



HUMAN GENOME EPIDEMIOLOGY (HuGE) REVIEW

Meta-Analysis of the Association of the Cathepsin D Ala224Val Gene Polymorphism with the Risk of Alzheimer's Disease: A HuGE Gene-Disease Association Review

Christos Ntais¹, Anastasia Polycarpou¹, and John P. A. Ioannidis^{1,2,3}

¹ Clinical and Molecular Epidemiology Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece.

² Biomedical Research Institute, Foundation for Research and Technology-Hellas, Ioannina, Greece.

³ Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, MA.

Received for publication August 5, 2003; accepted for publication October 10, 2003.

A C-to-T polymorphism in exon 2 of the cathepsin D gene encoding cathepsin D (*CTSD*) has been implicated as a risk factor for Alzheimer's disease. The authors performed a meta-analysis of 14 studies (16 comparisons) with *CTSD* genotyping (3,174 Alzheimer's disease cases and 3,298 controls). Overall, the random effects odds ratio for the *T* versus the *C* allele was 1.17 (95% confidence interval (CI): 0.95, 1.44), with some between-study heterogeneity ($p < 0.01$). There was significant between-study heterogeneity but no evidence of a significant association when the first hypothesis-generating study was excluded from the calculations (odds ratio (OR) = 1.11, 95% CI: 0.91, 1.35; $p = 0.29$). The summary odds ratio for *T* carriers versus *T* noncarriers was similar in subjects carrying or not carrying an apolipoprotein E $\epsilon 4$ allele (*APOE**4). The increased susceptibility to Alzheimer's disease conferred by *APOE**4 carriage tended to be more prominent in the presence of the *T* allele (random effects OR = 6.07, 95% CI: 4.19, 8.79, and OR = 4.09, 95% CI: 3.15, 5.31, in *T* carriers and noncarriers, respectively). The meta-analysis shows that the *CTSD* polymorphism is not a major risk factor for Alzheimer's disease, although a small effect or an enhancement of the *APOE**4 effect cannot be excluded.

Alzheimer disease; cathepsin D; *CTSD*; epidemiology; genetics; meta-analysis; polymorphism (genetics)

Abbreviations: CI, confidence interval; OR, odds ratio.

Editor's note: This paper is also available on the website of the Human Genome Epidemiology Network (<http://www.cdc.gov/genomics/hugenet/default.htm>).

GENE

Cathepsin D, an intracellular acid protease which exhibits beta-secretase activity in vitro, has been implicated in the

processing of the amyloid precursor protein and the tau protein (1, 2). The cathepsin D gene (*CTSD*) is located on the short arm of chromosome 11 (11p15.5) and consists of nine exons. The synthesis of beta-amyloid peptide is a putative key event in the pathogenesis of Alzheimer's disease. Beta-amyloid derives from its precursor protein via proteolytic cleavage by secretases. Therefore, it has been postulated that variants in the genes coding for enzymes involved in the

Correspondence to Dr. John P. A. Ioannidis, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece (e-mail: jioannid@cc.uoi.gr).

TABLE 1. Characteristics of studies included in the meta-analysis

Country	Ethnic group	Selection/characteristics of cases and controls (age range (mean))		Eligible subjects*		First author, year (reference)
		Alzheimer's disease cases	Controls	Alzheimer's disease cases (no.)	Controls (no.)	
Germany	German	AD† according to NINCDS-ADRDA† criteria. Medical and family history, general medical and neurologic examination, psychiatric interview, neuropsychological testing, blood and CSF† studies, and CT† scans were performed to exclude other forms of dementia. 66.7% female (age range: 51–101 (74.4) years).	A. Randomly selected healthy subjects (age: >50 years) from general population ($n = 191$). No psychiatric disorders and dementia on psychiatric interview and neuropsychological testing. 51.8% female (age range: 50–100 (70.6) years). B. Nondemented, depressed hospitalized patients ($n = 160$). Same clinical evaluation as the AD group. 65.6% female (age range: 50–88 (68.0) years).	102	351	Papassotiropoulos, 1999 (9)
Germany	German	AD according to NINCDS-ADRDA criteria. Medical and family history, general medical and neurologic examination, psychiatric interview, neuropsychological testing, blood studies, and CT scans were performed to exclude other forms of dementia. 60.6% developed AD after the age of 65 years, and 62.2% had no family history for dementia. 61% female (mean age: 72.0 years).	Nondemented subjects. Same clinical evaluation as the AD group. 61% female (mean age: 69.0 years).	127	184	Papassotiropoulos, 2000 (7)
Germany	Caucasian‡	AD according to NINCDS-ADRDA criteria. 71.6% female (mean age: 74.2 years).	Patients or healthy volunteers with no signs of dementia and MMSE† score of ≥ 27 , interviewed by professional examiners. 59.9% female (mean age: 67.8 years).	324	302	Menzer, 2001 (8)
Italy	Italian	AD diagnosis based on DSM† criteria. A. Sporadic AD cases ($n = 131$). 63.4% female (age range: 45–88 (71.1) years). B. AD cases belonging to autopsy-proven AD families ($n = 66$): 33 early onset AD (mean age: 66.4 years) and 33 late-onset AD (mean age: 76.2 years) cases.	Subjects assessed to exclude diagnosis of any neurologic disorder. 66.6% female (age range: 26–108 (72.9) years).	197	126	Bagnoli, 2002 (16)
Italy	Italian	AD according to NINCDS-ADRDA criteria. Clinical examination, including neuropsychological testing, laboratory studies, and neurologic examination. 76.8% female (mean age: 76.3 years).	Community-dwelling elderly people with MMSE score of ≥ 28 . Clinical examination performed as in cases. 75% female (mean age: 71.7 years).	142	120	Ingegneri, 2003 (12)
Japan	Japanese	AD according to NINCDS-ADRDA criteria. 71.6% female (mean age: 74.4 years).	Community-dwelling elderly people judged cognitively normal by MMSE. 46.3% female (mean age: 74.9 years).	275	479	Matsui, 2001 (15)
	American	Autopsy-confirmed AD cases (mean age: 77.8 years).	Not clarified (mean age: 61.1 years).	69	50	

Table continues

proteolytic cleavage of amyloid precursor protein or in the degradation and clearance of beta-amyloid from the central nervous system may be potential risk factors for Alzheimer's disease.

GENE VARIANTS

The *CTSD* gene contains a polymorphic C-to-T transition site at position 224 in exon 2. This polymorphism results in

TABLE 1. Continued

Country	Ethnic group	Selection/characteristics of cases and controls (age range (mean))		Eligible subjects*		First author, year (reference)
		Alzheimer's disease cases	Controls	Alzheimer's disease cases (no.)	Controls (no.)	
Poland	Polish	Late-onset AD according to NINCDS-ADRDA criteria. CT scan was obtained for each patient. 66% female (mean age: 76.4 years).	Not clarified. 58% female (mean age: 74.2 years).	100	100	Styczynska, 2003 (19)
Spain	Spanish	AD according to NINCDS-ADRDA criteria. 67% female (age range: 50–98 (75.3) years).	Subjects randomly selected from a nursing home. Free from significant illness on complete neurologic and medical examinations, MMSE score of ≥ 28 , verified on annual follow-up assessment. 70% female (age range: 63–100 (80.4) years).	311	346	Mateo, 2002 (11)
Sweden	Swedish	AD according to NINCDS-ADRDA criteria. Clinical diagnosis based on medical history; physical, neurologic, and psychiatric examination; screening laboratory tests; ECG†; chest radiograph; EEG‡; and brain CT ($n = 111$). Neuropathologic diagnosis based on CERAD‡ criteria ($n = 93$). No family history of dementia. 61.3% female.	Healthy volunteers without history, symptoms, or signs of psychiatric or neurologic disease, malignant disease, or systemic disorders; MMSE score of ≥ 28 ($n = 76$). Autopsy group of patients who had died from cardiac or malignant disease; no history of dementia or psychiatric or neurologic diseases; negative autopsy for dementia ($n = 108$). 58.1% female.	204	186	Prince, 2001 (14)
Sweden	Scottish	Early onset AD according to DSM III-R criteria. No family history of dementia (age range: 30–65 years).	Not clarified.	121	152	Emahazion, 2001 (18)
United Kingdom	British	AD according to DSM IV and NINCDS-ADRDA criteria. Where possible, a CT scan was performed to aid diagnosis. 66% female (mean age: 77.7 years).	Healthy spouses or volunteers with unrevealing medical history and physical examination. 69% female (mean age: 77.1 years).	183	187	McIlroy, 1999 (10)
United States	American	Late-onset AD (mean age of onset: 71.8 years) according to NINCDS-ADRDA criteria. 35% autopsy-confirmed cases. 65% female (mean age: 76.1 years).	Controls recruited from dementia research center ($n = 89$) (mean age: 72.7 years) and survey study ($n = 248$) (mean age: 75.2 years). All survey subjects had an MMSE score of ≥ 28 .	531	337	Bhojak, 2000 (13)
United States	American	AD according to NINCDS-ADRDA criteria. AD clinic cases participating in a multicenter clinical drug trial and patients evaluated at university clinics. Community-based AD cases participating in dementia screening with extensive subsequent diagnostic evaluation. 58.6% female (mean age: 75.8 years) for Caucasians, 69.6% female (mean age: 73.7 years) for Hispanics.	Subjects evaluated in the community-screening program and found to be free from cognitive problems (MMSE score of > 27). 50.8% female (mean age: 75.7 years) for Caucasians, 58.9% female (mean age: 72.6 years) for Hispanics.	210	120	Crawford, 2000 (6)
	Hispanic			79	112	
United States	American	AD cases from dementia research center. 18.5% had neuropathologically confirmed AD (mean age: 70.8 years).	Cognitively normal subjects from the same dementia research center (mean age: 66.5 years).	200	182	Bertram, 2001 (17)

* All eligible subjects were genotyped with the exception of 21 controls in the study by Bhojak et al. (13), 12 controls in the study by Prince et al. (14), and one Alzheimer's disease case and three controls in the study by Emahazion et al. (18).

† AD, Alzheimer's disease; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association; CSF, cerebrospinal fluid; CT, computed tomography; MMSE, Mini-Mental State Examination; DSM, *Diagnostic and Statistical Manual of Mental Disorders*; ECG, electrocardiogram; EEG, electroencephalogram; CERAD, Consortium to Establish a Registry for Alzheimer's Disease.

‡ Germany: 173 cases, 217 controls; Switzerland: 44 cases, 55 controls; Italy: 107 cases, 30 controls.

an Ala38-to-Val substitution in the cathepsin D profragment (3). The polymorphism has been associated with increased secretion and altered intracellular maturation of the cathepsin D profragment in one study (3). Moreover, the *T* allele has been associated with a 50 percent decrease in beta-amyloid peptide 1–42 levels in the cerebrospinal fluid of patients with Alzheimer's disease (4). Finally, this polymorphism was recently reported to be significantly associated with general intelligence in healthy elderly people (5).

Molecular epidemiologic studies have presented seemingly contradictory results concerning a potential role of *CTSD* polymorphism in Alzheimer's disease (6–19). There is also controversy on whether this polymorphism may interact with the apolipoprotein E ϵ 4 allele (*APOE*4*), which is the best known genetic determinant for sporadic Alzheimer's disease (20). Single studies may have been underpowered to detect interactions or even overall effects. Given the amount of accumulated data, we deemed it important to perform a quantitative synthesis of the evidence using rigorous methods. Thus, we conducted a comprehensive meta-analysis of all available studies relating the *CTSD* polymorphism to the risk of Alzheimer's disease.

DISEASE

Alzheimer's disease is the most common cause of progressive cognitive impairment in the elderly, with an annual incidence of approximately 1 percent at 65–69 years increasing up to 40 percent in the very elderly (>85 years of age) (21, 22). Mutations in the amyloid precursor protein, presenilin 1, and presenilin 2 genes account for 5 percent of the cases and result in an autosomally dominant pattern that is expressed with complete penetrance and early manifestations (23). Alzheimer's disease is probably slightly more common in females (24). Other proven or postulated risk factors include head injury (in particular among males) (25), as well as family history, low income, low education, low occupational status, depression, exposure to aluminum in drinking water, hypertension, and Down's syndrome (26, 27). Conversely, the use of nonsteroidal antiinflammatory drugs to treat arthritis has been associated with a reduced risk of Alzheimer's disease, as has been estrogen use by postmenopausal women (27). Physical activity, diets with high levels of vitamins B₆, B₁₂, and folate, and red wine in moderate quantities may be protective (27). The prevalence of Alzheimer's disease varies considerably among different population groups (28). At least a few dozens of polymorphisms have been examined in relation to sporadic Alzheimer's disease, and published reviews are available (23, 29). Among them, there is conclusive evidence from several studies and meta-analysis thereof that *APOE*4* is a strong risk factor for developing Alzheimer's disease for both male and female subjects and for both early onset (<65 years) and late-onset disease (20), with an approximately fivefold increase in the odds of developing Alzheimer's disease. Single studies have also implicated other polymorphisms as being important, although the reported odds ratios are much smaller than those seen for *APOE*4*, and attempts at replication in subsequent research have not been conclusive. Meta-analyses for some other polymorphisms have already appeared in the

literature (30–35). They suggest no significant overall associations for several of these polymorphisms, including the myeloperoxidase gene promoter polymorphism (30), intronic or promoter region polymorphisms of presenilin 1 (for late-onset disease) (31), an insertion-deletion polymorphism or a missense mutation in the alpha-2 macroglobulin gene (32), and several polymorphisms of the protein tau gene (33). Associations of modest effect size (odds ratios (ORs) = 1.30–1.35) have been claimed in meta-analyses of the low density lipoprotein receptor-related protein gene exon 3 polymorphism (34) and of an insertion-deletion polymorphism in the angiotensin-converting enzyme I gene (35).

META-ANALYSIS METHODS

Identification and eligibility of relevant studies

We considered all studies that examined the association of the *CTSD* polymorphism with Alzheimer's disease. Sources included MEDLINE and EMBASE (from January 1994 to September 2003). The search strategy was based on combinations of "Alzheimer's disease," "*CTSD*," "cathepsin D," "polymorphism," "allele," and "genetics." References of retrieved articles were also screened.

Case-control studies were eligible if they had determined the distribution of *CTSD* genotypes in Alzheimer's disease cases (regardless of age of onset) and in a concurrent control group of dementia-free subjects using a molecular method for genotyping. Cases with Alzheimer's disease were eligible regardless of whether they had a family history of Alzheimer's disease or not. However, we excluded family-based studies of pedigrees with several affected cases per family, because their analysis is based on linkage considerations.

Data extraction

Two investigators independently extracted data and reached consensus on all items. The following information was sought from each report: authors, journal and year of publication, country of origin, selection and characteristics of Alzheimer's disease cases and controls, demographics, ethnic group of the study population, eligible and genotyped cases and controls, and number of cases and controls for each *CTSD* genotype. For studies including subjects of different ethnic groups, data were extracted separately for each ethnicity, whenever possible. Furthermore, we examined whether matching had been used, whether there was specific mention of blinding of the personnel who performed the genotyping to the clinical status of the subjects, and whether the genotyping method had been validated.

Meta-analysis

The primary analysis compared Alzheimer's disease cases with controls for the contrast of *T* versus *C* alleles. This analysis aims to detect overall differences. We also examined the contrast of extremes (homozygotes), *T/T* versus *C/C*. Finally, we examined the contrast of *T/T* versus (*C/T* + *C/C*) and the contrast of (*C/T* + *T/T*) versus *C/C*. These contrasts

TABLE 2. Distribution of CTSD alleles among Alzheimer's disease cases and controls in the included studies

First author, year (reference)	Ethnic group	T/T		C/T		C/C	
		Alzheimer's disease cases (no.)	Controls (no.)	Alzheimer's disease cases (no.)	Controls (no.)	Alzheimer's disease cases (no.)	Controls (no.)
Papassotiropoulos, 1999 (9)	German	1	0	27	47	74	304
McIlroy, 1999 (10)	British	0	1	29	16	154	170
Bhojak, 2000 (13)	American	2	0	98	56	431	260
Crawford, 2000 (6)	American	0	0	43	20	167	100
	Hispanic	0	2	13	28	66	82
Papassotiropoulos, 2000 (7)	German	0	0	30	18	97	166
Bertram, 2001 (17)	American	2	1	31	29	167	152
Matsui, 2001 (15)	Japanese	0	1	4	7	271	471
	American	1	1	8	6	60	43
Menzer, 2001 (8)	Caucasian*	3	1	43	33	278	268
Emahazion, 2001 (18)	Scottish	0	3	13	27	107	119
Prince, 2001 (14)	Swedish	0	0	27	22	177	152
Mateo, 2002 (11)	Spanish	2	8	54	54	255	284
Bagnoli, 2002 (16)	Italian	4	1	41	26	152	99
Ingegneri, 2003 (12)	Italian	4	1	29	21	109	98
Styczynska, 2003 (19)	Polish	1	0	11	9	88	91

* German, Swiss, and Italian.

correspond to the recessive and dominant effects, respectively, of the *T* allele.

The odds ratio was used as the metric of choice. For each genetic contrast, we estimated the between-study heterogeneity across all eligible comparisons using the chi-square-based *Q* statistic (36). Heterogeneity was considered significant for $p < 0.10$. Data were combined using both fixed-effects (Mantel-Haenszel) and random-effects (DerSimonian and Laird) models (37). Random effects incorporate an estimate of the between-study variance and tend to provide wider confidence intervals, when the results of the constituent studies differ among themselves. In the absence of between-study heterogeneity, the two methods provide identical results. Random effects are more appropriate when heterogeneity is present (37).

We also performed cumulative meta-analysis (38) and recursive cumulative meta-analysis (39, 40) to evaluate whether the summary odds ratio for the *T* versus *C* contrast changed over time as more data accumulated and whether the strength of the association changed when the first hypothesis-generating study was excluded from the calculations (41). Inverted funnel plots and the Begg-Mazumdar publication bias diagnostic (nonparametric τ correlation coefficient) (42) evaluated whether the magnitude of the observed association was related to the variance of each study, that is, whether large studies gave different results compared with smaller ones (43). Finally, we evaluated whether the summary results were different when the analysis was limited to studies with more intensive efforts to exclude Alzheimer's disease from controls (those that clearly performed neuropsychological testing for all controls).

Previous investigations have alluded to the possibility that the *T* allele may interact with the *APOE*4* allele in conferring susceptibility to Alzheimer's disease (7, 9). Thus, we also evaluated the effect of *T* allele carriage on the risk of Alzheimer's disease separately for *APOE*4*-positive and *APOE*4*-negative subjects. Moreover, we evaluated the genetic effect conferred by the presence of *APOE*4* separately in subjects carrying the *T* allele and those not carrying the *T* allele. Odds ratios were combined with fixed and random effects models, as described above. When these data were not reported, we communicated with the primary investigators to obtain this information, whenever possible.

Analyses were performed with SPSS 11.0 (SPSS, Inc., Chicago, Illinois) and Meta-Analyst (Joseph Lau, Boston, Massachusetts) software. Whenever there were 0 values in a 2×2 table, we added 0.5 to all four cells, so that an odds ratio could be calculated. All *p* values are two tailed.

META-ANALYSIS RESULTS

Eligible studies

Fourteen studies probing the relation between the *CTSD* polymorphism and Alzheimer's disease susceptibility were identified (6–19) and are profiled in table 1. Two of the eligible studies (6, 15) contained subjects of two different ethnic groups, so a total of 16 separate comparisons were considered. There was considerable diversity of ethnic groups. Eleven studies (6–15, 19) selected Alzheimer's disease cases according to criteria from the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA), two studies (16, 18) selected Alzhei-

TABLE 3. Summary odds ratios for various contrasts

Contrast	No. of study comparisons	Random effects		Fixed effects	
		OR*	95% CI*	OR	95% CI
<i>T</i> vs. <i>C</i>	16 (12,944)†	1.17	0.95, 1.44‡	1.16	1.01, 1.32
<i>T/T</i> vs. <i>C/C</i>	16 (5,552)	1.07	0.55, 2.10	0.98	0.52, 1.85
<i>T/T</i> vs. (<i>C/T</i> + <i>C/C</i>)	16 (6,472)	1.05	0.54, 2.06	0.97	0.51, 1.84
(<i>C/T</i> + <i>T/T</i>) vs. <i>C/C</i>	16 (6,472)	1.19	0.97, 1.47§	1.18	1.03, 1.36

* OR, odds ratio; CI, confidence interval.

† Numbers in parentheses, number of cases and controls combined.

‡ Significant between-study heterogeneity ($p < 0.01$).

§ Significant between-study heterogeneity ($p < 0.05$).

mer's disease cases according to *Diagnostic and Statistical Manual of Mental Disorders: DSM-III-R* and *DSM-IV* criteria, and one study (17) did not clarify the exact criteria used for the diagnosis of Alzheimer's disease. Five studies (13–17) also included autopsy-confirmed Alzheimer's disease cases. Two studies (7, 16) mentioned that they included cases with a family history of Alzheimer's disease, eight studies (9–14, 18, 19) specifically excluded such patients, and the remaining did not clarify the background of family history. One study (6) mentioned that 61 percent of Alzheimer's disease cases had late-onset disease (age of onset: >65 years), five studies (8, 12–14, 19) specifically included only late-onset Alzheimer's disease cases, one study (18) specifically excluded such patients, and the remaining did not clarify the age at onset. Controls did not have a diagnosis of Alzheimer's disease, but the amount of additional screening (general physical and neurologic examination, psychiatric interview, neuropsychological testing, blood and cerebrospinal fluid studies, computed tomography scan, and Mini-Mental State Examination score) to exclude Alzheimer's disease differed substantially across studies.

Specific matching for age was described in five studies (10, 12–15). One study also matched for sex (10). Only one study (7) specifically mentioned blinding of the personnel who performed the genotyping. Appropriate molecular methods for genotyping were used. All studies used polymerase chain reaction, and two studies (14, 18) also used dynamic allele-specific hybridization.

Meta-analysis database

The eligible studies summarized in table 2 included a total of 3,175 cases with Alzheimer's disease and 3,334 controls, of whom 3,174 and 3,298, respectively, had genotype data. The *T* allele was more highly represented among controls of American descent (overall prevalence of 8.6 percent, 95 percent confidence interval (CI): 7.1, 10.1) than in controls of European (7.8 percent, 95 percent CI: 7.0, 8.6) or Asian (0.9 percent, 95 percent CI: 0.3, 1.5) descent. There was significant heterogeneity in the prevalence rates of the *T* allele even across the control subjects of European descent, with a rate of 14.3 percent among Hispanic Americans, 10.4 percent in Italy, 10.1 percent in Spain, and lower rates in northern European countries (7.6 percent in the United Kingdom, 6.3 percent in Sweden, 6.1 percent in Germany,

and the lowest prevalence rate of 4.5 percent in a Polish population). Overall, the prevalence of *T/T* homozygosity was 0.3 percent, 0.8 percent, and 0.2 percent in control subjects of American, European, and Asian descent, respectively. The respective prevalence rates of *C/T* heterozygosity were 16.6 percent, 14.0 percent, and 1.5 percent, and the respective rates for *C/C* homozygosity were 83.1 percent, 85.2 percent, and 98.3 percent. The distribution of genotypes in control groups was consistent with Hardy-Weinberg equilibrium in all studies.

Overall effects

There was a trend suggesting that the *T* allele may confer increased susceptibility to Alzheimer's disease (figure 1). As shown in table 3, the summary odds ratio was 1.17 by random effects ($p = 0.14$), and there was significant heterogeneity among the 16 study comparisons ($p < 0.01$ for heterogeneity). We found no evidence of an association of the *T/T* genotype with the risk of Alzheimer's disease relative to the *C/C* genotype. There was no significant between-study heterogeneity. No evidence of an association with Alzheimer's disease was discerned also when the recessive model was examined for the effect of *T*, while a trend for an association was seen in the dominant model (by random effects, OR = 1.19, 95 percent CI: 0.97, 1.47; $p = 0.10$). There was no between-study heterogeneity in the recessive model contrast, while significant heterogeneity ($p < 0.05$) was still seen for the dominant model contrast. Subgroup analysis of studies with cases and controls of European descent yielded similar results (15 comparisons (11,436 alleles): OR = 1.18, 95 percent CI: 0.96, 1.46; $p = 0.12$) ($p < 0.01$ for heterogeneity).

Bias diagnostics

In cumulative meta-analysis and recursive cumulative meta-analysis, the magnitude of the summary odds ratio had not been stable over time, and it had changed considerably per year with an apparent dissipation of the postulated effect (by random effects, summary OR for *T* vs. *C*: 2.05 at the end of 1999, 1.41 at the end of 2000, 1.16 at the end of 2001, 1.14 at the end of 2002, and 1.17 in 2003). Excluding the first hypothesis-generating study (9), we found that the summary odds ratio became 1.11 (95 percent CI: 0.91, 1.35; $p = 0.29$)

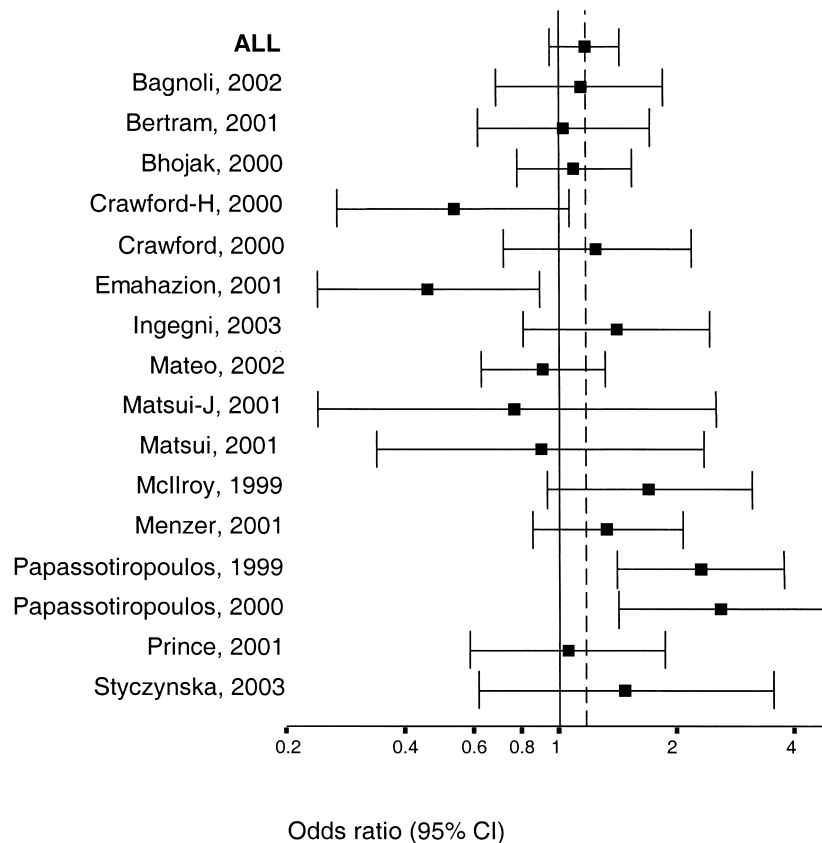


FIGURE 1. Meta-analysis for the effect of the *T* allele versus the *C* allele on the risk of Alzheimer's disease. Each comparison is presented by the name of the first author and the year of publication. "H" signifies Hispanic subjects, and "J" signifies Japanese subjects. For each comparison, the point estimate of the odds ratio and the accompanying 95% confidence interval (CI) are shown. "ALL" represents the summary random-effects estimate for the comparison along with the respective 95% confidence interval. Values above 1 denote an increased risk for Alzheimer's disease with the *T* allele.

with significant between-study heterogeneity ($p < 0.05$), and there was no evidence of any association even in the comparison of the homozygotes (*T/T* vs. *C/C*: OR = 0.96, 95 percent CI: 0.48, 1.90; $p = 0.91$), with no between-study heterogeneity. There was no relation between the effect size and the variance of each study, suggesting that larger studies agreed with the results of smaller studies. Analyses limited to studies with more intensive efforts to exclude Alzheimer's disease from controls yielded similar results (11 comparisons (10,382 alleles): OR = 1.22, 95 percent CI: 0.96, 1.54; $p = 0.10$) ($p = 0.02$ for heterogeneity).

Interaction with the *APOE* genotype

Nine studies (7–9, 11–14, 16, 18) obtained data on both *CTSD* and *APOE* genotypes. With one study (8) separating male and female subjects, 10 comparisons became available.

These nine comparative studies tended to give a slightly inflated effect for the *T* allele, as compared with the full meta-analysis of 14 comparisons (summary OR by random effects = 1.23 vs. 1.17 for the complete meta-analysis); thus, inferences should be cautious. Among carriers of the high-risk *APOE*4* allele, *T* allele carriers had a random-effects

odds ratio of 1.38 (95 percent CI: 0.89, 2.15) for Alzheimer's disease compared with subjects not carrying the *T* allele. Among subjects not carrying the *APOE*4* allele, the respective odds ratio was 1.13 (95 percent CI: 0.90, 1.42). There was significant between-study heterogeneity in the *APOE*4*-positive group ($p = 0.07$). The two effect sizes overlapped widely.

Among carriers of the *T* allele, the presence of *APOE*4* increased the risk of Alzheimer's disease 6.07-fold (95 percent CI: 4.19, 8.79), with no between-study heterogeneity. Among subjects without the *T* allele, the presence of *APOE*4* increased the risk of Alzheimer's disease 4.09-fold (95 percent CI: 3.15, 5.31), with significant between-study heterogeneity ($p < 0.01$). The two estimates overlapped widely in individual studies and overall, but typically the odds ratios were larger in the group of *T* allele carriers (figure 2).

DISCUSSION

This meta-analysis includes data from 14 case-control studies with over 6,000 genotyped Alzheimer's disease cases and controls, and it proves that the *CTSD* polymorphism is

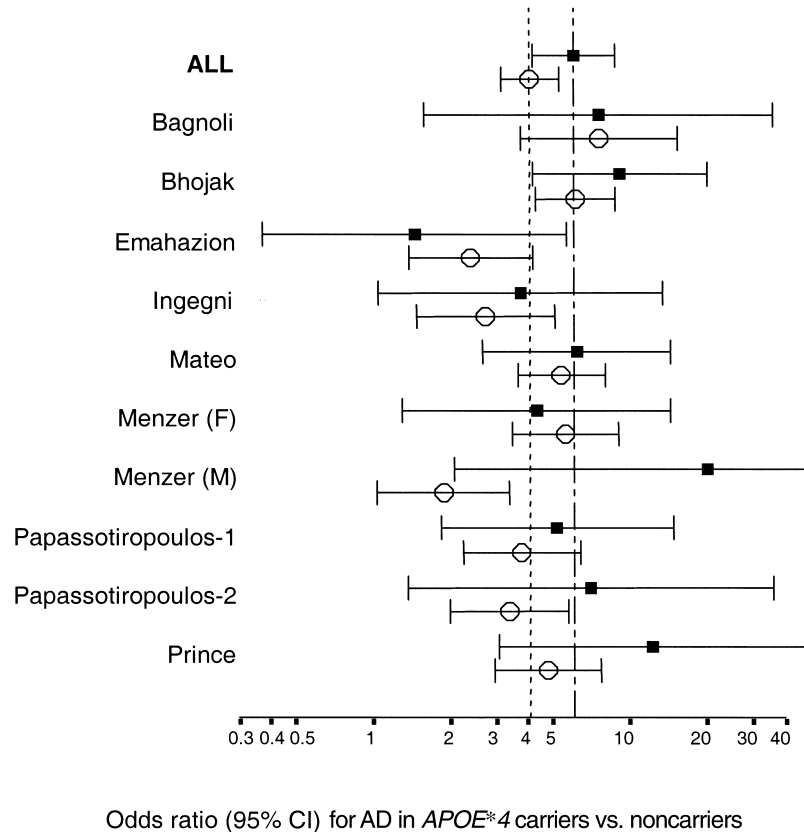


FIGURE 2. Meta-analysis for the effect of the *APOE**4 allele in *T* carriers and in *T* noncarriers on the risk of Alzheimer's disease (AD). Each comparison is presented by the name of the first author. "M" signifies male subjects, and "F" signifies female subjects. For each comparison, the point estimate of the odds ratio and the accompanying 95% confidence interval (CI) are shown. Filled squares represent *T* carriers, while open circles represent *T* noncarriers. "ALL" represents the summary random-effects estimate for the comparison along with the respective 95% confidence interval. Papassotiropoulos-2 and -1 pertain to the following respective references: (7, 9).

not a strong risk factor for Alzheimer's disease. The current evidence cannot exclude that the *T* allele of the *CTSD* polymorphism may increase modestly the risk of Alzheimer's disease, but there was significant heterogeneity in the results of various studies. However, we found that bias might exist, with a decreasing effect size, as more data accumulated over time. Moreover, the meta-analysis also cannot exclude the possibility that the presence of the *CTSD T* allele may enhance the increased susceptibility toward Alzheimer's disease conferred by *APOE**4. However, even if this is true, it would pertain to the relatively small group of subjects who carry both the *APOE**4 and *CTSD T* alleles. In all, the impact of the *T* allele on a population level would be small, if present at all.

The meta-analysis findings may be interpreted against the postulated biologic context of the *CTSD* polymorphism. Cathepsin D is an intracellular acid protease with beta-secretase activity in vitro (1, 2) that can cleave amyloid precursor protein and the tau protein to generate fragments with intact microtubule-binding domains (44), which might play a role in the pathogenesis of paired helical filaments. One study has shown that the *CTSD* polymorphism is associated with

increased secretion and altered intracellular maturation of procathepsin D (3). It should be noted that it is not clear whether the *CTSD* polymorphism has functional consequences for the mature form of the enzyme. Although we cannot exclude a biologic effect for the *CTSD* polymorphism, our findings are in accordance with the results of a recent full genome scan showing no significant linkage of Alzheimer's disease to the short arm of chromosome 11, the region where the *CTSD* gene is located (45-47).

Attention should be given to the design of individual studies. The results of meta-analyses may be affected by methodological problems and potential biases in the designs of the constituent studies. Nondifferential misclassification errors may dilute the strength of an observed association. Alzheimer's disease cases were generally selected according to appropriate criteria. However, some young control subjects may have developed Alzheimer's disease at older ages. The choice of an appropriate age window for assessing a postulated genetic risk factor for Alzheimer's disease is difficult. Studies of younger subjects may be more suitable for identifying risk factors that result in early onset Alzheimer's disease. Conversely, selection of younger subjects

may be less appropriate, if the influence of the postulated genetic risk factor is more important in late-onset Alzheimer's disease.

Subgroup effects and effect modification (e.g., differential effects of a genetic polymorphism on early vs. late-onset Alzheimer's disease cases or familial vs. sporadic disease) or complex interactions with other genes may also need to be considered (48). Our analyses addressed interactions with *APOE*4*, the major known genetic determinant of Alzheimer's disease. Interactions with other genetic or environmental factors have not been studied. The trend for a stronger effect of *APOE*4* in the presence of *T* allele carriage is interesting in the light of data suggesting that *T* carriage may affect the general intelligence (5). However, subgroup and interaction analyses should be interpreted cautiously, since differences between subgroups may occur by chance (49) and their validation would require studies with even larger sample sizes than the several thousand included in this meta-analysis. Finally, population stratification may theoretically have affected the results of the constituent studies in the meta-analysis (50), since we documented that the frequency of the *T* allele varied considerably among the different ethnic groups or even among the different ethnic groups of European descent. However, most studies were strictly ethnically defined, so the population stratification effect is unlikely to have been of any considerable magnitude.

Because of the increasing prevalence of Alzheimer's disease worldwide, it is crucial to identify genetic risk factors for this neurodegenerative disease. Thus, the list of identified polymorphisms that may influence the risk of Alzheimer's disease is continuously expanding (23, 29), but most of the reported associations of candidate genes to date remain nonreplicated or at least controversial after subsequent investigation. Early and small genetic association studies may come up with spurious findings (41, 51, 52). Even when genetic associations are replicated, usually they do not have a major public health impact that would lead to screening recommendations (53). Nevertheless, such knowledge could improve our understanding about the pathogenesis of complex diseases such as Alzheimer's disease.

LABORATORY TESTS

The methods used for *CTSD* genotyping in the analyzed studies are straightforward and include polymerase chain reaction (9) and dynamic allele-specific hybridization (14). The error rate due to misclassification is likely to be small. Future studies should nevertheless ensure and clearly report that assessment of genotyping has been performed while blinded to the clinical status of the patient.

POPULATION TESTING

To date there has been no population testing of the *CTSD* polymorphism. Based on the results of the meta-analysis, such testing would not be indicated given the currently available data.

ACKNOWLEDGMENTS

The authors thank Jonathan A. Prince, Patrizia Mecocci, and Anthony J. Brookes for providing additional data from their studies.

REFERENCES

1. McDermott JR, Gibson AM. Degradation of Alzheimer's beta-amyloid protein by human cathepsin D. *Neuroreport* 1996;7: 2163-6.
2. Chevalier N, Vizzavona J, Marambaud P, et al. Cathepsin D displays in vitro beta-secretase-like specificity. *Brain Res* 1997; 750:11-19.
3. Touitou I, Capony F, Brouillet JP, et al. Missense polymorphism (C/T224) in the human cathepsin D pro-fragment determined by polymerase chain reaction-single strand conformational polymorphism analysis and possible consequences in cancer cells. *Eur J Cancer* 1994;30A:390-4.
4. Papassotiropoulos A, Lewis HD, Bagli M, et al. Cerebrospinal fluid levels of β -amyloid(42) in patients with Alzheimer's disease are related to the exon 2 polymorphism of the cathepsin D gene. *Neuroreport* 2002;13:1291-4.
5. Payton A, Holland F, Diggle P, et al. Cathepsin D exon 2 polymorphism associated with general intelligence in a healthy older population. *Mol Psychiatry* 2003;8:14-18.
6. Crawford FC, Freeman MJ, Schinka J, et al. The genetic association between cathepsin D and Alzheimer's disease. *Neurosci Lett* 2000;289:61-5.
7. Papassotiropoulos A, Bagli M, Kurz A, et al. A genetic variation of cathepsin D is a major risk factor for Alzheimer's disease. *Ann Neurol* 2000;47:399-403.
8. Menzer G, Muller-Thomsen T, Meins W, et al. Non-replication of association between cathepsin D genotype and late onset Alzheimer disease. *Am J Med Genet* 2001;105:179-82.
9. Papassotiropoulos A, Bagli M, Feder O, et al. Genetic polymorphism of cathepsin D is strongly associated with the risk for developing sporadic Alzheimer's disease. *Neurosci Lett* 1999; 262:171-4.
10. McIlroy SP, Dynan KB, McGleenon BM, et al. Cathepsin D gene exon 2 polymorphism and sporadic Alzheimer's disease. *Neurosci Lett* 1999;273:140-1.
11. Mateo I, Sanchez-Guerra M, Combarros O, et al. Lack of association between cathepsin D genetic polymorphism and Alzheimer disease in a Spanish sample. *Am J Med Genet* 2002;114: 31-3.
12. Ingegini T, Nocentini G, Mariani E, et al. Cathepsin D polymorphism in Italian elderly subjects with sporadic late-onset Alzheimer's disease. *Dement Geriatr Cogn Disord* 2003;16: 151-5.
13. Bhojak TJ, DeKosky ST, Ganguli M, et al. Genetic polymorphisms in the cathepsin D and interleukin-6 genes and the risk of Alzheimer's disease. *Neurosci Lett* 2000;288:21-4.
14. Prince JA, Feuk L, Sawyer SL, et al. Lack of replication of association findings in complex disease: an analysis of 15 polymorphisms in prior candidate genes for sporadic Alzheimer's disease. *Eur J Hum Genet* 2001;9:437-44.
15. Matsui T, Morikawa Y, Tojo M, et al. Cathepsin D polymorphism not associated with Alzheimer's disease in Japanese. *Ann Neurol* 2001;49:544-5.
16. Bagnoli S, Nacmias B, Tedde A, et al. Cathepsin D polymorphism in Italian sporadic and familial Alzheimer's disease. *Neurosci Lett* 2002;328:273-6.
17. Bertram L, Guenette S, Jones J, et al. No evidence for genetic

- association or linkage of the cathepsin D (*CTSD*) exon 2 polymorphism and Alzheimer disease. *Ann Neurol* 2001;49:114–16.
18. Emahazion T, Feuk L, Jobs M, et al. SNP association studies in Alzheimer's disease highlight problems for complex disease analysis. *Trends Genet* 2001;17:407–13.
 19. Styczynska M, Religa D, Pfeffer A, et al. Simultaneous analysis of five genetic risk factors in Polish patients with Alzheimer's disease. *Neurosci Lett* 2003;344:99–102.
 20. Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* 1997;278:1349–56.
 21. Hebert LE, Scherr PA, Beckett LA, et al. Age-specific incidence of Alzheimer's disease in a community population. *JAMA* 1995;273:1354–9.
 22. Cummings JL, Cole G. Alzheimer disease. *JAMA* 2002;287:2335–8.
 23. St George-Hyslop PH. Molecular genetics of Alzheimer's disease. *Biol Psychiatry* 2000;47:183–99.
 24. Gao S, Hendrie HC, Hall KS, et al. The relationships between age, sex, and the incidence of dementia and Alzheimer disease: a meta-analysis. *Arch Gen Psychiatry* 1998;55:809–15.
 25. Fleming S, Oliver DL, Lovestone S, et al. Head injury as a risk factor for Alzheimer's disease: the evidence 10 years on; a partial replication. *J Neurol Neurosurg Psychiatry* 2003;74:857–62.
 26. Launer LJ, Andersen K, Dewey ME, et al. Rates and risk factors for dementia and Alzheimer's disease. *Neurology* 1999;52:78–84.
 27. McDowell I. Alzheimer's disease: insights from epidemiology. *Aging (Milano)* 2001;13:143–62.
 28. Hendrie HC, Ogunniyi A, Hall KS, et al. Incidence of dementia and Alzheimer's disease in 2 communities. *JAMA* 2001;285:739–47.
 29. Combarros O, Alvarez-Arcaya A, Sanchez-Guerra M, et al. Candidate gene association studies in sporadic Alzheimer's disease. *Dement Geriatr Cogn Disord* 2002;14:41–54.
 30. Combarros O, Infante J, Llorca J, et al. The myeloperoxidase gene in Alzheimer's disease: a case-control study and meta-analysis. *Neurosci Lett* 2002;326:33–6.
 31. Dermaut B, Roks G, Theuns J, et al. Variable expression of presenilin 1 is not a major determinant of risk for late-onset Alzheimer's disease. *J Neurol* 2001;248:935–9.
 32. Koster MN, Dermaut B, Cruts M, et al. The α 2-macroglobulin gene in AD: a population-based study and meta-analysis. *Neurology* 2000;55:678–84.
 33. Russ C, Powell JF, Zhao J, et al. The microtubule associated protein tau gene and Alzheimer's disease—an association study and meta-analysis. *Neurosci Lett* 2001;314:92–6.
 34. Sanchez-Guerra M, Combarros O, Infante J, et al. Case-control study and meta-analysis of low density lipoprotein receptor-related protein gene exon 3 polymorphism in Alzheimer's disease. *Neurosci Lett* 2001;316:17–20.
 35. Kehoe PG, Katzov H, Feuk L, et al. Haplotypes extending across ACE are associated with Alzheimer's disease. *Hum Mol Genet* 2003;12:859–67.
 36. Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med* 1997;127:820–6.
 37. Petitti DB. Meta-analysis, decision analysis and cost-effectiveness analysis. New York, NY: Oxford University Press, 1994.
 38. Lau J, Antman EM, Jimenez-Silva J, et al. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992;327:248–54.
 39. Ioannidis JP, Contopoulos-Ioannidis DG, Lau J. Recursive cumulative meta-analysis: a diagnostic for the evolution of total randomized evidence from group and individual patient data. *J Clin Epidemiol* 1999;52:281–91.
 40. Ioannidis J, Lau J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. *Proc Natl Acad Sci U S A* 2001;98:831–6.
 41. Ioannidis JP, Ntzani EE, Trikalinos TA, et al. Replication validity of genetic association studies. *Nat Genet* 2001;29:306–9.
 42. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–101.
 43. Ioannidis JP, Trikalinos TA, Ntzani EE, et al. Genetic associations in large versus small studies: an empirical assessment. *Lancet* 2003;361:567–71.
 44. Kenessey A, Nacharaju P, Ko LW, et al. Degradation of tau by lysosomal enzyme cathepsin D: implication for Alzheimer neurofibrillary degeneration. *J Neurochem* 1997;69:2026–38.
 45. Kehoe P, Wavrant-De Vrieze F, Crook R, et al. A full genome scan for late onset Alzheimer's disease. *Hum Mol Genet* 1999;8:237–45.
 46. Collins JS, Perry RT, Watson B Jr, et al. Association of a haplotype for tumor necrosis factor in siblings with late-onset Alzheimer's disease: the NIMH Alzheimer Disease Genetics Initiative. *Am J Med Genet* 2000;96:823–30.
 47. Farrer LA, Bowirrat A, Friedland RP, et al. Identification of multiple loci for Alzheimer disease in a consanguineous Israeli-Arab community. *Hum Mol Genet* 2003;12:415–22.
 48. Lehmann DJ, Williams J, McBroom J, et al. Using meta-analysis to explain the diversity of results in genetic studies of late-onset Alzheimer's disease and to identify high-risk subgroups. *Neuroscience* 2001;108:541–54.
 49. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992;116:78–84.
 50. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet* 2003;361:598–604.
 51. Lohmueller KE, Pearce CL, Pike M, et al. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;33:177–82.
 52. Ioannidis JPA. Genetic associations: false or true? *Trends Mol Med* 2003;9:135–8.
 53. Little J, Khoury MJ, Bradley L, et al. The Human Genome Project is complete. How do we develop a handle for the pump? *Am J Epidemiol* 2003;157:667–73.