Report

An NMR Blood Test for Uveal Melanoma?

Esther V. Smith,* Nancy H. Kolodny,* §§ Evangelos S. Gragoudas,† Kathleen M. Egan,† Susan Finn,† Donald J. D’Amico,* † and Johanna M. Seddon†

In a recent article a simple nuclear magnetic resonance (NMR) blood test was suggested for the detection of the presence of cancer. The test’s sensitivity to uveal melanoma of both pre- and posttreatment status has been investigated. Cases in this study were 95 patients with uveal melanoma, and controls were 70 participants in an ongoing case control study of retinal eye disease being conducted at the Massachusetts Eye and Ear Infirmary. Proton NMR evaluations at 4.7 T (200 MHz) were performed on plasma obtained from EDTA and citrated blood samples. The average line-width values were calculated from each spectrum. Statistical analysis revealed that mean proton NMR line widths were essentially equal for patients treated (18.7 Hz) and untreated tumors (18.4 Hz) and for controls (18.5 Hz). Results based on this data set suggest that proton NMR spectroscopy has little predictive power in the detection of uveal melanoma or in the monitoring of therapy. Invest Ophthalmol Vis Sci 31:993–997, 1990

In a 1986 New England Journal of Medicine article which attracted national attention, Fossel et al described a method for the detection of malignant tumors based on high-resolution proton nuclear magnetic resonance (1H NMR) spectroscopy of human blood plasma.1 It was suggested that the diagnosis of cancer could be made by measuring the mean line width at half the peak height of the plasma lipid methyl and methylene proton resonances, as shown in Figure 1. At field strengths of 8.45 T (360 MHz) and 9.4 T (400 MHz) in their study, spectra of plasma from patients with untreated malignancies resulted in mean line width values of 29.9 ± 2.5 Hz. These values were significantly different from those of normal controls (39.5 ± 1.6 Hz, P < 0.0001), of patients with nontumor diseases (36.1 ± 2.6 Hz, P < 0.0001), and of patients with benign tumors (36.7 ± 2 Hz, P < 0.0001).1 This NMR blood test for cancer has become controversial, and many research groups have not been able to reproduce the findings of Fossel et al.2-9 In this paper we present the results of an investigation of the test’s sensitivity to uveal melanoma of both pre- and posttreatment status.

Materials and Methods. Cases in this study were 95 patients with uveal melanoma as diagnosed by indirect ophthalmoscopy, fluorescein angiography, and ultrasound at the Massachusetts Eye and Ear Infirmary. Patients with untreated tumors and tumors after treatment with proton beam irradiation were included. Controls were 70 participants in an ongoing case control study of retinal eye disease. Informed consent was obtained from all subject and control patients. A total of 325 blood samples was obtained, in most cases two from each subject. One was anticoagulated in a tube containing EDTA and the other in a tube containing citrate. Tubes were labeled only with an identification number to ensure that the NMR evaluations would be carried out without knowledge of the participant’s disease status. Within 1 hr of collection, each whole-blood specimen was centrifuged for 5 min at 3500 rpm. The plasma supernatant was then stored in a Nunc soft plastic tube, frozen at −10 to −18°C, and transported to a freezer located at the laboratory in which the 1H-NMR evaluations would be performed. Samples were stored frozen for up to 12 weeks.

Prior to NMR analysis, each frozen plasma sample was completely thawed and shaken, and approximately 0.5 mL was pipetted into a 5-mm NMR tube (Wilmad Glass). A smaller tube containing deuterium oxide for field frequency locking and shimming was inserted into the NMR tube containing the plasma. The time between thawing of the specimen and 1H-NMR spectroscopy never exceeded 120 min. All 1H-NMR evaluations were performed at Wellesley College on an NR/200 FTNMR spectrometer (IBM) equipped with a 4.7 T magnet (Bruker) and a variable temperature unit (Bruker). The temperature within the bore was maintained at 24°C throughout each experiment.

1H-NMR spectroscopy was performed at 200.133 MHz with the following parameters: sweep width, 2000 Hz; pulse width, 6 μsec; number of scans, 64; dummy scans, 4; recovery delay, 0 sec; decoupling power, 5 L; irradiation time, 2 sec; line broadening, 2 Hz. Shimming was optimized for each sample, and “Solvsup”, a continuous-wave solvent suppression program, was used to saturate the water peak. The two peaks of interest, the methyl and methylene proton resonance peaks of the plasma lipids, occurring at approximately 0.8 and 1.2 parts per million (ppm), respectively, were manually phased until the baseline became as horizontal as possible without distortion of the peak shapes. Each spectrum was expanded, scaled...
to 0.044 ppm/cm (8.809 Hz/cm) and plotted in the range 0.5–1.6 ppm. Baselines were drawn as shown in Figure 1. Lines at half the peak heights ($h_1$ and $h_2$) were drawn manually, (parallel to the baselines) and their widths ($W_1$ and $W_2$) were measured to the nearest 0.1 cm (0.9 Hz). The mean line width for each spectrum was calculated as described by Fossel et al.1

Spectra of 99 samples, approximately one third of the total collected, were run in triplicate. Because the triplicate line-width measurements for individual samples were found not to vary by more than 1 mm (0.9 Hz) when performed by one observer, only one spectrum was obtained for all subsequent samples. The amount of sample degradation based on the level of lactate detected in the shape of the methylene proton resonance peak at 1.2 ppm was described as either “none” (not present), “shoulder” (some present), or “prominent” (outstanding). Spectral shoulders are indicated by the curved arrows in Figure 1.

Initial data analyses revealed that the outcome measure, mean line width at half height, varied by degree of degradation (none, shoulder, prominent), with higher line-width readings among more degraded samples. It was also noted that samples derived from younger persons were more likely to be scored as having prominent degradation. Because degradation was a source of imprecision in the data and because age was a factor to be evaluated, degraded samples were excluded from analysis. Thirty-five of 67 controls (52%) and 37 of 85 (44%) cases were excluded on this basis. Assessment of degradation had been made without knowledge of the disease status of the participant.

For each of the remaining participants, an EDTA or citrated sample without degradation was available. Among the 36 participants for whom both samples were available without degradation, the values obtained from the citrated samples (mean = 18.3 Hz) and EDTA samples (mean = 18.1 Hz) were similar (paired t-test, $P = 0.45$). Therefore, the mean line width used in analyses for these individuals was the average of the values obtained from the citrated and EDTA samples. For the remaining individuals, the mean line width measurement was taken to be the value from the EDTA or citrated sample without evidence of degradation.

The influence of age on the outcome measure, mean line width, was evaluated in univariate analysis using Pearson's correlation coefficient and simple linear regression. Discrete data, including sex and case/control status, were evaluated using analysis of variance. Multiple regression was used to assess the relationship between case and treatment status, ad-
Table 1. Distribution of age, sex, and mean line width values by comparison group

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases, pretreatment</td>
<td>8 58.3 (43-69) 17.8 (14.1-21.6)</td>
<td>8 56.1 (34-68) 18.8 (16.3-23.8)</td>
<td>16 57.2 (34-69) 18.3 (14.1-23.8)</td>
</tr>
<tr>
<td>Cases, posttreatment</td>
<td>19 64.4 (45-80) 18.0 (15.9-25.5)</td>
<td>13 63.1 (40-78) 19.5 (18.1-22.9)</td>
<td>32 63.9 (40-80) 18.7 (15.9-25.5)</td>
</tr>
<tr>
<td>Controls</td>
<td>14 58.5 (38-71) 17.8 (15.9-19.4)</td>
<td>18 65.4 (52-79) 18.9 (15.9-24.7)</td>
<td>32 62.4 (38-79) 18.4 (15.9-24.0)</td>
</tr>
</tbody>
</table>

In each of the three groups (pretreatment, posttreatment, and controls), females had significantly greater line width (LW) values than males ($P < 0.01$). N indicates the number of persons in each category. The range is indicated in parentheses below each value of mean age (yr) and mean line width (Hz).

Results. After exclusion of participants for whom both EDTA and citrated samples were degraded, 80 participants remained in the study: 16 preirradiation cases, 32 postirradiation cases, and 32 controls. The distribution of cases and controls by age and sex is presented in Table 1. The mean age of the patients with cancer was 60 yr and of the controls was 64 yr. Overall, the three groups did not differ significantly in age ($P = 0.08$), although pretreatment cases were somewhat younger on average than the other two groups. There were approximately equal numbers of males and females by comparison group ($P = 0.45$), although somewhat more posttreatment cases were male and more controls were female.

To evaluate whether these differences could bias results, further analyses examined whether mean line-width measurements varied according to age and sex. Figure 2 demonstrates that participant age was not significantly correlated with mean line width ($r = 0.08$); the slope of the regression line indicates that mean line width increased less than 0.2 Hz with every decade ($0.017 \times 10$). Females had significantly higher values compared to males ($P < 0.01$), and this pattern was observed in each of the three comparison groups (Table 1).

Figure 3 depicts the distribution of mean line-width values in each group. With no adjustment made for age and sex, values were similar for pre- (18.3 Hz) and posttreatment (18.7 Hz) cases and for controls (18.4 Hz) ($P = 0.79$). When the outcome measure was dichotomized using the median value of all participants (18.5 Hz) as the cut point, 56% (9/16) of pretreatment cases, 50% (16/32) of posttreatment cases, and 50% (16/32) of controls had values below the median ($P = 0.91$), indicating no tendency for any of the three groups to have higher or lower mean line-width values.

Multiple regression analysis was used to determine whether case status, pre- or posttreatment, significantly predicts mean line width, controlling for age and sex. Of the four variables in the model, only gender had a statistically significant slope ($1.256/0.452 > 1.96, P < 0.05$); mean line widths were, on average, 1.3 Hz higher in females than in males, controlling for age and case status. The slope corre-
justing for age and sex, pretreatment cases had values 0.02 Hz lower than controls, on average, whereas posttreatment cases had values 0.45 Hz higher than controls, neither significantly different from zero.

Discussion. Analysis of the data obtained in this study demonstrates that proton NMR spectroscopy does not serve as a screening test for uveal melanoma. There are many possible explanations for this negative result, ranging from uncertainty in the design of the experiments conducted to questions about the molecular basis for observed line width variations.

The great excitement caused by the possibility of a simple NMR blood test for cancer induced many groups to attempt to reproduce Fossel's results. Many of these groups found technical and experimental difficulties with the Fossel test. A few groups obtained results which were exactly opposite to those of Fossel et al; for example, two studies resulted in mean spectral line widths for the blood plasma specimens from patients with tumors that were greater than those of healthy volunteers. Conducted in response to Fossel's argument that these experiments were not performed properly, another study revealed no statistically significant differences among the mean line widths, as is the case in our investigation.

We designed our experiments to eliminate some of the features of Fossel's test most subject to criticism: complications in the line-width measurement, such as those found in the interobserver variability in drawing the baseline and the distortion of the Lorentzian line shape caused by the presence of lactate, a product of degeneration resulting from sample storage at elevated temperatures. In evaluating samples stored at 4°C, Chmurny et al found that Fossel Indices (line-width values) increased by an average of 2.0 Hz after being stored for 1.5–2.5 months. In our study, the following precautions were taken: NMR evaluations were performed blinded with respect to the disease status of the patient; one observer made all measurements; samples demonstrating the presence of lactate were excluded from statistical analysis; and samples were frozen below −10°C. In addition, the use of two different anticoagulants were compared, and age and sex were evaluated as variables.

Our experiments were conducted on an NMR spectrometer operating at 200 MHz. This spectrometer was chosen, in part, because its cost ($175,000) makes it somewhat more feasible as a piece of clinical laboratory equipment than a spectrometer operating at 400 or 500 MHz and costing between $400,000 and $500,000. Fossel's group suggested that experiments conducted at such low frequencies might not yield significant line-width differences; however, our experiments at 200 MHz should show little difference from theirs at 250 MHz if the technique is accurate, and they do not. Furthermore, since line widths depend on chemical shift in cases where more than one component contributes to a peak (vide infra), and chemical shift is directly proportional to field strength, at 200 MHz we should have seen a mean line-width difference of about 5 Hz if Fossel's value of 10 Hz at 400 MHz is correct. We found no difference in methyl and methylene resonance mean line width of plasma samples from cases with treated or untreated uveal melanoma and controls with no uveal melanoma.

The width of an NMR peak caused by a single proton or a group of magnetically equivalent protons is generally considered to be inversely proportional to $T_2$, the proton's spin–spin relaxation time. Fossel's original premise, based on long $T_2$ values obtained for isolated cancer cell lipids, was that high resolution $^1H$-NMR experiment on human blood plasma should yield results in which narrower mean line widths of the methyl and methylene proton resonances could be interpreted as predicting malignancy. However, line widths may also reflect the presence of overlapping peaks from similar protons in different molecules in the sample. It has been established recently that various plasma components with a range of chemical shifts are contributing to the NMR signals of interest. This has been shown by Chmurny in a comparison of mean line width values as a function of magnetic field strength. As field strength increases, the lines broaden and separate into a series of peaks which become well-resolved by...
620 MHz. Thus, the widths of the methyl and methylene peaks cannot be interpreted solely in terms of $T_2$ values.

Before a simple NMR blood test for the diagnosis of cancer becomes accepted, it is important to identify and investigate the blood components that are being measured. The major criticism of the NMR blood test is that little attempt appears to have been made to determine what the NMR experiment actually detects and what is contributing to the spectra. A careful quantitative analysis of the lipids and lipoproteins present in the plasma, as well as a $T_2$ evaluation of each component, would help in understanding what is contributing to the spectra. Such analysis may predict which plasma conditions may yield false positives or negatives, as well as give insight as to the general applicability of the method.

In conclusion, these results suggest that proton NMR spectroscopy shows little promise as a screening method for uveal melanoma and has little or no ability to discriminate between pre- and posttreatment uveal melanoma patients. These results are in accord with those of at least eight other investigators who for various other cancers were unable to duplicate the findings of Fossel et al.

Key words: uveal melanoma, proton NMR spectroscopy, NMR, plasma, proton beam irradiation

From the *Howe Laboratory of Ophthalmology and the ‡Retina Service of the Massachusetts Eye and Ear Infirmary, Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts; the ‡Magnetic Resonance Imaging Facility, Massachusetts Institute of Technology, Cambridge, Massachusetts; and the ‡Department of Chemistry, Wellesley College, Wellesley, Massachusetts. Supported in part by the Massachusetts Lions Eye Research Fund, Inc. (Drs. Kolodny and D'Amico) and the Uveal Melanoma Fund. Submitted for publication: September 8, 1989; accepted October 4, 1989. Reprint requests: Evangelos S. Gragoudas, Retina Service, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA 02114.

References