iMedPub Journals

JOURNAL OF NEUROLOGY AND NEUROSCIENCE

2010
Vol.1
No. 3:1
doi: 10:3823/310

Inheritance Of Alzheimer´s Disease Investigated By Complex Segregation Analysis

Manzano S, MD, PhD[1], Baron M, MD, PhD[2], Gomez-Tortosa E, MD, PhD[3], Barquero MS, MD, PhD[†1], Jimenez-Escrig A, MD, PhD[4].

1 Hospital Clinico San Carlos, Madrid, Spain.
2 Fundación Jimenez Diaz, Madrid, Spain.
3 Fundación Hospital Alcorcon, Madrid, Spain.
4 Hospital Ramon y Cajal, Madrid, Spain.

Address correspondence to:
Dr. Jimenez-Escrig S. de Neurologia
Hospital Ramon y Cajal 28034 Madrid, Spain
Email: adriano.jimenez@hrc.es

Acknowledgement:
Genetics research in our laboratory is done thanks to a grant research of the Fundacion Areces

Abstract

Background: *Complex segregation analysis (CSA) consists in the mathematical modeling of the hereditability of a transmissible condition. After generating a model, it can be known the most likely pattern of transmission, the frequency of the gene in that population and the penetrance of the condition.*

Objective: *To assess the inheritance for Alzheimer's disease in a Spanish population by CSA.*

Methods: *We ascertained 21 families (297 individuals) through probands, with 76 individuals affected with Alzheimer's disease fulfilling CERAD criteria. These families gave a total of 44 nuclear families to be included in the model. CSA was performed using the software POINTER examining the following models: non transmission, multifactorial (polygenic and environmental), Mendelian (dominant, recessive, codominant), polygenic, mixed (Mendelian plus polygenic) and a general model (Mendelian plus multifactorial). Four liability classes where defined according to the age of onset of the disease (<60 year-old; 60-69; 70-79; >80). Hypothesis testing was performed by comparing the fit of the specific model to the general unrestricted model.*

Results: *The model that best fitted the data in this population was the Mendelian dominant model with a gene frequency of 0.0164. This gene explains a 65.7% of the hereditability of this condition. Penetrance of the gene according to age followed an exponential pattern (2.47; 25.44; 27.88; 32.22).*

Conclusions: *Alzheimer's disease in these families is inherited due to a Mendelian dominant gene. The results support the importance of linkage efforts by suggesting that a Mendelian locus is segregating within a proportion of families with Alzheimer's disease ascertained through probands.*

Introduction

Alzheimer Disease (AD) is the most common neurodegenerative disease. This disease shares with other neurodegenerative diseases that following ageing, family history is the second risk factor for the disease. The growing understanding of AD genetics is being the key to the knowledge of the pathogenic mechanism driving to the disease.

Familial aggregation was recognized as a prominent characteristic in many neurodegenerative disorders decades ago (Bertram and Tanzi, 2005b). After the molecular genetic (Martin, 1999) and biochemical properties of these diseases have been unravelled, one of their characteristics which has emerged is the dichotomy between familial (rare) and seemingly non-familial (common) forms (sporadic or idiopathic) that is present in the genetic epidemiology of neurodegenerative diseases. Familial forms (Gail Pairitz J., 1998) have Mendelian patterns of transmission, while in seemingly sporadic forms a growing body of evidence suggests influence of multiple genetic traits that may associate an interaction with environmental factors. In AD, there are three rare fully penetrant autosomal dominant forms caused by mutations in APP (Goldgaber et al., 1987), PSEN1 (Barinaga, 1995) and PSEN2 (Levy-Lahad et al., 1995) genes, and a common

2010
Vol.1
No. 3:1
doi: 10:3823/310

incompletely penetrant susceptibility variant, namely, the ε4 allele in APOE gene (Chartier-Harlin et al., 1994), that significantly increases the risk by lowering the age of onset (AO) of the disease (Bertram and Tanzi, 2005a).

Familial aggregation in a disease does not necessarily imply a genetic etiology. When familial cases appear, genetic and/or environmental factors may be influencing the observed pattern of disease transmission in families. The genetic factors may be Mendelian with any mode of inheritance, polygenic, or any mixture of these ones. Various methods have been proposed for the statistical inference of gene effects in familial data. When examining a family with a certain disease present in several members, the issue is whether a genetic component or an environmental factor is the primary responsible for the trait. The simplest way to determine the genetic contribution to a trait is by examining the recurrence risk ratios. The most popular method is due to Risch (Risch, 1990) and is defined by

$$\lambda_R = k_R/k$$

where R denotes the relationship with the proband, k_R is the prevalence in relatives of type R, and k is the prevalence in the general population. In any genetic model

$$1 \leq \lambda_1 \leq \lambda_s \leq \lambda_M$$

where M, s and 1 are relationship subscripts that denote MZ twins, siblings and parents (or offspring) respectively. Typically, λ_R is calculated for siblings and λ_s is known as the *sibling relative risk*. Examples of λ_s for different diseases include Huntington's disease (where $k_s = 0.5$, $k = 0.0001$, and so $\lambda_s \approx 5{,}000$), recessive CMT (where $k_s = 0.25$, $k = 0.004$, and so $\lambda_s \approx 500$) and Parkinson's disease (where $k_s = 0.3$, $k = 0.1$, and so $\lambda_s = 3$). In general, the greater the value of λ_s, the greater the genetic influence on the trait. However, in itself, λ_s is not necessarily a reliable parameter for estimating the power of a proposed linkage study. For example, in some two locus models a λ_s as high as 10 does not guarantee that underlying genes will be easily mapped by linkage studies.

The power to detect genetic influence of a variant can also be defined in terms of *genotype relative risks (GRR's)*((Schaid and Sommer, 1993). Consider a biallelic locus with alleles of type A, a and relative frequencies $f(A)$, $f(a)$, where A is the disease susceptibility allele. The conditional probabilities that an individual with a particular genotype has a disease D are known as *penetrance* parameters and given by

$$f_{AA} = P(D|AA),\ f_{Aa} = P(D|Aa),\ f_{aa} = P(D|aa)$$

The genotype relative risks for D at this locus are

$$g_1 = f_{Aa} / f_{aa}\ ;\ g_2 = f_{AA} / f_{aa}$$

The relationship between the sibling relative risk ratio and genotype relative risks depends on both allele frequency and mode of inheritance (Rybicki and Elston, 2000) Explicit formulae relating *GRR* and λ_s for a dominant, recessive, additive and multiplicative models may be found in Wittke-Thompson et al. (Wittke-Thompson et al., 2005).

When discussing the heritability of a trait is worth to consider that there are two different measures that may be both referred to as heritability (Abney et al., 2001). Heritability in the broad sense (denoted H^2) is defined as the proportion of total variance in a trait that is due to all genetic components (additive, dominance and epistatic), while narrow heritability (denoted h^2) is defined as the proportion of phenotypic variance that can be attributed to additive genetic variance. The additive genetic variance at a locus measures the variance due to the mean effects of single alleles. Dominance variance of a trait at a locus measures the variance due to the interaction of alleles that constitute a genotype. Epistatic variance is due to the interaction effect between loci. Total additive (respectively, dominance) variance is the additive (respectively, dominance) variance at each locus summed over the genome. Similarly, total epistatic variance is the total variance obtained by summing the contribution of epistatic variance of all pairs of loci over the genome. Typically, one assumes that the additive effects are the primary contributors to the trait. A heritability score near zero suggests that almost all variation is due to environmental causes, whereas a heritability score near 1 implies that almost all variation is due to genetic factors.

It is important to bear in mind that heritability is a ratio and as such does not necessarily provide an accurate measure of how important genes are in determining the phenotype. Heritability reflects the proportion of total variation due to a gene variant, reflecting both the variant's frequency in the population and the size of the effects that the gene variant causes and is primarily used for assessing the genetic contribution to a quantitative trait. Sibling relative risk, on the other hand, assesses the increased disease risk to siblings that share one-half of their genes with affected probands and is used in connection with qualitative traits. For a fixed value of λ_R the corresponding heritability decreases with decreasing population prevalence (Risch, 2001).

A major point when considering the hereditability of a tract is the evaluation of the segregation pattern. Simple segregation analysis considers the proportion of affected and not affected in the offspring and examines this proportion against the theoreti-

2010
Vol.1
No. 3:1
doi: 10:3823/310

cal proportion of autosomal dominant (50%) or recessive (25% of affected / 75% of non-affected, in the simplest case) and considering the confidence intervals discloses whether a particular mode of transmission is possible or can be ruled out. A more general method for evaluating the transmission of a trait within pedigrees is complex segregation analysis (CSA), which test the fitting of the inheritance of the trait to different models, genetic and non genetic, allowing to select the model that obtains better fitting of the data. Whilst simple segregation analysis only evaluates whether the proportion of affected and unaffected offspring in families is consistent with Mendelian expectations, CSA can consider more complicated patterns of transmission and environmental perturbations. CSA can be applied to any pedigree structure and works with both qualitative and quantitative traits.

The parameters estimated in CSA are: 1) an underlying discrete risk trait (that may be present in double dose (*AA*), one dose (*Aa*) or absent (*aa*)) that influences a given individual's age-dependent risk for disease (in genetic models, this trait represents a high-risk allele, whereas in non-genetic models, the trait is interpreted more generally as levels of exposure to an unmeasured major environmental risk factor); 2) the transmission parameters which represent the probability that a parent transmits the risk trait to an offspring; and 3) the penetrance of the risk trait. CSA can also be used to further define the genetic features of a trait, such as the high risk allele frequency in the population. In addition, it can be used to evaluate etiologic heterogeneity in a trait, either by doing CSA in defined subsets or by contrasting the likelihoods under competing models for each family.

The mixed model, which is the one we have used here (another possibility for CSA is a regressive logistic model for disease (Bonney, 1986)) assumes that the liability to the disease (x) can be described by an underlying continuous liability scale in which a biallelic single major locus (g), a polygenic component (c), and environmental effects (e) operate independently. The liability (x) is then defined as $x = g + c + e$. The respective variances of these parameters are denoted as $V = G + C + E$. The relative contribution of the polygenic component is defined by H, the heritability, which reflects genetic transmission not ascribed to a major gene or cultural transmission ($H = C/V$).

Model parameters in the mixed model are:

A major locus has two alleles (*A*,*a*), whose genotype frequencies have to follow the Hardy-Weinberg equilibrium.

q, the frequency of the high risk allele *A*;

t, the genetic distance or displacement at the single major locus measured in standard deviations on the liability scale between the two homozygous genotypes (*AA* and *aa*);

d, degree of dominance at the major locus obtained by the equation $d = (\mu_{Aa} - \mu_{aa}) / (\mu_{AA} - \mu_{aa})$, such that $d = 0$ corresponds to a recessive gene, $d = 1$ corresponds to a dominant gene, $0 < d < 1$ corresponds to some degree of additivity and $d = 0.5$ is referred to as codominant;

H the polygenic heritability in the children (k); $H = C_k/V$

Z, the ratio of adult to childhood heritability; $Z = C_a/C_k$

and t_1, t_2 and t_3, the respective probabilities that genotypes *AA*, *Aa*, and *aa* transmit the allele *A*.

The general model contains the most parameters. This model is then compared with a Mendelian transmission model, an environmental transmission model, and a polygenic model. Under a Mendelian model, the transmission probabilities, namely, the probabilities that the *AA*, *Aa*, and *aa* genotypes will pass on an *A* allele, do not significantly differ from the Mendelian expectations of 1, 0.5, and 0, respectively, whereas in the general model these transmission probabilities can take any value. Under the environmental model, these probabilities are all equal because the phenotypic mode that a child is in is unrelated to the mode that the parent is in. Whilst the Mendelian and environmental models can contain multiple small genetic and environmental effects, a polygenic model considers only the multiple small genetic effects so it has no large deviation in the trait caused by either a major locus or the environment. Having a Mendelian model favoured in a data set, dominant and recessive Mendelian submodels can be evaluated.

There are several software packages that can perform CSA: PAP (Pedigree Analysis Package, Department of Medical Biophysics and Computing, University of Utah, Salt Lake City), SAGE (Case Western Reserve University Statistical Analysis for Genetic Epidemiology at http://darwin.cwru.edu/sage/), GAP (Genetic Analysis Package from Epicenter Software, at http://icarus2.hsc.usc.edu/epicenter/gap.html) and POINTER (ftp://cedar.genetics.soton.ac.uk/pub/PROGRAMS/pointer/). These variety of software aimed to do CSA perform a maximum likelihood analysis to find the combination of the parameter listed above values which gives the largest overall likelihood for the observed data. Within the variety of models considered, it proceeds usually by testing a general non-restricted model, which contain the maximum parameters that is fitted to the data and will give the best fit models of varying degrees of generality, both to determine whether a Mendelian locus is likely to exert a large effect on the phenotype of interest and to estimate the magnitude of genetic sources of variation in the trait (Gail Pairitz J., 1998). This model is then compared with restricted models such as the Mendelian transmission models (Mendelian dominant, Mendelian recessive and Mendelian co-dominant), the environmental transmission model, and the polygenic or 'no major gene model'. These models are built by testing the genetic hypotheses by keeping the relevant parameters from d, t, q, and H constant, whereas the remaining parameters are estimated by maximizing the likelihood of the phenotypes in the families.

2010
Vol.1
No. 3:1
doi: 10:3823/310

Mendelian models assume a major locus with two alleles that act either in a dominant, co-dominant or recessive fashion. The 'no major gene' model assumes that the baseline risk is not influenced by the risk trait (i.e. all persons have the same specific risk of disease). The environmental model assumes that an individual's phenotype depends on his or her environmental exposures and is independent of the phenotype of the parents. There are two parameters to test multifactorial heritability. The parameter *H* represents polygenic heritability in the offspring, where $H = C_k/V$, in which C_k is the multifactorial component and V is the overall variance. The second parameter is *Z* for which H_Z represents the multifactorial heritability in parents, where $Z = C_a/C_k$, the ratio of the multifactorial component in adults and children. Significant deviation of *Z* from 1 suggests a generational difference in multifactorial heritability.

In segregation analysis, it is incorrect to assume that the gene frequency is constant at all ages because any gene causing specific mortality must decrease with age. Risks *(R)* can then be determined using mortality figures that allow to calculate cumulative mortalities and risk, so that the R_j, the risk attributed to the j_{th} liability class, is

$$R_j = (I_j - M_{j-1})/(1 - M_{j-1})$$

where I_j is the cumulative incidence to the mid-point and M_{j-1} is the cumulative specific mortality to the end of the preceding class.

As we pointed above, models are compared by a likelihood ratio test. The difference between the minus twice the log likelihood plus a constant (*-2lnL + k*) calculated under a general model (with *m* parameters) and under a reduced model (with *n* parameters) is asymptotically distributed as χ^2 with *m - n* degrees of freedom. Another way to compare hypotheses is by using the Akaike information criterion (AIC) (Akaike HA, 1974). AIC is calculated as *-2lnL + k* plus twice the number of free parameters in the model. The model with the lowest AIC is taken to give the best fit to the data. Comparison by means of AIC values has the advantage that one model does not have to be a subset of the other so it can be used for examining non-nested models.

Finally, CSA not only allows to determine whether a major gene is involved in a familial trait but also to predict the pattern of inheritance of the hypothesized gene, the penetrance and the disease allele frequency. Taking the age-specific mortality into account, (Iselius et al., 1991) defined the penetrance in gene carriers (*G'*) as the approximate cumulative incidence for gene carriers in the j_{th} liability class, given by the following:

$$P_j = P(\text{aff} \mid G', j) + [1 - P(\text{aff} \mid G', j)]\, M'_{j-1}$$

where the genotype-specific mortality is,

$$M'_{j-1} = \Sigma P(G' \mid \text{aff}_{,i})(M_i - M_{i-1}) \,/\, \Sigma P(P' \mid \text{aff}_{,i})\,(I_i - I_{i-1})$$

The aim of our study was to assess the contribution of genetic factors in AD in an unselected large number of Spanish families, and to investigate a possible Mendelian inheritance as explanation for the reported familial aggregation of AD.

Patients and methods

In a prospective study, we ascertained through probands 21 multigenerational extended pedigrees (297 individuals), with 76 individuals affected with Alzheimer´s disease fulfilling CERAD criteria. These families gave a total of 44 nuclear families to be included in the model. Information was gathered on the probands themselves, as well as about the family history of two previous generations. Since probands were unable to give accurate answer to most questions, we interviewed the caregiver, usually a family member to ensure the accuracy of the information. Questions included the proband date of birth, sex, date of diagnosis, birthplace and birthplace of grandparents. The family history included any incidence of cognitive deterioration in the proband relatives, including cause and date of death for deceased relatives, cognitive status, type of cognitive status, date of diagnosis, and records of diagnosis.

CSA was carried out using the unified version of the mixed model of Morton and Mac-Lean (1974), implemented in the computer program POINTER (Morton et al., 1971). We analyzed the following models: non transmission (cohort effect), multifactorial (polygenic and environmental), Mendelian (dominant, recessive, codominant), polygenic, mixed (Mendelian plus polygenic) and a general model.

Liability classes

The POINTER program permits the construction of four male and four female liability classes, which describe age specific risks. To take into account age-specific mortality, all individuals whose age was known at the time of ascertainment were assigned to one of four liability classes according to its age at ascertainment, diagnosis, or death (Table 2). The liability indicator was calculated as previously described $R_j = (I_j - M_{j-1})/(1 - M_{j-1})$. Therefore, four classes were formed according to the age ranges given in Table 2. Cumulative incidence figures, to the mid-point of each class, were calculated given the rates per 100.000 as described in Bermejo (1987) and individuals were assigned to one of the four liability classes (<60 year-old; 60-69; 70-79; >80) (Table 2) according to their prior probability of affection based on the age

2010
Vol.1
No. 3:1
doi: 10:3823/310

specific prevalence rates for AD in Spain (Bermejo, 1987). Since the phenotypes were defined as dichotomies of affection status: normal versus affected, the liability to affection represented by *x* can be defined by a threshold on the liability scale, such that affection occurs when *x* is greater than a given threshold.

Table 1. Age and sex-specific prevalences of AD in Spain based on data by Bermejo 1987.

Liability class	Age of onset (years)	Mortality rates per 100,00 population
I	<60	0.3
II	60-69	30
III	70-79	230
IV	>80	1,300

Ascertainment probability

The ascertainment probability (π), as used in POINTER, is ~0 if the probability of ascertaining a family increases in proportion to the number of affected offspring (single selection) and close to 1 if the probability of ascertaining a family is independent of the number of affected offspring (complete selection). Since POINTER only accepts nuclear families as an input, extended pedigrees have to be analyzed by dividing them into their component nuclear families. Those nuclear families not containing affected probands though containing affected relatives of the "POINTER" (nominal probands) were codified in each sibling considering that the ascertainment probability value (π) is 1. Only nuclear families ascertained through pointers with at least one affected individual were included. This last approach was chosen because simulations and empirical results have shown similar results either including or not families with no affected members (Marazita et al., 1992). In this case, first-degree relatives of the proband were partitioned into nuclear families containing the proband as a parent (complete selection) or as a child (incomplete selection). There was only one proband in each family, and therefore an ascertainment probability (π) of 0.001 was used in the analysis, corresponding to single selection. Nevertheless, when all models are examined while varying the ascertainment probability over the range 0.001–0.2, the results found to be highly robust to changes in the specified ascertainment model (Figure 1).

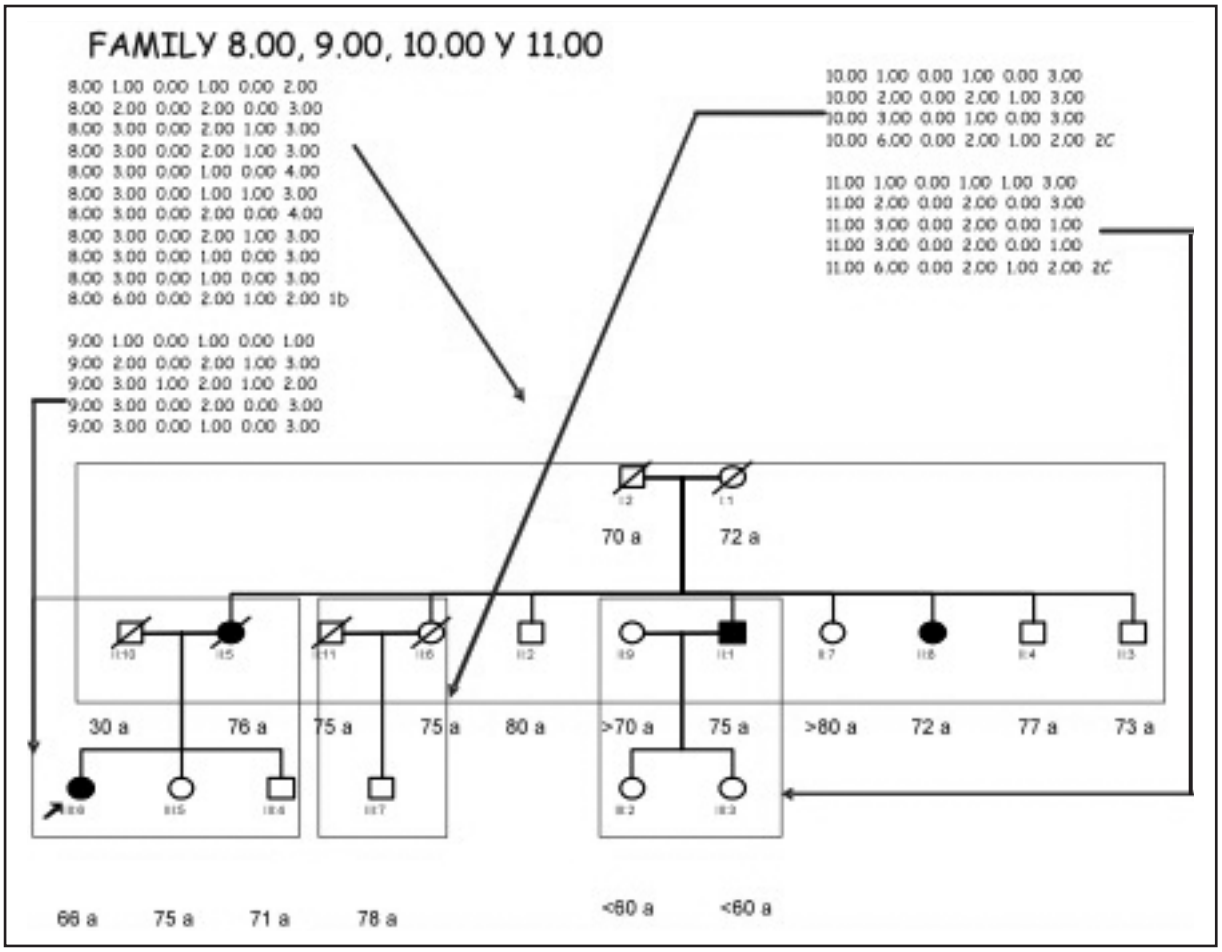


Figure 1. Example of multigenerational extended family divided into nuclear families.

2010
Vol.1
No. 3:1
doi: 10:3823/310

Test of genetic heterogeneity

The data set consisting of all nuclear families was analyzed first in order to determine whether polygenic or major locus models would explain the occurrence of AD entirely. In a second step of the study, we analyzed those individuals whose DNA was available to examine the APOE genotype. DNA was extracted from a blood sample using a phenol-chlorophorm extraction and ethanol precipitation method (Berának M., 2006). APOE genotype was determined by amplification of the exon 4 of APOE gene by polymerase chain reaction, followed by restriction fragment length polymorphism analysis using the restriction enzyme *Hha I* (Hixon JE, 1990). To determine whether the genetic background to AD was different depending on APOE genotype, overall data were subsequently analyzed in two subsets of families, those which have a proband APOE ε4 carrier versus those families whose proband was not an APOE ε4 carrier. Parameters for the polygenic, dominant, and recessive models were estimated separately in these two groups. Since the difference between the summed likelihoods in the partitioned analysis and the likelihood of the total data set is asymptotically distributed as χ^2 with *p(g - 1)* degrees of freedom, where *p* is the number of iterated parameters and *g* is the number of subgroups, heterogeneity χ^2 test (Khoury et al., 1993;Williams and Anderson, 1984) compared the sum of -2lnL of a particular model, computed on the subsets, with the -2lnL computed on the total 21 (44 nuclear families) families. This statistic was computed as follows: χ^2 = -2 [ΣlnL (best-fitted model/subgroup i) -lnL (bestfitted model/all family data)], where Σ is the sum overall i subgroups.

Results

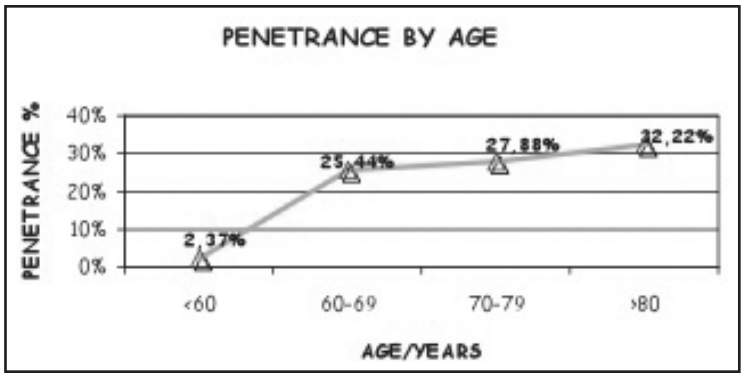
The total number of individuals included in the study was 76 (23% males and 77% females) with an average age of onset of 70 years-old. The results from the CSA for all families are given in the Table 2. The familial aggregation of AD was not due to chance, since the sporadic model was rejected (χ^2 = 143 df = 1, P<0.001). All models incorporating a major gene for genetic transmission gave a better fit to these data than the multifactorial model (χ^2 = 6 df = 1, p<0.025). The best fit among the Mendelian models was for the dominant model with a gene frequency of 0.0164 and a penetrance that increases with age (about 32.29% >80 years old, see Figure 2). When we examined the general model, we found that a gene explains the 66.7% of the heritability with t_2 = 0.18, that is below 0.5 what means that there are epistatic interactions. Penetrance of a hypothesized gene in homozygous or heterozygous carriers are 32.29% (>80 year-old); 27.88% (70-79 year-old); 25.44% (60-69 year-old); 2.37% (<60 year-old) (Figure 2).

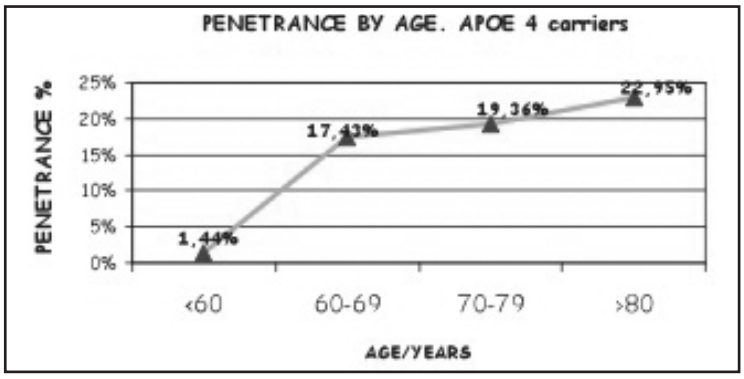
Table 2. Results of CSA of the overall data. Parameter estimates corresponding to maximum likelihood models under each set of constraints are shown for each examined model.

Models	-2ln(L)	d	t	q	H	t1	t2	t3	Z	df
(1) no transmisibility (cohort effect)	321.97	(0)	(0)	(0)	(0)	-	-	-	-	8
(2) multi-factorial	179.02	(0)	(0)	(0)	0.105	-	-	-	(1)	7
(3) dominant	163.61	(1)	3.06	0.0164	(0)	(1.0)	(0.5)	(0.0)	(1)	6
(4) codominant	163.82	(0.5)	6.14	0.016	(0)	(1.0)	(0.5)	(0.0)	(1)	6
(5) recessive	176.48	(0)	3.0	0.176	(0)	(1.0)	(0.5)	(0.0)	(1)	6
(6) t1=t2=t3 (polygenic-no transmission of major effect)	288.49	1.0	3.2	0.03	0.30	(0.97)	(0.97)	(0.97)	(1)	1
(7) mixed (d non restricted)	179.41	1.0	1.72	0.507	0.44	(1.0)	(0.5)	(0.0)	(1)	3
(8) general(non restricted)	172.95	1.0	1.64	0.354	0.343	1.2	0.18	0.01	0.348	0

In the second part of the study, we compared two subsets, those carrying an APOE ε4 allele and not carrier ones and the best-fitting model was evaluated separately in these two subgroups. The total number of families analyzed by APOE genotyping were 4 APOE3/3 and 13 APOE 3/4 or 4/4. CSA of the two separated groups (APOE ε4 carriers, APOE ε4 non carriers) concluded that the best model of inheritance was the dominant one for APOE ε4 carriers, and the codominant one for ε4 non carriers families, but we could not reject the other inheritance models. The penetrance for APOE ε4 carriers was 22.95% (>80 years old) and 20.37% (>80 years old) for APOE ε4 non carriers. A gene can explain the 48% and 34% of the heritability for APOE ε4 carriers and ε4 non carriers respectively, with epistatic interactions (t_2 = 0.35;<0.5). The difference between the 2lnL of the overall set and the sum of the 2lnL for the two groups yielded an χ^2 = 54 with df = 7; p< 0.01. This indicated evidence for etiologic heterogeneity between families ascertained.

2010
Vol.1
No. 3:1
doi: 10:3823/310





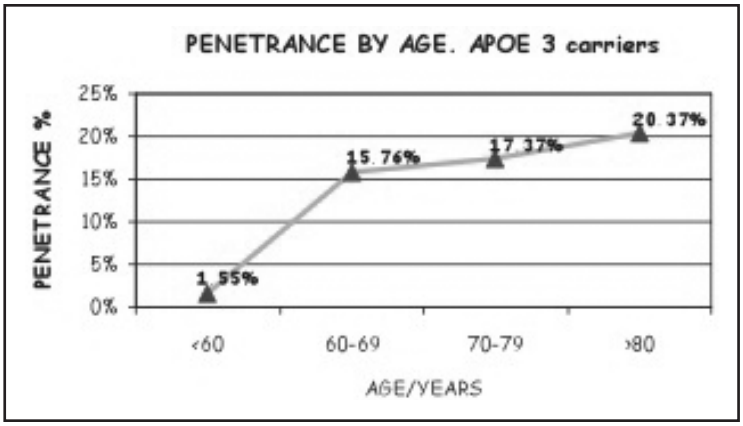


Figure 2. Penetrance of the candidate gene by age.

Table 3. Results of CSA of the families with APOE 4/4 or 3/4 carriers.

Models	-2ln(L)	d	t	q	H	t1	t2	t3	Z	df
(1) no transmisibility	181.61	(0)	(0)	(0)	(0)	-	-	-	-	8
(2) multi-factorial	88.61	(0)	(0)	(0)	0.105	-	-	-	(1)	7
(3) dominant	83.06	(1)	3.1	0.0137	(0)	(1.0)	(0.5)	(0.0)	(1)	6
(4) Codominant	83.31	(0.5)	6.3	0.013	(0)	(1.0)	(0.5)	(0.0)	(0)	6
(5) Recessive	84.77	(0)	3.45	0.149	(0)	(1.0)	(0.5)	(0.0)	(1)	6
(6) t1=t2=t3	159.68	1.0	3.2	0.03	0.30	0.97	0.97	0.97	(1)	1
(7) Mixed	88.31	1.0	1.42	0.56	0.687	(1.0)	(0.5)	(0.0)	0.49	3
(8) general(non restricted)	87.49	1.0	1.48	0.50	0.588	1.0	0.35	0.1	0.58	0

Table 4. Results of CSA of the families with APOE 3/3 carriers.

Models	-2ln(L)	d	t	q	H	t1	t2	t3	Z	df
(1) no transmisibility	57.84	(0)	(0)	(0)	(0)	-	-	-	-	8
(2) multi-factorial	28.46	(0)	(0)	(0)	0.999	-	-	-	(1)	7
(3) dominant	25.86	(1)	3.41	0.016	(0)	(1.0)	(0.5)	(0.0)	(1)	6
(4) Codominant	25.83	(0.5)	6.8	0.015	(0)	(1.0)	(0.5)	(0.0)	(0)	6
(5) Recessive	28.17	(0)	5.22	0.161	(0)	(1.0)	(0.5)	(0.0)	(1)	6
(6) t1=t2=t3	51.96	1.0	3.2	0.03	0.30	0.97	0.97	0.97	(1)	1
(7) Mixed	29.09	1.0	1.76	0.544	0.484	(1.0)	(0.5)	(0.0)	0.4	3
(8) general (non restricted)	28.66	1.0	1.27	0.461	0.667	1.0	0.35	0.1	0.409	0

2010
Vol.1
No. 3:1
doi: 10:3823/310

Discussion

Complex diseases such as AD are difficult to study from a genetic point of view. Nevertheless, a genetic approach to this disease through methodology as CSA that have demonstrated its practical usefulness in diverse genetic conditions can be very useful. In breast cancer, several different reports using CSA supported an autosomal dominant inheritance with a variable penetrance between 70 and 90 percent in gene carriers, at least in a subset of breast cancer cases (Williams and Anderson, 1984). These results provided the logical platform for additional linkage studies resulting in the discovery of the two breast cancer genes BRCA1 and BRCA2 (Hall et al., 1990;Wooster et al., 1994). These studies support that epidemiological approaches such as CSA allow to obtain preliminary data and make a good selection of the familial aggregates in order to obtain much more accuracy in further genetic studies, because if Mendelian segregation is not supported, analyses of candidate loci or random markers for linkage to the trait of interest would likely be unproductive, at least in the same data set.

We must assume that CSA has limitations in order to consider the results showed above. First of all, the major limitation of CSA is that a large amount of a very specific type of data is generally needed. Ascertainment of an appropriate sample is also necessary. Moreover, there is no reliable method to determine the sample size required for a desired level of power to detect a Mendelian locus by CSA (Gail Pairitz J., 1998).

Another practical limitation is the inability to distinguish the effect of a single locus that underlies a trait and the effects of two or more independently acting loci with similar transmission pattern (Gail Pairitz J., 1998). Since CSA cannot detect whether one phenotype is caused by different genotypes, i.e. genetic heterogeneity, a high impact of a small proportion of the families in which there was a strong genetic effect cannot be completely ruled out in our results. Third, the POINTER software assumes that any major gene inheritance occurs through a single two-allele autosomal locus, but actually, the inheritance pattern may be more complex, making the identification of a specific model more difficult.

A known limitation of CSA is that of a lack of assessment of statistical power. The effect of a rare major gene may remain masked, under the overwhelming number of 'sporadic' AD cases. Although 21 families were included in this study, lack of power may be an explanation for the findings, since none of the models examined could be rejected. The involvement of a genetic factor in AD seems obvious considering the striking reports of extensive families but the influence of this genetic factor cannot easily be unravelled by CSA. The inclusion of more individuals, especially larger-sized families (i.e. inclusion of second-degree relatives) may improve power to detect genetic mechanisms underlying transmission of AD in this cohort.

Although data on age at the onset of disease were available, the data are incomplete and possibly subject to error because of the difficulty in defining onset. Further CSA studies might usefully distinguish early and late onset as a route to discriminating between genetic and environmental etiology.

To date, apart from our study there is only one other CSA of AD (Farrer et al., 1991). The shortage of CSA studies in AD and other neurodegenerative diseases is caused by the theoretical difficulties of these studies and the troubles in collecting the data for doing it. Nevertheless, there is a need for more studies using this sort of analysis that will allow to know the real situation of the genetics of these diseases.

In conclusion, the results strongly support a Mendelian dominant or codominant susceptibility gene for AD, acting in a proportion of families. Nevertheless, Mendelian factors alone are not sufficient to fully explain the familial aggregation of this phenotype, and residual familial effects are necessary to adequately fit the data. This suggests that polygenic factors may also contribute to the etiology of AD. Parameters derived from this study may facilitate future linkage studies and have uncourageous to start searching of new genes for AD in this population.

2010
Vol.1
No. 3:1
doi: 10:3823/310

Bibliography

1. Abney M, McPeek MS, Ober C (2001) Broad and narrow heritabilities of quantitative traits in a founder population. Am J Hum Genet 68:1302-1307
2. Akaike HA (1974) New look at the statistical model identification IEEE. Transmission Automation Control 19:716-723
3. Barinaga M (1995) New Alzheimer's gene found. Science 268:1845-1846
4. Beránek M. (2006) Comparison of various methods used for extraction of cell-free genomic DNA from human plasma., 14 edn pp 21-24
5. Bertram L, Tanzi RE (2005b) The genetic epidemiology of neurodegenerative disease. J Clin Invest 115:1449-1457
6. Bertram L, Tanzi RE (2005a) The genetic epidemiology of neurodegenerative disease. J Clin Invest 115:1449-1457
7. Bonney GE (1986) Regressive logistic models for familial disease and other binary traits. Biometrics 42:611-625
8. Chartier-Harlin MC, Parfitt M, Legrain S, Perez-Tur J, Brousseau T, Evans A, Berr C, Vidal O, Roques P, Gourlet V,. (1994) Apolipoprotein E, epsilon 4 allele as a major risk factor for sporadic early and late-onset forms of Alzheimer's disease: analysis of the 19q13.2 chromosomal region. Hum Mol Genet 3:569-574
9. Farrer LA, Myers RH, Connor L, Cupples LA, Growdon JH (1991) Segregation analysis reveals evidence of a major gene for Alzheimer disease. Am J Hum Genet 48:1026-1033
10. Gail Pairitz J. (1998) Statistical genetics 98. Complex segregation analysis: Uses and limitations. Am J Hum Genet 63, 942-946. 1998.
11. Goldgaber D, Lerman MI, McBride OW, Saffiotti U, Gajdusek DC (1987) Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease. Science %20;235:877-880
12. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250:1684-1689
13. Hixon JE VD (1990) Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with Hhal. J Lipid Res 31:545-548 Iselius L, Slack J, Littler M, Morton NE (1991) Genetic epidemiology of breast cancer in Britain. Ann Hum Genet 55:151-159
14. Khoury MJ, Beaty TH, Cohen BH (1993) Scope and strategies of genetic epidemiology: analysis of articles published in Genetic Epidemiology, 1984-1991. Genet Epidemiol 10:321-329
15. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, Yu CE, Jondro PD, Schmidt SD, Wang K,. (1995) Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science 269:973-977
16. Marazita ML, Hu DN, Spence MA, Liu YE, Melnick M (1992) Cleft lip with or without cleft palate in Shanghai, China: evidence for an autosomal major locus. Am J Hum Genet 51:648-653
17. Martin JB (1999) Molecular basis of the neurodegenerative disorders. N Engl J Med 340:1970-1980
18. Morton NE, Yee S, Lew R (1971) Complex segregation analysis. Am J Hum Genet 23:602-611
19. Risch N (1990) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222-228
20. Risch N (2001) The genetic epidemiology of cancer: interpreting family and twin studies and their implications for molecular genetic approaches. Cancer Epidemiol Biomarkers Prev 10:733-741
21. Rybicki BA, Elston RC (2000) The relationship between the sibling recurrence-risk ratio and genotype relative risk. Am J Hum Genet 66:593-604
22. Schaid DJ, Sommer SS (1993) Genotype relative risks: methods for design and analysis of candidate-gene association studies. Am J Hum Genet 53:1114-1126
23. Williams WR, Anderson DE (1984) Genetic epidemiology of breast cancer: segregation analysis of 200 Danish pedigrees. Genet Epidemiol 1:7-20
24. Wittke-Thompson JK, Pluzhnikov A, Cox NJ (2005) Rational inferences about departures from Hardy-Weinberg equilibrium. Am J Hum Genet 76:967-986
25. Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D,. (1994) Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science 265:2088-2090