Relative Risk of Alzheimer Disease and Age-at-Onset Distributions, Based on APOE Genotypes among Elderly African Americans, Caucasians, and Hispanics in New York City

Ming-Xin Tang,1,2 Gladys Maestre,2,7 Wei-Yann Tsai,4 Xin-Hua Liu,1 Lin Feng,7 Wai-Yee Chung,7 Michael Chun,5 Peter Schofield,5 Yaakov Stern,1,4,5,6 Benjamin Tycko,2,7 and Richard Mayeux1,2,3,5,6

1The Gertrude H. Sergievsky Center, 2The Taub Center for Alzheimer’s Disease Research in the City of New York, the Divisions of 3Epidemiology and 4Biostatistics, School of Public Health, and the Departments of 5Neurology, 6Psychiatry, and 7Pathology, Columbia University College of Physicians and Surgeons, New York

Summary

Apolipoprotein-E e4 (APOE-e4) has been consistently associated with Alzheimer disease (AD) and may be responsible for an earlier age at onset. We have previously reported a diminished association between APOE-e4 and AD in African Americans. Using a new method, which allows inclusion of censored information, we compared relative risks by APOE genotypes in an expanded collection of cases and controls from three ethnic groups in a New York community. The relative risk for AD associated with APOE-e4 homozygosity was increased in all ethnic groups (African American relative risk [RR] = 3.0; 95% confidence interval [CI] = 1.5–5.9; Caucasian RR = 7.3, 95% CI = 2.5–21.6; and Hispanic RR = 2.5, 95% CI = 1.1–5.7), compared with those with APOE-e3/e3 genotypes. The risk was also increased for APOE-e4 heterozygous Caucasians (RR = 2.9, 95% CI = 1.7–5.1) and Hispanics (RR = 1.6, 95% CI = 1.1–2.3), but not for African Americans (RR = 0.6, 95% CI = 0.4–0.9). The age distribution of the proportion of Caucasians and Hispanics without AD was consistently lower for APOE-e4 homozygous and APOE-e4 heterozygous individuals than for those with other APOE genotypes. In African Americans this relationship was observed only in APOE-e4 homozygotes. These results confirm that APOE genotypes influence the RR of AD in Caucasians and Hispanics. Differences in risk among APOE-e4 heterozygote African Americans suggest that other genetic or environmental factors may modify the effect of APOE-e4 in some populations.

Introduction

The presence of the e4 allele of the apolipoprotein-E gene (APOE) has been consistently associated with an increased risk of Alzheimer disease (AD) (Borgaonkar et al. 1993; Corder et al. 1993, 1994; Mayeux et al. 1993; Noguchi et al. 1993; Payami et al. 1993; Poirier et al. 1993; Saunders et al. 1993a, 1993b; Strittmatter et al. 1993a, 1993b; Chartier-Harlin et al. 1994; Kuusisto et al. 1994; Liddell et al. 1994; Roses et al. 1994; Sorbi et al. 1994; Hendrie et al. 1995; Maestre et al. 1995). However, it is apparent that the APOE-e4 allele alone is neither necessary nor sufficient to cause the disease, because (1) AD develops in the absence of APOE-e4 (Lannfelt et al. 1994; Maestre et al. 1995) and (2) some individuals escape the disease despite having an APOE-e4 allele. The presence of APOE-e2 has been associated with both a decreased (Corder et al. 1994; Chartier-Harlin et al. 1994) and an increased risk of AD (Lannfelt et al. 1994; Maestre et al. 1995; van Duijn et al. 1995).

Roses et al. (Roses et al. 1994; Roses 1995) have explained the variability by proposing that APOE genotypes influence the age at onset of AD. In both familial and sporadic AD, they observed an earlier age at onset among APOE-e4 homozygous and APOE-e4 heterozygous AD cases than among those AD cases with other APOE genotypes. These observations have needed confirmation, because they imply that APOE genotypes may predict specific age-at-onset distributions. Commercial tests based on the APOE allele system have become available, allowing physicians to estimate the risk of AD and the average age at onset, by APOE genotype.

There may be important limitations, however. The association between AD and APOE-e4 is less consistent for individuals >80 years of age (Poirier et al. 1993; Roses et al. 1994), when disease risk is highest (Evans et al. 1989; Gurland et al. 1995; Hebert et al. 1995), and the association may be stronger in women than in men (Payami et al. 1994). The relationship between AD...
and APOE-e4 in ethnic groups other than Caucasians has not been thoroughly examined. For example, an association between AD and APOE-e4 homozygosity, but not between AD and APOE-e4 heterozygosity, was observed in African Americans in New York (Maestre et al. 1995), whereas both APOE-e4 homozygosity and APOE-e4 heterozygosity were associated with AD among African Americans in Indiana (Hendrie et al. 1995).

In this investigation we developed a new method to estimate relative risk (RR) of AD among individuals, with incomplete observations based on APOE genotypes in an expanded elderly group of African Americans, Hispanics, and Caucasians residing in the Washington Heights community of New York City. We also used a previously published method (Turnbull 1974) to plot the age distribution of the proportion of individuals who remained free of AD, by APOE genotype, within each ethnic group.

**Subjects and Methods**

**Subjects and Setting**

All data were derived from individuals residing in the same community. AD cases were identified as part of a comprehensive community-based registry of all dementias in northern Manhattan. We used information (clinical and billing records, the computerized daily census, monthly discharge lists, and the daily logs) from a number of sources, medical as well as nonmedical: regional hospitals (including inpatient and outpatient services), private practitioners and the single neurologist unaffiliated with the Columbia Presbyterian Medical Center, federal and state health agencies offering assistance to individuals in the area, nursing homes and other health-related facilities, an HMO (Health Insurance Plan of Greater New York), and all senior-citizen centers in the community. A list of individuals with home—health-care attendants in the community was obtained from the Human Resources Administration in New York City. All individuals were screened to identify those with and without cognitive impairment, as described elsewhere (Gurland et al. 1995). Those with, and a proportion of those without, cognitive impairment were examined for the presence or absence of AD. Controls were volunteers recruited from the same community, but most were identified in senior-citizen centers and not in medical facilities. As with AD cases, all controls at each community site were informed about a study of aging and health and were asked to participate. All controls received interviews and clinical assessments identical to those used for AD cases. The refusal rate, for both AD cases and controls identified at these centers, was <20%.

AD cases and controls were also identified from a random survey of Medicare recipients in the same geographic area (the population >65 years of age was 9,349; United States Bureau of the Census 1991). The Health Care Financing Administration (HCFA) provided access to a random sample of the names and addresses of recipients within three contiguous ZIP codes in the community. Potential subjects were sent a letter from HCFA that explained that they had been randomly selected as candidates for a study of aging. A 90-min in-person interview was conducted for those who agreed to participate, followed by a complete medical and neurological assessment by a physician. Names of AD cases and controls identified in the survey and the registry were cross-referenced to eliminate duplicates. Participation rate was 68% in the HCFA-based random sample and did not differ by ethnic group or gender.

The Columbia University institutional review board reviewed and approved this project. All individuals provided written informed consent. Both AD cases and controls were followed over a 4-year period.

**Diagnosis of AD**

Physicians elicited the medical history and conducted a standardized physical and neurological examination. Whenever possible, all ancillary information, including medical charts and reports of laboratory studies, were used in the evaluation. A standardized neuropsychological battery measuring performance in memory, orientation, abstract reasoning, language, and construction was administered. Independence in activities of daily living was measured to determine whether subjects met functional criteria for dementia. Except for APOE genotypes, all clinical information was reviewed at a diagnostic conference of physicians and neuropsychologists, to arrive at a consensus diagnosis based on research criteria (McKhann et al. 1984; American Psychiatric Association 1987). The development of our diagnostic methods in relation to the cultural and educational demographics of this community has been described elsewhere (Pittman et al. 1992; Stern et al. 1992). The reliability and consistency of these assessments have also been reported elsewhere (Schofield et al., in press). The majority of AD cases were alive at the time of this investigation, but data were also available from 16 AD cases with postmortem confirmation of AD that were from the same community.

**Age at Onset**

Although all informants for prevalent AD cases were asked about the duration of symptoms, we did not attempt to estimate the age at onset in these individuals, because of the difficulty in substantiating this type of information. To improve validity, we censored (left-censored) the age at onset in prevalent AD cases, on the
basis of methods described elsewhere (Turnbull 1974; Cupples et al. 1991; Tang et al., in press). In brief, for each prevalent case, the EM (expectation and maximization) algorithm, as implemented by Turnbull, was used to compute the expected age at onset (see appendixes). This value was less than the patient’s age on entry into the study and was within an age interval based on the distribution of ages either (1) at study entry, for all younger prevalent AD cases, or (2) ages at onset, for all younger incident AD cases. In the age interval, the lowest age for Caucasians and Hispanics was 50 years, whereas in African Americans the lowest age was 59 years. We used age 50 years as the lowest age for computing the expected age at onset for all prevalent AD cases, regardless of ethnic group. For example, in a patient with AD who was 75 of age, the estimated age at onset would be 50–75 years. Several controls in the registry became demented during the 4-year follow-up. The age at onset for these AD cases was their age on the date at which dementia was first documented. For controls, we right-censored at their age at the last follow-up examination.

**Ethnic Group**

Ethnic group was classified by self-report using the format of the 1990 United States Bureau of the Census (1991). This allows each individual to identify himself as belonging to a particular racial or ethnic group. Each individual is then asked whether he or she is of “Hispanic” origin. Using this information, we separated subjects into three ethnic groups: African American, Hispanic, and Caucasian (non-Hispanic), according to self-report, on the basis of direct interview with the subject or, if the individual was demented, a family member.

**APOE Genotyping**

Genomic DNA was amplified by PCR and was subjected to CfoI restriction analysis using APOE primers and conditions modified from those described by Hixson and Vernier (1991). Genotypes were determined without knowledge of patient-control status.

**Data Analysis**

APOE allele frequencies were determined by counting alleles and calculating sample proportions. For comparison across ethnic groups, APOE allele frequencies were calculated by use of data from all controls and from a random 10% sample of the AD cases within each ethnic group, to create a representative population of individuals ≥65 years of age that would include patients with AD (Evans et al. 1989). For comparison of AD cases and controls, within and across ethnic groups, APOE allele frequencies were calculated for all subjects and were compared by $\chi^2$ analysis (Fleiss 1981).

We used an iterative method (i.e., the EM algorithm) to obtain the nonparametric maximum-likelihood estimates of the age-at-onset distributions, by APOE genotypes, within each ethnic group, from the “self-consistent equation” (Kaplan and Meier 1958; Turnbull 1974; present study, appendix A). A parametric approach was used to develop a proportional hazards regression model to estimate the RR for AD, by APOE genotype (refer to appendix B). This analysis was similar to the method used by Cupples et al. (1989, 1991). Both univariate and multivariate RRs for AD associated with various APOE genotypes were derived from the proportional hazards model.

**Results**

**Demographics**

Data were available for 305 patients with AD and for 485 nondemented healthy elderly. Some 55% of the patients with AD were identified through the registry, compared with 45% from the survey. Some 31% of the controls were volunteers from the community sites, and 69% were identified through the survey. Patients and controls differed by gender (76% of AD cases were women, vs. 67% of controls; $\chi^2 = 6.8, P < .01$). In all ethnic groups combined, AD cases were older (76.4 ± 9.1 years [range 46–96 years]) than controls (72.9 ± 6.7 years [range 46–93 years]; $P < .05$). African Americans and Hispanics were generally older and less well educated than Caucasians (table 1). African-American and Hispanic AD cases also had less education than controls (7.3 ± 5.0 years for AD cases, vs. 9.0 ± 4.6 years for controls; $P < .05$).

**APOE Allele Frequencies**

Overall, the APOE allele frequencies differed significantly among the three ethnic groups (data not shown; $\chi^2 = 19.7, P < .0005$). The main difference was in the frequency of APOE-e4 among African Americans, which was most pronounced among individuals >72 years of age (median age), in whom the frequency of APOE-e4 was .21, compared with .09 both in Caucasians and in Hispanics (table 2). APOE allele frequencies did not differ by gender.

Among AD cases, the distribution of APOE alleles also differed across ethnic groups ($\chi^2 = 11.9, \text{df} 4, P < .01$), with higher frequencies of APOE-e4 in Caucasians (27%) than in African Americans (23%) or Hispanics (19%) and with higher frequencies of APOE-e2 in African Americans (11%) than in the other two ethnic groups (Caucasians, 3%; and Hispanics, 7%), as noted in table 1. In Caucasians and Hispanics, but not in African Americans, the APOE-e4 allele frequency was significantly higher among AD cases than among controls.
When the groups were subdivided into age quartiles, however, APOE-e4 frequency in African-American AD cases in the youngest quartile (age ≤69 years) differed significantly from that in controls in the youngest quartile (42% in AD cases, vs. 15% in controls), because of a greater number of APOE-e4 homozygotes in this age group. In the three other quartiles, African-American AD cases and controls had nearly identical APOE-e4 frequencies.

**RR and the Age-at-Onset Distribution**

The impact of APOE genotypes on RR was calculated, and the cumulative age-at-onset distributions were plotted for each ethnic group. Specific APOE genotypes for AD cases and controls, for each ethnic group, are shown in table 3. African Americans homozygous for the APOE-e4 allele had a significantly increased risk of AD (RR = 3.0, 95% CI = 1.5–5.9) compared with African Americans with APOE-e3/e3 genotypes, but the risk associated with APOE-e4 heterozygosity was significantly decreased (RR = 0.6, 95% CI = 0.4–0.9). The proportions, by age and APOE genotype, of individuals without AD are shown in figure 1. We isolated individuals with APOE-e4 homozygosity and individuals with APOE-e4 heterozygosity, to show the differences between the age distributions in the proportion remaining unaffected, compared with other APOE genotypes.

Caucasians, both those APOE-e4 homozygous and those APOE-e4 heterozygous, had a significantly increased RR for AD (e4 homozygous RR = 7.3, 95% CI = 2.5–21.6; and e4 heterozygosity RR = 2.9, 95% CI = 1.7–5.1), compared with Caucasians with APOE-e3/e3 genotypes. Hispanics with one or more copies of the APOE-e4 allele also had higher RRs of AD (e4 homozygous RR = 2.5, 95% CI = 1.1–5.7; and e4 heterozygous individuals RR = 1.6, 95% CI = 1.1–2.3), compared with Hispanics with APOE-e3/e3 genotypes. Figures 2 and 3 illustrate the age distribution of the proportion of individuals without AD, comparing those with at least one APOE-e4 allele to those with other APOE genotypes, among Caucasians and Hispanics.

We also examined the protective effect of APOE genotypes with one or more e2 alleles, again using APOE-e3/e3 genotypes as the reference. The disease risk associated with APOE-e2 was lower among Caucasians (RR = 0.2, 95% CI 0.03–1.7), but not in African Americans (RR = 1.3, 95% CI = 0.8–2.3) or Hispanics (RR = 1.0,

**Table 1**

Demographics and APOE Allele Frequency among Patients with AD and among Controls, by Ethnic Group

<table>
<thead>
<tr>
<th>Ethnic Group (N)</th>
<th>Age (years)</th>
<th>Education (years)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>APOE-e3</td>
</tr>
<tr>
<td>African American</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD cases (106)</td>
<td>78 ± 7.6</td>
<td>7.4 ± 4.0</td>
<td>.66</td>
</tr>
<tr>
<td>Controls (154)</td>
<td>74 ± 5.8*</td>
<td>9.4 ± 3.8*</td>
<td>.73</td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD cases (59)</td>
<td>71 ± 9.1</td>
<td>13.4 ± 4.2</td>
<td>.70</td>
</tr>
<tr>
<td>Controls (112)</td>
<td>72 ± 8.5</td>
<td>12.3 ± 4.0*</td>
<td>.84</td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD cases (140)</td>
<td>77 ± 8.3</td>
<td>4.8 ± 3.7</td>
<td>.74</td>
</tr>
<tr>
<td>Controls (219)</td>
<td>73 ± 6.1*</td>
<td>7.0 ± 4.3*</td>
<td>.83</td>
</tr>
</tbody>
</table>

Note.—Allelic frequencies differed significantly between AD cases and controls, within and across ethnic groups; for explanation, see the text.

* Significant difference (P < .05) exists within each stratum.

**Table 2**

APOE Allele Frequency, by Age and Ethnic Group

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Frequency (No. Counted)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>APOE-e3</td>
</tr>
<tr>
<td>African American:</td>
<td></td>
</tr>
<tr>
<td>≤72 years of age</td>
<td>.72 (107)</td>
</tr>
<tr>
<td>&gt;72 years of age*</td>
<td>.74 (136)</td>
</tr>
<tr>
<td>Caucasian:</td>
<td></td>
</tr>
<tr>
<td>≤72 years of age</td>
<td>.84 (112)</td>
</tr>
<tr>
<td>&gt;72 years of age</td>
<td>.85 (87)</td>
</tr>
<tr>
<td>Hispanic:</td>
<td></td>
</tr>
<tr>
<td>≤72 years of age</td>
<td>.81 (193)</td>
</tr>
<tr>
<td>&gt;72 years of age</td>
<td>.86 (195)</td>
</tr>
</tbody>
</table>

Note.—To compare ethnic groups across ages, we used all controls recruited and a 10% random sample of cases from each ethnic group, to form a representative sample of elderly ≥65 years of age.

* $\chi^2 = 19.7$, df = 4, P < .0005. APOE allele frequencies differ in African Americans relative to other ethnic groups >72 years of age, the median age of the cohort.
95% CI = 0.5–1.7). Each analysis was repeated, with either stratification by gender or inclusion of educational level as a covariate; but the results were unchanged.

We also examined differences between the number of individuals either left- or right-censored and the number of those with complete information, by APOE genotype and ethnic group; there were no statistically significant differences, in the number of individuals, by APOE genotype, but incident AD cases were significantly more numerous among African Americans and Hispanics than among Caucasians (table 4).

Discussion

We examined the RR for AD, on the basis of APOE genotypes among three distinct ethnic groups residing in New York City, using a new method that included both left- and right-censored observations. Using this method, we also plotted the age distribution of the proportions of individuals within each ethnic group remaining free of AD, by APOE genotype. As expected, the APOE-e4/e4 genotype was associated with the highest RR for AD in all three ethnic groups, compared with individuals with the APOE-e3/e3 genotype. Among Caucasians and Hispanics, APOE-e4 heterozygosity was also associated with a significantly higher RR for AD, compared with individuals with the APOE-e3/e3 genotype. However, among African Americans, APOE-e4 heterozygotes had a significantly lower RR for AD, compared with individuals with the APOE-e3/e3 genotype. Stratification by gender or educational level did not change the associations between APOE-e4 and AD in any ethnic group. There was only minimal evidence that AD risk was reduced for Caucasians with APOE-e2 genotypes, and there was no decrease in risk for African Americans or Hispanics with these genotypes.

An association between APOE-e4 and AD has been reported among African Americans in Indiana (Hendrie et al. 1995). African Americans were also included in a North Carolina study (Saunders et al. 1993a), but the specific association between APOE-e4 and AD was not described. In both studies the APOE-e4 frequency was lower than that reported here; in Indiana the APOE-e4 allele frequency in controls was .12, whereas in North Carolina study the APOE-e4 frequency was .17 overall. No association between APOE-e4 and AD was found in a small study of elderly Nigerians (Osuntokun et al. 1995), where the e4 allele frequency in the control population was .26. Differences in APOE allele frequencies by ethnic group are well established in adults and children (Davignon et al. 1988; Kamboh et al. 1989; Sepehrnia et al. 1989; Hallman et al. 1991; Srinivasan et al. 1993; Sandholzer et al. 1995; Zekraoui et al. 1995). Regional differences in the frequency of APOE alleles among African populations have also been described (Sepehrnia et al. 1989; Sandholzer et al. 1995; Zekraoui et al. 1995). Zekraoui et al. (1995) observed the APOE-e4 frequency to vary from .03 among Wolof in Senegal to .38 among Tutsi in Burundi, and a similarly elevated e4 frequency (.4) was also observed in South African bushman (Sandholzer et al. 1995). The APOE-e4 allele frequency among the African Americans in the present study and others (Kamboh et al. 1989; Srinivasan et al. 1993) was 20%–26%. Thus, it is possible that the individuals included in our study were genetically distinct from those in the Indiana study. It is also important to note that the APOE-e4 frequency did not decline with age, in contrast to the trend noted among Caucasians in the United States and Europe (Kervinen et al. 1994; Louhija et al. 1994; Schachter et al. 1994).

The decline in the frequency of APOE-e4 with age in Caucasians has been attributed to mortality from heart disease (Eichner et al. 1993; Kervinen et al. 1994; Luc et al. 1994; Wilson et al. 1994; Stengard et al. 1995), because
Figure 1 Proportion of individuals without AD, among African Americans, by APOE genotypes. APOE genotypes are divided into three groups: APOE-e4 homozygosity, APOE-e4 heterozygosity, and all other APOE genotypes as the reference. We retained three groups only among African Americans, to demonstrate the different effects of APOE homozygosity and APOE heterozygosity.

variations in the APOE locus are associated with cholesterol and triglyceride concentrations (Davignon et al. 1988; Reilly et al. 1992; Louhija et al. 1994). For example, in young and middle-aged Finns the allele frequency of APOE-e4 is ~.2, and it decreases by almost 50% with age, because of heart disease (Kuusisto et al. 1994; Stengard et al. 1995), but a strong association between AD and APOE-e4 persists (Kuusisto et al. 1994).

In contrast, APOE-e4 does not consistently predict cholesterol or triglyceride concentrations in adults or children of African origin (Sepehrnia et al. 1989; Srinivasan et al. 1993; Garjra et al. 1994; Sandhoizer et al. 1995).

Figure 2 Proportion of individuals without AD, among Caucasians, by APOE genotypes. APOE genotypes are divided into two groups: those with any APOE-e4 and all other APOE genotypes as the reference. For the method used, see the text.
1995). Age-related changes in the population frequency of APOE-e4 in African Americans have not been thoroughly investigated, but the frequency in our study remained elevated among the individuals >80 years of age. This suggests that the relationship of APOE-e4 to heart disease, longevity, and AD in African Americans may not parallel the relationship in Caucasians. Alternatively, other genes or environmental factors, which are unique to African Americans, may have a greater influence on mortality and survival.

Given the consistent association between APOE-e4 heterozygosity and AD, seen both in the literature and among Caucasians and Hispanics in the present study, we considered the possibility of a systematic bias in our investigation of African Americans. The vast majority of African-American AD cases and controls were born

![Figure 3](image-url)  
**Figure 3** Proportion of individuals without AD, among Hispanics genotypes. APOE genotypes are divided into two groups: those with any APOE-e4 and all other APOE genotypes as the reference. For the method used, see the text.

Table 4

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>African American:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalent-AD cases</td>
<td>23 (.41)</td>
<td>10 (.179)</td>
<td>2 (.036)</td>
<td>2 (.036)</td>
<td>13 (.232)</td>
<td>6 (.107)</td>
</tr>
<tr>
<td>Not demented</td>
<td>77 (.50)</td>
<td>16 (.104)</td>
<td>1 (.006)</td>
<td>3 (.019)</td>
<td>55 (.357)</td>
<td>2 (.013)</td>
</tr>
<tr>
<td>Incident-AD cases</td>
<td>26 (.52)</td>
<td>6 (.12)</td>
<td>1 (.02)</td>
<td>0</td>
<td>13 (.26)</td>
<td>4 (.08)</td>
</tr>
<tr>
<td>Caucasian:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalent-AD cases</td>
<td>23 (.46)</td>
<td>1 (.02)</td>
<td>0</td>
<td>2 (.04)</td>
<td>21 (.42)</td>
<td>3 (.06)</td>
</tr>
<tr>
<td>Not demented</td>
<td>78 (.696)</td>
<td>14 (.125)</td>
<td>0</td>
<td>1 (.009)</td>
<td>19 (.17)</td>
<td>0</td>
</tr>
<tr>
<td>Incident-AD cases</td>
<td>7 (.778)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1 (.11)</td>
<td>1 (.11)</td>
</tr>
<tr>
<td>Hispanic:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalent-AD cases</td>
<td>40 (.556)</td>
<td>6 (.083)</td>
<td>1 (.014)</td>
<td>2 (.028)</td>
<td>21 (.292)</td>
<td>2 (.028)</td>
</tr>
<tr>
<td>Not demented</td>
<td>148 (.676)</td>
<td>29 (.132)</td>
<td>1 (.005)</td>
<td>0</td>
<td>38 (.174)</td>
<td>3 (.104)</td>
</tr>
<tr>
<td>Incident-AD cases</td>
<td>40 (.588)</td>
<td>6 (.088)</td>
<td>0</td>
<td>3 (.044)</td>
<td>15 (.221)</td>
<td>4 (.059)</td>
</tr>
</tbody>
</table>

**Note.**—There was a difference in the no. of individuals with incident AD, by ethnic group ($\chi^2 = 21.1$, df = 4, $P < .001$); Caucasians had fewer individuals.
in the United States and have resided in its northeastern region for most of their adult lives. The two methods used to identify patients with AD and controls (a community-based registry and a survey of a random sample of Medicare recipients) identified a representative population, both healthy and diseased, of African Americans, Caucasians, and Hispanics residing in the community. Despite the fact that this is one of the larger studies of APOE and AD, the infrequency of subjects with certain APOE genotypes, such as e2/e2, makes some of our conclusions tentative. We have previously investigated the methods used to establish the diagnosis of AD both among individuals from different cultures and among individuals with lower educational attainment (Pittman et al. 1992; Stern et al. 1992). Our methods appear to be as free of bias as is possible (Gurland et al. 1995). Two African Americans were among the postmortem AD cases included in this investigation; a single patient was APOE-e4/e4, and another was APOE-e2/e3. In the data analyses, we adjusted for gender and education as potential confounders, but this did not change our primary observations. Further, we have been unable to attribute the lower AD risk associated with APOE-e4 heterozygosity to comorbidity from heart disease (Chun et al. 1995).

For the estimation of RR, more complete information was available from African Americans and Hispanics than was available from Caucasians. Although the maximum-likelihood method does not depend on the presence of uncensored information, the estimates of RR would have improved if we had had more complete information.

Therefore, we consider it possible that some as yet unidentified genetic or environmental factor reduces the risk of AD among APOE-e4 heterozygous African Americans. Alternatively, genetic variations in other proteins that interact with apolipoprotein E directly or indirectly at the cellular level might account for the observations in the present study. Because APOE-e4 heterozygosity (primarily APOE-e4/e3) occurs in ~25%~30% of most populations in the United States and Europe (Davignon et al. 1988), understanding this protective effect, whether genetic or environmental, would seem an important point of investigation.

Acknowledgments

The authors wish to thank Marcel Kempowski, Adria Gonzales, John Shinin, Maria Gonzales, Harold Brown, and Jean Denaro for their technical assistance in this project. We also acknowledge the efforts of Drs. Barry Gurland, David Wilder, and Rafael Lantigua, from the Center for Geriatrics and Gerontology at Columbia University, in the design and conduct of the registry and Medicare survey. We also thank Drs. Karen Marder, Ruth Ottman, and Nicole Schupf for their help with the interpretation and preparation of the manuscript. The work was supported by Federal Grants AG07232, AG10963, AG08702, and RR00645 and by the Charles S. Robertson Memorial Gift for Alzheimer’s Disease Research (from the Banbury Fund). M.-X.T. is supported by Faculty Scholar Award 95-045 from the Alzheimer’s Disease and Related Disorders Association. The authors also wish to thank Drs. Nicole Schupf and Ruth Ottman for helpful critique of the manuscript.

Appendix A

Estimating the Survival Distribution of Age at Onset

Let X be the age at onset of AD, and let

\[ Y = \begin{cases} 
\text{age at first visit,} & \text{if it was prevalent AD} \\
\text{age at last visit,} & \text{if it was not AD} \\
\text{age at onset,} & \text{if it was newly diagnosed or incident AD, where age at onset is known.}
\end{cases} \]

We can express the available information on X by a pair of random variables \((Y, \Delta)\), where

\[ \Delta = \begin{cases} 
0, & \text{if } X < Y \text{ (left-censored)} \\
1, & \text{if } X > Y \text{ (right-censored)} \\
2, & \text{if } X = Y \text{ (uncensored)}.
\end{cases} \]

Our goal is to estimate the survival distribution function \(S(x) = 1 - F(x) = P(X > x)\) of X from a sample of \(n\) independent pairs \((y_i, \delta_i), i = 1, \ldots, n\).

Turnbull (1974) proved that the nonparametric maximum-likelihood estimate (NMLE) \(\hat{S}\) satisfies the following “empirical self-consistent equation”:

\[
\hat{S}(x) = \hat{S}_0(x) + \hat{S}_1(x) + \hat{S}_2(x) - \int_0^x \frac{\hat{S}(t)}{S(t)} dS_1(t) + \int_x^\infty \frac{1 - \hat{S}(x)}{1 - \hat{S}(t)} dS_0(t),
\]

where

\[
\hat{S}_i(x) = \frac{1}{n} \sum_{i=1}^n I(y_i > x, \delta_i = i), \quad i = 0, 1, 2,
\]

and the indicator function \(I(\cdot)\) is defined by

\[
I(y_i > x, \delta_i = i) = \begin{cases} 
0, & \text{if } y_i \leq x \text{ or } \delta_i \neq i \\
1, & \text{if } y_i > x \text{ and } \delta_i = i.
\end{cases}
\]
Therefore the NMLE $\hat{S}$ is also called a “self-consistent estimator.” We can use the EM algorithm to obtain the NMLE $\hat{S}$.

Appendix B

Regression Model—Parametric Approach

Regression models proposed in survival analysis often involve the assumption of proportional hazard functions; that is, $h(x, z) = r(z)h_0(x)$, where $h(x, z)$ is the hazard function for a subject with a $p$ vector measure of covariates $z$, $r(z)$ expresses the relationship between $z$ and regression parameter $\beta$ and can be interpreted as RR. $h_0(x)$ is the hazard function of the underlying failure-time distribution. In some AD cases $h_0(x)$ has a simple form, $h_0(x) = h$ (constant) if the underlying distribution is exponential, whereas $h_0(x) = \lambda x^\gamma$ for Weibull distribution (Lawless 1982, p. 16). The most commonly used proportion function is $r(z) = \exp(\beta z)$. The assumption can be empirically checked by using computer-graphic displays.

The parameters in the regression model are composed of two parts; one is the parameter $\beta$ involved in the proportion function, and the other one, $\theta$, characterizes the underlying distribution. The model parameter $\eta = (\theta, \beta)$ can be estimated by the maximum-likelihood method.

If it is assumed that there are no uncensored AD cases, the likelihood function can be written as

$$L(\theta, \beta) = \prod_{i=1}^{n} S(Y_i; \theta, \beta)^{s_i} [1 - S(Y_i; \theta, \beta)]^{1-s_i}.$$  

With the proportional hazard model, the score function, which is the first derivative of the log likelihood, is

$$Q_1(\theta, \beta) = \frac{\partial \log L(\theta, \beta)}{\partial \theta} = \sum_{i=1}^{n} \frac{\delta_i - S(y_i; \theta, \beta)}{1 - S(y_i; \theta, \beta)} \exp(\beta z_i) \frac{\partial \log S_0(y_i; \theta)}{\partial \theta};$$

$$Q_2(\theta, \beta) = \frac{\partial \log L(\theta, \beta)}{\partial \beta} = \sum_{i=1}^{n} \frac{\delta_i - S(y_i; \theta, \beta)}{1 - S(y_i; \theta, \beta)} \exp(\beta z_i) \log S_0(y_i; \theta) z_i.$$  

where $S_0$ is the underlying survival function. The maximum-likelihood estimate $\eta$ can be obtained by numerically solving the equation $Q = 0$. The asymptotic variance of $\eta$ is estimated by the inverse of the observed information matrix of second partials of the log likelihood for Newton's method (Isaacson and Keller 1966, p. 97). Note that, when $X$ has Weibull or exponential distribution (a special case of $\gamma = 1$), the proportion hazard model is $h(x, z) = \lambda x^\gamma \exp(\beta z)$. The score function can be simplified as

$$Q(\eta) = \sum_{i=1}^{n} \frac{\delta_i - S(y_i; \eta)}{1 - S(y_i; \eta)} \exp(\beta z_i) \xi_i,$$

where

$$S(y_i; \eta) = \exp[-\exp(\eta z_i)],$$

$$\eta = [\log(\lambda), \gamma, \beta],$$

$$\xi_i = (1, \log y_i, z_i).$$

The stated method can be easily modified for the data that include the uncensored AD cases.

References


Eichner JE, Keller LH, Orchard TJ, Grandits GA, McCallum


Tang et al.: APOE-Based Risk of Alzheimer Disease


