



# An estimate of amyotrophic lateral sclerosis heritability using twin data

A Al-Chalabi,<sup>1</sup> F Fang,<sup>2</sup> M F Hanby,<sup>1</sup> P N Leigh,<sup>1</sup> C E Shaw,<sup>1</sup> W Ye,<sup>2</sup> F Rijdsdijk<sup>3</sup>

See Editorial commentary, p 1299

<sup>1</sup>King's College London, MRC Centre for Neurodegeneration Research, Institute of Psychiatry, London, UK

<sup>2</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

<sup>3</sup>MRC Centre for Social, Genetic and Developmental Psychiatry, NIHR Biomedical Research Centre for Mental Health, King's College London, London, UK

## Correspondence to

Professor Ammar Al-Chalabi, MRC Centre for Neurodegeneration Research, P 041, Institute of Psychiatry, King's College London, London SE5 8AF, UK; ammar.al-chalabi@kcl.ac.uk

Received 27 January 2010

Revised 21 July 2010

Accepted 22 July 2010

Published Online First

22 September 2010

## ABSTRACT

**Background** Causative gene mutations have been identified in about 2% of those with amyotrophic lateral sclerosis (ALS), often, but not always, when there is a strong family history. There is an assumption that there is a genetic component to all ALS, but genome-wide association studies have yet to produce a robustly replicated result. A definitive estimate of ALS heritability is therefore required to determine whether ongoing efforts to find susceptibility genes are worth while.

**Methods** The authors performed two twin studies, one population- and one clinic-based. The authors used structural equation modelling to perform a meta-analysis of data from these studies and an existing twin study to estimate ALS heritability, and identified 171 twin pairs in which at least one twin had ALS.

**Results and discussion** Five monozygotic twin pairs were concordant-affected, and 44 discordant-affected. No dizygotic twin pairs were concordant-affected, and 122 discordant-affected. The heritability of sporadic ALS was estimated as 0.61 (0.38 to 0.78) with the unshared environmental component 0.39 (0.22 to 0.62). ALS has a high heritability, and efforts to find causative genes should continue.

## INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of motor neurons with a median survival of about 2 years. The lifetime prevalence is about 1 in 400,<sup>1</sup> but the appalling prognosis means the point prevalence is only 5 per 100 000 persons.<sup>2</sup> In 95% of cases, there is no family history of ALS, but a study has found that the risk to siblings or offspring of a proband is elevated between nine and 17 times.<sup>3</sup> There is thus a familial tendency to ALS, which is best regarded as a complex disease.

There have been many genetic studies of ALS,<sup>4</sup> but the only consistent findings have been the identification of six genes in which mutation predisposes to ALS, both in families and in sporadic cases, accounting for about 2% of ALS overall; these are *SOD1*, *TARDBP*, *FUS*, *VAPB*, *ANG* and *OPTN*.<sup>5–12</sup> Genome-wide association studies have not resulted in independent replication of findings, although replication of a large study in a second large internal cohort has been successful.<sup>13</sup> There is, nevertheless, the assumption of a genetic component to all ALS. A twin study would help determine whether this assumption is reasonable by allowing an estimate of ALS heritability. The British Motor Neuron Disease (MND) Twin Study<sup>14</sup> made a start on this process by reporting the likely range in which the ALS heritability value would lie: 0.38 to 0.85. We have therefore collected twin data from two further sources: the Swedish Twin Registry<sup>15</sup>

and the King's College Hospital Motor Neuron Disease Clinic Register.<sup>16</sup> We have analysed these two new data sets with the existing data used for the British MND Twin Study, independently but simultaneously in a meta-analysis framework, to generate for the first time an estimate of ALS heritability.

## METHODS

### Data collection

Data for the British MND Twin Study were extracted from the published manuscript.<sup>14</sup> Methods for the Swedish Twin Registry and King's Register have been previously described.<sup>3, 16</sup> For the King's Register, pedigree records as obtained at the initial visit of each participant were used to identify twins. Zygosity status was self-reported, a method shown in the Swedish Twin Registry to be correct in 99% of twin pairs.<sup>15</sup> Concordance was defined as reported in the pedigree at the time of last follow-up. The two new studies were approved by the Research Ethics Committees of each institution. Informed written consent to participate in research was obtained for all participants.

### Familiality

Although the distinction between familial and sporadic ALS is to a large extent artificial, the population under study consists of two groups: those with a strong family history of ALS in whom we can assume a strong genetic component, and those without, in which the demonstration of a heritable component would be most useful. We therefore tested the effect of including and excluding individuals with a family history of ALS in a first-degree relative. For the British MND Twin Study, this information was provided in the published information. For the Swedish Twin Registry, data were linked to the Swedish Multi-Generational Register, which has family history information for those born after 1932 who were still alive in or after 1967, the Swedish Inpatient Register (1973 to 2004) and the Swedish Causes of Death Register (1961 to 2003) to identify relatives with ALS. For the King's Register, pedigree structure and family history taken both at the initial visit and at subsequent follow-up were used to identify families with affected first degree relatives.

### Statistical methods

Genetic model-fitting is based on biometrical genetic theory where contributions of additive genetic (A), dominant genetic (D), shared environmental (C) and unshared environmental (E) effects on the trait are estimated by examining how often twins of different types correlate. Monozygotic



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://jnnp.bmj.com/site/about/unlocked.xhtml>

(MZ) twins share all their genes, while dizygotic (DZ) twins share only half on average. A trait that is strongly genetic would therefore be expected to be strongly correlated between MZ twins, but show about half as strong a correlation between DZ twins. Calculation of correlations between twins for a continuous trait, such as height, is straightforward, but when the trait under study is discrete, as is ALS, a statistical measure called tetrachoric correlation is used, which assumes an underlying normal distribution of liability with a threshold effect determining if someone is affected or not. A DZ correlation which is less than half that of the MZ suggests the effect of dominance, whereas a DZ correlation which is higher than half that of the MZ suggests the effect of shared environment.<sup>17</sup> As these two sources of variance are confounded, ACE and ADE models are fitted separately. Model fit is indicated by  $\chi^2$  statistics and related p values.

We performed a meta-analysis to maximise power. Twins of unknown zygosity were excluded from analysis. Tetrachoric correlations for ALS for MZ and DZ twins were calculated from contingency tables summarising the number of concordant-affected, concordant-unaffected and discordant pairs, using the Mx program.<sup>18</sup> The calculations assume that the joint liability distribution is bivariate normal, and a threshold determines affection or normality. Data for the British MND Twin Study and the King's Register were derived from local case register information on ALS, rather than being direct population-based observations, as is the case for the Swedish Twin Registry. The number of concordant-unaffected pairs is therefore unknown, requiring ascertainment correction and the liability threshold to be fixed to the prevalence of ALS. Population studies put the median age of onset of ALS at about 70 years of age,<sup>19</sup> so the prevalence of ALS was taken as about 1 in 740 (z-value threshold=3.00), which is the approximate lifetime risk to this age.<sup>1</sup> No such corrections were necessary for the Swedish data. The combined data were then used to estimate the overall MZ and DZ twin correlations. Because there were no concordant-affected DZ twin pairs in any study, we also modelled the effect of the next DZ twin pair sampled being concordant-affected.<sup>20</sup>

Different sex twin pairs were included in the analysis, even though there is a differential risk for males and females with ALS, to avoid a loss of power and to allow comparison with the British MND Twin Study, which did not differentiate between twins by sex.

## RESULTS

### Numbers studied

The British MND Twin Study examined 10 872 death certificates of people with MND, identifying 75 twin pairs for further study. It excluded seven twin-pairs where one of the pair had died before the proband developed disease, and also excluded one pair

on the basis they were from a family with known familial ALS. It is arguable whether these are valid reasons for exclusion, and so we did not exclude such pairs. Of 26 MZ probands identified, four were double twin pairs and therefore counted twice. Thus, the original study analysed 20 discordant and one concordant MZ pair, and 44 discordant and no concordant DZ pairs. We analysed 22 discordant and two concordant-affected MZ pairs, and 51 discordant and no concordant-affected DZ pairs overall, and excluded one MZ concordant pair for analysis of sporadic ALS.

The Swedish Twin Registry study examined 86 441 twin pairs, of which 14 824 were MZ, 43 070 were DZ and 28 547 were of unknown zygosity. Of the MZ pairs, 15 were discordant and two concordant for ALS. Of the DZ pairs, 58 were discordant and none concordant for ALS. For six twins, no family history data were available. For the remainder, no family history of ALS was identified in first-degree non-twin relatives.

The King's Register contained records of 1502 individuals with ALS and 3480 first-degree relatives. There were 21 twin pairs in which at least one twin had ALS. There were eight MZ pairs, one concordant and seven discordant for ALS, and 13 DZ pairs, all discordant. The concordant MZ pair and one of the DZ pairs had a family history of ALS in a first degree relative and were excluded from the analysis of sporadic ALS.

### Model results

Results for all fitted genetic models are presented in table 1.

The model estimate of prevalence for ALS in the Swedish data (threshold z-value=3.21) was close to the fixed threshold chosen for the British MND Twin Study and King's Register based on life-time prevalence by age 70. The MZ and DZ twin correlations were 0.73 (0.55 to 0.85) and 0.16 (−0.24 to 0.47), respectively for the twins including familial cases and 0.65 (0.42 to 0.81) and 0.16 (−0.51 to 0.47) for sporadic cases only, indicating dominance genetic effects in both cases. For sporadic ALS, although an ADE model fitted the data well ( $\chi^2_{(2)}=3.16$ ,  $p=0.21$ ), the heritability of ALS could not be meaningfully decomposed into additive and dominance genetic effects, most likely because of the lack of concordant-affected DZ twins. To resolve this, we explored two alternative procedures. First, we interpreted an AE model. This fitted the data well ( $\chi^2_{(3)}=4.60$ ,  $p=0.20$ ) with heritability estimated at 0.61 (0.38 to 0.78). Second, we adopted an approach used by the British MND Twin Study and others: all studies were assumed to have a single concordant-affected DZ twin pair as the next studied. In this situation, the MZ and DZ twin correlations were 0.65 (0.42 to 0.81) and 0.47 (0.27 to 0.63) respectively, indicating shared environmental effects. Although the ACE model fitted the data well ( $\chi^2_{(2)}=3.65$ ,  $p=0.16$ ), the 95% CI for the shared environmental component included zero. The reduced AE model

**Table 1** Genetic model-fitting results: variance components (with 95% CI)

Genetic model	Twins included	Modelled as next dizygotic pair concordant	Additive genetic component (A)	Common environmental (C) or dominant genetic (D) component	Unique environmental component (E)
ADE	All	No	0.00 (0.00 to 0.80)	0.73 (0.00 to 0.85)	0.27 (0.15 to 0.45)
AE	All	No	0.71 (0.53 to 0.84)	—	0.29 (0.16 to 0.47)
ACE	All	Yes	0.52 (0.04 to 0.85)	0.21 (0.00 to 0.58)	0.27 (0.15 to 0.45)
AE	All	Yes	0.76 (0.60 to 0.86)	—	0.24 (0.14 to 0.40)
ADE	Sporadic	No	0.00 (0.00 to 0.75)	0.64 (0.00 to 0.81)	0.36 (0.19 to 0.59)
AE	Sporadic	No	0.61 (0.38 to 0.78)	—	0.39 (0.22 to 0.62)
ACE	Sporadic	Yes	0.36 (0.00 to 0.80)	0.30 (0.00 to 0.63)	0.35 (0.19 to 0.57)
AE	Sporadic	Yes	0.70 (0.52 to 0.83)	—	0.30 (0.17 to 0.48)

(dropping C) showed an additive genetic component of 0.70 (0.52 to 0.83).

## DISCUSSION

We estimate the heritability of sporadic ALS to be 0.61 (0.38 to 0.78). The strength of this study lies in the extra power available from combining the two new studies and the previous British study in a meta-analysis. This has allowed us to independently but simultaneously model the different studies, despite their different designs. The value we obtain and the confidence interval lie within the range predicted by the British MND Twin Study. Although it can be difficult to estimate heritability in relatively rare diseases, our study combines data from 10 872 MND death certificates over a 10-year period, a country-wide study of 86 441 twin pairs and a clinic-based study of 4982 individuals with ALS or their family members, generating results from 171 ALS twin pairs.

Our analysis differs from the British MND Twin study in an important aspect. We used structural equation modelling.<sup>21</sup> This is a powerful likelihood-based technique that allows an estimate of heritability despite the lack of concordant-affected DZ twin pairs, without resorting to statistical workarounds. The British MND Twin study used an accepted method in which the next affected DZ twin observed is assumed to come from a concordant pair.<sup>20</sup> When we adopted this approach, the results were not very different from our unmodified analysis (table 1). This is important, as it suggests that a moderate increase in sample size designed to identify more twin pairs will not greatly change our results.

When familial samples were included in our analysis, the heritability estimate was, as expected, higher, but the overall findings are not greatly different from those for sporadic ALS, whether or not we allowed for the next DZ twins identified to be concordant-affected.

We have previously estimated the heritability of age of onset of *SOD1*-mediated ALS as 0.29 (0 to 0.42) using a variance components approach in ALS pedigrees.<sup>22</sup> This contrasts with the much higher estimate for susceptibility heritability we obtain here, indicating independent genetic effects on age of onset and susceptibility.

Comorbidity of ALS and frontotemporal dementia, and pathological overlap is well recognised. No formal twin study has been performed for frontotemporal dementia, but in more than 41% of cases, there is some family history, suggesting a high heritability. We did not distinguish between individuals with and without frontotemporal dementia concurrent with ALS, and our heritability estimate will therefore also have implications for frontotemporal dementia.

This study has important implications for sporadic ALS. Clinically, the knowledge of a significant genetic component means that the question of why ALS has developed can be more confidently explained in terms of complex disease genetics. For research, despite the difficulties with replication of genome-wide association studies,<sup>4</sup> we can pursue more highly powered genome-wide association studies boldly, explore statistical methods for interactions and epistasis, and use techniques to find rare genetic variants, copy number variants, microsatellites and epigenetic changes.

**Acknowledgements** We thank the NIHR Centre for Biomedical Research.

**Funding** FF and WY were supported by Bjorklunds Fond 2008-20892. MFH was supported by the British Geriatric Society. We thank the Medical Research Council (UK), the Wellcome Trust, the Motor Neurone Disease Association of Great Britain and Northern Ireland, The American ALS Association, the ALS Therapy Alliance and the Angel Fund for support.

**Competing interests** None.

**Patient consent** Obtained.

**Ethics approval** Ethics approval was provided by the King's College Hospital and SLAM/IOP Research Ethics Committees for the UK participants. For the Swedish Twin Registry, the ethical advice from the Research Ethics Committee of the Karolinska Institutet was that specific approval was not required to use the data. For the British MND Twin Study (1997) we used published data.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1. Johnston CA, Stanton BR, Turner MR, *et al*. Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. *J Neurol* 2006;**253**:1642–3.
2. Beghi E, Logroscino G, Chio A, *et al*. The epidemiology of ALS and the role of population-based registries. *Biochim Biophys Acta* 2006;**1762**:1150–7.
3. Fang F, Kamel F, Lichtenstein P, *et al*. Familial aggregation of amyotrophic lateral sclerosis. *Ann Neurol* 2009;**66**:94–9.
4. Dion PA, Daoud H, Rouleau GA. Genetics of motor neuron disorders: new insights into pathogenic mechanisms. *Nat Rev Genet* 2009;**10**:769–82.
5. Rosen DR, Siddique T, Patterson D, *et al*. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 1993;**362**:59–62.
6. Sreedharan J, Blair IP, Tripathi VB, *et al*. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* 2008;**319**:1668–72.
7. Kabashi E, Valdmanis PN, Dion P, *et al*. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat Genet* 2008;**40**:572–4.
8. Vance C, Rogelj B, Hortobagyi T, *et al*. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 2009;**323**:1208–11.
9. Kwiatkowski TJ Jr, Bosco DA, Leclerc AL, *et al*. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 2009;**323**:1205–8.
10. Nishimura AL, Mitne-Neto M, Silva HC, *et al*. A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. *Am J Hum Genet* 2004;**75**:822–31.
11. Greenway MJ, Andersen PM, Russ C, *et al*. ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nat Genet* 2006;**38**:411–13.
12. Maruyama H, Morino H, Ito H, *et al*. Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* 2010;**465**:223–6.
13. van Es MA, Veldink JH, Saris CG, *et al*. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* 2009;**41**:1083–7.
14. Graham AJ, Macdonald AM, Hawkes CH. British motor neuron disease twin study. *J Neurol Neurosurg Psychiatry* 1997;**62**:562–9.
15. Lichtenstein P, De Faire U, Floderus B, *et al*. The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies. *J Intern Med* 2002;**252**:184–205.
16. Wijesekera LC, Mathers S, Talman P, *et al*. Natural history and clinical features of the flail arm and flail leg ALS variants. *Neurology* 2009;**72**:1087–94.
17. Christian JC, Norton JA Jr, Sorbel J, *et al*. Comparison of analysis of variance and maximum likelihood based path analysis of twin data: partitioning genetic and environmental sources of covariance. *Genet Epidemiol* 1995;**12**:27–35.
18. Neale MC, Boker SM, Xie G, *et al*. *Mx: Statistical Modeling*. 6th edn. Richmond, VA: VCU Box 900126 23298: Department of Psychiatry, 2003.
19. Chio A, Mora G, Calvo A, *et al*. Epidemiology of ALS in Italy: a 10-year prospective population-based study. *Neurology* 2009;**72**:725–31.
20. Bailey A, Le Couteur A, Gottesman I, *et al*. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 1995;**25**:63–77.
21. Karlin S, Cameron EC, Chakraborty R. Path analysis in genetic epidemiology: a critique. *Am J Hum Genet* 1983;**35**:695–732.
22. Fogh I, Rijsdijk F, Andersen PM, *et al*. Age at onset in *sod1*-mediated amyotrophic lateral sclerosis shows familiarity. *Neurogenetics* 2007;**8**:235–6.