

Multiple Etiologies for Alzheimer Disease Are Revealed by Segregation Analysis

Valluri S. Rao,¹ Cornelia M. van Duijn,⁴ Linda Connor-Lacke,³ L. Adrienne Cupples,² John H. Growdon,³ and Lindsay A. Farrer^{1,2,3}

¹Department of Neurology, Boston University School of Medicine, ²Department of Epidemiology and Biostatistics, Boston University School of Public Health, and ³Department of Neurology, Harvard Medical School, Boston; and ⁴Department of Epidemiology and Biostatistics, Erasmus University, Rotterdam

Summary

We have evaluated several transmission models for Alzheimer disease (AD), using the logistic regressive approach in 401 nuclear families of consecutively ascertained and rigorously diagnosed probands. Models postulating no major gene effect, random environmental transmission, recessive inheritance, and sporadic occurrence were rejected under varied assumptions regarding the associations among sex, age, and major gene susceptibility. Transmission of the disorder was not fully explained by a single Mendelian model for all families. Stratification of families as early- and late-onset by using the median of family mean onset ages showed that, regardless of the model studied, two groups of families fit better than a single group. AD in early-onset families is transmitted as an autosomal dominant trait with full penetrance in both sexes and has a gene frequency of 1.5%. Dominant inheritance also gave the best fit of the data in late-onset families, but this hypothesis was rejected, suggesting the presence of heterogeneity within this subset. Our study also revealed that genetically nonsusceptible males and females develop AD, indicating the presence of phenocopies within early-onset and late-onset groups. Moreover, our results suggest that the higher risk to females is not solely due to their increased longevity.

Introduction

Alzheimer disease (AD) is a common dementing illness that affects >10% of the population > 65 years of age (Evans et al. 1989). Although head trauma, hypothyroidism, depression, parental age, and nonsmoking have been identified as possible risk factors (Breteler et al. 1992), biological mechanisms underlying these associations are largely unknown.

Several epidemiological surveys (reviewed by Breteler et al. 1992) and genetic studies (reviewed by St George-Hyslop et al. 1989) have suggested familial aggregation and multiple etiologies for AD. Recent studies demonstrate the existence of susceptibility genes for familial AD (FAD) on chromosome 14 (Mullan et al. 1992; St George-Hyslop et al. 1992; Schellenberg et al. 1992; Van Broeckhoven et al. 1992), chromosome 19 (Pericak-Vance et al. 1991; Borraonkar et al. 1993; Corder et al. 1993; Strittmatter et al. 1993), and chromosome 21 (Goate et al. 1991). If AD occurs in response to environmental factors only, then the incidence of AD is expected to vary geographically and demographically more widely than is observed in most studies (Rocca et al. 1991). Evidently, a strong positive family history and advanced age are the most consistently observed significant risk factors for AD.

Previous segregation analyses using the mixed model approach of Morton and MacLean (1974) indicated that AD in clinic-based (Farrer et al. 1991*b*) and population-based (van Duijn et al. 1993) patient series is transmitted as an autosomal dominant characteristic with reduced penetrance and a multifactorial component. However, the best fitting model in the Farrer et al. (1991*b*) study indicated that parents heterozygous for the disease allele transmit the disease more frequently than was expected under a Mendelian model. Van Duijn et al. (1993) were unable to resolve whether a cohort effect for heritability accounts for the evidence for a nonmajor gene component. Moreover, neither of these studies quantified susceptibility, considered random environmental models, or allowed for heterogeneity of disease transmission.

An earlier study on age at onset and lifetime risk for FAD suggested that the early-onset form is compatible with autosomal dominant inheritance, whereas there is a possibility of both genetic and shared environmental factors in the late-onset form (Farrer et al. 1990). Molecular genetic studies have revealed heterogeneity between early- and late-onset FAD (St George-Hyslop et al. 1990) and even among families with early-onset disease, by virtue of the fact that the trait is linked to chromosome 14 in some families (Mullan et al. 1992; St George-Hyslop et al. 1992; Schellenberg et al. 1992; Van Broeckhoven et al. 1992) and to chromosome 21 in others (Goate et al. 1991).

Received November 8, 1993; accepted for publication July 14, 1994.

Address for correspondence and reprints: Dr. Lindsay A. Farrer, Neurogenetics Laboratory, Boston University School of Medicine, 80 East Concord Street, Boston, MA 02118-2394.

© 1994 by The American Society of Human Genetics. All rights reserved.

0002-9297/94/5505-0019\$02.00

In order to quantify susceptibility and allow for heterogeneity of disease transmission, we have evaluated several transmission models for the disease, under various assumptions regarding the association of sex, age, and major gene susceptibility, by using the logistic regressive approach of Bonney (1984, 1986). The present study is based on >400 families, including the families studied by Farrer et al. (1991). The enlarged sample enabled more precise estimation of the gene frequency and phenocopy rates, exploration of models allowing for the interaction of disease liability, age at onset, and gender effects, and stratification of the data set to test for heterogeneity.

Subjects, Material, and Methods

Subjects

Diagnostic information and family history on 231 consecutively ascertained subjects who sought diagnostic evaluation of a memory disorder were collected between April 1986 and August 1990 at the Massachusetts Alzheimer Disease Research Center (MADRC). Although this is not a population-based sample, we believe that the clinical characteristics of patients with AD seen in the MADRC are identical to diagnosed cases throughout the U.S. For the purpose of this study, the important subject criteria are diagnostic accuracy and longitudinal follow-up. These are the strengths of our sample. Details of these cases and the methods of data collection were reported earlier (Farrer et al. 1991b).

Standard NINCDS/ADRDA (McKhann et al. 1984), NIA (Khachaturian 1985), and Multi-Institutional Research in Alzheimer Genetic Epidemiology (MIRAGE) (Farrer et al. 1994) research criteria were used for the diagnosis of AD. This devised AD rating scale incorporates the existing research criteria as well as a rating of the reliability of information and pertinent information regarding comorbidity. A study of reliability of diagnosis among MIRAGE sites also indicated that this scale is more appropriate for family studies and multicenter investigations (Farrer et al. 1994). Accordingly, all probands had extensive laboratory and neuropsychological examinations at the MADRC. Families were excluded if the proband had a history of head trauma, stroke, depressive illness, thyroid disorder, B12 or folate deficiency, or toxin exposure preceding onset of dementia, or a single cognitive deficit or lack of progression. Diagnoses in relatives were established by assessment of information obtained from interview of multiple informants, medical records including autopsy reports if available, death certificates, and nursing home records (Farrer et al. 1991b).

Upon follow-up of the 231 cases in our earlier study (Farrer et al. 1991b), the clinical diagnosis in 29 individuals changed from probable AD to possible AD or dementia other than AD, and hence these subjects were excluded, leaving 202 patients with probable AD. Information on

175 new cases of probable AD was collected in a similar manner by the MADRC between September 1990 and February 1993. An additional 24 families ascertained during the period of the original study in which the diagnosis in the proband was changed from possible AD to probable AD were also included, yielding a total sample of 401 families. AD was neuropathologically confirmed in 29 (90.6%) of the 32 cases that went to autopsy. Families of the three cases who did not meet autopsy criteria for AD were excluded from the analyses.

Statistical Methods

Segregation analysis was performed using the regressive logistic model for family data (Bonney 1984, 1986). AD was considered as a dichotomous trait with age-dependent penetrance. The major gene component was modeled as a diallelic locus. In this model, consider a variable Y for an individual that has a value of 1 if affected and 0 otherwise. Then, the odds of having the disease for the person can be defined as $\Pr(Y = 1)/\Pr(Y = 0)$. The logit (θ) (i.e., the logarithm of odds of having the disease) may be expressed as a function of disease phenotypes of the individual's spouse (Z_S), the father (Z_F) and the mother (Z_M) as follows: $\theta = \beta + \delta_S Z_S + \delta_F Z_F + \delta_M Z_M$, with Z 's taking on the values 0 or 1 for the individual if unaffected or affected, respectively. The coefficients δ_S , δ_F , and δ_M are the regressive familial components, and β is the risk of developing AD. The observed risk of an individual is modified depending on the affection status of parents and the spouse. Although we could have estimated δ_S , δ_F , and δ_M independently for affected and unaffected spouses, fathers, and mothers respectively, an adequate number of informative pairs of each type (spouse pairs, father-offspring, mother-offspring) was not available to estimate one or all the regressive familial coefficients. Hence, all regressive familial coefficients (δ_S , δ_F , and δ_M) in these analyses were fixed to zero and all genetic variation including sibling correlations was attributed to the major locus component. Spouses and offspring of probands were ignored in this study because the offspring are too young to express the disease, and hence these individuals provide little information on transmission of AD. Potential loss of information from spouses on the estimation of gene frequency and penetrance parameters was compensated by the incorporation of cumulative incidence data from the general population (see below). Random mating and Hardy-Weinberg equilibrium for population frequencies of genotypes were also assumed.

All families were ascertained through a single proband, and analyses were corrected for this ascertainment bias following the approach proposed by Elston and Sobel (1979). Furthermore, it was necessary to impose independent and unbiased estimates of cumulative incidence because the risk of AD in first-degree relatives in a clinic based sample is not comparable to the risk among persons

in the general population (Breitner et al. 1988; Farrer et al. 1989; van Duijn et al. 1993). The cumulative incidence of AD in the general population—estimated to be .09 for males and .145 for females, respectively, by age 80 years—was derived from population incidence data (Schoenberg et al. 1987; Kokmen et al. 1988). Cumulative incidence was extrapolated linearly to .11 for men and .20 for women at age 102 years, which was the oldest age observed in the data set.

Age at onset was assumed to follow a logistic distribution with age coefficient α and baseline parameter β . Ages at onset among AD cases and censoring ages (i.e., current age or age at death) of all subjects (including affected individuals) were incorporated in the estimation of α . Susceptibility was defined as the cumulative probability of being affected if one lives to the maximum age of 102 years. Two different models are biologically plausible, depending on the influence of genotypes on the individuals. The first model (model I) assumes β , the initial risk of an individual at birth, to be genotype specific with common susceptibility parameter γ , i.e., the genotype primarily influences the age at onset. The second model (model II) assumes common (to all genotypes) initial risk β with genotype dependent susceptibility γ , i.e., the genotype determines the risk of the disease. However, both models allow for sex dependence on α , β , and/or γ . Because consistently better likelihoods and meaningful parameter estimates were obtained using model II for equivalent models, results presented in this report were derived under the assumptions of model II only.

Segregation analysis was carried out by fitting several hypotheses—including dominant, recessive, additive, arbitrary major gene, sporadic, environmental, and general (unrestricted) models—to the data using the computer program REGTL of the SAGE package (Bailey-Wilson and Elston 1987). In these analyses, sex-dependent genotype susceptibilities were calculated (model II); cumulative incidence values for males and females were fixed at the above values; and no sex effect on β or α was assumed. In addition, because sex differences in the expression of AD may be mediated through age instead of through underlying susceptibility, models imposing sex dependence on α and β separately as well as jointly were derived. The hybrid maximization technique (Atwood et al. 1992) consisting of a few iterations of direct search followed by estimation by the variable metric method was employed to derive parameter estimates. Hypotheses were tested against a more general model using a likelihood ratio test in a hierarchical manner. The test statistic follows a χ^2 distribution with df equal to the difference in number of independent parameters of the two models under comparison. Likelihoods were also compared using Akaike's (1974) information criteria (AIC).

To test for differences in transmission between early-onset and late-onset AD, we stratified the families in three

Table 1**Mean Age at Onset among 401 Probands, by Sex and Onset Age Group**

ONSET AGE GROUP	MALES		FEMALES		TOTAL	
	No.	mean \pm sd	No.	mean \pm sd	No.	mean \pm sd
Early-onset ...	68	63.0 \pm 6.4	130	62.4 \pm 6.8	198	62.6 \pm 6.7
Late-onset	73	74.9 \pm 4.9	129 ^a	74.1 \pm 5.4	202	74.4 \pm 5.2
Total	141	69.1 \pm 8.2	259	68.2 \pm 8.5	400	68.5 \pm 8.4

^a Onset age in one woman was unknown, but the family was assigned to the late-onset group since the proband was 83 years old at the time of examination.

ways: (1) by age at onset of the probands using age 65 years as the cut-off; (2) by the mean age at onset among the probands; and (3) by the median of the mean onset ages of each family. Since there is no universally accepted cut-off age to distinguish early-onset from late-onset AD, we chose the median of the family mean onset ages because this stratification yielded a better balance of sample sizes within each group and has been shown to differentiate early-onset FAD from late-onset FAD (Farrer et al. 1990). Although, for each of these stratification methods, segregation analyses were carried out for the early-onset and late-onset groups separately, consistently better likelihoods were obtained using median as the stratifier as compared to procedures (1) or (2). Consequently, we report here the results obtained by using median of family means alone as the stratifier. For a given model, whether two groups of families fit better than a single group was compared by using the log likelihood value $-2 \ln L = -2 \ln L_{\text{total}} - (-2 \ln L_{\text{early}} - 2 \ln L_{\text{late}})$ which is assumed to follow a χ^2 distribution asymptotically with $df = df_{\text{early}} + df_{\text{late}} - df_{\text{total}}$. In performing the tests of heterogeneity, one additional df was added to adjust for the estimation of the age at onset used to stratify the families.

Results

The 401 probands (260 women and 141 men) had 2,112 first-degree relatives, of whom 188 (8.9%) were affected, and the mean age at onset among probands was 68.5 \pm 8.4 years (range 46–87 years). Additional onset age characteristics of the probands are given in table 1. Table 2 shows that there is an approximate threefold increase in the percentage of mothers affected with AD, as compared with fathers in the total sample. Affection status was unknown for 2% of the relatives. AD was observed 50% more frequently in mothers, fathers, and sisters of late-onset probands than in the group of relatives of early-onset probands.

The results of segregation analysis performed on 401 families are presented in table 3. The arbitrary major gene

Table 2
Distribution of First-Degree Relatives of Probands, by Affection Status

	Fathers	Mothers	Brothers	Sisters	Total
Overall:					
Affected	35	89	26	38	188
Unaffected	343	304	633	598	1,878
Unknown	23	8	8	7	46
% Affected	8.7	22.2	3.9	5.9	8.9
Early onset:					
Affected	13	34	11	14	72
Unaffected	172	161	286	280	899
Unknown	13	3	2	4	22
% Affected	6.6	17.2	3.7	4.7	7.2
Late onset:					
Affected	22	55	15	24	116
Unaffected	171	143	347	318	979
Unknown	10	5	6	3	24
% Affected	10.8	27.1	4.1	7.0	10.4

model with Mendelian transmission was rejected ($\chi^2_3 = 10.4$; $P = .02$) as compared to the general model. Furthermore, we found that τ_{AB} is significantly different from expected value of .5 ($\tau_{AB} = .17$; $\chi^2_1 = 10.5$; and $P < .005$)

indicating that a simple major gene model does not fit the data. All Mendelian hypotheses—dominant ($\chi^2_3 = 11.9$; $P = .04$), recessive ($\chi^2_3 = 29.9$; $P < .005$), additive ($\chi^2_3 = 29.4$; $P < .005$) and non-Mendelian hypotheses—no major gene ($\chi^2_8 = 132.2$, $P < .005$), sporadic ($\chi^2_2 = 121.2$; $P < .005$) and environmental ($\chi^2_3 = 132.2$; $P < .005$)—were rejected in favor of the general model. Notably, transmission from the heterozygote in the general model is \ll than 50%. These results indicate a more complex transmission mechanism (mixed or polygenic model) for AD or a possible presence of heterogeneity in the sample. Further, the estimated susceptibilities for male and female noncarriers (γ_{BB}) to develop AD are greater than zero across differing models, suggesting the existence of phenocopies or new mutations in some families.

Stratification by Age at Onset

Separation of families into early and late onset by the median of the family means for age at onset (71 years) gave a better fit ($P < .005$) for every model tested. Models for early- and late-onset families are presented in tables 4 and 5, respectively. In early-onset families, the arbitrary major gene model was not significantly different from the general model ($\chi^2_3 = 2.7$; $P = .47$) or the τ_{AB} relaxed model ($\chi^2_1 = 2.0$; $P = .22$) suggesting that a single major gene model

Table 3
Segregation Analysis of Alzheimer Disease in 401 Families

MODELS ^a	q_A	τ_{AA}^b	τ_{AB}^b	τ_{BB}^b	β	α	μ	σ^2	φ			δ			$-2 \ln L$	n_c^c	P
									γ_{AA}	γ_{AB}	γ_{BB}	γ_{AA}	γ_{AB}	γ_{BB}			
Non-Mendelian:																	
1. General1010	1.0 ^d	.1786	.0003	-14.08	.1698	82.93	114.10	1.0 ^d	1.0 ^d	.0198	1.0 ^d	.5731	.0	1,936.4	10	...
2. General (τ_{AA} and τ_{BB} fixed)1023	[1.0]	.1665	[.0]	-13.98	.1683	83.05	116.08	1.0 ^d	1.0 ^d	.0174	1.0 ^d	.5664	.0	1,936.4	8	...
3. No major gene	-14.78	.1888	78.27	92.32	.20231112	2,068.6	2	<.005
4. Sporadic0894	.1952 ^e	.1952 ^e	.1952 ^e	-15.10	.1905	79.28	90.64	.5782	.4549	.1495	.6575	.6217	.0059	2,057.6	8	<.005
5. Environmental0790 ^f	.0790 ^f	.0790 ^f	.0790 ^f	-14.78	.1888	78.27	92.32	.9056	.7000	.1117	.9000	.7175	.0014	2,068.6	7	<.005
Mendelian:																	
6. Arbitrary major gene0790	[1.0]	[.5]	[.0]	-14.45	.1732	83.42	109.69	1.0 ^d	1.0 ^d	.0663	1.0 ^d	.6733	.0120	1,946.9	7	.02
7. Dominant0733	[1.0]	[.5]	[.0]	-14.50	.1742	83.25	108.40	1.0 ^d	1.0	.0774	.7350	.7350	.0121	1,948.3	5	.04
8. Recessive3517	[1.0]	[.5]	[.0]	-15.09	.1842	81.90	96.95	1.0 ^d	.0927	.0927	.9112	.0	.0	1,966.3	5	<.005
9. Additive1135	[1.0]	[.5]	[.0]	-14.66	.1806	81.13	100.82	1.0 ^d	.5514	.1028	.9917	.4958	.0	1,965.8	5	<.005
Sex dependent:																	
10. α :																	
(φ)																	
(δ)																	
.....	.0839	[1.0]	[.5]	[.0]	-14.50	.1729	83.86	110.06	1.0 ^d	1.0 ^d	.0571	1.0 ^d	.6202	.0133	1,946.3	8	.48
11. β :																	
(φ)																	
(δ)																	
.....	.0819	[1.0]	[.5]	[.0]	-14.52	.1736	83.67	109.21	1.0 ^d	1.0 ^d	.0607	1.0 ^d	.6415	.0128	1,946.6	8	.71
12. α and β :																	
(φ)																	
(δ)																	
.....	.0886	[1.0]	[.5]	[.0]	-13.46	.1593	84.49	129.61	1.0 ^d	1.0 ^d	.0518	1.0 ^d	.5730	.0131	1,942.6	9	.14

^a Models 3-9 compared against model 1; models 10-12 compared against model 6.
^b Data in brackets ([]) = parameter fixed.
^c n_c = no. of independently estimated parameters = (no. of estimated - no. of dependent) parameters.
^d Fixed at the boundary by the maximization function.
^e $\tau_{AA} = \tau_{AB} = \tau_{BB}$.
^f $\tau_{AA} = \tau_{AB} = \tau_{BB} = q_A$.

Table 4

Segregation Analysis of Alzheimer Disease in Early Onset Families

MODELS ^a	q _A	τ _{AA} ^b	τ _{AB} ^b	τ _{BB} ^b	β	α	μ	σ ²	♀			♂			-2 ln L	n _c ^c	P
									γ _{AA}	γ _{AB}	γ _{BB}	γ _{AA}	γ _{AB}	γ _{BB}			
Non-Mendelian:																	
1. General	.0626	1.0 ^d	.1128	.0296	-15.04	.2042	73.67	78.89	1.0 ^d	1.0 ^d	.0902	1.0 ^d	.8982	.0011	753.5	10	...
2. General (τ _{AA} and τ _{BB} fixed)	.0487	[1.0]	.1150	[.0]	-15.05	.2057	73.15	77.73	1.0 ^d	1.0 ^d	.1166	1.0 ^d	.7662	.0408	754.2	8	.22
3. No major gene	-15.64	.2167	72.18	70.08	.20031102	772.6	2	.02
4. Sporadic	.0153	.0300 ^e	.0300 ^e	.0300 ^f	-15.63	.2159	72.43	70.61	.8000	.6200	.1872	.8900	.5982	.0949	772.3	8	<.005
5. Environmental	.0395 ^f	.0395 ^f	.0395 ^f	.0395 ^f	-15.64	.2167	72.18	70.08	.8100	.6600	.1615	.9000	.6000	.0686	772.6	7	<.005
Mendelian:																	
6. Arbitrary major gene	.0153	[1.0]	[.5]	[.0]	-15.18	.2046	74.18	78.60	1.0 ^d	1.0 ^d	.1757	1.0 ^d	1.0 ^d	.0826	756.2	7	.47
7. Dominant	.0153	[1.0]	[.5]	[.0]	-15.18	.2046	74.18	78.60	1.0 ^d	1.0	.1757	1.0 ^d	1.0	.0826	756.2	5	>.995
8. Dominant with τ _{AB} relaxed	.0457	[1.0]	.1192	[.0]	-15.04	.2056	73.19	77.86	1.0 ^d	1.0	.1221	.8283	.8283	.0399	754.4	6	.27
9. Recessive	.2880	[1.0]	[.5]	[.0]	-15.18	.2046	74.19	78.61	.7899	.1474	.1474	.7756	.0502	.0502	761.5	5	.07
10. Additive	.0524	[1.0]	[.5]	[.0]	-15.17	.2051	73.96	78.20	1.0 ^d	.5782	.1564	1.0 ^d	.5306	.0611	759.6	5	.25
Sex dependent:																	
11. α:																	
(♀)																	
(♂)																	
12. β:																	
(♀)																	
(♂)																	
13. α and β:																	
(♀)																	
(♂)																	

NOTE.—Notes b–f are as in table 3.

^a Models 3–5 compared against model 1; models 2, 7, and 9–13 compared against model 6; model 8 compared against model 7.

Table 5

Segregation Analysis of Alzheimer Disease in Late-Onset Families

MODELS ^a	q _A	τ _{AA} ^b	τ _{AB} ^b	τ _{BB} ^b	β	α	μ	σ ²	♀			♂			-2 ln L	n _c ^c	P
									γ _{AA}	γ _{AB}	γ _{BB}	γ _{AA}	γ _{AB}	γ _{BB}			
Non-Mendelian:																	
1. General	.1043	1.0 ^d	.1083	.0	-18.36	.2196	83.61	68.22	1.0	1.0	.0072	1.0	.5409	.0	1,110.3	10	...
2. General (τ _{AA} and τ _{BB} fixed)	.1066	[1.0]	.1096	[.0]	-18.49	.2207	83.79	67.56	1.0 ^d	1.0 ^d	.0021	1.0 ^d	.5280	.0	1,109.6	8	<.005
3. No major gene	-19.22	.2383	80.65	57.93	.20121107	1,224.4	2	<.005
4. Sporadic	.0005	.1050 ^e	.1050 ^e	.1050 ^e	-19.64	.2408	81.58	56.74	1.0	.7005	.2010	1.0	.8361	.1101	1,213.6	8	<.005
5. Environmental	.0377 ^f	.0377 ^f	.0377 ^f	.0377 ^f	-18.37	.2276	80.72	63.51	.9899	.6001	.1692	.9001	.6429	.0680	1,224.7	7	<.005
Mendelian:																	
6. Arbitrary major gene	.0937	[1.0]	[.5]	[.0]	-18.87	.2242	84.15	65.43	1.0 ^d	1.0 ^d	.0304	1.0 ^d	.5961	.0024	1,120.6	7	.02
7. Dominant	.0876	[1.0]	[.5]	[.0]	-18.94	.2254	84.03	64.78	1.0 ^d	1.0	.0431	.6603	.6603	.0015	1,122.6	5	NA ^g
8. Recessive	.3734	[1.0]	[.5]	[.0]	-18.82	.2262	83.18	64.29	1.0 ^d	.0737	.0737	.8000	.0	.0	1,138.7	5	NA ^g
9. Additive	.1205	[1.0]	[.5]	[.0]	-18.47	.2238	82.52	65.66	1.0 ^d	.5466	.0933	.9246	.4623	.0	1,143.6	5	NA ^g
Sex dependent:																	
10. α:																	
(♀)																	
(♂)																	
11. β:																	
(♀)																	
(♂)																	
12. α and β:																	
(♀)																	
(♂)																	

NOTE.—Notes b–f are as in table 3.

^a Models 3–6 compared against model 1; models 2 and 10–12 compared against model 6.

^g NA = not applicable. Since model 6 is significantly different from model 1, comparisons of models 7, 8, and 9 with model 6 are not applicable.

adequately fits the data. Further evidence for a single major gene in this group was provided by the rejection of no major gene ($\chi^2_8 = 19.1$; $P = .02$), sporadic ($\chi^2_2 = 18.8$; $P < .005$) and environmental ($\chi^2_3 = 19.1$; $P < .005$) models. Although dominant ($\chi^2_2 = .0$; $P > .995$), recessive ($\chi^2_2 = 5.3$; $P = .07$), and additive ($\chi^2_2 = 3.4$; $P = .25$) models could not be rejected in comparison with the arbitrary major gene model in early onset families, the dominant model has the lowest AIC value (766.2). In late-onset families, the arbitrary major gene model with Mendelian transmission was rejected as compared with the general model ($\chi^2_3 = 10.3$; $P = .02$) and the τ_{AB} relaxed model ($\tau_{AB} = .11$; $\chi^2_1 = 11.0$; and $P < .005$). Further, Mendelian hypotheses—dominant ($\chi^2_3 = 11.9$, $P = .04$), recessive ($\chi^2_3 = 29.9$; $P < .005$), and additive ($\chi^2_3 = 29.4$; $P < .005$)—and non-Mendelian hypotheses—no major gene ($\chi^2_8 = 114.1$; $P < .005$), sporadic ($\chi^2_2 = 103.3$; $P < .005$) and environmental ($\chi^2_3 = 114.4$; $P < .005$)—were also rejected among the tested models. These results indicate that there are differences in transmission mechanisms across different onset groups and also suggest possible heterogeneity within late-onset families.

Sex-dependent Age at Onset Parameters

In order to ascertain whether the effect of an individual's sex on expression of AD is mediated through the genotype only, models allowing for sex differences in the age at onset distribution (α) and/or baseline risk common to all genotypes (β) were also evaluated. Analyses were carried out under conditions identical to the ones used without sex dependence on α or β . In the total group of families, none of the models that iterated α or β were significantly better than the corresponding models that assumed no sex effect on α or β (table 3). The conclusions were unchanged when analogous comparisons were made in early- (table 4) and late-onset (table 5) families, with the exception of α and β iterated simultaneously in the early-onset families, which showed slight improvement. These findings suggest that the association between sex and genotype (through sex-specific susceptibilities, γ_s) accounts for most of the sex-related variation in expression of AD.

Discussion

This study of AD among 401 consecutively ascertained and rigorously diagnosed cases in a clinic-based population confirms our previous findings (Farrer et al. 1991b; van Duijn et al. 1993) that transmission of the disorder cannot be explained fully by a single Mendelian transmission model. However, our findings extend and in some ways differ from the earlier studies. Like Farrer et al. (1991b) and van Duijn et al. (1993), we rejected models postulating no major gene effect, recessive inheritance, and sporadic occurrence. A random environmental transmission model was also tested, and it too was rejected. All Mendelian models were rejected, although the parameters of the general model resemble the dominant model.

Our findings appear to be robust, because similar results were obtained from analyses of data sets collected over different time periods in identical fashion (results not shown). However, the conclusions are predicated on the assumption that the major gene influences one's susceptibility to the disease but not age at onset. Our findings are also dependent on the accuracy of the cumulative-incidence estimates derived from epidemiological studies. For example, estimations of cumulative incidence from our pedigree data (results not shown) rendered similar conclusions regarding preference of models; however, several parameter estimates—most notably gene frequency—were unrealistic because the disease was determined to be more than twice as frequent in males and females in these families than in the general population (Schoenberg et al. 1987; Kokmen et al. 1988). Conversely, the assumption that the disease is rarer in the population than in first-degree relatives of AD probands led to better discrimination of the dominant model over the recessive model (Farrer et al. 1991b; present study).

Age-at-Onset Heterogeneity

Although dominant inheritance may explain the transmission of AD among the tested Mendelian models (table 3), the general model, which has a definite non-Mendelian character, is significantly better. One explanation for this observation is that there exists a mixture of families with different modes of transmission. In an attempt to resolve this question, we stratified the families by age at onset, a parameter that has been postulated to discriminate AD cases etiologically (Breitner et al. 1988; Farrer et al. 1990; St George-Hyslop et al. 1990). Because there is no universally accepted cut-off between early- and late-onset disease, we applied several approaches to stratify the families. Segregation analysis performed separately for families classified according to the mean onset age of affected members in the family also showed that the disease is transmitted differently in early- and late-onset AD. Regardless of the genetic or nongenetic models studied, two groups of families always fit better than a single group.

The major gene for early-onset AD appears to be fully penetrant in both sexes by age 102 years and has a frequency of 1.5%. The dominant model also suggests that in these families genetically nonsusceptible males and females develop the disease at rates of $\sim 8.3\%$ and $\sim 17.6\%$, respectively. Such persons may be phenocopies, which can be due to multifactorial inheritance, new mutations, inheritance of a second major gene, or environmental insult. It is noteworthy that the additive model in this group could not be rejected, suggesting a possible alternate mechanism for AD. In contrast, although dominant inheritance was the best explanation for transmission of AD in late-onset families, all models were rejected in comparison to the general model suggesting the possibility of heterogeneity in this group. These findings are consistent with the con-

clusions from a study of age at onset and risk of AD in families ascertained for genetic linkage studies (Farrer et al. 1990). These investigators observed that risk to first-degree relatives of early-onset cases was $\sim 50\%$ (consistent with dominant inheritance), whereas risk to relatives of late-onset probands exceeded 85% (consistent with heterogeneity).

Our results may prove unsatisfactory if AD in a substantial proportion of cases follows a more complex genetic mechanism, such as oligogenic inheritance. This explanation is compatible with the observation of reduced transmission from the heterozygote in a single-gene model. A multifactorial inheritance model for AD has been proposed (McGuffin et al. 1991), but our nuclear family data proved insufficient to estimate the regressive familial coefficients from which a multifactorial or a polygenic model can be deduced (Demenais et al. 1992). Consequently, all regressive familial components (δ_S , δ_F , and δ_M) were fixed to zero and complex genetic models (e.g., polygenic, multifactorial, and mixed models) could not be tested in a meaningful way. Further, despite the availability of methodology for studying two-locus qualitative traits (Elandt-Johnson 1970), differentiation between these models and single gene models will be difficult until the theory is extended to late-onset disorders and implemented in computer programs.

Do Women Really Have an Increased Risk?

The literature regarding sex differences in risk for AD is controversial. Some studies suggest that AD is more prevalent in women because more women survive to an age when one is likely to become affected (Kay et al. 1964; Treves et al. 1986; Schoenberg et al. 1987; Farrer et al. 1989). This argument implies that if men and women had an equal life expectancy, then there would be no sex difference in risk. It is also possible, however, that men and women respond differently to the same risk factor (Nee et al. 1987; Breitner et al. 1988; van Duijn et al. 1993). In this case, risk is expected to be higher in women at all ages. Our analysis afforded an opportunity to explore this question more precisely than previous studies because we simultaneously estimated sex-specific parameters for age at onset (α and β) and underlying susceptibility (γ). Our results demonstrate that the sex effect is significant when varying the γ s, but not α and β , in the total group as well as in early- and late-onset groups separately. These findings favor the hypothesis that women are innately more susceptible to AD than are men. One explanation may be that men and women differ in exposure to environmental risk factors such as smoking and head trauma.

Limitations of Segregation Analysis for Resolving the Genetics of AD

Uncertainty in diagnosis continues to pose significant problems in genetic studies of AD. Even with rigorous

diagnostic procedures, the maximum accuracy of diagnosis based on autopsy confirmation is $\sim 90\%$ (Joachim et al. 1988; present study). In an attempt to improve diagnostic certainty, we have developed and tested in a multicenter study a new diagnostic rating scale for AD that incorporates existing research diagnostic criteria (Farrer et al. 1994). Using this approach, patients with variant phenotypes such as AD with parkinsonism or AD with vascular dementia may be distinguished. Future genetic studies might exploit the opportunity afforded by the regressive model approach, to consider the disease as a polychotomous trait. Of the few studies which have examined reliability of diagnosis in relatives (Heston et al. 1981; Pericak-Vance et al. 1988), secondary cases of dementia detected by interview of multiple informants and review of supporting documentation proven by autopsy are nearly always AD.

Diagnostic accuracy is substantially worse in unexamined relatives beyond the immediate family (Farrer et al. 1991b). Therefore, we limited our analysis to nuclear family data. Consequently, familial components (δ s), which account for resemblance not due to the major gene, could not be estimated in our analysis, due to lack of information about transmission from spouses of probands; this information would become available only after follow-up of their children when they surpass the critical age period for that family for developing AD.

Even with our large data set, Mendelian models could not be estimated very easily or discriminated from each other in some cases, because the likelihood surface was relatively flat. One explanation is that a much larger data set is required. To overcome this difficulty and study more complex models of disease transmission, we are currently assembling a data set of appropriate magnitude in our MIRAGE study (Farrer et al. 1994). Alternatively, the censoring bias associated with this disorder and/or heterogeneity in the mode of transmission may mask certain genetic patterns, regardless of the size of the sample. This study made an earnest attempt to address both of these issues and suggests that future analyses would benefit by an advanced methodology allowing more effective testing of oligogenic and heterogeneity models.

Correspondence between Segregation Analysis and Specific Risk Factors

Evidence for dominant transmission of AD in early-onset families is consistent with the demonstration of linkage of a susceptibility locus for FAD to chromosome 14 in early-onset families (Mullan et al. 1992; St George-Hyslop et al. 1992; Schellenberg et al. 1992; Van Broeckhoven et al. 1992). Although the dominant model is a better fit statistically and is supported by molecular genetic studies on chromosomes 14 and 21, the additive model seems to be more consistent with the observation that the ϵ -4 allele of apolipoprotein E (ApoE) has a significant dose-dependent

relationship with susceptibility and age at onset of late-onset AD (Corder et al. 1993). While this factor as documented in late-onset families (Corder et al. 1993; Noguchi et al. 1993; Payami et al. 1993; Poirier et al. 1993; Saunders et al. 1993; Strittmatter et al. 1993) is not fully predictive, it is possible that the ϵ -4 allele of ApoE may act as an additive or codominant major gene in some early-onset AD families without mutations in either chromosome 14 or 21 (Borgaonkar et al. 1993). A screen of ApoE in a large cohort of early-onset subjects suggested that the association between ϵ -4 and AD is much stronger among affected individuals having a positive family history (van Duijn et al. 1994), however additional studies using methods for distinguishing likely genetic cases (Farrer and Cupples 1994) are needed to assess this hypothesis.

Because our analyses strongly indicate the existence of multiple etiologies for AD, one may speculate that heterogeneity between early-onset and late-onset AD corresponds to subgroups of patients having defects in the chromosome 14 AD gene or the amyloid precursor protein (APP) gene on chromosome 21 (early onset) and ApoE (early onset and late onset). However, until the chromosome 14 gene is identified, it is not possible to determine whether most families have a defect in this gene, because they lack sufficient power for linkage analysis. Screening for APP mutations would not be efficacious either, because these are extremely rare (Tanzi et al. 1992). In contrast to autosomal dominant inheritance of AD cases linked to chromosomes 14 and 21, having the ϵ -4 allele is apparently not deterministic, but rather it is a risk factor for manifestation of AD. Thus, it is not feasible to distinguish with certainty most AD cases linked to ApoE. An alternative approach would be to evaluate the segregation of ApoE alleles in AD families, but this requires data on relatives of the proband. Unfortunately, in a clinic-based or population-based sample, parents and affected siblings, who provide most genetic information, are often deceased when the proband is ascertained. However, it is still possible to investigate the influence of ApoE on transmission of AD by stratifying families according to the ApoE genotype of the proband. We are evaluating the ApoE genotypes in patients currently followed in MADRC. Of the 401 cases in this study, we have typed 57 probands. The proportion of affected first-degree relatives in early-onset families did not differ among probands with and without at least one ϵ -4 allele. A comparable proportion was observed among late-onset probands with ϵ -4. However, late-onset probands without ϵ -4 had very few affected relatives. These findings are consistent with our conclusion from segregation analysis of etiological heterogeneity among late-onset AD. A much larger set of families will be needed to have sufficient power for testing this hypothesis more formally by segregation analysis.

One major strength of our analysis was the ability in all circumstances to rule out the environment alone as re-

sponsible for transmission of AD in these families. However, it is still plausible that certain nongenetic risk factors may cause AD in some persons or exacerbate the risk for disease in a individual who is genetically susceptible. Investigation of more complex models that include covariates such as paternal age (Farrer et al. 1991a), smoking (believed to confer a protective effect) (van Duijn et al. 1991), and head trauma (Mayeux et al. 1993) needs to be further pursued, which may help delineate important gene-environment interactive mechanisms.

Acknowledgments

We thank Drs. Richard H. Myers and Joan E. Bailey-Wilson for their valuable discussions, and the anonymous reviewers for their critical comments. We also thank Dan K. Kiely for technical assistance and Jemma Williams for preparing the manuscript. This work was supported by NIH grants R01-AG09029 and P50-AG05134. L.A.F. was supported by a fellowship from Alfred P. Sloan Foundation. The results reported in this paper were obtained by using the program package SAGE, which is supported by U.S. Public Health Service Resource grant RR03655 from the Division of Research Resources.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automatic Control* 19:716-723
- Atwood LD, Wilson AF, Elston RC, Bailey-Wilson JE (1992) Computational aspects of fitting a mixture of two normal distributions using maximum likelihood methods. *Commun Stat Simul Comput* 21:769-781
- Bailey-Wilson JE, Elston RC (1987) *Statistical Analysis for Genetic Epidemiology* (SAGE). LSU Medical Center, New Orleans
- Bonney GE (1984) On the statistical determination of major gene mechanisms in continuous human traits: regressive models. *Am J Med Genet* 18:731-749
- (1986) Regressive logistic models for familial disease and other binary traits. *Biometrics* 42:611-625
- Borgaonkar DS, Schmidt LC, Martin SE, Kanzer MD, Edelsohn L, Growdon J, Farrer LA (1993) Linkage of late-onset Alzheimer disease with apolipoprotein E type 4 on chromosome 19. *Lancet* 342:625
- Breitner JCS, Silverman JS, Mohs RC, Davis KL (1988) Familial aggregation in Alzheimer's disease: comparison of risk among relatives of early and late-onset cases and among male and female relatives in successive generations. *Neurology* 38:207-212
- Breteler MMB, Claus JJ, van Duijn CM, Launer LJ, Hofman A (1992) Epidemiology of Alzheimer's Disease. *Epidemiol Rev* 14:59-82
- Corder EH, Saunders AM, Strittmatter WJ, Schmechel D, Gaskell P, Small GW, Roses AD, et al (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261:921-923
- Demnais FM, Liang AE, Bonney GE (1992) Numerical comparisons of two formulations of the logistic regressive models

- with the mixed model in segregation analysis of discrete traits. *Genet Epidemiol* 9:419-435
- Elandt-Johnson RC (1970) Segregation analysis for complex modes of inheritance. *Am J Med* 22:129-194
- Elston RC, Sobel E (1979) Sampling considerations in the gathering and analysis of pedigree data. *Am J Hum Genet* 31:62-69
- Evans DA, Funkenstein HH, Albert MS, Scherr PA, Cook NR, Chown MJ, Hebert LE, et al (1989) Prevalence of Alzheimer's disease in a community population of older persons. *JAMA* 262:2551-2556
- Farrer LA, Cupples LA (1994) Estimating the probability for major gene Alzheimer disease. *Am J Hum Genet* 54:374-383
- Farrer LA, Cupples LA, Blackburn S, Kiely D, Auerbach S, Growdon JH, Connor L, et al (1994) Interrater agreement for diagnosis of Alzheimer disease: the MIRAGE study. *Neurology* 44:652-656
- Farrer LA, Cupples LA, Connor L, Wolf PA, Growdon JH (1991a) Association of decreased paternal age and late-onset Alzheimer disease: an example of genetic imprinting? *Arch Neurol* 48:599-604
- Farrer LA, Myers RH, Connor L, Cupples LA, Growdon JH (1991b) Segregation analysis reveals evidence of a major gene for Alzheimer disease. *Am J Hum Genet* 48:1026-1033
- Farrer LA, Myers RH, Cupples LA, St George-Hyslop PH, Bird TD, Rossor MN, Mullan MJ, et al (1990) Transmission and age at onset patterns in familial Alzheimer's disease: evidence for heterogeneity. *Neurology* 40:395-403
- Farrer LA, O'Sullivan DM, Cupples LA, Growdon JH, Myers RH (1989) Assessment of genetic risk for Alzheimer's disease among first-degree relatives. *Ann Neurol* 25:485-493
- Goate A, Chartier-Harlin M-C, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, et al (1991) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349:704-706
- Heston LL, Mastri AR, Anderson E, White J (1981) Dementia of the Alzheimer type: clinical genetics, natural history, and associated conditions. *Arch Gen Psychiatry* 31:1085-1090
- Joachim CL, Morris JH, Selkoe DJ (1988) Clinically diagnosed Alzheimer's disease: autopsy results in 150 cases. *Ann Neurol* 24:50-56
- Kay DWK, Beamish P, Roth M (1964) Old age mental disorders in Newcastle upon Tyne. Part I. A study of prevalence. *Br J Psychiatry* 110:146-158
- Khachaturian ZS (1985) Diagnosis of Alzheimer's disease. *Arch Neurol* 22:1097-1105
- Kokmen E, Chandra V, Schoenberg BS (1988) Trends in incidence of dementing illness in Rochester, Minnesota, in three quinquennial periods, 1960-1974. *Neurology* 38:975-980
- McGuffin P, Sargeant M, Weppner G (1991) The genetics of Alzheimer's disease and the ethical implications for prevention. In: Sram R, Bulyzhenkov V, Prilipko L, Christen Y (eds) *Ethical issues of molecular genetics in psychiatry*. Springer, Berlin, pp 42-56
- McKhann G, Drachmann D, Folstein M, Katzman R, Price D, Stadlan D (1984) Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer Disease. *Neurology* 34:939-945
- Mayeux R, Ottman R, Tang M, Noboa-Bauza L, Marder K, Gurland B, Stern Y (1993) Genetic susceptibility and head injury as risk factors for Alzheimer disease among community-dwelling elderly persons and their first-degree relatives. *Ann Neurol* 33:494-501
- Morton NE, MacLean CJ (1974) Analysis of family resemblance. III. Complex segregation of quantitative traits. *Am J Hum Genet* 26:489-503
- Mullan M, Houlden H, Windelspecht M, Fidani L, Lombardi C, Diaz P, Rossor M, et al (1992) A locus for familial early-onset Alzheimer disease on the long arm of chromosome 14, proximal to the α 1-antichymo-trypsin gene. *Nature Genet* 2:340-342
- Nee LE, Eldridge E, Sunderland T, Thomas CB, Ktaz D, Thompson KE, Weingartner H, et al (1987) Dementia of the Alzheimer type: clinical and family study of 22 twin pairs. *Neurology* 37:359-363
- Noguchi S, Murakami K, Yamada N (1993) Apolipoprotein E genotype and Alzheimer's disease. *Lancet* 342:737
- Payami H, Kaye J, Heston LL, Bird TD, Schellenberg GD (1993) Apolipoprotein E genotype and Alzheimer's disease. *Lancet* 342:738
- Pericak-Vance MA, Bebout JL, Gaskell PC Jr, Yamaoka LH, Hung W-Y, Alberts MJ, Walker AP, et al (1991) Linkage studies in familial Alzheimer disease: Evidence for chromosome 19 linkage. *Am J Hum Genet* 48:1034-1050
- Pericak-Vance MA, Yamaoka LH, Haynes CA, Speer MC, Haines JL, Gaskell PC, Hung W-Y, et al (1988) Genetic linkage studies in Alzheimer's disease families. *Exp Neurol* 102:271-279
- Poirier J, Davignon J, Bouthillier D, Kogan S, Bertrand P, Gauthier S (1993) Apolipoprotein E polymorphism and Alzheimer's disease. *Lancet* 342:697-699
- Rocca WA, Hofman A, Brayne C, Breteler MMB, Clarke M, Copeland JRM, Dartigues J-F et al (1991) Frequency and distribution of Alzheimer's disease in Europe: A collaborative study of 1980-1990 prevalence findings. *Ann Neurol* 30:381-390
- St George-Hyslop PH, Haines JL, Farrer LA, Polinsky RJ, Van Broeckhoven C, Goate A, Crapper-McLachlan D, et al (1990) Genetic linkage studies suggest that Alzheimer's disease is not a single homogeneous disorder. *Nature* 347:194-197
- St George-Hyslop PH, Haines JL, Rogaev E, Mortilla M, Vaula G, Pericak-Vance M, Foncin J-F, et al (1992) Genetic evidence for a novel familial Alzheimer disease locus on chromosome 14. *Nature Genet* 2:330-334
- St George-Hyslop PH, Myers RH, Haines JL, Farrer LA, Tanzi RE, Koji A, Chaker A, et al (1989) Familial Alzheimer's disease: progress and problems. *Neurobiol Aging* 10:417-425
- Saunders AM, Strittmatter WJ, Schmechel D, St George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, et al (1993) Association of apolipoprotein E allele E4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43:1467-1472
- Schellenberg GD, Bird TD, Wijsman EM, Orr HT, Anderson L, Nemens E, White JA, et al (1992) Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. *Science* 258:668-671
- Schoenberg BS, Kokmen E, Okazaki H (1987) Alzheimer's disease and other dementing illnesses in a defined United States population: incidence rates and clinical features. *Ann Neurol* 22:724-729
- Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M,

- Enghild J, Salvesen GS, Roses AD (1993) Apolipoprotein E: High avidity binding to β -amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci USA* 90:1977-1981
- Tanzi RE, Vaula G, Romano DM, Mortilla M, Huang TL, Tupler RG, Wasco W, et al (1992) Assessment of amyloid β -protein precursor gene mutations in a large set of familial and sporadic Alzheimer disease cases. *Am J Hum Genet* 51:273-282
- Treves T, Korczyn AD, Zilber N, Kahana E, Leibowitz Y, Alter M, Schoenberg BS (1986) Presenile dementia in Israel. *Arch Neurol* 43:26-29
- van Broeckhoven C, Backhovens H, Cruts M, Winter GD, Bruyland M, Cras P, Martin J-J (1992) Mapping of a gene predisposing to early-onset Alzheimer's disease to chromosome 14q24.3. *Nature Genet* 2:335-339
- van Duijn CM, de Knijff P, Cruts M, Wehnert A, Havekes LM, Hofman A, Broeckhoven CV (1994) Apolipoprotein E4 allele in a population-based study of early-onset Alzheimer's disease. *Nature Genet* 7:74-78
- van Duijn CM, Farrer LA, Cupples LA, Hofman A (1993) Genetic transmission for Alzheimer disease among patients identified in a Dutch population based survey. *J Med Genet* 30:640-646
- van Duijn CM, Hoffman A (1991) Relation between nicotine intake and Alzheimer disease? *Br Med J* 302:1491-1494