



Neurodegenerative disease in *C9orf72* repeat expansion carriers: population risk and effect of *UNC13A*

✉ Jiali Gao,¹ Andrew G. L. Douglas,^{1,2} Christos V. Chalitsios,¹ ✉ Jakub Scaber,¹
✉ Kevin Talbot,¹ ✉ Martin R. Turner¹ and ✉ Alexander G. Thompson¹

The *C9orf72* hexanucleotide repeat expansion (HRE) is the most common monogenetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Neurodegenerative disease incidence in *C9orf72* HRE carriers has been studied using cohorts from disease-affected families or by extrapolating from population disease incidence, potentially introducing bias.

Age-specific cumulative incidence of ALS and dementia was estimated using Kaplan–Meier and competing risk models in *C9orf72* HRE carriers compared to matched controls in UK Biobank. Risk modification by *UNC13A* genotype was examined.

Of 490 331 individuals with valid genetic data, 701 had >100 repeats in *C9orf72* [median age 55 (interquartile range 48–62), follow-up 13.4 years (12.3–14.1)]. The cumulative incidence of ALS or dementia was 66% (95% confidence interval 57%–73%) by age 80 in *C9orf72* HRE carriers versus 5.8% (4.5%–7.0%) in controls, or 58% (50%–64%) versus 5.1% (4.1%–6.4%), accounting for the competing risk of other-cause mortality. Forty-one per cent of dementia incidence accrued between age 75–80. C-allele homozygosity at rs12608932 in *UNC13A* increased ALS or dementia risk in *C9orf72* HRE carriers [hazard ratio 1.81 (1.18–2.78)].

C9orf72 HRE disease was incompletely penetrant in this population-based cohort, with risk modified by *UNC13A* genotype. This has implications for counselling at-risk individuals and modelling expected phenoconversion for prevention trials.

1 Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK

2 Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Foundation Trust, Oxford OX3 9DU, UK

Correspondence to: Alexander G. Thompson

Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital

Level 6 West Wing, Oxford OX3 9DU, UK

E-mail: alexander.thompson@ndcn.ox.ac.uk

Keywords: amyotrophic lateral sclerosis; frontotemporal dementia; *C9orf72*; neurodegeneration; *UNC13A*

Introduction

A hexanucleotide repeat expansion (HRE) in an intronic region of *C9orf72* is the most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) in European-ancestry populations, present in around 10% of ALS

and FTD cases.^{1–3} Individuals with *C9orf72* ALS or FTD are indistinguishable from their sporadic counterparts, but, in aggregate, have a younger onset,^{2,4} more equal sex distribution⁵ and shorter survival.^{2,4,5} The effect of repeat length on phenotype remains uncertain.⁶ *C9orf72* HRE has also been reported in parkinsonism,^{7,8} amnesic dementia^{9,10} and Huntington's disease-like syndrome,⁹

Received March 05, 2025. Revised June 16, 2025. Accepted June 19, 2025. Advance access publication July 19, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of the Guarantors of Brain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

but, given a population prevalence of 1 in 700,⁹ the association between *C9orf72* HRE and these phenotypes is unresolved.

Accurately defining disease incidence in *C9orf72* HRE carriers has implications for genetic counselling, the mechanistic understanding of *C9orf72*-related neurodegeneration and the planning of preventative trials.^{11,12} Estimates have varied widely due to methodological differences. Cohort studies have estimated the penetrance of ALS-FTD in *C9orf72* HRE carriers as near-complete by age 80,^{1,13} but are limited by sample size and biased by recruitment from families with higher penetrance. Indirect estimates calculated from the incidence of ALS in first-degree relatives of *C9orf72* HRE carriers or from the population incidence of ALS-FTD and population prevalence of *C9orf72* HRE have suggested a penetrance of 24%–33%,^{14,15} but are reliant on the accuracy of these underlying statistics and do not include other potential disease phenotypes.

Family history is not universal in *C9orf72*-associated ALS and FTD cases. The marked variability in penetrance between families implies that genetic modifiers, particularly common variants, may influence disease risk.¹⁴ Among known ALS and FTD risk loci, common polymorphisms in the *UNC13A* gene are notably associated with a modestly increased risk of both diseases,^{16,17} with convincing mechanistic evidence of pathogenicity in TDP-43 proteinopathies.¹⁸ In particular, homozygosity for the C allele at rs12608932 has been linked to more aggressive disease and represents a strong candidate modifier of ALS-FTD risk in *C9orf72* HRE carriers.¹⁹ This study aimed to determine the incidence of ALS, dementia and parkinsonism in *C9orf72* HRE carriers in a prospective population-based cohort study, thereby addressing the limitations of previous study designs. The effect of rs12608932 in *UNC13A* on this risk was also examined.

Materials and methods

Participants

The UK Biobank is a prospective, population-based cohort study of 502 625 participants, aged 40–69.²⁰ Approximately 9.2 million invitations were sent to individuals living within 40 km from one of 22 assessment centres and 502 625 individuals (5.5% response rate) attended a baseline assessment, consisting of touchscreen questionnaire, interview, physical assessments and biological sampling (blood, urine and saliva). Whole genome sequencing was performed on blood samples from 490 640 participants using Illumina NovaSeq technology, from which *C9orf72* HRE carriers were identified using ExpansionHunter,²¹ with a cut-off of >100 repeats to prioritize specificity (Supplementary material, 'Methods' section, 'ExpansionHunter'). These individuals are referred to as '*C9orf72* HRE carriers' henceforth. As a sensitivity analysis, analyses were repeated, including all individuals with >30 repeats, the commonly used pathogenicity threshold. Controls were restricted to individuals with <30 repeats throughout. Genotyping was also performed using Affymetrix (now part of ThermoFisher Scientific), from which rs12608932 genotype was extracted (Supplementary material, 'Methods' section).

Participants also consented to linkage of hospital episode statistics (HES) and death registry data. These data are censored on 31 Oct 2022 in England, 31 Aug 2022 in Scotland and 31 May 2022 in Wales. A small proportion of individuals (1297, 0.26%) were knowingly lost to follow-up prior to this, mostly due to withdrawal of consent or migration overseas.

Outcomes

Incident ALS, FTD, any-cause dementia (including FTD) and parkinsonism diagnoses were extracted from linked HES and death registry data (Supplementary material, 'Methods' section, 'Neurodegenerative disease outcomes'). Participants with prevalent ALS, dementia or parkinsonism at recruitment were excluded.

Statistical analysis

C9orf72 HRE carriers and non-carriers were compared using chi-squared tests for categorical variables and Mann-Whitney U-tests for continuous variables, with Bonferroni adjustment for multiple testing.

Age-specific cumulative incidences of neurodegenerative diseases were estimated using both Kaplan–Meier models, censoring at the end of follow-up or death, and models considering other-cause mortality as a competing risk. *C9orf72* HRE carriers were compared by log-rank testing to controls matched by age, sex, Townsend deprivation index (TDI) and ethnicity (white, non-white; Supplementary material, 'Methods' section, 'Matching' and Supplementary Table 1). Analyses were also stratified by sex, BMI, smoking, alcohol and highest education level to examine potential effect modification by these factors and subset by data source of diagnosis. The association between expansion size and disease variables was examined using simple linear regression.

The impact of rs12608932 genotype on neurodegenerative disease incidence was investigated using Kaplan–Meier analyses and Cox regression, in the whole cohort and restricted to *C9orf72* HRE carriers. Cox models utilized age as the timescale and were adjusted for sex, age at recruitment and Townsend deprivation index, with robust standard errors calculated using a sandwich estimator clustered by assessment centre. As a sensitivity analysis, models were additionally adjusted for the first 15 genetic principal components and clustered by genetic batch. The interaction between rs12608932 genotype and *C9orf72* HRE status was investigated by inclusion of an interaction term in the whole cohort.

Results

Baseline characteristics

Among 490 331 individuals with valid genetic data, 701 had at least one *C9orf72* allele with >100 repeats, indicating a prevalence of 143 per 100 000 [95% confidence interval (CI) 132–154], and two were homozygous for the expansion. Median follow-up of *C9orf72* HRE carriers was 13.4 years [interquartile range (IQR) 12.3–14.1]. Three had prevalent neurodegenerative diagnoses and were excluded.

Compared to unmatched controls with <30 repeats, *C9orf72* HRE carriers without prevalent neurodegenerative diagnoses were younger [median 55 (IQR 48–62) versus 58 (50–63) years, $P < 0.001$], more likely of white ethnicity (97.9% versus 94.2%, $P < 0.001$), never-smokers (61.2% versus 54.4%, $P = 0.011$), non-alcohol-drinkers (10.2% versus 8.0%, $P = 0.003$) and of lower body mass index [BMI; 25.3 (23.0–28.7) versus 26.7 (24.1–29.9) kg/m², $P < 0.001$] (Supplementary Table 2). After matching by age, sex, TDI and ethnicity, differences in smoking, alcohol and BMI remained (Supplementary Table 3). Five individuals with missing matching data were excluded, leaving 693 carriers for further analysis.

Age-specific incidence of *C9orf72* HRE-associated diseases

Compared to matched controls, *C9orf72* HRE carriers had a higher cumulative incidence by age 80 of ALS [21% (95% CI 16%–27%) versus

Table 1 Age-specific cumulative incidences of ALS and dementias in C9orf72 HRE carriers compared to matched controls

C9orf72 status	Outcome	Incidence by 65 years % (95% CI)	Incidence by 70 years % (95% CI)	Incidence by 75 years % (95% CI)	Incidence by 80 years % (95% CI)
No C9 exp	ALS	<0.1 (0–<0.1)	<0.1 (0–<0.1)	0.3 (<0.1–0.5)	0.4 (0.1–0.7)
C9 exp	ALS	6.5 (4.3–8.5)	12 (8.4–15)	17 (13–21)	21 (16–27)
No C9 exp	FTD	<0.1 (0–0.1)	<0.1 (0–0.1)	<0.1 (0–0.2)	<0.1 (0–0.2)
C9 exp	FTD	0.9 (0.1–1.8)	3.2 (1.4–5.0)	7.1 (3.9–10)	10 (5.4–15)
No C9 exp	Dementia	0.2 (<0.1–0.3)	0.4 (0.2–0.5)	1.4 (1.0–1.9)	5.4 (4.1–6.7)
C9 exp	Dementia	2.2 (1.0–3.5)	9.8 (6.7–13)	29 (23–35)	59 (49–67)
No C9 exp	ALS or Dementia	0.2 (<0.1–0.3)	0.4 (0.2–0.6)	1.7 (1.2–2.2)	5.8 (4.5–7.0)
C9 exp	ALS or Dementia	7.7 (5.4–9.9)	19 (15–23)	39 (33–45)	66 (57–73)
No C9 exp	ALS and Dementia	<0.1 (0–<0.1)	<0.1 (0–<0.1)	<0.1 (0–<0.1)	<0.1 (0–<0.1)
C9 exp	ALS and Dementia	1.1 (0.2–2.1)	2.4 (0.8–3.9)	3.9 (1.6–6.2)	6.1 (2.3–9.7)
No C9 exp	Parkinsonism	0.1 (<0.1–0.2)	0.2 (<0.1–0.4)	0.9 (0.6–1.3)	2.4 (1.5–3.2)
C9 exp	Parkinsonism	0.3 (0–0.8)	0.6 (0–1.4)	3.0 (0.6–5.4)	4.3 (0.8–7.7)

Numbers rounded to two significant figures. C9 exp = participants carrying at least one C9orf72 HRE allele of over 100 repeats ($n = 693$); No C9 exp = participants homozygous for unexpanded C9orf72 (<30 repeats) matched by age, sex, Townsend deprivation index and ethnicity ($n = 6930$); ALS = amyotrophic lateral sclerosis; CI = confidence interval; FTD = frontotemporal dementia; HRE = hexanucleotide repeat expansion.

0.4% (0.1%–0.7%), $P_{\log\text{-rank}} < 0.001$], FTD [10% (5.4%–15%) versus <0.1% (0%–0.2%), $P_{\log\text{-rank}} < 0.001$] and any-cause dementia [59% (49%–67%) versus 5.4% (4.1%–6.7%), $P_{\log\text{-rank}} < 0.001$] (Table 1 and Fig. 1). Adjusting for competing risk of other-cause mortality before disease onset reduced this to 18% (14%–22%) for ALS, 7.3% (4.7%–11%) for FTD and 44% (37%–51%) for any-cause dementia (Supplementary Table 4 and Supplementary Fig. 1). Fifteen C9orf72 HRE carriers [6.1% (2.3%–9.7%)] had recorded diagnoses of both ALS and dementia by age 80 compared to only one control [<0.1% (0%–<0.1%)]. Of the total dementia incidence linked to C9orf72, 41% was accrued between ages 75–80.

Overall, the combined cumulative incidence of ALS and any-cause dementia by age 80 was 66% (57%–73%) in C9orf72 HRE carriers compared to 5.8% (4.5%–7.0%) in controls, assuming survival to this age, or 58% (50%–64%) in carriers compared to 5.1% (4.1%–6.4%) in controls accounting for other cause mortality.

C9orf72 HRE carriers showed significantly different parkinsonism incidence compared to controls by log rank testing ($P = 0.034$), but no higher cumulative incidence by age 80 [4.3% (0.8%–7.7%) versus 2.4% (1.5%–3.2%); Supplementary Fig. 2].

Among ALS or dementia patients, C9orf72 HRE carriers were younger at first recorded diagnosis compared to age and sex-matched non-carriers [ALS: median age 64.9 (95% CI 63.4–67.8) versus 68.7 (66.2–70.8), $n = 72$ /group, $P_{\log\text{-rank}} = 0.020$; dementia: 72.2 (71.3–73.7) versus 74.7 (74.0–75.6), $n = 120$ /group, $P_{\log\text{-rank}} < 0.001$; Supplementary Fig. 3]. Only eight C9orf72 HRE carriers had a diagnosis of parkinsonism and no significant difference in age of diagnosis was detected (Supplementary Fig. 3).

Stratification by sex, BMI, smoking, alcohol and highest educational qualification demonstrated no interaction between these variables and C9orf72 status in determining ALS and dementia incidence (Supplementary Fig. 4). Analyses using only hospital diagnoses or only death certificate diagnoses are presented in Supplementary Fig. 5 and Supplementary Table 5.

Among 110 individuals identified to carry between 30 and 100 repeats in C9orf72, the cumulative incidence by age 80 of ALS was only marginally higher than in matched controls by Kaplan–Meier modelling [2.8% (0%–6.7%) versus 0.5% (0%–1.5%), $P_{\log\text{-rank}} < 0.001$; Supplementary Table 6 and Supplementary Fig. 6]. The cumulative incidence of any-cause dementia was 16% (1.2%–29%) compared to 5.4% (2.45%–8.2%) in matched controls ($P_{\log\text{-rank}} = 0.004$). No cases of FTD were recorded in the 30–100 repeat group.

Cumulative incidence of ALS and dementia in individuals carrying more than 30 repeats in C9orf72 was 58% (50%–65%) at age 80 compared to 5.8% (4.6%–7.0%) in controls ($P_{\log\text{-rank}} < 0.001$) (Supplementary Table 7 and Supplementary Fig. 7), assuming survival to this age.

Expansion size

C9orf72 expansion size was not significantly associated with age of first ALS or FTD diagnosis or age of death in people with ALS or FTD but increased with age at study entry in C9orf72 HRE carriers ($\beta = 4.5$ repeats/year, 95% CI 1.9–5.2, $P < 0.001$; Fig. 2). Expansion size was similar in those with ALS, parkinsonism, dementia and no neurodegenerative diagnosis.

Effect of UNC13A genotype

Among UK Biobank participants with valid genotyping data, homozygosity for the C allele of rs12608932 in UNC13A ($n = 60\,581/485\,114$) was associated with higher risk of ALS, FTD and the combined outcome of ALS or any-cause dementia, compared with other genotypes [ALS: $P_{\log\text{-rank}} < 0.001$, HR = 1.81 (1.50–2.19); FTD: $P_{\log\text{-rank}} = 0.002$, HR = 1.58 (1.18–2.13); ALS or dementia: $P_{\log\text{-rank}} = 0.007$, HR = 1.08 (1.02–1.15); Fig. 3 and Supplementary Tables 8 and 9]. No significant association was observed for any-cause dementia alone [$P_{\log\text{-rank}} = 0.18$, HR = 1.04 (0.98–1.11)]. A/C heterozygosity did not differ in risk from A/A homozygosity for any outcome (Supplementary Table 8).

In carriers of the C9orf72 HRE, the CC genotype at rs12608932 ($n = 81/690$) was associated with significantly higher risk of all four outcomes: ALS [$P_{\log\text{-rank}} = 0.038$, HR = 2.00 (1.09–3.66)], FTD [$P_{\log\text{-rank}} = 0.063$, HR = 2.03 (1.10–3.73)], dementia [$P_{\log\text{-rank}} = 0.011$, HR = 2.11 (1.24–3.58)] and ALS or dementia [$P_{\log\text{-rank}} = 0.008$, HR = 1.85 (1.20–2.85)], although for FTD only Cox analyses reached statistical significance, likely due to lower power. The effect sizes for ALS and FTD in C9orf72 HRE carriers were comparable to those observed in the full cohort. Accordingly, models exploring an interaction between C9orf72 HRE status and UNC13A genotype on risk indicated a potential interaction between genotypes for risk of any-cause dementia and ALS or any-cause dementia, but not ALS or FTD alone (Supplementary Table 10).

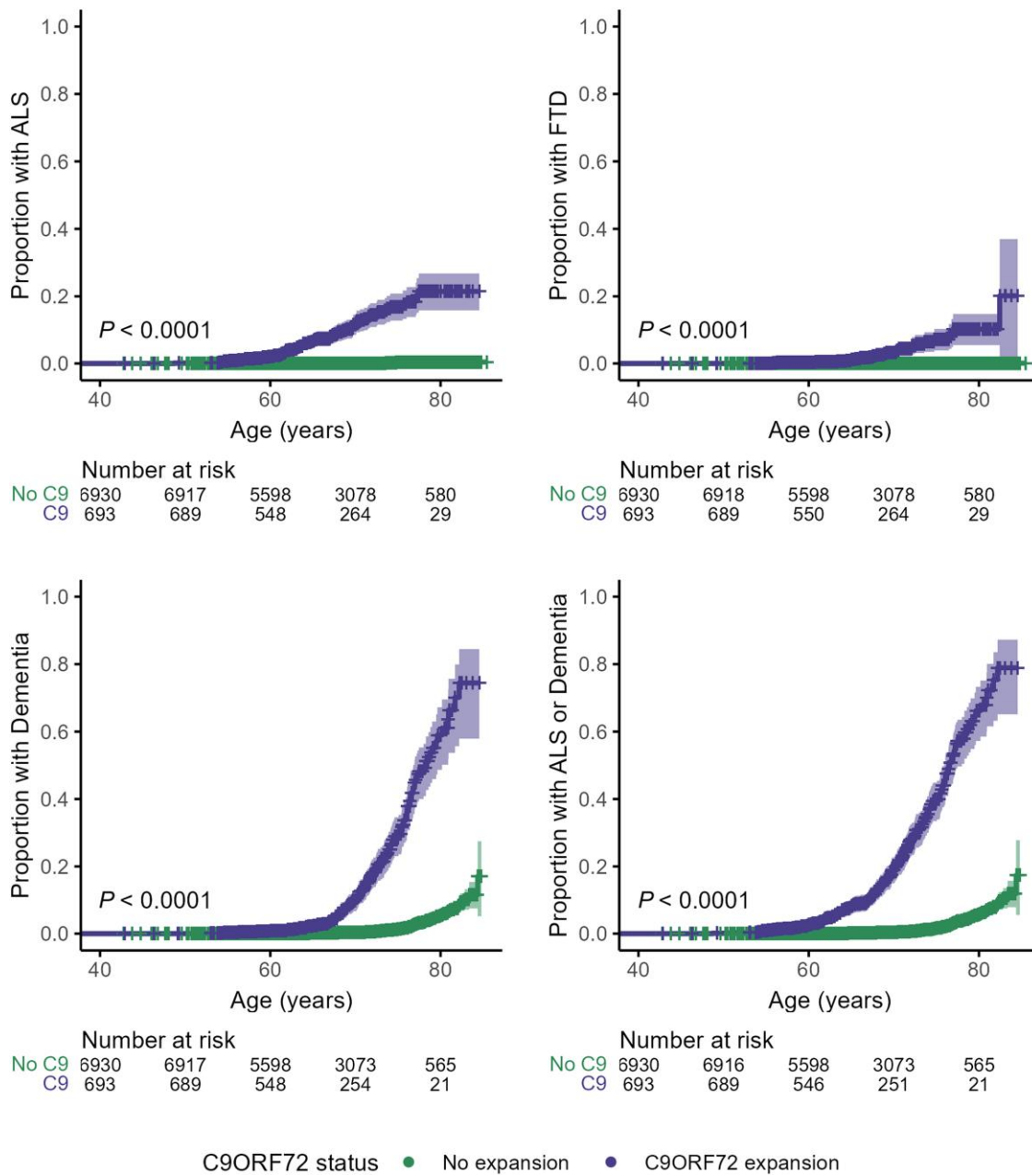


Figure 1 Age-specific cumulative incidences of ALS and dementias in *C9orf72* HRE carriers compared to matched controls. Kaplan–Meier plots of cumulative incidences of ALS, FTD, dementia and a combined outcome of ALS and dementia by age in individuals carrying the *C9orf72* HRE (purple, $n = 693$) compared to age, sex, Townsend deprivation index and ethnicity-matched controls (green, $n = 6930$). Shaded bands indicate 95% confidence intervals. P -values were calculated by the log-rank test. Tables below represent the number of individuals at risk at a given time point (No C9 = No expansion; C9 = *C9orf72* HRE expansion). Dementia refers to any-cause dementia (including FTD). ALS = amyotrophic lateral sclerosis; FTD = frontotemporal dementia; HRE = hexanucleotide repeat expansion.

Discussion

In a population-based cohort of 693 people carrying the *C9orf72* HRE, the cumulative incidence of ALS and dementia by age 80 was 66% in those surviving to this age, or 58% accounting for other-cause mortality. No robust association with parkinsonism incidence was detected. ALS incidence peaked between ages 65–75 at around 1% per year, whereas dementia diagnoses were greatest after age 70 at around 5% per year. Additionally, the prevalence of *C9orf72* HRE in this study was 143 per 100 000 (1 in 699), consistent with existing studies suggesting 1 in 517–699.^{1,9} Risk of ALS, FTD

and any-cause dementia in *C9orf72* HRE carriers was approximately doubled by the homozygosity for the C allele of rs12608932 in *UNC13A*.

These estimates improve upon previous data by utilizing a prospective population-based cohort design, placing the penetrance of neurodegenerative disease in *C9orf72* HRE carriers lower than estimates from disease-enriched cohorts (>95% by age 80),^{1,13} but higher than indirect estimates (24%–33%).^{14,15} This further emphasizes the incomplete penetrance of *C9orf72* HRE-associated disease, raising interesting aetiological questions regarding the genetic and environmental factors determining the individualized risk profile of

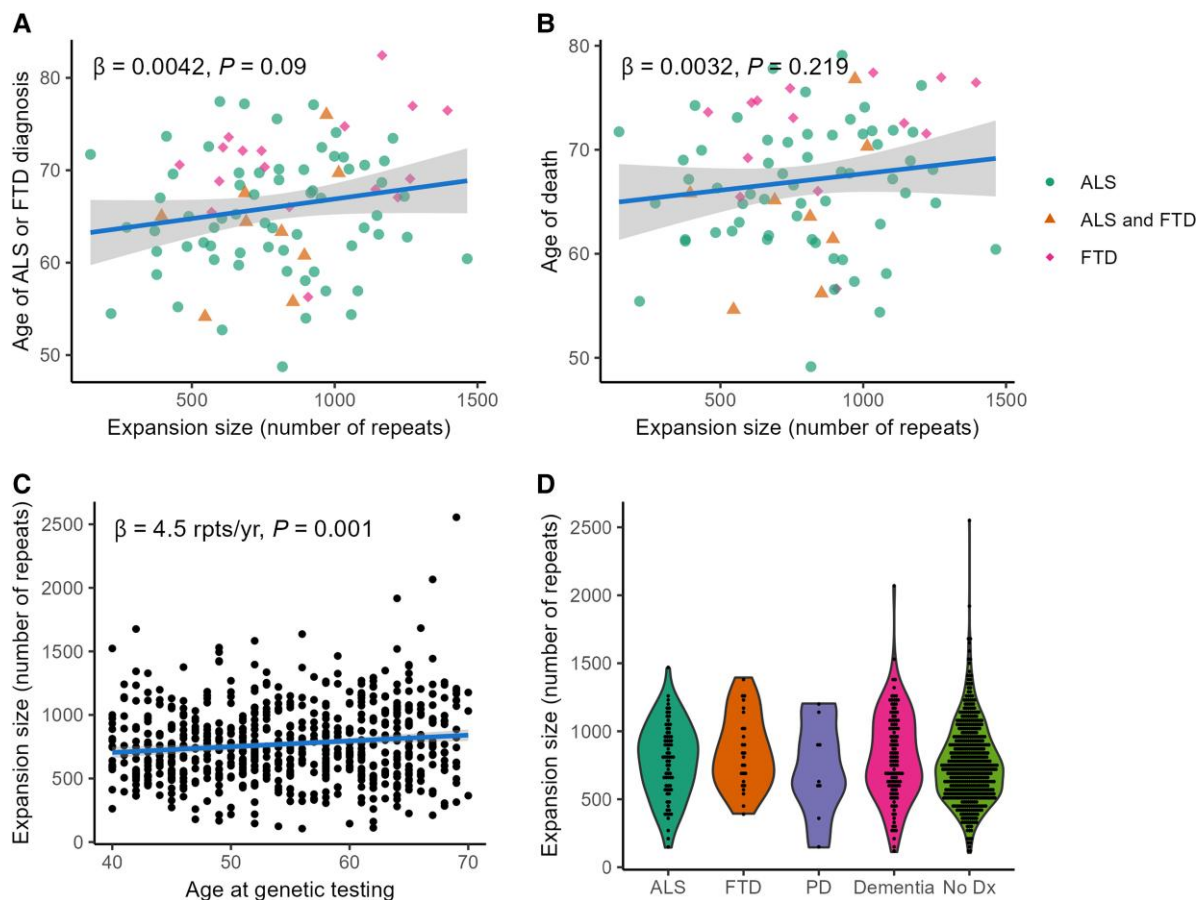


Figure 2 *C9orf72* expansion size and phenotype. (A) Association between *C9orf72* expansion size and age at first recorded ALS or FTD diagnosis ($n = 91$). Points are coloured by diagnosis. (B) Association between *C9orf72* expansion size and age of death in individuals with a recorded ALS or FTD diagnosis who have died. Points are coloured by diagnosis ($n = 81$). (C) Association between age at genetic testing and expansion size in individuals carrying a *C9orf72* HRE ($n = 701$) (D) Expansion size and diagnosis ($n = 74$ for ALS, 26 for FTD, 8 for PD, 121 for any-cause dementia, 521 for No Dx). ALS = amyotrophic lateral sclerosis; FTD = frontotemporal dementia; HRE = hexanucleotide repeat expansion; No Dx = no neurodegenerative diagnosis recorded; PD = parkinsonism. Labels represent fitted linear regression models and associated P -values.

each *C9orf72* HRE carrier. These factors must be carefully considered when counselling asymptomatic individuals considering predictive genetic testing.

In particular, common variants in *UNC13A*—most notably the C allele of rs12608932—have been associated with higher ALS and FTD risk in genome-wide association studies^{16,17} and shorter survival in ALS cohorts.¹⁹ Mechanistic studies have persuasively linked TDP-43 proteinopathies to *UNC13A* loss of function through cryptic exon inclusion, potentiated by the rs12608932 risk C allele.¹⁸ An antisense oligonucleotide therapy targeting *UNC13A* as a means to slow disease progression in symptomatic ALS is now also in clinical trials.²² This study adds the crucial link between *UNC13A* and *C9orf72* HRE-associated disease in a population-based cohort. These genes appear to exert a proportional effect on risks of ALS and FTD, with *UNC13A* rs12608932 C-allele homozygosity doubling risk in both *C9orf72* HRE carriers and non-carriers. In contrast, for any-cause dementia, there is evidence of a statistical interaction between genotypes, which likely reflects the enrichment of TDP-43-related dementias in those with either *C9orf72* HRE or *UNC13A* risk genotypes compared with the general population, rather than a true synergistic effect.

UNC13A genotype may partially explain inter-familial differences in penetrance and inter-individual differences in disease

risk,¹⁴ inform the design of future clinical trials and disease models,²³ and suggest a targetable pathway for future preventative therapies in *C9orf72* HRE carriers. Further work is required to explore its role in other genetic forms of ALS and FTD.

The finding of a high incidence of recorded non-FTD dementia in *C9orf72* HRE carriers adds to the debate surrounding the spectrum of *C9orf72* HRE-associated disease.²⁴ These cases may represent misdiagnosis or mis-recording of FTD.²⁵ Alternatively, they may suggest an association between *C9orf72* HRE and amnesic dementias, such as limbic-predominant age-related TDP-43 encephalopathy (LATE) or Alzheimer's disease, hitherto under-detected due to the skewed selection of individuals for genetic testing. Against the latter possibility is the lack of genetic evidence linking *C9orf72* to dementia in genome-wide association studies.²⁶ Adequately resolving these uncertainties requires further study of the neuropsychological and neuropathological phenotypes of these individuals, including analysis of post-mortem tissue.

To reduce the risk of false positive detection given the low prior probability of carrying a *C9orf72* HRE in a population-based cohort, a conservative threshold of 100 repeats was applied for the primary analysis. Analyses of individuals with 30–100 repeats demonstrated a lower incidence of ALS and dementia compared with people with

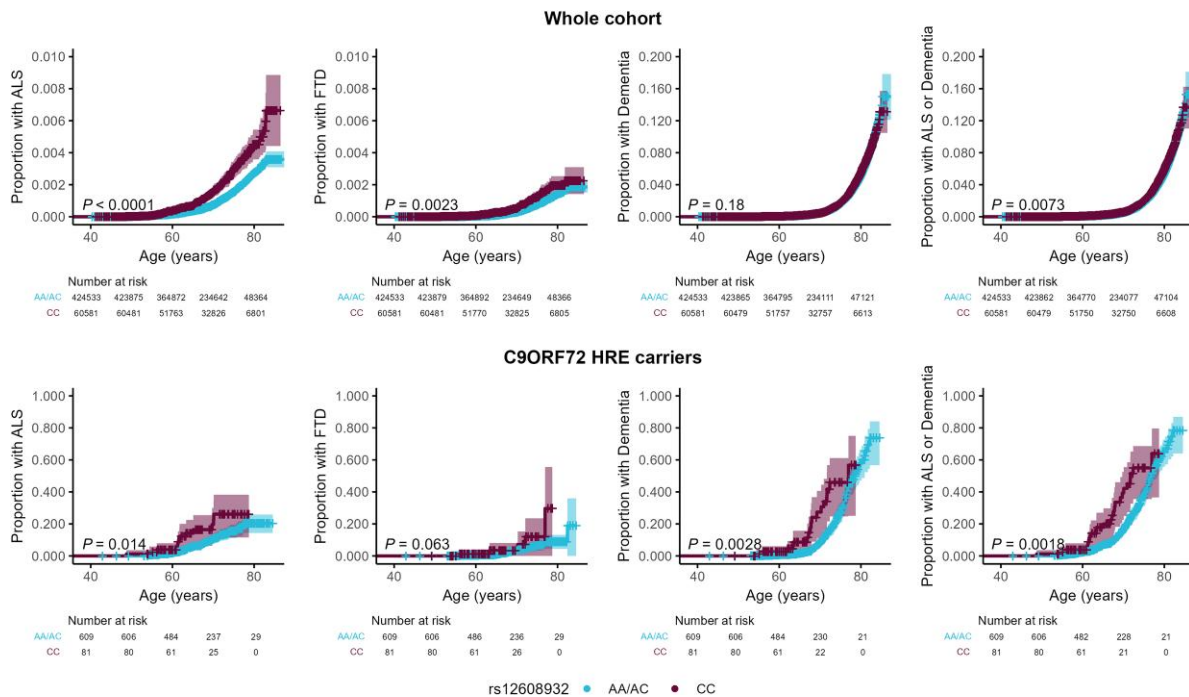


Figure 3 *UNC13A* genotype and neurodegenerative disease incidence. Kaplan–Meier plots of cumulative incidences of ALS, FTD, dementia and a combined outcome of ALS and dementia by age in individuals stratified by rs12608932 genotype, in the whole UK biobank cohort (top, $n = 485\,114$) and in individuals carrying the *C9orf72* HRE (bottom, $n = 690$). Shaded bands indicate 95% confidence intervals. P-values were calculated by the log-rank test. Tables below represent the number of individuals at risk at a given time point. ALS = amyotrophic lateral sclerosis; FTD = frontotemporal dementia; HRE = hexanucleotide repeat expansion.

over 100 repeats, which could indicate lower penetrance at shorter repeat lengths. However, given that ExpansionHunter has a reported specificity of 99.9%,²¹ a significant proportion of false positives must be expected, preventing conclusions in this population-based analysis. It is also notable that no association between expansion length and risk of ALS or dementia is observed above 100 repeats.

This study has several limitations. Despite a median follow-up of over 13 years, data are available for fewer individuals at older ages, leading to greater uncertainty in incidence estimates at these ages. Moreover, the UK Biobank cohort is subject to healthy volunteer bias,²⁷ the recruitment age (40–69) effectively excludes younger onset disease, and hospital and death registry diagnoses likely significantly post-date symptom onset. The methodology for obtaining diagnostic data in UK Biobank is imperfect and both false positive and false negative cases are likely. Any-cause dementia diagnoses from hospital and mortality data have an estimated positive predictive value of 84.5% (95% CI 74.5–88.8), leading to potential false positive diagnoses and an overestimate of cumulative incidence.²⁵ Adjusting for this would reduce the estimated cumulative incidence of any-cause dementia by age 80 to 50% as opposed to the 59% presented. False negatives are difficult to estimate, but the reported dementia incidence of 5.4% by age 80 is far lower than population estimates of around 11%,²⁸ though similar to other studies in UK Biobank.²⁹ In contrast, ALS diagnoses are generally believed accurate.³⁰

Conclusions

Analysis of *C9orf72* HRE-related disease in a population-based cohort is consistent with incomplete penetrance, with a large

proportion manifesting as late-onset dementia. Risk of ALS and dementia in *C9orf72* HRE carriers is influenced by common variants in *UNC13A*. This has implications for genetic counselling and modelling of expected phenoconversion in future preventative trials.

Data availability

The data that support the findings of this study are available from UK Biobank, subject to a registration and application process (see <https://www.ukbiobank.ac.uk>).

Acknowledgements

This research has been conducted using the UK Biobank Resource under application number 57629 and was carried out at the National Institute for Health and Care Research (NIHR) Oxford Biomedical Research Centre (BRC).

Funding

A.G.T. received funding from the Medical Research Council (MR/T006927/1) and Motor Neurone Disease Association (Thompson/Jan20/952-795). C.C. received funding from the Motor Neurone Disease Association (Thompson/Apr23/896-791). J.S. received funding from the My Name's 5 Dottie Foundation.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at [Brain](#) online.

References

- Majounie E, Renton AE, Mok K, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: A cross-sectional study. *Lancet Neurol.* 2012;11:323-330.
- Byrne S, Elamin M, Bede P, et al. Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: A population-based cohort study. *Lancet Neurol.* 2012;11:232-240.
- Mahoney CJ, Beck J, Rohrer JD, et al. Frontotemporal dementia with the C9ORF72 hexanucleotide repeat expansion: Clinical, neuroanatomical and neuropathological features. *Brain.* 2012; 135:736-750.
- Chiò A, Borghero G, Restagno G, et al. Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. *Brain.* 2012;135:784-793.
- Umoh ME, Fournier C, Li Y, et al. Comparative analysis of C9orf72 and sporadic disease in an ALS clinic population. *Neurology.* 2016;87:1024-1030.
- van der Ende EL, Jackson JL, White A, Seelaar H, van Blitterswijk M, Swieten JC. Unravelling the clinical spectrum and the role of repeat length in C9ORF72 repeat expansions. *J Neurol Neurosurg Psychiatry.* 2021;92:502-509.
- Lesage S, Le Ber I, Condroyer C, et al. C9orf72 repeat expansions are a rare genetic cause of parkinsonism. *Brain.* 2013;136(Pt 2): 385-391.
- Xi Z, Zinman L, Grinberg Y, et al. Investigation of C9orf72 in 4 neurodegenerative disorders. *Arch Neurol.* 2012;69:1583-1590.
- Beck J, Poulter M, Hensman D, et al. Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *Am J Hum Genet.* 2013;92:345-353.
- Murray ME, Bieniek KF, Banks Greenberg M, et al. Progressive amnesic dementia, hippocampal sclerosis, and mutation in C9ORF72. *Acta Neuropathol.* 2013;126:545-554.
- Dharmadasa T, Scaber J, Edmond E, et al. Genetic testing in motor neurone disease. *Pract Neurol.* 2022;22:107-116.
- Benatar M, Heiman-Patterson T, Cooper-Knock J, et al. Guidance for clinical management of pathogenic variant carriers at elevated genetic risk for ALS/FTD. *J Neurol Neurosurg Psychiatry.* 2025;96:e334339.
- Murphy NA, Arthur KC, Tienari PJ, Houlden H, Chiò A, Traynor BJ. Age-related penetrance of the C9orf72 repeat expansion. *Sci Rep.* 2017;7:2116.
- Van Wijk IF, Van Eijk RPA, Van Boxmeer L, et al. Assessment of risk of ALS conferred by the GGGGCC hexanucleotide repeat expansion in C9orf72 among first-degree relatives of patients with ALS carrying the repeat expansion. *Amyotroph Lateral Scler Frontotemporal Degener.* 2024;25:188-196.
- Douglas AGL, Baralle D. Reduced penetrance of gene variants causing amyotrophic lateral sclerosis. *J Med Genet.* 2024;61: 294-297.
- Van Es MA, Veldink JH, Saris CGJ, et al. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet.* 2009;41:1083-1087.
- Diekstra FP, Van Deerlin VM, van Swieten JC, et al. C9orf72 and UNC13A are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: A genome-wide meta-analysis. *Ann Neurol.* 2014;76:120-133.
- Brown A-L, Wilkins OG, Keuss MJ, et al. TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. *Nature.* 2022; 603:131-137.
- Chiò A, Mora G, Restagno G, et al. UNC13A influences survival in Italian amyotrophic lateral sclerosis patients: A population-based study. *Neurobiol Aging.* 2013;34:357.e1-357.e5.
- Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12:e1001779.
- Dolzhenko E, van Vugt JJFA, Shaw RJ, et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. *Genome Res.* 2017;27:1895-1903.
- QurAlis Corporation. QurAlis Corporation. Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of QRL-201 in Participants With Amyotrophic Lateral Sclerosis (ALS). ClinicalTrials.gov. 2022. Accessed 2 May 2025. <https://clinicaltrials.gov/study/NCT05633459>
- Staffaroni AM, Quintana M, Wendelberger B, et al. Temporal order of clinical and biomarker changes in familial frontotemporal dementia. *Nat Med.* 2022;28:2194-2206.
- Cooper-Knock J, Shaw PJ, Kirby J. The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. *Acta Neuropathol.* 2014;127:333-345.
- Wilkinson T, Schnier C, Bush K, et al. Identifying dementia outcomes in UK Biobank: A validation study of primary care, hospital admissions and mortality data. *Eur J Epidemiol.* 2019;34: 557-565.
- Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet.* 2022;54:412-436.
- Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol.* 2017;186:1026.
- Prince M, Knapp M, Guerchet M, et al. *Dementia UK: Update.* 2nd ed. Alzheimer's Society; 2014.
- Gong J, Harris K, Peters SAE, Woodward M. Serum lipid traits and the risk of dementia: A cohort study of 254,575 women and 214,891 men in the UK Biobank. *EClinicalMedicine.* 2022;54: 101695.
- Horrocks S, Wilkinson T, Schnier C, et al. Accuracy of routinely-collected healthcare data for identifying motor neurone disease cases: A systematic review. *PLoS One.* 2017;12:e0172639.