcortical and subcortical left cerebellar hemisphere infarcts (arrows).

D–F, Right frontal, subcortical white matter infarct (arrow). Also note the left internal capsule infarct (curved arrow).

G–I, Left cerebral hemisphere cortical and subcortical infarct in a watershed distribution (arrows).
servance and image quality. Fewer than 3% of studies were judged to be technically inadequate for any reason, and the rate of nonacceptability did not vary among field centers. The readers were blinded to sources of images and reported no detectable quality differences among scans from different field centers, except lower signal-to-noise ratios on the 0.35-T scans, which also may have slightly less T1 contrast related to small lesions.

The \( \kappa \) statistics for intrareader and interreader agreement on the presence of infarctlike lesions are shown in Table 1. Kappa scores were acceptable (\( >0.7 \)) for infarct detection, except for a \( \kappa \) of 0.3 for interreader agreement on small infarcts.

Kappa statistics for semiquantitative grading (on a scale of 0 to 9) of atrophy (as reflected by ventricle and sulcal size) and extent of white matter disease are shown in Table 2. All \( \kappa \) scores were 0.5 or greater, except for sulcal size, which was less than 0.3. However, a relaxed \( \kappa \), based on a variation of plus and minus 1 in sulcal grade, was greater than 0.8. Similarly relaxed \( \kappa \) scores for lateral ventricle grade and white matter disease were 1.0.

Table 3 summarizes the detection of infarctlike lesions in the group with clinical history of stroke and the control group. Seventy-two percent of the clinically positive group had MR scans with infarctlike lesions, 83% of which were larger than 3 mm. Twenty-three percent of the control group had MR studies with infarctlike lesions, 72% of which were larger than 3 mm.

**Discussion**

Initial concern about the ability of an elderly population to tolerate this exam proved unwarranted; 80% of the eligible participants completed the exam. Of the 85 participants who refused the examination, 15 (20%) gave claustrophobia as the reason. Ninety-eight percent of the participants who arrived at the scanning facility successfully completed the exam. Furthermore, fewer than 3% of these studies were judged technically inadequate for any reason, including excessive motion. In addition, the results of a follow-up questionnaire indicate that more than 90% of the participants scanned under this protocol would agree to second studies. In summary, these results indicate that a prospective MR study of a large elderly population is feasible, and that approximately 80% of the full Cardiovascular Health Study cohort (approximately 4000 subjects) will be able to be studied.

Little experience exists with the use of MR imaging in multicenter studies. For such studies it is essential to address two key factors: the comparability of studies obtained in different locations and the reproducibility of image interpretation. The recent report of the Consortium to Establish a Registry for Alzheimer’s Disease illustrates many of the difficulties encountered in a large-population multicenter study of brain MR scans (15). Major problems encountered in the latter study included noncomparability of studies, primarily related to inconsistent scan angles, and variations in reader evaluations of cerebral atrophy, extent of white matter disease, and presence or absence of infarcts. We designed the Cardiovascular Health Study protocol to minimize as many of these problems as possible, given the practical constraints (including financial) of such a study.

**TABLE 1: Reproducibility of infarct interpretation**

<table>
<thead>
<tr>
<th>Infarct</th>
<th>Intrareader</th>
<th>Interreader</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>( \kappa )</td>
<td>n</td>
</tr>
<tr>
<td>Infarct &gt; 3 mm</td>
<td>29</td>
<td>0.71</td>
</tr>
<tr>
<td>Infarct ≤ 3 mm</td>
<td>29</td>
<td>0.71</td>
</tr>
</tbody>
</table>

**TABLE 2: Reproducibility of atrophy and white matter disease grades**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intrareader</th>
<th>Interreader</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>( \kappa )</td>
</tr>
<tr>
<td>Lateral ventricle</td>
<td>29</td>
<td>0.64</td>
</tr>
<tr>
<td>Relaxed(^a)</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>Sulci</td>
<td>29</td>
<td>0.29</td>
</tr>
<tr>
<td>Relaxed(^a)</td>
<td></td>
<td>0.86</td>
</tr>
<tr>
<td>White matter disease</td>
<td>29</td>
<td>0.61</td>
</tr>
<tr>
<td>Relaxed(^a)</td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

\(^a\) ±1 grade allowance.

**TABLE 3: Cardiovascular Health Study participants with MR-defined infarctlike lesions**

<table>
<thead>
<tr>
<th></th>
<th>Clinical Stroke (n = 100)</th>
<th>Nonstroke Control (n = 203)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR infarct present</td>
<td>72 (72%)</td>
<td>46 (23%)</td>
</tr>
<tr>
<td>&gt;3 mm</td>
<td>60</td>
<td>33</td>
</tr>
<tr>
<td>≤3 mm</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>MR infarct absent</td>
<td>28 (28%)</td>
<td>157 (77%)</td>
</tr>
</tbody>
</table>

Note.—See "Methods" for definition of "infarct."
Reliability of interpretative results was a concern on this project because of its experimental importance and limited prior reproducibility data (18). In terms of stroke, no report of MR interpretation reproducibility of nonacute stroke could be found, although a report on acute stroke detection indicated an interobserver correlation coefficient of greater than 0.9 (3).

In this study, the “large infarct” category had satisfactory reproducibility, but the “small infarct” category did not. The Reading Center reexamine the studies in which readers disagreed whether infarctlike lesions were present. In nearly all cases the studies were adjudicated as “infarct present.” This suggests that if a reader reports a small infarct, it almost certainly will be seen on rereview, but sometimes a first-time reader may miss infarcts, particularly small ones. The low interreader agreement on small infarcts has prompted the Reading Center to revise its protocol so that in the main study, all scans will be read in duplicate for the presence or absence of infarcts. This procedure takes very little time and should improve the sensitivity of the study, possibly at the expense of specificity.

Numerous papers have reported on the reproducibility of grading cerebral atrophy and the extent of white matter disease, most using fewer grade levels in their scoring schema. Hendrie et al reported a $\kappa$ value of 0.69 for the grading of white matter disease on a scale of 0 to 3 (19), whereas Leys et al reported $\kappa$ values of 0.43 and 0.5 for periventricular and subcortical white matter hyperintensities, respectively (20). In the study by the Consortium to Establish a Registry for Alzheimer’s Disease, an intraclass correlation coefficient of 0.77 was determined for white matter disease and 0.64 to 0.82 for measures of cortical atrophy, with ventricle sizes having the highest coefficients and cerebral sulci the lowest coefficient. Earlier x-ray computed tomographic studies showed similar levels of reproducibility of ventricular and sulcal size (21, 22). These studies suggest that readers can subjectively grade ventricular size into roughly five levels and sulcal size into three to four levels with reasonable reliability. The ability to grade the relative degree of white matter disease is more difficult to assess from the literature, because most previous studies have attempted to subdivide white matter disease into subcategories such as “periventricular or subcortical” and “patchy or confluent.” In general, the more subcategorization attempted, the poorer the reproducibility (15). For this reason, we chose to grade white matter disease on a global scale, not attempting to differentiate the various patterns of the signal changes.

The 0-to-9 scale used in this study for grading ventricular and sulcal size and extent of white matter disease was intended to be overly fine in terms of dynamic range so as not to compress the data excessively and to diminish the influence of digit preference. The results indicate that even on a scale of 10, absolute agreement within and between readers is satisfactory ($\kappa > 0.5$) for ventricle size and white matter disease. The grading of sulcal size is less reproducible on this scale, but the relaxed $\kappa$ of greater than 0.8 with an allowance of plus and minus 1 suggests that only four to five grades of sulcal size can be reproducibly reported with our system.

Probably the most important research design factors incorporated into the Cardiovascular Health Study-MR Study are a relatively rigid scanning protocol, the use of a reference library of “standard cases,” and the preservation of digital image data. Loosely defined scanning protocols consistently have resulted in noncomparability of images (15). In terms of image comparability, fewer than 3% of the exams were inadequate for any reason, including deviations from the protocol. Readers did not detect any consistent differences in images from the various field centers, except lower signal-to-noise ratios on the 0.35-T images.

The use of the anterior commissure–posterior commissure coordinates is an important part of the protocol design. It greatly increases the comparability of studies between subjects. It also increases the accuracy of anatomic location (17). It not only aids the pooling and comparison of intersubject image data, but will greatly increase the value of future MR studies under the Cardiovascular Health Study scanning protocol. Sequential exams of the same subject can be directly compared without compromising normalization procedures.

Most previous studies have relied on “hard copies” and limited results categories. These approaches intrinsically lose much of the initially available data. Hard copies of medical digital images reflect fewer than 75% of the digital data, primarily because of gray-scale and windowing limitations (23). Furthermore, when only hard copies are archived, the original, richer digital data are lost. Image data analysis
relying on a limited number of nonquantitative results categories further limits the analysis by greatly compressing the data into a few arbitrary fields.

Rather than strive for a perfect method, we aimed for immediate practicality while allowing future flexibility. The latter is addressed by use of a completely digital design with preservation of reconstructed digital image data. The long-term archiving of digital data is now feasible with optical and magnetic storage media. The complete Cardiovascular Health Study image data file of approximately 4000 studies will total approximately 40 gigabytes, which will be contained on approximately 60 optical disks contained in the Image Archival and Retrieval System. These data will be available for future studies on any Cardiovascular Health Study-approved workstation, within or outside of the MR Imaging Reading Center, which is connected to the network.

MR-defined stroke (infarctlike lesions) was detected in 72% of participants with clinical histories of stroke. In those subjects without clinical histories of stroke, 23% were found to have infarctlike lesions on MR. Most of the infarctlike lesions in both groups were larger than 3 mm (83% in the clinical stroke group, 72% in the control group). These results are consistent with, but do not prove, the hypothesis that MR is more sensitive to stroke than is clinical history. Although lacking pathologic confirmation, the far greater prevalence of lesions in the group with positive clinical histories of stroke suggests that there is reasonable specificity of the findings. Previous radiologic-pathologic correlations of MR findings similar to those used in our definition of infarctlike lesions also indicate that most, if not all, MR-defined infarcts in this study are indeed infarcts (4, 5). Whether the cases with clinical histories of stroke and negative MR studies were false-negative MR results or false-positive clinical results may not be determinable, but this important group of cases is being further investigated in detail and will be reported separately. In terms of false-negative MR studies, white matter lesions with normal T1-weighted and increased T2-weighted signal intensities are of particular concern. We have not classified such lesions as infarcts on the assumption that they are regions of ischemic demyelination, not regions of ischemic necrosis (6, 7, 8). This controversy needs to be further evaluated with additional clinical-radiologic-pathologic correlative studies.

The 23% prevalence of infarctlike lesions in the control group is consistent with the hypothesis that there is significant subclinical cerebrovascular disease detectable by MR. This group of subjects is also being further investigated and will be reported with the clinical-positive-MR-negative group. If one assumes that these are indeed subclinical strokes, then the projected prevalence of stroke (clinical and subclinical) in the entire Cardiovascular Health Study cohort is 26%, six times greater than the prevalence of clinical stroke. Even if only infarcts larger than 3 mm are counted, the projected prevalence of Cardiovascular Health Study clinical and subclinical stroke is 20%, four times greater than clinical stroke alone. These results should increase the power of the study to detect risk factors of cerebrovascular disease significantly.

The results of this study suggest that the MR protocol used in this pilot project is a practical and reproducible method for identifying subclinical cerebrovascular disease in the elderly. These additional data should significantly increase the power of the Cardiovascular Health Study to determine associated risk factors, which it is hoped will lead to better understanding and management of cerebrovascular disease risk factors. This hope is supported by a separately published paper from this same study, which demonstrates strong correlations between age, various cardiovascular risk factors (particularly hypertension), and the grades of cerebral atrophy and extent of white matter disease (2). As a result of this feasibility trial, this MR protocol is now being applied to the full Cardiovascular Health Study cohort. The MR method also may be applicable to other epidemiologic studies of cerebrovascular disease.

Acknowledgments

We thank Grace M-H. Lee, MD, Linda Wilkins, RT, Chul Paik, PhD, and Marge Eddy for their contributions and help in preparing this manuscript.

References


