MR Image Texture Analysis Applied to the Diagnosis and Tracking of Alzheimer's Disease

Peter A. Freeborough* and Nick C. Fox

***Abstract*— We assess the value of magnetic resonance (MR) image texture in Alzheimer's disease (AD) both as a diagnostic marker and as a measure of progression. T_1-weighted MR scans were acquired from 40 normal controls and 24 AD patients. These were split into a training set (20 controls, 10 AD) and a test set (20 controls, 14 AD). In addition, five control subjects and five AD patients were scanned repeatedly over several years. On each scan a texture feature vector was evaluated over the brain; this consisted of 260 measures derived from the spatial gray-level dependence method. A stepwise discriminant analysis was applied to the training set, to obtain a linear discriminant function. In the test set, this function yielded significantly different values for the control and AD groups $(p < 10^{-4})$ with only small group overlap; a classification rate of 91% was obtained. For the repeatedly scanned control subjects, the median increment in the discriminant function between successive scans of 0.12 was not significantly different from zero $(p > 0.05)$; for the repeatedly scanned AD patients the corresponding median increment of 1.4 was significantly different from zero $(p < 0.05)$. MR image texture may be a useful aid in the diagnosis and tracking of Alzheimer's disease.**

***Index Terms*— Alzheimer's disease, classification, magnetic resonance imaging (MRI), texture.**

I. Introduction

A definite diagnosis of Alzheimer's disease (AD) relies on histological confirmation at post-mortem or cerebral biopsy, and, therefore, in the majority of cases a definitive diagnosis is not available in life. The diagnosis of AD remains a clinical one with an accuracy of approximately 80%–90% [1]–[2]. With the development of new drug therapies for AD, there has, therefore, been a need to improve the accuracy of diagnosis, particularly in the early stages of the disease when diagnosis is most difficult. Furthermore, to assess potential drug therapies it is desirable to be able to measure the progression of the disease accurately. The quantitative analysis of magnetic resonance (MR) images may offer such measures.

Studies of intracranial tumors have demonstrated that MR image texture may be used to determine the tumor type [3]–[5]. However, to the best of our knowledge, there have been no reported studies of MR image texture changes in AD. Post-mortem studies reveal that AD is a destructive process affecting the majority of the brain. If one considers the brain structure observed on magnetic resonance imaging (MRI) to represent a macro-texture, one would expect the degeneration of the brain in AD to be reflected in a simplification of this texture. In this paper we assess the value of MR image texture both as a diagnostic marker for AD, and as a measure of disease progression.

Manuscript received April 28, 1997; revised January 12, 1998. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was M. W. Vannier. *Asterisk indicates corresponding author.*

*P. A. Freeborough is with the Dementia Research Group, National Hospital for Neurology and Neurosurgery, 8-11 Queen Square, London WC1N 3BG UK (e-mail: paf@ic.ac.uk).

N. C. Fox is with the Dementia Research Group, National Hospital for Neurology and Neurosurgery, London WC1N 3BG UK

Publisher Item Identifier S 0278-0062(98)06454-4.

II. Methods

A. Patients and Scans

Forty normal subjects and 24 AD patients were recruited into the study. The AD group consisted of ten patients with familial AD (each of whom had a family member with autopsy proven AD) and 14 patients with a clinical diagnosis of probable AD (none of whom had a family history of the disease).

MRI head scans were acquired for each of the 40 normal subjects and the 24 AD patients on a General Electric 1.5-T Signa unit using a spoiled gradient-echo technique. The sequence utilized a $256 \times 128 \times 124$ matrix over a $24 \times 24 \times 19.2$-cm field of view, with parameters TR $= 35$ ms, TE $= 5$ ms, 35° flip angle and one excitation. Prior to inverse Fourier transformation the scanner automatically zero padded the matrix to $256 \times 256 \times 124$, yielding final voxel dimensions of $0.94 \times 0.94 \times 1.5$ mm. The resultant scans possessed T_1-weighted contrast.

A "training set" was formed from 20 randomly selected normal scans in conjunction with the ten scans from the patients with familial AD. The familial cases were used in this set as their diagnoses were more certain than those of the sporadic AD patients. The remaining 20 normal scans and 14 scans from patients with sporadic AD were utilized as a "test set."

B. Texture Measurement

The brain was identified on each scan using an interactive three-dimensional (3-D) morphological segmentation tool [6]–[7], and voxel intensities were rescaled to yield a mean of 128 over the resultant brain region.

The texture was quantified over all coronal slices of the brain region using the spatial gray-level dependence method (SGLDM) [8] with measures 1–13 (see Appendix). SGLDM involves the estimation of the discrete second-order probability function $C(i, j|\Delta x, \Delta y)$, which represents the probability of occurrence of a pixel pair with gray levels i and j given the spacing between the pixels is Δx and Δy in the x and y dimensions respectively. This is referred to as a *cooccurrence matrix*, which we estimated from a scan as in (1), shown at the botom of the following page, where $I(x, y)$ represents the image intensity at voxel (x, y) in a given slice; a single matrix was estimated over *all* coronal slices of the scan. The brain region identified with the segmentation tool was used to determine which voxels were within brain.

We estimated symmetric cooccurrence matrices by averaging the matrices $C(i, j|\Delta x, \Delta y)$ and $C(i, j|-\Delta x, -\Delta y)$; so eliminating the distinction between opposite offset directions, which is appropriate as textures are homogenous by definition. Each of the latter matrices were estimated using (1), the previous brain identification being used to determine which voxels were within brain.

The offset $(\Delta x, \Delta y)$ is conventionally represented as an offset angle and distance [8]. For a given offset distance, we calculated symmetric cooccurrence matrices for offset angles of 0°, 45°, 90°, and 135°. From each of these matrices we derived the 13 SGLDM measures described in the appendix. The mean and range of each measure over the four offset angles were used as features; this yielded 26 features for each offset distance.

The features were calculated for offset distances of 1, 2, 3, 4, 6, 8, 12, 16, 24, and 32; a large range was chosen so that the derived measures may be sensitive to texture differences at many scales. This yielded a total of 260 measures.

C. Feature Selection and Classification

A subset of the 260 features was selected using the sequential forward selection method based on the Mahalanobis distance between the control and patient groups in the training set [9]–[10]. Starting with an empty subset of features, at each stage of this procedure the feature which yielded the largest significant increase in the Mahalanobis distance was added to the subset; the procedure iterated until none of the remaining features would significantly $(p < 0.05)$ increase the Mahalanobis distance.

The resultant subset of features was considered to form a feature vector, which we denote $\boldsymbol{y}$. The instances of $\boldsymbol{y}$ in the training set were used to parameterize the linear discriminant function

$$D(\boldsymbol{y}) = \boldsymbol{b}^T \boldsymbol{y} + c \tag{2}$$

where

$$\boldsymbol{b} = 2\boldsymbol{K}^{-1}(\boldsymbol{m}_{\mathrm{AD}} - \boldsymbol{m}_C) \tag{3}$$

$$c = \boldsymbol{m}_C^T \boldsymbol{K}^{-1} \boldsymbol{m}_C - \boldsymbol{m}_{\mathrm{AD}}^T \boldsymbol{K}^{-1} \boldsymbol{m}_{\mathrm{AD}} \tag{4}$$

and where $\boldsymbol{m}_C$ is the mean $\boldsymbol{y}$ over the control group, $\boldsymbol{m}_{\mathrm{AD}}$ is the mean $\boldsymbol{y}$ over the AD group and $\boldsymbol{K}$ is the covariance matrix of $\boldsymbol{y}$ over both groups. Classification using this discriminant function was performed using a decision boundary of $-2\ \ln(\lambda)$, where λ is the assumed ratio of the prior probabilities of each class, i.e., P(AD) divided by P(control); this corresponds to the standard linear classifier [10]. Prior to application to the test set, we arbitrarily chose a value of $\lambda = 2$, so as to place the decision boundary closer to the control group mean than to the AD group mean. This was appropriate because we expected the discriminant function to cover a large range in the AD group (reflecting a range of disease severities) and a relatively small range in the control group; a classifier should, therefore, place more significance on deviations from the control group mean than on deviations from the AD group mean.

D. Application to Test Set

The discriminant function was then calculated for each member of the test set; the necessary subset of texture features being calculated for each scan as required. A decision boundary of $-2\ \ln(2)$ was used for classification. Difference in the value of the discriminant function for the patient group compared to the control group was tested for statistical significance using the Mann-Whitney U test [11].

In addition, five of the control subjects and five of the AD patients had several further scans over a 2–4-year time interval (these groups included members of both test and training sets); the discriminant function was evaluated for each of these additional scans. The median change in the discriminant function between successive scans was tested for statistical significance using a two-tailed sign test [11].

III. Results

The discriminant function (determined using the training set) is shown in (5), where the parameters a–q are the SGLDM features specified in Table I

$$\begin{aligned} D = {} & -0.101a + 1.57b - 0.193c + 51.0d + 6.56e - 21.8f \\ & + 122g + 36.9h - 23.7i + 342j - 3.43k - 1.42l \\ & - 26.0m - 0.0373n + 0.007\,67o - 173p + 0.264q + 866. \end{aligned} \tag{5}$$

TABLE I
The SGLDM [8] Texture Features Used as Parameters to the Texture Discriminant Function of (5)

Parameter	SGLDM Feature	Mean/Range	Distance
a	Difference Variance	Range	4
b	Variance	Range	1
c	Difference Variance	Range	6
d	Entropy	Range	8
e	Sum Average	Range	12
f	Correlation	Mean	16
g	Difference Entropy	Range	16
h	Entropy	Range	1
i	Correlation	Mean	24
j	Inverse Difference Moment	Range	24
k	Sum Average	Mean	24
l	Sum Average	Range	24
m	Sum Entropy	Range	24
n	Difference Variance	Range	24
o	Sum Variance	Range	32
p	Sum Entropy	Range	3
q	Difference Variance	Range	3

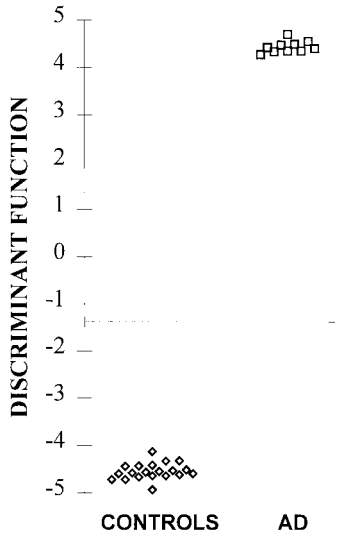

Fig. 1. The texture discriminant function evaluated for each member of the training set; the x-axis is at the level of the classification decision boundary i.e., $D = -2\ ln(2)$. The wide separation of the groups merely demonstrates that the discriminant function models the training data closely, and is not indicative of the true discriminatory power of the technique.

Fig. 1 shows the value of the discriminant function for each member of the training set. Fig. 2(a) shows the value of the discriminant function for the test set. In the test set, the values of the discriminant function for the control and AD groups were significantly different $(p < 10^{-4})$, with respective means of -4.18 and 1.60, and respective standard deviations of 1.71 and 3.82. Classification using the discriminant function with a decision boundary at $-2\ \ln(2)$, resulted in the correct classification of all control subjects but the misclassification of three AD patients; this corresponded to a classification rate of 91% (79% sensitivity and 100% specificity).

$$C(i,j|\Delta x, \Delta y) = \frac{\text{No. of } (x,y) \text{ for which: } I(x,y) = i, I(x+\Delta x, y+\Delta y) = j, \text{ both } (x,y) \text{ and } (x+\Delta x, y+\Delta y) \text{ are within brain}}{\text{No. of } (x,y) \text{ for which both } (x,y) \text{ and } (x+\Delta x, y+\Delta y) \text{ are within brain}} \tag{1}$$

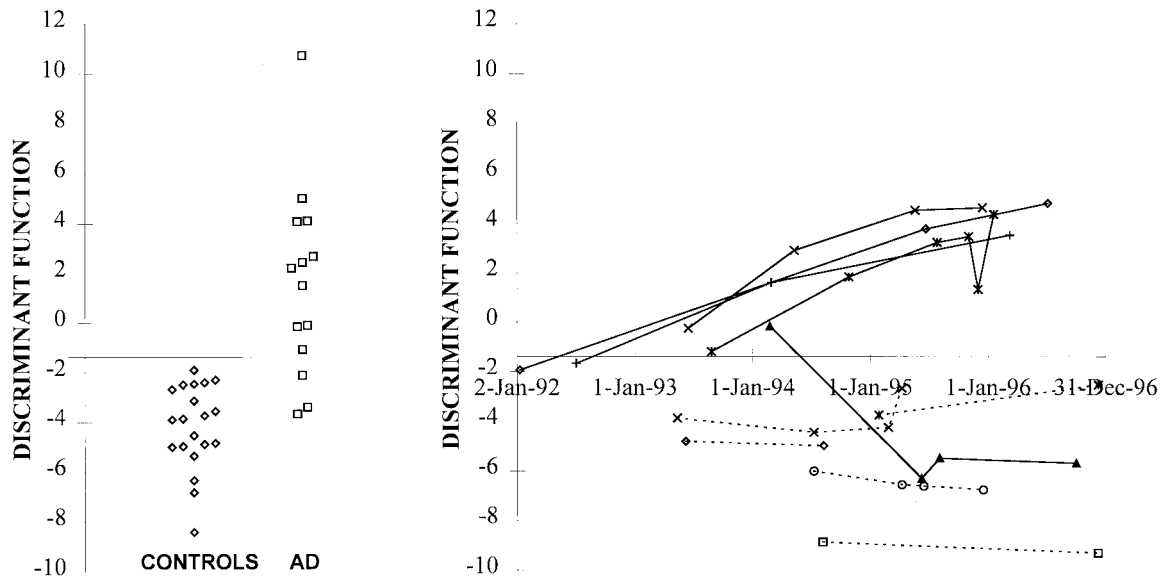

Fig. 2. (a) Evaluation of the texture discriminant function for each member of the test set and (b) changes in the value of the discriminant function over time for the five control subjects (dotted lines) and five AD patients (solid lines) who were scanned repeatedly over several years. In each case the x-axis is at the level of the classification decision boundary i.e., $D = -2\ \ln(2)$.

Fig. 2(b) shows changes in the value of the function over time for the five control subjects and five AD patients who were scanned repeatedly; this figure excludes measures derived from scans used in the testing or training sets. The median change in the discriminant function between successive scans of -0.15 for the control subjects was not significantly different from zero $(p > 0.05)$; whereas the median change of 1.4 for the AD patients was significantly different from zero $(p < 0.05)$.

IV. Discussion

The experiment focused upon assessing the value of MR image texture as a measure of change in Alzheimer's disease. Texture features were calculated over a large range of offset distances (1–32) as this allowed the "resolution" of the relevant texture differences to be automatically determined as part of the discriminant analysis. The data was split into training and test sets as the test set could then be utilized to assess the discriminant function in an unbiased manner.

The calculation of the texture discriminant function over the training set (Fig. 1) demonstrated that the function modeled the training data well. However, a discriminant function generally matches training samples better than samples from the general population and, therefore, this result does not indicate the true discriminatory power of the function. The true discriminatory power of the function is demonstrated by its evaluation over the test set [Fig. 2(a)]. In this set there is only small overlap between the control and AD groups; all control subjects were correctly classified, but three AD patients were misclassified. Two of the misclassified AD patients were in the early stages of the disease at the time of their scans; in these cases it is likely that the accumulated change since the onset of the disease was insufficient to bring the texture measure outside the normal range. The scan of the third misclassified patient suffered from significant motion related artifact, which may have influenced the texture measure. Misclassification may also reflect uncertainty in the "gold standard" clinical diagnosis; confirmation of this diagnosis will not be available until post-mortem. The overall classification rate of 91% suggests the texture discriminant function may be useful as a complementary marker to aid diagnosis.

For the five control subjects who were repeatedly scanned, the value of the discriminant function changed only by small amounts and in no consistent direction (i.e., the median change was not significantly different from zero); in all cases the function remained within the range of the test set control group [Fig. 2(b)]. In contrast, for the five AD patients who were repeatedly scanned the median change of 1.4 was found to be significantly greater than zero, indicating substantial and consistent increments over time. This suggests that the discriminant function is sensitive to the progression of the disease. It is notable that for four of the AD patients the discriminant function increased steadily with time, whereas in one case a downward and less consistent trend was observed—the discriminant function taking a value within the AD range on the first scan, but falling into the control range on the subsequent scans [Fig. 2(b)]. The source of this anomaly is not clear.

The discriminant analysis selected cooccurrence features calculated over a wide range of offset distances (see Table I). It is, therefore, unlikely that the discriminant function measures small scale features at the limit of the image resolution—this often being the motivation for the use of texture analysis. We instead suggest that the function measures large scale structure and image contrast. The way in which these features change in AD is complex and diffuse; the texture discriminant function attempts to summarize such changes in a single measure.

In this study we applied the discriminant analysis to 260 features using only 30 samples for training. The large number of degrees of freedom compared to the number of data samples, resulted in a discriminant function which was heavily influenced by "chance" differences between the groups in the training set, hence, the much tighter groupings in the training set (Fig. 1) compared to those in the test set [Fig. 2(a)]. We opted for this approach as the alternative of using only a few features (ten or fewer), would likely involve the arbitrary rejection of the most discriminating features. However, the sensitivity of the discriminant analysis to chance differences means that repeating the analysis on another training set is unlikely to yield a

discriminant function based on exactly the same set of features. This in no way undermines the experiment, because we chose to test the discriminant function on a completely independent set and, therefore, the assessment of the technique is not influenced by these issues.

This study has ignored the effects of texture changes associated with motion artifact. The development of methods for detecting motion artifact, and using these as the basis for a quality-control procedure, could further improve the robustness of this method.

The texture discriminant function developed here is based upon two-dimensional texture measures calculated over a 3-D structure. More concise and, therefore, effective descriptors may be obtained by developing 3-D measures of texture. Indeed, the generalization of SGLDM into 3-D is trivial. One might expect that a further improvement in discrimination may be obtained by basing the discriminant analysis on such 3-D measures.

V. Conclusions

We have demonstrated that a texture discriminant function derived from MRI brain scans, yields significantly different values for Alzheimer's disease patients compared to normal controls, and furthermore that progression of the disease over time is reflected in this measure. Such measures may be useful aids in the diagnosis and tracking of Alzheimer's disease.

Appendix
Derivation of Texture Features from a Cooccurrence Matrix

In this study we derived 13 texture measures (after [8]) from each cooccurrence matrix. We denote the cooccurrence matrix $C(i,j)$, and the number of distinct gray levels in the quantized image as N. We also define

$$C_x(i) = \sum_{j=1}^{N} C(i,j)$$

$$C_y(i) = \sum_{i=1}^{N} C(i,j)$$

$$C_{x+y}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{N} \sum_{j=1}^{N} C(i,j), \qquad k = 2, 3, \cdots 2N$$

$$C_{x-y}(k) = \sum_{\substack{i=1 \\ |i-j|=k}}^{N} \sum_{j=1}^{N} C(i,j), \qquad k = 0, 1, \cdots N-1$$

and μ_x, μ_y, σ_x, and σ_y as the means and standard deviations of C_x and C_y, respectively. We utilized the following 13 features:

1) Angular second moment:

$$f_1 = \sum_{i=1}^{N} \sum_{j=1}^{N} \{C(i,j)\}^2.$$

2) Contrast:

$$f_2 = \sum_{n=0}^{N-1} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N} \sum_{j=1}^{N} C(i,j) \right\}.$$

3) Correlation:

$$f_3 = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} ij\, C(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}.$$

4) Variance:

$$f_4 = \sum_{i=1}^{N} \sum_{j=1}^{N} (i-\mu)^2 C(i,j).$$

5) Inverse difference moment:

$$f_5 = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{1+(i-j)^2} C(i,j).$$

6) Sum average:

$$f_6 = \sum_{i=2}^{2N} i C_{x+y}(i).$$

7) Sum variance:

$$f_7 = \sum_{i=2}^{2N} (i-f_6)^2 C_{x+y}(i).$$

8) Sum entropy:

$$f_8 = \sum_{i=2}^{2N} C_{x+y}(i) \log\big(C_{x+y}(i)\big).$$

9) Entropy:

$$f_9 = -\sum_{i=1}^{N} \sum_{j=1}^{N} C(i,j) \log\big(C(i,j)\big).$$

10) Difference variance:

$$f_{10} = \text{variance of } C_{x-y}.$$

11) Difference entropy:

$$f_{11} = -\sum_{i=0}^{N-1} C_{x-y}(i) \log\{C_{x-y}(i)\}.$$

12) Information measures of correlation:

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

$$f_{13} = \sqrt{1 - \exp\{-2(HXY2 - HXY)\}}$$

where HX and HY are the entropies of C_x, and C_y and

13)

$$HXY = -\sum_{i=1}^{N} \sum_{j=1}^{N} C(i,j) \log\{C(i,j)\}$$

$$HXY1 = -\sum_{i=1}^{N} \sum_{j=1}^{N} C(i,j) \log\{C_x(i) C_y(j)\}$$

$$HXY2 = -\sum_{i=1}^{N} \sum_{j=1}^{N} C_x(i) C_y(j) \log\{C_x(i) C_y(j)\}.$$

References

[1] J. Wade, T. Mirsen, V. Hachinski, M. Fisman, C. Lau, and H. Merskey, "The clinical diagnosis of Alzheimer's disease," *Arch. Neurol.*, vol. 44, pp. 24–29, 1987.

[2] L. A. Klatka, R. B. Schiffer, J. M. Powers, and A. M. Kazee, "Incorrect diagnosis of Alzheimer's disease," *Arch. Neurol.*, vol. 53, pp. 35–42, 1996.

[3] L. R. Schad, S. Bluml, and I. Zuna, "MR tissue characterization of intracranial tumors by means of texture analysis," *Magn. Reson. Imag.*, vol. 11, pp. 889–896, 1993.

[4] R. A. Lerski, K. Straughan, L. R. Schad, D. Boyce, S. Bluml, and I. Zuna, "MR image texture analysis—An approach to tissue characterization," *Magn. Reson. Imag.*, vol. 11, pp. 873–887, 1993.
[5] C. Rossmanith, H. Handels, S. J. Poppl, E. Rinast, and H. D. Weiss, "Characterization and classification of brain tumors in three-dimensional MR image sequences," *Lecture Notes in Computer Science no. 1131 (Proc. VBC)*, K. H. Hohne and R. Kikinis, Eds., 1996, pp. 420–429.
[6] P. A. Freeborough, N. C. Fox, and R. I. Kitney, "Interactive algorithms for the segmentation and quantitation of 3-D MRI brain scans," *Comput. Meth., Programs in Biomed.*, vol. 53, pp. 15–25, 1997.
[7] K. H. Hohne and W. A. Hanson, "Interactive segmentation of MRI and CT volumes using morphological operations," *J. Comput. Assist. Tomogr.*, vol. 16, pp. 285–294, 1992.
[8] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 610–621, 1973.
[9] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. New York: Academic, 1988.
[10] C. W. Therrien, *Decision Estimation and Classification*. New York: Wiley, 1989.
[11] G. M. Clarke and D. Cooke, *A Basic Course in Statistics*. London, U.K.: Edward Arnold, 1992.

Magnetocardiographic Localization of Arrhythmia Substrates: A Methodology Study with Accessory Pathway Ablation as Reference

Pär L. Ågren,* Håkan Göranson, Tomas Hindmarsh, Evert Knutsson, Dag Mohlkert, Mårten Rosenqvist, and L. Bergfeldt

***Abstract*—In magnetocardiographic (MCG) localization of arrhythmia substrates, a model of the thorax as volume conductor is a crucial component of the calculations. In this study, we investigated different models of the thorax, to determine the most suitable to use in the computations.**

Our methods and results are as follows. We studied 11 patients with overt Wolff–Parkinson–White syndrome, scheduled for catheter ablation. The MCG registrations were made with a 37-channel "superconducting quantum interference device" system. The underlying equivalent current dipole was computed for the delta-wave. Three models of the thorax were used: the infinite halfspace, a sphere and a box. For anatomical correlation and to define the suitable sphere and box, magnetic resonance images were obtained. As reference we used the position of the tip of the catheter, at successful radio-frequency-ablation, documented by cine-fluoroscopy. Nine patients could be evaluated. The mean errors (range) when using the infinite halfspace, the sphere and the box were 96 (49–125), 21 (5–39), and 36 mm (20–58 mm), respectively $(p < 0.0001)$.

In conclusion, the sphere was significantly better suited than the other models tested in this study, but even with this model the accuracy of MCG localization must further improve to be clinically useful. More realistic models of the thorax are probably required to achieve this goal.

***Index Terms*—Biomagnetism, magnetocardiography, noninvasive localization, volume conductors.**

I. Introduction

The identification and localization of the underlying substrate is crucial when attempting to cure cardiac arrhythmias. So far, the methods used include 12-lead electrocardiography, body surface potential mapping, vectorcardiography, echocardiography (ECG), scintigraphic techniques, computed tomography (CT), magnetic resonance imaging (MRI), signal averaged ECG [1], and invasive peroperative epi- and endocardial mapping [2], [3]. These methods are, however, imprecise and as with endocardial mapping, also exposes the patient to potential hazards including X-rays. Cardiac electrical events create weak magnetic fields which are not significantly affected when passing through overlaying tissues. These magnetocardiographic (MCG) signals are, therefore, well suited for localization computations. The magnetic

Manuscript received November 21, 1997. This work was supported by the Swedish National Heart and Lung Foundation and the Karolinska Institute. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was R. L. Barbour. *Asterisk indicates corresponding author.*

*P. L. Ågren is with the Department of Cardiology, Karolinska Hospital, Karolinska Institute, Stockholm, S-791 82 FALUN, Sweden (e-mail: pl_agren@algonet.se).

H. Göranson is with the Department of Radiology, Karolinska Hospital, Karolinska Institute, Stockholm, S-791 82 FALUN, Sweden.

T. Hindmarsh is with the MR Center, Karolinska Hospital, Karolinska Institute, S-791 82 FALUN, Stockholm, Sweden.

E. Knutsson is with the Department of Neurophysiology, Karolinska Hospital, Karolinska Institute, Stockholm, S-791 82 FALUN, Sweden.

D. Mohlkert is with the Department of Thoracic Radiology, Karolinska Hospital, Karolinksa Institute, Stockholm, S-791 82 FALUN, Sweden.

M. Rosenqvist and L. Bergfeldt are with the Department of Cardiology, Karolinska Hospital, Karolinska Institute, Stockholm, S-791 82 FALUN, Sweden.

Publisher Item Identifier S 0278-0062(98)06444-1.