Complex phenotypic analysis to identify genes which contribute to or modify the development of Alzheimer's disease

Paul Hollingworth

Thesis submitted for the degree of Doctor of Philosophy.
Cardiff University, December 2005.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Contents

List of Figures

List of Tables

Acknowledgements

I would first like to thank my supervisors, Professor Julie Williams and Dr Marian Hamshere. Professor Williams provided me with the opportunity to work within the Department of Psychological Medicine, among an interdisciplinary team investigating the genetic aetiology of Alzheimer's disease (AD). She inspired me to diversify my research from my background in psychology into the exciting, but often complex, field of genetics and was a constant source of ideas and guidance. Dr Marian Hamshere is a statistical geneticist who works to meticulously high standards. Throughout my time in Cardiff she has always found time, no matter how busy, to discuss my work and offer advice.

Data from this thesis arise from collections in the United Kingdom, National Institute of Mental Health and National Institute of Aging. I was directly involved in the data collection funded by the Medical Research Council and the Alzheimer's Research Trust. I offer my thanks to the principal investigators and research teams at the Institute of Psychiatry, London; Cambridge University and Trinity College, Dublin. I reserve particular mention for my colleagues working on the 'field team' in Cardiff, particularly Kimberley Dowzell who worked tirelessly, without complaint, to minimise the stress I endured whilst combining the demands of writing up and coordinating data collection for the AD research initiative. I would also like to offer thanks to the families and AD sufferers who kindly offered their time to help with this research.

I should not forget my parents, Carol and Steve, whose guidance and patience has kept me focused and provided me with the confidence and necessary foundation to achieve academic success. I would like to dedicate this thesis to my partner and best friend, Melanie, with whom I have lived since the very beginning of my undergraduate studies. Throughout this time she has made me smile every single day and I sincerely believe I wouldn't have completed my undergraduate studies, or progressed further, without her unconditional support.

I would also like to thank Professor Peter Holmans, Dr Gary Hill, Dr Valentina Moskvina, Eliot Barnett, Luke Jehu, Dr Richard Abraham, Dr Angharad Morgan and all those in the department for helpful discussions, performing laboratory work, helping with programming and offering advice.

General abstract

Late-onset Alzheimer's disease (LOAD) is a heritable disorder. It is invariably characterised by a decline in cognitive abilities, however, marked variation in behavioural symptoms and age at onset are observed between sufferers. This clinical heterogeneity may be genetically modified, hence, may provide a productive avenue of exploration for those seeking to unravel the genetic aetiology of LOAD. This thesis employed a sequential three stage approach to search for loci implicated in the development of genetically influenced features of the disease.

Behavioural symptoms in 1,120 unrelated individuals with LOAD were assessed using the Neuropsychiatric Inventory. The 12 symptom domain scores were subjected to principal components analysis. Three interpretable components were identified, comprising: "frontal lobe dysfunction", "psychosis" and "mood". These components remained stable when taking account of disease severity.

The familiality of clinical variation was assessed. Affected siblings from 388 families were characterised in terms of aggression, psychosis and mood disturbances. Age at onset data were available for affected siblings from 458 families. Familial clustering was found for age at onset, psychosis, aggression and mild depression, with the strongest evidence noted for age at onset and psychosis. Major depression and a combined phenotype of depression with anxiety showed limited evidence of familial aggregation.

Covariate linkage analysis was employed to search for loci which may influence clinical variation in LOAD. This included a sample of 513 affected relative pairs. Increases in LOD were observed with age at onset (chromosome 1, 2, 12, 19 and 21), aggression (chromosome 9), psychosis (chromosome 7 and 15) and minor depression (chromosome 21).

Understanding factors associated with behavioural symptoms and age at disease onset may lead to the achievable goal of disease modification. These findings support the hypothesis that clinical variation in AD is genetically modified, setting the stage for future linkage and association studies.

Chapter 1
General Introduction

Alzheimer's disease (AD) is a devastating neurodegenerative disorder and is the primary cause of dementia in elderly populations. It is estimated that AD accounts for about two thirds of subjects with dementia (Nussbaum and Ellis 2003) and that it afflicts approximately 15 million individuals worldwide (Fratiglioni et al. 1999). In the United States alone it is estimated that as many as 4.5 million individuals are currently suffering with AD (Hebert et al. 2003). In England and Wales there are thought to be approximately 180,000 new cases of dementia each year (Matthews and Brayne 2005).

Aside from the obvious detrimental effects to sufferers, AD also causes severe distress for family members and caregivers, along with placing a huge burden on the economy (Lowin et al. 2001). In the United Kingdom, the direct costs of AD are between 7 and 14 billion pounds per year (Lowin et al. 2001), whilst in the United states the total annual cost has been estimated at around $76 billion (Rice et al. 1993). Current projections suggest that the number of elderly people will double in the next generation, resulting in more than a billion people over the age of 60 by 2025 in Europe alone (Taket 1992). Given that AD is age dependent, the escalating growth of the elderly population, particularly in the oldest age group, means that the economic and societal costs of the disease will increase over the coming years (Villareal and Morris 1999). Herbert and colleagues (2003) estimated that, baring a cure, the number of individuals with AD in the United States will increase almost threefold by 2050, to over 13 million people. The predicted increase in the prevalence will undoubtedly have an enormous impact on society (Souetre et al. 1999).

There is currently no cure for AD. Cholinesterase inhibitors and *N*-methyl d-aspartate receptor-targeted therapies are, at present, the only treatments available in the UK. These forms of medication provide modest benefits to cognition, activities of daily living and behaviour, and can provide temporary stabilisation of the rate of decline (Desai and Grossberg 2005). However, they do not benefit all

AD sufferers and their positive affects are usually temporary (Cummings and Cole 2002). As such, there is an urgent need for more effective therapeutic interventions. Even treatments which reduce the incidence of AD by only 1% would offset the projected increase in costs due to an aging population (Brookmeyer et al. 1998). It is therefore essential that we gain a comprehensive understanding of its aetiology. Such an understanding could lead to the achievable goal of disease-modification. In recent years major advances in understanding the causes and pathogenesis of AD have arisen from molecular genetic research and this powerful tool will undoubtedly continue to provide important conceptual and practical advances. Therefore, gaining a more complete understanding of the genetic contribution to AD must be a priority, as it will provide a strong platform for the development of future preventative and therapeutic strategies.

1.1 Clinical definition of Alzheimer's disease

The diagnosis of AD is difficult. To date there are no specific neuroimaging or biological markers for the disease. Histopathological examination of brain tissue at post mortem is the only way to diagnose AD definitively. However, even this is problematic, as only 50% to 60% of individuals meeting neuropathological criteria for AD will have experienced cognitive decline during life (Knopman et al. 2003).

Diagnosis of AD is generally based on physical examination, patient history and detailed cognitive assessment. These approaches serve to compare each individual against a series of inclusion criteria and also allow competing causes of dementia to be systematically excluded (e.g. vascular dementia, dementia with lewy bodies etc.). These approaches can be complimented by the use of neuroimaging which generally serves to identify vascular contributions to dementia and rule out potentially treatable causes (Kantarci and Jack 2003). Typically, AD is characterised by an insidious onset and gradual decline in cognition and functional abilities (Desai and Grossberg 2005). An initial phase of forgetfulness is usually accompanied by difficulty learning, recalling new information and progressive language disorder, from anomia in the early stages to complete aphasia as the illness progresses. Visuospatial difficulties can often become apparent,

manifesting as geographical disorientation or difficulty with copying figures in cognitive testing (Cummings and Cole 2002). Furthermore, deficits in executive function usually occur over the course of the illness (Baudic et al. 2005; Rainville et al. 2002). Motor disturbances are common in the later stages of disease development, including gait changes, rigidity and seizures (McKhann et al. 1984). The advanced stages of the disease are characterised by total dependence on others for assistance in activities of daily living. At this point patients often lose all semblance of speech and are frequently bed bound (Villareal and Morris 1999). Without treatment those with AD usually survive for between 7 and 10 years after onset of symptoms (Bracco et al. 1994; Larson et al. 2004).

DSM-IV (American Psychiatric Association, 1994) and the National Institute of Neurological and Communication Disorders and Stroke and the Alzheimer's disease and Related Disorders Associations (NINCDS-ADRDA) (McKhann et al. 1984) criteria are generally used in both research and clinical settings to make a diagnosis of Alzheimer's disease. These criteria were mainly developed to attain uniformity of classification for research and treatment purposes (Clarfield and Foley 1993). They both require deficits in memory and one other area of cognition, including aphasia, apraxia, agnosia and executive functioning, further stipulating that these difficulties should cause significant decline in functional abilities and activities of daily living. DSM-IV and NINCDS-ADRDA criteria also stipulate that the illness should be characterised by an insidious onset and gradual decline in cognitive and functional abilities, and that competing causes of dementia are ruled out. NINCDS–ADRDA guidelines specify separate criteria for possible and definite AD. A diagnosis of *possible AD* is reserved for those with an atypical course of illness or individuals who might have some other co-morbid illness that can lead to dementia but that is not considered to be the primary cause of the disease. *Definite AD* is reserved for cases where pathological evidence is available, either by autopsy or brain biopsy, which shows an excess abundance of neurofibrillary tangles and senile plaques compared to what would be expected among healthy age matched individuals (McKhann et al. 1984).

The reliability and validity of DSM and NINCDS-ADRDA criteria for AD have been found to be good. To assess the inter-rater reliability of diagnostic criteria for AD

O'Conner and colleagues (1996) compared diagnoses of 100 elderly people, with a variety of diagnoses, within and between five research centres based in Australia, Germany, the Netherlands, United Kingdom and United States. The within centre inter-rater reliability, using DSM criteria for AD, was high, whilst between centre reliability was reported to be moderate to good. To assess the validity of these diagnostic criteria a number of studies have attempted to confirm, via autopsy, diagnoses made using NINCDS-ADRDA and DSM criteria during life. In general, the accuracy of clinical diagnosis relative to neuropathology has been reported to be between 86% and 93% (Becker et al. 1994; Gearing et al. 1995; Holmes et al. 1999). However, despite the validity of the diagnostic criteria it is still notable that 'pure AD' only accounts for between 50% and 60% of all dementia cases, with a further 20% to 30% showing AD pathology in conjunction with other pathological lesions (Desai and Grossberg 2005; Gearing et al. 1995; Holmes et al. 1999).

Both DSM-IV and NINCDS-ADRDA criteria support a cut-off of 65 to differentiate between early- and late- onset AD. Both early- and late- onset forms of AD are often categorised as sporadic or familial (Ashford and Mortimer 2002). The term familial AD is usually reserved for cases in which a clear pattern of autosomal dominant inheritance is observed, or for cases carrying genetic mutations known to cause early-onset forms of the disease. Familial cases usually present with the disease before 65 years of age. Sporadic, or non-familial AD, which constitutes around 95% of cases of the disease, includes those in which no clear mode of inheritance is observed (Ashford and Mortimer 2002).

The neuropathological hallmarks of the disease include extracellular deposits of β-amyloid in senile plaques, neurofibrillary tangles of phosphorylated tau protein, neuron degeneration and synaptic loss (Cummings and Cole 2002; Kamboh 2004). The current criteria for pathological diagnosis of definite AD require the presence of both senile plaques and neuropathological tangles (Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease 1997). The main constituent of the extracellular senile plaques is the 42 amino-acid amyloid β peptide (Aβ) that is derived from the amyloid precursor protein (APP) (Kamboh 2004). The APP protein is present in

almost all tissues, and undergoes three alternative steps of cleavage by α-, β- and γ-secretase enzymes. When cut by α-secretase and then γ-secretase, APP generates a harmless peptide. However, when cut by β-secretase and then γ – secretase, APP generates peptides of 39 to 43 amino acids, of which Aβ-42 accounts for about 10%. Aβ-42 is neurotoxic and involved in the formation of senile plaques in AD brains (Selkoe 2001). Tangles are the second major histopathological feature of AD. They contain paired helical filaments of abnormally phosphorylated tau protein which occupy the cell body and extend into the dendrites. Senile plaques and neurofibrillary tangles can occur independently of one another (Selkoe 2001). AD is also characterised by reduction in synaptic density and loss of neurons. Neuronal loss or atrophy in the nucleus basalis, locus ceruleus and raphe nuclei of the brainstem leads to deficits in cholinergic, noradrenergic and serotonergic transmitters, respectively (Cummings and Cole 2002). Synaptic loss is the best current pathologic correlate of cognitive decline, and synaptic dysfunction is evident long before synapses and neurons are lost (Coleman et al. 2004).

1.2 Epidemiology of late-onset Alzheimer's disease

AD accounts for between 45% and 76% of cases of dementia in those over 65 years of age (Bachman et al. 1992; Gautrin et al. 1990; Kokmen et al. 1989; Ostbye and Crosse 1994; von Strauss et al. 1999). It is estimated that between 2.3 and 4.5 million individuals are currently suffering with AD in the United states (Brookmeyer et al. 1998; Hebert et al. 2003). In England and Wales it is estimated that there are approximately 180,000 new cases of dementia each year (Matthews and Brayne 2005). Perhaps the most notable risk factors for late-onset Alzheimer's disease (LOAD) are a family history of dementia and the presence of the Apolipoprotein (APOE) ε4 allele. Numerous population based twin studies have been conducted and generally support a heritability estimate of around, or in excess of, 60% (Bergem et al. 1997; Gatz et al. 1997; Raiha et al. 1996), whilst family based studies consistently find the presence of one or more affected family members to be a strong risk factor for the disease (Fratiglioni et al. 1993; Jarvik et al. 1996; Martinez et al. 1998; Sleegers et al. 2004; van Duijn et al. 1991). Twin and family studies of late-onset AD are discussed in more detail in sections 3.1.2

and 3.1.3. Despite displaying substantial heritability, APOE is currently the only gene to show consistent association with LOAD. The association between APOE and AD will be discussed in section 1.3. In addition to genetics, numerous other risk factors have been reported, of which increasing age, female gender and low levels of education are the most consistent and will be discussed briefly in this chapter. Other, less well supported, risk factors have been reported, including smoking (Bowirrat et al. 2001; Lee 1994; Letenneur et al. 1994b), head injury (Launer et al. 1999; Nicoll et al. 1995; Canadian Study of Health and Aging 1994), depression (Devanand et al. 1996; Speck et al. 1995), cardiovascular risk factors (Luchsinger et al. 2004) and oestrogen replacement therapy (Henderson et al. 1994; Schmidt et al. 1996). However, a full review of these is beyond the scope of this introduction.

1.2.1 Aging and late-onset Alzheimer's disease

Increasing age is the most notable non-genetic risk factor for AD. Numerous studies have aimed to estimate the risk attributable to aging. Comparisons between these studies are hampered by the use of different methodology and study populations. Some have sought to determine the prevalence of AD (e.g. the proportion of affected individuals in the population at a specific point in time), according to age and other risk factors. Alternatively, others have focused on disease incidence over time (e.g. the number of newly acquired cases among previously healthy individuals over a given period of time). In theory, cumulative incidence at specific ages should equal the age specific prevalence.

Prevalence estimates have tended to vary across studies. The Rotterdam study, employed a cross sectional, population based, design incorporating 7,528 participants from a suburb of Rotterdam (Ott et al. 1995). The prevalence of AD was estimated to be 0.9%, 7.4% and 26.8% among those in the 65 to 74, 75 to 84 and ≥ 85 year age groups respectively. These estimates are somewhat lower than those reported elsewhere. For example, in a population based study of all individuals aged 65 years or older in three urban Chicago neighbourhoods, Hebert and colleagues (2003) reported a steady rise in AD prevalence from around 5%

among those between 65 and 74 years of age, to nearly 50% in those aged 85 years or older.

Some have suggested that the increasing risk of AD with age may be exaggerated by studies looking at the age specific prevalence rates, as prevalence is determined by both incidence and duration of the disease. Hence, differential survival within specific age groups could affect the age distribution of dementia prevalence (Drachman 1994). As such longitudinal based studies which focus on the incidence of AD may be better placed to delineate the relationship between aging and dementia.

A number of larger, multi centre and meta-analyses have been performed, which are likely to provide the most reliable and informative sources of data. Recently, Matthews and Brayne (2005) reported data from the MRC-CFAS study, in which individuals from England and Wales were assessed longitudinally for 2 years. They found that the incidence of AD rose with age, from 7.4 per 1,000 person years at 65 to 69 years of age to 84.9 per 1,000 person years in those aged over 85. These findings are comparable with other large population based analyses of AD incidence. For example, Jorm and Jolley (1998) performed a meta-analysis using data from 23 published studies, concluding that incidence rates for both dementia and AD rose exponentially up to the age of 90 years. Incidence rates differed according to ethnicity and with diagnostic criteria for dementia, with lower incidence in Eastern Asians, but were largely comparable to those reported by Matthews and Brayne (2005). Similarly, consistent results were reported by the authors of the European studies of dementia network (EURODEM) (Launer 1992). EURODEM was formed in 1988 to harmonise protocols used in a number of population based follow up studies on incident dementing illness. In 1999, analyses of 528 incident dementia patients were presented, incorporating over 28,000 years of person follow up (Launer et al. 1999). Of these, 352 patients were diagnosed with AD. They reported that disease incidence increased steeply with age, from 2.5 per 1000 person years at 65 years of age to 85.6 in those aged 90+ years.

Further to these larger studies, findings from numerous population based samples at single sites have been used to determine age specific incidence rates for AD in a variety of populations, including American rural (Ganguli et al. 2000) and urban communities (Bachman et al. 1993; Fillenbaum et al. 1998; Havlik et al. 2000; Hebert et al. 1995; Kawas et al. 2000; Newman et al. 2005; Tang et al. 2001), Canada (Canadian Study of Health and Aging 1994), Italy (Di Carlo et al. 2002), Nigeria (Hendrie et al. 2001), Sweden (Fratiglioni et al. 1997; Guo et al. 1999; von Strauss, 1999), Rotterdam, Holland (Ott et al.), Cambridge, UK (Paykel et al. 1998), Australia (Waite et al. 2001) and Japan (Yoshitake et al. 1995).

On the whole most studies report incidence rates of between 1 and 7.4 per 1000 person years in those aged around 65, raising to between 20 and 85 per 1000 person years among those aged over 85. There is some variation in incidence rates reported in these studies. For example, compared to findings elsewhere, lower levels of AD were reported in the Framingham Study (Bachman et al. 1993) and in a study of community dwelling Italians (Di Carlo et al. 2002). Variable incidence rates could be attributable to differences in study design, population sampling methods, diagnostic criteria or real geographical incidence variations. For example, the study of an Italian population, reported by Di Carlo and colleagues, employed a two stage assessment approach, in which all individuals were screened using the mini mental state examination (MMSE). Those who scored below a certain cut-off were referred for further diagnostic evaluation. However, such an approach assumes perfect screening test sensitivity (Rogan and Gladen 1978). This is a particular problem as the MMSE is known to have low sensitivity for cognitive decline, especially in the early stages of disease development (Cummings and Cole 2002). Furthermore, they did not correct for educational level. Those with a higher education generally perform better on cognitive tests (Qiu et al. 2001; Stern et al. 1994b), increasing the possibility that those with a high education could have been falsely recorded as not having dementia. It is advantageous to perform detailed evaluations with a random sample of those who pass and fail initial screening, thus, the proportion of 'false negatives' can be estimated and adjusted for (Matthews and Brayne 2005). In addition, a number of studies have used different diagnostic criteria and reported contrasting response rates, which further hinders comparisons between findings.

In the Framingham study only those with moderate to severe dementia were considered, which is likely to explain the relatively low incidence of AD reported. It should also be noted that non participation in population based samples is likely to bias incidence estimates. Participation is likely to be biased towards those not suffering with dementia. As such, one would expect lower participation rates to be associated with an underestimation of dementia prevalence. However, moderate to high participation rates have generally been reported, usually between 70% and 80%.

Despite these methodological concerns, there is a general consensus that the risk of AD increases exponentially with age, with incidence rates approximately doubling every five years up to 90+ years of age. Kukull and colleagues (2002) plotted the incident rates per 1000 person years against age for 7 independent population studies. Although, incidence rates differed between studies, the *slope* of the lines plotting incidence against age were remarkably similar, confirming that age is a strong risk factor for AD. Increasing risk with age could be taken to imply that the disease is inevitable in those who live long enough. However, some have hypothesised that there is an 'extreme survivor' effect, where the risk of developing AD begins to reduce after a certain age (Ritchie and Kildea 1995). Analyses of this hypothesis are problematic as most studies do not include sufficient numbers of very old people to draw reliable conclusions. Studies which have included adequate numbers of those aged over 80 and 90 years have reported contradictory findings. Ritchie and colleagues (1995) carried out a meta-analysis of nine epidemiological studies of senile dementia, including samples of elderly individuals over 80 years of age. They reported that the rate of increase in dementia prevalence reduced after the age of 80 years, levelling of around 40% at approximately 95 years of age. However, these estimates were based on varying forms of dementia. Studies which have differentiated between AD and vascular dementia, seem to show that the age related increase in risk is stronger in AD (Ruitenberg et al. 2001; von Strauss et al. 1999). Studies which have specifically sought to investigate age specific incidence of AD among the very elderly have generally not reported a reduction in risk among the those aged over 90 (Gussekloo et al. 1995; von Strauss et al. 1999). For example, Von Strauss and colleagues (1999) assessed the prevalence of AD among the very elderly in the

Kungsholmen study. They examined 1,424 individuals aged over 77 years and identified 358 cases of dementia, of which 274 were considered to have AD. They found that the risk of AD developing rose exponentially even after the age of 85, with those aged 95 or over being approximately 9 times more likely to develop AD compared to those aged less than 84 years. However, they did provide some evidence that the risk of AD began to plateau after the age of 90 among males. This observation has been noted elsewhere. Miech and colleagues (2002) reported that among the Cache County population the incidence of AD increased exponentially until ages 85 to 90 years, but appeared to decline after 93 years in men and 97 years in women. However, the majority of incidence studies have shown an increase in risk of dementia after the age of 90 (Jorm and Jolley 1998; Matthews and Brayne 2005; Paykel et al. 1998), although a number of these have been limited to small numbers of nonagenarians.

1.2.2 Gender and late-onset Alzheimer's disease

Female gender has been reported as a risk factor for AD. However the findings are somewhat contradictory. In general AD is more prevalent in females; however these estimates are largely biased by the differential life expectancy among males and females. Epidemiological population based studies have reported differing results regarding the effect of gender on AD risk. A number of studies have reported different incidence rates among males and females. Brayne and colleagues (1995) performed a 2.4 year follow up of a cohort aged 75 years and over. They reported that the incidence of AD was 1.5 and 3.3 per 1000 person years among males and females, respectively. Likewise, Fratiglioni and colleagues (1997) reported a positive association between gender and AD risk in the Kungsholmen study, finding increased incidence rates for AD at all ages in women compared to men. This effect became more notable with increasing age. In total, women were found to be over 3 times more likely to develop AD than males. Hagnell and colleagues (1992) also reported a significant association between gender and lifetime risk of developing AD (25.5% among males, compared to 31.9% among females)

Several studies have reported no difference in the prevalence of dementia according to gender. For example, no notable sex differences were found in the Framingham study (Bachman et al. 1993), the Rochester population based study (Rocca et al. 1998), in a population based study of a French community (Letenneur et al. 1994a), or in the studies reported by Ganguli and colleagues (2000) and Paykel and colleagues (1994). These studies have generally included no more than 2,000 participants, followed up for only 2 to 4 years, hence prevalence estimates are likely to be imprecise (Ruitenberg et al. 2001). This is particularly problematic in the later age groups, with the majority of these studies incorporating only a small number of participants in the later stages of life. Also follow up was often hampered by incomplete response rates due to refusal and/or death of participants.

Larger studies, have offered useful insights into the relationship between gender and AD. For example, Reitenburg and colleagues (2001) assessed the relationship between gender and AD risk in the population based Rotterdam study. They included a cohort of nearly 8,000 individuals over 55 years of age at baseline, and completed one follow up after 3 to 4 years with almost 80% of participants, and a further follow up 4 to 6 years later with over 60% of participants. They reported that the incidence of AD for males and females was generally similar. However, in those aged 90 to 94 years the incidence rate of AD among males and females was 11% and 53%, respectively. In those aged ≥95 years the incidence rate was 86% among females, whereas no males were reported to develop AD in this age group. As such, females were found to have over a five-fold increase in risk of developing AD after the age of 90, compared to males. These differences could not be explained by differential response rates. Similarly, in the EURODEM study, 528 incident cases of dementia were identified (Launer et al. 1999). Significant gender differences in the incidence of AD were reported in those aged above 85 years. At 90 years of age, the incidence rate among women was 81.7%, compared to just 24.0% in men. These findings suggest that the differential risk to males and females may be restricted to very old age.

1.2.3 Education and late-onset Alzheimer's disease

Low levels of education have been reported to increase the risk of developing AD.

Katzman (1993) proposed that education might enhance the brains reserves by increasing synaptic density in the neocortical association cortex. Others have extended the cognitive reserve hypothesis to incorporate the possible beneficial effects of mental activity throughout the lifespan, taking into account other lifestyle and occupational factors associated with mental activity (Stern et al. 1994b).
The effect of education on AD risk has been examined in a number of population based longitudinal studies. For example, in the Rotterdam study, Ott and colleagues (1995) found that the relative risk of dementia decreased in a dose dependent manner, with increasing educational status. The effect could not be explained by the confounding effect of cardiovascular disease. Likewise in the Kungsholmen study, Qui and colleagues (2001) followed a dementia free cohort of 1,296 individuals over 75 years of age, identifying 109 cases of AD. Low levels of education were associated with over a two-fold increase in risk for disease development. Findings from the EURODEM study also suggest that less time in education increases risk of AD, but only among females (Launer et al. 1999). This effect remained unchanged after accounting for the confounding effects of cardiovascular disease or stroke. The gender specific association could have reflected a lack of power in this study as only 96 males were available for hypothesis testing. Further epidemiological longitudinal studies (Canadian Study of Health and Aging 1994; Di Carlo et al. 2002; Mortimer et al. 2003; Schmand et al. 1998; Zhang et al. 1990) and cross sectional studies (Hill et al. 1993; Katzman 1993; Mortel et al. 1995; Risch 2000; Stern et al. 1994b) have reported an association between AD and education.

These findings suggest that education related factors, operating in childhood, could be associated with cognitive function in later life and the subsequent development of AD. Indeed, Plassman and colleagues (1995) have provided strong evidence that intelligence and education in early adulthood correlate with cognitive function in later life. They combined data regarding education and intelligence, assessed on enrolment into the US armed forces in the 1940's, with cognitive function in a group of elderly male twins. Cognition in later life was correlated with both intelligence and education assessed in early adulthood.

Despite the preponderance of evidence suggesting that education is associated with AD, a number of large studies have failed to find a relationship (Cobb et al. 1995; Fratiglioni et al. 1991; Paykel et al. 1994). A firm association between AD and education is difficult to establish for a number of reasons. There are numerous lifestyle factors which differ according to education (Winkleby et al. 1992) which are likely to bias findings if not successfully accounted for. Perhaps the most obvious confounding variable would be socioeconomic status (SES). Most studies show a moderate to high correlation between education and SES (Evans et al. 1997b; Winkleby et al. 1992). However, studies which have simultaneously assessed the effect of education and SES on AD have generally concluded that education remains a strong predictor of AD development after controlling for SES. For example, Evans and colleagues (1997b) investigated the effect of education, occupational prestige and income on AD risk, concluding that education was the strongest predictor of disease development. Low income and lower occupational status were both associated with AD risk in individual analysis, however were not additionally predictive after controlling for education. Likewise, De Ronchi and colleagues (1998) tested the association between education and AD, after controlling for occupational level, in a sample of 495 elderly subjects with middle to high SES, among which a large proportion had received no formal education. Those with no education were reported to have a four-fold increase in risk of developing AD, compared to those who had been educated. This effect was particularly strong in those between 61 and 69 years of age. More recently, Karp and colleagues (2004) have extended their earlier work (Qiu et al. 2001), by determining whether their previously reported association between educational level and AD could be explained by occupation based socioeconomic status. In univariate analysis education was a strong predictor of AD development, associated with over a three-fold increase in risk. This effect remained significant after controlling for socioeconomic status. Low levels of education were associated with an increased AD risk in both those with a low and high SES. However, one report has suggested that occupational level is a stronger predictor of disease development than education (Bonaiuto et al. 1995). However, the sample used in this analysis was relatively small, comprising only 48 cases and 96 matched controls; hence interpretation of these findings should be made with caution.

The relationship between education and AD is complicated by issues surrounding the diagnosis of the disease. Those with a higher education perform better on neuropsychological tests, thus are less likely to be diagnosed with dementia (Qiu et al. 2001; Stern et al. 1994b). A number of positive findings to date have originated from studies which have aimed to determine the incidence of AD within given populations (Di Carlo et al. 2002; Launer et al. 1999; Mortimer et al. 2003; Ott et al. 1995; Qiu et al. 2001; Schmand et al. 1998; Zhang et al. 1990). These studies largely rely on relatively basic means of cognitive screening. Those who perform poorly on screening are generally referred for further diagnostic evaluation. As such, those with a low education are more likely to be evaluated extensively, if indeed they do perform worse on screening tests. However, given these limitations, the majority of studies appear to show an association between education and AD. This would suggest that increasing education either directly, or indirectly, reduces the risk of developing AD, maybe by increasing cognitive reserves.

Research into risk factors for AD has been plagued with inconsistent results, owing largely to methodological variations between studies (Hendrie 1998). A large number of studies have relied on cross sectional, case-control, designs, which are hindered by case ascertainment biases and differential survival. As discussed in sections 1.2.1 to 1.2.3 numerous population based studies have been conducted, which may provide more useful insights. However, population based methods still suffer from problems related to sample stratification, and the extent to which findings from specific populations can be generalised to the wider population is often not clear. There are numerous difficulties inherent in the search for risk factors for AD. For example, information about risk factors may be systematically biased between cases and controls. On the whole data regarding AD sufferers generally comes from an informant, or proxy. It is possible that the proxy of an AD case may recall previous medical history differently to the proxy of a control, or the control themselves (Launer et al. 1999). Also, studies are often biased as it is difficult to ascertain whether a particular risk factor is associated with the disease per se or whether it is associated with increased survival after the onset of dementia. In summary, it would appear that increasing age is the only non-genetic factor, which is widely acknowledged to increase risk of developing

AD, whilst female gender and low levels of education are perhaps the next most consistent contributory factors to the disease.

1.3 A brief overview of the genetics of Alzheimer's disease

The most significant genetic advances have come from the rare autosomal dominant forms of AD, characterised primarily by lower ages at onset compared to LOAD. To date, mutations in three genes including the gene encoding APP on chromosome 21 (Goate et al. 1991), presenilin 1 (PSEN1) on chromosome 14 (Sherrington et al. 1995) and presenilin 2 (PSEN2) on chromosome 1 (Levy-Lahad et al. 1995b; Rogaeva et al. 1995) have been found to cause AD in families with early-onset autosomal dominant forms of the disease. Together, mutations in these genes account for approximately 50% of early onset AD cases, with the main contribution from PSEN1 (Tandon et al. 2000). The identification of these genes has provided useful insights in understanding the biological mechanisms in AD as a whole. For example, most of the pathogenic mutations in the APP and presenilin genes are associated with abnormal processing of APP, which leads to the overproduction of toxic Aβ42 found in senile plaques (Kamboh 2004). However, these genes are not believed to be implicated in the more common form of late-onset AD.

As mentioned in section 1.2 and discussed further in sections 3.1.2 and 3.1.3, LOAD is thought to have a substantial genetic component, with twin studies reporting heritability estimates of around 60% (Bergem et al. 1997; Gatz et al. 1997; Raiha et al. 1996). To date, the Apolipoprotein gene (APOE), located on chromosome 19, is the only widely acknowledged gene associated with LOAD. APOE has three major isoforms (apoE2, apoE3 and apoE4), which differ in amino acid sequence at two sites, codon 112 and 158.These isoforms are coded for by alleles, APOE ε2, ε3 and ε4. In normal populations the ε3 allele is the most frequent, whilst ε4 occurs slightly more often than the ε2 allele (Hendrie 1998). The ε4 allele increases the risk of LOAD in a dose dependent manner, whilst ε2 is thought to be protective (Farrer et al. 1997). The association between APOE and AD was first reported in a series of publications in 1993 (Corder et al. 1993;

Saunders et al. 1993; Schmechel et al. 1993). Saunders and colleagues (1993) reported an association between the APOE ε4 allele and AD using a small prospective series of sporadic AD cases and spouse controls, which replicated in a sample of autopsy confirmed cases. Authors from the same group reported that the APOE ε4 allele was associated with both late-onset familial and sporadic forms of AD, increasing risk for the disease from 20% to 90% and reducing the age at onset (AAO) from 94 to 68 years with increasing ε4 alleles (Corder et al. 1993). They concluded that homozygosity for the APOE ε4 allele was almost sufficient to cause AD by 80 years of age. Around the same time, the same group also reported that the ε4 allele was associated with increased Aβ deposition in senile plaques, which are a major neuropathological feature of AD (Schmechel et al. 1993).

Taken together these findings provided strong evidence that the APOE gene was implicated in the development of LOAD. Since then hundreds of studies, using divergent populations, have demonstrated an association between AD and the pathogenic ε4 allele. The findings in relation to APOE have been remarkably consistent with only a few failing to find an association, largely in selective populations. An excellent review of these findings can be found at: http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=107741, alternatively, Raber and colleagues (2004) and Ashford and colleagues (2002) have also provided reviews.

The strength of the relationship varies among epidemiological studies, however the APOE ε4 allele is generally not found to be necessary nor sufficient to cause AD. For example, in the Framingham study cohort, comprising 1,030 elderly individuals, Myers and colleagues (1996) reported that 45% of ε4 homozygotes had not developed dementia by the age of 80. In addition, they reported that about 50% of AD in their cohort was not attributable to APOE genotype. Likewise, in a further population based study Evans and colleagues (1997) found that APOE only accounted for a small proportion of the incidence of AD. Indeed, they reported that if the allele did not exist or had no effect on disease risk, the incidence would be reduced by only 13.7%. These findings refute claims that the ε4 allele is either necessary or sufficient to cause AD. It should be noted that epidemiological

samples often require study populations in which enough individuals will develop AD for the study to be economically and practically viable, as such populations are often restricted to those over 65 years of age. However, many individuals with one or more APOE ε4 alleles are likely to develop AD at a younger age than the study criterion. Also, many such studies are based on samples that are dementia free at onset. Those with AD risk alleles are therefore more likely to be excluded from the study, which could lead to an underestimation of the APOE effect. In general, estimates of population risk attributable to APOE are between 20% to 57% (Nalbantoglu et al. 1994; Seshadri et al. 1995; Slooter et al. 1998).

The APOE ε2 allele is reportedly protective against AD (Bickeboller et al. 1997; Corder et al. 1994; Talbot et al. 1994). Despite being nearly as common as the ε4 allele in the general population, there are relatively few AD patients studied with the ε2 allele (Raber et al. 2004). Talbot and colleagues (1994) presented evidence that the ε2 allele may confer protection to AD, and that its effect is not simply due to the absence of the ε4 allele. These findings have been replicated by others, for example Corder and colleagues (1994) found that the risk of AD was lowest in subjects with the ε2/ε3 genotype, reporting that although a substantial proportion of AD is due to the presence of the ε4 allele, up to 23% of AD was attributable to the absence of APOE ε2 in their sample.

The neuropathological pathway by which APOE increases the risk of developing AD is not well understood. A number of studies have reported that the APOE ε4 allele is associated with increased senile plaque and neurofibrillary tangle formation in brains of AD sufferers studied at autopsy (Ghebremedhin et al. 2001). Furthermore, Bennett and colleagues (2003) found that after controlling for the effect of AD pathology, including senile plaques and neurofibrillary tangles, the association between APOE and AD no longer remained, suggesting that the ε4 allele is related to AD through an association with the pathological hallmarks of the disease, rather than via some other mechanism. In addition to these findings the APOE ε4 allele has been associated, in a dose dependent manner, with elevated rate of hippocampal atrophy in longitudinally assessed patients (Mori et al. 2002).

The effect of the APOE ε4 allele is thought to reduce with increasing age, with some authors suggesting that is has little effect on risk of developing AD after 90 years of age (Farrer et al. 1997; Meyer et al. 1998). However, in a sample of 109 cases and 303 controls aged over 85 Skoog and colleagues (1998) reported the APOE ε4 allele remained predictive of AD. Numerous studies have reported a lower AAO among those with increasing numbers of APOE ε4 alleles (Corder et al. 1993; Lucotte et al. 1994). Indeed, Meyer and colleagues suggested that APOE genotype does not appear to influence whether subjects will develop AD, but rather *when* susceptible individuals will develop the disease. However, others have reported that APOE explains <10% of the variance in AAO (Slooter et al. 1998). The ε4 allele is also thought to reduce AAO among some, but not all, familial forms of early onset AD (Levy-Lahad et al. 1995a). In addition to the ε2 and ε4 alleles, polymorphisms in the APOE promoter region have been implicated with the disease, however the results of these associations are often contradictory (Bullido et al. 1998; Lambert et al. 2002; Lambert et al. 1998; Song et al. 1998; Wang et al. 2000).

Despite the robust association between APOE and AD, the ε4 allele is neither necessary nor sufficient to cause the disease. Several studies have indicated that a number of other genes are implicated in the development of LOAD (Jarvik et al. 1996; Martinez et al. 1998; Steffens et al. 2000). Tremendous effort has been put into identifying these genes. Linkage analysis offers a means of identifying regions which are likely to contain disease loci. To date, findings from linkage analysis studies of LOAD have generally been inconclusive, with perhaps the most convincing region of linkage located on chromosome 10 (Bertram et al. 2000; Ertekin-Taner et al. 2000; Farrer et al. 2003; Li et al. 2002; Myers et al. 2000). A further explanation of linkage analysis is presented in section 4.1.2 of this thesis and linkage findings in relation to late-onset AD are discussed in sections 4.1.4 and 4.1.5. Numerous functional and positional candidate genes have been identified. Functional candidate genes are those which have a known biological function which could be implicated in the development of AD (e.g. those that are involved in the production, degradation and clearance of Aβ within the brain), whereas positional candidate genes are those located in genetic regions identified through linkage analyses. Most studies have been restricted to genes which are

both functional and positional candidates for AD (e.g. functional candidate genes within regions of linkage). Like linkage studies, results from association analyses have generally been negative, inconclusive or contradictory. Since the year 2000, it is estimated that over 200 publications, reporting positive associations in over 50 genes with AD have been published (Becker et al. 2004). However, very few of these associations have positively replicated when analysed in independent samples, increasing the possibility that they are 'false positives' (Brookes and Prince 2005). In general, it would appear that LOAD is not likely to be due to a mutation in a single gene. However, it is more probable that a number of genes, interacting with environmental risk factors, cause AD (Brookes and Prince 2005). A full review of association studies with AD as a whole is beyond the scope of this thesis, although reviews are provided by Brookes and Prince (2005), Bertram and Tanzi (2004) and Kamboh (2004).

Association and linkage studies of LOAD are plagued by issues of locus heterogeneity and phenocopies. Heterogeneity is the term used when identical phenotypes arise from different mutations at the same or different loci, whereas the term phenocopies refers to subjects with clinically indistinguishable non-genetic forms of the disease. Rare genetic variation may be associated with a small proportion of cases, and therefore undetectable using current experimental designs (Pritchard 2001). It is possible that those linked to specific susceptibility loci could display phenotypically distinguishable forms of AD (e.g. as observed in those carrying mutations for early-onset AD). Thus, the challenge for geneticists is to identify the phenotype which is associated with variation in a particular gene (Freimer and Sabatti 2003). Clinical variation commonly observed in AD may offer a suitable candidate to identify genetically homogenous forms of the disease.

1.4 A brief overview of clinical variation observed in late-onset Alzheimer's disease

The clinical phenotype of LOAD is invariably associated with deficits in several areas of cognition. As noted in section 1.1 deficits in memory and two other areas, including aphasia, apraxia, agnosia and executive functioning, are generally required to formulate a diagnosis of probable AD occurring to DSM-IV (American

Psychiatric Association 1994) and NINCDS-ADRDA (McKhann et al. 1984) criteria. However, a number of behavioural symptoms are commonly displayed by disease sufferers. The following quote is taken from the case report of the first incidence of what is known as Alzheimer's disease:

> 'Sometimes she greets the doctor as if he were a visitor... on other occasions she screams that he wants to cut her open... on others yet she fears him as a threat to her honor as a woman... she seems to have auditory hallucinations. Often she screams for many hours in a horrible voice' (Alzheimer 1907)

The description was of a 51-year-old woman who presented with focal cognitive deficits, but also delusions of jealousy and auditory hallucinations (Alzheimer 1907). Behavioural disturbances in AD can include affective symptoms, agitation, aggression and psychosis (Burns et al. 1990a, b, c). The type, severity and prevalence of behavioural symptoms vary greatly and they are not generally believed to be an inevitable consequence of disease progression (Cummings 2000; Sweet et al. 2003). As such they are not a diagnostic requirement for probable AD.

There is controversy about how best to categorise behavioural symptoms. DSM-IV guidelines suggest the use of additional coding to encapsulate AD with depressive mood and AD with delusions. However, specific guidelines for the diagnosis and classification of behavioural symptoms in AD are not available. Behavioural problems are associated with many serious consequences, including increased functional deficits (Stern et al. 1994a), cognitive impairment (Jeste et al. 1992), increased rate of decline (Neumann et al. 2001; Wilson et al. 2000), earlier institutionalisation (Borson and Raskind 1997; Steele et al. 1990), and increased caregiver distress (Donaldson et al. 1998; Craig et al. 2005a). As such, gaining a more comprehensive understanding of their aetiology is essential.

Another notable aspect of clinical variation in AD is AAO. Symptoms of AD can present at anytime from 30 to 90+ years of age. Despite sharing major clinical and neuropathological features a distinction is often made between those with disease

onset before and after the age of 65. Those with a disease onset less than this are widely termed as 'early onset AD', which is often taken to represent a distinct disease 'sub-phenotype' (Raskind et al. 1995; Whitehouse 1995). However, the classification of early onset AD as a clinical subtype remains controversial. Evidence for this dichotomy comes from numerous studies which have reported clinical differences between those with early- and late- onset AD. For example, early onset forms of the disease have been found to be characterised by shorter survival, more rapid cognitive deterioration, more severe language disturbances and more severe AD related neuropathology (Koss et al. 1996; Sevush et al. 1993; Villareal and Morris 1999). The most compelling evidence for a distinction between early- and late-onset AD comes from genetic studies. As already noted in section 1.3 genetic mutations which cause autosomal dominant AD are almost entirely restricted to those with an early age of disease onset (Villareal and Morris 1999). Others have hypothesised that further categorisation by AAO may be useful for genetic studies of AD. For example, Olson and colleagues (2001) have reported evidence that suggests those with a disease onset over the age 80 years may be linked to a genetic locus on chromosome 21.

It is clear that Alzheimer's disease is a clinically heterogeneous disorder. Currently, little is known about the underlying causes of the clinical differences observed in AD. Gaining a more comprehensive understanding may aid both the study of the clinical heterogeneity and AD as a whole. A more detailed overview of the clinical variation observed in AD can be found in sections 2.1.1 to 2.1.3 of this thesis.

1.5 General aims and outline of this thesis

Genetic and epidemiological studies of LOAD rely largely on comparisons made between 'cases', diagnosed to published criteria for AD, and healthy elderly individuals. Implicit in such an approach is the concept that LOAD is clinically homogeneous and can be defined categorically as either present or absent. However, LOAD is a clinically heterogeneous illness and increasing attention is now being paid to utilizing defined subgroups in the hope of unpicking the complex aetiology of the illness (Olson et al. 2001; Pericak-Vance et al. 2000; Sweet et al. 2003)..

Clinical variation in the disease phenotype offers a means of characterising sub-phenotypes or limiting the effect that 'phenocopies' have on genetic analyses. Using aspects of clinical variation to identify sub-phenotypes has proved successful in identifying genes for other complex disorders (Rioux et al. 2001; Van Eerdewegh et al. 2002), whilst the categorisation by AAO proved crucial to the identification of mutations which cause early onset forms of AD (Lendon et al. 1997).

The aim of this thesis is to investigate the underlying genetic aetiology of the clinical variation observed among AD sufferers. It is important to consider clinical variation from two distinct standpoints. First, clinical variation could identify distinct 'sub-phenotypes', or more homogeneous forms of the disease. Under such a model, genetic variation would increase disease risk but only within a specific sub group of sufferers, for example, among the very elderly. Alternatively, genetic variants may act as 'disease modifiers'. According to such a model, genetic variation would not increase the risk of developing AD, but could influence disease processes and presentation in the presence of neurodegeneration owing to genetic variation at another locus, or environmental influences.

Two aspects of clinical variation will be considered in this thesis. First, behavioural symptoms, which represent a substantial problem in AD (Cummings 2000). Second, age at disease onset, which has already been used to define clinical homogeneous subsets of the disease, will be investigated. Both age at disease onset and behavioural symptoms may fluctuate as a result of genetic variation or may act as clinical markers for disease sub-phenotypes.

In this thesis a three stage approach to identifying genes which are implicated in the clinical heterogeneity observed in AD is presented. In the first stage, the emphasis is placed upon characterising behavioural disturbances. As already noted, a wide variety of symptoms are common among AD sufferers and a number of these can often appear in tandem. This represents a methodological problem to studies aiming to delineate their underlying causes (Borson and Raskind 1997). The evidence to date suggests that certain symptoms in AD occur more frequently

together than one would expect by chance and could, therefore, represent behavioural components (Frisoni et al. 1999). In chapter 2 of this thesis data regarding 12 common symptoms, assessed using the Neuropsychiatric Inventory (Cummings 1997), in a large sample of 1,120 AD cases, will be subjected to principal components analysis. The primary aim of this study is to elucidate behavioural components which will be useful in further genetic analyses, under the assumption that behavioural problems may reflect differing manifestations of common underlying neuropathology. This approach has the further advantage of reducing the dimensionality of the data. This is particularly beneficial in this exploratory investigation as it reduces the number of statistical tests required in subsequent analyses. Before proceeding to investigate the genetic underpinnings of behavioural components, or age at disease onset, it is important to determine if they are likely to be subject to genetic influence. In chapter 3 of this thesis, the familiality of AAO and behavioural components will be assessed using a large sample of affected sibling pairs. Familial clustering of age at disease onset and behavioural components would suggest that they are genetically modified and provide justification for future genetic studies.

In chapter 4, aspects of clinical variation which show evidence of being genetically influenced will be used to test for linkage using a regression based method of covariate analysis. Linkage analysis provides a means of locating regions which are likely to harbour genes which increase susceptibility to a particular phenotype. As noted in section 1.3 and discussed in more detail in sections 4.1.4 and 4.1.5 linkage studies have yielded a number of regions which could harbour genes that increase susceptibility to AD. However, studies to date have been characterised by a lack of consistency and failure to replicate positive findings (Bertram and Tanzi 2004). Incorporating covariates into linkage analysis has two main advantages. First, genetic variation which does not increase susceptibility to the disease, but rather modifies its progression can be identified. Second, covariate linkage analysis allows for locus heterogeneity owing to the covariates. Using partly overlapping samples, Myers and colleagues (2002) and Blacker and colleagues (2003) have reported two of the largest genome screens for LOAD to date. In chapter 4, data from these two studies is combined to provide a large sample of relative pairs, well characterised in terms of phenotypic variation,

genotyped on a dense grid of markers. The identification of regions harbouring loci which increase susceptibility to genetically modified aspects of clinical variation will set the stage for future linkage and association studies.

Chapter 2

Phenotypic characterisation of late-onset Alzheimer's disease

2.1 Introduction

2.1.1 Clinical presentation of late-onset Alzheimer's disease.

Alzheimer's disease (AD) is characterised by an insidious onset and progressive decline in memory and cognitive abilities (McKhann et al. 1984). Individuals with AD show fluctuations in the severity of cognitive impairment over days or weeks, but over a number of years the pattern is one of unavoidable decline (Mohs 2005). Rate of disease progression varies, however on average one would expect to observe progression from disease onset to terminal stages in 7 to 10 years (Larson et al. 2005; Jost and Grossberg 1995; Knopman et al. 1988).

Symptoms of AD can present from the age of 30 up to 90+ years of age. DSM-IV (American Psychiatric Association 1994) criteria for the diagnosis of Alzheimer's disease supports the widely used, but somewhat arbitrary, cut-off of 65 years for distinguishing between early- and late- onset AD. The early-onset form of the disease is usually familial and follows an autosomal dominant pattern of inheritance with a high penetrance. Mutations in the gene encoding the Amyloid Precursor Protein (Goate et al. 1991), Presenilin 1 (PSEN1) (Sherrington et al. 1995) and the Presenilin 2 (PSEN2) (Levy-Lahad et al. 1995b) genes account for around 50% of early-onset AD cases (Tandon et al. 2000). AD prevalence increases exponentially between the ages of 60 and 90 (Jorm and Jolley 1998). The late-onset form of AD (LOAD) accounts for around 99% of all cases of the disease (Rocca et al. 1991).

Disease progression in AD is associated with decline in numerous areas of cognition, including deficits in short term memory, attention (Petersen et al. 2001), aphasia (Carlomagno et al. 2005; Grossman et al. 2004), visuospatial ability (Henderson et al. 1989; Lineweaver et al. 2005) and executive functioning

(Cummings 2000). Increasing disease severity and cognitive impairment are also associated with a notable and catastrophic decline in functional abilities and activities of daily living (Harwood et al. 2000; Matsuda and Saito 2005). Long-term memory, general intelligence, vocabulary, reading ability, perceptual abilities and the capability to perform previously well-learned activities, are more severely affected in the later stages of disease development (Mohs 2005). These changes are widely believed to be invariable consequences of disease progression, with AD sufferers eventually losing all semblance of cognitive function (Cummings 2000; Mohs 2005; Morris et al. 1989).

A wide range of behavioural symptoms can also occur during the illness, including delusions, hallucinations, agitation, depression, apathy and irritability. Such symptoms are common but vary greatly among disease sufferers (Cummings 2000). A number of studies have attempted to determine the prevalence of these symptoms in AD. Depression, anxiety and apathy are widely reported to be the most common behavioural symptoms (Lyketsos et al. 2000; Mega et al. 1996; Craig et al. 2005a). Depressive symptoms occur more frequently in AD than they do in the healthy population (Burns et al. 1990c). However, the relationship between depression and AD is a complex one. Such symptoms are common in the early stages of the disease and increased prevalence rates of depression have even been reported in the pre-clinical stages of AD (Gatz et al. 2005a). However, it is unclear whether depression is a preclinical symptom, occurring before the onset of cognitive decline, or whether it acts as a risk factor for AD. The relationship between AD and depression is further complicated as cognitive deficits are often associated with depression in the absence of dementia (Abas et al. 1990), which can make the distinction between depression and AD difficult, especially in the early stages of disease development. Although symptoms of depression are common in AD they often occur in the absence of a major depressive episode (Purandare et al. 2001), with the prevalence of dysthymia being approximately double that of major depression. Starkstein and colleagues (1997) performed longitudinal assessments with a consecutive series of AD patients and found major depression to be a longer lasting mood change. Those who had dysthymia were less likely to be depressed after 18 months of follow up than those who met criteria for major depression at study entry.

Apathy is also among the most common behavioural disturbances observed in patients with AD (Craig et al. 2005a). It is broadly defined as a loss of motivation and manifests in behaviours such as diminished initiation, poor persistence, lack of interest, indifference, low social engagement, blunted emotional response and lack of insight. Apathy often becomes apparent early in the clinical course and has been shown to increase in severity in tandem with worsening cognitive abilities (Mega et al. 1996). Prevalence estimates indicate that around a half of AD cases experience symptoms of apathy (Starkstein et al. 1995; Weiner et al. 2005; Craig et al. 2005a), however it has been suggested that as many as 92% of patients in the later stages of disease development will have displayed apathetic behaviour at some point during their illness (Mega et al. 1996). There is a considerable overlap between symptoms associated with depression and apathy, for example loss of interest or pleasure and fatigue could be used as indicators for both symptoms. However, certain behaviours are specific to either apathy or depression; for example, suicidal ideation and pessimism are indicators of depression whereas poor persistence and indifference are more in coupling with apathy (Landes et al. 2001). As such it is possible for trained observers to delineate between the two symptoms in most cases. There is evidence that the close association noted between apathy and depression results from the symptoms sharing similar neurobiological underpinnings, rather than being due to an overlap in non-specific symptoms (Landes et al. 2001). However, there are numerous reports where depression is present in patients without symptoms of apathy, and vice versa (Marin et al. 1994), therefore they can be viewed as distinct and discriminable symptoms.

Anxiety can take numerous forms in AD. In its mildest form it manifests as irrational avoidance of places or situations and worry when left alone. In its more severe form it can be devastating, causing nervousness resulting in hyperventilation, shaking and marked increases in heart rate. As such it can impact on the patient's ability to function socially and in activities of daily living. Increased anxiety occurs in up to 70% of AD patients during the course of their illness (de Toledo et al. 2004; Ferretti et al. 2001; Teri et al. 1999) and is associated with worsening abilities to perform activities of daily living (Teri et al.

1999). It is frequently viewed as synonymous with other behaviours, notably agitation and aggression, or as a component of a broader syndrome, such as psychosis or depression (Mega et al. 1996; Mintzer et al. 2001). However, several studies have reported anxiety to occur as a distinct syndrome in up to 45% of patients (Burns et al. 1990d; Mega et al. 1996; Teri et al. 1999).

Psychotic symptoms are commonly experienced by AD sufferers, especially in the later stages of disease development (Paulsen et al. 2000; Sweet et al. 2003). Delusions, misidentification syndromes and hallucinations are the common types of psychotic symptoms displayed (Cummings 2000). A wide variety of delusions have been reported, including persecutory delusions, delusions of infidelity, delusions of abandonment and delusions that deceased individuals are still living (Sweet et al. 2003; Tariot et al. 1995). Misidentification delusions are also frequent in AD patients, for example the belief that a family member is someone else, or that one's home is not one's own. Delusions in AD are typically non-bizarre and simple, and seem to differ somewhat from the more complex and bizarre delusions seen in patients with schizophrenia (Jeste and Finkel 2000). They are reported to occur in between 16% to 70% of AD patients (Bassiony and Lyketsos 2003). Hallucinations are less common, occurring in around 7% to 30 % of patients (Marin et al. 1997). They can occur in any sensory modality, but visual and auditory hallucinations are more common (Jeste and Finkel 2000; Tariot et al. 1995) occurring in around 19% and 12% of patients, respectively (Bassiony and Lyketsos 2003). In addition to problems arising from population heterogeneity, prevalence estimates for delusions and hallucinations vary greatly due to differences in how AD and symptoms are diagnosed. Population based studies (de Toledo et al. 2004; Steinberg et al. 2003) and studies which assess symptom development longitudinally (Cohen-Mansfield et al. 1998; Deutsch et al. 1991; Lopez et al. 1991; Paulsen et al. 2000; Rosen and Zubenko 1991) are more likely to provide accurate estimates of symptom prevalence. Reviewing this literature it seems likely that delusions occur in around 45% of patients, whereas hallucinations are less common, occurring in approximately 25% of AD sufferers.

Other common symptoms which occur during the course of AD are agitation, sleep disruption, appetite disturbances and disinhibition. Agitation, with aggressiveness

and physical combativeness, is of particular concern as it represents one of the most challenging behaviours observed in the course of AD (Rabins et al. 1982; Reisberg et al. 1987; Ryden 1988). Between 30% to 70% of patients display notable symptoms of agitation (Mirakhur et al. 2004; Senanarong et al. 2004), which become more prominent in the moderate to severe stages of disease development (Chen et al. 1998; Levy et al. 1996a). Sleep and appetite disturbances have afforded less attention in the literature to date. Sleep disturbances can take the form of insomnia, early morning awakening and waking many times during the night. The relationship between sleep disturbances and AD is complicated as normal aging is often associated with disruption in circadian rhythm (Harper et al. 2005). However, sleep disturbances in AD occur more frequently than in the normal aging population, affecting around 18% to 53% of patients (Aalten et al. 2003; Harper et al. 2005; Mirakhur et al. 2004). As with many behavioural disturbances sleep disruption becomes more apparent in concert with increasing functional and cognitive decline (McCurry et al. 2004). Eating disturbances can take the form of alterations in food preferences, appetite and eating habits. These symptoms are more common in frontotemporal type dementia (FTD) (Bozeat et al. 2000), with one report suggesting that some form of eating disturbance is an invariable component of FTD (Ikeda et al. 2002). They are less marked in AD, occurring in 24% to 63% of patients (Aalten et al. 2003; Bozeat et al. 2000; Ikeda et al. 2002; Mirakhur et al. 2004).

Disinhibition is one of the least studied behavioural symptoms occurring in AD. It can take numerous forms, including abnormal motor behaviour, hypomania, loss of insight, egocentrism, and poor self-care (Starkstein et al. 2004). Once present it has been shown to be persistent through the course of the disease (Starkstein et al. 2004). Cross sectional studies suggest that between 12% and 30% of patients experience symptoms of disinhibition (Aalten et al. 2003; Frisoni et al. 1999).

It is clear that behavioural symptoms represent a significant problem in AD. As already highlighted, they occur in a substantial proportion of cases. They are also associated with many serious consequences for the patient, caregivers and the wider society. Many behavioural symptoms, including apathy, psychosis and agitation (Jeste et al. 1992; Landes et al. 2001; Stern et al. 1994a; Teri et al. 1990)

are associated with increased functional and cognitive deficits (Craig et al. 2005a). High levels of behavioural disturbance are reported to be associated with increased rate of decline in AD (Jeste et al. 1992; Levy et al. 1996a; Neumann et al. 2001; Wilson et al. 2000), which suggests that such symptoms may be a marker for a more severe and aggressive subtype of the disease. Behavioural symptoms are also associated with increased caregiver distress (Donaldson et al. 1998; Craig et al. 2005a). Shin and colleagues (2005) reported a negative correlation between caregiver quality of life and agitation, anxiety, disinhibition, irritability and a global measure of behavioural disturbance. Others have reported sleep problems to be among the most troublesome behaviours for caregivers (Gaugler et al. 2000; Hope et al. 1998). Psychosis, aggression and sleep disturbances are among the primary reasons for institutionalisation of AD patients (Gaugler et al. 2000; Magni et al. 1996). This is of particular concern as nursing home placement necessitates a difficult psychosocial adjustment for older adults and their families (Gaugler et al. 1999) and is also associated with massive economic costs, which are a significant burden to the wider society (Gaugler et al. 2000).

2.1.2 Factors associated with behavioural symptoms in late-onset Alzheimer's disease.

It is important to consider variables which are associated with clinical variation in AD. To date the aetiology and pathophysiology of behavioural disturbances are largely unknown (Cummings 2000; Esiri 1996), however, numerous biological and clinical correlates have been reported with individual behavioural symptoms.

The relationship between global measures of cognitive function and behavioural symptoms is widely reported, whilst the relationship between behavioural symptoms and selective aspects of cognition is less studied (Spalletta et al. 2004). Numerous symptoms, including psychosis, anxiety and agitation have been shown to be associated with general measures of cognitive impairment (Levy et al. 1996a; Paulsen et al. 2000; Porter et al. 2003; Turvey et al. 2001). However, such associations are not reported in all studies (Marin et al. 1997; Migliorelli et al. 1995a; Migliorelli et al. 1995b). Behavioural symptoms have also been associated

with increased rate of decline (Burns et al. 1990b; Neumann et al. 2001; Rosen and Zubenko 1991; Spalletta et al. 2004). The relationship between behavioural problems and cognitive impairment could point towards the existence of a more aggressive form of the disease, characterised by a more rapid decline in cognition and excess behavioural symptoms. However, others have reported that cognition and behavioural symptoms represent independent dimensions of AD (Spalletta et al. 2004; Wobrock et al. 2003), suggesting that different neurobiological systems could be implicated in the pathogenesis of cognitive and behavioural disturbances seen in AD. Discrepancies between studies could result from the different aspects of cognition measured. It is possible that only certain cognitive abilities are associated with particular behavioural symptoms, if these are not assessed in a particular study then the association will be missed. Indeed, studies which have looked at specific forms of cognitive deficit have reported associations between psychotic symptoms and frontally mediated cognitive abilities such as conceptualisation, verbal fluency and abstraction (Flynn et al. 1991; Jeste et al. 1992; Perez-Madrinan et al. 2004). Behavioural symptoms that increase in prevalence among patients with greater cognitive impairment may be linked with the cholinergic dysfunction associated with AD (Cummings and Back 1998).

The prevalence and severity of a number of symptoms have been shown to vary as a function of both age at disease onset and current age. For example, anxiety (Porter et al. 2003), misidentification syndromes (Burns et al. 1990b) and depression (Lawlor et al. 1994) are reportedly more common in those with a younger AAO. Whereas increased psychosis, less agitation and less depression have been shown to be associated with increasing age (Levy et al. 1996a). Small but inconsistent gender differences have been reported with individual behavioural symptoms and global measures of behavioural disturbance (Landes et al. 2001; Ott et al. 1996). Some have suggested that increased behavioural problems, in particular agitation, are more common among females (Levy et al. 1996a). However, others have found that symptoms such as delusions and misidentification syndromes are more common among males (Burns et al. 1990a, b).

Behavioural problems are likely to be a manifestation of underlying structural brain damage (Fairburn and Hope 1988), however the neuropathology of behavioural symptoms is far less understood than the neuropathology of the cognitive deficits observed in AD, or even mild cognitive impairment (MCI) (Kordower et al. 2001). A relatively small number of patients have been comprehensively assessed in life to allow the study of neuropsychiatric symptoms in relation to autopsy findings. In the few available studies, only a small number of brain regions were investigated and only a few neuropathological and neurochemical parameters were assessed (Cummings 2000). Studies using magnetic resonance imaging (MRI), computed tomography (CT) and single photon emission tomography (SPECT) have identified areas of the brain associated with behavioural symptoms, however there is currently little consensus among the findings (Cummings 2000). This heterogeneity requires further investigation in larger, well characterised, samples. A full review of the neuropathological and neurochemical studies of behavioural symptoms in AD is beyond the scope of this thesis, however comprehensive reviews are provided by Cummings and colleagues (2000) and McIlroy and Craig (2004).

Evidence that certain symptoms result from specific neuropathological deficits can also be gleaned from studies comparing the prevalence rates of behavioural symptoms across different forms of dementia. For example, there is a suggestion that depression is more common in vascular dementia and that delusions occur more frequently in AD (Lyketsos et al. 2000), whereas visual hallucinations are more common in dementia with Lewy bodies (DLB) (Geser et al. 2005). Frontotemporal dementia is characterised by changes in eating preference, increased levels of disinhibition, apathy and stereotyped behaviours (Ames et al. 1994; Bozeat et al. 2000). It is plausible and likely that the differing profile of behavioural changes observed between dementia types is a direct consequence of differential neuropathological changes associated with the different diseases.

Genetics of behavioural symptoms in AD have received increased attention over the last 10 years. A number of positive associations have been reported, for example polymorphisms in the serotonin and dopamine receptor genes have been significantly associated with psychotic symptoms in AD (Assal et al. 2004; Holmes

et al. 1998a; Holmes et al. 2001; Lam et al. 2004; Nacmias et al. 2001; Sweet et al. 1998), however, such associations have failed to be replicated elsewhere (Craig et al. 2004b). A few studies have reported relationships between Apolipoprotein E (APOE) genotype and depressive symptoms (Ballard et al. 1997; Ramachandran et al. 1996; Scarmeas et al. 2002), although the majority found no relationship (Craig et al. 2005b; Forsell et al. 1998; Gabryelewicz et al. 2002; Holmes et al. 1998b; Levy et al. 1999; Steffens et al. 2003). On the whole the study of behavioural genetics in AD is in its infancy and the findings to date are inconsistent. This thesis is primarily concerned with genetic variation which is associated with clinical heterogeneity observed in AD.

It is important to consider clinical variation, including the presence or absence of behavioural symptoms, from two distinct standpoints. First, clinical variation may identify 'sub-phenotypes', or less heterogeneous forms of AD, which have a distinct aetiology from AD in general. Early onset forms of familial AD provide a primary example of this, with both AAO and disease duration differing between those carrying the PSEN1, PSEN2 and APP early onset AD mutations (Lippa 1999). Under the 'sub-phenotype hypothesis' genetic variation that increases risk of developing a particular sub-phenotype (e.g. AD with psychosis) is unlikely to increase risk of developing other distinct types of AD (e.g. AD without psychosis). Behavioural symptoms are a good candidate for defining homogeneous disease sub-phenotypes. Although common, such symptoms are very rarely reported to occur in *all* patients and as such it would appear that they are not merely a consequence of increasing disease severity. It therefore follows that patients from clinically heterogeneous groups may be affected by differing underlying aetiologies, or may show variable distribution of burden in brain pathology (Larner 2005). This level of heterogeneity suggests that Alzheimer's disease, as currently defined, may more closely resemble a syndrome with multiple contributing aetiologies rather than a disease with a unitary cause (Zubenko 1997). To such ends psychosis (Sweet et al. 2003) and depression (Zubenko 1997) have been suggested as symptoms which may mark disease sub-phenotypes that could be useful for further analysis.

An alternative hypothesis is that genetic variation may influence the phenotypic expression of the disease without being directly involved in its aetiology. According to such a model, genetic variation would not increase the risk of developing AD but alter aspects of the clinical presentation in the presence of neurodegeneration resulting from other genetic and/or environmental causes. For example, polymorphisms in the serotonin and dopamine receptor genes have been significantly associated with psychotic symptoms in AD (Assal et al. 2004; Holmes et al. 1998a; Holmes et al. 2001; Lam et al. 2004; Nacmias et al. 2001; Sweet et al. 1998), but do not appear to increase the risk of developing AD itself. Such genetic variation could also increase the prevalence of similar symptoms in other disorders. Indeed, there is growing evidence that genetic variation can modify clinical presentation across disease boundaries. For example, AAO in both AD and Parkinson's disease have been linked to the same region of chromosome 10 (Li et al. 2002), whilst linkage has been shown to the same region of chromosome 6 in studies looking at AD with psychosis (Bacanu et al. 2002), bipolar disorder (McQueen et al. 2005) and schizophrenia (Levinson et al. 2000). These findings are largely preliminary, it therefore remains to be seen if these genomic regions actually contain genetic variation which influences disease presentation across these diseases.

The traditional approach of classifying psychiatric disorders as categorical, homogeneous entities have been questioned in other psychiatric illnesses, including depression (Korszun et al. 2004; Parker 2000), schizophrenia (Hallmayer et al. 2005) and bipolar disorder (Fisfalen et al. 2005). The analysis of disease sub-phenotypes has already proven successful in identifying genes for other complex disorders. For example, the inclusion of bronchial hyperresponsiveness as a covariate in linkage analysis of asthma refined the linkage region and resulted in the identification of ADAM33 as the first susceptibility gene which increases risk of developing asthma (Van Eerdewegh et al. 2002). Likewise, studies of Crohn's disease have used reduced AAO to define a disease subtype in linkage analysis and succeeded in reducing their LOD-1 regions from 50cM to 5cM (Rioux et al. 2001). As a result increasing attention is now being paid to defining and analysing sub-phenotypes. These studies have primarily relied on clinical variation of

disease, or co-morbid illness, to identify distinct sub-phenotypes. For many diseases this approach is in its infancy.

In order to effectively study both the disease modifying gene and the sub-phenotype hypotheses it is important to accurately characterise the clinical differences observed within AD and then demonstrate familial aggregation. To date, there is little consistency between studies aiming to identify the correlates of behavioural symptoms in AD. Some have argued that these inconsistencies arise from the number of symptoms that can appear in tandem in some patients, which leads to methodological problems (Borson and Raskind 1997). Frisoni and colleagues (1999) hypothesised that the many behavioural symptoms observed in AD represent a lesser number of underlying components, and that symptoms of the same component are more likely to co-occur. This is supported by the numerous reports of symptoms which are observed together more commonly than would be expected by chance. For example, agitated behaviour is often associated with delusions and hallucinations (Burns et al. 1990d; Lyketsos et al. 1999; Ott et al. 1996; Rapoport et al. 2001), whilst others have reported that depressive symptoms are associated with both aggression (Lyketsos et al. 1999) and psychosis (Holmes et al. 1998b). The identification of behavioural components has many benefits. First, they could be beneficial to the management and treatment of behavioural symptoms (Lawlor and Bhriain 2001), as they may help clinicians in their choice of treatment since groups of related symptoms are likely to respond to the same therapeutic intervention (Aalten et al. 2003; Lawlor and Bhriain 2001; Street et al. 2001). Second, the existence of components may point to possible common underlying neurobiological determinants that cut across a number of dementia subtypes and even other disease entities. Studying such components may offer vital clues about the underlying pathology of AD and could provide more consistent findings than looking at individual symptoms.

Factor analysis offers a means of assessing whether certain behavioural symptoms occur more frequently together, and is useful for detecting behavioural components present in AD. It assumes that behavioural components exist and that some symptoms, and not others, characterise these components. Cluster analysis and latent class analysis offer alternative methods of characterising behavioural

symptoms. Both of these types of analyses examine the relationship between the chosen variables and then place participants into particular groups or classes. Factor, cluster and latent class analysis all have the advantage that they reduce the dimensionality of the observed data, which has practical benefits for researchers as they produce a smaller set of variables which require analysis. Such approaches also allow pertinent questions to be asked of the data, for example should delusions and hallucinations be viewed as separate entities, or do they make up a broader behavioural component of psychosis. Cluster and latent class analysis are well suited to studying the sub-phenotype hypothesis, as they place participants into groups which are optimally different. However, they are less compatible with studying disease modifying hypotheses, especially if one assumes that phenotypic variation is a quantitative trait which exists in the whole study population. Factor analysis is less concerned with sorting participants into groups or classes. Instead it provides a means of reducing a large number of variables into a smaller number of underlying components, under the assumption that the original variables are in some way related and represent differing manifestations of a common underlying cause. Participants can then be scored quantitatively against *each* component. As such factor analysis provides information which can be used to inform the study of sub-phenotypes, but does not preclude investigation of disease modifying hypotheses.

2.1.3 Factor analysis studies of behavioural symptoms in late-onset Alzheimer's disease.

Numerous studies have used methods of factor and cluster analysis to identify behavioural components in AD (Aalten et al. 2003; Amer-Ferrer et al. 2005; Devanand et al. 1992; Frisoni et al. 1999; Fuh et al. 2001; Gauthier et al. 2005; Harwood et al. 1998; Haupt et al. 1998; Hope et al. 1997; Mack et al. 1999; Marin et al. 1997; Matsuoka et al. 2003; McShane 2000; Mirakhur et al. 2004; Moran et al. 2004; Ott et al. 1996; Spalletta et al. 2004). Analyses have been based upon a variety of different scales used to characterise behavioural symptoms, of which the Neuropsychiatric Inventory (NPI) (Cummings 1997) is the most common.

A summary of the factor analytic studies performed using the NPI to assess AD patients can be seen in figure 2.1.

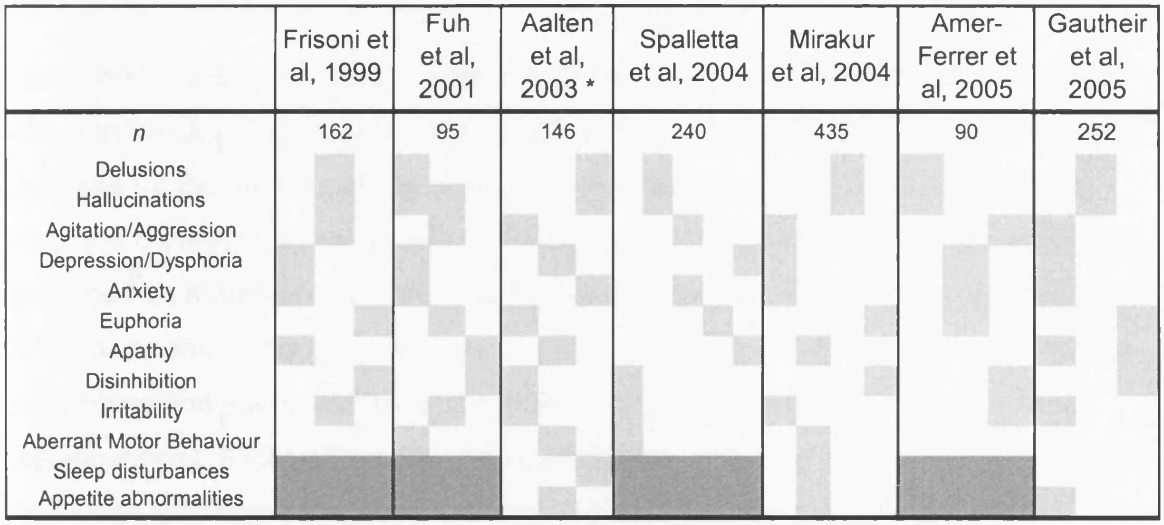


Figure 2.1 Summary of factor analytic studies using the NPI to study behavioural symptoms in AD. Each column represents a component, with "shaded" symptoms loading onto that component. The studies presented by Frisoni, Fuh, Spalletta and Amer-Ferrer used the earlier 10 item version of the NPI which does not assess sleep and appetite disturbances.
* Results from sub-analysis using just cases with probable AD presented.

Frisoni and colleagues (1999) used the 10 item version of the NPI and reported dimensions of 'mood' (containing items depression/dysphoria, anxiety and apathy), 'psychosis' (containing items agitation, hallucinations, delusions and irritability) and 'frontal behavioural features' (containing items relating to disinhibition and euphoria). Fuh and colleagues (2001) used the Chinese version of the NPI (Leung et al. 2001) and reported a three-factor solution, with factors defined as 'mood and psychosis' (including items relating to delusions, hallucination, depression/dysphoria, anxiety and aberrant motor behaviour), 'psychomotor regulation' (that included hallucinations, agitation, euphoria and irritability) and 'social engagement' (which included disinhibition and apathy). However, the samples used by Frisoni and colleagues and Fuh and colleagues were relatively small (n=162 and n=95 respectively) and no corrections were made for either disease duration or severity, which are known to influence the occurrence of behavioural symptoms (Finkel 2001). Also, Fuh and colleagues used a sample of patients still in their home environment or living with family. This is likely to have biased their findings, as those with more severe and frequent behavioural

symptoms often create a greater burden for caregivers (Kaufer et al. 1998) and are more likely to be placed in residential or nursing care (Steele et al. 1990).

Aalten and colleagues (2003) used the 12 item version of the NPI in a sample of 199 individuals suffering with various types of dementia. They performed separate analyses for the whole sample, those in the mild to moderate and more severe stages of dementia. They reported a three-factor solution, which comprised factors described as hyperactivity (including agitation, euphoria, disinhibition and aberrant motor behaviour), mood/apathy (depression/dysphoria, apathy, anxiety, night time behaviours and eating abnormalities) and 'psychosis' (including delusions and hallucinations). Slight differences were reported in the mild and severe groups. In those with mild dementia, aberrant motor behaviour loaded higher with the 'mood/apathy' factor, not with 'hyperactivity', whereas, sleep disturbances were in the 'psychosis' factor, rather than the 'mood/apathy' factor. It is possible that these symptoms may be more related to the severity of dementia. The sample consisted of individuals with differing dementia diagnoses. Using a sub sample of 146 cases diagnosed with probable AD, factor loadings similar to the mild dementia group were reported.

Spalletta and colleagues (2004) analysed the factor structure of the NPI, taking into account cognitive function, in a sample of 240 AD patients. They proposed that the variance contained in the NPI could be accounted for by five factors, including 'hyperactivity', 'psychosis', 'anxiety', 'mood' and 'mood/anxiety'. However, they selected factors based on eigenvalues greater than 0.8. Using an eigenvalue cut-off of 1, fewer factors were used to account for the common variance in the NPI and the symptom loadings were increasingly comparable to those reported elsewhere. In a much smaller sample comprising 90 AD patients, Amer-Ferrer and colleagues (2005) reported that the NPI comprised three components, including psychosis, affective and discontrol syndromes. In their study, hallucinations and delusions loaded together to form the 'psychosis' syndrome, depression/dysphoria, anxiety and euphoria formed the 'affective' syndrome whilst agitation, irritability and disinhibition made up a 'discontrol' component. The compounding effect of disease severity was not controlled for in the studies reported by Spalletta and colleagues and Amer-Ferrer and colleagues.

Using the same sample as used in their earlier study (Craig et al. 2005a), Mirakhur and colleagues (2004) have reported the largest factor analysis of the NPI using AD patients (n=465). Four components were identified representing 'affect' (including depression/dysphoria, anxiety, irritability and agitation), 'physical behaviour' (comprising symptoms of apathy, aberrant motor behaviour, sleep and appetite disturbance), 'psychosis' (including delusions and hallucinations) and 'hypomania' (including disinhibition and euphoria). Of particular note was the high loading of items relating to aggression/agitation within a depressive symptom factor, which had not previously been observed in factor analytical studies using the NPI. In the study by Mirakhur and colleagues no allowance was made for disease severity, although patients were required to have suffered with dementia for greater than 3 years, with a mean disease duration at assessment of 5.7 years. Hence, these data offer a fair reflection of symptom development over the lifetime of the illnesss. Furthermore, the component solution was stable to various methods of rotation, indicating that they are unlikely to represent artefacts of a single rotation.

Finally, Gauthier and colleagues (2005) used the NPI to assess behavioural symptoms in response to memantine. They reported three factors labelled 'hyperactivity' (comprising symptoms of agitation/aggression, depression/dysphoria, anxiety, irritability, apathy and eating disturbances), 'psychosis' (including delusions and hallucinations) and 'mood/apathy' (with high loadings in elation, disinhibition and apathy). These results were largely comparable to those reported elsewhere, with hallucinations and delusions forming a psychosis factor and depression/dysphoria and anxiety loading highly together onto a separate factor. In coupling with a number of previous studies disinhibition and euphoria loaded highly together, this indicates that a behavioural component characterised by lack of control may exist. The sample used in this analysis differed somewhat from those used elsewhere as all participants were in the moderate to severe stages of disease severity. This is likely to have biased their results, however it increases the likelihood that participants will have passed through the period of risk for developing many behavioural symptoms, increasing the reliability of the resulting factor loadings. In all of the analyses using the NPI most symptoms had high factor loading. Also, each study reported a core set of

symptoms which loaded onto conceptually similar factors across analyses, for example hallucinations and delusions consistently formed one factor, as did depression/dysphoria and anxiety. This supports the hypothesis that components of behavioural symptoms exist.

Similar findings have been observed in factor analytic studies based on instruments other than the NPI (Devanand et al. 1992; Haupt et al. 1998; Hope et al. 1997; Marin et al. 1997; Matsuoka et al. 2003; McShane 2000; Ott et al. 1996). Hope and colleagues (1997) used the Present Behavioural Examination to assess 97 participants with either AD or vascular dementia. Three behavioural components were revealed, representing 'over activity', 'aggressive behaviour' and 'psychosis' (including items relating to anxiety, persecutory delusions and hallucinations). In an extension to this study, including 104 participants they reported these components remained stable over a 24 month period (McShane 2000). Behavioural symptoms vary in AD and vascular dementia (Lyketsos et al. 2000), therefore combining patients with differing diagnoses is likely to affect the factor structure in such analyses. Using the dementia behaviour disturbance scale, Ott and colleagues (1996) assessed 125 cases with probable AD and reported components labelled 'antisocial', 'disorientation', 'apathy', 'amnesia', 'agitation' and 'reclusiveness'. It is difficult to draw comparisons between this and other studies as several common behavioural symptoms were not assessed in their sample, including delusions and hallucinations.

A number of studies have used the BEHAVE-AD (Reisberg et al. 1996) scale to assess behavioural symptoms in AD. The BEHAVE-AD assesses a narrower range of symptoms than the NPI (for example, it does not cover symptoms relating to changes in appetite or irritability) however, different types of delusions, hallucinations, aggression and depressed mood are rated separately, for example delusions of abandonment are rated independently from paranoid delusions. Haupt and colleagues (1998) performed factor analysis of the BEHAVE-AD using data from a limited sample of 48 patients with probable AD and reported four behavioural components. These components represented 'depression', 'apathy', 'psychotic symptoms/aggression' (including items relating to delusions, hallucinations and aggressive behaviour) and 'misidentifications/agitation'.

Interestingly, the factor structure remained stable over 3 weeks of assessments. However, it should be noted that the sample used in this analysis is considerably smaller than recommended for performing factor analyses (Kline 1994). In a larger sample of 151 AD sufferers, assessed using the BEHAVE-AD rating scale, Harwood and colleagues (1998) reported five behavioural components, characterised by 'agitation and anxiety', (agitation, anxiety of upcoming events and other anxieties), 'psychosis' (delusions of theft, suspiciousness/paranoia and visual hallucinations), 'aggression' (verbal aggressiveness, physical threats/violence, fear of being left alone and other delusions), 'depression' (tearfulness and depressed mood) and 'activity disturbance' (wandering and the delusion that one's house is not one's home). It is notable that the sample used in this analysis comprised entirely outpatients and relatively low levels of cognitive impairment were reported. As such the prevalence of behavioural symptoms in this sample was lower than reported elsewhere (Aalten et al. 2003; Cummings 1997; Moran et al. 2004), therefore the factor structure obtained cannot be generalised to the broader population of AD sufferers, especially those in the later stages of the disease. Similarly, Moran and colleagues (2004) used the BEHAVE-AD to assess a sample of 240 AD patients in the mild stages of the disease. They performed latent class analysis and defined three classes, characterised by low behavioural disturbances; affective disturbance and anxiety; and aggression, delusions and agitation. However, the sample was not large enough in this study to perform latent class analysis therefore analyses were restricted to a subset of the items covered by the BEHAVE-AD.

It is difficult to draw comparisons between studies which have used different rating scales and different study populations. However, it is notable that a number of studies using rating scales other than the NPI have identified the behavioural component representing psychosis. Also, in the study by Moran and colleagues anxiety was found to load onto the same behavioural component as depression, which is consistent with several studies which have used the NPI. Harwood and colleagues found anxiety to load onto the same behavioural component as agitation, which has been reported elsewhere (Gauthier et al. 2005; Mirakhur et al. 2004; Spalletta et al. 2004). The finding of similar behavioural components across samples, using different rating scales, increases the possibility that robust

behavioural components have been identified. However, there are numerous inconsistencies between the component structures in factor analytical studies reported to date. These differences could emerge from the use of different sample populations and relatively small sample sizes for factor analytical purposes. It is notable that the majority of studies performed to date have not taken account of disease severity when performing factor analyses. It is likely that a proportion of those in the early stages of the disease will go onto display new behavioural symptoms as their illness develops. Also, almost all symptoms increase in prevalence and severity with worsening disease progression (Cummings 2000), as such correlations between symptoms will be inflated if disease severity is not accounted for. This will undoubtedly affect the factor structure obtained in such analyses. Indeed, it would appear that ambiguous symptoms, which do not load reliably onto any factor, are dependent on the stage of dementia. Given the limitations of previous investigations, further studies are required which rely on large, adequately powered, samples in which behavioural symptoms have been carefully characterised. Furthermore, it is unclear to what extent current findings are biased by the confounding effects of disease severity. As such, future studies may provide more reliable results if this confounding variable is suitably accounted for.

2.1.4 Study design and aims

Behavioural symptoms may fluctuate as a result of genetic variation or may act as clinical markers for disease sub-phenotypes. To identify genetic variation associated with such symptoms it is essential that they are well characterised. The evidence from the foregoing literature review, regarding behavioural symptoms in AD, supports the view that they are common and may reflect differing underlying neuropathology. In addition, the evidence to date suggests that certain symptoms in AD occur more frequently together and could, therefore, represent behavioural components. Future analyses may be better served by focusing on these components rather than individual symptoms. Methods of identifying behavioural components to date have relied on samples which are generally underpowered to perform such analysis, which could account for differences in the results between studies.

In this chapter, the NPI is subjected to principal components analysis using data from the largest sample studied to date, comprising data from 1,120 AD sufferers. Factor analytic studies are largely dependent on the variables that are measured. The Neuropsychiatric Inventory provides data from a broad range of symptoms and was therefore selected to assess behavioural problems in this study. In addition, direct comparisons with previous research are facilitated by using the NPI, as it has been widely used elsewhere. The majority of studies to date have not taken account of disease severity when performing factor analyses. Almost all behavioural symptoms increase in prevalence during the later stages of disease development (Cummings 2000). Hence, the confounding effect of disease severity is likely to affect the component structure obtained in factor analyses. This is a particular problem in cross-sectional studies. Incorporating disease severity as a covariate in such analysis will limit the effect that it has on the component structure. Two methods of incorporating disease severity will be used in this analysis. First, it will be controlled for when calculating the correlation matrix between all symptoms included in the principal components analysis. This has the effect of removing the shared variance between all variables which is attributable to disease severity. Second, separate analysis will be performed for those in the mild to moderate and those in the moderately severe to severe stages of dementia. As the prevalence of behavioural symptoms largely increase with advancing disease progression, those assessed in the early stages of disease development may still harbour a moderate risk of going on to develop particular symptoms. As such, analysis of those in the later stages may provide more reliable results.

Numerous correlates of individual symptoms have been reported, however the findings to date have been largely inconsistent. Analyses may be better served by focusing on behavioural components rather than individual symptoms. As such, supplementary analysis will be performed to investigate the relationship between cognitive function, gender, AAO, APOE genotype, age, years in education and family history of dementia.

In summary, this chapter is primarily concerned with three aims, outlined below:

1) To identify components which represent the variety of behavioural symptoms commonly observed among AD sufferers.
2) To perform supplementary analyses controlling for the potential confounding effect of disease severity.
3) To investigate the relationship between behavioural components with clinical, demographic and genetic variables.

2.2 Methodology

2.2.1 Participants

A total of 1,120 individuals diagnosed with late-onset probable AD were used for the analysis. The sample comprised individuals ascertained from both community and hospital settings in the UK and Ireland, collected as part of the National Alzheimer's Research Initiative, funded by the Medical Research Council and the Alzheimer's Research Trust. AD sufferers were ascertained by four collaborating centres, comprising: Cardiff University School of Medicine, Cardiff; Institute of Psychiatry, London; Trinity College, Dublin; and Cambridge University, Cambridge, who contributed data on 420, 272, 125 and 303 participants, respectively. Data collectors from all centres received comprehensive training on all aspects of the assessment battery before the study commenced. A two-day training exercise was held at the Institute of Psychiatry, led by Professor Simon Lovestone. Regular meetings were held between contributing centres and futher training days were organised periodically to train additional data collectors. The author (PH) was involved in co-ordinating data collection between the 4 centres, including the facilitation of communication between centres, organising regular meetings, training and supervision of data collectors, obtaining and collating clinical information. The author also performed assessments with 270 AD sufferers and/or their informant between July 2001 and December 2004. Ethical permission was obtained from the Multi-centre Research Ethics Committee (MREC), relevant local ethics committees and NHS trusts.

All individuals were Caucasian and diagnosed with probable AD in accordance with the National Institute of Neurological and Communication Disorders and Stroke and the Alzheimer's disease and Related Disorders Associations (NINCDS–ADRDA) clinical diagnostic criteria for AD (McKhann et al. 1984). All diagnoses were made based on a semi-structured interview, with known validity for AD pathology (Holmes et al. 1999) (i.e. positive predictive value of 92%) which included: 1) The Mini Mental State Examination (Folstein et al. 1975); 2) The Cambridge Mental Disorders of the Elderly Examination (CAMDEX; informant interview) (Roth et al. 1986); 3) The Blessed Dementia Scale (Blessed et al. 1968); 4) The Bristol Activities of Daily Living Scale (Bucks et al. 1996); 5)

Webster Rating Scale (Webster 1968); 6) Global Deterioration Scale (Reisberg et al. 1982); 7) Cornell Scale for Depression in Dementia (Alexopoulos et al. 1988a); 8) Neuropsychiatric Inventory (12 Item version) (Cummings 1997). Interviews were primarily conducted with the AD sufferer's next of kin in their own home. All informants were required to have a regular contact with the AD sufferer and be familiar with their illness.

2.2.2 Measures

The Neuropsychiatric Inventory (NPI) (Cummings 1997) was used to assess the prevalence and severity of behavioural symptoms in all participants. The NPI is an informant-based rating scale which evaluates 12 common behavioural symptoms in AD including delusions, hallucinations, agitation, depression/dysphoria, anxiety, euphoria, apathy, disinhibition, irritability, aberrant motor behaviour, night-time behaviour disturbances, and appetite/eating abnormalities. The severity of each symptom is rated categorically from 0-3, with anchor points for 'does not occur' (0), 'mild' (1), 'moderate' (2) and 'severe' (3). The frequency with which the symptom occurs is rated categorically from 0-4, with points for 'never' (0), 'less than once per week' (1), 'about once per week' (2), 'several times per week' (3) and 'once or more per day' (4). Frequency and severity scores are then multiplied to give an overall *domain score* for each symptom ranging from 0 to 12. Content validity, concurrent validity, inter-rater reliability, and test-retest reliability of the NPI are well established and it is commonly used in both research and clinical settings (Cummings 1997).

The Mini Mental State Examination (MMSE) (Folstein et al. 1975) is a widely used measure of cognitive function consisting of 20 questions. The maximum score is 30, with scores of 24 or below being indicative of cognitive impairment. The global severity of dementia in all individuals was rated using the Global Deterioration Scale (GDS) (Reisberg et al. 1982). The GDS is a well established scale used to stage the magnitude of functional and cognitive deficits in dementia. Scores range from 1 to 7, with anchor points for 'no cognitive decline' (1), 'very mild' (2), 'mild' (3), 'moderate' (4), 'moderately severe' (5), 'severe' (6) and 'very severe' (7) cognitive decline. To aid subsequent analysis participants were divided into two

groups based on the severity of their illness at assessment; the groups comprised mild to moderate (GDS 2-4) and moderately severe to severe (GDS 5-7) stages of dementia. Age at onset (AAO) was used in supplementary analyses and was defined as the age at which first symptoms of Alzheimer's disease were observed by friends or family.

A detailed family history was obtained from all participants or their next of kin where necessary. Individuals were characterised as having a positive family history if at least one first degree relative was reported as being diagnosed with dementia. Those individuals with a family history of dementia were further partitioned into those with a family history of AD, where at least one family member was reported to have been diagnosed specifically with AD. Individuals were classified as having no family history of dementia if all their first degree relatives were reported to be cognitively intact and at least two first degree relatives had lived passed the AAO in the AD patient. Individuals were classified as unknown if they had missing data, were adopted, if any first degree relatives had memory problems but were not formally diagnosed with dementia, or if they had less than 2 first degree relatives who had lived passed their age at disease onset. All participants, or their informant, were asked about the total number of years spent in full time education. This included basic schooling, further education, apprenticeships, technical college and higher education.

2.2.3 APOE genotyping

A subset of 878 individuals had APOE genotypes available. Genotyping was performed using standardised methods (Saunders et al. 1993). For the purposes of statistical analysis individuals were categorised as either having no (x / x), one (ε4/ x) or two (ε4/ε4) ε4 alleles.

2.2.4 Statistical analysis

The prevalence of the 12 behavioural symptoms among those in the mild to moderate stages of dementia were compared to the reported prevalence among those with moderately severe to severe AD by constructing 2x2 contingency tables and applying the chi-square goodness of fit test with one degree of freedom. NPI

domain scores (frequency x severity) were considered to be ordinal therefore Kendall's tau-b was used to test the association between disease severity and 'domain scores' on the NPI. Due to the number of tests employed here Bonferroni significance criteria were applied. The association between disease severity *and* prevalence for each of the twelve symptoms on the NPI was assessed (24 tests in total) therefore the Bonferroni-adjusted criterion for significance was 0.002.

Domain scores for each section of the NPI were used to test the component structure of the NPI and determine clusters of symptoms which occur together. As symptom domain scores were considered to be ordinal polychoric correlations were calculated using the POLYCHOR function in SAS 8.02 (SAS Institute Inc., 1999-2001). The resulting NPI polychoric correlation matrix was submitted to Principal Components Analysis (PCA). PCA is a method of reducing a large number of variables into a smaller number of underlying components, under the assumption that the original variables are in some way related and represent differing manifestations of a common underlying cause. Three criteria were used to select the number of components to retain in further rotated analyses, including eigenvalues greater than 1, inspection of the scree plot and theoretical relevance of the resulting components (Kline 1994). Selecting eigenvalues greater than 1 was proposed by Kaiser (1960) and ensures that each component extracts as much, or more, variance from the correlation matrix as would be expected from any one of the original variables. Non-orthogonal rotation of the components was conducted as orthogonal rotation methods (such as varimax rotation) maximise loadings onto components in a way which does not allow for inter correlations between components. Orthogonal rotation was considered inappropriate for this analysis as some relationship would be expected between components due to the underlying disease process in AD. Absolute component loadings greater than, or equal to, 0.4 were deemed substantial and used in the interpretation of components. Cronbach's α coefficients were used to assess the internal consistency of the factors.

A number of behavioural symptoms occur later in the illness; therefore separate analyses were performed for the sample as a whole, those in the mild to moderate stage of disease development and those in the moderately severe to severe

stages of the illness. To further investigate the relationship between disease severity and the component structure, supplementary analyses were performed. First order polychoric correlations were calculated for the 12 NPI domains, controlling for GDS. The resulting correlation matrix was then submitted to PCA as described above.

T-tests were performed to compare component scores, derived from PCA using the full sample, of those in the mild to moderate stages of dementia (GDS 2-4) with those in the moderately severe to severe stages (GDS 5-7) of disease development. Linear regression analysis was used to determine the relationship between individual component scores derived from PCA with gender, AAO, age at assessment, level of cognitive functioning (as measured by the Mini Mental State Examination (Folstein et al. 1975)), years in education and the number of APOE ε4 alleles. A separate analysis was performed to assess the relationship between family history of dementia and AD with the component scores as a substantial proportion of the sample were classified as 'unknown' for family history.

2.3 Results

2.3.1 Characteristics of the sample

Basic characteristics of the sample can be found in Table 2.1.

Table 2.1 Basic characteristics of the sample used in the principal components analysis (n=1120).

		Range
Gender n (%)		
Male	334 (29.8)	-
Female	786 (70.2)	-
GDS n (%)		
Mild to Moderate 2-4	434 (38.7)	-
Moderately Severe to Severe 5-7	686 (61.3)	-
APOE Genotype n (%)		
2/2	6 (0.7)	-
2/3	36 (4.1)	-
2/4	31 (3.5)	-
3/3	299 (34.1)	-
3/4	399 (45.4)	-
4/4	107 (12.2)	-
Family History of dementia, n (%)		
Positive family history of dementia	171 (15.3)	-
Negative family history of dementia	259 (23.1)	-
Unknown	690 (61.6)	-
Mean age at assessment, years (sd)	81.2 (6.5)	62 - 99
Mean age at onset, years (sd)	75.4 (6.8)	60 - 93
Mean number of years in education (sd)	10.4 (2.7)	0 - 23
Mean disease duration, months (sd)	69.5 (44.9)	3 - 312
Mean MMSE Score (0 - 30) (sd)	12.8 (9.0)	0 - 28
Mean NPI Total (0 - 144) (sd)	30.5 (22.2)	0-112

Table 2.2 shows descriptive statistics for the NPI in the full sample, those in the mild to moderate stages of disease development (GDS 2-4) and those in the moderately severe to severe (GDS 5-7) stages of dementia. Apathy, depression/dysphoria and agitation/aggression were the most commonly reported symptoms occurring in 73.2%, 56.8% and 56.1% of patients, respectively. Euphoria was the least common behavioural disturbance, apparent in only 10.5% of patients. All symptoms were more prevalent in the moderately severe to severe patients, compared to those in the mild to moderate stages of dementia ($p<0.002$),

with the exception of depression/dysphoria ($\chi^2(1)= 1.369$, p=0.242) and anxiety ($\chi^2(1)= 2.025$, p=0.155). Similarly, all symptoms showed increased domain scores in the moderately severe to severe patients, compared to those in the mild to moderate stages of dementia (p<0.002), with the exception of depression/dysphoria ($r= 0.069$, n = 1120, p=0.009), euphoria ($r = 0.073$, n=1120, p=0.008) and anxiety ($r = 0.064$, n = 1120, p=0.018).

Table 2.2 Prevalence and severity of the Neuropsychiatric Inventory symptom domains in the full sample, those in the mild to moderate and moderately severe to severe stages of dementia.

Item	Full Sample: GDS 2-7 (n=1120)		Mild to moderate patients: GDS 2-4 (n=434)		Moderately severe to severe patients: GDS 5-7 (n=686)	
	Mean domain score (sd)	% of patients with symptom	Mean domain score (sd)	% of patients with symptom	Mean domain score (sd)	% of patients with symptom
Delusions	2.5 (3.7)	44.8	1.1 (2.4)	27.4	3.5 (4.1)	55.8
Hallucinations	1.2 (2.7)	24.2	0.4 (1.3)	11.3	1.7 (3.2)	32.4
Agitation/Aggression	2.9 (3.8)	56.1	1.5 (2.7)	39.4	3.8 (4.0)	66.6
Depression/Dysphoria	2.6 (3.5)	56.8	2.2 (3.2)	54.6	2.9 (3.5)	58.2
Anxiety	2.1 (3.3)	40.2	1.6 (2.7)	37.6	2.2 (3.6)	41.8
Euphoria	0.4 (1.5)	10.5	0.3 (1.2)	7.6	0.5 (1.7)	12.4
Apathy	5.3 (4.2)	73.2	3.4 (3.5)	61.8	6.3 (4.2)	80.5
Disinhibition	1.3 (2.6)	31.4	0.7 (1.7)	22.6	1.8 (3.1)	37.0
Irritability	2.4 (3.5)	43.0	1.6 (2.8)	37.1	2.8 (3.8)	46.8
Aberrant Motor Behaviour	3.2 (3.8)	51.8	1.7 (2.9)	33.9	4.0 (4.0)	63.1
Sleep disturbances	3.2 (4.0)	50.4	1.8 (3.0)	35.3	4.2 (4.3)	60.1
Appetite abnormalities	3.5 (4.1)	52.9	2.7 (3.6)	44.5	4.3 (4.3)	58.2

2.3.2 Principal components analysis

Inspection of the scree plot, shown in figure 2.2, indicated that either a one, three or five component solution could be appropriate for these data. Eigenvalues derived from the principal components analysis (PCA) can be seen in Table 2.3. Three components had an eigenvalue which exceeded the cut-off of one and each of these components represented clinically meaningful domains, hence a three component solution was deemed most appropriate for these data.

The matrix of the component loadings from the PCA, rotated to oblimin criteria, is presented in Table 2.4. Loadings greater than 0.4 were deemed substantial and included in the following interpretations.

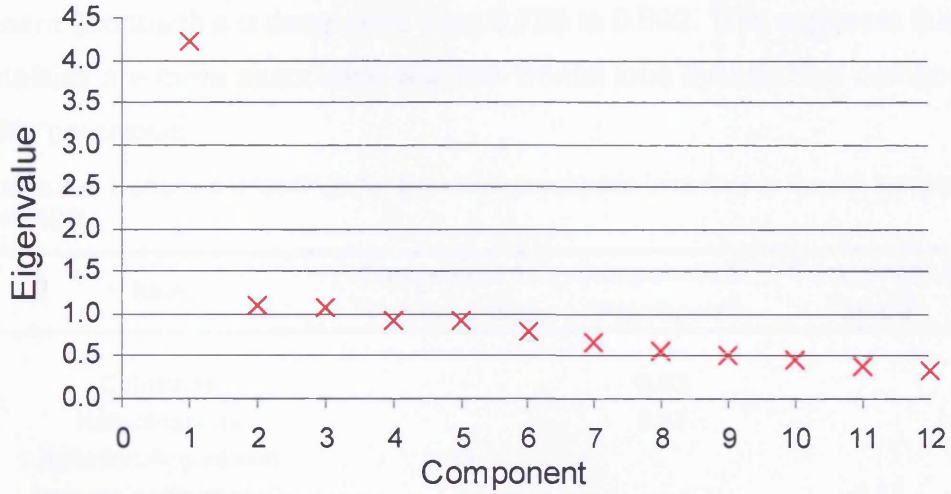


Figure 2.2. Scree plot from principal components analysis of the whole sample

Table 2.3 Eigenvalues and percentage of variance explained for components in the full sample principal components analysis

Component	Eigenvalue	% of Variance
1	4.23	35.22
2	1.09	9.11
3	1.08	9.01
4	0.93	7.77
5	0.93	7.71
6	0.79	6.58
7	0.67	5.59
8	0.57	4.74
9	0.52	4.35
10	0.46	3.83
11	0.39	3.27
12	0.34	2.81

The first component identified accounted for 35.2% of the variance and represented 'frontal lobe dysfunction', with agitation/aggression, euphoria, apathy, disinhibition, irritability, aberrant motor behaviour, sleep and appetite disturbances loading onto this component. The second component reflects 'psychosis', accounting for 9.1% of the variance in the correlation matrix. Delusions and hallucinations loaded highly onto this component, sleep disturbances also had a moderate loading on this factor. Component three, 'mood', accounted for 9.0% of the variance. Reliability analysis found that if sleep disturbances were removed from the 'psychosis' component the Cronbach's α increased slightly from 0.593 to

0.611. Whereas, if they were removed from the 'frontal lobe dysfunction' component Cronbach's α decreased from 0.725 to 0.692. This suggests that sleep abnormalities are more associated with the 'frontal lobe dysfunction' component than with 'psychosis'.

Table 2.4 Component loadings for the Neuropsychiatric Inventory in the full sample (n=1120).

Item	Component 1: "Frontal Lobe"	Component 2: "Psychosis"	Component 3: "Mood"
Delusions	.	0.83	.
Hallucinations	.	0.81	.
Agitation/Aggression	0.53	.	.
Depression/Dysphoria	.	.	0.87
Anxiety	.	.	0.75
Euphoria	0.69	.	.
Apathy	0.52	.	.
Disinhibition	0.77	.	.
Irritability	0.54	.	.
Aberrant Motor Behaviour	0.53	.	.
Sleep disturbances	0.43	0.41	.
Appetite abnormalities	0.51	.	.
Eigenvalue	4.23	1.09	1.08
% of Variance	35.2	9.1	9.0

NB: Absolute values less than 0.4 are not shown

A secondary analysis was performed controlling for the effect of disease severity. The component structure was largely analogous to the one reported above. The only notable differences were that apathy loaded with depression/dysphoria and anxiety, rather than the frontal lobe dysfunction, whilst the loading for appetite disturbances on the frontal lobe dysfunction reduced to 0.324. The matrix of the component loadings from the PCA, rotated to oblimin criteria, controlling for disease severity can be found in appendix 2a.

To further investigate the relationship between dementia severity and the component structure, separate analyses were performed for those in the mild to moderate stages of dementia (GDS 2-4) and those in the moderately-severe to severe stages (GDS 5-7). In the mild to moderate group (n=434) four eigenvalues exceeded the cut-off of one, however the scree plot (figure 2.3) indicated that a three-component solution may also be compatible. To aid comparison between the

analyses of those in the mild to moderate stages and the full sample both the three- and four-component solutions are reported here.

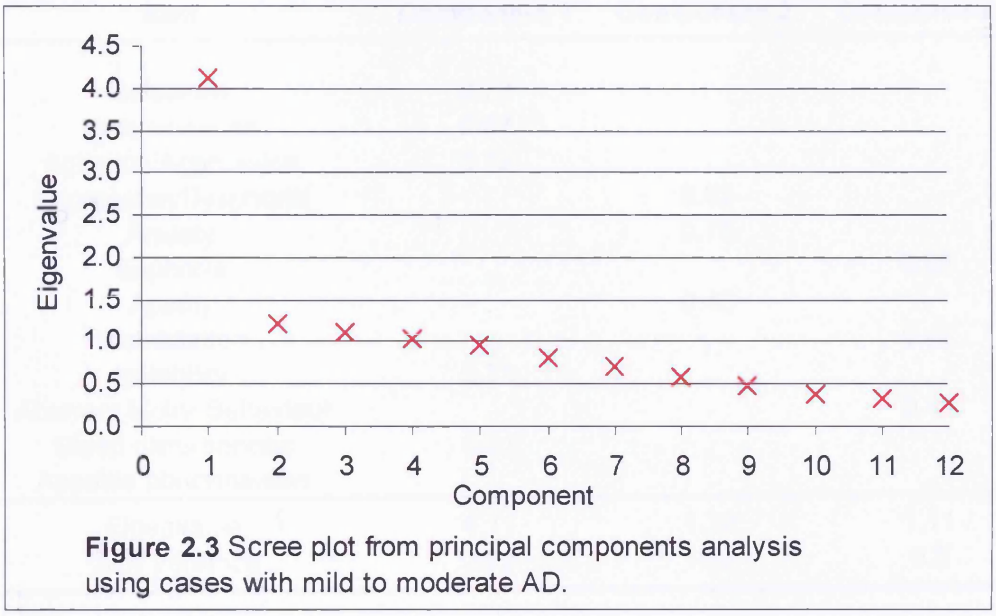


Figure 2.3 Scree plot from principal components analysis using cases with mild to moderate AD.

Component loadings for the three-component solution can be seen in table 2.5. Component one of the three-component solution accounted for 34.3% of the variance and included delusions, hallucinations, agitation/aggression, irritability and sleep disturbances. Apathy loaded with anxiety and depression/dysphoria to make up component two, explaining 10.2% of the variance in the correlation matrix. Component three explained 9.3% of the variance and contained items relating to euphoria, disinhibition and aberrant motor behaviour. Appetite disturbances did not have a substantial loading on any component. The four-component solution differed slightly; delusions and hallucinations loaded with aberrant motor behaviours, sleep and appetite disturbances. Irritability and agitation/aggression loaded together onto component two. Depression/dysphoria, anxiety and apathy again loaded together onto component three. The fourth component was made up of euphoria and disinhibition and explained 8.8% of the variance in the correlation matrix. Component loadings for the four-component solution in the mild to moderate group can be seen in table 2.6.

Table 2.5 Component loadings for the Neuropsychiatric Inventory in mild to moderate AD cases (n=434). Three-component solution.

Item	Component 1	Component 2	Component 3
Delusions	0.73	.	.
Hallucinations	0.67	.	.
Agitation/Aggression	0.72	.	.
Depression/Dysphoria	.	0.82	.
Anxiety	.	0.75	.
Euphoria	.	.	0.86
Apathy	.	0.46	.
Disinhibition	.	.	0.42
Irritability	0.71	.	.
Aberrant Motor Behaviour	.	.	0.41
Sleep disturbances	0.61	.	.
Appetite abnormalities	.	.	.
Eigenvalue	4.11	1.22	1.11
% of Variance	34.3	10.2	9.3

NB: Values less than 0.4 are not shown

Table 2.6 Component loadings for the Neuropsychiatric Inventory in mild to moderate AD cases (n=434). Four-component solution.

Item	Component 1	Component 2	Component 3	Component 4
Delusions	0.60	.	.	.
Hallucinations	0.79	.	.	.
Agitation/Aggression	.	0.83	.	.
Depression/Dysphoria	.	.	0.80	.
Anxiety	.	.	0.70	.
Euphoria	.	.	.	0.87
Apathy	.	.	0.45	.
Disinhibition	.	.	.	0.47
Irritability	.	0.88	.	.
Aberrant Motor Behaviour	0.42	.	.	.
Sleep disturbances	0.58	.	.	.
Appetite abnormalities	0.67	.	.	.
Eigenvalue	4.11	1.22	1.11	1.04
% of Variance	34.3	10.2	9.3	8.8

NB: Values less than 0.4 are not shown

In the moderately severe to severe group (n = 686) three components had eigenvalues greater than the cut-off of one and inspection of the scree plot (figure 2.4) indicated that a three component was appropriate. The three-component solution produced a comparable structure to that obtained in the full sample. The only notable difference was that the loading for sleep disturbances on component

two reduced from 0.412 to 0.398. Component loadings for the NPI among patients in the moderately severe to severe stages of AD can be seen in table 2.7.

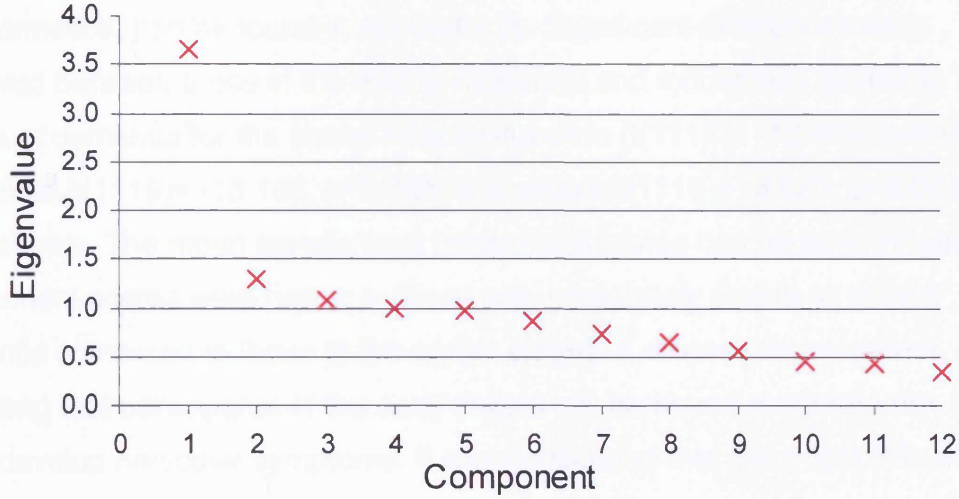


Figure 2.4 Scree plot from principal components analysis using cases with moderately severe to severe AD

Table 2.7. Component loadings for the Neuropsychiatric Inventory in moderately severe to severe AD cases (n=686).

Item	Component 1	Component 2	Component 3
Delusions	.	0.82	.
Hallucinations	.	0.80	.
Agitation/Aggression	0.60	.	.
Depression/Dysphoria	.	.	0.86
Anxiety	.	.	0.69
Euphoria	0.54	.	.
Apathy	0.62	.	.
Disinhibition	0.70	.	.
Irritability	0.58	.	.
Aberrant Motor Behaviour	0.56	.	.
Sleep disturbances	0.40		.
Appetite abnormalities	0.47	.	.
Eigenvalue	3.66	1.29	1.08
% of Variance	30.5	10.8	9.0

NB: Values less than 0.4 are not shown

2.3.3 Component score analysis

Component scores derived from PCA were used to determine the relationship between behavioural components of the NPI and disease severity, gender, AAO, cognitive function, age at assessment, years in education and the number of

APOE ε4 alleles. Component scores were positively skewed; therefore log transformations were performed to normalise the data before further analyses. The distribution of component scores for each component, before and after transformation, can be found in appendix 2b. Significant differences were observed between those in the mild to moderate and moderately severe to severe stages of dementia for the frontal lobe dysfunction ($t(1118)= -12.484$, $p<0.001$), psychosis ($t(1118)= -13.168$, $p<0.001$) and mood ($t(1118)= -3.023$, $p<0.003$) components. The mean transformed component scores can be seen in table 2.8. Component scores were higher in those with moderately severe to severe dementia compared to those in the earlier stages of disease development, indicating that participants in the early stages still harbour a moderate risk of going on to develop particular symptoms. It was considered that this might influence the findings when looking for features associated with the component scores, hence only data relating to participants in the moderately severe to severe stages of dementia were considered in further analyses.

Table 2.8 Mean transformed component scores for those in the mild to moderate and moderately severe to severe stages of dementia.

Component	Mild to moderate patients: GDS 2-4 (n=434)		Moderately severe to severe patients: GDS 5-7 (n=686)	
	Mean	sd	Mean	sd
Frontal lobe dysfunction	0.41	0.36	0.72	0.42
Psychosis	0.69	0.24	0.95	0.36
Mood	0.54	0.40	0.62	0.45

Linear regression models were used to simultaneously assess the effect of clinical variables on component scores. The following variables were entered into the model at step 1: gender, AAO, age at assessment, years in education and number of APOE ε4 alleles. Backward removal was used to remove variables that did not explain a significant proportion of variance in component scores. Due to the obvious relationship between cognitive function and disease duration in AD ($r(1116) = -0.437$, $p < 0.001$ in this dataset), the effect of illness duration was controlled when assessing the relationship between cognitive function and component scores. This was done by entering disease duration at step 2 before cognitive function (MMSE) was entered at step 3.

The linear regression analysis used to assess the impact of clinical variables on the frontal lobe dysfunction scores can be seen in table 2.9.

Table 2.9 Linear regression model for frontal lobe dysfunction component scores as predicted by age at onset, gender, age at assessment, APOE ε4 status, years in education and MMSE (controlling for disease duration) (n=499).

		B†	R^2	R^2 Change	Sig.
Step 1	Age at onset (years)	-.011	.038	.038	<.001
Step 2	Disease Duration (years)	-.010	.042	.004	.149
Step 3	MMSE	.007	.110	.065	<.001
Excluded variable(s)					
Gender *		-076	-	-	.089
Age at assessment (years)		.121	-	-	.123
Years in Education (years)		-.080	-	-	.070
APOE ε4 status **		-	-	-	-
	ε4 / x	.038	-	-	.393
	ε4 / ε4	.050	-	-	.264

Dependent Variable = Component 1 Scores - Frontal lobe dysfunction

† Unstandardised regression coefficient

* Gender: 0 = Female, 1 = Male

** reference group = No APOE ε4 alleles (x / x)

AAO was significantly associated with 'frontal lobe dysfunction' component scores ($\beta = -0.011$, $p < 0.001$). Higher scores were found with an earlier AAO. AAO explained 3.8% of the variance in component 1 scores. After controlling for disease duration cognitive function significantly improved the prediction of component scores, explaining 6.5% of the variance ($R^2_{increase} = 0.065$, $p < 0.001$). Higher component scores were found among those with more severe cognitive impairment. Neither gender ($\beta = -0.076$, p=0.089), age at assessment ($\beta = 0.121$, p=0.123) nor APOE ε4 status significantly predicted 'frontal lobe dysfunction' component scores. Those with more years in education ($\beta = -.080$, p=0.070) showed a trend towards lower component scores which did not meet criteria for statistical significance. Given the sample size it is reasonable to assume that years in education does not have a notable effect on the frontal lobe dysfunction component in this sample.

Table 2.10 shows the linear regression analysis used to assess the impact of the clinical variables on 'psychosis' component scores.

Table 2.10 Linear regression model for the psychosis component scores as predicted by age at onset, gender, age at assessment, APOE ε4 status, years in education and MMSE (controlling for disease duration) (n=499).

		B†	R^2	R^2 Change	Sig.
Step 1	Disease Duration (years)	.003	.002	.002	.384
Step 2	MMSE	-.012	.056	.054	<.001
Excluded variable(s)					
Gender *		-.068	-	-	.127
Age at onset (years)		.011	-	-	.831
Age at assessment (years)		.017	-	-	.704
Years in Education (years)		-.049	-	-	.277
APOE ε4 status **		-	-	-	-
	ε4 / x	-.046	-	-	.308
	ε4 / ε4	.044	-	-	.328

Dependent Variable = Component 2 Scores - Psychosis

† Unstandardised regression coefficient

* Gender: 0 = Female, 1 = Male

** reference group = No APOE ε4 alleles (x / x)

Neither gender (β = -0.068, p=0.127), AAO (β = 0.011, p=0.831), age at assessment (β = -0.017, p=0.704), years in education (β = -0.049, p=0.277) nor APOE ε4 status had a significant effect on 'psychosis' component scores. After controlling for disease duration, cognitive impairment explained 5.4% of the variance in component 2 scores ($R^2_{increase}$ = 0.054, p < 0.001). Higher psychosis component scores were found among those with more severe cognitive impairment.

Neither gender (β = -0.070, p=0.117), AAO (β = -0.058, p=0.240), age at assessment (β = -0.058, p=0.200), years in education (β = -0.045, p=0.312) nor APOE ε4 status had a significant effect on 'mood' component scores. Cognitive impairment had a small effect on 'mood' scores (β = -0.010, p=0.050). Those presenting with more cognitive impairment, after controlling for disease duration, displayed higher mood component scores. However, the effect was small with cognitive impairment explaining only 1% of the variance in 'mood' component scores. A summary of the analysis can be seen in table 2.11.

Table 2.11 Linear regression model for the mood component scores as predicted by age at onset, gender, age at assessment, APOE ε4 status, years in education and MMSE (controlling for disease duration) (n=499).

		B†	R^2	R^2 Change	Sig.
Step 1	Disease Duration (years)	-.012	.003	.003	.188
Step 2	MMSE	-.010	.011	.007	.050
Excluded variable(s)					
Gender *		-.070	-	-	.117
Age at onset (years)		-.058	-	-	.240
Age at assessment (years)		-.058	-	-	.200
Years in Education (years)		-.045	-	-	.312
APOE ε4 status **		-	-	-	-
	ε4 / x	.038	-	-	.398
	ε4 / ε4	.039	-	-	.385

Dependent Variable = Component 3 Scores - Mood

† Unstandardised regression coefficient

* Gender: 0 = Female, 1 = Male

** reference group = No APOE ε4 alleles (x / x)

A separate analysis was performed to assess the relationship between family history of dementia and diagnosed AD with the component scores. In the moderately severe to severe dementia group 176 individuals were classified as having at least one first degree relative with diagnosed dementia, 101 of these being reported as Alzheimer's disease. 115 met criteria for no history of diagnosed dementia. The mean transformed component scores for those with no family history of dementia, those with a positive family history of diagnosed dementia and AD can be seen in table 2.12.

Table 2.12 Mean transformed component scores for those with no family history of dementia, those with ≥1 1st degree relative with diagnosed dementia and those with ≥1 1st degree relative with diagnosed AD

Component	No family history dementia (n=115)		Family history of dementia (n=176)†		Family history of AD (n=101)	
	Mean	sd	Mean	sd	Mean	sd
Frontal lobe dysfunction	0.75	0.47	0.71	0.41	0.69	0.38
Psychosis	0.95	0.34	0.97	0.35	0.96	0.36
Mood	0.64	0.47	0.59	0.46	0.59	0.47

† This group also contains those with a familiy history of diagnosed AD

Those with a family history of dementia or AD showed a trend towards lower scores for the frontal lobe component, however, the differences were not significant, $t(289)=1.197$, $p=0.232$ and $t(214)=1.420$, $p=0.157$ respectively. There was no evidence that psychosis component scores differed in those with and without a family history of dementia, $t(289)= -0.672$, $p=0.502$, or diagnosed AD, $t(214)= -0.214$, $p=0.831$. Likewise, mood component scores did not differ among those with and without a family history of dementia, $t(289)=0.777$, $p=0.438$, or diagnosed AD, $t(214)=.728$, $p=0.467$.

2.4 Discussion

2.4.1 Summary of findings

This is the largest study to date to assess the component structure of behavioural symptoms in AD. The findings show that behavioural symptoms are common and that the majority increase in prevalence and severity among patients in the later stages of dementia compared to those in the mild to moderate stages. Principal components analysis indicated that the twelve symptoms covered by the NPI represent three behavioural components: 'frontal lobe dysfunction', 'psychosis' and 'mood'. These components remained stable after controlling for disease severity and in a separate analysis restricted to data from those in the later stages of disease development. The component structure differed somewhat when restricting analysis just to those in the mild to moderate stages of disease development. Higher scores on the frontal lobe dysfunction component were associated with a lower AAO, whilst both frontal lobe dysfunction and psychosis component scores were elevated in those with more severe cognitive impairment after controlling for disease duration. None of the behavioural components identified were associated with gender, current age, years in education, family history of dementia or number of APOE ε4 alleles.

2.4.2 Discussion of behavioural symptoms and component structure

Apathy was the most frequent and severe symptom reported on the NPI, which is consistent with several other reports (Aalten et al. 2003; Frisoni et al. 1999; Mega et al. 1996; Spalletta et al. 2004; Craig et al. 2005a). In accordance with previous findings, high levels of depression/dysphoria and anxiety were reported (Aalten et al. 2003; Moran et al. 2004; Senanarong et al. 2004; Spalletta et al. 2004; Craig et al. 2005a). Euphoria was the least common symptom reported on the NPI. The prevalence and severity of agitation, disinhibition, irritability, aberrant motor behaviour, sleep and appetite disturbances reported in this sample were comparable to those reported elsewhere (Cummings 2000; Frisoni et al. 1999; Fuh et al. 2001; Craig et al. 2005a; Senanarong et al. 2004). Delusional behaviour and hallucinations were more common than in a number of previous studies (Cummings 2000; Spalletta et al. 2004) and were more in accordance with prevalence rates observed in AD patients residing in special units and nursing

homes (Lange et al. 2004; Matsuoka et al. 2003). This may reflect the large proportion of cases in the moderate to severe stages of dementia represented in this sample.

The component structure obtained in this analysis is similar to those reported elsewhere. For example, Aalten and colleagues (2003) reported a factor labelled 'hyperactivity' which contained items agitation, euphoria, disinhibition, irritability and aberrant motor behaviour, which was very similar to the 'frontal lobe dysfunction' component reported here. When a two-factor solution of the NPI was considered by Spalletta and colleagues (2004) a factor very similar to the 'frontal lobe dysfunction' component was reported, the only differences being the inclusion of anxiety and exclusion of euphoria. Frisoni and colleagues (1999) described a factor that contained items relating to disinhibition, euphoria, apathy and aberrant motor behaviour. In addition, Senanarong and colleagues (2004) recently reported strong correlations between items relating to agitation, irritability, disinhibition and aberrant motor behaviour which persisted after controlling for MMSE score. In the study by Fuh and colleagues (2001), the 'frontal lobe dysfunction' component found in this analysis was separated into two factors, with apathy and disinhibition loading onto a separate factor to agitation, euphoria and irritability. However, Fuh and colleagues used the 10-item version of the NPI and only included data from 95 AD patients. The symptoms included in the 'frontal lobe dysfunction' component have also been reported to load together in other studies that have not used the NPI. For example, Ott and colleagues (1996) used the dementia behaviour disturbance scale, reporting that symptoms comparable to those represented by aberrant motor behaviour, sleep disturbances, appetite disturbances and apathy loaded onto one factor.

'Psychosis' represents the most consistent behavioural component reported to date. Symptoms of delusions and hallucinations have loaded together across all of the studies that have explored the factor structure of the NPI in samples of AD patients. Hallucinations and delusions have also loaded onto the same behavioural component in factor analyses of data from rating scales other than the NPI (Harwood et al. 1998; Hope, 1997; Haupt et al. 1992; Mack et al. 1999; McShane 2000). These symptoms have also been found to constitute a single

component in a study of frontotemporal dementia (Mourik et al. 2004). In this analysis hallucinations and delusions made up an exclusive component, with the questionable inclusion of sleep disturbances. Other studies have reported items of irritability (Frisoni et al. 1999; Mourik et al. 2004), sleep disturbances (Aalten et al. 2003; Matsuoka et al. 2003), agitation (Frisoni et al. 1999), depression/dysphoria (Fuh et al. 2001), anxiety (Fuh et al. 2001) and aberrant motor behaviours (Fuh et al. 2001) to load highly with 'psychosis'.

The 'mood' component has been noted in a number of previous studies (Aalten et al. 2003; Amer-Ferrer et al. 2005; Frisoni et al. 1999; Gauthier et al. 2005; Lange et al. 2004; Mirakhur et al. 2004; Mourik et al. 2004). Also, using the BEHAVE-AD scale for determining presence of behavioural and psychological symptoms, Moran and colleagues (2004) identified a class of AD patients whose illness was characterised by high incidence of depression and anxiety. A number of studies have found apathy to load highly with symptoms of depression and anxiety (Aalten et al. 2003; Frisoni et al. 1999), however these findings are not supported by this analysis.

On the whole it would appear that there are a core set of symptoms that consistently load together. For example, hallucinations and delusions; depression/dysphoria and anxiety; agitation, irritability, disinhibition, aberrant motor behaviour and euphoria. The fact that these symptoms load together across studies, using different samples, and a variety of rating scales, increases the probability that robust constructs have been identified and that behavioural components reflect an underlying pathophysiological process.

A number of symptoms do not appear to load consistently across or within studies. For example, Aalten and colleagues (2003) found that the assignment of the symptoms aberrant motor behaviour, sleep disturbances and anxiety to one particular component was questionable. The attribution of symptoms such as apathy, sleep and eating disturbances to behavioural components tends to be less reliable across studies. This may reflect the use of different samples, study designs and the relatively small numbers of patients included for factor analytical purposes. Also, studies using different scales are often difficult to compare.

However, it may be that these 'floating symptoms' are more associated with disease severity or other unmeasured variables, which differ across studies.

Eleven of the twelve symptoms loaded exclusively onto one component, with many symptoms having very high loadings, increasing the reliability of the components. In addition, the three-component solution remained stable to different analytical methods, controlling for the confounding effect of disease severity, and when analysis was restricted to just those in the later stages of dementia. Differences were observed between those in the mild to moderate stages of dementia and those in the later stages of disease development. If one were to assume an underlying pathological cause for behavioural symptoms, analysis of those in the early stages might be unreliable as a large proportion of participants who had not displayed particular symptoms could still harbour a sizable risk of going on to develop them over the course of the illness. Hence, the component structure obtained among those in the later stages of the disease is likely to be more representative of symptom development over the lifetime of the illness. It should be noted that even though differences did occur between those in different stages of dementia, certain symptoms loaded together across the different analyses. For example, hallucinations and delusions reliably loaded together, as did euphoria, disinhibition and aberrant motor behaviour. Depression/dysphoria and anxiety also had high loadings onto the same component across the different analyses.

2.4.3 Discussion of component score analysis

None of the behavioural components were associated with gender, current age, years in education, family history of dementia or number of APOE ε4 alleles. Findings in relation to APOE genotype are largely consistent with previous research. A small number of studies have reported a relationship between APOE genotype and specific behavioural symptoms in AD, such as aggression (Craig et al. 2004c), delusions (Scarmeas et al. 2002), depression (Ballard et al. 1997; Ramachandran et al. 1996) and psychosis (Ramachandran et al. 1996). However, the majority of studies have found no relationship (Craig et al. 2005b; Gabryelewicz et al. 2002; Steffens et al. 2003). Therefore, it seems unlikely that APOE is implicated in the prevalence or severity of behavioural symptoms in AD.

Likewise, previous studies have reported relationships between gender and behavioural problems (Ott et al. 1996), however, these findings were inconsistent. The results from this analysis do not support a link between gender and behavioural components of AD. In addition, years spent in full time education was not related to any of the components.

Interestingly, higher scores on the frontal lobe dysfunction component were associated with a lower AAO. Although, only a small proportion of variance in component scores was attributable to AAO, suggesting it is likely to have little predictive value in determining which patients are likely to experience the symptoms represented in the frontal lobe component. The lack of relationship between current age and component scores has been shown before (Haupt et al. 1998). However, in cross sectional studies it is difficult to draw conclusions from the analysis of age at assessment, as it is largely dependent on when patients were ascertained. Longitudinal studies which follow symptom development through the course of the illness are better placed to determine if behavioural disturbances vary as a function of age.

Frontal lobe dysfunction and psychosis component scores were both elevated in those with more severe cognitive impairment after controlling for disease duration. There is a general consensus that increased burden from behavioural symptoms is associated with lower cognitive ability, with the suggestion that such problems are also associated with an increased rate of cognitive deterioration (Ballard et al. 1995; Drevets and Rubin 1989; McShane et al. 1997; Paulsen et al. 2000). This could implicate behavioural symptoms as a marker for a more rapid and aggressive form of the disease, which is supported by studies which have shown increased neuropathological burden in those with excess behavioural symptoms (Farber et al. 2000; Tekin et al. 2001b). It would also appear that behavioural disturbances do not directly lead to worsening cognitive impairment as the increased rate of decline has been shown to precede the onset of symptoms like psychosis in AD (Paulsen et al. 2000).

2.4.4 Implications and future research

It is important to consider the validity and the possible underlying causes for components identified from this type of analysis. A number of symptoms in component one appear to be associated with frontal lobe function. Appetite disturbances, apathy, disinhibition, euphoria and aberrant motor behaviour are more common in FTD than AD, linking them with frontal lobe function (Levy et al. 1996b). In addition, a higher burden of frontal lobe neurofibrillary tangles have been reported in patients with agitation (Cummings and McPherson 2001; Tekin et al. 2001b), whilst others have also observed associations between agitation and disinhibition with frontal lobe pathology (Chen et al. 1998; Sultzer et al. 1995; Tekin et al. 2001a).

It has been hypothesised that psychotic symptoms in AD are underpinned by genetic influences (Sweet et al. 2003), indeed familial aggregation of psychosis in siblings with AD has been demonstrated (Sweet et al. 2002; Tunstall et al. 2000), with the odds of both AD and psychosis developing in siblings increasing over two-fold if the proband has psychotic symptoms. Furthermore, polymorphisms in the serotonin and dopamine receptor genes have been significantly associated with psychotic symptoms in AD (Holmes et al. 1998a; Holmes et al. 2001; Nacmias et al. 2001; Sweet et al. 1998). Recently, a genome screen was performed using AD with psychotic symptoms as the phenotype of interest (Bacanu et al. 2002). One significant and two suggestive linkage peaks were identified, adding increased evidence to the hypothesis that psychotic symptoms are genetically modified. The genetic basis of behavioural components in AD will be discussed further in subsequent chapters of this thesis.

The co-occurrence of depression/dysphoria and anxiety is a finding that supports what is often observed clinically, suggesting that these symptoms are differing manifestations of a common underlying cause. An alternate hypothesis draws on a causal explanation, with one of the symptoms causing, or lowering the threshold for the expression of the other (Gorwood 2004). Middeldorp and colleagues (2005) reviewed twenty-three twin studies and twelve family studies which assessed the co-morbidity of anxiety disorders and depression. They concluded that twin studies

provide evidence that the co-existence of anxiety disorders and depression is explained by a shared genetic vulnerability for both disorders. They also report that family studies show some support for this hypothesis. Such a scenario is referred to as *pleiotropism*, where a common gene is implicated in the vulnerability to distinct syndromes, or symptoms. Given that these symptoms also appear to make up a behavioural component in AD, genetic analyses may be facilitated by considering anxiety and depression in tandem, rather than treating them as distinct symptoms in hypothesis testing. However, it should be noted that Middeldorp and colleagues also found evidence from family studies which is consistent with the alternate hypothesis that anxiety is an epiphenomenon of depression (or vice versa).

These findings have implications for future research aiming to delineate factors associated with behavioural disturbances in AD. The reduction of symptoms into a smaller number of underlying components is likely to facilitate the identification of variables associated with behavioural problems (Frisoni et al. 1999). This thesis is primarily concerned with investigating the genetic aetiology of clinical differences observed between AD sufferers. Subsequent analyses will be guided by the three behavioural components identified in this chapter. This has a number of advantages, for example it reduces methodological problems arising from the number of symptoms which can co-occur among AD patients. Also, the identification of components reduces the dimensionality of the data, decreasing the number of statistical tests required in subsequent analyses. Before seeking to identify loci implicated in the development of behavioural components it is important to determine if they are likely to be genetically influenced. In chapter 3, results from the principal components analysis will be used to guide the classification of behavioural symptoms in a family based sample, to determine if such symptoms aggregate within families. Symptoms aggregating within families are more likely to be genetically influenced, and will be subjected to linkage analysis in chapter 4 to search for loci implicated in their aetiology.

2.4.5 Methodological critique

There are a number of methodological limitations in this area of research. First, this study uses a cross sectional design, hence the findings relate to the frequency

and severity of the symptoms but do not take into account their stability over time. This study also relies on information obtained from an informant or caregiver, which could lead to bias if they lack or distort information. However, the NPI is a widely used scale with proven validity and reliability (Cummings and McPherson 2001) and the informant interview is generally acknowledged to be the best way of assessing behavioural symptoms in AD. Studies of this type aim to identify behavioural components, however they are constrained to the variables that are measured and included in the analysis. This could potentially lead to misrepresentation of behavioural components as important variables may be omitted. However, the NPI covers a broad spectrum of behavioural symptoms common in AD and evidence suggests that cognitive impairment and behavioural aspects represent independent domains (Spalletta et al. 2004). It is also possible that the identification of frontal lobe and psychosis components could stem from the inclusion of cases with frontotemporal and lewy body type dementia in this sample. However, given that such cases are generally believed to make up a relatively small proportion of dementia sufferers (Neary et al. 2005; Zaccai et al. 2005) this is unlikely to have substantially affected the results.

2.4.6 Conclusions

Moran and colleagues (2004) pointed out the umbrella term of 'behavioural and psychological syndromes of dementia' (BPSD) is too broad a target for treatment and research. Behavioural symptoms rarely occur in isolation, therefore targeting individual symptoms is probably too narrow an approach to take. Hence, the identification of components is vitally important both in terms of AD research and treatment. This is the most powerful study presented to date which has assessed behavioural disturbances in AD. Evidence for three components of behavioural symptoms was provided, representing frontal lobe dysfunction, psychosis and mood. These components may be subject to genetic variation. Before seeking such loci it is necessary to determine whether clinical heterogeneity in AD is likely to be genetically underpinned. In chapter 3 the familial aggregation of clinical variables will be assessed, guided by the components identified in this chapter.

Chapter 3

Familial influences on clinical variation in late-onset Alzheimer's disease

3.1 Introduction

3.1.1 Heritability and familiality of phenotypic variation

As identified in chapter 2, AD is a clinically heterogeneous disorder, in which marked variation in AAO and variable presentation of behavioural symptoms are typical. Numerous authors have hypothesised that the clinical variation observed in AD (Sweet et al. 2003; Tunstall et al. 2000) and other psychiatric disorders (Fisfalen et al. 2005; Korszun et al. 2004; Wickham et al. 2002) may act as a marker for disease sub-phenotypes associated with increased genetic liability. Similarly, it has been suggested that the clinical presentation of AD may be modified by genetic variation that does not confer increased vulnerability to the disease as a whole.

Before proceeding to investigate either hypothesis it is necessary to understand what is implied by the term '*phenotype*'. The Oxford English dictionary defines the term as 'the observable characteristics of an individual resulting from the interaction of its genotype with the environment'. As such the major task underlying psychiatric genetics is to identify the behavioral phenotype that fluctuates according to variations in one, or several, genes (Tsuang et al. 1993). To date, genetic studies of psychiatric disorders have largely relied on disease phenotypes categorized in accordance with standard diagnostic criteria and structured psychiatric interviews (Merikangas et al. 1989). Disease phenotypes defined in accordance with standard diagnostic criteria have been shown to be heritable in twin studies of numerous diseases, including Alzheimer's disease (see section 3.1.2), schizophrenia (Cannon et al. 1998; Cardno et al. 1999) and bipolar disorder (Smoller and Finn 2003). However, limited success has been achieved in mapping genes for familial, non-Mendelian, diseases based on these phenotypes (Risch 2000). Some have argued that although diagnostic criteria are often found

to be highly reliable, they have little biological validity. For example, Kendler (1990) proposed that making reliability a priority in the evaluation of psychiatric disorders might reduce validity. As such Leboyer and colleagues (1998) questioned whether modern definitions of clinical syndromes, which are currently considered as phenotypes, accurately reflect underlying genetic variation. They coined the term 'candidate symptom approach' as an analogy to the biologically used term 'candidate gene'. The major goal for psychiatric geneticists following the 'candidate symptom approach' is to identify, and accurately characterise, clinical (or 'phenotypic') characteristics associated with genetic variation.

Tsuang and colleagues (1993) proposed a number of specific guidelines for identifying phenotypic indicators. In summary, they argued that phenotypic variation should show some degree of stability, be specific to the disease of interest and have biological and clinical relevance in order to be useful for further genetic analysis. Tsuang and colleagues (1993), along with others (Lander and Schork 1994), have also stipulated that phenotypic variation must show clear familial transmission if it is to be considered useful for further genetic analyses (e.g. increased heritability, co-segregation or familial aggregation).

Twin and adoption studies are the gold standard in determining the extent to which phenotypic variation is genetically influenced (Cardno and McGuffin 2002). Twin studies require that the phenotype of interest is measured appropriately in monozygotic (MZ) and dizygotic (DZ) twins. Subsequent analyses can be performed to differentiate between effects resulting from individual specific environment, shared environment and genetic variation. Genetic effects are estimated by comparing the similarity, or concordance rates, of MZ twins, who inherit identical genes from their parents, with DZ twins, who share only 50% of their genes in common. If concordance rates are elevated in MZ twins it emphasises the importance of genetic effects. MZ and DZ are considered to be influenced equally by environmental conditions. As such genetic, individual environmental and shared environment effects can be estimated. However there is some evidence that the equal environment assumption may be invalid. For example, MZ twins socialise together more and others emphasise their similarities more frequently than is the case among DZ twins (Kendler and Gardner 1998).

Furthermore, MZ twins are also exposed to increased pre-birth shared environment with around 65% being monochorionic, which never occurs among dizygotic twin pairs. This could affect growth rates and also increases the possibility of shared *in utero* viral infections (Cardno and McGuffin 2002). Adoption studies compare the risk of disease in adopted children of affected and unaffected biological parents. Alternatively, the disease risk is assessed in biological and adoptive relatives of affected adoptees. A less frequently used approach compares the risk of a disorder in adoptees who have affected biological parents and unaffected adoptive parents, with the risk among adoptees with unaffected biological parents and affected adoptive parents.

Familiality is concerned with the extent to which phenotypic variation clusters within families, for example do individuals within certain families present with clinical characteristics which are more alike than one would expect by chance. Familiality is often easier to assess as it is not dependent on the ascertainment and careful assessment of large twin based samples. It is important to remember that familiality represents the proportion of phenotypic variance accounted for by genetic *and* shared environmental effects, and is not directly comparable to genetic *heritability* (the proportion of phenotypic variance accounted for by genes alone). Tunstall and colleagues (2000) argued that sibling pairs are exposed to very little shared environment in later life, therefore familial effects are likely to reflect underlying heritability in studies of late life diseases, like AD. However, the role of early shared environment, for example influences from shared domestic experiences, school, peers and shared life events, cannot be assessed using familiality studies. Estimates of shared environment in twin studies assessing the heritability of AD have generally been low (Gatz et al. 1997; Pedersen et al. 2004; Pedersen et al. 2001; Raiha et al. 1996). However, two studies have reported considerable shared environment effects (Bergem et al. 1997; Breitner et al. 1995). This illustrates how common experiences among relatives can still play a role in phenotypic presentation even in later life, therefore interpretation of familiality studies should be made with caution. However, the majority of findings support the hypothesis that shared environment effects are limited in later life; hence familiality studies offer an appropriate means of identifying aspects of clinical variation in AD that are likely to be genetically influenced. On the whole,

sibling pairs are easier to ascertain than twins. In addition, sibling pairs can be used for linkage analysis. This offers a distinct advantage over twin samples, allowing the direct assessment of the genetic epidemiology of clinical variation showing familiality. Monozygotic pairs are not informative for genetic linkage analysis, whilst samples of affected dizygotic twin pairs alone are rarely used as sufficient sample sizes are difficult to ascertain.

3.1.2 Twin studies of late-onset Alzheimer's disease

A number of twin studies have been performed to assess the heritability of late-onset Alzheimer's disease, of which the majority have been of Scandinavian origin. Several early reports document monozygotic twin pairs who have remained discordant for over 15 years. However, these reports are of relatively little value as it is now widely acknowledged that LOAD arises from a combination of environmental and genetic risk factors, which are unlikely to demonstrate full penetrance. As such the goal of twin studies has generally been in determining the *extent* to which genetic heritability is implicated in AD.

In an early twin study of AD, Nee and colleagues (1987) did not report any evidence for heritability. They used a sample of 22 twin pairs and reported similar concordance rates among MZ and DZ twins. However, their study relied on volunteer twin pairs which is likely to have biased their results, making interpretation more difficult. This is perhaps illustrated in the over representation of MZ twins in their sample, with only 5 DZ pairs included. In addition, the average age of assessment was 70, which increases the likelihood that new cases of AD would develop in the years following the study.

A number of studies have reported population based approaches to twin studies. For example, Gatz and colleagues (1997) reported data from a population based sample of Swedish twins ascertained through an existing registry. They used twins who were reared apart, and a similar number of pairs who were reared together. A total of 781 twin *pairs* aged over 53 were considered in their study, along with an additional sample of 416 in which the co-twin was deceased. Just fewer than 1000 individuals completed screening and subsequent clinical assessments, among

which 75 were considered to be suffering with dementia. The relatively low prevalence of dementia probably represents the number of individuals under the age of 60 years included in their analyses; in fact the prevalence rate was much higher in a subset of their sample over 65 years of age. The concordance rate for MZ twins for AD was 67% compared to 22% among DZ twins, resulting in a heritability estimate of between 75% and 85%. The sample used was relatively small with only 10 MZ pairs and 30 DZ pairs in which one member had probable AD. In addition, the sample of twins was relatively young and the onset of new dementia in years succeeding the study was not taken into account in the analysis method used.

Gatz and colleagues (2005b) have recently extended this work, reporting preliminary analyses from the HARMONY study. They incorporated data from all twin pairs on the Swedish twin registry who were aged 65 years and over at baseline in 1998. As such it is more representative of the population as a whole and less effected by biases introduced by soliciting volunteers. In total the HARMONY cohort included 20,269 twins, of which 14,498 performed at least a basic telephone screen for cognitive decline. Consequently, a total of 4,537 twin *pairs* were assessed, with only one twin from each pair being assessed on 5,424 occasions. The main reason for only one member of a pair participating was death of the co-twin before screening commenced. Of those screened by telephone 2,139 twins were referred for clinical assessment, primarily because either they, or their twin, were considered to be cognitively impaired during telephone screening. The prevalence of dementia in the study population was estimated to be 7.7%, which is comparable with other studies. As such 187 and 288 demented monozygotic and dizygotic twins were ascertained, respectively. The concordance rate for AD among MZ twins was 59%, compared to 32% and 24% among like and unlike sex DZ twins respectively. It is interesting that the evidence for a genetic loading on AD remains stable in a very large population based study. However, it should be noted that more MZ twins were successfully ascertained as both members of DZ pairs were less likely to be living at the time of the study, which could have potentially biased these findings.

Breitner and colleagues (1995) reported a twin study conducted in a sample of 9,213 male twins ascertained in the USA through a registry of war veterans. Of these 2,888 (31%) were either deceased, out of the country or untraceable. A total of 1,589 twins were thought to show signs of cognitive decline after telephone screening and were followed up in person. Those showing signs of other dementias (e.g. arising from multiple strokes or alcoholism) were excluded. As such, 90 twins were identified with suspected dementia, of which 38 met criteria for AD after assessment. The prevalence of AD in this sample was very low, estimated to be between 0.4% and 0.7%, which reflects the low age range of the sample ascertained (ranging from 62-73 years of age at baseline screening). The concordance rate for AD among MZ twins was 21% compared to 11% among DZ twins. These findings suggest a genetic component to AD, but the low prevalence of dementia, low age range of twins and lack of concordant pairs make interpretation difficult. Furthermore, the study was restricted to males, therefore the results cannot be generalised to the population as whole.

Raiha and colleagues (1996) reported another population based study using Finnish Twins. They studied birth records of all twins born before 1958 and still alive in 1967, identifying a total of 13,888 pairs. This twin registry was linked with hospital discharge records to identify those with dementia or related disease, yielding a total of 178 twins with confirmed dementia, primarily of the Alzheimer's type. The pairwise concordance among MZ twins was 31% compared to 9% among DZ pairs, however this difference was not statistically significant. Wide confidence intervals were reported for the MZ and DZ concordance rates which could reflect a lack of power in this study. Also, it is likely that dementia sufferers in the Finnish population were under ascertained due to the inclusion of only those cases with a hospital discharge record. As such those with moderate to severe dementia, requiring hospitalisation, were more likely to be included in analyses. However, with adequate social support AD sufferers can remain in their home environment for many years without the need for medical treatment. This is probably reflected in the low incidence of dementia reported in this study. However, it is unlikely that these selection biases operated according to zygosity; therefore they are unlikely to influence the conclusions drawn. In addition, the rate of AD was higher among MZ twins than in DZ pairs which could reflect

ascertainment bias. Given these limitations, the use of medical records is particularly useful as it incorporates a long follow up time of twin pairs.

Bergem and colleagues (1997) matched data on 23,000 cognitively impaired individuals with 26,000 individuals on the Norwegian Twin Register. They identified 151 'index cases', of which 72 were suitable for further study (i.e. both twins were living and willing to participate). They assessed each twin pair using standardised diagnostic criteria and identified 29 concordant pairs in which both had dementia. They reported a concordance rate of 57% among MZ twins compared to 33% among DZ pairs. When restricting analyses to AD, the concordance rates were 87% and 46% among MZ and DZ twins, respectively, corresponding to heritability estimate of between 55% and 61%. Furthermore, no evidence was found for a genetic component in the aetiology of vascular dementia. This supports the view that genetic variation plays a major role in AD susceptibility; whereas, vascular dementia is more likely to result from exposure to environmental risk factors.

It has been reported that concordant MZ twins are more likely to have a positive family history of dementia than discordant pairs (Rapoport et al. 1991). This is consistent with the hypothesis that concordant twins are genetically influenced. Steffens and colleagues (2000) have extended this research to determine if the increase risk to family members of concordant twins could be attributed to Apolipoprotein genotype (APOE) (discussed in section 1.3). They incorporated data from 15 concordant and 79 discordant twin pairs, reporting a positive family history in 21% and 9.5% of concordant and discordant twins respectively. This finding suggests a stronger genetic mode of disease transmission among concordant twins. In addition the effect is likely to be underestimated in these analyses, as relatives of concordant pairs were younger than those of discordant twins, as such the onset of new dementia among family members over the coming years is likely to be more marked for the concordant group. Tentative evidence was provided which suggests that the increased risk to family members of concordant twins is mediated, but not wholly explained, by APOE genotype. However, this analysis was not sufficiently powered to draw firm conclusions, as only one concordant twin pair did not possess an APOE ε4 allele

In addition to risk for AD as a whole, Barak and colleagues (2003) have reported a twin study which aimed to assess the heritability of cognitive impairment in late life. They ascertained and performed a basic telephone screen of cognitive function with 32 MZ and 18 DZ twin pairs aged between 65 and 86 years. They reported no differences in concordance rates for cognitive impairment, however there study was hindered by potentially imprecise measures of zygosity and the use of a rather crude measure of cognitive function.

Of particular concern for twin studies is the failure to detect all cases of dementia, which in general would require all members of a population to be assessed rigorously and longitudinally. This is beyond the means of most studies. However, prevalence rates in twin studies have been largely comparable to those reported in population based surveys. Furthermore, the majority of studies find DZ concordance rates which are comparable to those reported among sibling pairs who also share half of their genes in common (van Duijn et al. 1991). In addition, it is unlikely that undetected and detected cases would differ in a way which would affect concordance. A similar problem derives from the late-onset nature of AD. Many genetically predisposed partners of affected twins may die of other causes before developing dementia. However, incomplete ascertainment of all affected twins is unlikely to differ according to zygosity, therefore it does not influence the conclusions which can be drawn from twin studies. Further methodological issues arise from the use of small sample sizes and the calculation of shared environment affects. Even in large studies, due to the relatively low prevalence of AD, sample sizes tend to be low. As a result wide confidence intervals are generally reported and the statistical power to identify group differences is limited. Also, as AD is age related, shared environment affects reported in twin analyses are often over estimated as twins share the same age. Despite these limitations, the evidence from twin studies is remarkably consistent. Perhaps the best evidence comes from the large population based studies of Scandinavian origin (Bergem et al. 1997; Gatz et al. 1997; Raiha et al. 1996), which generally support a heritability estimate of between 60% and 80%.

3.1.3 Family studies of late-onset Alzheimer's disease

In addition to twin research, a number of studies have attempted to elucidate the proportion of AD risk attributable to genetic factors by assessing the familial aggregation of the disease. For example, Sleegers and colleagues (2004) investigated the epidemiology of dementia in a genetically isolated Dutch population in which they ascertained 122 patients with probable AD. They reported that individuals with AD were more closely related than healthy individuals. Consistent with what is already known, they reported that clustering was strongest among those with early onset forms of the disease. However, they found that over 60% of late-onset AD sufferers had a family history of dementia, with the pattern of transmission being comparable with an autosomal dominant disease in 14% of late-onset AD cases. Others have reported family history of dementia to be a risk factor for late-onset AD. For example Van Duijn and colleagues (1991) reported that the relative risk of AD to first degree relatives of dementia sufferers was 3.5. The relative risk to first degree relatives reduced with increasing age at onset (AAO), however among those with AAO greater than 80 years there were still more participants with a family history of dementia compared to controls. This finding has been replicated elsewhere. Fratiglioni and colleagues (1993) assessed the risk of late-onset AD in relation to family history of dementia in a sample of 98 cases and 216 healthy individuals. They reported that the presence of at least one first degree relative with dementia was associated with a three-fold increase in the risk of developing AD. Similarly, the Canadian Study of Health and Aging found those with a family history of dementia to have a two to three-fold increase in risk for AD (Canadian Study of Health and Aging 1994).

Given that the APOE ε4 allele is widely believed to increase risk of developing LOAD in a dose dependent manner, it is reasonable to speculate that this could explain findings of heritability and familial clustering in AD. Jarvik and colleagues (1996) assessed the relationship between family history of dementia and AD, accounting for APOE genotype. Family history was assessed by asking informants of AD sufferers and non-demented controls about the prevalence of memory problems among their family members. They found that family history remained a

significant predictor of AD status, regardless of APOE genotype (apart from ε4 homozygotes, of which almost all were affected). Payami and colleagues (1997) report similar findings, they performed longitudinal assessments over a total of 4 years, with a sample of 114 healthy, cognitively intact, individuals aged 75 at baseline. Those with an APOE ε4 allele and family history of dementia were found to have a nine-fold increase in disease risk compared to those without a family history. In a larger study, Martinez and colleagues (1998) collected family data for 290 subjects with probable AD, performing clinical examinations with living relatives where possible. They reported that the APOE ε4 allele increased lifetime risk of developing AD in a dose dependent manner, however, familial clustering of AD was largely due to factors other than APOE genotype. These findings are consistent with the conclusion that genetic variation which increases the risk of developing AD remains to be found.

Family based, non-twin, methods have a number of disadvantages. First, they cannot differentiate between shared environment and genetic effects, therefore increased familial risk could be attributable to environmental conditions common to members of the same family. Also, family history effects may be exaggerated due to healthy individuals being less likely to know about memory difficulties among other family members compared with informants of AD sufferers. However, they do have the particular advantage that a greater proportion of the population can be ascertained, providing more powerful samples for hypothesis testing. The evidence to date is largely consistent, suggesting a combination of environmental and genetic risk factors increase susceptibility to late-onset AD.

3.1.4 Heritability and familiality of age at onset in Alzheimer's disease

The familiality of clinical variation in AD has received very little attention. AAO represents perhaps the most studied clinical variation in relation to genetic influence. The effect that genetic variation can have on AAO is demonstrated clearly in the early onset forms of AD. Those harbouring presenilin 1, presenilin 2 and amyloid precursor protein mutations show marked fluctuations in age at disease onset (Mullan et al. 1993), which corresponds largely to the specific mutation present in each family. It has been suggested that APOE ε4 allele is

associated with when, rather than if, late-onset AD develops (Meyer et al. 1998), with AAO reducing in a dose-dependent manner with increasing ε4 alleles (Corder et al. 1993). However, a number of studies have reported that APOE does not account for all the genetic variation in AAO (Duara et al. 1996; Jarvik et al. 1996; Tunstall et al. 2000). Warwick Daw and Colleagues (2000) performed segregation analysis of late-onset AD families and estimated that between four and seven unidentified genes modify AAO , of which four were estimated to have an effect size greater than, or equal to, that of APOE. Furthermore, Tunstall and colleagues assessed the relationship between ages of onset among 106 affected sibling pairs. They reported a moderate relationship between ages at onset among siblings, which was not wholly accounted for by APOE genotype. It should be noted that as case finding in this sample relied on the identification of two or more affected *living* relatives, it generally excluded data from siblings who were either deceased or below the age of risk for developing dementia. As such the correlation between siblings may have been inflated. Li and colleagues (2002), using a much larger sample of 1121 individuals with AD from 449 families, estimated the 'heritability' of AAO in AD to be 42%, further supporting the notion that AAO is modified by genetic factors.

A number of studies have analysed familial risk to siblings, conditioning on the probands AAO (Silverman et al. 2005; Silverman et al. 2003; Wu et al. 1998b). They reported that the genetic risk to siblings decreases with increasing onset age in the proband. This supports the view that non-genetic, environmental influences have more of a role in very late-onset AD. As such, genetic studies of sporadic AD may be better served by analysis of those with an earlier disease onset. However, others have reported that genetic heritability does not differ significantly with age (Pedersen et al. 2004). As noted in section 3.1.2, Pederson and colleagues assessed the heritability of AD in a sample of Swedish twins, estimating that 48% of the liability to AD could be attributed to genetic variation. This estimate did not differ significantly between twins younger than 80 years of age and those aged 80 and over at baseline. It should be noted that the actual heritability estimate among those aged younger than 80 years was 59%, compared to 40% in those aged over 80. However, the sample used was relatively small. Consequently, large confidence intervals were reported for each heritability estimate. It is therefore

unsurprising that no statistically significant differences were reported between the groups. It is interesting to note that increased heritability and familiality has been associated with lower age at disease onset in studies of schizophrenia (Sham et al. 1994), bipolar disorder (Strober et al. 1988), major depressive disorder (Weissman et al. 1984) and obsessive compulsive disorder (Pauls et al. 1995). This suggests that the genetic loading across a number of diseases is greater in cases with lower ages of onset, with later ages of onset being more influenced by other, environmental factors.

3.1.5 Heritability and familiality of behavioural features in Alzheimer's disease

In addition to AAO there is evidence that other aspects of the clinical phenotype in AD may be genetically modified. In 2000 Tunstall and colleagues reported the first study to directly assess the familial aggregation of behavioural symptoms observed in AD. They reported a significant correlation in current mood state, as measured by the Cornell scale for depression in dementia (Alexopoulos et al. 1988a), in a sample of 86 sibling pairs with probable AD. They also reported excess pair wise concordance among siblings for symptoms of agitation. Symptoms of aggression and psychosis showed very slight increases in pairwise concordance among siblings, compared to what would be expected by chance. However, no formal significance testing of these increases was reported. These findings should be viewed with caution as the compounding effect of disease severity among siblings was not controlled for. This is particularly important in cross sectional studies of behavioural symptoms. As identified in chapter 2, most behavioural symptoms become more prevalent and severe in the later stages of disease development, therefore failure to account for severity of dementia could influence the findings from such analyses. The relationship between disease severity and behavioural symptoms could also induce another bias. Given that AAO has been shown to be correlated among sibling pairs (Tunstall et al. 2000), siblings of a similar age at assessment are likely to have suffered with dementia for around the same number of years. This increases the possibility that siblings of a similar age will be at a comparable stage of dementia at assessment. As behavioural symptoms vary as a function of disease severity this could falsely

inflate correlations among pairs. Studies of this type, which assess symptom development cross sectionally should control for the confounding effect of disease severity.

Familial effects on depression are of particular interest as a number of studies have reported that individuals with a family history of an affective disorder are more likely to experience a depressive episode in the course of AD. Pearlson and colleagues (1990) assessed a sample of 122 AD sufferers. They found that those experiencing their first episode of major depression had significantly more first and second degree relatives with mood disorders than were reported among non-depressed AD sufferers. This finding has been replicated elsewhere in larger samples, using differing methodology (Lyketsos et al. 1996; Strauss and Ogrocki 1996). In these studies AD sufferers were generally excluded if their first episode of depression occurred before the onset of dementia. Butt and Strauss (2001) simultaneously analysed personal (e.g. previous lifetime history of depression before the onset of AD) and family history of depression in AD. They reported that both a personal lifetime history and a family history were risk factors for depression in AD, and that the increase in risk attributable to personal history of depression may be, in part, explained by familial influences. In sum, Butt and colleagues found that those with a family history of depression had a three-and-a-half-fold increase in risk of developing depressive symptoms, compared to AD sufferers without a family history of depression. Given these findings it is likely that those with a family history of mood disorders carry a previously unexpressed genetic risk for depression, which manifests in the presence of neurodegeneration observed in AD.

Familial aggregation of psychosis in AD has also been demonstrated (Bacanu et al. 2005; Sweet et al. 2002). Sweet and colleagues studied the relationship between psychotic symptoms in a sample comprising 461 affected siblings of 371 probands with possible, probable or definite AD. They generated two classifications of psychosis. A broad definition required that patients had demonstrated either delusions or hallucinations at any time point during their illness, whereas a more restrictive classification required the presence of more than one psychotic symptom, or the presence of psychotic symptoms during more than one assessment. Using the broad definition, the odds of both AD and

psychosis (AD+P) developing in siblings increased over two-fold if the proband had psychotic symptoms (Sweet et al. 2002). This effect remained stable after controlling for sibling age, AAO and disease severity. Furthermore, the effect appeared more marked in a separate analysis using the more restrictive definition of psychosis, with siblings of those with AD+P exhibiting over a three-fold increase in risk for developing psychotic symptoms, compared to if the proband was psychosis free. Sweet and colleagues have since followed up these results by categorising individuals as either non psychotic, having single or multiple psychotic symptoms (Bacanu et al. 2005). They estimated that the heritability for AD characterised by multiple psychotic symptoms was 61%, compared to 30% for AD+P defined by the occurrence of any psychotic symptoms. However, these estimates of heritability represent the upper limit as they assume that shared environment among siblings does not influence susceptibility to AD as a whole, or psychotic symptoms occurring during the illness. However, given these limitations, the heritability estimates of AD+P compare favourably with heritability estimates of both AD (Gatz et al. 1997) and schizophrenia (Cannon et al. 1998; Cardno et al. 1999). These findings indicate two plausible scenarios. First, separate genetic vulnerabilities for AD and psychosis may exist. In such a case the genetic vulnerability to psychosis may manifest as a result of the neurodegeneration associated with AD. Alternatively, there may be a particular vulnerability to a refined phenotype of AD with psychosis.

Together with the earlier report by Tunstall and colleagues (2000) these remain the only studies to have investigated the familial aggregation of psychotic symptoms observed in AD. Independent replication in well characterised samples is necessary before firm conclusions can be drawn. Reviewing the broader literature it is interesting to note that familial influences on psychotic symptoms have also been reported in a number of other psychiatric disorders. For example, family history of psychosis has been identified as a risk factor for psychotic symptoms in Huntington's disease (Tsuang et al. 2000), major depression (Leckman et al. 1984) and bipolar disorder (Potash et al. 2003; Potash et al. 2001). It could be that particular genetic variation which contributes major risk of developing psychotic disorders such as schizophrenia could be associated with psychosis in AD. Indeed, polymorphisms in dopamine and serotonin receptor

genes, which have been implicated in the aetiology of schizophrenia, have been found to be associated with psychosis in AD (Holmes et al. 1998a; Nacmias et al. 2001; Sweet et al. 1998).

3.1.6 Study design and aims

The vast majority of studies find that late-onset AD is a heritable disorder. However, there is substantial clinical heterogeneity between disease sufferers, perhaps most notably in the age at disease onset and the development of behavioural symptoms. Evidence from the literature reviews presented in section 3.1.4 and 3.1.5 suggest that the clinical variation observed in AD is subject to genetic influence. However, there is a scarcity of studies which have looked at the familiality of clinical variation in AD using large, well characterised, samples. There are only three studies which have assessed the familiality of psychosis in AD (Bacanu et al. 2005; Sweet et al. 2002; Tunstall et al. 2000), of which two are largely overlapping. Whereas the analysis of familial aggregation of depression and agitation in AD has only been presented once (Tunstall et al. 2000). Further study is required to delineate aspects of clinical variation observed in AD which may be subject to genetic influence.

As identified in chapter 2 certain symptoms occur together more frequently than would be expected by chance. This suggests that certain behaviours may result from common underlying neurobiological determinants. Familial aggregation of behavioural disturbances in AD would suggest that this underlying neuropathology is genetically influenced, which manifests as an increased susceptibility to certain symptoms. If a number of symptoms are associated with the same underlying cause, studying the familiality of behavioural *components* may provide more consistent results. In chapter 2, three components were identified which represented the broad spectrum of behavioural disturbances commonly, but not inevitably, observed among AD sufferers. Specifically these were 'frontal lobe dysfunction', 'psychosis' and 'mood'. These behavioural components are largely comparable with the three symptoms for which familiality has already been suggested (e.g. psychosis, depression and agitation).

In this chapter, these behavioural components will be used to guide the classification of behavioural symptoms in a large sample of affected sibling pairs. For example, in the PCA presented in chapter 2, hallucinations and delusions loaded together, as did depression/dysphoria and anxiety. This supports the analysis of a broader syndrome of psychosis, rather than treating hallucinations and delusions as distinct entities. Likewise, symptoms of anxiety may be a differing manifestation of the underlying pathology associated with the development of depression in AD. As such, analyses of a combined depression and anxiety component may be more appropriate.

Family based samples offer an appropriate means of determining whether aspects of phenotypic variation are likely to be genetically influenced. Three such samples were available, comprising sibling pairs collected by the National Institute of Mental Health AD Genetic Consortium ('NIMH Sample'), the Indiana Alzheimer Disease Centre ('NIA Sample') and a sample collected in the United Kingdom ('UK Sample'). The NIMH sample is one of the largest publically available family based samples in the world which is accessible to those studying the genetics of AD. Both the NIMH and UK families are well characterised, with standardised assessment batteries performed with most individuals. As such, detailed information regarding age at disease onset and the presence of behavioural symptoms are available for those in these families. The NIMH and UK samples have previously been used to investigate the familial clustering of clinical heterogeneity in AD. Tunstall and colleagues (2000) assessed the familial influence on AAO, mood disturbances, agitation, psychosis and aggression in the UK sample. However, they did not report significance levels for their findings in relation to behavioural symptoms. Furthermore, they did not control for the confounding effect of disease severity. Sweet and colleagues (2002) incorporated data from over 90% of the NIMH sample used in this chapter, however there study was restricted to psychotic symptoms. The NIMH sample has not previously been used to investigate the familial aggregation of age at disease onset, mood disturbances and aggression in AD. The NIA families were collected in a clinical setting primarily for genetic studies; as such data regarding behavioural symptoms is currently unavailable. However, data regarding AAO, suitable for investigating familiality, were available.

This chapter presents the largest study to date which assesses the familiality of clinical variation in AD, focusing on the relationship between aggression, psychosis and mood disturbances among affected sibling pairs. In addition, the familiality of age at disease onset will be analysed. Previous findings and those reported in chapter 2 of this thesis suggest that behavioural symptoms may vary according to disease severity, AAO and gender. As such these variables will be controlled for when assessing the familiality of aggression, psychosis and mood disturbances. The APOE ε4 allele is widely believed to reduce age at disease onset in a dose dependent manner. To determine whether further genes are likely to exist which influence AAO supplementary analysis will be performed removing the effect of APOE.

To summarise, the main aims of this chapter are:

1) To investigate the familiality of AAO in AD among affected sibling pairs.
2) To perform supplementary analyses investigating the familiality of AAO controlling for the effect of the APOE ε4 allele.
3) To investigate the familiality of aggression, psychosis and mood disturbances in AD among affected sibling pairs, controlling for disease severity, AAO, current age and gender.

3.2 Methodology

3.2.1 Sample description

Families were selected from those collected by the National Institute of Mental Health AD Genetic Consortium (referred to in this thesis as the 'NIMH Sample'), the Indiana Alzheimer Disease Centre (referred to in this thesis as the 'NIA Sample') and a sample collected in the United Kingdom (referred to in this thesis as the 'UK Sample'). A total of 452 families were selected using the following criteria: 2 or more siblings diagnosed with probable or definite AD according to NINCDS-ADRDA diagnostic criteria (McKhann et al. 1984) with onset ages greater than or equal to 65 years. To reduce potential genetic heterogeneity caused by ethnic origin only Caucasian families were selected.

A total of 287 families were selected from the NIMH sample, comprising 240 affected sibling pairs, 37 affected sibling trios and 10 affected siblings quartets. The NIMH Sample was ascertained in the United States by three collaborating centres, comprising: the University of Alabama, Birmingham; Harvard Medical School, Massachusetts and John Hopkins University, Baltimore. Participants were ascertained through systematic screening of patients in clinical settings and also via media advertising and local AD association referrals. Data collection began in 1991 and continued until 1997. Participants in the NIMH sample were assessed using a detailed assessment battery which included the following scales: Bristol Activities of Daily Living Scale (Bucks et al. 1996), Mini-mental State Examination (Folstein et al. 1975), Brief Psychiatric Rating Scale (Overall and Gorman 1962), Hachinski Ischaemic Scoring System (Hachinski et al. 1975) and the CERAD Neuropsychological Battery (Morris et al. 1989).

A total of 101 families were selected from the UK sample, comprising 89 affected sibling pairs, 10 affected sibling trios and 2 families with four affected siblings. Siblings were ascertained through contact with clinical services. All participants were interviewed using an assessment battery which has been validated against post-mortem diagnosis and shows a positive predictive value of over 90% for detecting AD pathology (Holmes et al. 1999). Assessment scales included: Blessed Dementia Scale (Blessed et al. 1968), Bristol Activities of Daily Living

Scale (Bucks et al. 1996), Manchester and Oxford Scale for Psychological Assessment of Dementia (MOUSEPAD) (Allen et al. 1996), the Behavioural Pathology in AD Rating Scale (BEHAVE-AD) (Reisberg et al. 1987), Cornell Scale for Depression in Dementia (Alexopoulos et al. 1988a), Mini-mental State Examination (Folstein et al. 1975), Webster Rating Scale (Webster 1968), Cambridge Mental Disorders of the Elderly Examination (CAMDEX; informant interview, physical examination sections and CAMCOG) (Roth et al. 1986) and the Global Deterioration Scale (Reisberg et al. 1982).

Sixty four families were selected from the NIA sample, comprising 56 affected sibling pairs and 8 affected sibling trios. All participants were ascertained in the United States by the Indiana Alzheimer Disease Centre. In the NIMH and UK samples diagnosis was based on semi structured interview, while in the NIA sample affection status was based on clinical diagnoses. The Clinical Dementia Rating Scale and the Global Deterioration Scale were used to stage disease progression in the NIMH and UK samples, respectively.

3.2.2 Definitions of clinical features of Alzheimer's disease

AAO was defined as the age which first symptoms of Alzheimer's disease were observed and was available for subjects from the UK, NIMH and NIA samples. It was considered that the method of family selection outlined above would inflate the correlation between siblings, as they all had to have an AAO greater than 65 years. To reduce this bias families were selected based on the proband having an AAO $\geq$ 65. Data for all available siblings were included in the analyses, regardless of AAO.

Clinical data were available for siblings from the UK and NIMH samples which allowed the categorisation of psychosis, mood and aggression to be made. Psychotic symptoms in subjects from the NIMH sample were assessed using the Brief Psychiatric Rating Scale (BPRS) (Overall and Gorman 1962), which assesses behavioural symptoms present at any time during an individual's illness. Individuals were scored positive for psychosis if they scored highly on items relating to grandiosity, suspiciousness, unusual thought content or hallucinations.

Those in the UK sample were assessed using the Manchester & Oxford Universities Scale for the Psychopathological Assessment of Dementia (MOUSEPAD) (Allen et al. 1996), which measures the presence of psychiatric symptoms at any point during the course of an individual's illness. Subjects were classified as having psychosis if they demonstrated either two or more delusional symptoms or more than one type of hallucinatory behaviour. More caution was shown with regards to delusions in both samples as apparently delusional behaviour among AD sufferers is often a form of confabulation secondary to amnesia and therefore likely to have a different aetiology to other psychotic features.

Three variables were created to assess the familiality of mood, representing minor depression, major depression and depression with anxiety. Major Depression in subjects from the NIMH sample was defined in accordance with DSM-IV (American Psychiatric Association 1994) criteria for a 'major depressive episode'. Subjects were defined as meeting criteria for major depression if they had been depressed for a period of two weeks and had displayed five or more of the following symptoms:

- Depressed mood most of the day.
- Diminished interest or pleasures in usual activities.
- Fluctuations in weight.
- Insomnia or hypersomnia.
- Psychomotor agitation or retardation.
- Fatigue or loss of energy.
- Feelings of worthlessness or excessive inappropriate guilt.
- Diminished ability to concentrate or indecisiveness.
- Recurrent thoughts of death, suicidal ideation.

In accordance with DSM-IV criteria, these symptoms had to cause clinically significant distress or impairment in social, occupational or other areas of functioning. Those meeting these criteria were classified as 'AD with major depression' (AD+MajD). Individuals who had not displayed any symptoms of depression were classified as 'AD without depression' (AD-MajD). Those who displayed depressive symptoms that did not meet criteria for major depression

were excluded from familiality analyses of major depression (e.g. those displaying between 1 and 4 symptoms of depression). DSM-IV criteria for 'major depressive episode' can be found in appendix 3a.

The Cornell scale for depression in dementia (CSDD) (Alexopoulos et al. 1988a) was used to assess symptoms of depression in subjects from the UK sample. The CSDD is a 19-item, clinician administered depression scale that was developed specifically to measure depressive symptom severity in older adults. It has been validated in populations with and without dementia (Alexopoulos et al. 1988a, b). Each item is rated on a three point scale, with anchor points for absent (0), mild (1), and severe (2), providing a total summary score of 0 to 38, with higher values representing greater severity of depressive symptoms. A cut-off of 13 on CSDD was taken to be indicative of major depression as suggested by Alexapoulos and colleagues (Alexopoulos et al. 1988a). Those scoring less than 5 on the CSDD were classified as 'AD without depression' (AD-MajD). Those scoring between 5 and 13 were excluded from familiality analyses of major depression. The Cornell scale for depression in dementia can be found in appendix 3b.

A broader definition of depression was utilised to assess cases in which depressive symptoms were present that did not meet criteria for major depression. Participants in the NIMH sample were classified as 'AD with depression' (AD+D) if they had experienced depressed mood for a period lasting two weeks or more. Those in the UK sample were classified as AD+D if they scored 5 or above on the CSDD. All other individuals with complete clinical data were classified as 'AD without depression' (AD-D).

The principal components analysis presented in chapter 2 identified a behavioural component including depression/dysphoria and anxiety. To assess familiality, a combined depression and anxiety component was defined in the siblings. DSM-IV criteria for a 'major depressive episode' and 'generalized anxiety disorder' were used to guide the classification of depression and anxiety, respectively. Major depressive episode was classified in accordance with the criteria noted above (AD+MajD). Anxiety was defined as a period of time lasting at least two weeks

when anxiety and/or tension were present most of the time, and two or more of the following symptoms had occurred at the same time:

- Feeling anxious or tense (jittery, nervous, restless, uptight).
- Difficulty sleeping.
- Sweating, blushing, dizziness, palpitations or shortness of breath.
- Muscular tension.
- Excessive worrying.
- Fidgeting or being unable to sit still.

In accordance with DSM-IV criteria, these symptoms had to cause clinically significant distress or impairment in social, occupational or other important areas of functioning. Those meeting these criteria were classified as 'AD with anxiety' (AD+Anx). Individuals who had not displayed any symptoms of anxiety were classified as 'AD without anxiety' (AD-Anx). Those who had displayed symptoms of anxiety, but who did not meet the full criteria for AD+Anx were classified as unknown. DSM-IV criteria for 'generalised anxiety disorder' can be found in appendix 3c.

Classifications for AD with major depression and AD with anxiety were concatenated to produce a combined variable representing AD with depression and anxiety. Subjects who were classified as AD+MajD *or* AD+Anx were classified as having 'AD with depression/anxiety' (AD+DepAnx). Subjects who were classified as AD-MajD *and* AD-Anx were categorised as 'AD without depression/anxiety' (AD-DepAnx). All other subjects were excluded from analyses of depression and anxiety. Insufficient data were available for subjects from the UK sibling sample regarding symptoms of anxiety; therefore analyses of depression and anxiety were restricted to the NIMH sample.

The Principal components analysis presented in chapter 2 found a behavioural component which contained items relating to aggression/agitation, euphoria, apathy, disinhibition, irritability, aberrant motor behaviour, sleep disturbances and appetite abnormalities. Reliable and complete data were available from the UK and NIMH sibling samples regarding symptoms of aggression. However, the NPI was not used to assess the subjects in the family based samples. Therefore, data

regarding euphoria, apathy, disinhibition, irritability, aberrant motor behaviour, sleep disturbances and appetite abnormalities were limited. As such, data analysis was restricted to symptoms of aggression. Aggression had a moderate loading of 0.53 on the frontal lobe component in PCA of the full sample in chapter 2. This increased to 0.60 when analysis was restricted to just those in the moderately severe to severe stages of dementia and increased further, to 0.66, in analysis controlling for the effects of disease severity.

Participants in the NIMH and UK samples were coded as 'AD with aggression' (AD+Aggr) if they had displayed unprovoked physical or verbal aggression. Symptoms of aggression in the NIMH sample were assessed by means of a semi structured interview, in which the informant was asked whether the subject had made either verbal or physical threats, or displayed violent behaviour. Aggressive symptoms were also assessed as part of the Brief Psychiatric Rating Scale (BPRS) (Overall and Gorman 1962). Individuals were scored positive for AD+Aggr if they had made physical or verbal threats, or scored highly on the hostility item scored as part of the BPRS. Participants in the UK sample were assessed using the Behavioural Pathology in AD Rating Scale (BEHAVE-AD) (Reisberg et al. 1987) and were coded as 'AD with aggression' (AD+Aggr) if they had displayed unprovoked physical or verbal aggression. Those who had not displayed these symptoms were classified as 'AD without aggression' (AD-Aggr).

In all analyses of behavioural symptoms, subjects were excluded from the analyses if they had a previous history of mood disorders, bipolar disease, unipolar disease or an anxiety disorder. Subjects were also excluded if they did not meet criteria for either the presence or absence of each symptom (for example, those with incomplete data).

The Clinical Dementia Rating (CDR) Scale (Hughes et al. 1982) and Global Deterioration Scale (GDS) (Reisberg et al. 1982) were used to rate severity of dementia in the NIMH and UK samples, respectively. The GDS is described in section 2.2.2. The CDR offers an alternate means of staging dementia, providing ratings on a 7 point scale, with anchor points for unaffected (0), questionable (0.5), mild (1), moderate (2), severe (3), profound (4) and terminal (5). For the purposes

of further analyses GDS and CDR scores were recoded to make them more compatible. Combined severity ratings comprised very mild (CDR 0.5, GDS 2) mild (CDR 1, GDS 3) moderate (CDR 2, GDS 4) moderately severe (CDR 3, GDS 5) severe (CDR 4, GDS 6) and very Severe (CDR 5, GDS 7).

3.2.3 Statistical analysis

Intraclass correlation coefficients (ICC) offer a means of estimating the level of correlation between two variables (X and Y) among unordered pairs (e.g. where the attribution of siblings to X and Y is arbitrary) and were estimated to assess the relationship between AAO among siblings. Intraclass correlation coefficients are based on analysis of variance methods (ANOVA), which allow the variation in continuous traits (e.g. AAO) to be partitioned into between- and within- subjects variation, which in this example is analogous to between- and within- family variation in AAO. Intraclass correlations were estimated to quantify the proportion of variance in AAO accounted for by within-family (or familial) variation. Data originated from families with 2, 3, 4 and 5 affected siblings, which provided 1, 2, 3 and 4 independent pairs respectively (for example, a family with four affected members contributes 3 pairs, comprising the proband with siblings two, three and four). Intraclass correlations, using a one-way random effects method, were estimated for these four subsets of data in *separate* analyses using SPSS 12.0.1 (SPSS Inc. 2003). To combine these four ICCs into one overall estimate of the correlation coefficient, the estimates of the intraclass correlation coefficients were subjected to Fisher's z-transformation, with the resulting normal variates having the approximate sampling variance of $1/n(-2)$, where n is the number of pairs (Donner 1986). These four estimates were then averaged to give an overall intraclass correlation, weighted by the inverse of sampling variances. The APOE ε4 allele has been shown to reduce age at disease onset in a dose dependent manner (Corder et al. 1993) therefore when considering AAO, separate analyses were performed after adjusting AAO according to the number of ε4 APOE alleles possessed. This was achieved by performing a linear regression model to predict AAO using the number of ε4 alleles as the independent variable. The residuals were stored and used to calculate intraclass correlations as outlined above. SPSS 12.0.1 (SPSS Inc. 2003) was used to complete these analyses.

Generalised Estimating Equations (GEE) were used to assess the relationship of behavioural symptoms among siblings. The focus of this analysis was to determine the association of behavioural symptoms between probands and siblings. As there can be multiple pairs of probands and siblings from a single pedigree all pairs in the data set cannot be considered to be independent. GEE, as implemented in the SAS Proc Genmod (SAS Institute Inc., 1999-2001), were used to adjust for the non-independence of pairs in the data. In GEE, the regression coefficients in the model are estimated assuming that the observations are statistically independent. The standard errors of these coefficients are estimated in such a way that takes into account the correlation of the observations within families, and they do not tend to be greatly affected if this correlation structure is specified incorrectly. A binomial distribution was assumed for the presence or absence of behavioural symptoms and hence the logit-link function was used. Presence of the behavioural symptom of the proband was treated as the predictor variable and the presence of the symptom in the sibling was the outcome variable.

This method is akin to creating a 2x2 contingency table of behavioural symptoms for probands (rows) and siblings (columns). It would then be possible to identify whether there is a relationship between proband-sibling pairs and their behavioural symptoms. Using the 2x2 contingency table an odds ratio could be calculated to represent the odds of the sibling displaying the symptom if the proband was positive for the behaviour; compared to if the proband had not displayed the behaviour. Such an approach assumes that all sibling pairs are independent. The GEE method takes into account the non-independence owing to the fact that multiple pairs can arise from a single family. In addition, the GEE method allows covariates to be included in the analysis. As behavioural symptoms in AD may vary as a function of disease stage, sibling disease severity was entered as a covariate, along with sibling's AAO and gender.

3.3 Results

3.3.1 Characteristics of the sample

Basic characteristics of all affected sibling pair samples can be found in table 3.1. Both probands and siblings were predominantly female, 324/452 (71.7%) and 405/531 (76.3%) respectively. AAO ranged from 65 to 93 years in probands, with a mean (sd) of 74.86 (5.76) years. The mean (sd) AAO in siblings was 75.06 (6.05) years, ranging from 65 to 97 years. Participants were relatively well distributed with regard to severity with 3 (0.4%), 135 (16.6%), 222 (27.4%), 247 (30.4%), 164 (20.2%) and 41 (5.1%) in the very mild, mild, moderate, moderately severe, severe and very severe stages of dementia respectively. Classifications of behavioural symptoms in the NIMH, UK and combined sample can be seen in table 3.2.

3.3.2 Familiality of age at onset

Given that family selection was slightly different for analyses of AAO the total sample size comprised 307 probands and 381 siblings from the NIMH sample; 90 probands and 101 siblings from the UK sample; and 62 probands and 76 siblings from the NIA sample. The UK sample contained 8 siblings with an AAO less than 65 (range 61-64 years), the NIMH sample comprised 38 (range 42 to 64 years), whilst the NIA sample included 5 (range 57 to 64 years). After including siblings who presented with AD before the age of 65 the mean AAO (sd) among siblings were 73.48 (7.16), 75.29 (7.39) and 72.95 (7.03) years in the NIMH, UK and NIA samples, respectively. The full sample comprised 378 affected sibling pairs, 64 affected sibling trios, 16 affected sibling quartets and 1 affected sibling quintet, contributing 378, 128, 48 and 4 independent pairs, respectively. The intraclass correlation coefficient (ICC) for AAO among sibling pairs was 0.37 (95% CI = 0.22-0.50, $p < 0.001$), 0.39 (95% CI = 0.14-0.60, $p < 0.001$) and 0.44 (95% CI = 0.12-0.69, $p < 0.001$) in the NIMH, UK and NIA samples respectively. In the combined sample the intraclass correlation was 0.38 (95% CI = 0.26-0.49, $p < 0.001$).

Table 3.1 Basic characteristics of the NIMH, UK and NIA sibling samples used in familiality analysis

	NIMH Sample		UK Sample		NIA Sample		Combined Sample	
	Proband	Sibling	Proband	Sibling	Proband	Sibling	Proband	Sibling
n	287	344	101	115	64	72	452	531
Gender n (%)								
Male	78 (27.2)	84 (24.4)	25 (24.8)	21 (18.3)	25 (39.1)	21 (29.2)	128 (28.3)	126 (23.7)
Female	209 (72.8)	260 (75.6)	76 (75.2)	94 (81.7)	39 (60.9)	51 (70.8)	324 (71.7)	405 (76.3)
Mean age at assessment, years	80.95	81.35	82.19	82.85	-	-	81.23	81.67
range	68-100	68-101	68-93	69-102	-	-	68-100	68-102
missing Ages	0	0	16	22	-	-	16	22
Mean age at onset, years	74.70	74.94	76.68	76.52	73.20	73.74	74.86	75.06
range	65-93	65-97	65-89	65-91	65-84	65-95	65-93	65-97
missing AOO	0	1	16	20	0	0	16	21
Mean disease duration, years	6.25	6.43	5.36	6.27	-	-	6.06	6.40
range	1-27	0.5-24	0.5-17	0.5-20	-	-	0.5-27	0.5-24
missing durations	0	1	19	26	-	-	19	27
Disease Severity n (%) [a]								
Very Mild (CDR 0.5, GDS 2)	0 (0.0)	0 (0.0)	2 (2.3)	1 (1.0)	-	-	2 (0.5)	1 (0.2)
Mild (CDR 1, GDS 3)	59 (20.6)	48 (14.1)	10 (11.6)	18 (18.4)	-	-	69 (18.5)	66 (15.0)
Moderate (CDR 2, GDS 4)	84 (29.3)	115 (33.7)	11 (12.8)	12 (12.2)	-	-	95 (25.5)	127 (28.9)
Moderately Severe (CDR 3, GDS 5)	94 (32.8)	103 (30.2)	32 (37.2)	18 (18.4)	-	-	126 (33.8)	121 (27.6)
Severe (CDR 4, GDS 6)	37 (12.9)	55 (16.1)	28 (32.6)	44 (44.9)	-	-	65 (17.4)	99 (22.6)
Very Severe (CDR 5, GDS 7)	13 (4.5)	20 (5.9)	3 (3.5)	5 (5.1)	-	-	16 (4.3)	25 (5.7)
Missing	0	3	15	17			15	20

[a] Severity was rated using the Clinical Dementia Rating Scale (CDR) in the NIMH Sample and Global Deterioration Scale (GDS) in the UK Sample

Table 3.2 Classification of behavioural symptoms in the NIMH, UK and combined sibling samples.

	NIMH Sample		UK Sample		Combined Sample	
	Proband	**Sibling**	**Proband**	**Sibling**	**Proband**	**Sibling**
Psychosis						
n	287	344	101	115	388	459
AD with Psychosis n (%)	154 (58.6)	185 (60.9)	51 (62.2)	50 (56.8)	205 (59.4)	235 (59.9)
AD without psychosis n (%)	109 (41.4)	119 (39.1)	31 (37.8)	38 (43.2)	140 (40.6)	157 (40.1)
Unspecified - Mild delusions	0	0	3	8	3	8
Unspecified - History of psychiatric illness	2	8	2	0	4	8
Unspecified - Incomplete data	22	32	14	19	36	51
Minor Depression						
n	287	344	101	115	388	459
AD with Minor Depression n (%)	85 (33.5)	91 (31.2)	43 (57.3)	31 (37.3)	128 (38.9)	122 (32.5)
AD without Minor Depression n (%)	169 (66.5)	201 (68.8)	32 (42.7)	52 (62.7)	201 (61.1)	253 (67.5)
Unspecified - History of psychiatric illness	2	8	2	0	4	8
Unspecified - Incomplete data	31	44	24	32	55	76
Major depression						
n	287	344	101	115	388	459
AD with Major Depression n (%)	43 (20.3)	40 (16.6)	13 (28.9)	9 (14.8)	56 (21.8)	49 (16.2)
AD without Major Depression n (%)	169 (79.7)	201 (83.4)	32 (71.1)	52 (85.2)	201 (78.2)	253 (83.8)
Unspecified - Mild depression	42	51	26	21	68	72
Unspecified - History of psychiatric illness	2	8	2	0	4	8
Unspecified - Incomplete data	31	44	28	33	59	77
Depression & Anxiety						
n	287	344	-	-	287	344
AD with depression & anxiety n (%)	64 (31.1)	58 (25.8)	-	-	64 (31.1)	58 (25.8)
AD without depression & anxiety n (%)	142 (68.9)	167 (74.2)	-	-	142 (68.9)	167 (74.2)
Unspecified - Mild depression/anxiety	35	47	-	-	35	47
Unspecified - History of psychiatric illness	2	8	-	-	2	8
Unspecified - Incomplete data	44	64	-	-	145	179
Aggression						
n	287	344	101	115	388	459
AD with aggression n (%)	134 (51.9)	160 (52.3)	32 (44.4)	31 (40.3)	166 (50.3)	191 (49.9)
AD without aggression n (%)	124 (48.1)	146 (47.7)	40 (55.6)	46 (59.7)	164 (49.7)	192 (50.1)
Unspecified - Very mild aggression	3	3	12	16	15	19
Unspecified - History of psychiatric illness	2	8	2	0	4	8
Unspecified - Incomplete data	24	27	15	22	39	49

Increasing APOE ε4 alleles were associated with reduced AAO in a dose-dependent manner. Compared to those with no ε4 alleles the regression coefficient for those with one ε4 allele was $\beta = -3.137$, $p < 0.001$, whilst for those with two ε4 alleles it was $\beta = -6.560$, $p < 0.001$. Intraclass correlations were calculated after removing the effects of APOE. Complete APOE data were available for 344 affected sibling pairs, 52 affected sibling trios and 6 affected sibling quartets. In this sub sample the intraclass correlation for AAO among sibling pairs was 0.41 (95% CI = 0.29-0.52, $p < 0.001$). After removing the effect of

APOE on AAO for both probands and siblings the intraclass correlation reduced slightly to 0.38 (95% CI = 0.25-0.49, $p < 0.001$).

3.3.3 Familiality of psychosis

Classification of psychotic symptoms was possible for 345 (88.9%) probands and 392 (85.4%) siblings. Three probands and eight siblings did not meet criteria for AD+P or AD-P, as they had displayed mild delusional behaviour which was not sufficient to meet criteria for AD+P. Forty probands and 59 siblings were excluded because they either did not have data regarding psychotic symptoms or had a previous history of psychiatric disturbance. The prevalence of psychotic symptoms was comparable in the UK and NIMH samples, and among probands and siblings.

Demographic and clinical characteristics of the proband's siblings can be found in table 3.3. There was a significant association between proband psychosis status and the occurrence of AD+P in siblings in both the UK and NIMH samples. After fitting a GEE-based logit model, the estimated odds ratios for AD+P in siblings who had probands with AD+P in the NIMH sample was 3.37 (95% CI, 2.06-5.51; $p < 0.001$). This finding was replicated in the UK sample, with an OR of 3.10 (95% CI, 1.21-7.96; $p < 0.019$). In the combined sample the estimated odds ratio was 3.32 (95% CI, 2.15-5.13; $p < 0.001$). A significant association between proband psychosis status and the occurrence of AD+P in siblings was observed after controlling for sibling disease severity, OR = 3.25 (95% CI, 2.10-5.02; $p < 0.001$). In supplementary analysis, AD+P was significantly more prevalent among female siblings compared to males, OR = 1.80 (95% CI, 1.07-3.02; $p < 0.026$). Proband psychosis status remained a significant predictor of AD+P in siblings after controlling for sibling AAO, sibling disease severity and gender, OR = 3.30 (95% CI, 2.12-5.13; $p < 0.001$).

Table 3.3 Clinical characteristics of siblings of AD probands with and without psychotic symptoms.

Sibling Characteristics	Proband psychosis status	
	AD-Psychosis	AD+Psychosis
n	165	245
Gender n (%)		
Male	43 (26.1)	53 (21.6)
Female	122 (73.9)	192 (78.4)
Mean age at assessment, years	81.38	81.98
range	68-102	68-101
Mean age at onset, years	75.09	75.57
range	65-91	65-97
Mean disease duration, years	6.32	6.46
range	1-24	0.5-21
Disease Severity n (%) a		
Very Mild (CDR 0.5, GDS 2)	1 (0.6)	0 (0.0)
Mild (CDR 1, GDS 3)	29 (17.8)	33 (13.6)
Moderate (CDR 2, GDS 4)	49 (30.1)	67 (27.6)
Moderately Severe (CDR 3, GDS 5)	38 (23.3)	75 (30.9)
Severe (CDR 4, GDS 6)	35 (21.5)	55 (22.6)
Very Severe (CDR 5, GDS 7)	11 (6.7)	13 (5.3)
Missing	2	2
Psychosis status n (%)		
AD-P	88 (57.9)	66 (28.9)
AD+P	64 (42.1)	162 (71.1)
Unknown - Mild Delusions	2	6
Unknown - Insufficient Data	11	11

[a] Severity was rated using the Clinical Dementia Rating Scale (CDR) in the NIMH Sample and Global Deterioration Scale (GDS) in the UK Sample

3.3.4 Familiality of mood

Minor Depression

Data regarding symptoms of minor depression were available for subjects from the NIMH and UK samples. 329 (84.8%) probands and 375 (81.7%) siblings met criteria for either AD-D or AD+D. 55 probands and 76 siblings did not have data regarding depression. Four probands and eight siblings were classified as unknown as they had a previous history of mood disorders, bipolar disease, unipolar disease or an anxiety disorder. The prevalence of AD+D was significantly

higher in UK probands, compared to NIMH probands ($\chi^2(1)= 13.89$, $p < 0.001$) and UK siblings ($\chi^2(1)= 6.32$, $p=0.012$).

Table 3.4 Clinical characteristics of siblings of AD probands with and without minor depression.

Sibling Characteristics	Proband minor depression status	
	AD-Minor Depression	AD+Minor Depression
n	238	147
Gender n (%)		
Male	58 (24.4)	31 (21.1)
Female	180 (75.6)	116 (78.9)
Mean age at assessment, years	81.84	81.85
range	68-102	69-101
Mean age at onset, years	75.7	74.88
range	65-92	65-91
Mean disease duration, years	6.23	6.81
range	1-22	0.5-24
Disease Severity n (%) [a]		
Very Mild (CDR 0.5, GDS 2)	0 (0)	0 (0)
Mild (CDR 1, GDS 3)	36 (15.3)	18 (12.3)
Moderate (CDR 2, GDS 4)	73 (31.1)	39 (26.7)
Moderately Severe (CDR 3, GDS 5)	69 (29.4)	40 (27.4)
Severe (CDR 4, GDS 6)	42 (17.9)	42 (28.8)
Very Severe (CDR 5, GDS 7)	15 (6.4)	7 (4.8)
Missing	3	1
Minor Depression status n (%)		
AD-Minor Depression	160 (73.7)	76 (57.1)
AD+Minor Depression	57 (26.3)	57 (42.9)
Unspecified - History of depression	3	3
Unspecified - Insufficient Data	18	11
Major Depression status n (%)		
AD-Major Depression	160 (86.5)	76 (79.2)
AD+Major Depression	25 (13.5)	20 (20.8)
Unspecified - Mild Depression	32	35
Unspecified - History of depression	3	3
Unspecified - Insufficient Data	18	12

[a] Severity was rated using the Clinical Dementia Rating Scale (CDR) in the NIMH Sample and Global Deterioration Scale (GDS) in the UK Sample

Demographic and clinical characteristics of the proband's siblings can be found in table 3.4. AD+D in probands was significantly associated with AD+D in siblings in the NIMH sample, OR = 2.07 (95% CI, 1.21-3.55; p=0.008). AD+D in probands was associated with over a two-fold increase of AD+D in siblings from the UK sample, however this result did not meet criteria for statistical significance, OR = 2.39 (95% CI, 0.89-6.45; p=0.086). An odds ratio of 2.17 (95% CI, 1.36-3.46; p < 0.001) was observed in the combined UK and NIMH sample. In the combined sample sibling's disease severity, gender and AAO were entered into the model as covariates but did not significantly explain any of the variance in AD+D in siblings. Furthermore, AD+D status in probands remained a significant predictor of AD+D in siblings after controlling for siblings disease severity, gender, AAO and age at collection, OR = 2.16 (95% CI, 1.34-3.46; p=0.002).

Major Depression

56 probands and 49 siblings met criteria for AD with major depression (AD+MajD), 201 probands and 253 siblings had displayed no symptoms of depression. 68 probands and 72 siblings had mild depression but did not meet criteria for AD+MajD and were therefore categorised as unknown in analyses of major depression. A further 63 probands and 85 siblings were excluded from the analysis as they either had insufficient data or had a previous history of psychiatric illness.

After excluding families in which the proband had insufficient data, only 4 siblings In the UK sample met criteria for AD with major depression (AD+MajD). This sample was inadequate for independent hypothesis testing; therefore analysis was performed on the combined sample only. Demographic and clinical characteristics of the proband's siblings can be found in table 3.5. Major depression in probands was associated with around a two-fold increase in the corresponding symptom among siblings, but this result did not meet criteria for statistical significance, OR = 1.96 (95% CI, 0.87-4.45; p < 0.106). In supplementary analyses, neither sibling's AAO, gender nor disease severity significantly predicted siblings AD+MajD status.

Table 3.5 Clinical characteristics of siblings of AD probands with and without major depression.

Sibling Characteristics	Proband major depression status	
	AD-Major depression	AD+Major depression
n	238	63
Gender n (%)		
Male	58 (24.4)	13 (20.6)
Female	180 (75.6)	50 (79.4)
Mean age, years	81.84	81.61
range	68-102	71-101
Mean age at onset, years	75.70	74.44
range	65-92	65-88
Mean disease duration, years	6.23	7.19
range	1-22	1-24
Disease Severity n (%) [a]		
Very Mild (CDR 0.5, GDS 2)	0 (0.0)	0 (0.0)
Mild (CDR 1, GDS 3)	36 (15.3)	11 (17.5)
Moderate (CDR 2, GDS 4)	73 (31.1)	14 (22.2)
Moderately Severe (CDR 3, GDS 5)	69 (29.4)	20 (31.7)
Severe (CDR 4, GDS 6)	42 (17.9)	14 (22.2)
Very Severe (CDR 5, GDS 7)	15 (6.4)	4 (6.3)
Missing	3	0
Major depression status n (%)		
AD-Major depression	160 (86.5)	33 (76.7)
AD+Major depression	25 (13.5)	10 (23.3)
Unspecified - History of depression	3	2
Unspecified - Mild Depression	32	12
Unspecified - Insufficient Data	18	5
Minor depression status n (%)		
AD-Minor Depression	160 (73.7)	33 (58.9)
AD+Minor Depression	57 (26.3)	23 (41.1)
Unspecified - History of depression	3	2
Unspecified - Insufficient Data	18	5

[a] Severity was rated using the Clinical Dementia Rating Scale (CDR) in the NIMH Sample and Global Deterioration Scale (GDS) in the UK Sample

Depression and Anxiety

Data regarding AD with depression and anxiety were only available for individuals in the NIMH sample. Sixty-four probands and 58 siblings met criteria for AD+DepAnx, whilst 142 probands and 167 siblings were classified as AD-DepAnx.

35 probands and 47 siblings had displayed symptoms of either depression or anxiety, but these were not severe enough to meet full criteria for AD with depression and anxiety, whilst 2 probands and 8 siblings were classified as unknown as they had previous history of mood disorders, bipolar disease, unipolar disease or an anxiety disorder. The prevalence of AD+DepAnx was comparable in probands and siblings, $\chi^2(1)= 1.48$, p=0.223.

Clinical characteristics of siblings of AD+DepAnx and AD-DepAnx probands can be seen in table 3.6. AD+DepAnx in probands was associated with just over a two-fold increase in AD+DepAnx in siblings, but this result did not reach criteria for statistical significance, OR = 2.09 (95% CI, 1.00-4.38; $p < 0.052$). In supplementary analyses, sibling gender, AAO and disease severity were added as covariates, which revealed that male siblings were significantly more likely to have displayed depression and anxiety than females, OR = 2.48 (95% CI, 1.14-5.40; $p < 0.023$). Furthermore, when gender was added as a covariate, proband AD+DepAnx status significantly predicted AD+DepAnx status in siblings, OR = 2.40 (95% CI, 1.12-5.12; $p < 0.024$). Sibling severity and AAO were not associated with AD+DepAnx in siblings.

3.3.5 Familiality of aggression

Sufficient data were available to classify 330/388 probands (85.1%) and 383/459 siblings (83.4%) as either AD+Agg or AD-Agg. 15 probands and 19 siblings were classified as unknown for AD with aggression as they had displayed very mild aggressive behaviour which could have been provoked, and therefore did not meet criteria for AD+Agg or AD-Agg. 43 probands and 57 siblings were classified as unknown as they either had a previous history of psychiatric illness or had insufficient data. No difference was observed in the prevalence of aggression in probands and siblings, $\chi^2(1)= 0.013$, p=0.908, however a significantly higher prevalence of aggression was noted among NIMH families compared to those from the UK, $\chi^2(1)= 4.570$, p=0.033. Clinical characteristics of siblings of AD+Agg and AD-Agg probands can be seen in table 3.7.

Table 3.6 Clinical characteristics of siblings of AD probands with and without depression & anxiety.

Sibling Characteristics	Proband Depression & Anxiety status	
	AD-Depression & Anxiety	AD+Depression & Anxiety
n	172	75
Gender n (%)		
Male	49 (28.5)	14 (18.7)
Female	123 (71.5)	61 (81.3)
Mean age at assessment, years	81.35	81.83
range	68-95	72-101
Mean age at onset, years	75.33	74.45
range	65-90	65-88
Mean disease duration, years	6.06	7.38
range	1-21	0.5-24
Disease Severity n (%) [a]		
Very Mild (CDR 0.5, GDS 2)	0 (0)	0 (0)
Mild (CDR 1, GDS 3)	23 (13.6)	11 (14.7)
Moderate (CDR 2, GDS 4)	63 (37.3)	18 (24)
Moderately Severe (CDR 3, GDS 5)	49 (29)	29 (38.7)
Severe (CDR 4, GDS 6)	21 (12.4)	14 (18.7)
Very Severe (CDR 5, GDS 7)	13 (7.7)	3 (4)
Missing	3	0
Depression/Anxiety status n (%)		
AD-Depression & Anxiety	97 (78.2)	31 (63.3)
AD-Depression & Anxiety	27 (21.8)	18 (36.7)
Unspecified - History of depression	3	2
Symptoms too mild to meet criteria	22	12
Unspecified - Insufficient Data	23	12
Minor Depression status n (%)		
AD-Minor Depression	115 (74.2)	43 (62.3)
AD+Minor Depression	40 (25.8)	26 (37.7)
Unspecified - History of depression	3	2
Unspecified - Insufficient Data	14	4
Major Depression status n (%)		
AD-Major Depression	115 (85.2)	43 (78.2)
AD+Major Depression	20 (14.8)	12 (21.8)
Unspecified - Mild Depression	20	14
Unspecified - History of depression	3	2
Unspecified - Insufficient Data	14	4

[a] Severity was rated using the Clinical Dementia Rating Scale (CDR) in the NIMH Sample and Global Deterioration Scale (GDS) in the UK Sample

In the NIMH sample aggression in probands was associated with a significant increase in AD+Agg prevalence among siblings, OR = 2.35 (95% CI, 1.47-3.75; p < 0.001). This finding replicated in the UK sample, OR = 2.75 (95% CI, 1.01-7.44; p=0.047). The combined sample yielded an odds ratio of 2.38 (95% CI, 1.56-3.63; p < 0.001). When entered as covariates sibling disease severity, gender, AAO and current age did not contribute to the model, whilst proband aggression status remained a significant predictor of AD+Agg in siblings after controlling for all covariates, OR = 2.33 (95% CI, 1.52-3.57; p < 0.001). A summary of the Generalized Estimating Equation models for all behavioural symptoms can be seen in table 3.8.

Table 3.7 Clinical characteristics of siblings of AD probands with and without aggressive symptoms.

	Proband Aggression status	
Sibling Characteristics	AD-Aggression	AD+Aggression
n	196	194
Gender n (%)		
Male	44 (22.4)	44 (22.7)
Female	152 (77.6)	150 (77.3)
Mean age at assessment, years	81.67	81.94
range	68-99	69-102
Mean age at onset, years	75.35	75.56
range	65-97	65-92
Mean disease duration, years	6.27	6.42
range	0.5-24	0.5-20
Disease Severity n (%) [a]		
Very Mild (CDR 0.5, GDS 2)	0 (0.0)	1 (0.5)
Mild (CDR 1, GDS 3)	32 (16.6)	25 (13.0)
Moderate (CDR 2, GDS 4)	62 (32.1)	50 (25.9)
Moderately Severe (CDR 3, GDS 5)	44 (22.8)	64 (33.2)
Severe (CDR 4, GDS 6)	42 (21.8)	44 (22.8)
Very Severe (CDR 5, GDS 7)	13 (6.7)	9 (4.7)
Missing	3	1
Aggression status n (%)		
AD-aggression	108 (60.0)	71 (38.6)
AD+aggression	72 (40.0)	113 (61.4)
Unspecified - Very mild aggression	7	2
Unspecified - Insufficient Data	9	8

[a] Severity was rated using the Clinical Dementia Rating Scale (CDR) in the NIMH Sample and Global Deterioration Scale (GDS) in the UK Sample

Table 3.8 Summary of generalised estimating equation models for behavioural symptoms.

	NIMH Sample		UK Sample		Combined Sample					
	OR[1]	95% CI	OR[1]	95% CI	OR[1]	95% CI	OR[1] | Sibling Severity	95% CI	OR[1] | Sibling Severity, AAO, Gender	95% CI
Psychosis	3.37**	2.06-5.51	3.10*	1.21-7.96	3.32**	2.15-5.13	3.25**	2.10-5.02	3.30**	2.12-5.13
Minor Depression	2.07**	1.21-3.55	2.39	0.89-6.45	2.17**	1.36-3.46	2.22**	1.39-3.54	2.16**	1.34-3.46
Major Depression	-	-	-	-	1.96	0.87-4.45	1.96	0.87-4.40	1.88	0.84-4.21
Depression/Anxiety	2.09	1.00-4.38	-	-	2.09	1.00-4.38	2.07	0.99-4.33	2.40*	1.12-5.12
Aggression	2.35**	1.47-3.75	2.75*	1.01-7.44	2.38**	1.56-3.63	2.32**	1.52-3.54	2.33**	1.52-3.57

[1] Odds Ratio's (OR) represent the risk that each symptom will occur in siblings if the proband has displayed the symptom, compared to if the proband has not displayed the symptom.

* $p < 0.05$

** $p < 0.01$

3.4 Discussion

3.4.1 Summary of findings

The UK, NIMH and NIA siblings were similar to the non-family based sample reported in chapter 2 in terms of AAO, age of assessment, disease duration, gender ratio and disease severity. Symptoms of aggression, psychosis and depression were common in both the NIMH and UK families, with comparable prevalence rates to those reported in the non-family based sample reported in chapter 2. A significant familial effect on age at disease onset was observed in the three independent samples of sibling pairs. The APOE ε4 allele was associated with a lower AAO. However, the familial effect on AAO remained significant after controlling for APOE ε4 status. Significant familial aggregation was noted with symptoms of psychosis and aggression in the NIMH siblings, which also replicated in the independent sample collected in the UK. In the NIMH sample, minor depression clustered in families more often that would be expected by chance, whilst a trend towards familiality was observed among the UK sibling pairs. Neither depression and anxiety, nor major depression, were significantly related among sibling pairs in univariate analyses. However, depression and anxiety showed some evidence of familiality in a model including gender. Perhaps the most notable result was the strong relationship between psychotic symptoms observed among sibling pairs with AD. Those with psychotic probands were over three times as likely to develop symptoms of psychosis, compared to siblings of non-psychotic probands. This finding was highly significant in both the UK and NIMH sample, reducing the possibility of a type II error.

3.4.2 Discussion of findings in relation to previous literature

Strong evidence was found for familial effects on AAO in AD, which was not attributable to APOE. This supports previous findings by Tunstall and colleagues (2000) who noted familial influence on AAO in a sample which contained data from a large proportion of the UK siblings used here. The results also support the conclusion drawn by Warwick-Daw and Colleagues (2000) who suggested that there are a number of unidentified genes which exert a sizable effect on AAO in AD.

The findings in relation to psychotic symptoms offer support to what has been observed previously in AD, however it should be noted that there is considerable overlap between the UK sample and that used by Tunstall and colleagues (2000), whereas the previous report of familial aggregation of psychosis in AD by Sweet and colleagues (2000b) includes a proportion of the NIMH siblings used in this analysis. In fact, only 8% of probands, and 10% of the siblings in the NIMH sample were exclusive to the analyses reported here. Sweet and colleagues included individuals diagnosed with possible AD and those with a disease onset less than 65 years; as such 30% of probands and 31% of siblings selected by Sweet and colleagues were not used in the analysis presented here. It is interesting to note that the familial influence on psychotic symptoms remained prominent in analyses restricted to sibling pairs with late-onset probable, or definite, AD. This, along with the highly comparable effect sizes noted in the UK and NIMH sample, imply that the finding of familiality of psychosis in AD is robust.

Using a different method to the one used here, Tunstall and colleagues (2000) reported familiality of aggressive symptoms in AD, represented by a slight elevation in pairwise concordance for aggression among sibling pairs compared to what would be expected by chance. This is the first study to statistically assess the relationship between aggressive symptoms among relatives suffering with AD. Siblings of those with aggressive symptoms were more than twice as likely to display aggression. The relationship was less marked than that observed with psychosis, however the finding appears to be robust, replicating in the UK sample.

Three variables were used to assess the behavioural component of mood in the sibling pairs. In the main analyses only the broad definition of depression showed significant evidence of familial aggregation, with the prevalence of depression doubling in siblings of those with AD+D. This effect did not replicate in the smaller UK sample, however there was a trend in the same direction. The familial effect appears to be small; as such the UK sample may have been inadequately powered to detect this effect. No significant familial effects were noted with *major* depression. Likewise analysis of the combined depression and anxiety variable did not initially reveal statistically significant evidence for familial aggregation. However, after including gender as a covariate, depression and anxiety in

probands was associated with a significant two-fold increase in siblings. Using an overlapping sample, Tunstall and colleagues reported a moderate familial influence on mood. Others have reported a relationship between family history of depression and the prevalence of mood disturbances in AD (Butt and Strauss 2001; Lyketsos et al. 1996; Pearlson et al. 1990; Strauss and Ogrocki 1996). These studies strongly suggest a genetic influence in the aetiology of depressive symptoms in AD. The findings here offer some suggestive evidence in support of the previous findings, but the results are far from conclusive.

3.4.3 Implications and future research

It is important to consider whether familiality observed in studies such as this represents increased genetic liability. As discussed in section 3.1.1, familiality represents the total effect of genetic heritability and shared environment among relative pairs; however this analysis is not able to differentiate between these two effects. Some have suggested that shared environment effects are limited in later life (Sweet et al. 2002; Tunstall et al. 2000), which is supported by several twin studies which have reported limited shared environment effects when assessing the risk of developing AD (Gatz et al. 1997; Pedersen et al. 2004; Pedersen et al. 2001; Raiha et al. 1996). Sullivan, Neale and Kendler (2000) provided a meta-analysis of five family based studies of major depression, not related to AD, and found shared environment to have a negligible effect in explaining susceptibility to the disease. It therefore seems reasonable to assume that shared environment does not exert a substantial influence on depression observed in AD, although it cannot be completely ruled out. Familial aggregation of psychosis, aggression and age at disease onset may also be explained by environmental influences shared between family members in these analyses. However, as already discussed, genetic variation has been shown to modify AAO (Meyer et al. 1998) and susceptibility to psychotic and aggressive symptoms in AD (Sweet et al. 1998), illustrating that genetic variation may account for elements of clinical heterogeneity. It therefore seems more plausible to assume that genetic effects are the most important contributor to familiality, however further study is required to confirm this.

Conclusions from the findings in relation to AAO should be made with caution. Methods of sample ascertainment in sibling pair collections raise the possibility that siblings with a similar AAO are more likely to be sampled together, as they are more likely to be living with the disease at the same time. The optimal estimate of familiality would be gained through collection of population based studies with complete ascertainment of affected family members. It is encouraging that the familial effect remained strong in the NIMH sample, which was collected over a wider time frame than the NIA and UK siblings, with a total period of collection lasting over 6 years. As such sampling bias among NIMH sibling pairs would be expected to have less influence on familiality estimates.

Given the limitations outlined, these findings provide evidence that age at disease onset and behavioural components of AD aggregate within families. The evidence regarding AAO and psychosis are perhaps the most convincing, displaying the strongest familial influence. Future studies may seek to further unravel the genetic aetiology of clinical heterogeneity in AD. Although replication in larger, population based twin and family samples would be advantageous these, and previous, findings provide adequate justification for studies aiming to map genes associated with clinical variation observed in AD. The findings in relation to aggression and depression were not as strong, indeed minor depression was the only mood related variable assessed to show significant familial aggregation. However, due to issues of power, these findings do not preclude a genetic influence on mood state in AD.

It is likely that familial effects identified here are genetically underpinned; hence genetic variation may modify AAO, psychotic symptoms, and to a lesser extent mood and aggression in AD. Alternatively these clinical characteristics may provide suitable 'candidate phenotypes' to assist in the delineation of more homogeneous subsets for future research. A number of authors have already performed genetic association studies to search for genes associated with the clinical variation observed in AD. They have generally focused on candidate genes for other neuropsychiatric disorders, for example dopamine (Craig et al. 2004b; Holmes et al. 2001; Sweet et al. 1998) and serotonin (Assal et al. 2004; Holmes et al. 1998a; Lam et al. 2004; Nacmias et al. 2001) receptor genes. Some positive

associations have been reported, but the findings to date are not entirely consistent. Linkage analysis offers an alternative to genetic association studies and can provide a powerful first step in identifying regions which are likely to harbour loci associated with disease phenotypes. The family based samples used in this chapter originate from a number of linkage studies of LOAD, therefore offer an opportunity to search for genes associated with clinical variation showing familial aggregation. In chapter 4, AAO, psychosis, aggression and minor depression will be used as covariates in the linkage analysis of affected relative pairs from the NIMH, UK and NIA samples. The primary goal is to identify genetic regions which are likely to contain genes which could modify the clinical presentation of AD, or alternatively increase susceptibility to genetically homogeneous forms of the disease characterised by different clinical presentations.

3.4.4 Methodological critique

This study suffers from a number of limitations which are worthy of discussion. First, there were concerns regarding the power of this sample to detect familial effects, which was particularly a problem when the prevalence of the behavioural components was low (e.g. major depression, depression and anxiety). As such it is difficult to draw firm conclusions from these analyses. However, it should be noted that this is among one of the largest samples to date which has been used to assess the familiality of clinical features in AD. In addition, power to detect familial effects may have been reduced by creating binary variables to determine presence or absence of behavioural components. If one assumes that clinical variation observed in AD is a quantitative trait, it would be beneficial to construct a continuous measure of behavioural components. Indeed, Tunstall and colleagues reported a significant familial effect on current mood state when mood was rated on a quantitative scale using a sample much smaller than the one reported here. This approach also negates the requirement for researchers to define criteria for presence or absence of particular symptoms. These data were collected for a genome screen of AD; therefore data regarding behavioural symptoms was limited and largely incompatible with defining quantitative measures of aggression, psychosis and mood. This study offers a good insight into the familial influence on

clinical variation observed in AD. However, this field would be advanced by the collection of large well powered family based samples, where the primary goal was to determine familial aggregation of clinical variation.

This study relies largely on information obtained using a retrospective interview with the AD sufferers informant. It is possible that certain informants may have a particular propensity to report certain symptoms more than others. If the same informant was used for multiple members of the same family this may falsely inflate the correlation between siblings. However, reviewing the limited data available it appears that the informant often differed for each sibling within a family (i.e. it was usually the participants next of kin), which is likely to limit the extent to which reporting bias affected familiality estimates. Unfortunately, insufficient data were available to assess statistically the effect of reporting bias by informants. It is noteworthy that all the behavioural components displayed either a trend towards, or significant, familial influence. However, given previous findings, there was an a priori assumption that these symptoms would aggregate within families, thus these concerns may be unfounded. In addition, the familial effects on psychosis and AAO where much stronger than those for other symptoms, hence they are less likely to have resulted merely from a reporting bias.

This study employs the behavioural components identified in the UK non family based sample in chapter 2, to inform the categorisation of behavioural symptoms. This could represent a flaw in the current study design. Ideally, factor analysis would have been performed directly in the sibling pair sample, adjusting for family effects. Consequently, the familiality of factor scores could have been assessed directly as a quantitative trait. However, the data collected for the sibling pair samples was not originally intended for this purpose. It was therefore less than ideal for factor analytical procedures (e.g. the assessment of a number of common symptoms including, euphoria, disinhibition, sleep and appetite disturbances was either limited or absent). It is possible that the behavioural components identified in chapter 2 were not appropriate to transfer into the affected sibling pair sample, as the participants used for the PCA comprised primarily sporadic, non familial, cases of AD. However, there is no evidence, either in chapter 2 or the literature (Haupt et al. 1992; Holmes and Lovestone 2002), to suggest that behavioural

symptoms differ among those with and without a family history of dementia. In addition, the prevalence of behavioural symptoms reported in the UK and NIMH sibling pairs was comparable to those reported in the sample used in chapter 2. Furthermore, the behavioural components identified in chapter 2 have been reported elsewhere on numerous occasions, using different rating scales and study populations. As such they can be regarded as robust behavioural components commonly observed in AD. It therefore seems reasonable to generalise findings from the UK non family based dataset to the sibling pair sample.

3.4.5 Conclusions

Principal components analysis identified three components which represent the broad range of behavioural symptoms commonly observed among AD sufferers. These findings were used to guide the characterisation of behavioural symptoms in a sample of affected sibling pairs, with the primary goal of assessing the familial influences on the clinical presentation of AD. Strong evidence was found to suggest that the presence of psychotic symptoms and age at disease onset in AD are familial. Likewise, aggression and depression showed some evidence of familial aggregation. The correlation between ages at onset among siblings was largely independent of the APOE ε4 allele. These findings support the hypothesis that genetic factors, at least in part, contribute to the clinical differences observed among AD sufferers. They also provide justification for studies aiming to map genes associated with the phenotypic heterogeneity observed in AD.

Chapter 4

Linkage analysis of late-onset Alzheimer's disease

4.1 Introduction

4.1.1 Genetics background

Linkage analysis provides a means of localising regions that may harbour genetic variants which increase susceptibility to disease. The human genome consists of a sequence of approximately 3,000 million bases. Long molecules of double stranded DNA located in the nucleus of every cell form chromosomes. In general, humans inherit 23 chromosomes, containing a copy of the human genome, from each parent. These comprise 22 pairs of autosomes, and a pair of sex chromosomes, X and a Y in males and two X's in females.

DNA provides the genetic blueprint which is used to produce the proteins that form part of the structure, and perform most of the cellular activities, of living organisms. *Genes* are segments of DNA that specify the amino acid sequence of proteins. They account for roughly 3% of the DNA sequence in the human genome and occur in approximately the same physical genomic location in all human beings (International Human Genome Sequencing Consortium 2004). The remaining DNA sequence is thought to be non-coding and is often referred to as 'junk DNA', however it is now known that some of these regions are involved in the regulation of nearby coding sequences (Dicks and Savva 2003). The entire human genome is currently thought to consist of approximately 20,000 to 25,000 genes (International Human Genome Sequencing Consortium, 2004; Stein 2004). Linear chains of amino acids are used to make up proteins. There are 20 such amino acids that are used as 'building blocks'. As such an enormous number of different and diverse proteins can be constructed. The structure of DNA allows accurate copies to be made which can be passed on from parents to offspring. The accuracy of replication is crucial as changes in the DNA sequence disrupt the coding sequence. This can lead to proteins which differ in structure and function, which can have a potentially harmful effect on the cell and the organism.

During reproduction a single cell, called the zygote, is formed from two gametes, the ovum and the sperm. Gametes are formed by a special type of cell division known as meiosis, which gives rise to cells which, rather than containing two sets of all chromosomes (diploid) have a single (haploid) set of the 22 autosomes and a sex chromosome. Therefore, the resulting zygote comprises the correct number of 23 diploid chromosomes. Two chromosomes in the zygote are said to be homologous if they belong to the same chromosomal family (e.g. both chromosome 1). Homologous chromosomes are similar in both length and sequence meaning that humans possess two copies of every gene – one from each parent. During meiosis crossing over of genetic material can occur, achieved by the exchange of genetic information between homologous chromosomes of paternal and maternal origin. As a result alternating segments of paternal and maternal DNA are produced, which are then passed onto the next generation (Sham and McGuffin 2002).

Ancestral mutations lead to variation which occurs at numerous positions, or loci, across the genome, meaning that the DNA sequence can take different forms. These different forms are called *alleles* (e.g. a_1, a_2, a_3, a_4 and so on). Since human cells are diploid there are two alleles at each locus. A pair of alleles is called a *genotype* (e.g. a_1/a_1, a_1/a_2, a_2/a_4 etc). If the two alleles at a locus are identical then the individual is said to be homozygous (e.g. a_1/a_1, a_2/a_2) for that genotype, if they are different they are said to be heterozygous (e.g. a_1/a_2). Changes in the DNA sequence can occur through the process of mutations. Allelic variations in the DNA sequence are called *polymorphisms*, which can take the form of substitutions, deletions, translocations and insertions. Substitutions refer to replacement of one base pair by another, deletions and insertions occur when one or more base pairs are deleted or inserted from the DNA sequence. DNA can also be translocated from one segment of the sequence to another. It is such mutations within genes which give rise to the genetic variability observed within the population. Mutations can cause, or increase, susceptibility to particular diseases, which are then passed down in families from generation to generation.

The basic nature of the gene was defined by Gregor Mendel over a century ago. He observed that many traits in plants and animals are heritable and coined the term 'gene' to refer to discrete units of information transmitted from parent to offspring. Mendel proposed the laws of dominance, segregation and independent assortment. Mendel's law of dominance stated that hereditary units passed on from each parent could be either *dominant* or *recessive*. Both dominant and recessive alleles could be inherited by offspring; however the dominant trait will mask the expression of the recessive trait. For example, if a_1 is dominant and a_2 is recessive, then homozygotes a_1/a_1 and heterozygotes a_1/a_2 have identical phenotypes. The trait coded for by the recessive allele is only observed among those with the homozygous a_2/a_2 genotype. As such, autosomal dominant disease would occur in the offspring of those with an a_1/a_2 mother and a_2/a_2 father at an affected/unaffected ratio of 1:1. The offspring of two unaffected 'carriers' of a recessive disease gene (e.g. a_1/a_2) would be expected to display a ratio of unaffected/affected individuals of 3:1. However, Mendelian segregation ratios are not observed in many genetic disorders, and they also do not generalise to continuously distributed traits such as height. As such Mendel's law of dominance needs a slight conceptual modification to account for *penetrance*, which refers to the probability of a given phenotype conditional on a particular genotype (Cardno and McGuffin 2002). Those carrying a genotype which conveys a penetrance to a particular trait of 1 will inevitably express the corresponding phenotype, conversely a genotype with a penetrance of 0 will have no effect on the trait. Under Mendel's law of dominance particular genotypes always have a penetrance of 0 or 1. However, this is generally not the case for complex disorders, with penetrance usually lying between 0 and 1. As such, genetic variation at a particular locus may increase the probability that a particular trait or phenotype will be displayed, but will not make it inevitable. Mendel's law of segregation stated that during the formation of reproductive cells, a parent's paired hereditary units are randomly separated; each gamete containing only one of the two traits. Hence all individuals obtain one randomly selected gene from each parent. Finally, he proposed the law of independent assortment which suggested that genes were inherited independently of each other. However, this was later proven to be incorrect.

4.1.2 An introduction to genetic linkage analysis

Mendel's law of assortment states the transmission of alleles at two or more loci are independent. It has since been shown that this is only true for loci on separate chromosomes. Some alleles are inherited together more often than would be expected by chance. Within families genes which are located in close proximity to one another tend to be transmitted together. This provides an opportunity for a form of research called 'genetic linkage analysis'. Linkage analysis measures the co-segregation of specific marker alleles and a disease phenotype within individual families. Markers are usually short tandem repeat polymorphisms (STRs). STRs are highly polymorphic and are therefore more informative for linkage (Oetting et al. 1995). If a particular marker genotype co-segregates with a disease within families it is possible that mutations at this marker, or a locus nearby, increase susceptibility to the disease. Linkage analysis has the particular advantage that it does not require any prior knowledge of the pathophysiological mechanisms of the disorder. As such it has proven hugely successful in mapping mutations responsible for Mendelian diseases (Risch 2000) and has also been used in the study of genetically complex disorders.

Linkage analysis makes use of genetic *markers* spread across the genome. Such markers are DNA polymorphisms which have two or more alleles in the population of interest. They are anchored to a specific locus and their order and location can be identified using a genetic map. Linkage analyses can be performed using two loci, (i.e. two marker loci, or one marker locus and the proposed disease locus), which is called *twopoint* (or single point) *linkage analysis*. Alternatively, a number of marker loci spread across a chromosome can be analysed simultaneously, referred to as *multipoint linkage analysis*.

Linkage analysis can be performed by formulating specific inheritance models (parametric) or by employing methods that do not require disease models to be predefined (model-free). Parametric methods are model dependent and require researchers to specify the mode of transmission (e.g. dominant or recessive), penetrance of the susceptibility locus, marker and disease frequencies. On the whole they are a powerful means of testing for linkage; however the power is

drastically reduced if the model parameters are mis-specified (Clerget-Darpoux et al. 1986), which can often lead to false positive results. This is a particular concern in complex diseases where genetic models are difficult to determine. In the absence of accurate parameters model-free methods are more robust; however they have less power to detect linkage if in fact the correct disease model is known (Holmans 2003). Model-free methods can also be performed using samples of affected relative pairs, rather than complex pedigrees which are preferred for parametric methods. Due to difficulty in ascertaining large pedigrees, especially in diseases of late-onset, relative pair approaches are often advantageous as they allow a greater proportion of the population to be sampled.

Model-free methods using relative pairs rely on information about alleles shared in the pedigree. If two individuals inherit the same allele from a common ancestor it is said to be shared 'identical by descent' (IBD). As each individual carries two alleles at a marker locus, a pair of individuals can either share 0, 1 or 2 alleles. In outbred populations it is possible to determine the prior probabilities that a pair of relatives will share 0, 1 or 2 alleles IBD. For example, if a mother and father possess genotypes a_1/a_2 and a_3/a_4, respectively, at locus A, assuming independent assortment of alleles their offspring can possess the following genotypes with equal probability: a_1/a_3, a_1/a_4, a_2/a_3 or a_2/a_4. It therefore follows that two siblings have the probabilities 0.25, 0.50, 0.25 of sharing 0, 1 or 2 alleles IBD respectively. It is possible to calculate prior IBD sharing probabilities for any given pair of relatives. As already noted, alleles that are in close proximity have a higher chance of being transmitted together from one generation to the next than alleles at loci which are unlinked (e.g. far apart). Hence, relatives co-segregating the same genetic disorder will have a higher probability of sharing alleles IBD at, or near, the locus influencing susceptibility to the disease. By genotyping markers in chromosomal regions of interest a test for linkage between the markers and the disease phenotype can be performed. To search for regions of interest genetic markers can be selected at set intervals which cover large chromosomal regions, or even the whole genome.

Unfortunately, determining alleles shared IBD is often difficult when samples are restricted to sibling or relative pairs. A lack of information means it is sometimes

not possible to infer whether 0, 1 or 2 alleles are shared IBD with certainty when only sibling pairs are genotyped. In some cases, IBD can remain ambiguous even when parental genotypes are available. For example, it is difficult to determine IBD sharing among siblings where one parent is homozygous for a particular allele, or if both parents have the same genotype. When IBD cannot be determined explicitly allele frequencies are employed to give the best estimate of the IBD sharing probabilities. The use of multiple markers provides more accurate IBD information.

The LOD score is a commonly used measure of linkage. It is defined as the $\log_{10}$ of the ratio between two probabilities, L_{HA} and L_{H0}. Where, L_{H0} is the probability of the data under the null hypothesis, i.e. assuming that the there is no excess allele sharing. L_{HA} is the probability of the data under the alternative hypothesis, i.e. given the estimated IBD sharing probabilities. Dependent on the availability of marker genotypes the LOD score can be estimated at any point across the genome. The maximum LOD score (MLS) obtained across a given region provides the best estimate of the gene location. In the presence of a disease gene, the size of the MLS depends on the magnitude of the genetic effect and the proximity of the disease locus to the nearest point of analysis.

In the context of a full genome screen using sibling pairs, linkage signals are declared suggestive when they are expected to occur less than once per genome screen and significant if they are expected to occur less than 0.05 times per screen. Lander and Krugylak (1995) proposed that a MLS >2.2 is indicative of suggestive linkage, whereas a signal can be deemed significant if the MLS exceeds 3.6. However, these criteria are based on a particular number ASPs, genotyped on an infinite grid of markers and are not directly applicable to most studies. However, Lander and Krugylak do suggest that each genome screen should report empirically derived significance criteria based on the sample and data from that particular study.

A number of concerns are inherent in linkage methods. For example, the mis-specification of marker allele frequencies can lead to increased type 1 error rates when using model-free methods (Knapp et al. 1993). Issues of power also plague

linkage studies. For example, to give an 80% chance of identifying a locus with a relative risk to siblings of 1.5 (the estimated effect size of APOE in AD) at the 0.05 significance level a total of 600 sibling pairs would be needed (Myers et al. 2002). Family based samples remain difficult and costly to ascertain, especially for diseases of late-onset. As such many of the linkage studies of AD reported to date have been underpowered.

Heterogeneity and phenocopies present a substantial problem for linkage analysis. *Heterogeneity*, where identical phenotypes arise from different mutations at different loci, is a particular concern as many complex disorders are likely to result from the combination of many genes and environmental factors. This is likely to underpin the lack of, and inconsistent, linkage findings reported in the majority of studies of complex diseases (Altmuller et al. 2005; Bertram and Tanzi 2004; Gillanders et al. 2004). *Phenocopies*, subjects with clinically indistinguishable non-genetic forms of the disease, can lead to broad, ill defined linkage peak regions or increased type 2 error rate, even when using reasonable sample sizes of 500 to 1000 families (Risch 2000). Large sample sizes reduce the impact that locus heterogeneity and phenocopies have on linkage analyses (Schulze and McMahon 2003). However, adequate sample sizes are often not available, and the impact of these factors is still not clear even when large family based samples are ascertained. An alternative approach is to include covariates, which are likely to reduce the impact of locus heterogeneity and phenocopies on linkage analyses (Altmuller et al. 2005; Gillanders et al. 2004; Hamshere et al. 2005).

4.1.3 Methods of covariate linkage analysis

Covariates are generally used to reduce the effect of locus heterogeneity and phenocopies on linkage evidence by identifying homogenous subsets of the disease, or to search for genetic variation which acts as a disease modifier in complex disorders. Numerous methods of incorporating covariates in linkage analyses have been utilised. At this point, it is important to consider the uses of the different methods, and their associated advantages and disadvantages.

The simplest method of incorporating additional phenotypic characteristics into linkage analyses is to perform sample stratification. As such a primary analysis is performed using the whole dataset, followed by a secondary analysis on a subset according to the covariate of interest. For example, this method has been employed to perform linkage analysis of a refined phenotype of AD with psychosis (Bacanu et al. 2002). It has been suggested that this approach will reduce the impact of locus heterogeneity and phenocopies, however methods which rely on sample splitting often lead to substantial reductions in sample size. This is a particular concern as sample size is one of the most important factors for identifying significant linkage evidence (Altmuller et al. 2001). Furthermore, using such an approach makes it difficult to statistically assess the relative difference between primary and secondary analyses, as the sample sizes are likely to differ.

An alternative to methods which rely on sample stratification is to perform ordered subset analysis (Ghosh et al. 2000; Hauser et al. 2004). This involves ordering families in relation to a continuous variable (e.g. age at onset (AAO)) and then sequentially incorporating them into the analysis to identify the subset with the maximum evidence for linkage at a given set of markers. Using this method a prior 'cut point' for continuous variables is not required (e.g. AAO above or below 65), which is an advantage as such cut-offs are usually arbitrary and are not likely to be biologically valid. Also, it allows for detection of linkage peaks in subsets, even in the presence of heterogeneity in the full sample (Scott et al. 2003). Whilst this method is well suited to searching for loci that are strongly linked to disease in a small portion of the covariate distribution, it is less suitable for identifying disease modifying genes for covariates which are defined categorically. In addition, ordered subset analysis is presented with difficulty if the locus is only a risk factor among those in the mid range of a covariate distribution, as is the case with APOE which exerts its largest effect in those between 60 and 70 years of age.

An alternative approach has been suggested by Devlin and Colleagues (2002) who suggested the use of 'mixture models' to incorporate covariates into linkage analysis. They were primarily concerned with locus heterogeneity which presents a problem in the study of complex disease. As such they proposed two methods of assigning ASPs to groups who are either 'linked' or 'unlinked' to a specific region.

The first method utilises cluster analysis, based on phenotypic information, to assign sibling pairs into groups. The other uses estimated IBD information at the region of interest to assign sibling pairs to groups. They provide evidence that this method leads to a considerable increase in power when the covariate carries information about membership to linked and unlinked groups. However, the second method of clustering, based on estimated IBD sharing, was shown to perform badly when data were drawn from affected sibling pairs, as apposed to larger pedigrees, as sibling do not provide adequate IBD information to discriminate between linked and unlinked groups. This method is not suited to studying disease modifying hypotheses as IBD sharing among families is constrained to be greater than, or equal to, 50% (Hamshere et al. 2005).

A number of methods of covariate analysis have been proposed which aim to regress the IBD scores among affected pairs onto covariates (Dorr et al. 1997; Goddard et al. 2001; Greenwood and Bull 1997, 1999; Olson 1999; Rice et al. 1999). The method proposed by Greenwood and Bull (1997), and since modified extended by Olson (1999) and Goddard and colleagues (2001), uses regression methods to assess the dependence of the IBD sharing proportions on covariates. These methods have been shown to increase power to detect linkage, provided that the covariate reflects underlying locus heterogeneity. These methods also allow the inclusion of both categorical and continuous covariates.

An alternative is to treat covariates as quantitative traits. Linkage methods used to localise quantitative trait loci (QTL) focus on the within- and between-family variance of the trait, and is therefore driven more by the similarity, rather than the mean, of covariate values among relative pairs. This provides a distinctly different question to approaches that look at the mean variation in the covariate, such as ordered subset analysis. Indeed, ordered subsets and QTL analysis of the same dataset have been shown to provide contrasting results (Li et al. 2002; Scott et al. 2003). QTL methods are more specifically a method of identifying loci which are directly associated with the variable of interest (e.g. height, Aβ42 levels) rather than a covariate method. As such, QTL methods are less well suited to studying covariates, as comparisons cannot be made between the evidence from QTL analysis and overall linkage (in the absence of covariates).

In general covariate linkage methods can be divided into two categories. First methods which treat the family as the unit of analysis (e.g. QTL methods) and secondly, those that are able to treat the relative pair as the unit of interest (e.g. sample splitting and regression based methods). Compared to methods that focus on relative pairs, those that treat the family as the unit of interest are more likely to be disadvantaged if a number of phenocopies, or locus heterogeneity exist within the family. In general, covariate methods of linkage analysis have been shown to increase power to detect linkage, especially in the presence of heterogeneity. As such, they are likely to provide invaluable insights in the study of complex disorders, indeed, covariate linkage methods have already contributed to the identification of numerous novel linkage regions and also to the identification of susceptibility genes. These findings will be discussed in more detail in section 4.1.5.

4.1.4 Linkage analysis studies of late-onset Alzheimer's disease

The importance of genetic factors in the aetiology of LOAD has been discussed in previous chapters. As such a number of genome screens have been published in an attempt to localise susceptibility genes for AD. The evidence suggests that LOAD, like many other complex diseases, does not follow a simple mode of inheritance. It is therefore difficult to accurately specify a disease model when using parametric forms of linkage analysis, which leads to a decrease in power (Clerget-Darpoux et al. 1986). Linkage studies of LOAD have therefore relied mainly on model-free methods of analysis, which protect against loss of power due to mis-specified model parameters. Owing to the late-onset nature of Alzheimer's disease linkage studies have been largely based on samples of sibling pairs, as parents are very rarely available for genotyping.

In 1997 Pericak-Vance and colleagues (1997) reported the first full LOAD genome screen, using a sample of 54 large multigenerational families where a substantial number of members had late-onset AD. They used parametric and model-free linkage analysis, and divided the sample into a genomic screening set, comprising 16 families, and a follow up sample of 38 families. They identified four regions of interest, on chromosomes 4, 6, 12 and 20, with the strongest evidence for linkage

observed over a 30cM region of chromosome 12 (incorporating markers D12S373, D12S1057, D12S1042 & D12S390, located between 36cM and 67cM). In addition, it appeared that linkage to chromosome 12 was strongest in families where at least one individual did not possess an APOE ε4 allele, suggesting that APOE and the potential chromosome 12 locus might act independently.

Linkage to chromosome 12 by Pericak-Vance and colleagues in 1997 was followed up a year later in two independent studies; however, the findings are not straightforward to interpret. Firstly, in 1998 Rogaeva and colleagues replicated the chromosome 12 linkage in an independent sample, comprising 172 AD patients and 146 non-demented relatives taken from 53 families. They genotyped 6 markers in the region previously implicated on chromosome 12. When the data were analysed, assuming a single homogeneous locus on chromosome 12 caused AD in all 53 pedigrees, no evidence for linkage was found. However, significant evidence for linkage was revealed between D12S358 (26cM) and D12S373 (35cM) and in a region ~50cM distal to this, between D12S1090 (48cM) and D12S96 (68cM) using non-parametric linkage methods. The strongest evidence in this study appeared to originate from families in which at least 75% of affected members carried one or more APOE ε4 alleles.

A further attempt at replication was published by Wu and colleagues in 1998. They used a sample of 230 families collected by the National Institute of Mental Health (NIMH) AD Genetics Consortium. Each family contained at least 2 affected siblings with probable or definite late-onset AD (AAO > 60 years). No evidence for linkage on chromosome 12 was reported in their full sample. However, moderate linkage evidence was observed in a subset of the sibships in which neither pair possessed an APOE ε4 allele, with a maximum multipoint LOD score of 1.91 found at marker D12S98 around 24cM (chromosome wide p-value = 0.09, after correcting for multiple testing).

Pericak-Vance and colleagues followed up their initial findings by sampling more individuals in their original families and genotyping a tighter grid of markers in their region of interest (Scott et al. 2000). A total of 26 microsatellite markers were genotyped, located between 24cM and 152cM on chromosome 12. Weightings for

clinical and neuropathological factors, along with sibship size and APOE genotyped were also included to address issues of heterogeneity. Moderate evidence for linkage was observed at 9 of the 14 markers genotyped in the initial region of interest. Evidence for linkage was strongest in families who had at least one member with no APOE ε4 alleles. However, this is contradictory to their previous report in which linkage evidence was strongest in families where at least 75% of affected members carried one or more APOE ε4 alleles.

These preliminary efforts highlight the difficulties in identifying further novel risk loci for complex disorders, like LOAD. In review Craddock and Lendon (1998) concluded that these three studies offered encouraging evidence of the existence of a novel susceptibility gene, or genes, in the chromosome 12 region. These studies also highlighted the need for further well powered linkage studies in independent samples.

Using largely the same sample as in their earlier chromosome 12 paper (Wu et al. 1998a), Kehoe and colleagues reported a full genome screen in 1999. They employed a model-free method of analysis, incorporating 237 microsatellite markers separated by an average distance of 16.3cM. In either the full sample, or sub samples characterised by the absence or presence of ε4 alleles, 16 peaks with a maximum multipoint LOD score ≥ 1 were reported, which exceeded the number that would have been expected by chance. Only peaks on chromosome 1, 5, 9, 10 and 19 gave a MLS ≥ 1 in the full sample. No evidence for linkage was found in the region implicated by Pericak-Vance and colleagues in this study, although some evidence for linkage was found on 12p, at around 24cM. This was also the first published report of suggestive linkage to chromosome 10 (MLS=2.27).

Further to the initial findings (Kehoe et al. 1999) three articles published in *'Science'* in December 2000 reported significant linkage to Alzheimer's disease, and a related phenotype, on chromosome 10. In an extension to the Kehoe and colleagues genome screen Myers and colleagues (2000) reported significant linkage on chromosome 10 using a widely overlapping, but extended sample comprising 168 additional affected sibling pairs, genotyped on a tighter grid of

markers. The combined sample included 429 sibling pairs diagnosed with probable or definite AD, all with an age at disease onset ≥ 65 years. Fifteen additional markers on chromosome 10 were genotyped in all sibling pairs, giving a total of 26 microsatellite markers with an average interval of 5cM. Non-parametric multipoint linkage analysis yielded a peak LOD score of 3.83 at 82cM (between D10S1227 and D10S1225). The MLS was higher in the subset of the sample where both siblings had at least one APOE ε4 allele, compared to sibships in which neither member had an APOE ε4 allele. This is a likely reflection of the increased number of APOE ε4 positive sibling pairs in their sample. Allele sharing was similarly elevated in both groups, suggesting that stratification by APOE ε4 did not change the proportion of allelic sharing at the linkage peak. Furthermore, Holmans and colleagues (2005) have since reported covariate linkage using this sample and did not report a significant effect of APOE on chromosome 10.

A further report of linkage was reported in December 2000 by Bertram and colleagues (2000). They presented genetic linkage analysis of 7 markers on chromosome 10 in 435 multiplex AD families, in which all affected members had an AAO equal to or greater than 50 years. They performed parametric and model-free analysis, stratified by AAO and APOE genotype. They found significant evidence for linkage at a locus ~40cM distal to that reported previously (Ertekin-Taner et al. 2000; Kehoe et al. 1999; Myers et al. 2000). Significant evidence for linkage was found around marker D10S583 (115cM) under a dominant and recessive model in the full sample, and around marker D10S1671 (124cM) when restricting analysis just to those with late-onset disease. Model-free linkage results yielded similar results, with the strongest signals for markers D10S583, D10S1671 and D10S1710 in the late-onset families, with a peak linkage score at 124cM (D10S1710). However, none of the analyses yielded significant findings for marker D10S1225 (located around 80cM) that had previously been linked with LOAD (Ertekin-Taner et al. 2000; Kehoe et al. 1999; Myers et al. 2000).

Additional support for a chromosome 10 locus was reported in a study using plasma amyloid β42 peptide (Aβ42) levels as a surrogate marker for AD (Ertekin-Taner et al. 2000). Autosomal dominant mutations that cause early-onset familial AD all increase Aβ42 in plasma and the brain (Cai et al. 1993; Citron et al. 1992;

Duff et al. 1996). It is also found to be elevated in unaffected family members of patients with late-onset AD, whilst heritability estimates for Aβ42 are as high as 73% (Ertekin-Taner et al. 2001). Aβ42, therefore, appears to be a good biological marker for AD, and was used in linkage analysis to follow up previously reported linkage regions for late-onset AD on chromosomes 1, 5, 9, 10 and 19 (Kehoe et al. 1999). Using this method a MLS of 3.93 was found on chromosome 10 around 81cm, between markers D10S1227 and D10S1211. This finding coincides almost directly with the chromosome 10 linkage peak reported previously by Myers and colleagues (2000). Maximum multipoint LOD scores in all other genotyped regions were less than 0.5. It should be noted that the study was restricted to 10 families, focusing on five which had an AD proband with extremely high plasma Aβ42 or Aβ40. Linkage evidence was less striking in the unstratified sample, with a maximum multipoint LOD score of 1.82. Therefore, the results cannot be used to evaluate the contribution of the chromosome 10 locus to the AD population in general.

It should be noted that there is a considerable overlap between the samples used by Myers and colleagues and Bertram and colleagues. A total of 188 families were included in both analyses. To assess the impact of this overlap Bertram and colleagues performed separate analyses for families included in both studies and families independent to their investigations. Linkage evidence was most pronounced at three of the seven markers in those families who were present in the analysis by Myers and colleagues. However, in their analysis of the 'Myers families' they did not find linkage to D10S1225. They postulated that this discrepancy could have arisen from a number of factors, including sampling issues, inclusion of genetic information from all relative pairs (as opposed to just siblings) or differences in analytic methods. However, in their attempt to confirm the previously reported linkage peak, Bertram and colleagues genotyped marker D10S1225, leaving an interval of ~30cM to the next genotyped marker. Although, this marker was the peak multipoint marker reported in previous studies by Myers and colleagues and Erekin-Taner and colleagues, the *twopoint* LOD scores in those studies were only 1.54 and 0.57, respectively. The elevated multipoint score in the Myers and colleagues analysis appears to be driven largely by markers within 10cM of D10S1225, of which two had twopoint LOD scores of 2.12

(D10S1220) and 4.85 (D10S1211). A similar interpretation of the Ertekin-Taner and colleagues data is also appropriate. It would therefore seem reasonable that the analysis by Bertram and colleagues was inadequately placed to attempt replication of the proximal linkage peak and it remains unclear whether the two reported chromosome 10 peaks are independent. It is also unfortunate that Bertram and colleagues did not report the twopoint model-free analysis of the 'Myers' families. It remains a possibility these three linkage signals on chromosome 10 result from a single susceptibility gene, indeed varying results from linkage analyses were similarly reported on chromosome 19 prior to the identification of APOE as a risk factor for LOAD (Liu et al. 1996).

Extensions to these three studies have since been reported to cover the whole genome or include additional samples and marker genotypes (Blacker et al. 2003; Myers et al. 2002; Pericak-Vance et al. 2000). Myers and colleagues (2002) (Myers et al. 2002) reported stage 2 of a genome screen following up their prior stage 1 analyses (Kehoe et al. 1999). In the second stage 451 affected sib pairs were genotyped at a total of 327 microsatellite markers, incorporating an additional 91 markers mainly located within the 16 regions with a MLS ≥ 1 in stage 1. Ten of their stage 1 regions maintained LOD scores in excess of 1. Aside from the chromosome 10 finding already reported (Myers et al. 2000) minor linkage peaks were observed on chromosomes 5, 9 (2 peaks), 12, 19, and X in the full sample, with additional loci of interest on chromosomes 1, 6 and 21 when stratification by APOE ε4 was performed. However, other than the chromosome 10 finding none of these linkage regions met published criteria for suggestive linkage (Lander and Kruglyak 1995).

Blacker and colleagues (2003) reported a follow up to the linkage evidence reported by Bertram and colleagues (2000). They used the AD sample collected as part of the National Institute of Mental Health AD genetics initiative, which has been widely used in other linkage studies. Therefore, this study contained a sample overlap of up to 60% with other published genome screens (Bertram et al. 2000; Ertekin-Taner et al. 2000; Kehoe et al. 1999; Li et al. 2002; Myers et al. 2000; Myers et al. 2002; Olson et al. 2001; Pericak-Vance et al. 2000). However, sample selection and analysis method vary between studies. The sample used by

Blacker and colleagues consisted of 437 families. An AAO cut-off of 50 years was adopted, however analysis was also performed on a sub sample of 320 families where all affected members had an AAO ≥ 65 years.

They identified one 'highly significant' linkage region located in close proximity to the APOE locus on chromosome 19, along with 10 regions that met criteria for suggestive linkage. Including, 1q23, 3p26, 4q32, 6p21, 6q27, 9q22, 10q24, 14q22, 15q26 and 21q22. Only regions on chromosome 1, 9 and 10 showed evidence of suggestive linkage in the sub-stratum of families segregating late-onset AD. Linkage regions on chromosome 3, 14, 15 were restricted to the subset of 117 families which contained affected individuals with an AAO < 65 years. The other regions were identified in the unstratified sample.

Interestingly, Blacker and colleagues also reported a secondary analysis performed in families not used in previous reports of linkage. Most regions of suggestive evidence identified in the full dataset remained in the sub sample, whilst LOD score increases were noted on chromosome 3, 6 and 14. No evidence for linkage was observed on either chromosome 9, 10 or 12 in the independent families.

It should be noted that 65 cases with possible AD were included in the analysis. This was controlled for to some extent in parametric analyses by allowing for a phenocopy rate of 20%. However, such concessions are not attributable to multipoint model-free analyses, which is a particular concern as linkage evidence is known to be sensitive to mis-specification of the phenotype (Whittemore and Halpern 2001). It is also noteworthy that Blacker and colleagues used genetic maps and inter-marker distances provided by CIDR (www.cidr.jhmi.edu) which in some cases differ for those from the Marshfield Centre for Medical Genetics (http://research.marshfieldclinic.org/genetics) which have been used in the majority of other genome screens to date. It is currently unclear which map is most accurate, however parametric/model-free analyses are not generally robust to changes in map location, hence, this could account for differing results between studies (Daw et al. 2000). The final point of note is that the Blacker and colleagues study contained 26 non-Caucasian families, which could lead to problems

associated with population stratification. However, they did report that analyses excluding these families did not differ notably.

Pericak-Vance and colleagues (2000) reported a follow up to their earlier genome screen (Pericak-Vance et al. 1997) performing a genome scan using a sample of 466 families with late-onset AD (including 739 sibling pairs) genotyped on a 10cM grid of markers. A model-free affected sibling pair (ASP) method was employed, along with a complimentary approach using parametric analyses incorporating affected relative pairs. The analysis was also stratified into a subset of families that contained at least one affected member whose AD diagnosis had been confirmed by autopsy. They found interesting results on chromosomes 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, and 19, with chromosome 9 showing the highest MLS of 2.97, rising to 4.3 in the sub-sample of autopsy confirmed affected individuals. The APOE locus on chromosome 19 showed significant linkage, as did a region at around 125cM on chromosome 7 with a peak LOD of 1.97. Chromosome 10 only demonstrated a LOD score >1 in twopoint parametric analysis of autopsy confirmed cases. The peak chromosome 10 marker mapped closer to the peak identified by Myers and colleagues than it did to the one identified by Bertrem and colleagues. The previous finding on chromosome 12 was not replicated with a MLS < 1 over the whole region. As with many full genome screens it is likely many of the interesting regions represent false positives, as a number of tests are applied to the data and testing sub samples, without correction for multiple testing, ultimately leads to an increased type 1 error rate. It should also be noted that the study by Pericak-Vance and colleagues (2000) used families with a mean AAO ≥ 60 years. Therefore, it remains possible that a number of affected individuals had an AAO < 60 years, which are commonly excluded from studies of LOAD and may have a different disease aetiology. However, the number of individuals with an AAO < 60 years was not reported.

The studies reported by Pericak Vance and colleagues, Myers and colleagues and Blacker and colleagues remain the largest genome screens reported for late-onset AD. However, there is considerable sampling overlap between the studies which makes the comparison of results difficult. All three analyses have sampled different families collected as part of the NIMH AD Genetics Initiative, with only a

proportion of families selected being unique to each investigation. The studies by Pericak-Vance and colleagues and Myers and colleagues both use DNA from AD families taken from the Indiana Alzheimer Disease Centre National Cell Repository. In addition, the Myers and colleagues study contains a unique sample of 94 affected sibling pairs collected in the United Kingdom. Whereas, Pericak-Vance and colleagues used data from 216 affected and 164 unaffected family members, taken from 62 families recruited as part of the Collaborative Alzheimer Project (CAP), which are exclusive to their study. Furthermore, different, but overlapping, microsatellite markers have been used in these analyses.

The most consistent signals across samples are located on chromosomes 9 and 10. Myers and colleagues and Pericak-Vance and colleagues both report linkage to marker D9S741, with MLS of 1.8 and 2.97 (raising to 4.31 in autopsy confirmed AD samples), respectively. Blacker and colleagues report linkage to a region 12cM distal with a peak MLS of 1.3, however this reduces to 0.7 in families not used in the Myers and colleagues analysis. Linkage to chromosome 10 is reported at 59cM (Pericak-Vance et al. 2000), 82cM (Myers et al. 2002) and 135cM (Blacker et al. 2003). It would seem unlikely that such a broad linkage region results from the presence of one susceptibility gene. Pericak-Vance and Blacker both report minor evidence for linkage around chromosome 4q32 and 6q26-27. The locus on chromosome 6 also showed linkage evidence in the Blacker and colleagues full sample and in the substratum of families independent to their analyses. Linkage to this region has also been demonstrated in a more recent reanalysis of the NIMH dataset (Olson et al. 2002). Comparisons between the studies are simplified as Blacker and colleagues present data arising just from families unique to their study. Comparing the results of the genome screens reported by Myers and colleagues with the analyses of the families independent to the study reported by Blacker and colleagues would suggest that concordant linkage peaks between the two studies on chromosome 1, 5, 6, 9 10 and 21 are driven by overlapping samples.

Given the interesting, but largely conflicting, results from these analyses it is clear that appropriately powered replication studies are required in independent samples. To date, the only independent genome screens have come from two

non-Caucasian family based samples (Farrer et al. 2003; Lee et al. 2004) and one non-family based sample collected in Finland (Hiltunen et al. 2001). Hiltunen and colleagues (2001) performed a genome screen using linkage disequilibrium mapping in a Finnish sample of 47 AD cases and 41 age matched controls. Their sample consisted of individuals from a geographically isolated region of Finland, where all members have descended from a small group of original founders, which makes the use of LD mapping more appropriate. Twenty-one of the 366 polymorphic microsatellite markers across the whole genome showed association with AD (6 of these showed association at a more stringent significance level of 0.01). Additional genotyping of markers flanking these LD regions identified seven chromosomal loci that contained more than one marker associated with AD. Strong evidence was found on chromosome 6, around marker D6S1017 (~63cM), in a region which has been implicated by others (Blacker et al. 2003; Lee et al. 2004; Myers et al. 2002). Two markers on chromosome 10 were found to be associated with AD, however, the LD locus was located ~40cM to 75cM from the previously reported linkage evidence, suggesting that these two loci are independent (Blacker et al. 2003; Ertekin-Taner et al. 2000; Lee et al. 2004; Myers et al. 2002). It is interesting that microsatellite markers located 2.5cM and 1cM from the APOE locus did not show significant association with AD. This could have resulted from a lack of power, or alternatively may reflect the evolutionary ancient nature of the APOE ε4 allele, which questions the ability of LD mapping in this sample to detect other genetic mutations associated with AD which are of similar origin to APOE.

A further genome screen using LD mapping was performed by Farrer and colleagues in 2003, using a genetically isolated Israeli-Arab population. Particular high prevalence rates of AD have been noted in the Wadi Ara community with an overall prevalence rate of 20.5% in those over 60 years of age, raising to over 50% in those aged 80 years or older. Farrer and colleagues describe familial clustering of AD in this population and hypothesise that the increased risk of AD is due to genetic factors. To search for susceptibility genes in this sample they performed a LD screen between markers spread throughout the genome and AD. A screening set of 5 cases and controls from families with the highest AD prevalence was genotyped across 375 markers spread throughout the genome.

Markers showing significant difference in allelic frequency between cases and controls were genotyped in an enlarged sample of 100 cases and 100 controls. Significant allelic association was noted on chromosomes 2, 9 and 10. The region showing association on chromosome 2 was located at ~40cM and had not previously been reported in linkage studies, however this region has been implicated in the Finnish linkage disequilibrium screen (Hiltunen et al. 2001). Significant association was noted with 4 markers on chromosome 9, notably alleles with sizes of 244bp or greater were observed in 16.5% of AD cases and in none of the controls, resulting in a highly significant association ($p < 0.001$). This finding is of particular interest as this region has also been implicated in all three previous genome screens of AD using conventional linkage methods (Blacker et al. 2003; Myers et al. 2002; Pericak-Vance et al. 2000). To follow up previous reports of linkage Farrer and colleagues also performed intensive analyses of chromosomes 10 and 12. They reported significant evidence for allelic association at ~59cM and ~80cM on chromosome 10, which could represent a replication of the previous findings (Ertekin-Taner et al. 2000; Myers et al. 2000; Pericak-Vance et al. 2000), however the most consistent region of allelic association on chromosome 10 was observed in a more distal position between 105cM and 115cM. Four contiguous markers showed evidence of allelic and genotypic association in this region, which maps closely to the region of linkage reported by Bertram and colleagues (Bertram et al. 2000). Six markers were associated with AD on chromosome 12, with the strongest signal observed at ~83cM around the linkage region identified by Rogeava and colleagues (1998).

It is interesting that a number of the regions identified in the Israeli-Arab community sample have previously been implicated in AD, largely in samples of Caucasian origin, suggesting that they may be authentic loci rather than chance findings. It should be noted that the sample used was non-Caucasian and largely inbred. It is therefore unclear how generalisable these results are to other populations. However, previous studies have found success in mapping genes responsible for disease susceptibility in Caucasian outbred populations by using genetically isolated samples (Baldwin et al. 1995; Frydman et al. 1985).

A particular concern when assessing the study by Farrer and colleagues arises from the relatively small screening sample of 5 cases and controls. Such an approach is likely to lack power, leading to an increase in false negatives. This is illustrated in their chromosome 12 analyses. During their initial screen this region did not warrant further follow up, however due to previous reports of linkage and association in this region a number of markers were genotyped in their larger sample set, with several showing significant allelic association. Results from these analyses are therefore only useful in identifying or replicating regions of interest, rather than excluding regions from further study.

A further linkage study of late-onset AD has been reported, using a family based sample of Caribbean-Hispanic descent (Lee et al. 2004). The strongest evidence for linkage was observed on 18q, 10q and 12p in the first stage of their analyses. These findings were followed up through fine mapping using a slightly extended sample of 527 individuals from 104 families. Maximum multipoint LOD scores of 3.15 and 2.70 were obtained in their stage 1 analyses, on chromosome 10 and 18 respectively. The region on chromosome 10 was located around 142cM which is somewhat distal to those reported previously, with the closest region of previously identified linkage being ~20cM proximal (Bertram et al. 2000; Li et al. 2002). In their stage 2 analyses the LOD scores on chromosome 18 increased to 3.65, whereas the linkage evidence on chromosome 10 was reduced somewhat, with a maximum multipoint LOD score of 2.02. Suggestive evidence for linkage in this sample was also observed on chromosome 12 in close proximity to the region identified in previous reports (Kehoe et al. 1999; Pericak-Vance et al. 1997)

Reviewing the linkage findings to date provides an illustration of the difficulties in replicating linkage and in estimating the location of genes underlying complex disorders. Comparing studies reveals a total of 16 regions on 11 chromosomes that yield positive signals ($p < 0.01$) across at least two studies with markers no further than 25Mb apart. The strongest evidence seems to be around 5p13-15, 9p21, 9q22, 10q21-22, 12p11-12 and the APOE locus located around 19q13. Perhaps the strongest region of linkage reported to date is on chromosome 10. A number of studies, based on different but related phenotypes (e.g. familial AD (Bertram et al. 2000; Blacker et al. 2003; Farrer et al. 2003; Lee et al. 2004; Myers

et al. 2000), AAO in AD (Li et al. 2002) and plasma amyloid β42 peptide (Ertekin-Taner et al. 2000)) have reported evidence for linkage to chromosome 10q. The peak regions in these studies stretch from ~81cM to ~138cM. Three studies find strong evidence for linkage around 81cM (Ertekin-Taner et al. 2000; Kehoe et al. 1999; Myers et al. 2000), whereas others have reported linkage to a more distal region between 115cM and 138cM (Bertram et al. 2000; Blacker et al. 2003; Farrer et al. 2003; Li et al. 2002). Methodological or sampling differences across studies could be responsible for discrepant localisation of the two linkage peaks which could all be caused by the same gene (Bertram and Tanzi 2004). Alternatively, more than one loci could be present, which could either influence disease risk or act as a modifier for AAO (Li et al. 2002). At present it is impossible to distinguish between these two possibilities, but it seems likely that at least one major late-onset AD locus resides on chromosome 10q.

Like studies of many complex disorders, issues of power are a persistent problem for linkage studies of late-onset AD. Family based samples remain difficult and costly to ascertain, especially for diseases of late-onset, meaning that many of the linkage studies of AD reported to date have been underpowered. However, it is encouraging that the majority of linkage and association genome screens to date have found evidence for a susceptibility gene near the APOE locus (Blacker et al. 2003; Curtis et al. 2001; Li et al. 2002; Myers et al. 2002; Pericak-Vance et al. 2000; Zubenko et al. 1998). Further linkage and association studies are clearly required to unravel the complex issues surrounding the genetics of late-onset AD. Linkage studies to date have largely served to highlight that AD is likely to be characterised by substantial locus heterogeneity. As discussed earlier, covariate analysis can account for locus heterogeneity and aid in the identification of novel linkage signals. Hence, recent years have seen an increase in studies which have aimed to include aspects of clinical variation as covariates in linkage analysis.

4.1.5 Covariate linkage analysis studies of late-onset Alzheimer's disease

The inclusion of covariates, as briefly discussed in section 4.1.3, has two distinct advantages over standard methods of linkage analysis, both of which are pertinent

to this thesis. First, inclusion of covariates may facilitate the identification of regions which contain disease modifying genes. Second, covariate linkage analysis allows for locus heterogeneity owing to the covariates, consequently genes that are risk factors for certain 'sub-phenotypes' of AD are more likely to be localised (e.g. genes which cause AD in the very old). It is possible that the clinical heterogeneity observed in AD could be responsible for differing linkage and association results observed between studies.

The most commonly used covariate at present is AAO. It is hoped that if genetic mechanisms that influence AAO can be identified it may be possible to modify their processes to prevent or delay disease onset, which is of particular importance for late-onset disorders such as Alzheimer's disease. As found in section 3.3.2, the APOE ε4 allele has been shown to reduce AAO in a dose-dependent manner (Meyer et al. 1998), however there is substantial evidence which suggests that further, unidentified, genes are associated with AAO in AD (Li et al. 2002; Tunstall et al. 2000; Warwick Daw et al. 2000).

Olson and colleagues (2001) reported linkage to chromosome 21 when incorporating AAO as a covariate. Their sample comprised 252 ASPs with a disease onset ≥ 60 years collected as part of the NIMH AD genetic initiative, which has previously been used in genome screens of AD (Bertram et al. 2000; Blacker et al. 2003; Kehoe et al. 1999; Myers et al. 2000; Myers et al. 2002; Pericak-Vance et al. 2000). Sibling pairs with a higher mean AAO and current age showed increased evidence for linkage around the region containing the amyloid precursor protein (APP) gene. This is of particular interest as APP is known to confer high risk of AD to a number of families afflicted with early onset AD (Goate et al. 1991). Olson and colleagues (2002) have since extended this analysis, presenting findings from a whole genome scan on the same dataset, using the same covariates. In addition to the previously reported linkage evidence on chromosome 21 they reported significant AAO effects on chromosome 14 (increase in LOD of 1.89, p=0.003). Significant APOE effects were reported on chromosome 9 (increase in LOD of 1.52, p=0.008) and 14 (increase in LOD of 1.62, p=0.006), whilst current age had the largest effects on chromosome 20 (increase in LOD of 2.69, $p < 0.001$). Analyses of 'current age' and disease duration should be

interpreted with caution. Current age reflected age at interview in 33% of the cases and age at death in 67% of the sample. Hence, disease duration and current age in a proportion of individuals were reliant on when they were ascertained, and it makes little sense to suggest that this could be determined by genetic/biological factors. Also, cause of death was not taken into account; it therefore remains possible that participants may have died of a condition unrelated to their dementing illness, which questions the use of duration from disease onset to death as a surrogate marker for rate of disease progression.

In a slightly overlapping sample, Li and colleagues (2002) treat AAO as a QTL, using a sample of 449 and 174 families affected with AD and Parkinson's disease (PD), respectively. When looking at AD alone they found interesting regions on chromosomes 4, 6, 8, 10 (at 139cM, close to marker D10S1237), 13 & 17. Interestingly, linkage (MLS=1.55) was also found with AAO in Parkinson's disease at 132cM on chromosome 10. To further investigate this common region they combined the AD and PD samples, considering both disorders as a common disease in one model, and then considering them separately to distinguish the contribution of AAO between AD and PD. They found a peak MLS of 2.62 when both diseases were considered together at 133cM (between markers D10S1239 and D10S1237), indicating that a common modifier for AAO in AD and PD may be located on chromosome 10. The peak on chromosome 10 is distal to the more commonly observed linkage peak located around 80cM (Ertekin-Taner et al. 2000; Kehoe et al. 1999; Myers et al. 2000), but is in close proximity to the linkage region at ~130cM (Bertram et al. 2000; Blacker et al. 2003; Farrer et al. 2003)

It is noteworthy that AAO in families has been characterised in different ways. Olson and colleagues analysed the relationship between alleles shared IBD in ASPs and the *total* AAO in each sibling pair. In contrast, the variance components analysis by Li and colleagues focuses on the relationship between alleles shared IBD and the *similarity* of AAO. These are different questions, both with biological relevance. Increases in linkage evidence observed when using AAO mean as a covariate could indicate that a gene in the region contributes to AD in a subset of the population within a particular AAO range. In contrast, it is possible that susceptibility genes give rise to differing levels of disease risk with age, in this

case increased IBD sharing would be observed in sibling pairs with a closer AAO. It is also likely that sibling pairs who display more clinically distinguishing factors, for example, difference in AAO, represent phenocopies. Use of AAO similarity as a covariate is likely to reduce the impact of such phenocopies on analyses.

Holmans and colleagues (2005) used the difference between, and mean, AAO in sibling pairs as covariates in a single analysis. A novel measure of rate of decline (ROD) (mean and difference) was also considered. The analysis was based on the sample and genotypes used in the genome scan published by Myers and colleagues (2002). As such, there was some overlap with the samples used for previous genome screens of AAO in AD (Li et al. 2002; Olson et al. 2002), however a significant proportion of sibling pairs collected in the UK and the Indiana Alzheimer's Disease Centre (IADC), along with additional marker genotypes, were unique to their analysis.

The covariate method employed by Holmans and colleagues tested the effect of quantitative variables by modelling the IBD sharing probabilities in relation to covariates using regression. The most significant finding was on chromosome 21 in the NIMH sample, where elevated IBD sharing was observed among those with increasing AAO, supporting the findings of Olson and colleagues. However this did not replicate in the additional non-NIMH sibling pairs. A significant increase in LOD with AAO difference was observed on chromosome 12. As in previous analyses (Kehoe et al. 1999) increased IBD sharing was also observed on chromosome 12 in sibling pairs who did not possess an APOE ε4 allele, however the AAO effect was independent of pairwise APOE ε4 status. This effect was driven by sibling pairs from the non-NIMH sample; however a small, insignificant, increase in LOD (from 0 to 1.14) was noted when analysis was restricted to NIMH sibling pairs. A smaller increase in LOD was also reported on chromosome 12 using ROD difference in the combined NIMH+UK sample, with a univariate LOD score of 0.36 being increased to 1.51.

Genome screens which have incorporated AAO as a covariate have generally provided rather inconclusive results. Strong evidence for linkage has been observed on chromosome 21 in those with a later AAO, which could implicate APP

as a susceptibility gene for very late-onset AD (≥80 years). However, this finding failed to replicate in a completely independent sample (Holmans et al. 2005). The use of AAO as a covariate has provided further insight into the previously reported linkage region on chromosome 10, suggesting that the distal peak observed between 130 and 140cM may be related to age at disease onset. Methodological issues across studies also make comparisons difficult. For example, Li and colleagues and Olson and colleagues treat AAO in sibling pairs differently, asking distinct biological questions. Furthermore, Li and colleagues used a method of analysis which did not allow a test for overall linkage to be performed (e.g. to test for linkage to disease risk). It therefore remains possible that increased IBD sharing in their analysis is related to overall disease risk, independent of AAO. Indeed, despite finding evidence for overall linkage, Holmans and colleagues reported that AAO did not contribute to linkage evidence on chromosome 10.

Holmans and colleagues also incorporated a crude measure of rate of decline in their analysis. The most significant effect was found on chromosome 9 (at 103cM) when using mean ROD in the NIMH sample, with a univariate LOD score of 1.76 being increased to 3.58. IBD sharing was greater in those showing an increased ROD, an effect that remained significant even after allowing for APOE. However, this finding failed to replicate in a sample of sibling pairs collected in the UK. Mean ROD was also associated with other suggestive findings on chromosomes 1 (UK sample), 2 (NIMH sample) and 8 (NIMH sample). It should be noted that the ROD measure used by Holmans and colleagues was only based on two time points, which, whilst giving a good general picture of ROD may in some cases be inaccurate. In addition, a number of regions where suggestive linkage was found in the NIMH sample were not genotyped in the non-NIMH families, precluding replication.

Two genome screens have been published to date which have used the presence of psychotic symptoms in LOAD as a covariate (Avramopoulos et al. 2005; Bacanu et al. 2002). Sample stratification was employed by Bacanu and colleagues (2002) to perform linkage analysis on a subset of sibling pairs concordant for AD and presence of psychotic symptoms (AD+P). These data were further partitioned into sibling pairs where both members possessed at least one APOE ε4 allele

(AD+P+ε4). Significant linkage was observed on chromosome 2, along with two suggestive signals on chromosome 6 and 21. The significant and suggestive linkage signals, on chromosome 2 and 21 respectively, were found in the subset of siblings who were concordant for AD+P+ε4. In the subset of siblings sharing solely AD+P only the chromosome 6 finding met criteria for suggestive linkage. By restricting their analysis to those sibling pairs who were concordant for AD diagnosis, presence of psychotic symptoms and possessing at least one APOE ε4 allele, it was difficult to assess the individual contribution of AD, psychosis and APOE ε4 genotype. In addition, the method of sample stratification limited the sample size to 140 individuals taken from 65 families, resulting in a substantial loss of information.

Avramopoulos and colleagues (2005) also used a subset of the sample collected as part of the NIMH AD Genetics Initiative. Using a covariate method of linkage analysis they found linkage to chromosome 14 in sibling pairs who were non psychotic. However, they included families with early onset AD (AAO < 65 years) and subsequent analysis indicated that the finding on chromosome 14 was being primarily driven by families with at least one member with an earlier AAO.

The studies outlined above represent the only analyses of late-onset AD to date which have incorporated phenotypic covariates. Without replication is it difficult to draw conclusions from these studies, however they have been successful in identifying novel linkage regions. The inclusion of covariates leads to issues of multiple testing, which undoubtedly increases the likelihood that a number of these regions are false positives. Further analyses of these regions will provide more information about the value of covariate linkage studies in AD. Results obtained from covariate analyses of other disorders are encouraging. For example, covariate linkage analyses have facilitated the successful mapping of susceptibility genes in a number of other disorders, including asthma (Van Eerdewegh et al. 2002) and Crohn's disease (Rioux et al. 2001). The use of covariates has also been instrumental in refining, or amplifying, existing linkage peaks and identifying novel linkage signals in other disorders, including studies of autism (Shao et al. 2003), schizophrenia (Hallmayer et al. 2005) and prostate cancer (Gillanders et

al. 2004; Goddard et al. 2001). Promising results in other areas provide adequate justification for the use of similar approaches in the study of late-onset AD.

4.1.6 Study Design and aims

Aspects of clinical heterogeneity (e.g. AAO) have previously been used in AD, and other disorders, to define suitable covariates for genetic linkage analysis. Clinical variation which shows familial aggregation, or genetic heritability, is most likely to facilitate mapping of genes which either modify disease presentation or lead to genetically homogenous sub-phenotypes. In chapter 3, age at disease onset, psychosis, minor depression and aggression were found to aggregate within families. Familial clustering of AAO is consistent with previous findings (Li et al. 2002; Tunstall et al. 2000), and supports the hypothesis that further genetic loci which modify AAO remain to be identified (Warwick Daw et al. 2000). The findings of familial aggregation of psychosis and aggression are also consistent with previous findings (Bacanu et al. 2005; Sweet et al. 2002; Tunstall et al. 2000). Given the available data it is not possible to determine if familial aggregation results from genetic factors, or shared environment. However, it is reasonable to suggest that these symptoms are likely to be genetically modified. The findings from the previous chapter regarding the mood component were less reliable; however familial aggregation of minor depression was shown in two independent samples. As such this chapter will present linkage analysis using a regression based method to include covariates representing AAO, psychosis, aggression and minor depression.

The method of covariate linkage analysis employed in this chapter is particularly useful as it allows the statistical contribution of covariates and APOE to be assessed independently. It also incorporates information from relative pairs who have not displayed the symptom of interest (e.g. psychosis). Furthermore, it allows various hypotheses to be explored. For dichotomous traits, such as psychosis, individuals are coded as '+' if they had displayed the symptom, and '-' if it was absent. Therefore, pairwise trait categorisations are restricted to -/-, -/+ and +/+ in relative pairs. Given a linkage signal the pattern of IBD sharing among pairs can therefore inform the likely role of the loci. For example, in the presence of a

disease modifying locus, which does not confer risk to AD as a whole, increased IBD sharing would be expect in -/- and +/+ relative pairs, with reduced sharing in -/+ pairs. However, if the loci increased risk to a disease sub-phenotype, characterised by the symptoms of interest (e.g. AD with psychosis), elevated IBD sharing would be expected in +/+ relative pairs. Whereas -/+ and -/- pairs would be expected to show IBD sharing probabilities similar to those expected in the in the absence of linkage. For AAO, a continuous trait, it is possible to characterise affected relative pairs according to the pairwise mean AAO, and difference in AAO between members of each pair. Linkage observed with mean AAO implies the presence of a gene which increases risk of AD within a particular age range. Alternatively linkage observed with AAO difference would imply the presence of a locus which exhibits different levels of risk according to age. Therefore, those with a closer AAO would be likely to share the same genotype and therefore show evidence of linkage.

An extended dataset consisting of 513 affected relative pairs (ARPs) from the NIMH, UK and NIA samples introduced in section 3.2.1 will be used. All of these samples have data available regarding AAO and APOE genotype, whilst both the NIMH and UK samples are well characterised in terms of behavioural symptoms observed in AD. Genotypes from genome screens presented by Myers and colleagues (2002) and Blacker and colleagues (2003) will be combined providing genome wide coverage, with an average marker spacing of 5.5cM. As such this represents one of the largest samples used to date to perform linkage analysis of LOAD.

In summary, this chapter is primarily concerned with the following aim:

- To perform linkage analysis of late-onset AD, incorporating the following covariates, whilst controlling for the confounding effect of APOE:
    - Mean AAO among affected relative pairs
    - AAO difference between affected relative pairs
    - Psychotic symptoms
    - Symptoms of minor depression
    - Aggressive symptoms

4.2 Methodology

4.2.1 Sample description.

Data presented here originate from the NIMH, NIA and UK family based samples described in chapter 3. A more detailed description of these samples can be found in section 3.2.1. Families were selected based on the following criteria: 2 or more *relatives* diagnosed with probable or definite AD according to NINCDS-ADRDA diagnostic criteria (McKhann et al. 1984) with onset ages greater than or equal to 65 years. AAO was defined as the age at which the first symptoms of AD were observed. To reduce potential genetic heterogeneity caused by ethnic origin, only Caucasian families were selected. Analysis was restricted to affected relatives pairs (ARPs) who were genotyped for linkage.

Genotypes came from two overlapping genome screens of late-onset AD published by Myers and colleagues (2002) and Blacker and colleagues (2003). A total of 651 individuals were selected from 295 families from the NIMH sample. 160 individuals from 75 families collected in the UK had genotypes available for linkage and were selected for this analysis, whilst 136 individuals from 64 families were selected from the NIA sample. The total sample comprised 513 ARPs. The sample were predominantly sibling pairs, but also included 3 half siblings, 22 pairs of cousins and 12 pairs of avuncular relationship. Eleven parent-offspring pairs were excluded from the analysis as they are uninformative for linkage (i.e. they will always share 1 allele IBD).

4.2.2 Genotyping

Genotypes from the Kehoe and colleagues (1999), Myers and colleagues (2002) and Blacker and colleagues (2003) analyses were combined. A total of 610 markers were included, with an average spacing of 5.5cM. These comprised 237 marker loci with an average spacing of 15cM genotyped by Kehoe and colleagues. Myers and colleagues followed up the 16 regions giving a maximum LOD score > 1 by genotyping a further 91 marker loci in the UK, NIA and NIMH sample. As such the UK and NIA samples are only genotyped on 11 autosomes (genotypes were available for the X chromosome, however these were not used in these analyses), and markers were restricted to areas showing a maximum LOD score >

1 in the genome screen reported by Kehoe and colleagues. More details of the chromosomal regions genotyped in the UK and NIA samples can be found in appendix 4a. Blacker and colleagues (2003) genotyped 381 markers with an average spacing of 9cM. Ninety-nine of the markers genotyped by Blacker and colleagues and Myers and colleagues were overlapping. Markers were selected from the CHLC (http://lpg.nci.nih.gov/CHLC/index.html) and CEPH (http://www.cephb.fr/cgi-bin/wdb/ceph/systeme/forme) databases and marker positions were determined from the Marshfield maps (http://research.marshfieldclinic.org/genetics). Marker positions for those not included in the Marshfield maps were obtained through the deCode genetic maps (Kong et al. 2002). Where positions of markers could not be mapped accurately due to scarcity of synonymous markers within both maps, the physical distances obtained from the UCSC website (http://genome.ucsc.edu/) were used to map the markers relative to each other.

4.2.3 Description of variables used in covariate linkage analysis

AAO, psychosis, aggression and minor depression all showed evidence of familial clustering in section 3.3 and were therefore included as clinical covariates. Minor depression was taken to represent the mood dimension rather than major depression or depression/anxiety, as it showed the most evidence of familial clustering. Also, minor depression was the most prevalent mood related symptom, and data regarding depression/anxiety were not available for the UK sample. It was therefore considered that analysis of minor depression offered the optimal means of identifying genetic variation associated with mood in AD. AAO data were available for all samples; however data regarding other clinical features were only available for those in the NIMH and UK families. All clinical covariates were defined as described in section 3.2.2. Several studies have reported an association between aggressive symptoms and severity of dementia (Cummings 1997; Kloszewska 1998; Lopez et al. 2003; Senanarong et al. 2004). In section 2.3, symptoms of aggression were found to be more prevalent and severe among those with moderately severe to severe dementia, compared to those in the mild to moderate stages of disease development. It is therefore likely a proportion of those in the earlier stages of the disease will go on to develop aggressive

symptoms as the disease progresses. Therefore, those in the mild to moderate stages of dementia who had not displayed symptoms of aggression (GDS 2-4 and CDR 0.5-2 in UK and NIMH samples, respectively) were classified as unknown and only used in the estimation of allele frequencies and determining IBD sharing within families.

Paulsen and colleagues (2000) reported an increase in new psychotic symptoms of just 1.8% in the fourth year following onset of AD, suggesting that if they are to develop in the course of an individual's illness they will usually appear within 4 years of onset. Therefore, individuals with a disease duration ≥ 4 years who had not developed psychotic symptoms were categorised as 'Alzheimer's with no psychosis' (AD-P). Those with disease duration < 4 years who had not presented psychotic symptoms were coded as 'unknown', and were only used in the estimation of allele frequencies and determining IBD sharing within families.

4.2.4 Statistical analysis

Genotyping error checks A number of methods were used to detect genotyping errors. Non-Mendelian inheritance errors were determined for each marker using PEDCHECK (O'Connell and Weeks 1998). However this is likely to produce an underestimate of the true genotyping error rate in the absence of genotyped parents. As such, replicate datasets were simulated as discussed by Myers and colleagues (2002). Each simulation was based on the actual people genotyped and the allele frequencies at each locus, randomly introducing errors at a chosen fixed rate. For a range of given error rates, the average number of non-Mendelian errors in the simulated dataset were obtained for each locus. The final stipulated error rate was that at which the average number of errors in the simulated dataset most closely matched the observed number of errors. Data were also checked using the program SIBMED (Douglas et al. 2000) to eliminate genotypes that did not give visible inconsistencies but were nevertheless unlikely given the allele frequencies and marker maps. In addition it was possible to compare individual genotypes for the 99 overlapping markers genotyped separately by Myers and colleagues and Blacker and colleagues. The average detectable error rate for

these markers was 2.6%, ranging from 0.2% to 14.0%. Discrepant genotypes were removed from the analysis.

Linkage Analysis The program SPLINK (Holmans and Clayton 1995) was used to estimate marker allele frequencies. Estimated multipoint and twopoint affected relative pair IBD sharing probabilities were obtained using MERLIN (Abecasis et al. 2002). MERLIN uses the relationship information from all individuals to estimate IBD sharing probabilities. In comparison to other linkage programs it provides a computationally economic means of analysis, allowing the inclusion of large pedigrees genotyped at many markers, which was required for this dataset.

Ninety-nine markers overlapped between the datasets provided by Blacker and colleagues and Myers and colleagues, these were included in the analysis as independent markers separated by an artificial gap of 0.001cM. This negates the requirement for allele numbering to be consistent between studies and allows sample specific allele frequencies to be used. The use of overlapping individuals typed at similar, or the same markers, in both datasets, meant that many markers were likely to be in linkage disequilibrium (LD). Many popular linkage software packages, including MERLIN, generate multipoint linkage statistics under the assumption of linkage equilibrium between markers. Violation of this assumption can falsely inflate linkage statistics when IBD estimates rely on allele frequencies. This happens when a limited number of family members have been genotyped, or for example, if parental genotypes are unknown (Webb et al. 2005). The *distance* function in MERLIN was used to account for LD between markers. Markers within 0.5cM were considered to be in LD, representing around 500k base-pairs.

Likelihood formation The multipoint likelihood of the marker data at any point in the genome is given as a function of the IBD sharing in the affected relatives pairs by:

$$L = \prod_i \left(\sum_{j=0}^{2} z_j \frac{\hat{f}_{ij}}{f_{ij}} \right)$$

Where z_j is the (unknown) probability that an affected relative pair share *j* alleles IBD, and f_{ij}, $\hat{f}_{ij}$ are the prior and posterior (conditional on the observed marker data) probabilities that pair *i* share *j* alleles IBD (Olson 1999; Risch 1990). These estimates were obtained every 2cM using MERLIN. p_{FS} is the probability that a pair of affected full siblings share a given parental allele IBD. As suggested by Rice (Rice 1997) and Rice and colleagues (Rice et al. 1999), probabilities of sharing maternal and paternal alleles were assumed to be equal and independent. Then $z_0 = (1-p_{FS})^2$, $z_1 = 2p_{FS}(1-p_{FS})$ and $z_2 = p_{FS}{}^2$. Other types of relative pair, *R*, can only share 0 or 1 allele IBD. For these, $z_0 = 1-p_R$, $z_1 = p_R$ and $z_2 = 0$ (where p_R is the IBD probability for affected relative pairs of type *R*).

Inclusion of categorical covariates Psychosis, minor depression, aggression and APOE ε4 were treated as dichotomous traits, with individuals being classed as '+' if they displayed the symptom or '–' if they had not. Individuals were classed as 'ε4+' if they had at least one APOE ε4 allele and 'ε4-' if they did not. As such ARPs were classified as ε4-/ ε4-, ε4-/ ε4+ or ε4+/ ε4+ for APOE ε4 status. The effect of a binary covariate on the IBD sharing probabilities was investigated by modelling p_R in a logistic regression framework including a 3-level factor $\underline{\beta}$, with levels corresponding to the status of the pair with respect to the covariate (-/-, -/+ or +/+). As such the inclusion of one covariate represented here by β_k is given as:

$$p_R = \frac{e^{O_R+\alpha+\beta_k}}{1+e^{O_R+\alpha+\beta_k}}$$

where O_R is a fixed offset, ensuring that p_R takes the correct value for a relative pair of type *R* in the absence of linkage (i.e. all other coefficients in the regression = 0). Under the null hypothesis of no covariate effect, α is a measure of the divergence of IBD from the null in the sample as a whole. The subscript *k* indexes the status of the particular relative pair with respect to the covariate. Multiple pairs from the same pedigree were analysed as if they were independent, with parameters α and $\underline{\beta}$ in common. To ensure identifiability of the parameters

$\beta_{-/-}$ was set to zero (making α a measure of IBD divergence from the null in -/- pairs). To ensure that the model makes sense biologically, the degree of IBD sharing for the discordant (-/+) pairs was constrained to be less than or equal to the maximum IBD in the concordant pairs. One might expect a gene that modified the expression of a clinical covariate (e.g. psychotic symptoms) in individuals affected with late-onset AD (but not AD risk itself), to present increased sharing in -/- or +/+ pairs (or both), with -/+ pairs showing reduced sharing. A gene which acts to cause a sub-phenotype of late-onset AD characterised by a particular symptoms presence or absence would cause increased sharing in either -/- or +/+ pairs, with the effects on IBD in the pairs of other types being unclear (dependent on penetrances, gene frequencies etc.). Caution should be applied to the interpretation of the allele sharing estimates as differences could arise for a number of reasons.

Inclusion of Quantitative Covariates AAO was characterised in terms of the mean and difference within ARPs and represented by the regression coefficient β, which measures the relationship between the AAO covariate and IBD sharing. When testing AAO, where the variable of interest was the *difference* between the two members of an ARP, maximisation was carried with β constrained to be less than or equal to zero. As such, ARPs who showed most phenotypic resemblance were constrained to have higher IBD sharing probabilities.

LOD scores, test Statistic and significance levels To estimate the LOD score in the absence of covariates (e.g. the univariate LOD score), the likelihood was maximised with respect to a) α alone and, b) $\alpha = 0$, at each position x. The LOD score was then estimated as the $\log_{10}$likelihood for a) divided by the likelihood for b). To determine the LOD score give the effect of covariates, the likelihood was maximised with respect to c) α and β, and d) $\alpha = \beta = 0$. The covariate LOD score was then estimated as the $\log_{10}$likelihood for c) divided by the likelihood for d). The test statistic was then calculated by subtracting the maximum univariate LOD score over all values of x from the maximum covariate LOD score for all values of x (i.e. the increase in LOD attributable to the covariate of interest).

The location of the maximum likelihood was allowed to change when the covariate was added. This reflects the fact that linkage peaks from standard analyses are often some way from the true disease locus (Cordell 2001). Thus, incorporating covariates may give a more accurate estimate of the disease locus. Chromosome-wide significance levels were obtained by creating *n* (1000) replicate samples, each with the observed values of the covariate randomly permuted among ARPs. The likelihood-ratio analysis was performed for each replicate sample and the increase in MLS owing to the covariate estimated. The significance level was defined as $p=(r+1)/(n+1)$, where *r* is the number of replicates for which the test statistic exceeded the observed value (North et al. 2002).

For the chosen test statistic, it was not possible to obtain a genome-wide significance level for covariate effects, as this depends not only on the increase in LOD score given by the covariate, but also on the linkage evidence present in univariate analyses. For example, an increase in LOD score from 2 to 3 is more significant than from 0 to 1, because the former is likely to occur by chance (in the absence of covariate effects) only in a linkage peak region, whereas the latter could occur anywhere on the chromosome. Separate analyses were performed including the following as covariates: AAO (mean and difference), psychosis, minor depression and aggression. APOE was included as a covariate in all analyses. Analyses were restricted to autosomal chromosomes due to software limitations.

4.3 Results

Basic characteristics of the sample can be seen in table 4.1.

Table 4.1 Basic characteristics of the family based samples used for linkage analysis.

	NIMH Sample	UK Sample	NIA Sample	Combined Sample
Number of Families	295	75	64	434
Number of Individuals	651	160	136	947
No. of affected members in each family				
2 n (%)	247 (83.7)	66 (88.0)	56 (87.5)	369 (85.0)
3 n (%)	36 (12.2)	8 (10.7)	8 (12.5)	52 (12.0)
4 n (%)	11 (3.7)	1 (1.3)	0 (0.0)	12 (2.8)
5 n (%)	1 (0.3)	0 (0.0)	0 (0.0)	1 (0.2)
Gender n (%)				
Male	163 (25.0)	30 (18.8)	46 (33.8)	239 (25.2)
Female	488 (75.0)	130 (81.3)	90 (66.2)	708 (74.8)
Mean age at assessment, years	81.16	82.07	-	81.31
range	68 - 101	69 - 97	-	68 - 101
Mean age at onset, years	74.64	76.30	73.49	74.73
range	65 - 97	65 - 91	65 - 95	65 - 97
missing AOO	29	20	0	49
Mean disease duration, years	6.51	5.58	-	6.36
range	0 - 27	0 - 17	-	0 - 27
Psychosis				
AD + P, n (%)	334 (66.8)	76 (74.5)	-	410 (68.1)
AD - P, n (%)	166 (33.2)	26 (25.5)	-	192 (31.9)
Unspecified, AD-P & Duration < 4yrs	52	19	-	71
Unspecified	99	39	-	138
Minor Depression				
AD + D, n (%)	171 (32.1)	44 (41.1)	-	215 (33.6)
AD - D, n (%)	362 (67.9)	63 (58.9)	-	425 (66.4)
Unspecified	118	53	-	171
Aggression				
AD + Agg, n (%)	285 (68.2)	43 (55.1)	-	328 (66.1)
AD - Agg, n (%)	133 (31.8)	35 (44.9)	-	168 (33.9)
Unspecified, AD-D & GDS < 5	131	26	-	157
Unspecified	102	56	-	158

AAO data was available for 94.8% of all samples. The mean AAO (sd) was 74.73 (5.90). The mean AAO (sd) was 77.44 (6.22), 74.09 (5.37) and 71.82 (5.38) among those with 0, 1 or 2 ε4 alleles respectively. The total sample comprised 102 ε4-/ε4- pairs, 20 ε4+/ε4- pairs and 331 ε4+/ε4+ pairs. Pairwise classification of psychosis, minor depression and aggression can be found in table 4.2.

Table 4.2 Pairwise symptom classifications among relative pairs in the UK, NIMH and combined sample.

	Total Pairs	Number of -/- Pairs (%)	Number of +/- Pairs (%)	Number of +/+ Pairs (%)
Psychosis				
UK Sample	41	6 (14.6)	10 (24.3)	25 (60.9)
NIMH Sample	266	42 (15.7)	83 (31.2)	141 (53.0)
Combined Sample	307	48 (15.6)	93 (30.2)	166 (54.0)
Minor Depression				
UK Sample	57	25 (43.8)	22 (38.5)	10 (17.5)
NIMH Sample	296	144 (48.6)	111 (37.5)	41 (13.8)
Combined Sample	353	169 (47.8)	133 (37.6)	51 (14.4)
Aggression				
UK Sample	28	6 (21.4)	13 (46.4)	9 (32.1)
NIMH Sample	204	21 (10.2)	78 (38.2)	105 (51.4)
Combined Sample	232	27 (11.6)	91 (39.2)	114 (49.1)

The following sections detail covariate linkage results for analyses including AAO, psychosis, depression and aggression.

4.3.1 Covariate linkage analysis with age at onset

Multipoint LOD score graphs are shown in figure 4.1 for all chromosomes with a LOD score increase owing to AAO greater than, or equal to, 1 and a total multipoint LOD score including covariates greater than, or equal to, 2. The corresponding linkage results are summarised in table 4.3. A summary of linkage results for all chromosomes can be found in appendix 4b.

Table 4.3 Summary of linkage results after including pairwise AAO mean, difference and APOE. Results are shown for all chromosomes with a total maximum LOD score ≥ 2 and an increase in LOD attributable to AAO ≥ 1

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Mean AAO	Incr. Diff AAO	Incr. APOE	Incr. Mean | APOE	Incr. Diff | APOE
1	Full	521	0.73	1.58 (+) p=0.039	0.07 p=0.806	1.06	1.28 (+) p=0.093	0.47 p=0.510
1	NIMH	380	0.87	2.00 (+) p=0.024	0.00 p=1.000	1.04	1.48 (+) p=0.144	0.06 p=0.736
1	UK+NIA	141	0.54	0.17 p=0.846	0.38 p=0.478	1.38	0.13 p=0.427	0.02 p=0.477
2	Full	511	0.37	1.33 (+) p=0.150	1.37 p=0.212	0.74	1.37 (+) p=0.048	0.80 p=0.217
2	NIMH	380	0.37	1.33 (+) p=0.137	1.01 p=0.366	0.74	1.37 (+) p=0.126	0.50 p=0.494
2	UK+NIA	131	0.06	0.16 p=0.854	0.99 p=0.294	0.36	0.18 p=0.306	1.16 p=0.057
12	Full	523	0.13	0.24 p=0.847	0.79 p=0.496	1.00	0.03 p=0.658	0.51 p=0.294
12	NIMH	380	0.04	0.32 p=0.694	0.88 p=0.290	0.81	0.13 p=0.654	0.21 p=0.593
12	UK+NIA	143	0.65	0.14 p=0.828	1.71 p=0.046	0.06	0.14 p=0.396	2.12 p=0.014
14	Full	528	0.03	0.47 p=0.668	1.18 p=0.281	0.39	0.19 p=0.393	2.14 p=0.058
14	NIMH	380	0.33	0.87 p=0.271	0.74 p=0.344	0.34	0.63 p=0.303	0.98 p=0.204
14	UK+NIA	148	0.03	0.24 p=0.726	0.30 p=0.574	0.85	0.06 p=0.429	0.85 p=0.162
19	Full	529	1.47	0.48 p=0.203	1.36 p=0.016	-	-	-
19	NIMH	380	0.95	0.34 p=0.390	1.02 p=0.058	-	-	-
19	UK+NIA	149	1.41	0.73 p=0.219	0.66 p=0.174	-	-	-
21	Full	521	0.39	1.45 (+) p=0.033	0.00 p=1.000	1.35	0.95 (+) p=0.514	0.00 p=1.000
21	NIMH	380	0.07	2.26 (+) p=0.008	0.00 p=1.000	1.74	0.89 (+) p=0.356	0.00 p=1.000
21	UK+NIA	141	0.99	0.00 p=0.999	0.00 p=0.999	0.69	0.05 p=0.604	0.00 p=0.999

The six columns of maximum LOD scores refer to (left to right):
1) The maximum multipoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) mean AAO, 3) AAO difference, 4) APOE genotype. The increase in MLS given by including 5) mean AAO and 6) AAO difference *after allowing for APOE effects.*
(+) in the mean AAO column indicates increased IBD sharing with increasing AAO, (-) indicates increased IBD sharing in pairs with lower mean AAO.

The largest increase in LOD with mean AAO was observed on **chromosome 21** in the NIMH sample where a univariate LOD of 0.07 increased to 2.33 at 14cM, chromosome-wide p-value = 0.008. Pairs with increasing AAO showed elevated IBD sharing. The maximum LOD score was lower than previously reported (Holmans et al. 2005; Olson et al. 2001). Neither AAO difference, nor mean, had an effect on the maximum LOD score on chromosome 21 in the UK+NIA sample. As such the effect in the combined sample was smaller, with a univariate LOD of 0.38 increasing to 1.83, at 14cM, after including AAO mean as a covariate. However, the effect of mean AAO in the full sample still met criteria for chromosome-wide significance, p=0.033. Holmans and colleagues reported a maximum LOD score of 3.62 when including AAO mean as a covariate (NIMH sample only), which is somewhat larger than the maximum LOD score identified in this analysis. To further investigate this difference a separate, sibling pair only, analysis was performed incorporating just the sample and marker genotypes used by Holmans et al. The results were largely comparable, indicating that differences between these two analyses did not result from the use of different analytical methods (e.g. IBD sharing was estimated using MAPMAKER/SIBS by Holmans and colleagues, whereas MERLIN was used in this analysis. Also, ARPs were incorporated here, whereas Holmans and colleagues restricted analyses to affected sibling pairs). (Multipoint maximum LOD score graphs for these analyses can be found in appendix 4c). The dataset used in this chapter incorporates genotypes from 2 extra markers on chromosome 21, along with additional genotypes for 4 markers which where unique to this analysis. Marker D21S2052 (located at 24.7cM), was unique to this analysis and was in the region of elevated linkage reported by Holmans and colleagues. The univariate twopoint LOD score for this marker was 0 in the NIMH sample, rising to 0.31 after including mean AAO as a covariate. Of the other 7 markers located between 9cM and 29cM on chromosome 21, 6 showed an increase in LOD ≥ 1 after the inclusion of AAO mean. Removal of marker D21S2052 from the multipoint analysis broadened the linkage peak; however it only had a marginal effect on the magnitude and location of the MLS obtain after including AAO as a covariate (peak MLS with AAO mean of 2.33 at 14cM before removing marker D21S2052, compared to 2.47 located at 14cM after it was removed). As such it is likely that the reduction in linkage evidence, compared to that found in the analysis reported by Holmans and

colleagues, was due to the inclusion of relative pairs, additional families, family members and marker genotypes in the linkage region. Further details and twopoint LOD scores can be found in appendix 4c.

A maximum LOD score of 3.86 was observed on **chromosome 1** in the combined sample after including APOE, AAO mean and AAO difference. Mean AAO was associated with an increase in LOD from 0.73 to 2.30 around 32cM (chromosome-wide p-value = 0.039), with pairs presenting with a higher mean AAO displaying elevated IBD sharing. The increase in LOD attributable to AAO difference was 0.07. Inclusion of APOE as a covariate resulted in a LOD score increase of 1.06 around 34cM. The AAO effect resulted primarily from those in the NIMH sample, where an increase in LOD from 0.87 to 2.87 (chromosome-wide p-value = 0.024) was observed after including AAO mean as a covariate. This did not replicate in the UK+NIA sample, with a maximum LOD score of 0.63 observed after including mean AAO.

On **chromosome 2** a maximum LOD score of 2.60 was observed when including AAO mean, difference and APOE. This represented an increase of 2.23 over the maximum univariate LOD score. Both AAO difference and mean resulted in LOD score increases > 1. The maximum LOD score when including AAO mean was at 258cM, with elevated IBD sharing observed in pairs with a higher mean AAO. The maximum LOD score when including AAO difference was 1.74 at 122cM; pairs with a closer AAO demonstrated increased IBD sharing at this locus. Neither the increase attributable to AAO mean nor difference met criteria for chromosome-wide significance, p=0.150 and p=0.212 respectively. However, after including APOE in the model AAO mean had a significant effect on linkage evidence, p=0.048, in the combined sample. It is notable that the increase in linkage evidence attributed to AAO mean on chromosome 2 was restricted to the NIMH sample, whereas increases in the maximum LOD score when including AAO difference appeared to be consistent across the NIMH and UK+NIA samples, with increases of 1.01 and 0.99, respectively.

On **chromosome 14** the maximum univariate LOD score was 0.03, which increased to 2.64 when including AAO mean, difference and APOE. The main

effect appeared to result from the inclusion of APOE and AAO difference, which resulted in a LOD score increase of 2.53. Elevated IBD sharing was observed among pairs who were discordant for possessing an APOE ε4 allele and those with a closer age at disease onset. The increase in lod attributable to AAO difference did not meet criteria for chromosome-wide significance, p=0.281.The increase in LOD attributable to AAO mean was minimal in the combined, NIMH and UK+NIA samples.

Interestingly, the inclusion of AAO difference led to a chromosome-wide significant increase in LOD close to the APOE locus on **chromosome 19**, with a univariate LOD score of 1.47 raising to 2.83 (p=0.016) in the combined sample. Relative pairs with a closer AAO showed elevated IBD sharing. This effect appears to be consistent between the NIMH and UK+NIA samples. When incorporating AAO difference the LOD score increased from 0.95 to 1.97 (at 80cM) and from 1.41 to 2.07 (at 70cM) in the NIMH and UK+NIA samples, respectively. However, these effects did not meet criteria for chromosome-wide significance when considered alone, p=0.058 and p=0.174 for the NIMH and UK+NIA samples respectively. An increase in LOD of 1.41 to 2.14 was also noted with mean AAO around 66cM in the UK+NIA sample, with pairs presenting with a lower age at disease onset showing elevated IBD sharing. However, this effect did not meet criteria for chromosome-wide significance, p=0.219. The overall LOD score when including both AAO difference and mean as covariates was 2.69 in the UK+NIA sample, representing an increase of 1.28 over the univariate LOD score (p=0.185) . The maximum LOD score increase attributable to mean AAO in the NIMH sample was 0.34 (p=0.390).

Maximum multipoint LOD scores of 2.45 and 4.23 were observed on **chromosome 9 and 10,** respectively, after including AAO mean, difference and APOE as covariates. However, the univariate LOD score on chromosome 9 was 2.32 and 3.44 on chromosome 10. As such the increases in LOD attributable to AAO mean and difference were <0.1.

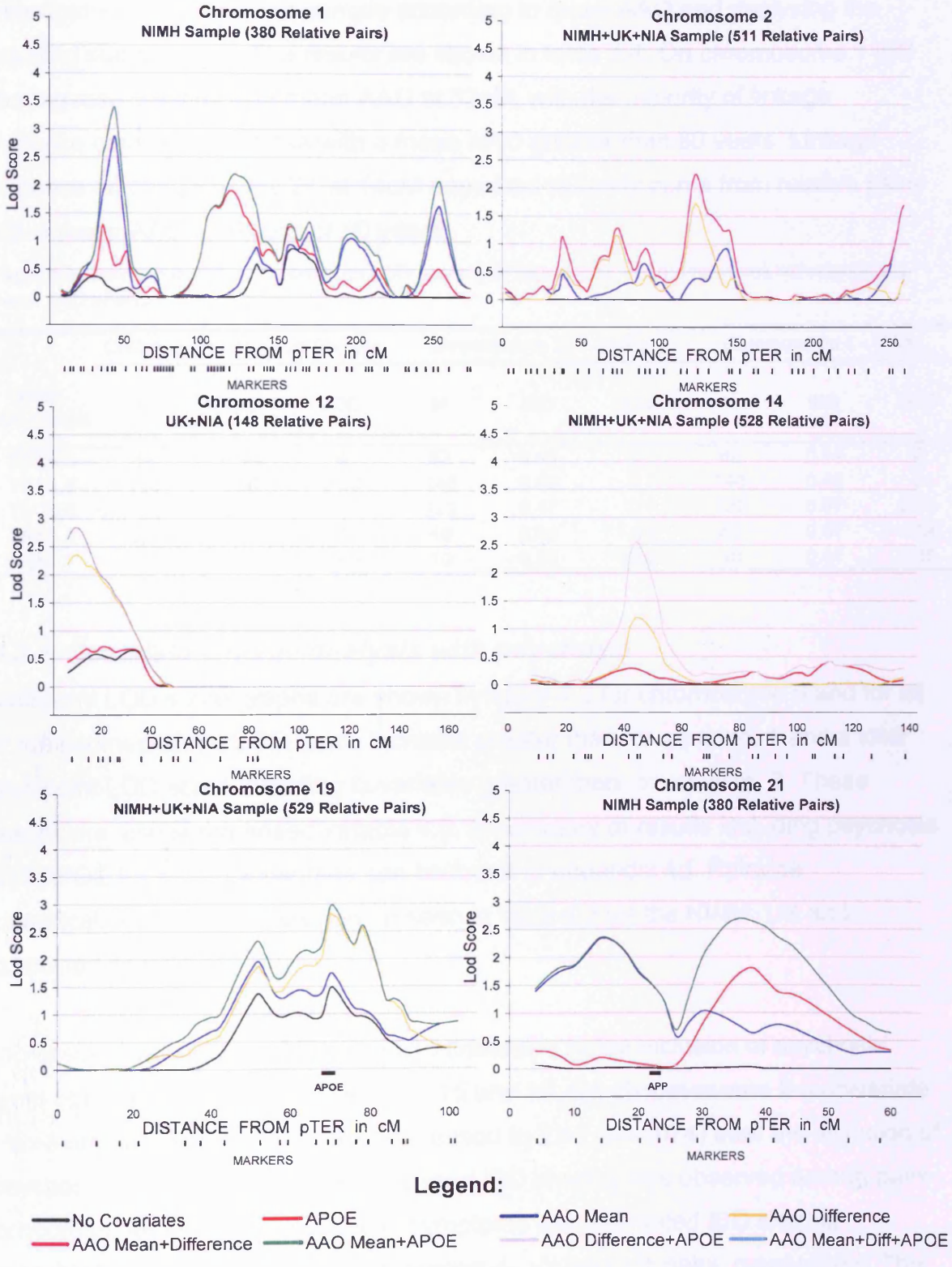


Figure 4.1 Multipoint maximum LOD score graphs for chromosomes where a total LOD score ≥ 2 and an increase attributable to AAO mean or difference ≥ 1 were observed.

The effect of mean AAO on IBD sharing on chromosomes 1 and 21 were further investigated by splitting the sample according to mean AAO and analysing the resulting sub samples. The results are shown in table 4.4. On chromosome 1 IBD sharing rose steadily with mean AAO at 32cM, with the majority of linkage evidence coming from those with a mean AAO greater than 80 years. Linkage evidence on chromosome 21 at 14cM appeared to solely come from relative pairs with a mean AAO in excess of 80 years.

Table 4.4 Variation in IBD and maximum LOD score with mean AAO for chromosomes with significant mean AAO effects

	Chromosome 21 - Full Sample			Chromosome 21 - NIMH			Chromosome 1 - NIMH		
Mean AAO range	N	IBD	LOD	N	IBD	LOD	N	IBD	LOD
65-69.9	89	0.45	0	62	0.43	0	62	0.44	0
70-74.9	196	0.51	0.03	146	0.50	0	146	0.45	0
75-79.9	158	0.49	0	122	0.47	0	122	0.57	0.70
80-84.9	61	0.58	0.58	40	0.68	1.93	40	0.67	1.54
85+	17	0.75	1.69	10	0.75	1.09	10	0.67	0.35

4.3.2 Covariate linkage analysis with psychosis

Multipoint LOD score graphs are shown in figure 4.2 for chromosome 6 and for all chromosomes with a LOD score increase greater than, or equal to, 1 and a total multipoint LOD score including covariates greater than, or equal to, 2. These results are also summarised in table 4.5. A summary of results including psychosis and APOE for all chromosomes can be found in appendix 4d. Pairwise classifications for psychosis can be seen in table 4.2 for the NIMH, UK and combined sample.

Increases in maximum LOD score ≥1 attributable to the inclusion of psychosis were observed on chromosomes 2, 7, 15 and 18. On **chromosome 2** a univariate maximum LOD score of 0.30 was increased to 2.42 (p=0.074) after the inclusion of psychosis in the NIMH sample. Increased IBD sharing was observed among pairs concordant for displaying psychotic symptoms with estimated IBD sharing probabilities of 0.52, 0.39 and 0.59 among -/-, -/+ and +/+ pairs, respectively. The maximum LOD score reduced to 1.68 after including an additional 37 relative pairs from the UK sample, representing an increase in LOD due to psychosis of 1.24,

p=0.271. In the NIMH and combined sample the peak LOD scores after the inclusion of psychosis were around 106cM and 110cM, respectively.

The largest, and most significant, increase in LOD score was observed on **chromosome 15**, where a univariate maximum LOD score of 0.20 was increased to 2.80 (p=0.007) around 80cM after psychosis was included as a covariate. Relative pairs who were concordant for the absence of psychotic symptoms provided the most evidence for linkage, with IBD estimates of 0.71, 0.42 and 0.52 among -/-, -/+ and +/+ pairs, respectively. No marker genotypes were available for those in the UK sample on chromosome 15; therefore analysis was restricted to 266 relative pairs from the NIMH sample.

Table 4.5 Summary of linkage results after including psychosis and APOE. Results are shown for all chromosomes with a total maximum LOD score ≥ 2 and an increase in LOD attributable to psychosis ≥ 1. In addition, results for chromosome 6 are shown.

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Psychosis	Incr. APOE	Incr. Psychosis | APOE
2	NIMH	266	0.30	2.12 (+) p=0.075	0.35	1.90 (+) p=0.164
2	NIMH+UK	303	0.45	1.24 (+) p=0.271	0.34	0.99 (+) p=0.507
6	NIMH+UK	303	1.09	0.58 p=0.249	1.11	0.77 p=0.092
7	NIMH	266	0.60	2.25 (+/-) p=0.015	0.32	2.00 (+/-) p=0.028
15	NIMH	266	0.20	2.60 (-) p=0.007	0.08	2.54 (-) p=0.017
18	NIMH	266	0.02	1.49 (-) p=0.200	1.14	1.23 (-) p=0.237

The four columns of maximum LOD scores refer to (left to right):
1) The maximum multipoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) Psychosis, 3) APOE genotype, 4) the increase in MLS given by including psychosis *after allowing for APOE effects.*
(+) indicates that IBD sharing was elevated in +/+ pairs, (-) indicates that IBD sharing was elevated in -/- pairs, (+/-) indicates that IBD sharing was elevated in both -/- and +/- pairs (compared to +/- pairs)
NB: Genotypes were not available in the UK sample for chromosomes where only NIMH results are shown

Increases in maximum LOD score of 2.25 and 1.49 were observed on chromosome 7 and 18. On **chromosome 7** a maximum LOD score of 2.85 after the inclusion of psychosis was observed. This increase reached criteria for chromosome-wide significance, p=0.015, and remained significant after controlling for APOE ε4 status (MLS increase controlling for APOE = 2.00, p=0.028). Consistent with a disease modifying hypothesis, IBD sharing was elevated in pairs concordant for absence *and* presence of psychotic symptoms, with IBD sharing estimates 0.67, 0.44 and 0.56 observed among -/-, -/+ and +/+ pairs, respectively. On **chromosome 18** a peak LOD score of 1.51 was observed around 106cM after incorporating psychosis. Linkage evidence was also elevated after including APOE, which resulted in an increase in the maximum LOD score from 0.02 to 1.16. However, neither the increase in LOD attributable to psychosis nor APOE met criteria for chromosome-wide significance, p=0.200 and p=0.383, respectively.

Bacanu and colleagues (2002) previously reported a genome screen using a subset of the NIMH sample presented here. They used sibling pairs who were concordant for AD, psychosis and possessing at least one APOE ε4 allele and reported significant linkage to **chromosome 2** and two suggestive linkage peaks on **chromosomes 6 and 21**. The significant linkage peak on chromosome 2 was ~50cM proximal to the linkage region identified in this chapter. The maximum LOD scores on chromosomes 6 and 21 reported here, after including psychosis status, were 1.67 and 0.52 respectively. As such these analyses do not appear to offer support for the AD+P linkage evidence identified by Bacanu and colleagues. Supplementary analyses were performed to assess these discrepancies.

Bacanu and colleagues reported a LOD score of 3.52 around marker D2S1356 when analysing a subset of 42 siblings concordant for AD, presence of psychotic symptoms and possessing at least one APOE ε4 allele (AD+P+ε4). The peak LOD score was <0.5 when they restricted their analysis to those with just AD+P (i.e. not stratified by APOE ε4 status). In the NIMH sample assessed in this chapter the maximum LOD score observed after including psychosis status at this locus was 0.69, increasing to 1.03 after including APOE, which does not support the previous report of linkage to AD+P+ε4 in this region.

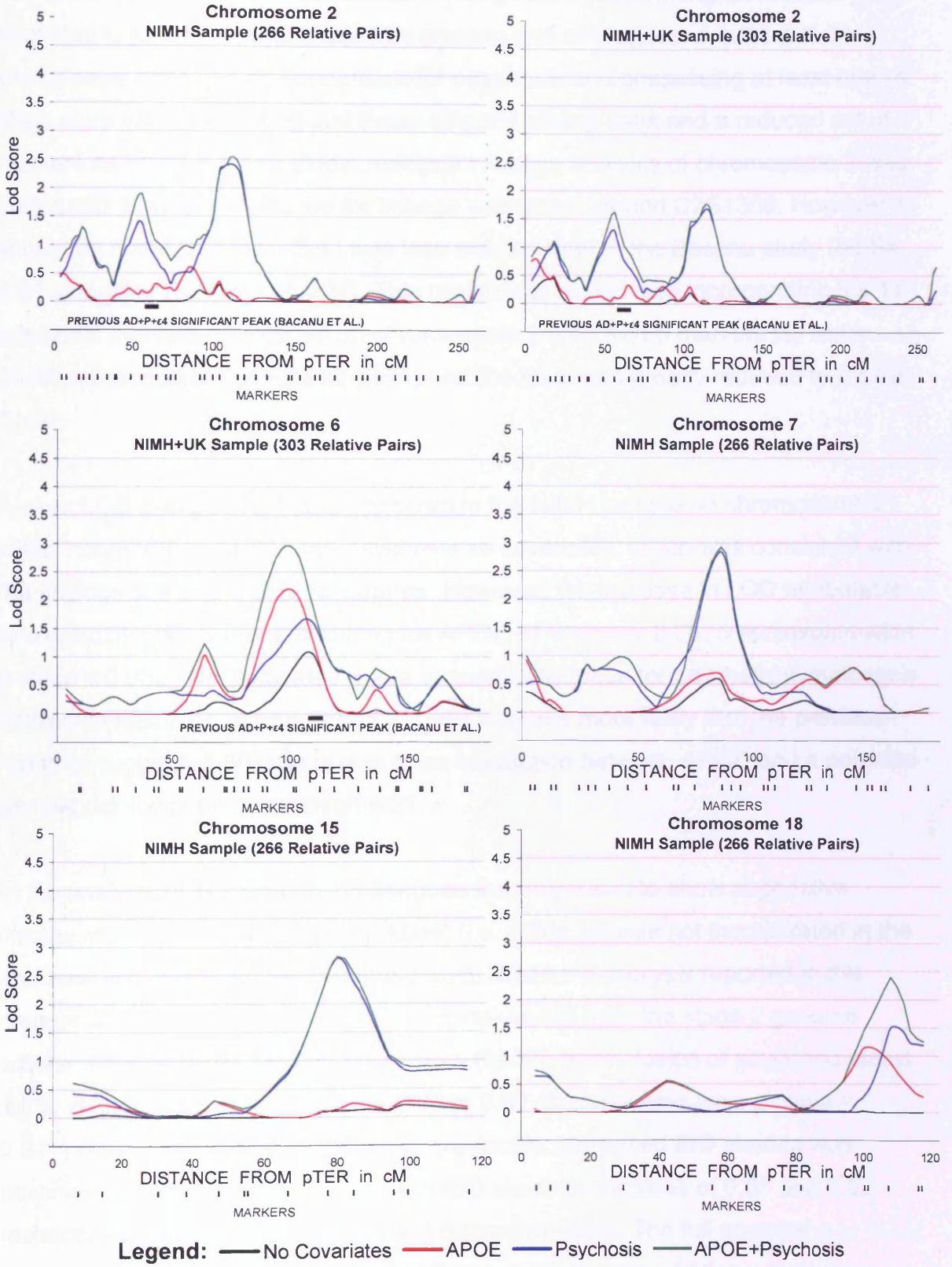


Figure 4.2 Multipoint maximum LOD score graphs for chromosomes where a total LOD score ≥ 2 and an increase attributable to psychosis ≥ 1 were observed. In additon the multipoint plot for chromosome 6 is shown.

The chromosome 2 locus was further investigated by performing an analysis restricted to just the markers used by Bacanu and colleagues. A subset of 51 sibling pairs classified as concordant for psychosis and possessing at least one ε4 allele were identified. Using just these affected sibling pairs and a reduced set of markers as in the Bacanu study, multipoint linkage analysis of chromosome 2 was performed. Elevated evidence for linkage was found around D2S1356. However it should be noted that this effect was less marked than in the Bacanu study (MLS= 1.84 at marker D2S1356, 64cM). This region was reanalysed incorporating the 11 additional markers genotyped on chromosome 2 (of which 6 markers lay within 25cM of the reported psychosis peak) and the MLS was greatly reduced to 0.84 at 61cM.

A peak LOD score of 4.01 was observed in the NIMH sample on chromosome 21 when incorporating APOE and psychosis as covariates, which was consistent with the findings of Bacanu and colleagues. However, the increase in LOD attributable to psychosis status after controlling for APOE ε4 was only 0.08, chromosome-wide p value = 0.282. This indicates that a susceptibility locus for psychotic symptoms is unlikely to reside on chromosome 21, and it seems more likely that the previous report of suggestive linkage is due to an interaction between APOE and a potential genetic risk locus on chromosome 21.

In the analysis by Bacanu and colleagues the only result to show suggestive linkage in the subset sharing only AD+P (i.e. APOE ε4 was not incorporated in the analysis) was observed on chromosome 6. When the analysis reported in this chapter was restricted to genotypes and individuals from the stage 2 genome screen reported by Myers and colleagues (2002), the inclusion of psychosis status led to an increase in the MLS from 0.60 to 2.62 (chromosome-wise p value = 0.011) at the same locus as identified previously. Increased IBD sharing was observed among -/- and +/+ pairs, with IBD sharing estimates of 0.57 and 0.63 respectively, compared to 0.42 among discordant pairs. The full analysis presented in this chapter incorporates 2 new markers and additional marker genotypes at 2 markers within this linkage region. Furthermore, an additional 59 relatives pairs were included in the full analysis. Consequently, the maximum LOD score when including psychosis reduced to 1.67, representing an increase of 0.58

over the maximum univariate LOD score of 1.08, chromosome-wide p value = 0.249. Additional maximum LOD score graphs for chromosome 6, restricted to marker genotypes used in the genome screen reported by Myers and colleagues, can be found in appendix 4e.

4.3.3 Covariate linkage analysis with aggression

Multipoint LOD score graphs are shown in figure 4.3 for all chromosomes with a total maximum LOD score ≥ 2 and an increase due to aggression ≥ 1. Covariate linkage results for analyses of aggression and APOE are summarised in table 4.6. Further details can be found in appendix 4f. Pairwise classifications for aggression can be seen in table 4.2 for the NIMH, UK and combined sample. Increases in maximum LOD score ≥1 attributable to inclusion of aggression were observed on chromosomes 3, 7, 9 and 12.

The largest LOD score increase attributable to aggression was observed on **chromosome 7** where a univariate maximum LOD score of 0.27 increased to 1.97 around 56cM. Increased IBD sharing was observed among concordant pairs, with IBD sharing estimates of 0.51, 0.35 and 0.51 for -/-, -/+ and +/+ pairs, respectively. The maximum LOD score also increased from 0.55 to 2.12 at 22cM on **chromosome 3,** with elevated IBD sharing among pairs concordant for aggression. However, increases in linkage evidence attributable to aggression on chromosomes 3 and 7 were not statistically significant, p=0.152 and p=0.108. Analyses on chromosomes 3 and 7 were restricted to those in the NIMH sample as linkage genotypes were not available for those in the UK families.

The most significant increase in MLS attributable to aggression was observed on **chromosome 9**. In the NIMH sample a univariate LOD of 2.12 increased to 3.64 at 80cM, chromosome-wide p value = 0.017. This effect reduced slightly when incorporating an additional 28 UK relative pairs, but remained significant (total MLS = 3.61, increase in MLS due to aggression = 1.43, chromosome wide p value = 0.019). Increased IBD sharing was observed in pairs in which neither member had displayed symptoms of aggression, with IBD sharing estimates of 0.77, 0.49 and 0.55 in -/-, -/+ and +/+ pairs respectively. The effect appeared stronger after

including APOE status into the model. The increase attributable to aggression was 1.85 after including APOE ε4 status, resulting in a total MLS after incorporating aggression and APOE of 4.09.

An increase of 1.20 was observed on **chromosome 12** in the combined sample when including aggression. The inclusion of APOE ε4 status also resulted in a maximum LOD score increase of 1.47. As such, the increase in LOD attributable to aggression after controlling for APOE ε4 status was 0.79, which did not meet criteria for chromosome wide significance, p=0.143.

Table 4.6 Summary of linkage results after including aggression and APOE. Results are shown for all chromosomes with a total maximum LOD score ≥ 2 and an increase in LOD attributable to aggression ≥ 1

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Aggression	Incr. APOE	Incr. Aggression | APOE
3	NIMH	204	0.55	1.57 (+/-) p=0.152	0.70	0.91 (+/-) p=0.317
7	NIMH	204	0.27	1.71 (+/-) p=0.108	0.34	1.43 (+/-) p=0.241
9	NIMH	204	2.12	1.53 (+/-) p=0.017	0.22	2.03 (+/-) p=0.017
9	NIMH+UK	232	2.18	1.43 (+/-) p=0.019	0.06	1.85 (+/-) p=0.021
12	NIMH	204	0.12	1.25 (+/-) p=0.308	1.59	0.67 (+/-) p=0.127
12	NIMH+UK	229	0.13	1.20 (+/-) p=0.339	1.47	0.79 (+/-) p=0.143

The four columns of maximum LOD scores refer to (left to right):
1) The maximum multipoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) Aggression 3) APOE genotype, 4) The increase in MLS given by including aggression *after allowing for APOE effects.*
(+) indicates that IBD sharing was elevated in +/+ pairs, (-) indicates that IBD sharing was elevated in -/- pairs, (+/-) indicates that IBD sharing was elevated in both -/- and +/- pairs (compared to +/- pairs)
NB: Genotypes were not available in the UK sample for chromosomes where only NIMH results are shown

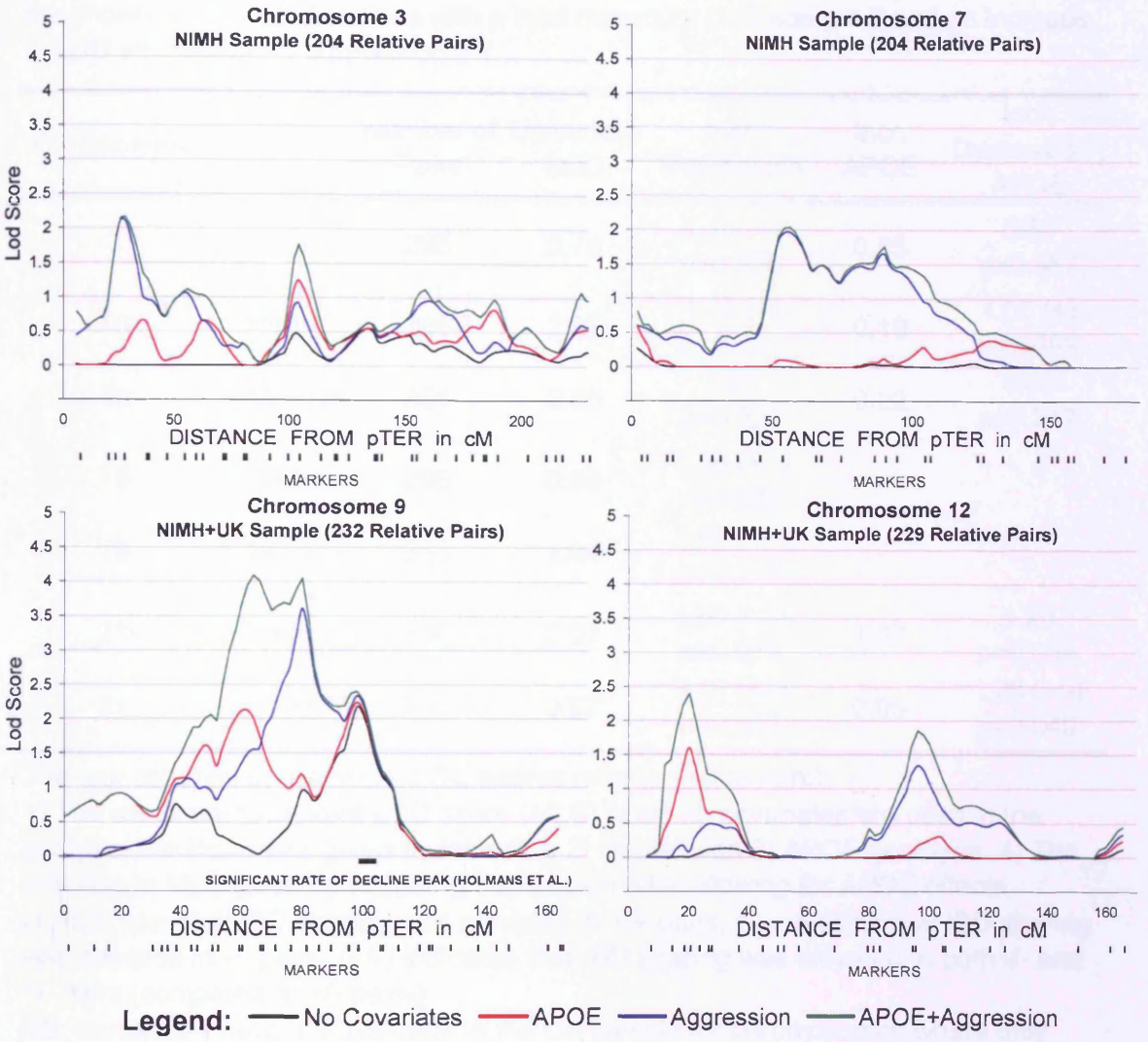


Figure 4.3 Multipoint maximum LOD score graphs for chromosomes where a total LOD score ≥ 2 and an increase attributable to aggression ≥ 1 were observed.

4.3.4 Covariate linkage analysis with depression

Multipoint LOD score graphs are shown in figure 4.4 for all chromosomes with a total maximum LOD score ≥ 2 and an increase due to depression greater than 1. Covariate linkage results for analyses of depression and APOE are summarised in table 4.7. Further details can be can be found in appendix 4g. Pairwise classifications for depression can be seen in table 4.2 for the NIMH, UK and the combined sample. Increases in maximum LOD score ≥1 attributable to inclusion of depression were observed on chromosomes 3, 10, 19 and 21.

Table 4.7 Summary of linkage results after including depression and APOE. Results are shown for all chromosomes with a total maximum LOD score ≥ 2 and an increase in LOD attributable to aggression ≥ 1

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Depression	Incr. APOE	Incr. Depression | APOE
3	NIMH	295	0.78	1.15 (+) p=0.225	0.35	0.80 p=0.367
10	NIMH	296	2.27	1.26 (+) p=0.081	0.19	1.07 (+) p=0.089
10	NIMH+UK	351	2.45	0.63 p=0.221	0.22	0.66 p=0.217
19	NIMH	296	0.96	1.56 (+/-) p=0.080	-	-
19	NIMH+UK	353	1.04	1.47 (+/-) p=0.115	-	-
21	NIMH	296	0.07	1.96 (+/-) p=0.079	1.47	1.20 p=0.034
21	NIMH+UK	351	0.27	2.19 (+/-) p=0.034	0.99	1.23 (+/-) p=0.040

The four columns of maximum LOD scores refer to (left to right):
1) The maximum multipoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) Depression 3) APOE genotype, 4) The increase in MLS given by including depression *after allowing for APOE effects.*
(+) indicates that IBD sharing was elevated in +/+ pairs, (-) indicates that IBD sharing was elevated in -/- pairs, (+/-) indicates that IBD sharing was elevated in both -/- and +/- pairs (compared to +/- pairs)
NB: Genotypes were not available in the UK sample for chromosomes where only NIMH results are shown

The largest and most significant increase in LOD attributable to depression was observed on **chromosome 21**. In the NIMH sample the univariate LOD increased from 0.07 to 2.02 at 3cM, chromosome-wide p value = 0.079. The effect became stronger after adding in 55 relative pairs from the UK sample, with a LOD score after incorporating depression of 2.47 observed in the combined NIMH+UK sample (MLS increase attributable to depression = 2.19, chromosome-wide p value = 0.034). The peak also moved to 25cM, although linkage evidence was still elevated around 3cM where an MLS of 1.93 was observed. Compared to -/+ pairs increased IBD sharing was noted in -/- and +/+ pairs. The depression peak in the combined samples was 11cM distal to the linkage peak identified with mean AAO in section 4.3.1. To further investigate the concordance of these results a

combined analysis was performed incorporating AAO and depression. In the linkage sample, AAO did not differ among those classified with and without depression, t(633)=-1.235, p=0.217. The multipoint LOD score graph for chromosome 21 including depression and AAO can be seen in figure 4.5. The inclusion of depression and mean AAO resulted in increases in MLS of 2.09 (chromosome-wide p value = 0.027) and 1.93 (chromosome-wide p value = 0.013) respectively, both around 20cM. Given that the effects of depression and mean AAO were not related, a combined maximum LOD score when both were added as covariates of 3.95 was observed.

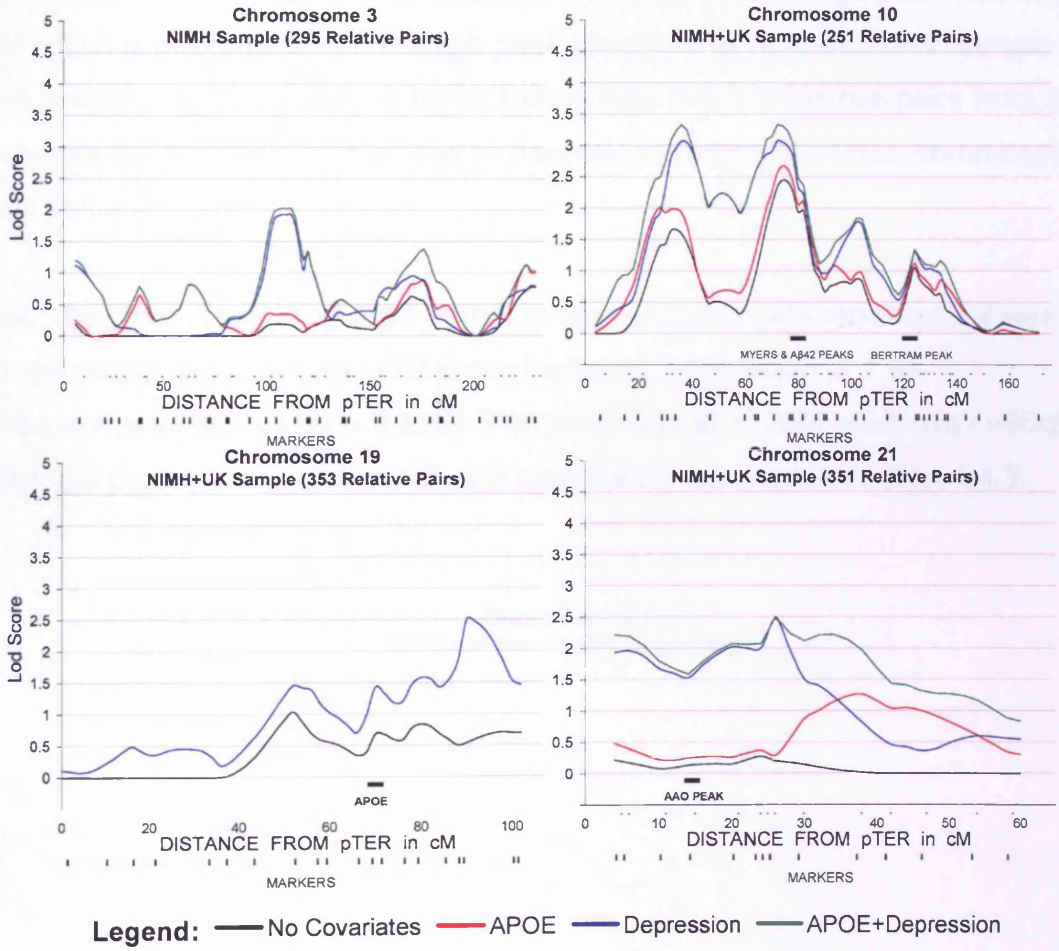


Figure 4.4 Multipoint maximum LOD score graphs for chromosomes where a total LOD score ≥ 2 and an increase attributable to depression ≥ 1 were observed.

The increase in LOD attributable to depression on **chromosome 19** was about 22cM distal to the APOE locus. The increase owing to depression was 1.47 in the combined sample, resulting in a maximum LOD score of 2.50. Increased IBD sharing was observed among pairs concordant for depression status, with IBD

sharing probabilities of 0.58, 0.45 and 0.61 among -/-, -/+ and +/+ pairs respectively. However, this effect did not meet criteria for chromosome-wide significance, p=0.115.

Interestingly depression status resulted in a MLS increase of 1.26 on **chromosome 10** in the NIMH sample, which approached chromosome wide significance, p=0.081. In the subset of 351 relative pairs with depression and APOE data the univariate LOD on chromosome 10 was 2.45, which was lower than that in the full, unstratified, sample. The maximum LOD score in the NIMH sample after including depression increased to 3.53. The linkage peak was at 72cM which is proximal to the linkage peak identified using part of this sample by Myers and colleagues (2002). After including data from 55 relative pairs from the UK sample the increase in LOD due to depression reduced to 0.63, chromosome wide p value = 0.221.

Finally, the increase in LOD attributable to depression on **chromosome 3** was 1.15 resulting in a maximum LOD score including depression of 1.93 (chromosome-wide p value = 0.225). The peak was at 112cM which was ~90cM to the linkage peak identified with AD and aggression reported in section 4.4.3.

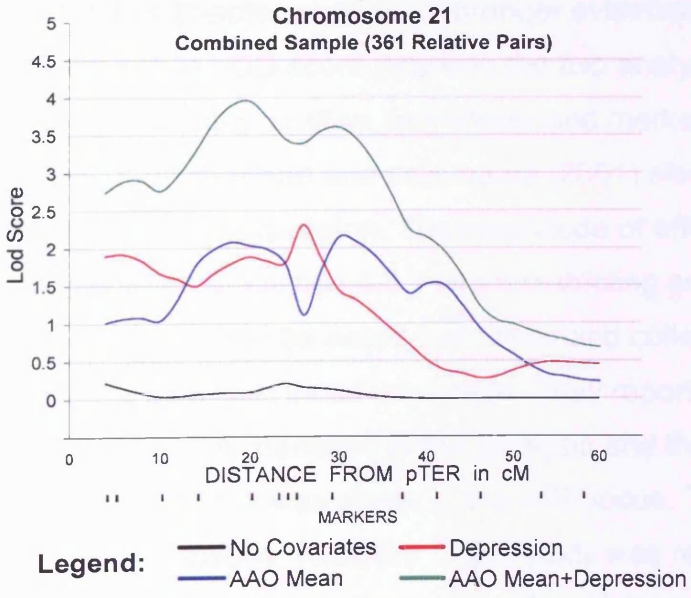


Figure 4.5 Multipoint maximum LOD score graph for chromosomes 21 including AAO mean and depression as covariates.

4.4 Discussion

4.4.1 Summary of findings

Covariate linkage analysis was performed in a large sample of relative pairs affected with late-onset AD, incorporating age at disease onset, psychosis, aggression and depression. Nine novel linkage regions were identified. Increases in linkage evidence, which reached criteria for chromosome wide significance, were observed on chromosomes 1, 2 and 21 when mean AAO was included as a covariate in linkage analysis. The inclusion of AAO difference led to an increase in MLS on chromosome 12 in the UK+NIA sample which met criteria for chromosome wide significance. In addition, a significant increase in LOD close to the APOE locus in the combined sample was noted with AAO difference. Significant increases in LOD were found after the inclusion of psychosis (chromosomes 7 and 15), aggression (chromosome 9) and depression (chromosome 21). Furthermore, numerous other linkage signals with covariate increases ≥ 1 were identified which failed to reach criteria for chromosome wide significance.

4.4.2 Discussion of linkage findings in relation to age at onset

The most significant finding was with AAO mean on chromosome 21p. This is consistent with previous analyses which have incorporated AAO as a covariate in genome screens of late-onset AD. Holmans and colleagues (2005) used a subset of the sample reported in this chapter and found stronger evidence for linkage than identified here. The decrease in LOD score between the two analyses was attributable to the inclusion of extra families, individuals and marker genotypes, but was mainly owing to one marker. Olson and colleagues (2001) also reported increased evidence for linkage in this region. The magnitude of effect was comparable to the one reported in section 4.3.1 when restricting analyses to just those in the NIMH sample. It should be noted that Olson and colleagues included many families which overlapped with those used here. They reported linkage evidence ~13cM distal to the peak identified in this analysis and the one reported by Holmans and colleagues, thus, it was closer to the APP locus. This discrepancy could be due to a number of reasons. First, the Olson study was restricted to genotypes from the stage one genome screen reported by Kehoe and colleagues (1999), whereas this analysis includes data from the stage 2 genome screen

reported by Myers and colleagues (2002). Second, Olson and colleagues included individuals with an AAO above 60 and required that all individuals had data regarding *current* age. Despite individuals in the UK and NIA samples having a similar distribution of AAO to those in the NIMH sample, linkage evidence decreased when including these families.

Modest evidence of linkage was observed ~14cM distal to the NIMH mean AAO linkage peak in univariate analysis of the UK+NIA sample (MLS= 0.99). APOE status also increased linkage evidence on chromosome 21 in this dataset; consistent with previous findings elevated IBD sharing was observed in relative pairs with no APOE ε4 alleles (Kehoe et al. 1999; Myers et al. 2002). This is of interest as evidence suggests that the APOE ε4 allele reduces AAO in a dose dependent manner. These findings suggest a role for APP, or a gene in this region, in the genetic susceptibility for AD. However, as pointed out by Olson and colleagues, the absence of ε4 alleles in those showing most evidence for linkage on chromosome 21 may not represent an interaction at the biological level, as genetic variation at this locus does not appear to confer risk to AD until very late in life. Individuals with APOE ε4 alleles may have developed 'ε4 type AD' prior to the age when genetic variation at the chromosome 21 begins to exert its effect.

The inclusion of depression as a covariate also increased linkage evidence at this locus. Elevated IBD sharing was observed among relative pairs in which both members were concordant for either the presence or absence of depression. The effects of depression and AAO appeared to be additive. It is possible to draw a number of conclusions from the linkage evidence on chromosome 21: (a) increased linkage evidence may result from epistasis between the APOE locus and the region on chromosome 21, (b) a gene in this region may increase susceptibility to AD among those ≥ 80 years of age, (c) a gene in this region modifies susceptibility to mood disturbances in AD. A combination of these three scenarios is possible. For example, genetic variation in this region may increase susceptibility to very late-onset AD, which modifies the risk of developing symptoms of depression.

The second largest increase in linkage evidence attributable to AAO was observed on chromosome 1. This finding was novel and was not reported in previous analyses including AAO (Holmans et al. 2005; Li et al. 2002; Olson et al. 2002). Similar to chromosome 21, increased IBD sharing was observed among pairs with a higher mean AAO, with the effect becoming particularly noticeable among those with a mean AAO ≥ 80 years. The finding was much stronger in the NIMH sample; however a very small effect was noted in the UK+NIA families. The effect attributable to AAO mean was mediated slightly by APOE, as such the effect of adding mean AAO after controlling for APOE was smaller, failing to meet chromosome wide significance.

Despite the known effects of APOE on AAO in AD, mean AAO did not have an effect on linkage in either the NIMH or UK+NIA sample around the APOE locus on chromosome 19. However, increases in LOD attributable to AAO difference were noted in both samples around the APOE locus. As a result the increase in LOD in the combined sample reached criteria for chromosome-wide significance. The effect was small, which is consistent with the findings in section 3.3.2 that suggest APOE accounts for only a small proportion of the familial effect on AAO.

The only other notable increase in linkage evidence attributable to AAO difference was on chromosome 12 in the UK+NIA sample. This finding was ~30cM proximal to the linkage peak identified by Pericak-Vance and colleagues (1997). This finding was not replicated in the NIMH families. Consequently, this finding was not significant in the combined sample, indicating that it might represent a false positive. The linkage peak on chromosome 12 with mean AAO in the NIMH sample was smaller than the one reported by Holmans and colleagues, and as such it did not meet criteria for chromosome wide significance.

Holmans and colleagues reported increased evidence for linkage on chromosome 6 and 15, which are not supported here after including additional markers genotypes and individuals. The maximum LOD score on chromosome 15 in the Holmans study was around marker D15S165, at ~20cM. It is notable that the analysis reported here incorporated additional marker genotypes for marker D15S165, along with genotypes from two additional markers on chromosome 15.

Similarly additional marker genotypes in the region previously identified on chromosome 6 were exclusive to this analysis, resulting in a decrease in linkage evidence.

The other major study of AAO in AD was reported by Li and colleagues (2002). Their analysis included data from a subset of families ascertained by the NIMH and NIA, which overlapped somewhat with those used in this chapter. As noted in section 4.1.5, their analysis differs to the one reported here, in that it treated AAO as a quantitative trait. As such Li and colleagues considered IBD sharing between pairs of relatives and AAO as the independent and dependent variables, respectively. Therefore, they were unable to test for overall linkage to AD. They reported evidence for linkage around marker D10S1237, at 139cM. Despite finding evidence for a highly significant linkage peak on chromosome 10 between 75cM and 86cM, the results from this chapter do not support the conclusion that the chromosome 10 loci is related to AAO. Li and colleagues also reported linkage peaks on chromosomes 4, 6, 8, 13 and 18. Those on 4, 8 and 13 did not replicate in this sample. Increases in linkage evidence were observed with AAO difference on chromosomes 6 and 18 which corresponds to those reported by Li and colleagues. However the maximum LOD scores on these chromosomes were less than 2, and did not meet criteria for chromosome wide significance.

4.4.3 Discussion of linkage findings in relation to psychosis

The strongest evidence for linkage to psychosis in AD was observed on chromosomes 15. Increased IBD sharing was observed among pairs in which neither member had displayed psychotic symptoms, which could be consistent with a sub-phenotype hypothesis. The increased evidence for linkage provided by psychosis status was independent of APOE. This provides evidence that a gene may be located in this region that is linked to a sub-phenotype of AD characterised by the absence of psychotic symptoms. In addition, the inclusion of psychosis as a covariate led to a significant increase in linkage evidence on chromosome 7. Increased IBD sharing was observed among pairs concordant for either the presence or absence of psychotic symptoms, whilst IBD sharing was decreased in pairs discordant for psychosis status. This finding is consistent with a disease

modifying hypothesis; however association studies would be required to confirm this. It is notable that the previous two analyses (Avramopoulos et al. 2005; Bacanu et al. 2002) which have performed genome screens, incorporating information regarding psychotic symptoms, have not reported increased linkage evidence in these regions. However, this could be attributable to methodological differences between these studies. For example, Bacanu and colleagues (2002) relied on sample stratification to incorporate data regarding psychotic symptoms into linkage analysis. As such, investigations were limited to sibling pairs in which both members had AD with psychotic symptoms. In this chapter increased linkage evidence was noted on chromosomes 7 and 15 among those *without* psychotic symptoms, which would not have been detected using the analysis method adopted by Bacanu and colleagues. Avramopoulos and colleagues treated hallucinations and delusions separately. It is therefore difficult to compare their results to those reported here, where hallucinations and delusions were viewed as symptoms of a larger behavioural component represented by psychosis.

Evidence from this genome screen does not support the previous findings, which suggested that a susceptibility locus for AD+P resides on chromosome 2 or 21 (Bacanu et al. 2002). This study suggests that the previous significant linkage signal identified on chromosome 2, found when analysis was restricted to sibling pairs where both members had psychotic symptoms and at least one ε4 allele, may represent a false positive. After incorporating additional marker genotypes in this region there was no evidence of linkage, either with APOE ε4 status or psychosis. Supplementary analyses, which were more comparable with those reported by Bacanu and colleagues, did identify a linkage peak with APOE on chromosome 2. However, the signal was greatly reduced when additional markers were genotyped close to the linkage peak. There was evidence for a potential locus on chromosome 21 when psychosis and APOE ε4 status were included in the analysis. However, the majority of the increase in MLS was accounted for by APOE ε4 status which agrees with previous findings in studies using part of this sample (Kehoe et al. 1999; Olson et al. 2001). Psychosis did not have a significant effect on linkage. This indicates that a susceptibility locus for psychotic symptoms is unlikely to reside on chromosome 21, and it seems more likely that the previous

report of suggestive linkage is due to an interaction between APOE and a potential genetic risk locus on chromosome 21.

Bacanu and colleagues also reported a linkage peak for AD+P on chromosome 6. In this chapter, the inclusion of psychosis as a covariate led to a chromosome wide significant increase in linkage evidence when analysis was restricted to genotypes and families from the stage 2 genome scan reported by Myers and colleagues (2002) ('stage 2' sample). This was of particular interest as the chromosome 6 finding was the only suggestive linkage peak reported by Bacanu and colleagues which was independent of APOE ε4 status. After the inclusion of additional families and markers the evidence for linkage in the absence of covariates increased, whilst the maximum LOD score in the model including psychosis status was reduced. As such the increase in linkage evidence attributable to psychosis reduced substantially; hence, did not meet criteria for significance. Given the contradictory results obtained within this sample these findings cannot be taken to exclude the possibility that a gene, or genes, in this region are implicated in the aetiology of psychosis in AD. However, further analyses in independent samples would be needed to investigate this region.

The chromosome 6 locus is of particular interest as evidence for linkage in this region has also been identified in studies of schizophrenia and bipolar disorder, both of which have prominent psychotic features. For example, a p-value of 0.00018 was observed at marker D6S474 (~6cM distal to the psychosis peak in the 'stage 2' sample) using 61 ASPs from 53 pedigrees (Cao et al. 1997). In an independent series of ASPs Cao and colleagues observed a p-value of 0.00095 at a locus ~2cM proximal to their previous linkage peak, a finding that was later replicated by Martinez and colleagues (1999) who observed a p-value of 0.013 at marker D6S424 (at 104cM). Further to this a maximum LOD score of 3.10 at 106.9cM (~6cM proximal to the psychosis peak identified in the 'stage 2' sample in this chapter) was found in a combined sample, collected at 8 centres, consisting of 824 independent sibling pairs with schizophrenia (Levinson et al. 2000). Also, the locus on chromosome 6 is arguably the most supported region of linkage implicated in the development of bipolar disorder. In their genome screen using 250 pedigrees affected with bipolar disorder, Dick and colleagues provided

evidence for a major locus on chromosome 6, with a MLS of 3.61 observed between D6S1021 and D6S474 (112cM-118cM). Pato and colleagues (2004) have replicated this finding in a sample of Portuguese bipolar disorder sufferers, reporting suggestive evidence for linkage at the same marker. More recently, Lambert and colleagues have found evidence for suggestive linkage to bipolar disorder in this region in an independent sample of sibling pairs from 135 families (Lambert et al. 2005). McQueen and colleagues (2005) have performed a meta-analysis of 11 studies and concluded that this region on chromosome 6 is the most important region in bipolar disorder. It is possible that the similarity of these findings in AD, schizophrenia and bipolar disorder could represent a locus that influences susceptibility to psychotic symptoms across a number of disorders. However, it should be noted that the common psychotic symptoms differ in AD and schizophrenia; for example, delusions in AD are typically non-bizarre and simple, and seem to differ somewhat from the more complex and bizarre delusions seen in patients with schizophrenia (Jeste and Finkel 2000). Hallucinations in AD patients are more frequently visual than auditory, which again differs from schizophrenia where the reverse is true (Jeste and Finkel 2000).

4.4.4 Discussion of linkage findings in relation to depression and aggression

Linkage analyses incorporating symptoms of depression and aggression in AD have not been reported previously. Significant increases in linkage evidence were observed on chromosomes 21 with depression, and chromosome 9 with aggression. The finding on chromosome 9 is of particular interest as this region has been implicated in numerous genome screens using the NIMH AD sample (Blacker et al. 2003; Myers et al. 2002; Pericak-Vance et al. 2000). In addition, Holmans and colleagues observed increased IBD sharing among sibling pairs with a faster rate of decline close to this locus. This is of interest as symptoms of aggression are generally believed to be a marker for a more rapid disease progression (Levy et al. 1996a; Lopez et al. 1999). However, linkage evidence on chromosome 9 appeared to be coming solely from those *without* symptoms of aggression, which is not consistent with the hypothesis that this region harbours a gene which is implicated in the development of a more severe form of AD

characterised by rapid disease progression and symptoms of aggression. In addition to the chromosome wide significant increase in LOD attributable to depression on chromosome 21 already discussed, a non-significant increase in LOD on chromosome 19 was also observed after including depression as a covariate. The peak was ~30cM from the APOE locus, located at ~60cM, therefore is unclear whether this reflects a role for APOE in the development of depression in AD.

Linkage evidence in AD is arguably most consistent on chromosome 10, with many studies finding evidence for linkage in this region. It is interesting that the covariate effects on this chromosome were limited. It is noteworthy that highly significant evidence for linkage in the combined sample was observed in univariate analyses of this region. As such, one would have to conclude that a genetic susceptibility locus for AD may reside on chromosome 10, which is unlikely to be restricted to specific disease sub-phenotypes included in these analyses. Furthermore, this region is not likely to be implicated in the modification of AAO (as suggested by Li and colleagues), psychotic symptoms, depression or aggression in AD. However, it should be noted that covariates will have the most effect when they are able to differentiate between those with high and low IBD sharing. In the presence of a high univariate linkage signal, where the majority of pairs have high IBD sharing, the power of covariates to exert an effect is likely to be limited.

4.4.4 Implications and future research

Linkage signals identified from this type of analysis could be explained in a number of ways. Firstly, linkage signals could represent disease modifying genes, which could influence susceptibility to psychotic symptoms in the presence of neurodegeneration caused by AD. For example, Psychotic symptoms have already been associated with the serotonin receptor genes, 5-HT_{2A} and 5-HT_{2C}, which are not associated with the risk of developing AD, hence are hypothesised to act as disease modifying genes (Assal et al. 2004; Holmes et al. 1998a; Lam et al. 2004; Nacmias et al. 2001). Alternatively, by including clinical information as a covariate, 'phenocopies' which do not share the same genetic aetiology will have

less effect on linkage results. For example, Bacanu and colleagues (2002) argue that AD+P represents a 'purer form of AD' that will be more useful in determining genuine linkage with the disease itself. It is particularly useful that the method of covariate linkage analysis used in this chapter gives some indication of which hypothesis is most likely to be correct for any given locus. The results from these analyses also demonstrate the utility of covariate linkage methods, both for the detection of linkage in the presence of locus heterogeneity and also in the identification of homogenous subgroups of patients who are likely to show linkage to a given loci. This information should be taken into account when performing further analyses based on linkage results from this study. For example, future studies of APP in late-onset AD should consider the role in the very elderly (e.g. those with an AAO ≥ 80), whilst also taking into account the occurrence of depressive symptoms during the course of the illness.

There were numerous examples where the inclusion of covariates led to an increase in LOD ≥ 1, which did not reach criteria for chromosome wide significance. It is unfortunate that a number of these regions resided on chromosomes which were only genotyped in the NIMH sample. Genotyping these regions in the UK and NIA samples, or in other independent datasets, will undoubtedly provide a clearer picture of how they are implicated in AD and with the covariates used in this chapter.

4.4.5 Methodological critique

There are a number of important factors to consider when interpreting the results from these analyses. First and foremost, these investigations are largely of an exploratory nature. As such, no adjustments for multiple testing were performed. Significance testing was performed at a chromosome-wide level. In theory it is possible to extrapolate these significance levels to the whole genome, in a Bonferroni style test, by multiplying by the total number of chromosomes investigated, adjusting for the length of the chromosome where the effect was observed. However, it was considered that this would have been too conservative for this explorative investigation. Furthermore, the inclusion of covariates undoubtedly increases the overall probability of type I error due to number of analyses performed on the same dataset. Olson and colleagues (2001) suggest

careful prior selection of covariates as one method to limit the problems caused by multiple testing. Two methods of limiting the number of covariates tested were employed here. First, analyses were limited to a small number of behavioural components, guided by the principal components analysis reported in chapter 2, rather than numerous individual symptoms. Second, linkage analysis was further restricted to aspects of phenotypic variation which showed evidence of familial clustering in chapter 3. Despite the use of these strategies a total of 5 covariates (AAO mean and difference, psychosis, depression and aggression) were analysed. As such one may choose to apply a Bonferroni adjustment to the chromosome or genome wide significance levels obtained for each covariate. This approach would be highly conservative and was therefore not adopted for these analyses. However, these considerations emphasise the need to replicate these findings in independent samples.

As mentioned in section 3.4, another possible limitation of this study stems from the data used to categorise individuals in terms of behavioural components. These data were primarily collected for use in a genome screen for late-onset AD; therefore limited information was available covering the presence, type and severity of behavioural symptoms. The Brief Psychiatric Rating Scale, BEHAVE-AD and the MOUSEPAD offer a good overview of behavioural disturbances. However, since data were collected at one time point in disease development, which varied between individuals, the possibility that behavioural components for some individuals may have changed over time, and were therefore mis-specified in this study, cannot be ruled out. It is reassuring that using these data it was possible to demonstrate familial aggregation of behavioural components (see section 3.3). It appears that the combination of clinical (e.g. AAO, psychosis etc.) and genetic information may offer vital clues in the search for genes associated with the development and progression of AD; therefore future collections of AD samples should aim to acquire information on behavioural symptoms experienced by disease sufferers longitudinally over the course of the illness.

The logistic regression approach employed in this study makes efficient use of covariate data within the context of an affected relative pair linkage design, as data from all informative related individuals contribute to the analysis. This contrasts with

the frequently used method in which a primary analysis is undertaken with the whole dataset followed by a secondary analysis on a subset selected according to the covariate. For example, this method was used by Bacanu and colleagues (Bacanu et al. 2002) in their genome screen of AD with psychosis. However, this approach overlooks a considerable amount of data, as the unselected individuals are either excluded totally or they remain in the analysis with phenotype designated as 'unknown'. Either way, a direct comparison between the results is hindered, as the sample sizes for the main and secondary analyses differ (Hamshere et al. 2005). For example, Bacanu and colleagues identified a significant linkage peak on chromosome 2 in sibling pairs concordant for AD, psychosis and possessing at least one ε4 allele. However, by performing analysis on a sample substratum it was difficult to determine whether the increased evidence for linkage resulted from AD alone, psychosis or an interaction with the APOE locus. The logistic regression approach also provides useful information about the underlying genetic effect. By comparing the pattern of IBD sharing among discordant and concordant pairs for each covariate, it is possible to make inferences about whether linkage regions contain genetic variation which is likely to modify disease progression or whether it contains a gene, or genes, which increase susceptibility to a specific disease sub-phenotype.

4.4.6 Conclusions

Familial aggregation of psychosis, depression, aggression and age at disease onset were reported in chapter 3 of this thesis, suggesting that aspects of clinical variation commonly observed in AD are genetically influenced. In this chapter a number of novel linkage regions have been identified after including clinical covariates in linkage analysis. This further supports the hypothesis that disease presentation in AD is genetically modified, or that clinically homogeneous sub-phenotypes exist which will be useful in untangling the complex genetics of late-onset AD. The strongest evidence for linkage was observed with mean AAO on chromosome 21. In addition, four significant increases in linkage evidence were observed after including psychosis, aggression and depression. While it is likely that a number of these loci represent false positives, future linkage and association analyses of these regions should consider incorporating the relevant phenotypic information. The obvious next step is to screen these regions for functional candidate genes which may be implicated in the susceptibility to AD, sub-phenotypes or disease modification.

Chapter 5

General discussion

Late-onset Alzheimer's disease (LOAD) is a clinically heterogeneous disorder. There is substantial evidence that LOAD is highly heritable, however, other than APOE, the identification of further susceptibility genes has remained problematic. Some have suggested that aspects of clinical variation may offer an appropriate means of delineating homogeneous disease subtypes which could be useful for genetic analysis (Sweet et al. 2003). This approach is becoming increasingly popular in genetic investigations and has facilitated the identification of susceptibility genes in studies of asthma (Van Eerdewegh et al. 2002) and Crohn's disease (Rioux et al. 2001). A growing number of linkage (discussed in section 4.1.4) and association studies (discussed briefly in section 1.3) have sought to identify genes associated with AD. However, studies which have aimed to determine which aspects of clinical variation are likely to be genetically modified have been relatively scarce (discussed in sections 3.1.4 and 3.1.5). This thesis has employed a more comprehensive approach to studying the genetic aetiology of clinical variation observed in AD. First, principal components analysis was used to elucidate behavioural components, or groups of symptoms which are likely to be related. These components were then used to inform the classification of behavioural symptoms in a sample of relative pairs, with the initial goal of determining which aspects of the clinical phenotype in AD are genetically influenced. A method of covariate linkage analysis was then used to search for loci linked to aspects of the disease phenotype showing evidence of familial clustering.

5.1 Summary of findings

The main aim in chapter 2 was to determine if the variety of behavioural symptoms observed in AD could be explained by a lesser number of components. Behavioural disturbances among a sample of 1,120 individuals suffering with AD were characterised using the Neuropsychiatric Inventory (NPI) (Cummings 1997). These data were then submitted to principal components analysis. This represents

the largest study of its type reported to date. Three components were identified which largely accounted for the 12 symptoms covered by the NPI, these were: 'frontal lobe dysfunction', 'psychosis' and 'mood'. They remain relatively stable after controlling for disease severity and in analysis restricted just to those in the later stages of disease development. Furthermore, the component structure was largely comparable to a number of previous reports (Aalten et al. 2003; Fuh et al. 2001; Mirakhur et al. 2004; Spalletta et al. 2004), providing further evidence that robust constructs had been identified which could reflect common underlying neuropathology. The identified components were independent of gender, current age, years in education, number of APOE ε4 alleles and family history of dementia. However, higher frontal lobe dysfunction component scores were associated with a lower AAO, whilst higher scores were reported for all components among those with more severe cognitive impairment.

It is likely that components identified through this kind of analysis have a closer biological link to the underlying causative mechanisms than symptoms which are categorised based on psychiatric classification systems. Research into the familial aggregation of behavioural symptoms is limited; however, the available evidence suggests that psychosis, depression and agitation in AD could be genetically influenced (Bacanu et al. 2005; Sweet et al. 2002; Tunstall et al. 2000). Similarly, a number of studies have suggested that AAO in LOAD may be heritable (Daw et al. 1999; Li et al. 2002; Tunstall et al. 2000). However, before searching for loci involved in the aetiology of clinical variation, further research were required to expand earlier family based studies. In chapter 3 of this thesis, data from the principal components analysis were used to inform the categorisation of symptoms in a large series of sibling pairs. As such individuals were rated for psychosis, aggression, depression with anxiety, minor and major depression, with the primary aim of determining whether these symptoms clustered in families more often than would be expected by chance. The familial influence on age at disease onset was also investigated.

A moderate correlation between ages at disease onset among siblings was observed. Furthermore, this effect was consistent across three independent samples of sibling pairs and was independent of the effects conferred by the

APOE ε4 allele. Of the behavioural symptoms assessed psychosis showed the strongest evidence for familial aggregation. Consistent with previous findings (Bacanu et al. 2005; Sweet et al. 2002), psychotic symptoms in probands were associated with over a three-fold increase in risk for developing psychosis in AD affected siblings. This effect was noted in both samples with available data. Significant familial influences were also observed for symptoms of aggression and minor depression. The results regarding major depression and depression with anxiety were less consistent. Both of these in probands were associated with around a two-fold increase in risk among siblings with AD. However, the effects on the whole were not statistically significant. This may have reflected a lack of power, given that both of these symptoms had a relatively low prevalence.

AAO, psychosis, aggression and minor depression showed the strongest evidence of being genetically influenced. As such they may be markers for more homogeneous disease subtypes or may be modified by genetic variation which does not increase susceptibility to the disease as a whole. Linkage analysis provides a means of searching for regions which are likely to contain loci linked to a particular phenotype. In chapter 4, age at disease onset (mean and difference among relative pairs), psychosis, aggression and minor depression were used as covariates in linkage analysis. Samples from two overlapping genome screens, presented by Myers and colleagues (2002) and Blacker and colleagues (2003), were combined to provide one of the largest samples used for linkage analysis of LOAD reported to date. A number of novel linkage signals were identified. Three peaks were observed with mean AAO, on chromosomes 1, 2 and 21. The linkage peak on chromosome 21 was in close proximity to the ones previously reported in overlapping samples (Holmans et al. 2005; Olson et al. 2001), and was largely due to increased IBD sharing probabilities among those aged over 80. Likewise, the linkage signals on chromosome 1 and 2 were driven by relative pairs with a later mean AAO. Linkage obtained with AAO difference has different, but equally important, biological implications. For example, one may expect to observe linkage among relative pairs with a similar AAO if different disease susceptibility genotypes give different rates of increase of disease risk with age. AAO difference led to a significant increase in LOD close to the APOE locus on chromosome 19 in the full sample. Also, it was associated with a significant increase in linkage

evidence on chromosome 12 in the UK+NIA sample, although this effect was not observed in analysis restricted to NIMH relative pairs.

The inclusion of behavioural covariates led to four chromosome wide significant increases in LOD. Two of these regions were attributable to psychosis, which would appear consistent with the findings in chapter 3 suggesting that the genetic influences on psychosis among AD sufferers are greater than those for aggressive and mood related disturbances. Specifically increases in linkage evidence were observed on chromosome 7 and 15 with psychosis, chromosome 21 with depression and chromosome 9 with aggression. Interestingly, the linkage peak observed with depression on chromosome 21 was in close proximity to the region showing increased linkage among pairs with a higher mean AAO, suggesting that a gene, or genes, in this region may be implicated in the development of late-late onset AD and the aetiology of depressive symptoms. No evidence was found to support the previously reported linkage peaks with psychotic symptoms in AD on chromosomes 2 and 21 (Bacanu et al. 2002), or on chromosome 14 (Avramopoulos et al. 2005). Chromosome 10 is arguably the most consistent region of linkage for late-onset AD reported to date (Bertram et al. 2000; Blacker et al. 2003; Ertekin-Taner et al. 2000; Farrer et al. 2003; Li et al. 2002; Myers et al. 2000). The data from these analyses come from two overlapping genome screens which both report linkage in this region, as such highly significantly linkage evidence was found on chromosome 10. However, none of the covariates had an effect on linkage in this region, hence it is likely that chromosome 10 harbours a gene, or genes, which increases risk of AD as a whole, irrespective of age at disease onset or behavioural disturbances.

5.2 Implications and future research

These findings have implications for those seeking to delineate the genetic underpinnings of AD. In particular, they underscore the utility of using covariates in genetic analysis to identify potentially homogeneous subsets of the disease. Further, they highlight the possibility that the clinical presentation of AD may be genetically modified.

This has particular implications as gaining a better understanding of the factors associated with behavioural symptoms and age at disease onset may lead to the achievable target of disease modification. Due to the late onset nature of AD, therapies which could delay the onset of symptoms, even if only briefly would have a major impact on public health. Indeed, one study has suggested that an intervention which could delay the onset of AD by just 2 years would result in nearly 2 million fewer cases of the disease in the United States over the next 50 years, whilst a delay of 5 years would reduce prevalence by a half (Brookmeyer et al. 1998). Behavioural symptoms, as discussed in section 2.1.2, are associated with many serious consequences, not only for the individual with AD (Burns et al. 1990b; Neumann et al. 2001; Rosen and Zubenko 1991; Spalletta et al. 2004), but also in terms of additional caregiver distress (Kaufer et al. 1998) and the impact on the economy (Steele et al. 1990). As such, the development of effective therapies to control these symptoms is of great importance. The identification of genes associated with the processes of disease development in AD is likely to offer useful insights into the aetiology of the clinical differences observed between patients. This may facilitate the development of therapeutic interventions aiming to modify disease presentation. Furthermore, it may serve as a means of identifying those at high risk of developing certain symptoms, which could assist the appropriate management of the illness by caregivers and clinicians.

It is likely that the analysis of disease sub-phenotypes will lead to the identification of loci which only account for a relatively small proportion AD cases. However, it remains likely that their identification would greatly advance our understanding of the pathophysiological disease process in AD as a whole. The findings from this thesis could be used to inform future genetic association studies aiming to find risk alleles which influence the clinical presentation of AD, and also for genes which increase disease susceptibility as a whole.

A number of studies have already been reported which have sought to identify genes associated with behavioural symptoms in AD. The most popular approach has been to search for genetic variation that *modifies* the clinical presentation of AD, rather than for genes which increase susceptibility to a disease sub-phenotype characterised by specific clinical features. As such they generally

compare the effect of genetic variation in samples of AD sufferers with and without clinically distinguishing features (e.g. AD cases with and without psychosis). Researchers have focused on genes that are associated with AD (e.g. APOE) or other psychiatric illnesses (e.g. dopamine and serotonin receptor genes).

The APOE ε4 allele is associated with more rapid disease progression, greater amyloid and neurofibrillary tangle burden, increased atrophy of the temporal lobes and more profound cholinergic loss in the frontal cortex (Gomez-Isla et al. 1996; Kanai et al. 1999; Lehtovirta et al. 1996b; Soininen et al. 1995). Behavioural symptoms are also associated with numerous neuropathological changes (Cummings 2000) and cholinergic deficits (Garcia-Alloza et al. 2005). As such, it is reasonable to hypothesise that APOE may play a role in the aetiology of behavioural disturbances in AD. In chapter 2 of this thesis none of the behavioural components identified through principal components analysis were associated with the APOE ε4 allele (see section 2.3.3)

Other studies which have investigated the effect of APOE on behavioural symptoms have generally reported contradictory results. A number of studies have reported an association between psychosis in AD and APOE (Cacabelos et al. 1997; Forsell et al. 1998; Ramachandran et al. 1996). Many of these have relied on relatively small samples and have reported trends, rather than statistically significant differences (Cacabelos et al. 1997; Forsell et al. 1998). However, an association between APOE and psychotic symptoms has been reported in longitudinally assessed (Scarmeas et al. 2002) and large cross sectional samples (Harwood et al. 1999). Despite these findings the majority of studies have not reported a relationship between APOE and psychosis. These studies have generally relied on samples of less than 300 AD sufferers (Hirono et al. 1999; Lehtovirta et al. 1996a; Lopez et al. 1997; Lyketsos et al. 1997; Nacmias et al. 2001). Levy and colleagues (1999) used a much larger sample of 605 AD sufferers and did not find an association between APOE and psychotic symptoms. However, their sample was recruited to test the effect of cholinesterase inhibitors on AD progression and as such was heavily biased toward those in the earlier stages of dementia. It is possible that APOE increases the susceptibility to psychotic symptoms expressed in the later stages of the disease. Indeed,

Harwood and colleagues (1999) noted that APOE had a stronger effect on psychotic symptoms in those in the severe stages of disease development.

There is some evidence to suggest that APOE is associated with the development of depressive illness in the absence of dementia, however the findings to date are somewhat contradictory (Holmes et al. 1998b; Krishnan et al. 1996; Rigaud et al. 2001; Zubenko et al. 1996). As such one may speculate that APOE is implicated in the development of depressive symptoms in AD. Indeed, such an association has been noted in a number of studies (Ramachandran et al. 1996). However, these findings have not replicated in larger samples (Holmes et al. 1996). Others have reported an association between the APOE ε2 allele and depression in AD (Holmes et al. 1996), however, numerous other studies have not noted this effect (Craig et al. 2005b; Harwood et al. 1999; Hirono et al. 1999; Levy et al. 1999; Lopez et al. 1997; Lyketsos et al. 1997; Muller-Thomsen et al. 2002).

In general, the evidence for an association between APOE and behavioural symptoms is limited and hindered by a lack of consistent findings. In addition to APOE, a number of studies have focused on genes which are implicated in the development of other psychiatric conditions, such as schizophrenia. As such, the most intensively studies genes in relation to behavioural symptoms in AD are dopamine (DRD) and serotonin (5-HT) receptor genes. Numerous polymorphisms in DRD1, DRD2, DRD3 and DRD4 have been studied extensively in relation to neuropsychiatric illnesses (Sweet et al. 1998). The relationship between DRD genes and schizophrenia is complicated, with numerous contradictory findings. Meta-analyses and large, multicenter studies appear to support an association between polymorphisms in the DRD2, DRD3 and DRD4 genes with schizophrenia (Abdolmaleky et al. 2005). Sweet and colleagues (1998) were the first to report a relationship between behavioural symptoms in AD and variation within the dopamine receptor genes, reporting a significant excess of DRD1 *B2/B2* homozygotes among those with symptoms of aggression and psychosis. They also reported that psychotic symptoms, but not aggression, occurred more frequently among patients who were homozygous for either DRD3 allele. These findings have been replicated in an independent sample (Holmes et al. 2001). More recently, Craig and colleagues (2005c) have performed analysis of the

DRD3 polymorphism using a much larger sample of 416 individuals with LOAD. In contrast to previous findings they reported that DRD3 was not associated with the presence of either delusions or hallucinations. It is difficult to draw conclusions from these three studies. It would appear that DRD1 is implicated in the modification of behavioural symptoms in AD. However, the role of DRD3 is less clear, with current research suggesting that its role in symptom development is likely to be limited.

Others have sought to investigate common polymorphisms in the 5-HT_{2A} and 5-HT_{2c} serotonin receptor genes (102 T/C and Cys23Ser polymorphisms, respectively) in relation to behavioural symptoms observed in AD. These polymorphisms have been implicated the aetiology of a number of psychiatric conditions, including psychosis (Abdolmaleky et al. 2005; Gutierrez et al. 1996; Williams et al. 1997), bipolar affective disorders (Arranz et al. 1997), and eating disorders (Nacmias et al. 1999). Reduced serotonergic activities have also been linked to depression and psychosis among dementia sufferers (Chen et al. 1996; Lanctot et al. 2001), whilst, aggressive symptoms in AD have been associated with 5-HT_{2A} receptor loss at post-mortem (Procter et al. 1992). Despite not showing an association with disease risk as a whole, polymorphisms in the 5-HT_{2A} and 5-HT_{2c} serotonin genes have been shown to increase risk of developing psychotic symptoms during the course of AD (Assal et al. 2004; Holmes et al. 1998a; Lam et al. 2004; Nacmias et al. 2001). However, it should be noted that the data is somewhat contradictory, with two studies reporting the C102 allele increases risk of developing psychotic symptoms (Holmes et al. 1998a; Nacmias et al. 2001), whilst others have found the T102 allele to confer increased risk (Assal et al. 2004; Lam et al. 2004). In addition, others have reported that the long variant of an insertion/deletion polymorphism in the promoter region of the serotonin transporter gene (5-HTTPR) is associated with an increased risk of developing aggressive and psychotic symptoms in AD (Sukonick et al. 2001). However, Assal and colleagues (2001) failed to replicate these findings in an independent sample.

Others have hypothesised that the catechol-O-methyltransferase (COMT) gene may be associated with the presence of behavioural symptoms. COMT resides on

chromosome 22q11.2 and has been postulated as a functional candidate gene for schizophrenia (Glatt et al. 2003) due to is role in encoding a protein that enzymatically inactivates dopamine (Axelrod and Tomchick 1958). A number of authors have investigated the relationship between psychotic symptoms in AD and COMT (Borroni et al. 2004; Sweet et al. 2005). Borroni and colleagues reported that the widely studied G to A substitution in codon 108/158 of COMT (RS4680) was associated with the development of psychotic symptoms, despite not influencing risk for AD as a whole. Following this report, Sweet and colleagues (2005) have reported an investigation of four polymorphisms, namely RS4680, ERE6, RS737865 and RS165599. They reported a significant association between the RS4680 polymorphism and psychosis among females, but not males. In haplotype analysis they reported a highly significant association of a four locus haplotype with psychotic symptoms in AD among females, which appeared to result from the additive effects of alleles at ERE6 and RS737865 which were in very high linkage disequilibrium.

Craig and colleagues (2004a) recently reported that the tryptophan hydroxylase (TPH) gene, located on chromosome 11, is associated with aggression in AD. A bi-allelic polymorphism in intron 7 of TPH has previously shown association with both suicidality (Bellivier et al. 2004) and aggressive behaviour (Manuck et al. 1999). However, the effect reported by Craig and colleagues was largely confined to males, suggesting that TPH may have gender specific effects. An earlier report found no association between TPH and AD (Wang et al. 2001), suggesting that it may modify susceptibility to aggressive symptoms after disease onset. Further to this, Craig and colleagues (2004b) have also reported an association between psychosis and genetic variation within the interleukin 1β (IL-1β) gene, located on chromosome 2q. They focused on a polymorphism in the promoter region which has previously shown association with schizophrenia (Katila et al. 1999; Zanardini et al. 2003), but not AD (Ehl et al. 2003). Using a sample of 424 probable AD sufferers they reported a significant association with IL-1β and the presence of hallucinations and delusions separately, and with a combined phenotype of delusions or hallucinations. These findings are of particular interest as cytokines such as IL-1β have been reported to influence levels of numerous neurotransmitters, such as serotonin and dopamine (Wichers and Maes 2002),

which have been implicated in the development of psychosis in AD (Garcia-Alloza et al. 2005). They also support the hypothesis that common aetiological pathways exist in AD with psychosis and schizophrenia (Sweet et al. 2003).

It is clear from the research to date that the inclusion of clinical covariates in genetic association studies is likely to increase our ability to identify loci implicated in the development, and progression, of Alzheimer's disease. Current association based investigations of behavioural problems in AD support the findings from chapter 3 of this thesis, which suggest that the variation commonly observed among AD sufferers is, at least in part, genetically influenced. To date studies have largely focused on candidate genes for other major psychiatric disorders, such as schizophrenia. It is notable that none of the genes which have already shown association with behavioural symptoms are in regions of linkage identified in chapter 4 of this thesis. Analysis of genes in these regions is likely to reveal further loci which are implicated in the processes of disease development in AD. To date, studies have tended to view behavioural symptoms in isolation. More success may be obtained by looking at behavioural components, rather than individual symptoms. In addition, the majority of studies have investigated the relationship between genetic variation and psychotic symptoms in AD, with a lesser number focusing on symptoms of depression and aggression. Symptoms of anxiety, disinhibition and euphoria, and disturbances in circadian rhythm and appetite have been largely overlooked.

Together with the near completion of the human genome project, the ever increasing efficiency and lowering costs of high throughput genotyping, ensure that the coming years will undoubtedly be a productive time for those investigating the genetics of complex disorders. The findings from this programme of research serve to highlight the utility of including clinical covariates in genetic analyses of AD, providing adequate justification for future studies with the a priori goal of identifying genetic variation associated with clinical heterogeneity. They also have implications for prospective large scale collections of samples intended for use in genetic studies of AD as a whole. Such studies are likely to benefit from comprehensively characterising the disease phenotype, using established

measurement tools. Such data is likely to prove useful in secondary analyses of these datasets.

Future studies may seek to identify other distinguishing factors of AD that may be useful for delineating more homogenous disease sub-phenotypes. For example, a large proportion of those with AD also have substantial co-morbid cerebrovasucular pathology (Humpel and Marksteiner 2005). AD with less cerebrovasucular disease may represent a 'purer form' of the illness which may be useful for genetic analyses. Alternatively, researchers may seek to draw on "endophenotypes", or measurable components of the illness, which lie along the pathological pathway of the disease (Gottesman and Gould 2003). These can take the form of neurophysiological, biochemical, neuroanatomical, cognitive and neuropsychological measures (Leboyer et al. 1998). In comparison to disease status as a whole, such intermediate phenotypes may be more closely linked to the biological function of genes. Of particular interest maybe the use of biological markers obtained through use of advanced neuroimaging, such as magnetic resonance imaging (MRI) or single photon emission computed tomography (SPECT).

As pointed out by Lovestone and Hardy (2002) there is now an urgent need to replicate findings in relation to clinical variation in independent, well characterised samples. The NIMH families used in this thesis has been used extensively to perform linkage analyses using numerous covariates including AAO (Holmans et al. 2005; Li et al. 2002; Olson et al. 2001, 2002), rate of decline (Holmans et al. 2005) and psychosis (Avramopoulos et al. 2005; Bacanu et al. 2002). The findings from this thesis offer promise to those searching for genes implicated with age at disease onset, psychosis and to a lesser extent aggression and mood disturbances in AD. As such, there is now a growing need, not only for further prospective association studies, but also for the current linkage findings to be replicated in independent family-based samples.

5.3 Conclusions

Genetic variation is likely to play a role in the clinical heterogeneity commonly observed in AD. This observation supports the view that AD is a syndrome with multiple contributing aetiologies. Linkage analysis has identified numerous regions which are likely to contain genes which influence the development of behavioural disturbances occurring as part of the illness and the age at disease onset. Further study is required to genetically dissect these regions. The identification of genes associated with clinical features of AD, is likely to increase our understanding of the disease as a whole, which could ultimately lead to the development of more effective therapeutic interventions.

Appendices

2a *Rotated component loadings after controlling for disease severity (GDS)*

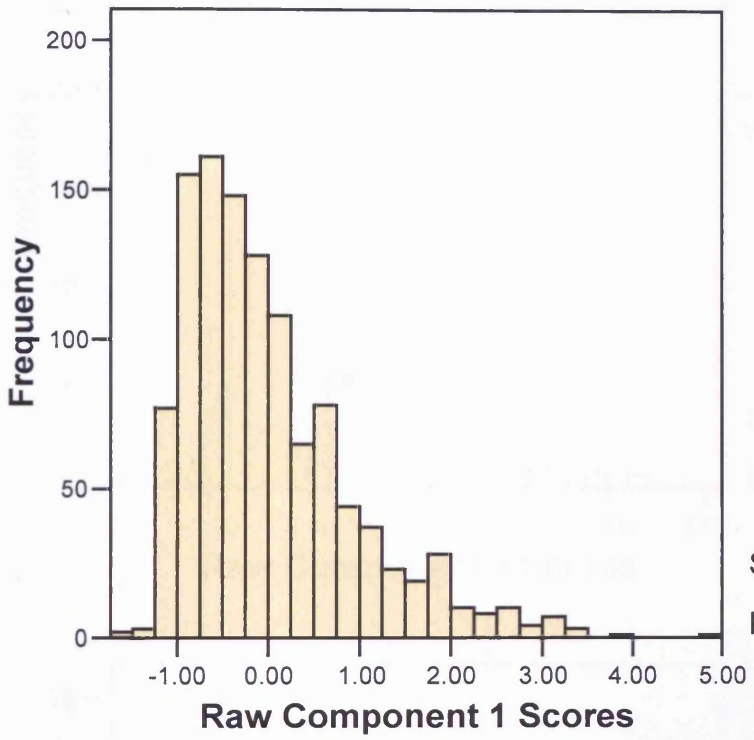
Component loadings for the NPI after controlling for GDS and years in education the full sample (n=1120).

Item	Component 1	Component 2	Component 3
Delusions	.	0.78	.
Hallucinations	.	0.78	.
Agitation/Aggression	0.66	.	.
Depression/Dysphoria	.	.	0.82
Anxiety	.	.	0.74
Euphoria	0.63	.	.
Apathy	.	.	0.42
Disinhibition	0.74	.	.
Irritability	0.67	.	.
Aberrant Motor Behaviour	0.52	.	.
Sleep disturbances	0.44	.	.
Appetite abnormalities	0.33	.	.
Eigenvalue	2.89	1.14	1.07
% of Variance	24.11	9.54	8.94

NB: Values less than 0.3 are not shown

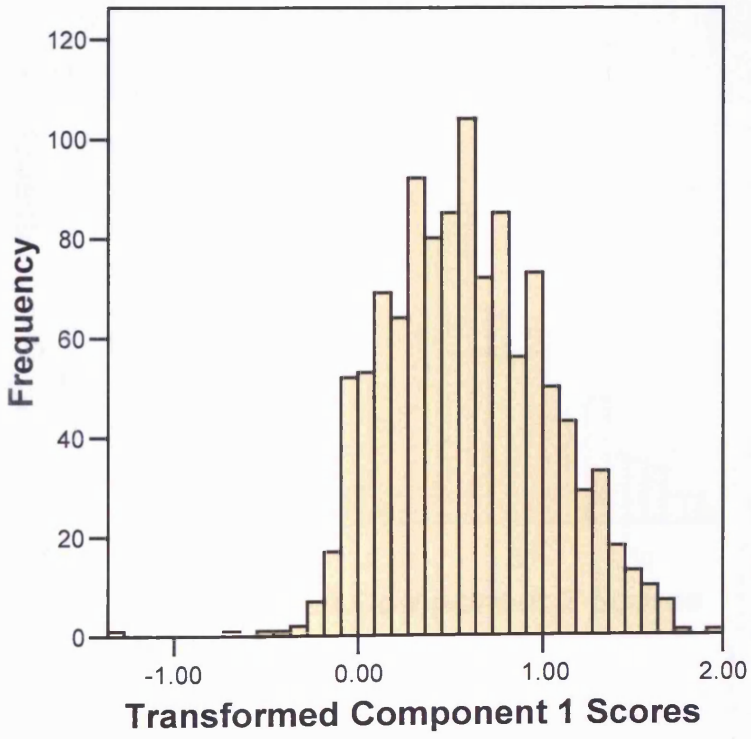
2b Distribution of component scores before and after log transformations

Component 1 - "Frontal lobe dysfunction"



Skewness: 1.297

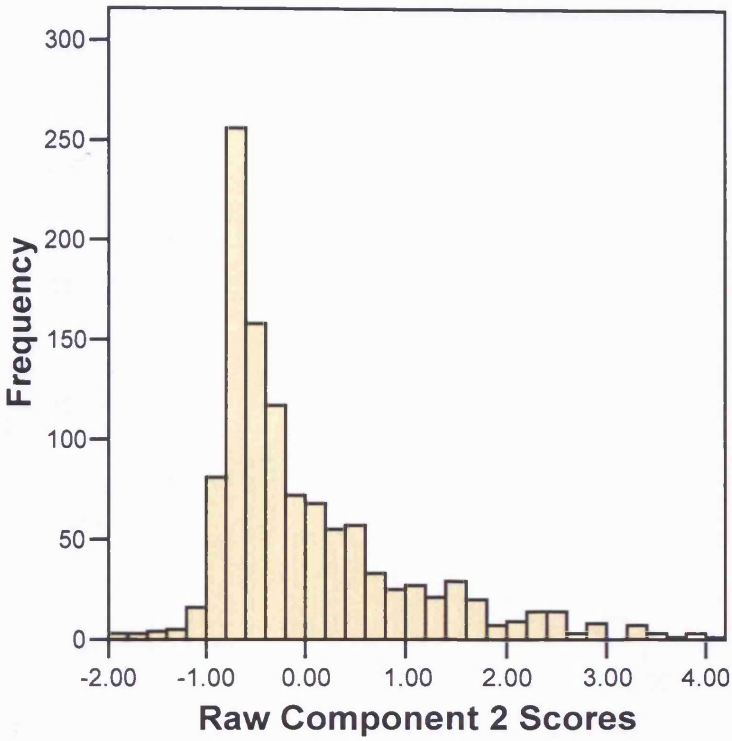
Kurtosis: 1.885



Skewness: 0.192

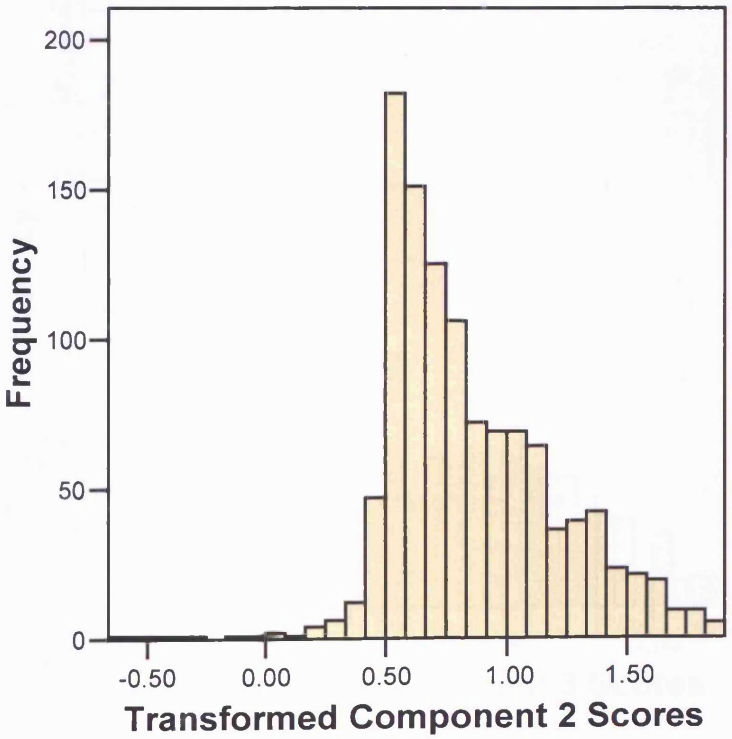
Kurtosis: -0.133

Component 2 - "Psychosis"



Skewness: 1.495

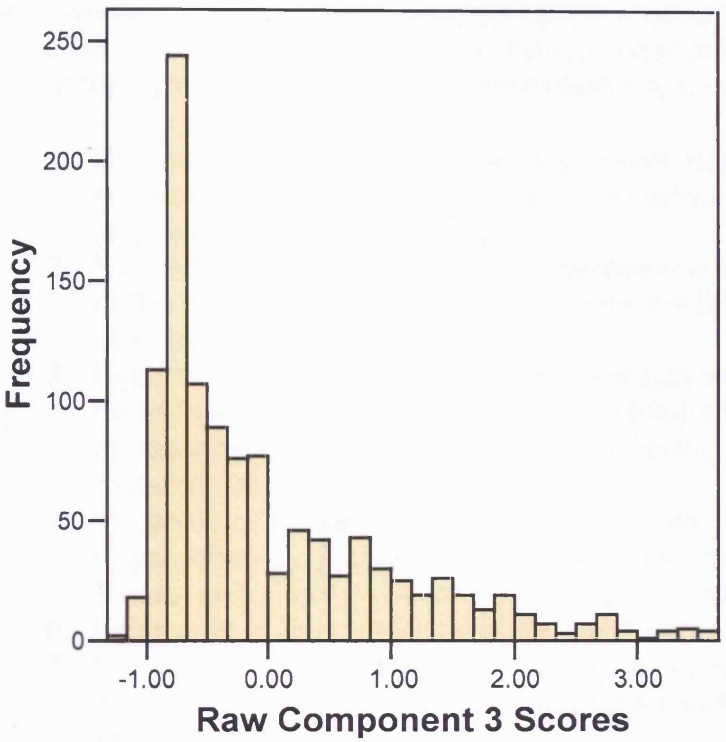
Kurtosis: 2.154



Skewness: 0.500

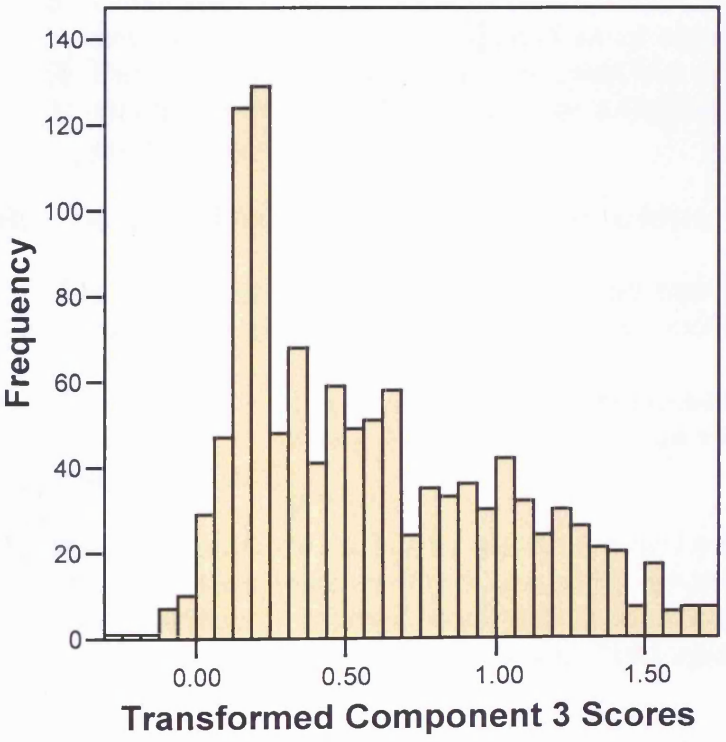
Kurtosis: 0.728

Component 3 - "Mood"



Skewness: 1.351

Kurtosis: 1.309



Skewness: 0.627

Kurtosis: -0.616

3a DSM-IV criteria for a major depressive episode

A. Five (or more) of the following symptoms have been present during the same 2-week period and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure.

1. Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad or empty) or observation made by others (e.g., appears tearful)
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation made by others)
3. Significant weight loss when not dieting or weight gain (e.g. a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day. Note: in children, consider failure to make expected weight gains.
4. Insomnia or hypersomnia nearly every day
5. Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness of being slowed down)
6. Fatigue or loss of energy nearly every day
7. feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick)
8. Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others)
9. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

B. The symptoms do not meet criteria for a Mixed Episode.

C. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.

D. The symptoms are not due to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition (e.g., hypothyroidism).

E. The symptoms are not better accounted for by Bereavement, i.e., after the loss of a loved one, the symptoms persist for longer than 2 months or are characterized by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation.

Taken from the "Diagnostic and statistical manual of mental disorders: DSM-IV" (American Psychiatric Association 1994)

3b Cornell scale for depression in dementia

Scoring: 0 absent 1 mild or intermittent 2 severe

Have you noticed him/her showing an anxious expression, ruminating, or worrying?	0	1	2
Showing a sad expression, sad voice, or being tearful?	0	1	2
Not reacting as positively as usual to pleasant events?	0	1	2
Being easily annoyed or short tempered?	0	1	2
Showing restlessness, hand wringing, or hair pulling?	0	1	2
Slowing of movements, slow speech, slow reactions	0	1	2
Complaining of many physical problems?	0	1	2
To be less involved in usual activities?	0	1	2
To be eating less than usual?	0	1	2
To show weight loss?	0	1	2
Fatiguing easily or being unable to sustain activities?	0	1	2
Being generally worse in the morning, in terms of mood or behaviour?	0	1	2
Falling asleep later than usual?	0	1	2
Waking many times during the night?	0	1	2
Waking early in the morning?	0	1	2
Saying he/she feels life is not worth living, talking of suicidal wishes, or making a suicide attempt?	0	1	2
Talking about blaming himself/herself, or talking about poor self-esteem or feelings of failure?	0	1	2
Appearing to expect the worst?	0	1	2
Speaking as though convinced he/she suffers from poverty, physical illness or loss, when this is clearly not the case?	0	1	2

3c DSM-IV criteria for generalised anxiety disorder

A. Excessive anxiety and worry (apprehensive expectation), occurring more days than not for at least 6 months, about a number of events or activities (such as work or school performance).

B. The person finds it difficult to control the worry.

C. The anxiety and worry are associated with three (or more) of the following six symptoms (with at least some symptoms present for more days than not for the past 6 months). Note: Only one item is required in children.

    1. Restlessness or feeling keyed up or on edge
    2. Being easily fatigued
    3. Difficulty concentrating or mind going blank
    4. Irritability
    5. Muscle tension
    6. Sleep disturbance (difficulty falling or staying asleep, or restless unsatisfying sleep)

D. The focus of the anxiety and worry is not confined to features of an Axis I disorder, for example, the anxiety or worry is not about having a Panic Attack (as in Panic Disorder), being embarrassed in public (as in Social Phobia), being contaminated (as in Obsessive-Compulsive Disorder), being away from home or close relatives (as in Separation Anxiety Disorder), gaining weight (as in Anorexia Nervosa), having multiple physical complaints (as in Somatization Disorder), or having a serious illness (as in Hypochondriasis), and the anxiety and worry do not occur exclusively during Posttraumatic Stress Disorder.

E. The anxiety, worry, or physical symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.

F. The disturbance is not due to the direct physiological effects of a substance (for example, a drug of abuse, a medication) or a general medical condition (for example, hyperthyroidism) and does not occur exclusively during a Mood Disorder, a Psychotic Disorder, or a Pervasive Developmental Disorder.

Taken from the "Diagnostic and statistical manual of mental disorders: DSM-IV" (American Psychiatric Association 1994)

4a Chromosomal regions genotyped in the UK and NIA samples.

Summary of chromosomal regions genotyped in the UK and NIA samples

Chromosome	Number of markers genotyped in UK+NIA sample	Chromosomal region genotyped in NIMH Sample (cM)	Region(s) genotyped in the UK and NIA samples (cM)
1	19	8.9 - 274.5	10.8 - 65.5, 125.5 -161.1
2	13	3.8 - 251.9	7.6 - 103.2
5	10	0.0 - 195.5	7.8 - 64.7
6	7	9.2 - 187.2	53.8 - 88.6
9	18	0.0 - 163.8	29.5 - 142.5
10	18	4.3 - 170.9	59.0 - 135.2
12	9	6.4 - 165.7	6.4 - 50.9
13	3	8.9 - 110.6	67.7 -82.9
14	8	9.4 - 138.2	9.4 - 66.8
19	7	0.0 - 100.6	56.7 - 100.0
21	11	3.0 - 57.8	3.0 - 57.8

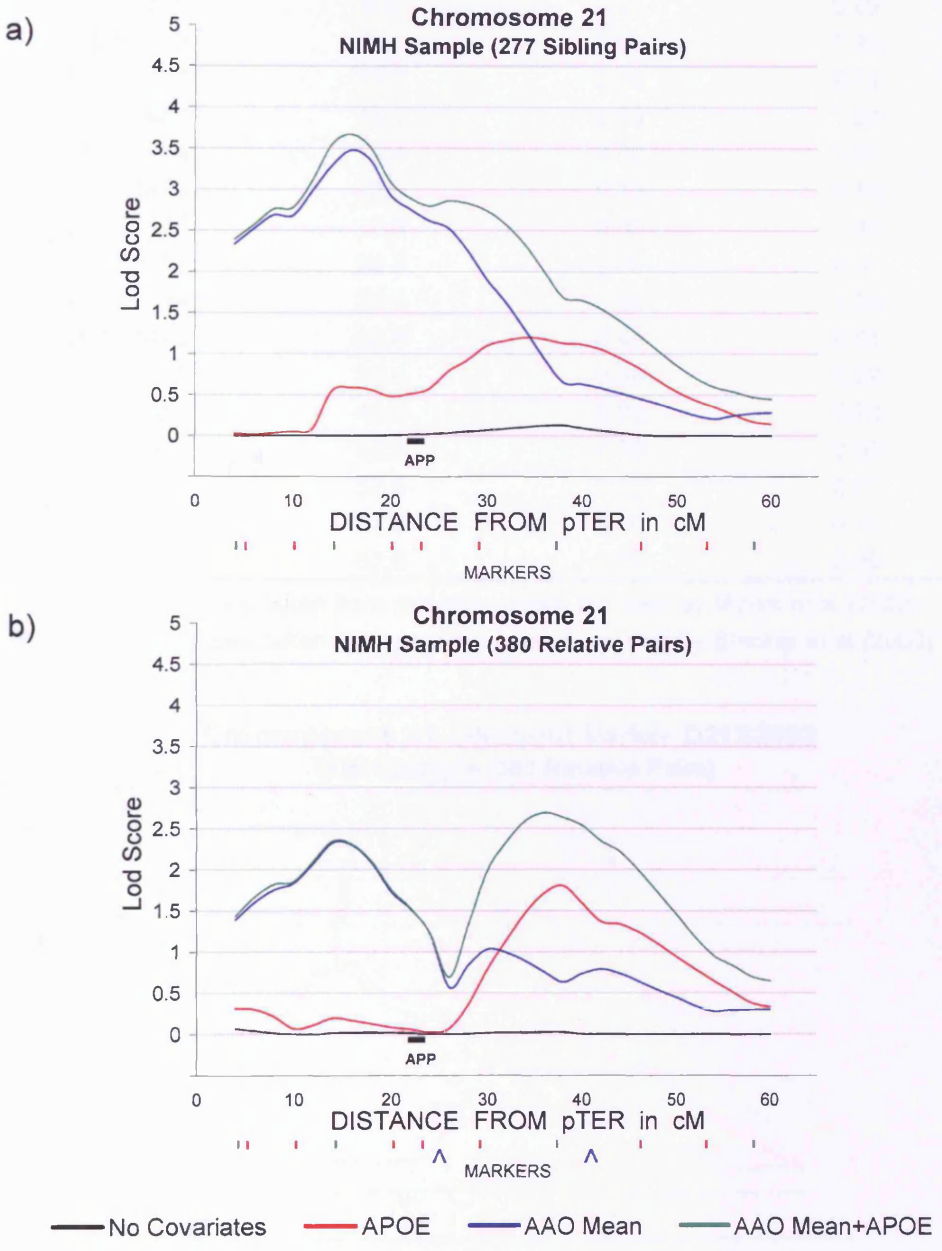
NB: Chromosomes that are not shown were not genotyped in the UK or NIA samples

4b *Summary table for linkage analyses including Age at Onset and APOE for all chromosomes.*

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Mean AAO	Incr. Diff AAO	Incr. APOE	Incr. Mean | APOE	Incr. Diff | APOE
1	Full	521	0.73	1.58 (+)	0.07	1.06	1.28 (+)	0.47
1	NIMH	380	0.87	2.00 (+)	0.00	1.04	1.48 (+)	0.06
1	UK+NIA	141	0.54	0.17	0.38	1.38	0.13	0.02
2	Full	511	0.37	1.33 (+)	1.37	0.74	1.37 (+)	0.80
2	NIMH	380	0.37	1.33 (+)	1.01	0.74	1.37 (+)	0.50
2	UK+NIA	131	0.06	0.16	0.99	0.36	0.18	1.16
3	NIMH	379	0.77	0.83	0.27	0.54	0.40	0.45
4	NIMH	379	0.40	0.49	0.13	0.94	0.02	0.00
5	Full	511	1.31	0.01	0.00	0.42	0.08	0.00
5	NIMH	380	1.36	0.02	0.00	0.60	0.00	0.00
5	UK+NIA	131	0.31	0.02	0.07	0.62	0.00	0.00
6	Full	523	0.72	0.30	0.93	0.80	0.04	0.26
6	NIMH	380	0.48	0.24	1.34	1.38	0.10	0.15
6	UK+NIA	143	1.15	0.14	0.03	0.22	0.01	0.04
7	NIMH	380	0.38	0.24	0.98	0.66	0.23	0.72
8	NIMH	379	0.75	0.02	0.59	0.82	0.00	0.50
9	Full	532	2.32	0.03	0.00	0.07	0.07	0.00
9	NIMH	380	3.32	0.24	0.00	0.14	0.32	0.00
9	UK+NIA	152	1.05	0.01	0.00	0.47	0.03	0.00
10	Full	523	3.44	0.01	0.00	0.76	0.00	0.02
10	NIMH	380	2.17	0.21	0.05	0.51	0.15	0.21
10	UK+NIA	143	1.42	0.34	0.00	0.21	0.28	0.00
11	NIMH	379	0.86	0.31	0.43	0.70	0.01	0.07
12	Full	523	0.13	0.24	0.79	1.00	0.03	0.51
12	NIMH	380	0.04	0.32	0.88	0.81	0.13	0.21
12	UK+NIA	143	0.65	0.14	1.71	0.06	0.14	2.12
13	Full	511	0.11	0.20	0.07	1.01	0.00	0.00
13	NIMH	380	0.19	0.23	0.12	0.89	0.00	0.00
13	UK+NIA	131	0.00	0.07	0.07	0.17	0.01	0.14
14	Full	528	0.03	0.47	1.18	0.39	0.19	2.14
14	NIMH	380	0.33	0.87	0.74	0.34	0.63	0.98
14	UK+NIA	148	0.03	0.24	0.30	0.85	0.06	0.85
15	NIMH	380	0.07	0.53	1.16	0.13	0.43	1.22
16	NIMH	380	0.11	0.47	0.30	0.51	0.26	0.04
17	NIMH	380	1.13	0.17	0.00	0.25	0.20	0.00
18	NIMH	380	0.16	0.23	1.64	0.66	0.21	1.01
19	Full	529	1.47	0.48	1.36	-	-	-
19	NIMH	380	0.95	0.34	1.02	-	-	-
19	UK+NIA	149	1.41	0.73	0.66	-	-	-
20	NIMH	379	0.35	0.70	0.00	0.33	0.52	0.02
21	Full	521	0.39	1.45 (+)	0.00	1.35	0.95 (+)	0.00
21	NIMH	380	0.07	2.26 (+)	0.00	1.74	0.89 (+)	0.00
21	UK+NIA	141	0.99	0.00	0.00	0.69	0.05	0.00
22	NIMH	379	0.50	0.42	0.37	0.61	0.45	0.56

The six columns of maximum mulitpoint LOD scores refer to (left to right):
1) The maximum mulitpoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) mean AAO, 3) AAO difference, 4) APOE genotype. The increase in MLS given by including 5) mean AAO and 6) AAO difference *after allowing for APOE effects.*
(+) in the mean AAO column indicates that IBD sharing increased with increasing AAO, (-) indicates increased IBD sharing in pairs with lower mean AAO
NB: Genotypes were not available in the UK + NIA sample for chromosomes where only NIMH results are shown

4c *Supplementary analyses of chromosome 21 incorporating age at onset data*



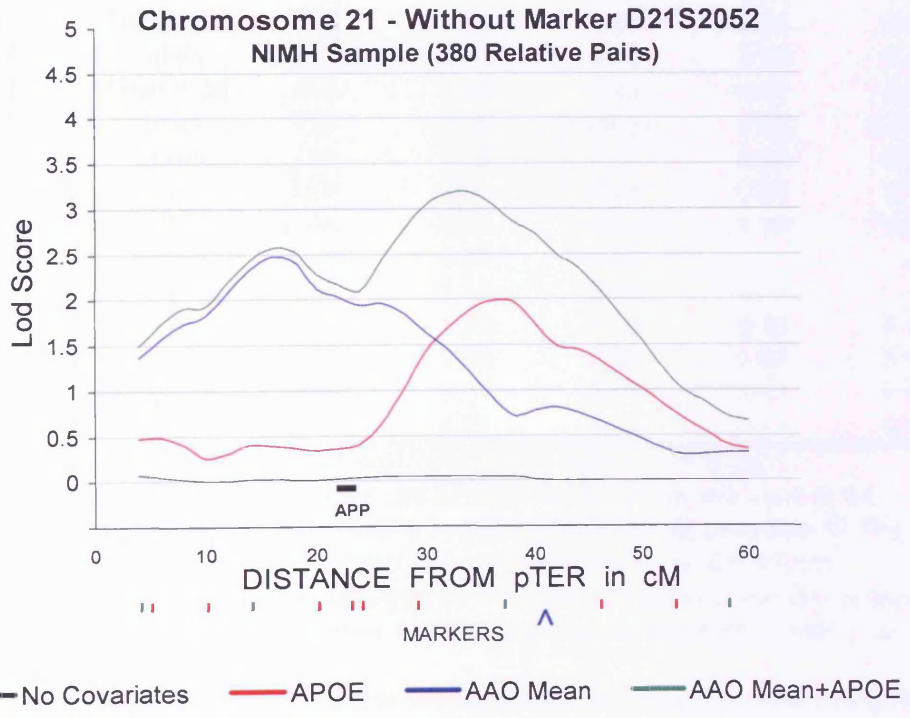
Maximum mulitpoint LOD score graphs for chromosome 21. Figure a) shows the analysis performed using using sibling pairs and marker genotypes analysed by Holmans and colleagues. Figure b) shows the analysis after incorportating additional families, individuals, relative pairs and marker genotypes. Markers highlighted in green and red were common to both analyses. However, additional marker genotypes were available for markers highlighted in green which were not used by Holmans and colleagues. The two markers highlighed in blue were not included in the analysis reported by Holmans and colleagues.

Two point LOD scores for chromosome 21 including mean AAO for the NIMH sample

Marker	cM	No Covariates	Incr. Mean AAO
D21S1432 [1]	3.0	0.17	0.73
D21S1432 [2]	3.0	0.11	0.63
D21S1201 [1]	3.0	0.27	0.99
D21S1899 [1]	9.7	0.02	1.23
D21S1437 [1]	13.1	0.00	2.19
D21S1437 [2]	13.1	0.13	1.24
D21S1914 [1]	19.4	0.00	1.62
D21S1435 [1]	22.6	0.15	0.83
D21S217 [1]	22.8	0.00	1.42
D21S2052 [2]	24.7	0.00	0.31
D21S1909 [1]	28.5	0.04	1.26
D21S1440 [1]	36.8	0.44	0.01
D21S1440 [2]	36.8	0.38	0.22
D21S2055 [2]	40.5	0.05	0.73
D21S266 [1]	45.9	0.00	0.06
D21S1885 [1]	52.5	0.00	0.01
D21S1446 [1]	57.8	0.00	0.11
D21S1446 [2]	57.8	0.00	0.09

[1] Marker genotypes taken from genome screen reported by Myers et al (2002)

[2] Marker genotypes taken from genome screen reported by Blacker et al (2003)



Multipoint LOD score graphs for chromosome 21, including AAO mean and APOE as covariates, after removing marker D21S2052.

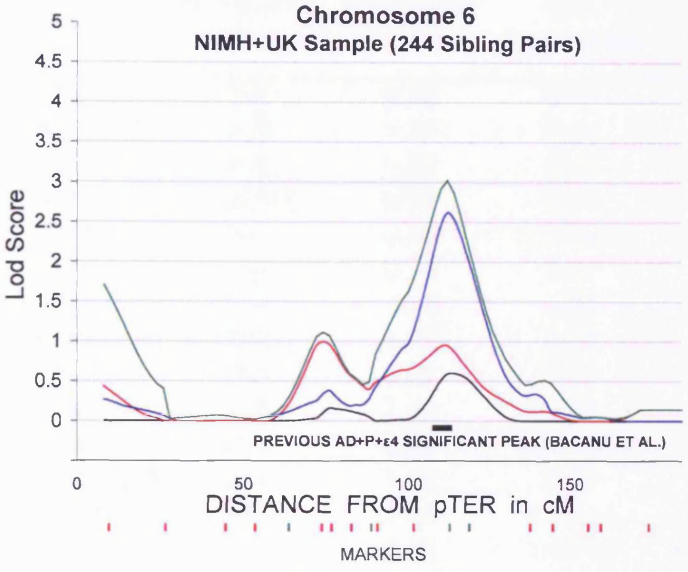
4d Summary table for linkage analyses including psychosis and APOE for all chromosomes.

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Psychosis	Incr. APOE	Incr. Psychosis | APOE
1	NIMH	266	1.11	0.60	1.34	0.48
1	NIMH+UK	303	1.30	0.84	0.75	0.36
2	NIMH	266	0.30	2.12 (+)	0.35	1.90 (+)
2	NIMH+UK	303	0.45	1.24 (+)	0.34	0.99 (+)
3	NIMH	265	0.35	0.54	0.72	0.04
4	NIMH	265	0.25	0.45	0.41	0.20
5	NIMH	266	0.82	0.47	0.09	0.69
5	NIMH+UK	303	0.53	0.76	0.07	1.00
6	NIMH	266	1.03	0.53	1.04	0.71
6	NIMH+UK	303	1.09	0.58	1.11	0.77
7	NIMH	266	0.60	2.25 (+/-)	0.32	2.00 (+/-)
8	NIMH	265	1.40	0.10	0.45	0.13
9	NIMH	266	1.98	0.47	0.01	0.47
9	NIMH+UK	307	1.31	0.71	0.07	0.64
10	NIMH	266	0.96	0.39	0.21	0.18
10	NIMH+UK	303	1.32	0.54	0.15	0.39
11	NIMH	265	0.91	0.10	1.44	0.09
12	NIMH	266	0.25	0.26	1.61	0.25
12	NIMH+UK	303	0.30	0.32	1.09	0.20
13	NIMH	266	0.89	0.06	0.34	0.06
13	NIMH+UK	303	0.91	0.07	0.35	0.06
14	NIMH	266	0.39	0.48	0.12	0.36
14	NIMH+UK	306	0.38	0.49	0.11	0.38
15	NIMH	266	0.20	2.6 (-)	0.08	2.54 (-)
16	NIMH	266	0.15	0.57	0.49	0.05
17	NIMH	266	0.41	0.21	0.52	0.17
18	NIMH	266	0.02	1.49 (-)	1.14	1.23 (-)
19	NIMH	266	1.34	0.56	-	-
19	NIMH+UK	307	1.00	0.65	-	-
20	NIMH	265	1.38	0.18	0.30	0.19
21	NIMH	266	0.05	0.25	3.87	0.08
21	NIMH+UK	304	0.11	0.41	3.63	0.36
22	NIMH	265	0.27	0.26	0.71	0.17

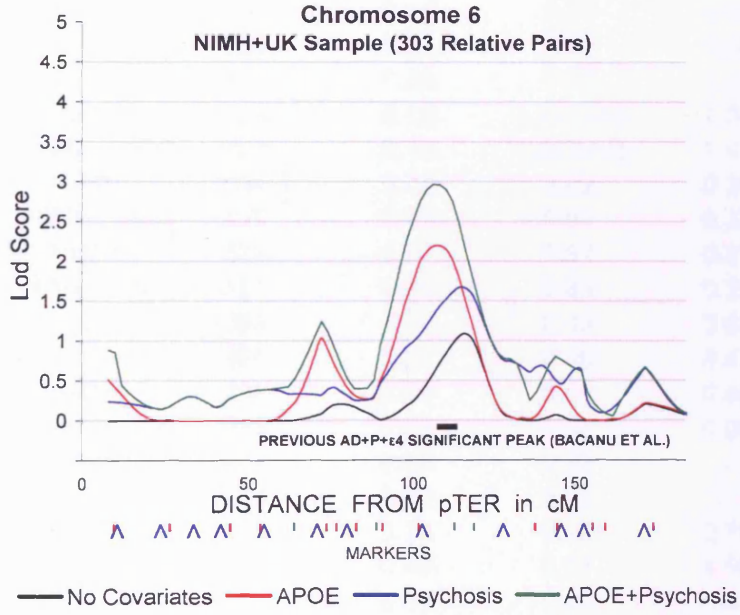
The four columns of maximum mulitpoint LOD scores refer to (left to right):
1) The maximum multipoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) psychosis, 3) APOE genotype, 4) The increase in MLS given by including psychosis *after allowing for APOE effects.*
(+) indicates that IBD sharing was elevated in +/+ pairs, (-) indicates that IBD sharing was elevated in -/- pairs, (+/-) indicates that IBD sharing was elevated in both -/- and +/+ pairs (compared to -/+ pairs)
NB: Genotypes were not available in the UK sample for chromosomes where only NIMH results are shown

4e *Supplementary linkage analyses of chromosome 6 incorporating psychosis data*

a)



b)



Maximum mulitpoint LOD score graphs for chromosome 6. Figure a) shows the analysis performed using sibling pairs and marker genotypes from the genome screen presented by Myers and colleagues (Myers et al. 2002). Figure b) shows the analysis after incorportating additional families, individuals, relative pairs and marker genotypes. Markers highlighted in green and red were common to both analyses. However, additional marker genotypes were available for markers highlighted in green which were not used in analysis (a). Markers highlighed in blue were not included in the analysis reported by Myers and colleauges.

4f Summary table for linkage analyses including aggression and APOE for all chromosomes.

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Aggression	Incr. APOE	Incr. Aggression | APOE
1	NIMH	204	0.90	0.44	0.28	0.16
1	NIMH+UK	229	1.24	0.46	0.27	0.19
2	NIMH	204	0.32	0.49	0.35	0.33
2	NIMH+UK	229	0.57	0.59	0.20	0.20
3	NIMH	204	0.55	1.57 (+/-)	0.70	0.91 (+/-)
4	NIMH	204	0.19	0.27	0.42	0.23
5	NIMH	204	0.29	0.21	0.65	0.22
5	NIMH+UK	229	0.20	0.55	0.74	0.23
6	NIMH	204	0.97	0.31	1.67	1.20 (+/-)
6	NIMH+UK	229	1.00	0.36	1.60	1.30
7	NIMH	204	0.27	1.71 (+/-)	0.34	1.43 (+/-)
8	NIMH	204	1.02	0.81	0.32	0.86
9	NIMH	204	2.12	1.53 (-)	0.22	2.03 (-)
9	NIMH+UK	232	2.18	1.43 (-)	0.06	1.85 (-)
10	NIMH	204	1.14	0.16	0.03	0.15
10	NIMH+UK	229	0.81	0.29	0.08	0.26
11	NIMH	204	0.55	0.67	1.50	0.06
12	NIMH	204	0.12	1.25 (+/-)	1.59	0.67 (+/-)
12	NIMH+UK	229	0.13	1.20 (+/-)	1.47	0.79 (+/-)
13	NIMH	204	0.09	0.62	0.34	0.71
13	NIMH+UK	229	0.09	0.81	0.35	0.75
14	NIMH	204	0.27	0.82	0.22	0.80
14	NIMH+UK	231	0.25	0.83	0.23	0.80
15	NIMH	204	0.09	0.43	0.05	0.44
16	NIMH	204	0.07	0.30	0.43	0.04
17	NIMH	204	0.59	0.05	0.46	0.00
18	NIMH	204	0.06	0.37	0.90	0.08
19	NIMH	204	0.36	0.08	-	-
19	NIMH+UK	232	0.49	0.09	-	-
20	NIMH	204	0.74	0.02	0.76	0.02
21	NIMH	204	0.49	0.21	4.59	0.04
21	NIMH+UK	230	0.63	0.04	4.41	0.01
22	NIMH	204	0.14	0.31	0.94	1.20

The four columns of maximum mulitpoint LOD scores refer to (left to right):
1) The maximum multipoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) aggression, 3) APOE genotype, 4) The increase in MLS given by including aggression *after allowing for APOE effects.*
(+) indicates that IBD sharing was elevated in +/+ pairs, (-) indicates that IBD sharing was elevated in -/- pairs, (+/-) indicates that IBD sharing was elevated in both -/- and +/+ pairs (compared to +/- pairs)
NB: Genotypes were not available in the UK sample for chromosomes where only NIMH results are shown

4g Summary table for linkage analyses including depression and APOE for all chromosomes.

Chromosome	Sample	Number of Pairs	Univariate LOD	Incr. Depression	Incr. APOE	Incr. Depression | APOE
1	NIMH	296	0.69	0.73	1.49	0.00
1	NIMH+UK	351	1.03	0.40	0.93	0.00
2	NIMH	296	0.10	0.53	0.36	0.17
2	NIMH+UK	351	0.20	0.44	0.42	0.65
3	NIMH	295	0.78	1.15 (+)	0.35	0.90
4	NIMH	295	0.37	0.31	0.39	0.09
5	NIMH	295	1.17	0.05	0.22	0.08
5	NIMH+UK	351	0.91	0.08	0.30	0.28
6	NIMH	296	0.82	0.92	0.52	0.57
6	NIMH+UK	351	0.87	0.52	0.42	0.36
7	NIMH	296	0.40	0.49	0.69	0.72
8	NIMH	295	0.64	0.60	0.32	0.84
9	NIMH	296	1.51	0.17	0.00	0.17
9	NIMH+UK	353	1.30	0.60	0.06	0.64
10	NIMH	296	2.27	1.26 (+)	0.19	1.25 (+)
10	NIMH+UK	351	2.45	0.63	0.22	0.66
11	NIMH	295	1.14	0.49	0.45	0.47
12	NIMH	296	0.28	1.03 (+/-)	0.68	0.92
12	NIMH+UK	351	0.28	1.11 (+/-)	0.78	0.91
13	NIMH	296	0.32	0.52	0.41	0.12
13	NIMH+UK	351	0.32	0.46	0.66	0.41
14	NIMH	296	0.26	0.41	0.08	0.33
14	NIMH+UK	353	0.25	0.41	0.03	0.41
15	NIMH	296	0.00	0.34	0.41	0.07
16	NIMH	296	0.09	0.32	1.23	0.00
17	NIMH	296	0.52	0.33	0.37	0.23
18	NIMH	296	0.05	0.99	1.95	0.70
19	NIMH	296	0.96	1.56 (+/-)	-	-
19	NIMH+UK	353	1.04	1.47 (+/-)	-	-
20	NIMH	295	0.63	0.88	1.13	0.73
21	NIMH	296	0.07	1.96 (+/-)	1.47	1.20
21	NIMH+UK	351	0.27	2.19 (+/-)	0.99	1.23 (+/-)
22	NIMH	294	0.24	0.85	0.62	0.42

The four columns of maximum mulitpoint LOD scores refer to (left to right):
1) The maximum multipoint LOD score (MLS) when no covariates are used in the analysis, the *increases* given by including 2) depression 3) APOE genotype, 4) The increase in MLS given by including depression *after allowing for APOE effects.*
(+) indicates that IBD sharing was elevated in +/+ pairs, (-) indicates that IBD sharing was elevated in -/- pairs, (+/-) indicates that IBD sharing was elevated in both -/- and +/+ pairs (compared to +/- pairs)
NB: Genotypes were not available in the UK sample for chromosomes where only NIMH results are shown

References

Aalten, P. et al. 2003. Behavioral problems in dementia: a factor analysis of the Neuropsychiatric inventory. *Dement Geriatr Cogn Disord* 15(2), pp. 99-105.

Abas, M. A. et al. 1990. Neuropsychological deficits and CT scan changes in elderly depressives. *Psychol Med* 20(3), pp. 507-520.

Abdolmaleky, H. M. et al. 2005. Genetics and epigenetics in major psychiatric disorders: dilemmas, achievements, applications, and future scope. *Am J Pharmacogenomics* 5(3), pp. 149-160.

Abecasis, G. R. et al. 2002. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30(1), pp. 97-101.

Alexopoulos, G. S. et al. 1988a. Cornell Scale for Depression in Dementia. *Biol Psychiatry* 23(3), pp. 271-284.

Alexopoulos, G. S. et al. 1988b. Use of the Cornell scale in nondemented patients. *J Am Geriatr Soc* 36(3), pp. 230-236.

Allen, N. H. et al. 1996. Manchester and Oxford Universities Scale for the Psychopathological Assessment of Dementia (MOUSEPAD). *Br J Psychiatry* 169(3), pp. 293-307.

Altmuller, J. et al. 2001. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 69(5), pp. 936-950.

Altmuller, J. et al. 2005. Phenotypic and genetic heterogeneity in a genome-wide linkage study of asthma families. *BMC Pulm Med* 5(1), p. 1.

Alzheimer, A. 1907. Uber eine Eigenartige Erkrankung der Hirnrinde. *Allgemeine Zeitschrift fur Psychiatrie und Psyschish-Gerichtlich Medicin* 64, pp. 146-148.

Amer-Ferrer, G. et al. 2005. [Main components of Neuropsychiatric Inventory in Alzheimer's disease. Definition of behavioral syndromes]. *Neurologia* 20(1), pp. 9-16.

American Psychiatric Association 1994. *Diagnostic and statistical manual of mental disorders : DSM-IV*. 4th ed. Washington, D.C.: American Psychiatric Association, pp. xxvii, 886.

Ames, D. et al. 1994. Repetitive and compulsive behavior in frontal lobe degenerations. *J Neuropsychiatry Clin Neurosci* 6(2), pp. 100-113.

Arranz, M. J. et al. 1997. 5-HT2A receptor and bipolar affective disorder: association studies in affected patients. *Neurosci Lett* 224(2), pp. 95-98.

Ashford, J. W. and Mortimer, J. A. 2002. Non-familial Alzheimer's disease is mainly due to genetic factors. *J Alzheimers Dis* 4(3), pp. 169-177.

Assal, F. et al. 2004. Association of the serotonin transporter and receptor gene polymorphisms in neuropsychiatric symptoms in Alzheimer disease. *Arch Neurol* 61(8), pp. 1249-1253.

Avramopoulos, D. et al. 2005. Linkage to chromosome 14q in Alzheimer's disease (AD) patients without psychotic symptoms. *Am J Med Genet B Neuropsychiatr Genet* 132(1), pp. 9-13.

Axelrod, J. and Tomchick, R. 1958. Enzymatic O-methylation of epinephrine and other catechols. *J Biol Chem* 233(3), pp. 702-705.

Bacanu, S. A. et al. 2002. Linkage analysis of Alzheimer disease with psychosis. *Neurology* 59(1), pp. 118-120.

Bacanu, S. A. et al. 2005. Heritability of psychosis in Alzheimer disease. *Am J Geriatr Psychiatry* 13(7), pp. 624-627.

Bachman, D. L. et al. 1992. Prevalence of dementia and probable senile dementia of the Alzheimer type in the Framingham Study. *Neurology* 42(1), pp. 115-119.

Bachman, D. L. et al. 1993. Incidence of dementia and probable Alzheimer's disease in a general population: the Framingham Study. *Neurology* 43(3 Pt 1), pp. 515-519.

Baldwin, C. T. et al. 1995. Linkage of early-onset osteoarthritis and chondrocalcinosis to human chromosome 8q. *Am J Hum Genet* 56(3), pp. 692-697.

Ballard, C. et al. 1995. Associations of psychotic symptoms in dementia sufferers. *Br J Psychiatry* 167(4), pp. 537-540.

Ballard, C. et al. 1997. Apolipoprotein E: non-cognitive symptoms and cognitive decline in late onset Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 63(2), pp. 273-274.

Barak, Y. et al. 2003. Concordance for cognitive impairment: a study of 50 community-dwelling elderly female-female twin pairs. *Compr Psychiatry* 44(2), pp. 117-120.

Bassiony, M. M. and Lyketsos, C. G. 2003. Delusions and hallucinations in Alzheimer's disease: review of the brain decade. *Psychosomatics* 44(5), pp. 388-401.

Baudic, S. et al. 2005. Executive function deficits in early Alzheimer's disease and their relations with episodic memory. *Arch Clin Neuropsychol.*

Becker, J. T. et al. 1994. The natural history of Alzheimer's disease. Description of study cohort and accuracy of diagnosis. *Arch Neurol* 51(6), pp. 585-594.

Becker, K. G. et al. 2004. The genetic association database. *Nat Genet* 36(5), pp. 431-432.

Bellivier, F. et al. 2004. Association between the TPH gene A218C polymorphism and suicidal behavior: a meta-analysis. *Am J Med Genet B Neuropsychiatr Genet* 124(1), pp. 87-91.

Bennett, D. A. et al. 2003. Apolipoprotein E epsilon4 allele, AD pathology, and the clinical expression of Alzheimer's disease. *Neurology* 60(2), pp. 246-252.

Bergem, A. L. et al. 1997. The role of heredity in late-onset Alzheimer disease and vascular dementia. A twin study. *Arch Gen Psychiatry* 54(3), pp. 264-270.

Bertram, L. et al. 2000. Evidence for genetic linkage of Alzheimer's disease to chromosome 10q. *Science* 290(5500), pp. 2302-2303.

Bertram, L. and Tanzi, R. E. 2004. The current status of Alzheimer's disease genetics: what do we tell the patients? *Pharmacol Res* 50(4), pp. 385-396.

Bickeboller, H. et al. 1997. Apolipoprotein E and Alzheimer disease: genotype-specific risks by age and sex. *Am J Hum Genet* 60(2), pp. 439-446.

Blacker, D. et al. 2003. Results of a high-resolution genome screen of 437 Alzheimer's Disease families. *Hum Mol Genet* 12(1), pp. 23-32.

Blessed, G. et al. 1968. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry* 114(512), pp. 797-811.

Bonaiuto, S. et al. 1995. Education and occupation as risk factors for dementia: a population-based case-control study. *Neuroepidemiology* 14(3), pp. 101-109.

Borroni, B. et al. 2004. Catechol-O-methyltransferase gene polymorphism is associated with risk of psychosis in Alzheimer Disease. *Neurosci Lett* 370(2-3), pp. 127-129.

Borson, S. and Raskind, M. A. 1997. Clinical features and pharmacologic treatment of behavioral symptoms of Alzheimer's disease. *Neurology* 48(5 Suppl 6), pp. S17-24.

Bowirrat, A. et al. 2001. Prevalence of Alzheimer's type dementia in an elderly Arab population. *Eur J Neurol* 8(2), pp. 119-123.

Bozeat, S. et al. 2000. Which neuropsychiatric and behavioural features distinguish frontal and temporal variants of frontotemporal dementia from Alzheimer's disease? *J Neurol Neurosurg Psychiatry* 69(2), pp. 178-186.

Bracco, L. et al. 1994. Factors affecting course and survival in Alzheimer's disease. A 9-year longitudinal study. *Arch Neurol* 51(12), pp. 1213-1219.

Brayne, C. et al. 1995. Incidence of clinically diagnosed subtypes of dementia in an elderly population. Cambridge Project for Later Life. *Br J Psychiatry* 167(2), pp. 255-262.

Breitner, J. C. et al. 1995. Alzheimer's disease in the National Academy of Sciences-National Research Council Registry of Aging Twin Veterans. III. Detection of cases, longitudinal results, and observations on twin concordance. *Arch Neurol* 52(8), pp. 763-771.

Brookes, A. J. and Prince, J. A. 2005. Genetic association analysis: lessons from the study of Alzheimers disease. *Mutat Res* 573(1-2), pp. 152-159.

Brookmeyer, R. et al. 1998. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *Am J Public Health* 88(9), pp. 1337-1342.

Bucks, R. S. et al. 1996. Assessment of activities of daily living in dementia: development of the Bristol Activities of Daily Living Scale. *Age Ageing* 25(2), pp. 113-120.

Bullido, M. J. et al. 1998. A polymorphism in the regulatory region of APOE associated with risk for Alzheimer's dementia. *Nat Genet* 18(1), pp. 69-71.

Burns, A. et al. 1990a. Psychiatric phenomena in Alzheimer's disease. I: Disorders of thought content. *Br J Psychiatry* 157, pp. 72-76, 92-74.

Burns, A. et al. 1990b. Psychiatric phenomena in Alzheimer's disease. II: Disorders of perception. *Br J Psychiatry* 157, pp. 76-81, 92-74.

Burns, A. et al. 1990c. Psychiatric phenomena in Alzheimer's disease. III: Disorders of mood. *Br J Psychiatry* 157, pp. 81-86, 92-84.

Burns, A. et al. 1990d. Psychiatric phenomena in Alzheimer's disease. IV: Disorders of behaviour. *Br J Psychiatry* 157, pp. 86-94.

Butt, Z. A. and Strauss, M. E. 2001. Relationship of family and personal history to the occurrence of depression in persons with Alzheimer's disease. *Am J Geriatr Psychiatry* 9(3), pp. 249-254.

Cacabelos, R. et al. 1997. Behavioral changes associated with different apolipoprotein E genotypes in dementia. *Alzheimer Dis Assoc Disord* 11 Suppl 4, pp. S27-34.

Cai, X. D. et al. 1993. Release of excess amyloid beta protein from a mutant amyloid beta protein precursor. *Science* 259(5094), pp. 514-516.

Canadian Study of Health and Aging 1994. The Canadian Study of Health and Aging: risk factors for Alzheimer's disease in Canada. *Neurology* 44(11), pp. 2073-2080.

Cannon, T. D. et al. 1998. The genetic epidemiology of schizophrenia in a Finnish twin cohort. A population-based modeling study. *Arch Gen Psychiatry* 55(1), pp. 67-74.

Cao, Q. et al. 1997. Suggestive evidence for a schizophrenia susceptibility locus on chromosome 6q and a confirmation in an independent series of pedigrees. *Genomics* 43(1), pp. 1-8.

Cardno, A. and McGuffin, P. 2002. Quantitative genetics. In: McGuffin, P. et al. eds. *Psychiatric Genetic and Genomics.* New York: Oxford University Press, pp. 31-54.

Cardno, A. G. et al. 1999. Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. *Arch Gen Psychiatry* 56(2), pp. 162-168.

Carlomagno, S. et al. 2005. Referential communication in Alzheimer's type dementia. *Cortex* 41(4), pp. 520-534.

Chen, C. P. et al. 1996. Presynaptic serotonergic markers in community-acquired cases of Alzheimer's disease: correlations with depression and neuroleptic medication. *J Neurochem* 66(4), pp. 1592-1598.

Chen, S. T. et al. 1998. Executive dysfunction in Alzheimer's disease: association with neuropsychiatric symptoms and functional impairment. *J Neuropsychiatry Clin Neurosci* 10(4), pp. 426-432.

Citron, M. et al. 1992. Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. *Nature* 360(6405), pp. 672-674.

Clarfield, A. M. and Foley, J. M. 1993. The American and Canadian consensus conferences on dementia: is there consensus? *J Am Geriatr Soc* 41(8), pp. 883-886.

Clerget-Darpoux, F. et al. 1986. Effects of misspecifying genetic parameters in lod score analysis. *Biometrics* 42(2), pp. 393-399.

Cobb, J. L. et al. 1995. The effect of education on the incidence of dementia and Alzheimer's disease in the Framingham Study. *Neurology* 45(9), pp. 1707-1712.

Cohen-Mansfield, J. et al. 1998. Delusions and hallucinations in an adult day care population. A longitudinal study. *Am J Geriatr Psychiatry* 6(2), pp. 104-121.

Coleman, P. et al. 2004. A focus on the synapse for neuroprotection in Alzheimer disease and other dementias. *Neurology* 63(7), pp. 1155-1162.

Cordell, H. J. 2001. Sample size requirements to control for stochastic variation in magnitude and location of allele-sharing linkage statistics in affected sibling pairs. *Ann Hum Genet* 65(Pt 5), pp. 491-502.

Corder, E. H. et al. 1994. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet* 7(2), pp. 180-184.

Corder, E. H. et al. 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261(5123), pp. 921-923.

Craddock, N. and Lendon, C. 1998. New susceptibility gene for Alzheimer's disease on chromosome 12? *Lancet* 352(9142), pp. 1720-1721.

Craig, D. et al. 2004a. Allelic variation at the A218C tryptophan hydroxylase polymorphism influences agitation and aggression in Alzheimer's disease. *Neurosci Lett* 363(3), pp. 199-202.

Craig, D. et al. 2004b. Psychotic symptoms in Alzheimer's disease are not influenced by polymorphic variation at the dopamine receptor DRD3 gene. *Neurosci Lett* 368(1), pp. 33-36.

Craig, D. et al. 2004c. Apolipoprotein E e4 allele influences aggressive behaviour in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 75(9), pp. 1327-1330.

Craig, D. et al. 2004d. The interleukin 1beta gene promoter polymorphism (-511) acts as a risk factor for psychosis in Alzheimer's dementia. *Ann Neurol* 56(1), pp. 121-124.

Craig, D. et al. 2005a. A cross-sectional study of neuropsychiatric symptoms in 435 patients with Alzheimer's disease. *Am J Geriatr Psychiatry* 13(6), pp. 460-468.

Craig, D. et al. 2005b. Association analysis of apolipoprotein E genotype and risk of depressive symptoms in Alzheimer's disease. *Dement Geriatr Cogn Disord* 19(2-3), pp. 154-157.

Cummings, J. L. 1997. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. *Neurology* 48(5 Suppl 6), pp. S10-16.

Cummings, J. L. 2000. Cognitive and behavioral heterogeneity in Alzheimer's disease: seeking the neurobiological basis. *Neurobiol Aging* 21(6), pp. 845-861.

Cummings, J. L. and Back, C. 1998. The cholinergic hypothesis of neuropsychiatric symptoms in Alzheimer's disease. *Am J Geriatr Psychiatry* 6(2 Suppl 1), pp. S64-78.

Cummings, J. L. and Cole, G. 2002. Alzheimer disease. *Jama* 287(18), pp. 2335-2338.

Cummings, J. L. and McPherson, S. 2001. Neuropsychiatric assessment of Alzheimer's disease and related dementias. *Aging (Milano)* 13(3), pp. 240-246.

Curtis, D. et al. 2001. A novel method of two-locus linkage analysis applied to a genome scan for late onset Alzheimer's disease. *Ann Hum Genet* 65(Pt 5), pp. 473-481.

Daw, E. W. et al. 1999. Multipoint oligogenic analysis of age-at-onset data with applications to Alzheimer disease pedigrees. *Am J Hum Genet* 64(3), pp. 839-851.

Daw, E. W. et al. 2000. Bias in multipoint linkage analysis arising from map misspecification. *Genet Epidemiol* 19(4), pp. 366-380.

De Ronchi, D. et al. 1998. The effect of education on dementia occurrence in an Italian population with middle to high socioeconomic status. *Neurology* 50(5), pp. 1231-1238.

de Toledo, M. et al. 2004. [Behavioural disorders in Alzheimer's disease. Data from a populational study]. *Rev Neurol* 38(10), pp. 901-905.

Desai, A. K. and Grossberg, G. T. 2005. Diagnosis and treatment of Alzheimer's disease. *Neurology* 64(12 Suppl 3), pp. S34-39.

Deutsch, L. H. et al. 1991. Psychosis and physical aggression in probable Alzheimer's disease. *Am J Psychiatry* 148(9), pp. 1159-1163.

Devanand, D. P. et al. 1992. Behavioral syndromes in Alzheimer's disease. *Int Psychogeriatr* 4 Suppl 2, pp. 161-184.

Devanand, D. P. et al. 1996. Depressed mood and the incidence of Alzheimer's disease in the elderly living in the community. *Arch Gen Psychiatry* 53(2), pp. 175-182.

Devlin, B. et al. 2002. Mixture models for linkage analysis of affected sibling pairs and covariates. *Genet Epidemiol* 22(1), pp. 52-65.

Di Carlo, A. et al. 2002. Incidence of dementia, Alzheimer's disease, and vascular dementia in Italy. The ILSA Study. *J Am Geriatr Soc* 50(1), pp. 41-48.

Dicks, J. and Savva, G. 2003. Comparative Genomics. In: Balding, D.J. et al. eds. *Handbook of Statistical Genetics. 2nd Ed.* Chichester: Wiley, pp. 139-161.

Donaldson, C. et al. 1998. Determinants of carer stress in Alzheimer's disease. *Int J Geriatr Psychiatry* 13(4), pp. 248-256.

Donner, A. 1986. A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model. *International Statistical Review* 54, pp. 67-82.

Dorr, D. A. et al. 1997. A meta-analysis of chromosome 18 linkage data for bipolar illness. *Genet Epidemiol* 14(6), pp. 617-622.

Douglas, J. A. et al. 2000. A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet* 66(4), pp. 1287-1297.

Drachman, D. A. 1994. If we live long enough, will we all be demented? *Neurology* 44(9), pp. 1563-1565.

Drevets, W. C. and Rubin, E. H. 1989. Psychotic symptoms and the longitudinal course of senile dementia of the Alzheimer type. *Biol Psychiatry* 25(1), pp. 39-48.

Duara, R. et al. 1996. Alzheimer's disease: interaction of apolipoprotein E genotype, family history of dementia, gender, education, ethnicity, and age of onset. *Neurology* 46(6), pp. 1575-1579.

Duff, K. et al. 1996. Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin 1. *Nature* 383(6602), pp. 710-713.

Ehl, C. et al. 2003. Association of an interleukin-1beta gene polymorphism at position -511 with Alzheimer's disease. *Int J Mol Med* 11(2), pp. 235-238.

Ertekin-Taner, N. et al. 2001. Heritability of plasma amyloid beta in typical late-onset Alzheimer's disease pedigrees. *Genet Epidemiol* 21(1), pp. 19-30.

Ertekin-Taner, N. et al. 2000. Linkage of plasma Abeta42 to a quantitative locus on chromosome 10 in late-onset Alzheimer's disease pedigrees. *Science* 290(5500), pp. 2303-2304.

Esiri, M. M. 1996. The basis for behavioural disturbances in dementia. *J Neurol Neurosurg Psychiatry* 61(2), pp. 127-130.

Evans, D. A. et al. 1997a. Apolipoprotein E epsilon4 and incidence of Alzheimer disease in a community population of older persons. *Jama* 277(10), pp. 822-824.

Evans, D. A. et al. 1997b. Education and other measures of socioeconomic status and risk of incident Alzheimer disease in a defined population of older persons. *Arch Neurol* 54(11), pp. 1399-1405.

Fairburn, C. G. and Hope, R. A. 1988. Changes in behaviour in dementia: a neglected research area. *Br J Psychiatry* 152, pp. 406-407.

Farber, N. B. et al. 2000. Increased neocortical neurofibrillary tangle density in subjects with Alzheimer disease and psychosis. *Arch Gen Psychiatry* 57(12), pp. 1165-1173.

Farrer, L. A. et al. 2003. Identification of multiple loci for Alzheimer disease in a consanguineous Israeli-Arab community. *Hum Mol Genet* 12(4), pp. 415-422.

Farrer, L. A. et al. 1997. 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* 278(16), pp. 1349-1356.

Ferretti, L. et al. 2001. Anxiety and Alzheimer's disease. *J Geriatr Psychiatry Neurol* 14(1), pp. 52-58.

Fillenbaum, G. G. et al. 1998. The prevalence and 3-year incidence of dementia in older Black and White community residents. *J Clin Epidemiol* 51(7), pp. 587-595.

Finkel, S. I. 2001. Behavioral and psychological symptoms of dementia: a current focus for clinicians, researchers, and caregivers. *J Clin Psychiatry* 62 Suppl 21, pp. 3-6.

Fisfalen, M. E. et al. 2005. Familial variation in episode frequency in bipolar affective disorder. *Am J Psychiatry* 162(7), pp. 1266-1272.

Flynn, F. G. et al. 1991. Delusions in dementia syndromes: investigation of behavioral and neuropsychological correlates. *J Neuropsychiatry Clin Neurosci* 3(4), pp. 364-370.

Folstein, M. F. et al. 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12(3), pp. 189-198.

Forsell, Y. et al. 1998. Psychotic symptoms and apolipoprotein E genotypes in an elderly population. *Biol Psychiatry* 44(2), pp. 139-140.

Fratiglioni, L. et al. 1993. Risk factors for late-onset Alzheimer's disease: a population-based, case-control study. *Ann Neurol* 33(3), pp. 258-266.

Fratiglioni, L. et al. 1999. Worldwide prevalence and incidence of dementia. *Drugs Aging* 15(5), pp. 365-375.

Fratiglioni, L. et al. 1991. Prevalence of Alzheimer's disease and other dementias in an elderly urban population: relationship with age, sex, and education. *Neurology* 41(12), pp. 1886-1892.

Fratiglioni, L. et al. 1997. Very old women at highest risk of dementia and Alzheimer's disease: incidence data from the Kungsholmen Project, Stockholm. *Neurology* 48(1), pp. 132-138.

Freimer, N. and Sabatti, C. 2003. The human phenome project. *Nat Genet* 34(1), pp. 15-21.

Frisoni, G. B. et al. 1999. Behavioral syndromes in Alzheimer's disease: description and correlates. *Dement Geriatr Cogn Disord* 10(2), pp. 130-138.

Frydman, M. et al. 1985. Assignment of the gene for Wilson disease to chromosome 13: linkage to the esterase D locus. *Proc Natl Acad Sci U S A* 82(6), pp. 1819-1821.

Fuh, J. L. et al. 2001. Behavioral disorders and caregivers' reaction in Taiwanese patients with Alzheimer's disease. *Int Psychogeriatr* 13(1), pp. 121-128.

Gabryelewicz, T. et al. 2002. Behavioural pathology in Alzheimer's disease with special reference to apolipoprotein E genotype. *Dement Geriatr Cogn Disord* 14(4), pp. 208-212.

Ganguli, M. et al. 2000. Ten-year incidence of dementia in a rural elderly US community population: the MoVIES Project. *Neurology* 54(5), pp. 1109-1116.

Garcia-Alloza, M. et al. 2005. Cholinergic-serotonergic imbalance contributes to cognitive and behavioral symptoms in Alzheimer's disease. *Neuropsychologia* 43(3), pp. 442-449.

Gatz, J. L. et al. 2005a. Do depressive symptoms predict Alzheimer's disease and dementia? *J Gerontol A Biol Sci Med Sci* 60(6), pp. 744-747.

Gatz, M. et al. 2005b. Complete ascertainment of dementia in the Swedish Twin Registry: the HARMONY study. *Neurobiol Aging* 26(4), pp. 439-447.

Gatz, M. et al. 1997. Heritability for Alzheimer's disease: the study of dementia in Swedish twins. *J Gerontol A Biol Sci Med Sci* 52(2), pp. M117-125.

Gaugler, J. E. et al. 2000. Predictors of institutionalization of cognitively impaired elders: family help and the timing of placement. *J Gerontol B Psychol Sci Soc Sci* 55(4), pp. P247-255.

Gaugler, J. E. et al. 1999. Caregiving and institutionalization: perceptions of family conflict and socioemotional support. *Int J Aging Hum Dev* 49(1), pp. 1-25.

Gauthier, S. et al. 2005. Effects of memantine on behavioural symptoms in Alzheimer's disease patients: an analysis of the Neuropsychiatric Inventory (NPI) data of two randomised, controlled studies. *Int J Geriatr Psychiatry* 20(5), pp. 459-464.

Gautrin, D. et al. 1990. Canadian projections of cases suffering from Alzheimer's disease and senile dementia of Alzheimer type over the period 1986-2031. *Can J Psychiatry* 35(2), pp. 162-165.

Gearing, M. et al. 1995. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part X. Neuropathology confirmation of the clinical diagnosis of Alzheimer's disease. *Neurology* 45(3 Pt 1), pp. 461-466.

Geser, F. et al. 2005. How to diagnose dementia with Lewy bodies: state of the art. *Mov Disord* 20 Suppl 12, pp. S11-20.

Ghebremedhin, E. et al. 2001. Gender and age modify the association between APOE and AD-related neuropathology. *Neurology* 56(12), pp. 1696-1701.

Ghosh, S. et al. 2000. The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. *Am J Hum Genet* 67(5), pp. 1174-1185.

Gillanders, E. M. et al. 2004. Combined genome-wide scan for prostate cancer susceptibility genes. *J Natl Cancer Inst* 96(16), pp. 1240-1247.

Glatt, S. J. et al. 2003. Association between a functional catechol O-methyltransferase gene polymorphism and schizophrenia: meta-analysis of case-control and family-based studies. *Am J Psychiatry* 160(3), pp. 469-476.

Goate, A. et al. 1991. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349(6311), pp. 704-706.

Goddard, K. A. et al. 2001. Model-free linkage analysis with covariates confirms linkage of prostate cancer to chromosomes 1 and 4. *Am J Hum Genet* 68(5), pp. 1197-1206.

Gomez-Isla, T. et al. 1996. Clinical and pathological correlates of apolipoprotein E epsilon 4 in Alzheimer's disease. *Ann Neurol* 39(1), pp. 62-70.

Gorwood, P. 2004. Generalized anxiety disorder and major depressive disorder comorbidity: an example of genetic pleiotropy? *Eur Psychiatry* 19(1), pp. 27-33.

Gottesman, II and Gould, T. D. 2003. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 160(4), pp. 636-645.

Greenwood, C. M. and Bull, S. B. 1997. Incorporation of covariates into genome scanning using sib-pair analysis in bipolar affective disorder. *Genet Epidemiol* 14(6), pp. 635-640.

Greenwood, C. M. and Bull, S. B. 1999. Analysis of affected sib pairs, with covariates--with and without constraints. *Am J Hum Genet* 64(3), pp. 871-885.

Grossman, M. et al. 2004. What's in a name: voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. *Brain* 127(Pt 3), pp. 628-649.

Guo, Z. et al. 1999. Occurrence and progression of dementia in a community population aged 75 years and older: relationship of antihypertensive medication use. *Arch Neurol* 56(8), pp. 991-996.

Gussekloo, J. et al. 1995. A community based study of the incidence of dementia in subjects aged 85 years and over. *J Neurol Neurosurg Psychiatry* 59(5), pp. 507-510.

Gutierrez, B. et al. 1996. Allelic association analysis of the 5-HT2C receptor gene in bipolar affective disorder. *Neurosci Lett* 212(1), pp. 65-67.

Hachinski, V. C. et al. 1975. Cerebral blood flow in dementia. *Arch Neurol* 32(9), pp. 632-637.

Hagnell, O. et al. 1992. Incidence of dementia in the Lundby Study. *Neuroepidemiology* 11 Suppl 1, pp. 61-66.

Hallmayer, J. F. et al. 2005. Genetic evidence for a distinct subtype of schizophrenia characterized by pervasive cognitive deficit. *Am J Hum Genet* 77(3), pp. 468-476.

Hamshere, M. L. et al. 2005. Genome-wide significant linkage in schizophrenia conditioning on occurrence of depressive episodes. *J Med Genet.*

Harper, D. G. et al. 2005. Disturbance of endogenous circadian rhythm in aging and Alzheimer disease. *Am J Geriatr Psychiatry* 13(5), pp. 359-368.

Harwood, D. G. et al. 2000. Relationship of behavioral and psychological symptoms to cognitive impairment and functional status in Alzheimer's disease. *Int J Geriatr Psychiatry* 15(5), pp. 393-400.

Harwood, D. G. et al. 1999. Apolipoprotein-E (APO-E) genotype and symptoms of psychosis in Alzheimer's disease. *Am J Geriatr Psychiatry* 7(2), pp. 119-123.

Harwood, D. G. et al. 1998. The behavioral pathology in Alzheimer's Disease Scale (BEHAVE-AD): factor structure among community-dwelling Alzheimer's disease patients. *Int J Geriatr Psychiatry* 13(11), pp. 793-800.

Haupt, M. et al. 1998. Presentation and stability of noncognitive symptom patterns in patients with Alzheimer disease. *Alzheimer Dis Assoc Disord* 12(4), pp. 323-329.

Haupt, M. et al. 1992. Alzheimer's disease: identical phenotype of familial and non-familial cases. *J Neurol* 239(5), pp. 248-250.

Hauser, E. R. et al. 2004. Ordered subset analysis in genetic linkage mapping of complex traits. *Genet Epidemiol* 27(1), pp. 53-63.

Havlik, R. J. et al. 2000. APOE-epsilon4 predicts incident AD in Japanese-American men: the honolulu-asia aging study. *Neurology* 54(7), pp. 1526-1529.

Hebert, L. E. et al. 1995. Age-specific incidence of Alzheimer's disease in a community population. *Jama* 273(17), pp. 1354-1359.

Hebert, L. E. et al. 2003. Alzheimer disease in the US population: prevalence estimates using the 2000 census. *Arch Neurol* 60(8), pp. 1119-1122.

Henderson, V. W. et al. 1989. Spatial disorientation in Alzheimer's disease. *Arch Neurol* 46(4), pp. 391-394.

Henderson, V. W. et al. 1994. Estrogen replacement therapy in older women. Comparisons between Alzheimer's disease cases and nondemented control subjects. *Arch Neurol* 51(9), pp. 896-900.

Hendrie, H. C. 1998. Epidemiology of dementia and Alzheimer's disease. *Am J Geriatr Psychiatry* 6(2 Suppl 1), pp. S3-18.

Hendrie, H. C. et al. 2001. Incidence of dementia and Alzheimer disease in 2 communities: Yoruba residing in Ibadan, Nigeria, and African Americans residing in Indianapolis, Indiana. *Jama* 285(6), pp. 739-747.

Hill, L. R. et al. 1993. Functional status, education, and the diagnosis of dementia in the Shanghai survey. *Neurology* 43(1), pp. 138-145.

Hiltunen, M. et al. 2001. Genome-wide linkage disequilibrium mapping of late-onset Alzheimer's disease in Finland. *Neurology* 57(9), pp. 1663-1668.

Hirono, N. et al. 1999. Lack of effect of apolipoprotein E E4 allele on neuropsychiatric manifestations in Alzheimer's disease. *J Neuropsychiatry Clin Neurosci* 11(1), pp. 66-70.

Holmans, P. 2003. Nonparametric Linkage. In: Balding, D.J. et al. eds. *Handbook of Statistical Genetics. 2nd Ed.* Chichester: Wiley, pp. 919-938.

Holmans, P. and Clayton, D. 1995. Efficiency of typing unaffected relatives in an affected-sib-pair linkage study with single-locus and multiple tightly linked markers. *Am J Hum Genet* 57(5), pp. 1221-1232.

Holmans, P. et al. 2005. Genome screen for loci influencing age at onset and rate of decline in late onset Alzheimer's disease. *Am J Med Genet B Neuropsychiatr Genet* 135(1), pp. 24-32.

Holmes, C. et al. 1998a. 5-HT2A and 5-HT2C receptor polymorphisms and psychopathology in late onset Alzheimer's disease. *Hum Mol Genet* 7(9), pp. 1507-1509.

Holmes, C. et al. 1999. Validity of current clinical criteria for Alzheimer's disease, vascular dementia and dementia with Lewy bodies. *Br J Psychiatry* 174, pp. 45-50.

Holmes, C. et al. 1996. Apolipoprotein E: non-cognitive symptoms and cognitive decline in late onset Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 61(6), pp. 580-583.

Holmes, C. and Lovestone, S. 2002. The clinical phenotype of familial and sporadic late onset Alzheimer's disease. *Int J Geriatr Psychiatry* 17(2), pp. 146-149.

Holmes, C. et al. 1998b. Apolipoprotein E: depressive illness, depressive symptoms, and Alzheimer's disease. *Biol Psychiatry* 43(3), pp. 159-164.

Holmes, C. et al. 2001. Psychosis and aggression in Alzheimer's disease: the effect of dopamine receptor gene variation. *J Neurol Neurosurg Psychiatry* 71(6), pp. 777-779.

Hope, T. et al. 1997. Behaviour changes in dementia. 2: Are there behavioural syndromes? *Int J Geriatr Psychiatry* 12(11), pp. 1074-1078.

Hope, T. et al. 1998. Predictors of institutionalization for people with dementia living at home with a carer. *Int J Geriatr Psychiatry* 13(10), pp. 682-690.

Hughes, C. P. et al. 1982. A new clinical scale for the staging of dementia. *Br J Psychiatry* 140, pp. 566-572.

Humpel, C. and Marksteiner, J. 2005. Cerebrovascular damage as a cause for Alzheimer's disease. *Curr Neurovasc Res* 2(4), pp. 341-347.

Ikeda, M. et al. 2002. Changes in appetite, food preference, and eating habits in frontotemporal dementia and Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 73(4), pp. 371-376.

International Human Genome Sequencing Consortium 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431(7011), pp. 931-945.

Jarvik, G. et al. 1996. Influence of apolipoprotein E genotype on the transmission of Alzheimer disease in a community-based sample. *Am J Hum Genet* 58(1), pp. 191-200.

Jeste, D. V. and Finkel, S. I. 2000. Psychosis of Alzheimer's disease and related dementias. Diagnostic criteria for a distinct syndrome. *Am J Geriatr Psychiatry* 8(1), pp. 29-34.

Jeste, D. V. et al. 1992. Cognitive deficits of patients with Alzheimer's disease with and without delusions. *Am J Psychiatry* 149(2), pp. 184-189.

Jorm, A. F. and Jolley, D. 1998. The incidence of dementia: a meta-analysis. *Neurology* 51(3), pp. 728-733.

Jost, B. C. and Grossberg, G. T. 1995. The natural history of Alzheimer's disease: a brain bank study. *J Am Geriatr Soc* 43(11), pp. 1248-1255.

Kaiser, H. F. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20, pp. 141-151.

Kamboh, M. I. 2004. Molecular genetics of late-onset Alzheimer's disease. *Ann Hum Genet* 68(Pt 4), pp. 381-404.

Kanai, M. et al. 1999. Apolipoprotein E4 accelerates dementia and increases cerebrospinal fluid tau levels in Alzheimer's disease. *Neurosci Lett* 267(1), pp. 65-68.

Kantarci, K. and Jack, C. R., Jr. 2003. Neuroimaging in Alzheimer disease: an evidence-based review. *Neuroimaging Clin N Am* 13(2), pp. 197-209.

Karp, A. et al. 2004. Relation of education and occupation-based socioeconomic status to incident Alzheimer's disease. *Am J Epidemiol* 159(2), pp. 175-183.

Katila, H. et al. 1999. Polymorphisms of the interleukin-1 gene complex in schizophrenia. *Mol Psychiatry* 4(2), pp. 179-181.

Katzman, R. 1993. Education and the prevalence of dementia and Alzheimer's disease. *Neurology* 43(1), pp. 13-20.

Kaufer, D. et al. 1998. Differential neuropsychiatric symptom responses to tacrine in Alzheimer's disease: relationship to dementia severity. *J Neuropsychiatry Clin Neurosci* 10(1), pp. 55-63.

Kawas, C. et al. 2000. Age-specific incidence rates of Alzheimer's disease: the Baltimore Longitudinal Study of Aging. *Neurology* 54(11), pp. 2072-2077.

Kehoe, P. et al. 1999. A full genome scan for late onset Alzheimer's disease. *Hum Mol Genet* 8(2), pp. 237-245.

Kendler, K. S. 1990. Toward a scientific psychiatric nosology. Strengths and limitations. *Arch Gen Psychiatry* 47(10), pp. 969-973.

Kendler, K. S. and Gardner, C. O., Jr. 1998. Twin studies of adult psychiatric and substance dependence disorders: are they biased by differences in the environmental experiences of monozygotic and dizygotic twins in childhood and adolescence? *Psychol Med* 28(3), pp. 625-633.

Kline, P. 1994. *An Easy Guide to Factor Analysis*. London: Routledge.

Kloszewska, I. 1998. Incidence and relationship between behavioural and psychological symptoms in Alzheimer's disease. *Int J Geriatr Psychiatry* 13(11), pp. 785-792.

Knapp, M. et al. 1993. The effect of misspecifying allele frequencies in incompletely typed families. *Genet Epidemiol* 10(6), pp. 413-418.

Knopman, D. S. et al. 2003. Essentials of the proper diagnoses of mild cognitive impairment, dementia, and major subtypes of dementia. *Mayo Clin Proc* 78(10), pp. 1290-1308.

Knopman, D. S. et al. 1988. Longitudinal study of death and institutionalization in patients with primary degenerative dementia. *J Am Geriatr Soc* 36(2), pp. 108-112.

Kokmen, E. et al. 1989. Prevalence of medically diagnosed dementia in a defined United States population: Rochester, Minnesota, January 1, 1975. *Neurology* 39(6), pp. 773-776.

Kong, A. et al. 2002. A high-resolution recombination map of the human genome. *Nat Genet* 31(3), pp. 241-247.

Kordower, J. H. et al. 2001. Loss and atrophy of layer II entorhinal cortex neurons in elderly people with mild cognitive impairment. *Ann Neurol* 49(2), pp. 202-213.

Korszun, A. et al. 2004. Familiality of symptom dimensions in depression. *Arch Gen Psychiatry* 61(5), pp. 468-474.

Koss, E. et al. 1996. Clinical and neuropsychological differences between patients with earlier and later onset of Alzheimer's disease: A CERAD analysis, Part XII. *Neurology* 46(1), pp. 136-141.

Krishnan, K. R. et al. 1996. Apolipoprotein E-epsilon 4 frequency in geriatric depression. *Biol Psychiatry* 40(1), pp. 69-71.

Kukull, W. A. et al. 2002. Dementia and Alzheimer disease incidence: a prospective cohort study. *Arch Neurol* 59(11), pp. 1737-1746.

Lam, L. C. et al. 2004. 5-HT2A T102C receptor polymorphism and neuropsychiatric symptoms in Alzheimer's disease. *Int J Geriatr Psychiatry* 19(6), pp. 523-526.

Lambert, D. et al. 2005. Stage 2 of the Wellcome Trust UK-Irish bipolar affective disorder sibling-pair genome screen: evidence for linkage on chromosomes 6q16-q21, 4q12-q21, 9p21, 10p14-p12 and 18q22. *Mol Psychiatry* 10(9), pp. 831-841.

Lambert, J. C. et al. 2002. Contribution of APOE promoter polymorphisms to Alzheimer's disease risk. *Neurology* 59(1), pp. 59-66.

Lambert, J. C. et al. 1998. Pronounced impact of Th1/E47cs mutation compared with -491 AT mutation on neural APOE gene expression and risk of developing Alzheimer's disease. *Hum Mol Genet* 7(9), pp. 1511-1516.

Lanctot, K. L. et al. 2001. Role of serotonin in the behavioral and psychological symptoms of dementia. *J Neuropsychiatry Clin Neurosci* 13(1), pp. 5-21.

Lander, E. and Kruglyak, L. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11(3), pp. 241-247.

Lander, E. S. and Schork, N. J. 1994. Genetic dissection of complex traits. *Science* 265(5181), pp. 2037-2048.

Landes, A. M. et al. 2001. Apathy in Alzheimer's disease. *J Am Geriatr Soc* 49(12), pp. 1700-1707.

Lange, R. T. et al. 2004. Psychometric properties and factor structure of the Neuropsychiatric Inventory Nursing Home version in an elderly neuropsychiatric population. *Int J Geriatr Psychiatry* 19(5), pp. 440-448.

Larner, A. J. 2005. "Frontal variant Alzheimer's disease": A reappraisal. *Clin Neurol Neurosurg.*

Larson, E. B. et al. 2004. Survival after initial diagnosis of Alzheimer disease. *Ann Intern Med* 140(7), pp. 501-509.

Launer, L. J. 1992. Overview of incidence studies of dementia conducted in Europe. *Neuroepidemiology* 11 Suppl 1, pp. 2-13.

Launer, L. J. et al. 1999. Rates and risk factors for dementia and Alzheimer's disease: results from EURODEM pooled analyses. EURODEM Incidence Research Group and Work Groups. European Studies of Dementia. *Neurology* 52(1), pp. 78-84.

Lawlor, B. and Bhriain, S. N. 2001. Psychosis and behavioural symptoms of dementia: defining the role of neuroleptic interventions. *Int J Geriatr Psychiatry* 16 Suppl 1, pp. S2-6.

Lawlor, B. A. et al. 1994. Clinical symptoms associated with age at onset in Alzheimer's disease. *Am J Psychiatry* 151(11), pp. 1646-1649.

Leboyer, M. et al. 1998. Psychiatric genetics: search for phenotypes. *Trends Neurosci* 21(3), pp. 102-105.

Leckman, J. F. et al. 1984. Subtypes of depression. Family study perspective. *Arch Gen Psychiatry* 41(9), pp. 833-838.

Lee, J. H. et al. 2004. Fine mapping of 10q and 18q for familial Alzheimer's disease in Caribbean Hispanics. *Mol Psychiatry* 9(11), pp. 1042-1051.

Lee, P. N. 1994. Smoking and Alzheimer's disease: a review of the epidemiological evidence. *Neuroepidemiology* 13(4), pp. 131-144.

Lehtovirta, M. et al. 1996a. Clinical and neuropsychological characteristics in familial and sporadic Alzheimer's disease: relation to apolipoprotein E polymorphism. *Neurology* 46(2), pp. 413-419.

Lehtovirta, M. et al. 1996b. SPECT and MRI analysis in Alzheimer's disease: relation to apolipoprotein E epsilon 4 allele. *J Neurol Neurosurg Psychiatry* 60(6), pp. 644-649.

Lendon, C. L. et al. 1997. Exploring the etiology of Alzheimer disease using molecular genetics. *Jama* 277(10), pp. 825-831.

Letenneur, L. et al. 1994a. Incidence of dementia and Alzheimer's disease in elderly community residents of south-western France. *Int J Epidemiol* 23(6), pp. 1256-1261.

Letenneur, L. et al. 1994b. Tobacco consumption and cognitive impairment in elderly people. A population-based study. *Ann Epidemiol* 4(6), pp. 449-454.

Leung, V. P. et al. 2001. Validation study of the Chinese version of the neuropsychiatric inventory (CNPI). *Int J Geriatr Psychiatry* 16(8), pp. 789-793.

Levinson, D. F. et al. 2000. Multicenter linkage study of schizophrenia candidate regions on chromosomes 5q, 6q, 10p, and 13q: schizophrenia linkage collaborative group III. *Am J Hum Genet* 67(3), pp. 652-663.

Levy, M. L. et al. 1996a. Longitudinal assessment of symptoms of depression, agitation, and psychosis in 181 patients with Alzheimer's disease. *Am J Psychiatry* 153(11), pp. 1438-1443.

Levy, M. L. et al. 1999. Apolipoprotein E genotype and noncognitive symptoms in Alzheimer's disease. *Biol Psychiatry* 45(4), pp. 422-425.

Levy, M. L. et al. 1996b. Alzheimer disease and frontotemporal dementias. Behavioral distinctions. *Arch Neurol* 53(7), pp. 687-690.

Levy-Lahad, E. et al. 1995a. Apolipoprotein E genotypes and age of onset in early-onset familial Alzheimer's disease. *Ann Neurol* 38(4), pp. 678-680.

Levy-Lahad, E. et al. 1995b. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 269(5226), pp. 973-977.

Li, Y. J. et al. 2002. Age at onset in two common neurodegenerative diseases is genetically controlled. *Am J Hum Genet* 70(4), pp. 985-993.

Lineweaver, T. T. et al. 2005. Differential effects of Alzheimer's disease and Huntington's disease on the performance of mental rotation. *J Int Neuropsychol Soc* 11(1), pp. 30-39.

Lippa, C. F. 1999. Familial Alzheimer's disease: genetic influences on the disease process (Review). *Int J Mol Med* 4(5), pp. 529-536.

Liu, L. et al. 1996. Allelic association but only weak evidence for linkage to the apolipoprotein E locus in late-onset Swedish Alzheimer families. *Am J Med Genet* 67(3), pp. 306-311.

Lopez, O. L. et al. 1991. Alzheimer's disease with delusions and hallucinations: neuropsychological and electroencephalographic correlates. *Neurology* 41(6), pp. 906-912.

Lopez, O. L. et al. 2003. Psychiatric symptoms vary with the severity of dementia in probable Alzheimer's disease. *J Neuropsychiatry Clin Neurosci* 15(3), pp. 346-353.

Lopez, O. L. et al. 1997. The apolipoprotein E epsilon 4 allele is not associated with psychiatric symptoms or extrapyramidal signs in probable Alzheimer's disease. *Neurology* 49(3), pp. 794-797.

Lopez, O. L. et al. 1999. Psychiatric medication and abnormal behavior as predictors of progression in probable Alzheimer disease. *Arch Neurol* 56(10), pp. 1266-1272.

Lovestone, S. and Hardy, J. 2002. Psychotic genes or forgetful ones? *Neurology* 59(1), pp. 11-12.

Lowin, A. et al. 2001. Alzheimer's disease in the UK: comparative evidence on cost of illness and volume of health services research funding. *Int J Geriatr Psychiatry* 16(12), pp. 1143-1148.

Luchsinger, J. A. and Mayeux, R. 2004. Cardiovascular risk factors and Alzheimer's disease. *Curr Atheroscler Rep* 6(4), pp. 261-266.

Lucotte, G. et al. 1994. Apolipoprotein E-epsilon 4 allele doses in late-onset Alzheimer's disease. *Ann Neurol* 36(4), pp. 681-682.

Lyketsos, C. G. et al. 1997. Depression, delusions, and hallucinations in Alzheimer's disease: no relationship to apolipoprotein E genotype. *J Neuropsychiatry Clin Neurosci* 9(1), pp. 64-67.

Lyketsos, C. G. et al. 1999. Physical aggression in dementia patients and its relationship to depression. *Am J Psychiatry* 156(1), pp. 66-71.

Lyketsos, C. G. et al. 2000. Mental and behavioral disturbances in dementia: findings from the Cache County Study on Memory in Aging. *Am J Psychiatry* 157(5), pp. 708-714.

Lyketsos, C. G. et al. 1996. Major depression in Alzheimer's disease. An interaction between gender and family history. *Psychosomatics* 37(4), pp. 380-384.

Mack, J. L. et al. 1999. Behavior Rating Scale for Dementia: development of test scales and presentation of data for 555 individuals with Alzheimer's disease. *J Geriatr Psychiatry Neurol* 12(4), pp. 211-223.

Magni, E. et al. 1996. Risk of mortality and institutionalization in demented patients with delusions. *J Geriatr Psychiatry Neurol* 9(3), pp. 123-126.

Manuck, S. B. et al. 1999. Aggression and anger-related traits associated with a polymorphism of the tryptophan hydroxylase gene. *Biol Psychiatry* 45(5), pp. 603-614.

Marin, D. B. et al. 1997. Noncognitive disturbances in Alzheimer's disease: frequency, longitudinal course, and relationship to cognitive symptoms. *J Am Geriatr Soc* 45(11), pp. 1331-1338.

Marin, R. S. et al. 1994. Group differences in the relationship between apathy and depression. *J Nerv Ment Dis* 182(4), pp. 235-239.

Martinez, M. et al. 1998. Apolipoprotein E epsilon4 allele and familial aggregation of Alzheimer disease. *Arch Neurol* 55(6), pp. 810-816.

Martinez, M. et al. 1999. Follow-up study on a susceptibility locus for schizophrenia on chromosome 6q. *Am J Med Genet* 88(4), pp. 337-343.

Matsuda, O. and Saito, M. 2005. Functional competency and cognitive ability in mild Alzheimer's disease: relationship between ADL assessed by a relative/ carer-rated scale and neuropsychological performance. *Int Psychogeriatr* 17(2), pp. 275-288.

Matsuoka, K. et al. 2003. Relationship between behavioral disturbances and characteristics of patients in special units for dementia. *Psychiatry Clin Neurosci* 57(6), pp. 569-574.

Matthews, F. and Brayne, C. 2005. The incidence of dementia in England and Wales: findings from the five identical sites of the MRC CFA Study. *PLoS Med* 2(8), p. e193.

McCurry, S. M. et al. 2004. Anxiety and nighttime behavioral disturbances. Awakenings in patients with Alzheimer's disease. *J Gerontol Nurs* 30(1), pp. 12-20.

McIlroy, S. and Craig, D. 2004. Neurobiology and genetics of behavioural syndromes of Alzheimer's disease. *Curr Alzheimer Res* 1(2), pp. 135-142.

McKhann, G. et al. 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 34(7), pp. 939-944.

McQueen, M. B. et al. 2005. Combined analysis from eleven linkage studies of bipolar disorder provides strong evidence of susceptibility loci on chromosomes 6q and 8q. *Am J Hum Genet* 77(4), pp. 582-595.

McShane, R. 2000. What are the behavioral and psychological symptoms in dementia? *Int Psychogeriatr* 12(Suppl 1), pp. 147-153.

McShane, R. et al. 1997. Do neuroleptic drugs hasten cognitive decline in dementia? Prospective study with necropsy follow up. *Bmj* 314(7076), pp. 266-270.

Mega, M. S. et al. 1996. The spectrum of behavioral changes in Alzheimer's disease. *Neurology* 46(1), pp. 130-135.

Merikangas, K. R. et al. 1989. Linkage studies of bipolar disorder: methodologic and analytic issues. Report of MacArthur Foundation Workshop on Linkage and Clinical Features in Affective Disorders. *Arch Gen Psychiatry* 46(12), pp. 1137-1141.

Meyer, M. R. et al. 1998. APOE genotype predicts when--not whether--one is predisposed to develop Alzheimer disease. *Nat Genet* 19(4), pp. 321-322.

Middeldorp, C. M. et al. 2005. The co-morbidity of anxiety and depression in the perspective of genetic epidemiology. A review of twin and family studies. *Psychol Med* 35(5), pp. 611-624.

Miech, R. A. et al. 2002. Incidence of AD may decline in the early 90s for men, later for women: The Cache County study. *Neurology* 58(2), pp. 209-218.

Migliorelli, R. et al. 1995a. Neuropsychiatric and neuropsychological correlates of delusions in Alzheimer's disease. *Psychol Med* 25(3), pp. 505-513.

Migliorelli, R. et al. 1995b. Prevalence and correlates of dysthymia and major depression among patients with Alzheimer's disease. *Am J Psychiatry* 152(1), pp. 37-44.

Mintzer, J. et al. 2001. Olanzapine in the treatment of anxiety symptoms due to Alzheimer's disease: a post hoc analysis. *Int J Geriatr Psychiatry* 16 Suppl 1, pp. S71-77.

Mirakhur, A. et al. 2004. Behavioural and psychological syndromes in Alzheimer's disease. *Int J Geriatr Psychiatry* 19(11), pp. 1035-1039.

Mohs, R. C. 2005. The clinical syndrome of Alzheimer's disease: aspects particularly relevant to clinical trials. *Genes Brain Behav* 4(3), pp. 129-133.

Moran, M. et al. 2004. Syndromes of behavioural and psychological symptoms in mild Alzheimer's disease. *Int J Geriatr Psychiatry* 19(4), pp. 359-364.

Mori, E. et al. 2002. Accelerated hippocampal atrophy in Alzheimer's disease with apolipoprotein E epsilon4 allele. *Ann Neurol* 51(2), pp. 209-214.

Morris, J. C. et al. 1989. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 39(9), pp. 1159-1165.

Mortel, K. F. et al. 1995. Education and occupation as risk factors for dementias of the Alzheimer and ischemic vascular types. *Dementia* 6(1), pp. 55-62.

Mortimer, J. A. et al. 2003. Head circumference, education and risk of dementia: findings from the Nun Study. *J Clin Exp Neuropsychol* 25(5), pp. 671-679.

Mourik, J. C. et al. 2004. Frontotemporal dementia: behavioral symptoms and caregiver distress. *Dement Geriatr Cogn Disord* 18(3-4), pp. 299-306.

Mullan, M. et al. 1993. Age of onset in familial early onset Alzheimer's disease correlates with genetic aetiology. *Am J Med Genet* 48(3), pp. 129-130.

Muller-Thomsen, T. et al. 2002. Depression in Alzheimer's disease might be associated with apolipoprotein E epsilon 4 allele frequency in women but not in men. *Dement Geriatr Cogn Disord* 14(2), pp. 59-63.

Myers, A. et al. 2000. Susceptibility locus for Alzheimer's disease on chromosome 10. *Science* 290(5500), pp. 2304-2305.

Myers, A. et al. 2002. Full genome screen for Alzheimer disease: stage II analysis. *Am J Med Genet* 114(2), pp. 235-244.

Myers, R. H. et al. 1996. Apolipoprotein E epsilon4 association with dementia in a population-based study: The Framingham study. *Neurology* 46(3), pp. 673-677.

Nacmias, B. et al. 1999. 5-HT2A receptor gene polymorphisms in anorexia nervosa and bulimia nervosa. *Neurosci Lett* 277(2), pp. 134-136.

Nacmias, B. et al. 2001. Association between 5-HT(2A) receptor polymorphism and psychotic symptoms in Alzheimer's disease. *Biol Psychiatry* 50(6), pp. 472-475.

Nalbantoglu, J. et al. 1994. Predictive value of apolipoprotein E genotyping in Alzheimer's disease: results of an autopsy series and an analysis of several combined studies. *Ann Neurol* 36(6), pp. 889-895.

Neary, D. et al. 2005. Frontotemporal dementia. *Lancet Neurol* 4(11), pp. 771-780.

Nee, L. E. et al. 1987. Dementia of the Alzheimer type: clinical and family study of 22 twin pairs. *Neurology* 37(3), pp. 359-363.

Neumann, P. J. et al. 2001. Measuring Alzheimer's disease progression with transition probabilities: estimates from CERAD. *Neurology* 57(6), pp. 957-964.

Newman, A. B. et al. 2005. Dementia and Alzheimer's disease incidence in relationship to cardiovascular disease in the Cardiovascular Health Study cohort. *J Am Geriatr Soc* 53(7), pp. 1101-1107.

Nicoll, J. A. et al. 1995. Apolipoprotein E epsilon 4 allele is associated with deposition of amyloid beta-protein following head injury. *Nat Med* 1(2), pp. 135-137.

North, B. V. et al. 2002. A note on the calculation of empirical P values from Monte Carlo procedures. *Am J Hum Genet* 71(2), pp. 439-441.

Nussbaum, R. L. and Ellis, C. E. 2003. Alzheimer's disease and Parkinson's disease. *N Engl J Med* 348(14), pp. 1356-1364.

O'Connell, J. R. and Weeks, D. E. 1998. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63(1), pp. 259-266.

O'Connor, D. W. et al. 1996. Cross-national interrater reliability of dementia diagnosis in the elderly and factors associated with disagreement. *Neurology* 47(5), pp. 1194-1199.

Oetting, W. S. et al. 1995. Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics* 30(3), pp. 450-458.

Olson, J. M. 1999. A general conditional-logistic model for affected-relative-pair linkage studies. *Am J Hum Genet* 65(6), pp. 1760-1769.

Olson, J. M. et al. 2001. The amyloid precursor protein locus and very-late-onset Alzheimer disease. *Am J Hum Genet* 69(4), pp. 895-899.

Olson, J. M. et al. 2002. A second locus for very-late-onset Alzheimer disease: a genome scan reveals linkage to 20p and epistasis between 20p and the amyloid precursor protein region. *Am J Hum Genet* 71(1), pp. 154-161.

Ostbye, T. and Crosse, E. 1994. Net economic costs of dementia in Canada. *Cmaj* 151(10), pp. 1457-1464.

Ott, A. et al. 1995. Prevalence of Alzheimer's disease and vascular dementia: association with education. The Rotterdam study. *Bmj* 310(6985), pp. 970-973.

Ott, A. et al. 1998. Incidence and risk of dementia. The Rotterdam Study. *Am J Epidemiol* 147(6), pp. 574-580.

Ott, B. R. et al. 1996. Gender differences in the behavioral manifestations of Alzheimer's disease. *J Am Geriatr Soc* 44(5), pp. 583-587.

Overall, J. E. and Gorman, D. R. 1962. The Brief Psychiatric Rating Scale. *Psychological Reports* 10, pp. 799-812.

Parker, G. 2000. Classifying depression: should paradigms lost be regained? *Am J Psychiatry* 157(8), pp. 1195-1203.

Pato, C. N. et al. 2004. Genome-wide scan in Portuguese Island families implicates multiple loci in bipolar disorder: fine mapping adds support on chromosomes 6 and 11. *Am J Med Genet* 127B(1), pp. 30-34.

Pauls, D. L. et al. 1995. A family study of obsessive-compulsive disorder. *Am J Psychiatry* 152(1), pp. 76-84.

Paulsen, J. S. et al. 2000. Incidence of and risk factors for hallucinations and delusions in patients with probable AD. *Neurology* 54(10), pp. 1965-1971.

Payami, H. et al. 1997. A prospective study of cognitive health in the elderly (Oregon Brain Aging Study): effects of family history and apolipoprotein E genotype. *Am J Hum Genet* 60(4), pp. 948-956.

Paykel, E. S. et al. 1994. Incidence of dementia in a population older than 75 years in the United Kingdom. *Arch Gen Psychiatry* 51(4), pp. 325-332.

Paykel, E. S. et al. 1998. Incidence of dementia and cognitive decline in over-75s in Cambridge: overview of cohort study. *Soc Psychiatry Psychiatr Epidemiol* 33(8), pp. 387-392.

Pearlson, G. D. et al. 1990. Association between family history of affective disorder and the depressive syndrome of Alzheimer's disease. *Am J Psychiatry* 147(4), pp. 452-456.

Pedersen, N. L. et al. 2004. How heritable is Alzheimer's disease late in life? Findings from Swedish twins. *Ann Neurol* 55(2), pp. 180-185.

Pedersen, N. L. et al. 2001. Multiple-threshold models for genetic influences on age of onset for Alzheimer disease: findings in Swedish twins. *Am J Med Genet* 105(8), pp. 724-728.

Perez-Madrinan, G. et al. 2004. Alzheimer disease with psychosis: excess cognitive impairment is restricted to the misidentification subtype. *Am J Geriatr Psychiatry* 12(5), pp. 449-456.

Pericak-Vance, M. A. et al. 1997. Complete genomic screen in late-onset familial Alzheimer disease. Evidence for a new locus on chromosome 12. *Jama* 278(15), pp. 1237-1241.

Pericak-Vance, M. A. et al. 2000. Identification of novel genes in late-onset Alzheimer's disease. *Exp Gerontol* 35(9-10), pp. 1343-1352.

Petersen, R. C. et al. 2001. Current concepts in mild cognitive impairment. *Arch Neurol* 58(12), pp. 1985-1992.

Plassman, B. L. et al. 1995. Intelligence and education as predictors of cognitive state in late life: a 50-year follow-up. *Neurology* 45(8), pp. 1446-1450.

Porter, V. R. et al. 2003. Frequency and characteristics of anxiety among patients with Alzheimer's disease and related dementias. *J Neuropsychiatry Clin Neurosci* 15(2), pp. 180-186.

Potash, J. B. et al. 2003. Familial aggregation of psychotic symptoms in a replication set of 69 bipolar disorder pedigrees. *Am J Med Genet B Neuropsychiatr Genet* 116(1), pp. 90-97.

Potash, J. B. et al. 2001. The familial aggregation of psychotic symptoms in bipolar disorder pedigrees. *Am J Psychiatry* 158(8), pp. 1258-1264.

Pritchard, J. K. 2001. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 69(1), pp. 124-137.

Procter, A. W. et al. 1992. Serotonergic pathology is not widespread in Alzheimer patients without prominent aggressive symptoms. *Neurochem Res* 17(9), pp. 917-922.

Purandare, N. et al. 2001. Depressive symptoms in patients with Alzheimer's disease. *Int J Geriatr Psychiatry* 16(10), pp. 960-964.

Qiu, C. et al. 2001. The influence of education on clinically diagnosed dementia incidence and mortality data from the Kungsholmen Project. *Arch Neurol* 58(12), pp. 2034-2039.

Raber, J. et al. 2004. ApoE genotype accounts for the vast majority of AD risk and AD pathology. *Neurobiol Aging* 25(5), pp. 641-650.

Rabins, P. V. et al. 1982. The impact of dementia on the family. *Jama* 248(3), pp. 333-335.

Raiha, I. et al. 1996. Alzheimer's disease in Finnish twins. *Lancet* 347(9001), pp. 573-578.

Rainville, C. et al. 2002. Executive function deficits in patients with dementia of the Alzheimer's type: a study with a Tower of London task. *Arch Clin Neuropsychol* 17(6), pp. 513-530.

Ramachandran, G. et al. 1996. A preliminary study of apolipoprotein E genotype and psychiatric manifestations of Alzheimer's disease. *Neurology* 47(1), pp. 256-259.

Rapoport, M. J. et al. 2001. Relationship of psychosis to aggression, apathy and function in dementia. *Int J Geriatr Psychiatry* 16(2), pp. 123-130.

Rapoport, S. I. et al. 1991. Discordance and concordance of dementia of the Alzheimer type (DAT) in monozygotic twins indicate heritable and sporadic forms of Alzheimer's disease. *Neurology* 41(10), pp. 1549-1553.

Raskind, M. A. et al. 1995. Is early-onset Alzheimer disease a distinct subgroup within the Alzheimer disease population? *Alzheimer Dis Assoc Disord* 9 Suppl 1, pp. S2-6.

Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease 1997 Consensus recommendations for the postmortem diagnosis of Alzheimer's disease The National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease *Neurobiol Aging* 18(4 Suppl), pp S1-2

Reisberg, B et al 1996 Behavioral pathology in Alzheimer's disease (BEHAVE-AD) rating scale *Int Psychogeriatr* 8 Suppl 3, pp 301-308; discussion 351-304

Reisberg, B et al 1987 Behavioral symptoms in Alzheimer's disease: phenomenology and treatment *J Clin Psychiatry* 48 Suppl, pp 9-15

Reisberg, B et al 1982 The Global Deterioration Scale for assessment of primary degenerative dementia *Am J Psychiatry* 139(9), pp 1136-1139

Rice, D P et al 1993 The economic burden of Alzheimer's disease care *Health Aff (Millwood)* 12(2), pp 164-176

Rice, J P 1997 The role of meta-analysis in linkage studies of complex traits *Am J Med Genet* 74(1), pp 112-114

Rice, J P et al 1999 Covariates in linkage analysis *Genet Epidemiol* 17 Suppl 1, pp S691-695

Rigaud, A S et al 2001 Association of the apolipoprotein E epsilon4 allele with late-onset depression *Neuroepidemiology* 20(4), pp 268-272

Rioux, J D et al 2001 Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease *Nat Genet* 29(2), pp 223-228

Risch, N 1990 Linkage strategies for genetically complex traits III The effect of marker polymorphism on analysis of affected relative pairs *Am J Hum Genet* 46(2), pp 242-253

Risch, N J 2000 Searching for genetic determinants in the new millennium *Nature* 405(6788), pp 847-856

Ritchie, K and Kildea, D 1995 Is senile dementia "age-related" or "ageing-related"?--evidence from meta-analysis of dementia prevalence in the oldest old *Lancet* 346(8980), pp 931-934

Rocca, W A et al 1998 Incidence of dementia and Alzheimer's disease: a reanalysis of data from Rochester, Minnesota, 1975-1984 *Am J Epidemiol* 148(1), pp 51-62

Rocca, W A et al 1991 Frequency and distribution of Alzheimer's disease in Europe: a collaborative study of 1980-1990 prevalence findings The EURODEM-Prevalence Research Group *Ann Neurol* 30(3), pp 381-390

Rogaeva, E et al 1998 Evidence for an Alzheimer disease susceptibility locus on chromosome 12 and for further locus heterogeneity *Jama* 280(7), pp 614-618

Rogaeva, E I et al 1995 Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene *Nature* 376(6543), pp 775-778

Rogan, W J and Gladen, B 1978 Estimating prevalence from the results of a screening test *Am J Epidemiol* 107(1), pp 71-76

Rosen, J and Zubenko, G S 1991 Emergence of psychosis and depression in the longitudinal evaluation of Alzheimer's disease *Biol Psychiatry* 29(3), pp 224-232

Roth, M et al 1986 CAMDEX A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia *Br J Psychiatry* 149, pp 698-709

Ruitenberg, A et al 2001 Incidence of dementia: does gender make a difference? *Neurobiol Aging* 22(4), pp 575-580

Ryden, M B 1988 Aggressive behavior in persons with dementia who live in the community *Alzheimer Dis Assoc Disord* 2(4), pp 342-355

Saunders, A M et al 1993 Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease *Neurology* 43(8), pp 1467-1472

Scarmeas, N et al 2002 Association between the APOE genotype and psychopathologic symptoms in Alzheimer's disease *Neurology* 58(8), pp 1182-1188

Schmand, B et al 1998 Apolipoprotein E phenotype is not related to late-life depression in a population-based sample *Soc Psychiatry Psychiatr Epidemiol* 33(1), pp 21-26

Schmechel, D E et al 1993 Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease *Proc Natl Acad Sci U S A* 90(20), pp 9649-9653

Schmidt, R et al 1996 Estrogen replacement therapy in older women: a neuropsychological and brain MRI study *J Am Geriatr Soc* 44(11), pp 1307-1313

Schulze, T G and McMahon, F J 2003 Genetic linkage and association studies in bipolar affective disorder: a time for optimism *Am J Med Genet C Semin Med Genet* 123(1), pp 36-47

Scott, W K et al 2000 Fine mapping of the chromosome 12 late-onset Alzheimer disease locus: potential genetic and phenotypic heterogeneity *Am J Hum Genet* 66(3), pp 922-932

Scott, W K et al 2003 Ordered-subsets linkage analysis detects novel Alzheimer disease loci on chromosomes 2q34 and 15q22 *Am J Hum Genet* 73(5), pp 1041-1051

Selkoe, D J 2001 Alzheimer's disease: genes, proteins, and therapy *Physiol Rev* 81(2), pp 741-766

Senanarong, V et al 2004 Agitation in Alzheimer's disease is a manifestation of frontal lobe dysfunction *Dement Geriatr Cogn Disord* 17(1-2), pp 14-20

Seshadri, S et al 1995 Apolipoprotein E epsilon 4 allele and the lifetime risk of Alzheimer's disease What physicians know, and what they should know *Arch Neurol* 52(11), pp 1074-1079

Sevush, S et al 1993 Age at disease onset and pattern of cognitive impairment in probable Alzheimer's disease *J Neuropsychiatry Clin Neurosci* 5(1), pp 66-72

Sham, P and McGuffin, P 2002 Linkage and association In: McGuffin, P et al eds *Psychiatric Genetics and Genomics* New York: Oxford University Press, pp 55-76

Sham, P C et al 1994 Age at onset, sex, and familial psychiatric morbidity in schizophrenia Camberwell Collaborative Psychosis Study *Br J Psychiatry* 165(4), pp 466-473

Shao, Y et al 2003 Fine mapping of autistic disorder to chromosome 15q11-q13 by use of phenotypic subtypes *Am J Hum Genet* 72(3), pp 539-548

Sherrington, R et al 1995 Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease *Nature* 375(6534), pp 754-760

Shin, I S et al 2005 Neuropsychiatric Symptoms and Quality of Life in Alzheimer Disease *Am J Geriatr Psychiatry* 13(6), pp 469-474

Silverman, J M et al 2005 Variability of familial risk of Alzheimer disease across the late life span *Arch Gen Psychiatry* 62(5), pp 565-573

Silverman, J M et al 2003 Familial patterns of risk in very late-onset Alzheimer disease *Arch Gen Psychiatry* 60(2), pp 190-197

Skoog, I et al 1998 A population study of apoE genotype at the age of 85: relation to dementia, cerebrovascular disease, and mortality *J Neurol Neurosurg Psychiatry* 64(1), pp 37-43

Sleegers, K et al 2004 Familial clustering and genetic risk for dementia in a genetically isolated Dutch population *Brain* 127(Pt 7), pp 1641-1649

Slooter, A J et al 1998 Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study *Arch Neurol* 55(7), pp 964-968

Smoller, J W and Finn, C T 2003 Family, twin, and adoption studies of bipolar disorder *Am J Med Genet C Semin Med Genet* 123(1), pp 48-58

Soininen, H et al 1995 A severe loss of choline acetyltransferase in the frontal cortex of Alzheimer patients carrying apolipoprotein epsilon 4 allele *Neurosci Lett* 187(2), pp 79-82

Song, Y Q et al 1998 Absence of association between Alzheimer disease and the -491 regulatory region polymorphism of APOE *Neurosci Lett* 250(3), pp 189-192

Souetre, E et al 1999 Economic impact of Alzheimer's disease in the United Kingdom Cost of care and disease severity for non-institutionalised patients with Alzheimer's disease *Br J Psychiatry* 174, pp 51-55

Spalletta, G et al 2004 Cognition and behaviour are independent and heterogeneous dimensions in Alzheimer's disease *J Neurol* 251(6), pp 688-695

Speck, C E et al 1995 History of depression as a risk factor for Alzheimer's disease *Epidemiology* 6(4), pp 366-369

Starkstein, S E et al 1997 Prospective longitudinal study of depression and anosognosia in Alzheimer's disease *Br J Psychiatry* 171, pp 47-52

Starkstein, S E et al 2004 Prevalence and clinical correlates of disinhibition in dementia *Cogn Behav Neurol* 17(3), pp 139-147

Starkstein, S E et al 1995 Prevalence and clinical correlates of pathological affective display in Alzheimer's disease *J Neurol Neurosurg Psychiatry* 59(1), pp 55-60

Steele, C et al 1990 Psychiatric symptoms and nursing home placement of patients with Alzheimer's disease *Am J Psychiatry* 147(8), pp 1049-1051

Steffens, D C et al 2003 Apolipoprotein E genotype and major depression in a community of older adults The Cache County Study *Psychol Med* 33(3), pp 541-547

Steffens, D C et al 2000 APOE and AD concordance in twin pairs as predictors of AD in first-degree relatives *Neurology* 54(3), pp 593-598

Stein, L D 2004 Human genome: end of the beginning *Nature* 431(7011), pp 915-916

Steinberg, M et al 2003 The incidence of mental and behavioral disturbances in dementia: the cache county study *J Neuropsychiatry Clin Neurosci* 15(3), pp 340-345

Stern, Y et al 1994a Utility of extrapyramidal signs and psychosis as predictors of cognitive and functional decline, nursing home admission, and death in Alzheimer's disease: prospective analyses from the Predictors Study *Neurology* 44(12), pp 2300-2307

Stern, Y et al 1994b Influence of education and occupation on the incidence of Alzheimer's disease *Jama* 271(13), pp 1004-1010

Strauss, M E and Ogrocki, P K 1996 Confirmation of an association between family history of affective disorder and the depressive syndrome in Alzheimer's disease *Am J Psychiatry* 153(10), pp 1340-1342

Street, J S et al 2001 Long-term efficacy of olanzapine in the control of psychotic and behavioral symptoms in nursing home patients with Alzheimer's dementia *Int J Geriatr Psychiatry* 16 Suppl 1, pp S62-70

Strober, M et al 1988 A family study of bipolar I disorder in adolescence Early onset of symptoms linked to increased familial loading and lithium resistance *J Affect Disord* 15(3), pp 255-268

Sukonick, D L et al 2001 The 5-HTTPR*S/*L polymorphism and aggressive behavior in Alzheimer disease *Arch Neurol* 58(9), pp 1425-1428

Sullivan, P F et al 2000 Genetic epidemiology of major depression: review and meta-analysis *Am J Psychiatry* 157(10), pp 1552-1562

Sultzer, D L et al 1995 The relationship between psychiatric symptoms and regional cortical metabolism in Alzheimer's disease *J Neuropsychiatry Clin Neurosci* 7(4), pp 476-484

Sweet, R A et al 2005 Catechol-O-methyltransferase haplotypes are associated with psychosis in Alzheimer disease *Mol Psychiatry* 10(11), pp 1026-1036

Sweet, R A et al 2003 Psychotic symptoms in Alzheimer disease: evidence for a distinct phenotype *Mol Psychiatry* 8(4), pp 383-392

Sweet, R A et al 2002 Increased familial risk of the psychotic phenotype of Alzheimer disease *Neurology* 58(6), pp 907-911

Sweet, R A et al 1998 Dopamine receptor genetic variation, psychosis, and aggression in Alzheimer disease *Arch Neurol* 55(10), pp 1335-1340

Taket, S R 1992 Resource allocation problems and health services for the elderly *World Health Statistics Quarterly* 45, pp 89-94

Talbot, C et al 1994 Protection against Alzheimer's disease with apoE epsilon 2 *Lancet* 343(8910), pp 1432-1433

Tandon, A et al 2000 Molecular genetics of Alzheimer's disease: the role of beta-amyloid and the presenilins *Curr Opin Neurol* 13(4), pp 377-384

Tang, M X et al 2001 Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan *Neurology* 56(1), pp 49-56

Tariot, P N et al 1995 The Behavior Rating Scale for Dementia of the Consortium to Establish a Registry for Alzheimer's Disease The Behavioral Pathology Committee of the Consortium to Establish a Registry for Alzheimer's Disease *Am J Psychiatry* 152(9), pp 1349-1357

Tekin, S et al 2001a Activities of daily living in Alzheimer's disease: neuropsychiatric, cognitive, and medical illness influences *Am J Geriatr Psychiatry* 9(1), pp 81-86

Tekin, S et al 2001b Orbitofrontal and anterior cingulate cortex neurofibrillary tangle burden is associated with agitation in Alzheimer disease *Ann Neurol* 49(3), pp 355-361

Teri, L et al 1999 Anxiety of Alzheimer's disease: prevalence, and comorbidity *J Gerontol A Biol Sci Med Sci* 54(7), pp M348-352

Teri, L et al 1990 Cognitive deterioration in Alzheimer's disease: behavioral and health factors *J Gerontol* 45(2), pp P58-63

Tsuang, D et al 2000 Familial aggregation of psychotic symptoms in Huntington's disease *Am J Psychiatry* 157(12), pp 1955-1959

Tsuang, M T et al 1993 Identification of the phenotype in psychiatric genetics *Eur Arch Psychiatry Clin Neurosci* 243(3-4), pp 131-142

Tunstall, N et al 2000 Familial influence on variation in age of onset and behavioural phenotype in Alzheimer's disease *Br J Psychiatry* 176, pp 156-159

Turvey, C L et al 2001 Caregiver report of hallucinations and paranoid delusions in elders aged 70 or older *Int Psychogeriatr* 13(2), pp 241-249

van Duijn, C M et al 1991 Familial aggregation of Alzheimer's disease and related disorders: a collaborative re-analysis of case-control studies EURODEM Risk Factors Research Group *Int J Epidemiol* 20 Suppl 2, pp S13-20

Van Eerdewegh, P et al 2002 Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness *Nature* 418(6896), pp 426-430

Villareal, D T and Morris, J C 1999 The diagnosis of Alzheimer's disease *J Alzheimers Dis* 1(4-5), pp 249-263

von Strauss, E et al 1999 Aging and the occurrence of dementia: findings from a population-based cohort with a large sample of nonagenarians *Arch Neurol* 56(5), pp 587-592

Waite, L M et al 2001 The incidence of dementia in an Australian community population: the Sydney Older Persons Study *Int J Geriatr Psychiatry* 16(7), pp 680-689

Wang, J C et al 2000 Effect of APOE genotype and promoter polymorphism on risk of Alzheimer's disease *Neurology* 55(11), pp 1644-1649

Wang, Y C et al 2001 No association between tryptophan hydroxylase gene polymorphism and Alzheimer's disease *Neuropsychobiology* 43(1), pp 1-4

Warwick Daw, E et al 2000 The number of trait loci in late-onset Alzheimer disease *Am J Hum Genet* 66(1), pp 196-204

Webb, E L et al 2005 SNPLINK: multipoint linkage analysis of densely distributed SNP data incorporating automated linkage disequilibrium removal *Bioinformatics* 21(13), pp 3060-3061

Webster, D 1968 Critical analysis of the disability in Parkinson's disease *Modern Treatment*, pp 257-282

Weiner, M F et al 2005 Early behavioral symptoms and course of Alzheimer's disease *Acta Psychiatr Scand* 111(5), pp 367-371

Weissman, M M et al 1984 Psychiatric disorders in the relatives of probands with affective disorders The Yale University--National Institute of Mental Health Collaborative Study *Arch Gen Psychiatry* 41(1), pp 13-21

Whitehouse, P J 1995 Should we differentiate subtypes of Alzheimer disease? *Alzheimer Dis Assoc Disord* 9 Suppl 1, p S1

Whittemore, A S and Halpern, J 2001 Problems in the definition, interpretation, and evaluation of genetic heterogeneity *Am J Hum Genet* 68(2), pp 457-465

Wichers, M and Maes, M 2002 The psychoneuroimmuno-pathophysiology of cytokine-induced depression in humans *Int J Neuropsychopharmacol* 5(4), pp 375-388

Wickham, H et al 2002 Familiality of clinical characteristics in schizophrenia *J Psychiatr Res* 36(5), pp 325-329

Williams, J et al 1997 Meta-analysis of association between the 5-HT2a receptor T102C polymorphism and schizophrenia EMASS Collaborative Group European Multicentre Association Study of Schizophrenia *Lancet* 349(9060), p 1221

Wilson, R S et al 2000 Hallucinations, delusions, and cognitive decline in Alzheimer's disease *J Neurol Neurosurg Psychiatry* 69(2), pp 172-177

Winkleby, M A et al 1992 Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease *Am J Public Health* 82(6), pp 816-820

Wobrock, T et al 2003 [Presentation and stability of cognitive and noncognitive symptom patterns in patients with Alzheimer's disease Disease course over a two-year period under constant treatment conditions with rivastigmine] *Fortschr Neurol Psychiatr* 71(4), pp 199-204

Wu, W S et al 1998a Genetic studies on chromosome 12 in late-onset Alzheimer disease *Jama* 280(7), pp 619-622

Wu, Z et al 1998b Role of familial factors in late-onset Alzheimer disease as a function of age *Alzheimer Dis Assoc Disord* 12(3), pp 190-197

Yoshitake, T et al 1995 Incidence and risk factors of vascular dementia and Alzheimer's disease in a defined elderly Japanese population: the Hisayama Study *Neurology* 45(6), pp 1161-1168

Zaccai, J et al 2005 A systematic review of prevalence and incidence studies of dementia with Lewy bodies *Age Ageing* 34(6), pp 561-566

Zanardini, R et al 2003 Association between IL-1beta -511C/T and IL-1RA (86bp)n repeats polymorphisms and schizophrenia *J Psychiatr Res* 37(6), pp 457-462

Zhang, M Y et al 1990 The prevalence of dementia and Alzheimer's disease in Shanghai, China: impact of age, gender, and education *Ann Neurol* 27(4), pp 428-437

Zubenko, G S 1997 Molecular neurobiology of Alzheimer's disease (syndrome?) *Harv Rev Psychiatry* 5(4), pp 177-213

Zubenko, G S et al 1996 Association of the APOE epsilon 4 allele with clinical subtypes of late life depression *Biol Psychiatry* 40(10), pp 1008-1016

Zubenko, G S et al 1998 A genome survey for novel Alzheimer disease risk loci: results at 10-cM resolution *Genomics* 50(2), pp 121-128