Genetic correlation between amyotrophic lateral sclerosis and schizophrenia


Published in:
Nature Communications

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2017 the authors.
This is an open access article published under a Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date: 13. Dec. 2018
Genetic correlation between amyotrophic lateral sclerosis and schizophrenia

Russell L. McLaughlin\textsuperscript{1,2,*}, Dick Schijven\textsuperscript{3,4,*}, Wouter van Rheenen\textsuperscript{3}, Kristel R. van Eijk\textsuperscript{3}, Margaret O’Brien\textsuperscript{1}, Project MinE GWAS Consortium\textsuperscript{†}, Schizophrenia Working Group of the Psychiatric Genomics Consortium\textsuperscript{‡}, René S. Kahn\textsuperscript{4}, Roel A. Ophoff\textsuperscript{4,5,6}, An Goris\textsuperscript{7}, Daniel G. Bradley\textsuperscript{2}, Ammar Al-Chalabi\textsuperscript{8}, Leonard H. van den Berg\textsuperscript{3}, Jurjen J. Luykx\textsuperscript{3,4,9,**}, Orla Hardiman\textsuperscript{2,**} & Jan H. Veldink\textsuperscript{3,**}

We have previously shown higher-than-expected rates of schizophrenia in relatives of patients with amyotrophic lateral sclerosis (ALS), suggesting an aetiological relationship between the diseases. Here, we investigate the genetic relationship between ALS and schizophrenia using genome-wide association study data from over 100,000 unique individuals. Using linkage disequilibrium score regression, we estimate the genetic correlation between ALS and schizophrenia to be 14.3% (7.05–21.6; $P = 1 \times 10^{-4}$) with schizophrenia polygenic risk scores explaining up to 0.12% of the variance in ALS ($P = 8.4 \times 10^{-7}$).

A modest increase in comorbidity of ALS and schizophrenia is expected given these findings (odds ratio 1.08–1.26) but this would require very large studies to observe epidemiologically.

We identify five potential novel ALS-associated loci using conditional false discovery rate analysis. It is likely that shared neurobiological mechanisms between these two disorders will engender novel hypotheses in future preclinical and clinical studies.

\textsuperscript{1} Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin D02 DK07, Republic of Ireland. \textsuperscript{2} Smurfit Institute of Genetics, Trinity College Dublin, Dublin D02 DK07, Republic of Ireland. \textsuperscript{3} Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht 3584 CX, The Netherlands. \textsuperscript{4} Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht 3584 CX, The Netherlands. \textsuperscript{5} Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA. \textsuperscript{6} Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, California 90095, USA. \textsuperscript{7} Department of Neurosciences, Experimental Neurology and Leuven Research Institute for Neuroscience and Disease (LIND), KU Leuven—University of Leuven, Leuven B-3000, Belgium. \textsuperscript{8} Department of Basic and Clinical Neuroscience, Maurice Wohl Clinical Neuroscience Institute, King’s College London, London WC2R 2LS, UK. \textsuperscript{9} Department of Psychiatry, Hospital Network Antwerp (ZNA) Stuivenberg and Sint Erasmus, Antwerp 2020, Belgium. * These authors contributed equally to this work. ** These authors jointly supervised this work. Correspondence and requests for materials should be addressed to R.L.M. (email: mclaughr@tcd.ie).

\textsuperscript{†} A full list of Project MinE GWAS Consortium members appears at the end of the paper. \textsuperscript{‡} A full list of Schizophrenia Working Group of the Psychiatric Genomics Consortium members appears at the end of the paper.
Amyotrophic lateral sclerosis (ALS) is a late-onset neurodegenerative condition characterized by progressive loss of upper and lower motor neurons, leading to death from respiratory failure in 70% of patients within 3 years of symptom onset. Although ALS is often described as a primarily motor-system disease, extramotor involvement occurs in up to 50% of cases, with prominent executive and behavioural impairment, and behavioural variant frontotemporal dementia (FTD) in up to 14% of cases. A neuropsychiatric prodrome has been described in some people with ALS–FTD, and higher rates of schizophrenia and suicide have been reported in first and second degree relatives of those with ALS, particularly in kindreds associated with the C9orf72 hexanucleotide repeat expansion. These clinical and epidemiological observations suggest that ALS and schizophrenia may share heritability.

ALS and schizophrenia both have high heritability estimates (0.65 and 0.64, respectively), however the underlying genetic architectures of these heritable components appear to differ. Analysis of large genome-wide association study (GWAS) datasets has implicated over 100 independent risk loci for schizophrenia and estimated that a substantial proportion (23%) of the variance in underlying liability for schizophrenia is due to additive polygenic risk (many risk-increasing alleles of low individual effect combining to cause disease) conferred by common genetic variants. This proportion, the single nucleotide polymorphism (SNP)-based heritability, is lower in ALS (8.2%), in which fewer than ten risk loci have been identified by GWAS. Nevertheless, both diseases have polygenic components, but the extent to which they overlap has not been investigated.

Recently, methods to investigate overlap between polygenic traits using GWAS data have been developed. These methods assess either pleiotropy (identical genetic variants influencing both traits) or genetic correlation (identical alleles influencing both traits). Genetic correlation is related to heritability; for both measures, binary traits such as ALS and schizophrenia are typically modelled as extremes of an underlying continuous scale of liability to develop the trait. If two binary traits are genetically correlated, their liabilities covary, and this covariance is determined by both traits having identical risk alleles at overlapping risk loci. Studies of pleiotropy and genetic correlation have provided insights into the overlapping genetics of numerous traits and disorders, although none to date has implicated shared polygenic risk between neurodegenerative and neuropsychiatric disease. Here, we apply several techniques to identify and dissect the polygenic overlap between ALS and schizophrenia. We provide evidence for genetic correlation between the two disorders which is unlikely to be driven by diagnostic misclassification and we demonstrate a lack of polygenic overlap between ALS and other neuropsychiatric and neurological conditions, which could be due to limited power given the smaller cohort sizes for these studies.

Results

Genetic correlation between ALS and schizophrenia. To investigate the polygenic overlap between ALS and schizophrenia, we used individual-level and summary data from GWAS for ALS (36,052 individuals) and schizophrenia (79,845 individuals). At least 5,582 control individuals were common to both datasets, but for some cohorts included in the schizophrenia dataset this could not be ascertained so this number is likely to be higher. For ALS, we used summary data from both mixed linear model association testing and meta-analysis of cohort-level logistic regression. We first used linkage disequilibrium (LD) score regression with ALS and schizophrenia summary statistics; this technique models, for polygenic traits, a linear relationship between a SNP’s LD score (the amount of genetic variation that it captures) and its GWAS test statistic. This distinguishes confounding from polygenicity in GWAS inflation and the regression coefficient can be used to estimate the SNP-based heritability ($h^2_g$) for single traits. In the bivariate case, the regression coefficient estimates genetic covariance ($\rho_g$) for pairs of traits, from which genetic correlation ($r_g$) is estimated; these estimates are unaffected by sample overlap between traits. Using constrained intercept LD score regression with mixed linear model ALS summary statistics, we estimated the liability-scale SNP-based heritability of ALS to be 8.2% (95% confidence interval = 7.2–9.1; mean $\chi^2 = 1.13$; all ranges reported below indicate 95% confidence intervals), replicating previous estimates based on alternative methods.

Results based on ALS meta-analysis summary statistics and free-intercept LD score regression with mixed linear model summary statistics were lower (Supplementary Table 1), resulting in higher genetic correlation estimates (Supplementary Table 2); for this reason, we conservatively use constrained intercept genetic correlation estimates for ALS mixed linear model summary statistics throughout the remainder of this paper. Heritability estimates for permuted ALS data were null (Supplementary Table 1).

LD score regression estimated the genetic correlation between ALS and schizophrenia to be 14.3% (7.05–21.6; $P = 1 \times 10^{-4}$). Results were similar for a smaller schizophrenia cohort of European ancestry (21,856 individuals), indicating that the inclusion of individuals of Asian ancestry in the schizophrenia cohort did not bias this result (Supplementary Fig. 1). In addition to schizophrenia, we estimated genetic correlation with ALS using GWAS summary statistics for bipolar disorder, major depressive disorder, attention deficit-hyperactivity disorder, autism spectrum disorder, Alzheimer’s disease (Supplementary Note 1), multiple sclerosis and adult height, finding no significant genetic correlation between ALS and any secondary trait other than schizophrenia (Fig. 1; Supplementary Table 2).

Polygenic risk score analysis. We supported the positive genetic correlation between ALS and schizophrenia by analysis of genome-wide polygenic risk scores (PRS). PRS are an approach to summarizing genetic variation that can be used to predict disease liability when averaged across a large number of polygenic risk alleles. This method is particularly useful for polygenic traits, a linear relationship between a SNP’s LD score (the amount of genetic variation that it captures) and its GWAS test statistic. This distinguishes confounding from polygenicity in GWAS inflation and the regression coefficient can be used to estimate the SNP-based heritability ($h^2_g$) for single traits. In the bivariate case, the regression coefficient estimates genetic covariance ($\rho_g$) for pairs of traits, from which genetic correlation ($r_g$) is estimated; these estimates are unaffected by sample overlap between traits. Using constrained intercept LD score regression with mixed linear model ALS summary statistics, we estimated the liability-scale SNP-based heritability of ALS to be 8.2% (95% confidence interval = 7.2–9.1; mean $\chi^2 = 1.13$; all ranges reported below indicate 95% confidence intervals), replicating previous estimates based on alternative methods.

Results based on ALS meta-analysis summary statistics and free-intercept LD score regression with mixed linear model summary statistics were lower (Supplementary Table 1), resulting in higher genetic correlation estimates (Supplementary Table 2); for this reason, we conservatively use constrained intercept genetic correlation estimates for ALS mixed linear model summary statistics throughout the remainder of this paper. Heritability estimates for permuted ALS data were null (Supplementary Table 1).

LD score regression estimated the genetic correlation between ALS and schizophrenia to be 14.3% (7.05–21.6; $P = 1 \times 10^{-4}$). Results were similar for a smaller schizophrenia cohort of European ancestry (21,856 individuals), indicating that the inclusion of individuals of Asian ancestry in the schizophrenia cohort did not bias this result (Supplementary Fig. 1). In addition to schizophrenia, we estimated genetic correlation with ALS using GWAS summary statistics for bipolar disorder, major depressive disorder, attention deficit-hyperactivity disorder, autism spectrum disorder, Alzheimer’s disease (Supplementary Note 1), multiple sclerosis and adult height, finding no significant genetic correlation between ALS and any secondary trait other than schizophrenia (Fig. 1; Supplementary Table 2).

![Figure 1 | Genetic correlation between ALS and eight secondary traits.](image)

Error bars indicating 95% confidence intervals and $P$-values were calculated by the LD score regression software using a block jackknife procedure. Secondary traits are: AD, Alzheimer’s disease; ADHD, attention deficit-hyperactivity disorder; ASD, autism spectrum disorder; BPD, bipolar disorder; MDD, major depressive disorder; MS, multiple sclerosis; SCZ, schizophrenia.
polygenic risk for schizophrenia in the ALS cohort. Polygenic risk scores (PRS) are per-individual scores based on the sum of alleles associated with one phenotype, weighted by their effect size, measured in an independent target sample of the same or a different phenotype\(^{10}\). PRS calculated on schizophrenia GWAS summary statistics for twelve \(P\)-value thresholds (\(P_T\)) explained up to 0.12% \((P_T = 0.2, P = 8.4 \times 10^{-7}\)) of the phenotypic variance in a subset of the individual-level ALS genotype data that had all individuals removed that were known or suspected to be present in the schizophrenia cohort (Fig. 2; Supplementary Table 5). ALS cases had on average higher PRS for schizophrenia compared to healthy controls and harbouring a high schizophrenia PRS for \(P_T = 0.2\) significantly increased the odds of being an ALS patient in our cohort (Fig. 3; Supplementary Table 7), which we consider to be too high to be true genetic correlation of 0%, we estimated the required rate of misdiagnosis of ALS as schizophrenia to be 4.86% (2.47–7.13) to obtain the genetic correlation estimate of 14.3% (7.05–21.6; Supplementary Table 6), which we consider to be too high to be true genetic correlation of 0%.

Modelling misdiagnosis and comorbidity. Using BUMHBOX\(^{21}\), a tool that distinguishes true genetic relationships between diseases (pleiotropy) from spurious relationships resulting from heterogeneous mixing of disease cohorts, we determined that misdiagnosed cases in the schizophrenia cohort (for example, young-onset FTD–ALS) did not drive the genetic correlation estimate between ALS and schizophrenia \((P = 0.94)\). Assuming a true genetic correlation of 0%, we estimated the required rate of misdiagnosis of ALS as schizophrenia to be 4.86% (2.47–7.13) to obtain the genetic correlation estimate of 14.3% (7.05–21.6; Supplementary Table 7), which we consider to be too high to be likely. However, if ALS and schizophrenia are genetically correlated, more comorbidity would be expected than if the genetic correlation was 0%. Modelling our observed genetic correlation of 14.3% (7.05–21.6), we estimated the odds ratio for having above-threshold liability for ALS given above-threshold liability for schizophrenia to be 1.17 (1.08–1.26), and the same for schizophrenia given ALS (Supplementary Fig. 4). From a clinical perspective, to achieve 80% power to detect a significant \((\alpha = 0.05)\) excess of schizophrenia in the ALS cohort as a result of this genetic correlation, the required population-based incident cohort size is 16,448 ALS patients (7,310–66,670).
Pleiotropic risk loci. We leveraged the genetic correlation between ALS and schizophrenia to discover novel ALS-associated genomic loci by conditional false discovery rate (cFDR) analysis.\(^4\)\(^,\)\(^22\) (Fig. 4; Supplementary Table 8). Five loci already known to be involved in ALS were identified (corresponding to \(\text{MOBP}, \text{C9orf72}, \text{TBK1}, \text{SARM1}\) and \(\text{UNC13A}\)) along with five potential novel loci at \(\text{cFDR}<0.01\) (\(\text{CNTN6}, \text{TNPI1}, \text{PPP2R2D}, \text{NCKAP5L}\) and \(\text{ZNF295-AS1}\)). No gene set was significantly enriched (after Bonferroni correction) in genome-wide cFDR values when analysed using MAGENTA.

Discussion

There is evolving clinical, epidemiological and biological evidence for an association between ALS and psychotic illness, particularly schizophrenia. Genetic evidence of overlap to date has been based primarily on individual genes showing Mendelian inheritance, in particular the \(\text{C9orf72}\) hexanucleotide repeat expansion, which is associated with ALS and FTD, and with psychosis in relatives of ALS patients.\(^2\) In this study, we have replicated SNP-based heritability estimates for ALS and schizophrenia using GWAS summary statistics, and have for the first time demonstrated significant overlap between the polygenic components of both diseases, estimating the genetic correlation to be 14.3%. We have carefully controlled for confounding bias, including population stratification and shared control samples, and have shown through analysis of polygenic risk scores that the overlapping polygenic risk applies to SNPs that are modestly associated with both diseases. Given that our genetic correlation estimate relates to the polygenic components of ALS (\(\text{h}^2_\text{P} = 8.2\%) and schizophrenia (\(\text{h}^2_\text{SCZ} = 23\%)\) and these estimates do not represent all heritability for both diseases, the accuracy of using schizophrenia-heritability for both diseases, the accuracy of using schizophrenia-polygenic risk applies to SNPs that are modestly associated through analysis of polygenic risk scores that the overlapping polygenic risk applies to SNPs that are modestly associated with both diseases. Schizophrenia is a polygenic neurodevelopmental disorder characterized by a combination of positive symptoms (hallucinations and delusions), negative symptoms (diminished motivation, blunted affect, reduction in spontaneous speech and poor social functioning) and impairment over a broad range of cognitive abilities.\(^23\) ALS is a late onset complex genetic disease characterized by a predominantly motor phenotype with recently recognized extra-motor features in 50% of patients, including cognitive impairment.\(^1\) It has been suggested that the functional effects of risk genes in schizophrenia converge by modulating synaptic plasticity, and influencing the development and stabilization of cortical microcircuitry.\(^2\) In this context, our identification of \(\text{CNTN6}\) (contactin 6, also known as \(\text{NB-3}\), a neural adhesion protein important in axon development)\(^26\) as a novel pleiotropy-informed ALS-associated locus supports neural network dysregulation as a potential convergent mechanism of disease in ALS and schizophrenia.

No significantly enriched biological pathway or ontological term was identified within genome-wide cFDR values using MAGENTA. Low inflation in ALS GWAS statistics, coupled with a rare variant genetic architecture,\(^7\) render enrichment-based biological pathway analyses with current sample sizes challenging. Nevertheless, nine further loci were associated with ALS risk at cFDR <0.01. Of these, \(\text{MOBP}, \text{C9orf72}, \text{TBK1}, \text{SARM1}\) and \(\text{UNC13A}\) have been described previously in ALS and were associated by cFDR analysis in this study owing to their strong association with ALS through GWAS.\(^7\) The remaining four loci (\(\text{TNPI1}, \text{PPP2R2D}, \text{NCKAP5L}\) and \(\text{ZNF295-AS1}\)) are novel associations and may represent pleiotropic disease loci. \(\text{TNPI1}\) encodes TNFAIP3 interacting protein 1 and is involved in autophagy and tissue homeostasis.\(^27\) The protein product of \(\text{PPP2R2D}\) is a regulatory subunit of protein phosphatase 2 and has a role in PI3K-Akt signalling and mitosis.\(^28\) \(\text{NCKAP5L}\) is a homologue of \(\text{NCKAP5}\), encoding NAP5, a proline-rich protein that has previously been implicated in schizophrenia, bipolar disorder and autism.\(^29,30\) \(\text{ZNF295-AS1}\) is a noncoding RNA.\(^31\) Further investigation into the biological roles of these genes may yield novel insight into the pathophysiology of certain subtypes of ALS and schizophrenia, and as whole-genome and exome datasets become available in the future for appropriately large ALS case–control cohorts, testing for burden of rare genetic variation across these genes will be particularly instructive, especially given the role that rare variants appear to play in the pathophysiology of ALS.\(^7\)

Figure 4 | Pleiotropy-informed ALS risk loci determined by analysis of cFDR in ALS GWAS P-values given schizophrenia GWAS P-values \((\text{cFDR}_{\text{ALS:SCZ}})\). Each point denotes a SNP; its x axis position corresponds to its chromosomal location and its height indicates the extent of association with ALS by cFDR analysis. The solid line indicates the threshold cFDR = 0.01. Any gene whose role in ALS is already established is in bold. A complete list of all loci at cFDR ≤0.05 is provided in Supplementary Table 8.
Our data suggest that other neuropsychiatric conditions (bipolar disorder, autism and major depression) do not share polygenic risk with ALS. This finding contrasts with our recent observations from family aggregation studies and may be unexpected given the extensive genetic correlation between neuropsychiatric conditions\(^5\). This could relate to statistical power conferred by secondary phenotype cohort sizes, and future studies with larger sample sizes will shed further light on the relationship between ALS and neuropsychiatric disease. It is also possible that the current study underestimates genetic correlations due to the substantial role that rare variants play in the genetic architecture of ALS\(^\) and future fine-grained studies examining heritability and genetic correlation in low-minor allele frequency and low-LD regions may identify a broader relationship between ALS and neuropsychiatric diseases.

A potential criticism of this study is that the polygenic overlap between ALS and schizophrenia could be driven by misdiagnosis, particularly in cases of ALS–FTD, which can present in later life as a psychotic illness and could be misdiagnosed as schizophrenia. This is unlikely, as strict diagnostic criteria are required for inclusion in the schizophrenia GWAS dataset\(^\). Furthermore, since core schizophrenia symptoms are usually diagnosed during late adolescence, a misdiagnosis of FTD-onset ALS–FTD as schizophrenia is unlikely. In this study, we found no evidence for misdiagnosis of ALS as schizophrenia (BUHMBOX \(P = 0.94\)) and we estimated that a misdiagnosis of 4.86% of ALS cases would be required to spuriously observe a genetic correlation of 14.3%, which is not likely to occur in clinical practice. We are therefore confident that this genetic correlation estimate reflects a genuine polygenic overlap between the two diseases and is not a feature of cohort ascertainment, but the possibility of some misdiagnosis in either cohort cannot be entirely excluded based on available data.

A positive genetic correlation between ALS and schizophrenia predicts an excess of patients presenting with both diseases. Most neurologists and psychiatrists, however, will not readily acknowledge that these conditions co-occur frequently. Our genetic correlation estimate confers an odds ratio of 1.17 (1.08–1.26) for harbouring above-threshold liability for ALS given schizophrenia (or vice versa) and a lifetime risk of 1.34,336 for both phenotypes together. Thus, a very large incident cohort of 16,448 ALS patients (7,510–66,670), with detailed phenotype information, would be required to have sufficient power to detect an excess of schizophrenia within an ALS cohort. Coupled with reduced life expectancy in patients with schizophrenia\(^\), this may explain the relative dearth of epidemiological studies to date providing clinical evidence of excess comorbidity. Moreover, it has also been proposed that prolonged use of antipsychotic medication may protect against developing all of the clinical features of ALS\(^3\), which would reduce the rate of observed comorbidity. Considering our novel evidence for a genetic relationship between ALS and schizophrenia, this underscores the intriguing possibility that therapeutic strategies for each condition may be useful in the other, and our findings provide rationale to consider the biology of ALS and schizophrenia as related in future drug development studies. Indeed, the glutamate-modulating ALS therapy riuzole has shown efficacy as an adjunct to risperidone, an antipsychotic medication, in reducing the negative symptoms of schizophrenia\(^4\).

In conclusion, we have estimated the genetic correlation between ALS and schizophrenia to be 14.3% (7.05–21.6), providing molecular genetic support for our epidemiological observation of psychiatric endophenotypes within ALS kindreds. To our knowledge, this is the first study to show genetic correlation derived from polygenic overlap between neurodegenerative and neuropsychiatric phenotypes. The presence of both apparent monogenic C9orf72-driven overlap\(^2\) and polygenic overlap in the aetiology of ALS and schizophrenia suggests the presence of common biological processes, which may relate to disruption of cortical circuitry. As both ALS and schizophrenia are heterogeneous conditions, further genomic, biological and clinical studies are likely to yield novel insights into the pathological processes for both diseases and will provide clinical sub-stratification parameters that could drive novel drug development for both neurodegenerative and psychiatric conditions.

### Methods

#### Study population and genetic data.

For ALS, 7,740,343 SNPs genotyped in 12,577 ALS patients and 23,475 healthy controls of European ancestry organized in 27 platform- and country-defined strata were used\(^2\). The schizophrenia dataset comprised GWAS summary statistics for 9,444,230 SNPs originally genotyped in 34,241 patients and 45,604 controls of European and Asian ancestry\(^7\). For LD score regression, GWAS summary statistics were generated for the ALS cohort using mixed linear model association testing implemented in Genome-wide Complex Trait Analysis\(^11\) or logistic regression combined with cross-stratum meta-analysis using METAL\(^12\). To evaluate sample overlap for PRS and FTD-onset analyses, we also obtained individual-level genotype data for 27,647 schizophrenia cases and 33,675 controls from the schizophrenia GWAS (Psychiatric Genomics Consortium\(^2\) and dbGaP accession number phs000021.v3.p2). Using 88,971 LD-pruned (window size 200 SNPs; shift 20 SNPs; \(\gamma > 0.25\)) SNPs in both datasets (INFO score > 0.8; MAF > 0.2), with SNPs in high-LD regions removed (Supplementary Table 4), samples were removed from the ALS dataset if they were duplicated or had a cryptically related counterpart (PLINK \(r > 0.1; 5,582\) individuals) in the schizophrenia cohort and whole strata (representing Finnish and German samples; 3,811 individuals) were also removed if commonality with the schizophrenia cohort could not be ascertained (due to unavailability of individual-level genotype data in the schizophrenia cohort) and in which a sample overlap was suspected (Supplementary Table 3).

#### LD score regression.

We calculated LD scores using LDSC v1.0.0 in 1 centiMorgan windows around 13,307,412 non-singleton variants genotyped in 379 European individuals (CEU, FIN, GBR, IBS and TSI populations) in the phase 1 integrated release of the 1,000 Genomes Project\(^5\). For regression weights\(^13\), we restricted LD score calculation to SNPs included both in the GWAS summary statistics and HapMap phase 3; for \(r_g\) estimation in pairs of traits this was the intersection of SNPs for both traits and HapMap. Because population structure and confounding were highly controlled in the ALS summary statistics by the use of mixed linear model association tests, we constrained the LD score regression intercept to 1 for \(h_g^2\) estimation in ALS, and we also estimated \(h_g^2\) with a free intercept. For \(r_g\) estimation in pairs of traits and \(r_g\) estimation the intercept \(h_g^2\) was a free parameter. We also estimated \(r_g\) using ALS meta-analysis results\(^7\) with free and constrained intercepts and with permuted data conserving population structure. Briefly, principal component analysis was carried out for each stratum using smartpca\(^26\) and the three-dimensional space defined by principal components 1-3 was equally subdivided into 1,000 cubes. Within each cube, case–control labels were randomly swapped and association statistics were re-calculated for the entire stratum using logistic regression. Study-level \(P\)-values were then calculated using inverse variance weighted fixed effect meta-analysis implemented in METAL\(^12\). \(h_g^2\) were estimated for these meta-analysed permuted data using LD score regression (Supplementary Table 1).

#### Polygenic risk score analysis.

We calculated PRS for 10,032 cases and 16,627 healthy controls in the ALS dataset (duplicate and suspected or confirmed related samples with the schizophrenia dataset removed) and for 34,216 schizophrenia-associated alleles and effect sizes reported in the GWAS summary statistics for 6,843,674 SNPs included in both studies and in the phase 1 integrated release of the 1,000 Genomes Project\(^5\) (imputation INFO score < 0.3; minor allele frequency < 0.01; A/T and G/C SNPs removed). SNPs were clumped in two rounds (physical distance threshold of 250 kb and a LD threshold \((R^2) > 0.5\) in the first round and a distance of 5,000 kb and LD threshold of > 0.2 in the second round) using PLINK v1.90b3y, removing high-LD regions (Supplementary Table 4), resulting in a final set of 496,548 SNPs for PRS calculations. Odds ratios for autosomal SNPs reported in the schizophrenia summary statistics were log-converted to beta values and PRS calculated using PLINK’s score function for twelve schizophrenia GWAS \(P\)-value thresholds (\(P_j\)): \(5 \times 10^{-5}\), \(5 \times 10^{-5}\), \(5 \times 10^{-5}\), \(5 \times 10^{-5}\), \(5 \times 10^{-5}\), \(5 \times 10^{-5}\), \(5 \times 10^{-5}\), \(5 \times 10^{-5}\), 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5. A total of 100 principal components (PCs) were generated for the ALS sample using GCTA version 1.2.4.4. Using R version 3.2.2, a generalized linear model was applied to model the phenotype of individuals in the ALS dataset. PCs that had a significant effect on the phenotype \((P < 0.0005\); Bonferroni-corrected for 100 PCs) were selected (PCs 1, 4, 5, 7, 8, 10, 11, 12, 14, 36, 49).
To estimate explained variance of PRS on the phenotype, a baseline linear relationship including only sex and significant PCs as variables was modelled first:

\[ y = \beta_0 + \beta_x x + \sum \beta_p x_{p*}, \]

where \( y \) is the phenotype in the ALS dataset, \( x \) is the intercept of the model with a slope \( \beta_x \) for each variable \( x \).

Subsequently, a linear model including polygenic scores for each schizophrenia \( P_t \) was calculated:

\[ y = \beta_0 + \beta_x x + \sum \beta_p x_{p*} + \sum \beta_{pr} x_{pr*}. \]

A Nagelkerke \( R^2 \) value was obtained for every model and the baseline Nagelkerke \( R^2 \) value was subtracted, resulting in a \( \Delta \) explained variance that describes the contribution of schizophrenia-based PRS to the phenotype in the ALS dataset. PRS analysis was also performed in permuted case–control data (1,000 permutations, conserving case–control ratio) to assess whether the increased \( \Delta \) explained variance was a true signal associated with phenotype. \( \Delta \) explained variances and \( P \)-values were averaged across permutation analyses.

To ensure we did not over- or under-correct for population effects in our model, we tested the inclusion of up to a total of 30 PCs in the model, starting with the PC with the most significant effect on the ALS phenotype (Supplementary Fig. 2). Increasing the number of PCs initially had a large effect on the explained variance, but this effect levelled out after 11 PCs. On the basis of this test we are confident that adding the 11 PCs that had a significant effect on the phenotype sufficiently accounted for possible confounding demographic factors. For the schizophrenia \( P_t \) for which we obtained the highest \( \Delta \) explained variance (0.2), we subdivided observed schizophrenia-based PRS in the ALS cohort into deciles and calculated the odds ratio for being an ALS case in each decile compared to the first decile using a similar generalized linear model:

\[ y = \beta_0 + \beta_x x + \sum \beta_p x_{p*} + \beta_{pr} x_{pr*}. \]

Odds ratios and 95% confidence intervals for ALS were derived by calculating the exponential function of the beta estimate of the model for each of the deciles 2–10.

**Diagnostic misclassification.** To distinguish the contribution of misdiagnosis from true genetic pleiotropy we used BUMHBOX with 417 independent ALS risk alleles in a sample of 27,647 schizophrenia patients for which individual-level genotype data were available. We also estimated the required misdiagnosis rate \( M \) of FTD–ALS as schizophrenia that would lead to the observed genetic correlation estimate as \( C / (C + 1) \), where \( C = p_g N_{GZSF} / N_{ALS} \) and \( N_{GZSF} \) and \( N_{ALS} \) are the number of cases in the schizophrenia and ALS datasets, respectively (derived in Supplementary Methods 1).

**Expected comorbidity.** To investigate the expected comorbidity of ALS and schizophrenia given the observed genetic correlation, we modelled the distribution in liability for ALS and schizophrenia as a bivariate normal distribution with the liability-scale covariance determined by LD score regression (Supplementary Methods 2). Lifetime risks for ALS and schizophrenia were estimated using the binomial power equation.

**Pleiotropy-informed risk loci for ALS.** Using an adapted cFDR method that allows shared controls between cohorts\(^2\), we estimated per-SNP cFDR given LD score-corrected\(^2\) schizophrenia GWAS \( P \)-values for ALS mixed linear model summary statistics calculated in a dataset excluding Finnish and German cohorts (in which suspected control overlap could not be determined), but including all other cohorts\(^2\). cFDR values were calculated to identify loci above which individuals develop ALS or schizophrenia, or both. The expected proportions of individuals above these thresholds were used to calculate the odds ratio of developing ALS given schizophrenia, or vice versa (Supplementary Methods 2). The required population size to observe a significant excess of comorbidity was calculated using the binomial power equation.

**References**


Acknowledgements

We acknowledge helpful contributions from Mr Gert Jan van de Vendel in the design and execution of PRS analyses. This study received support from the ALS Association; Fondation Thierry Latran; the Motor Neurone Disease Association of England, Wales and Northern Ireland; Science Foundation Ireland; Health Research Board (Ireland), The Netherlands ALS Foundation (Project MinE), to J.H.V., L.H.v.d.B., the Netherlands Organisation for Health Research and Development (Vici scheme, L.H.v.d.B.) and ZonMW under the frame of E-Rare-2, the ERA Net for Research on Rare Diseases and the Alzheimer’s Association Grant ADGC-10–196728. The Project MinE GWAS Consortium included contributions from the PARALS registry, SLALOM group, SLAP registry, FAIS Sequencing Consortium, SLAGEN Consortium and NNIPPS Study Group; the Schizophrenia Working Group of the Psychiatric Genomics Consortium included contributions from the Psychosis Endophenotypes International Consortium and Wellcome Trust Case-control Consortium. Members of these eight consortia are listed in Supplementary Note 2.

Author contributions

O.H., J.H.V. and A.A.-C. conceived the study. R.L.McL., D.S., W.v.R., K.R.v.E., M.O.B., D.G.B., A.A.-C., L.H.v.d.B., J.J.L., O.H. and J.H.V. contributed to study design. R.L.McL., D.S. and W.v.R. conducted the analyses. R.L.McL., D.S., O.H., J.J.L. and J.H.V. drafted the manuscript. B.S.K., B.A.O. and A.G. provided data and critical revision of the manuscript. This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/
Project MinE GWAS Consortium

Schizophrenia Working Group of the Psychiatric Genomics Consortium

New York, USA. 206Department of Priority Centre for Translational Neuroscience and Mental Health, University of Newcastle, Newcastle, Australia. 207Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University in Szczecin, Szczecin, Poland. 208Department of Mental Health and Substance Abuse Services, National Institute of Health and Welfare, Helsinki, Finland. 209Department of Mental Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA. 210Department of Psychiatry, University of Bonn, Bonn, Germany. 211Department de la Recherche Scientifique, Laboratoire de Génétique Moléculaire de la Neurotransmission et des Processus Neurodégénératifs, Hôpital de la Pitié Salpêtrière, Paris, France. 212Department of Genomics Mathematics, University of Bonn, Bonn, Germany. 213Research Unit, Sarlandet Hospital, Kristiansand, Norway. 214Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, USA. 215Virginia Boston Health Care System, Brockton, Massachusetts, USA. 216Department of Psychiatry, National University of Ireland Galway, Ireland. 217Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK. 218Division of Psychiatry, University of Edinburgh, Edinburgh, UK. 219Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. 220Massachusetts Mental Health Center Public Psychiatry Division of the Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA. 221Estonian Genome Center, University of Tartu, Tartu, Estonia. 222School of Psychology, University of Newcastle, Newcastle, Australia. 223First Psychiatric Clinic, Medical University, Sofia, Bulgaria. 224Department P. Arhus University Hospital, Risskov, Denmark. 225Department of Psychiatry, Royal College of Surgeons in Ireland, Ireland. 226King’s College London, London, UK. 227Maastricht University Medical Centre, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht, The Netherlands. 228Institute of Translational Medicine, University of Liverpool, UK. 229Max Planck Institute of Psychiatry, Munich, Germany. 230Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. 231Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany. 232Department of Psychiatry, Queensland Brain Institute and Queensland Centre for Mental Health Research, University of Queensland, Brisbane, Queensland, Australia. 233Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. 234Department of Psychiatry, Trinity College Dublin, Dublin, Ireland. 235Eli Lilly and Company, Lilly Corporate Center, Indianapolis, Indiana, USA. 236Department of Clinical Sciences, Psychiatry, Umeå University, Umeå, Sweden. 237DETECT Early Intervention Service for Psychosis, Blackrock, Dublin, Ireland. 238Centre for Public Health, Institute of Clinical Sciences, Queens University Belfast, Belfast, UK. 239Lawrence Berkeley National Laboratory, University of California at Berkeley, Berkeley, California, USA. 240Institute of Psychiatry at King’s College London, London, UK. 241Melbourne Neuropsychiatry Centre, University of Melbourne & Melbourne Health, Melbourne, Australia. 242Department of Psychiatry, University of Helsinki, Finland. 243Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Helsinki, Finland. 244Medical Faculty, University of Belgrade, Belgrade, Serbia. 245Department of Psychiatry, University of North Carolina, Chapel Hill, North Carolina, USA. 246Institute for Molecular Medicine Finland, FIMM, Helsinki, Finland. 247Department of Epidemiology, Harvard University, Boston, Massachusetts, USA. 248Department of Psychiatry, University of Oxford, Oxford, UK. 249Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia, USA. 250Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, New York, USA. 251PharmaTherapeutics Clinical Research, Pfizer Worldwide Research and Development, Cambridge, Massachusetts, USA. 252Department of Psychiatry and Psychotherapy, University of Gottingen, Gottingen, Germany. 253Psychiatry and Psychotherapy Clinic, University of Erlangen, Erlangen, Germany. 254Hunter New England Health Service, Newcastle, Australia. 255School of Biomedical Sciences, University of Newcastle, Newcastle, Australia. 256Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA. 257University of Iceland, Landskapi, National University Hospital, Reykjavik, Iceland. 258Department of Psychiatry and Drug Addiction, Tbilisi State Medical University (TSMU), Tbilisi, Georgia. 259Research and Development, Bronx Veterans Affairs Medical Center, New York, New York, USA. 260Wellcome Trust Centre for Human Genetics, Oxford, UK. 261deCODE Genetics, Reykjavik, Iceland. 262Department of Clinical Neurology, Medical University of Vienna, Vienna, Austria. 263Lieber Institute for Brain Development, Baltimore, Maryland, USA. 264Department of Medical Genetics, University Medical Centre, Utrecht, The Netherlands. 265Rudolf Magnus Institute of Neuroscience, University Medical Centre Utrecht, Utrecht, The Netherlands. 266Berkshire Healthcare N.H.S. Foundation Trust, Bracknell, UK. 267Section of Psychiatry, University of Verona, Verona, Italy. 268Department of Psychiatry, University of Oulu, Oulu, Finland. 269University Hospital of Oulu, Oulu, Finland. 270Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland. 271Health Research Board, Dublin, Ireland. 272Department of Psychiatry and Clinical Neurosciences, School of Psychiatry and Clinical Neurosciences, Queen Elizabeth I.I. Medical Centre, Perth, Western Australia, Australia. 273Department of Psychological Medicine and Neurology, MRC. Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Cardiff, Wales, UK. 274Computational Sciences CoE, Pfizer Worldwide Research and Development, Cambridge, Massachusetts, USA. 275Human Genetics, Genome Institute of Singapore, Singapore. 276University College London, London, UK. 277Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York, USA. 278Department of Genetics, The Hebrew University of Jerusalem, Jerusalem, Israel. 279Neuroscience Discovery and Translational Area, Pharma Research and Early Development, F. Hoffman-La Roche, Basel, Switzerland. 280School of Psychiatry and Clinical Neurosciences, The University of Western Australia, Perth, Australia. 281The Perkins Institute of Medical Research, Perth, Australia. 282LWA Centre for Clinical Research in Neuropsychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Departments of Psychiatry and Human and Molecular Genetics, Virginia Commonwealth University, Richmond, Virginia, USA. 283The Feinstein Institute for Medical Research, Manhasset, New York, USA. 284The Hofstra NS-LIJ School of Medicine, Hempstead, New York, USA. 285The Zucker Hillside Hospital, Glen Oaks, New York, USA. 286Saw Swee Hock School of Public Health, National University of Singapore, Singapore. 287Queensland Centre for Mental Health Research, University of Queensland, Brisbane, Queensland, Australia. 288The Broad Institute of M.I.T. and Harvard, Cambridge, Massachusetts, USA. 289Center for Human Genetic Research and Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA. 290Department of Child and Adolescent Psychiatry, Erasmus University Medical Centre, Rotterdam, The Netherlands. 291Department of Complex Trait Genetics, Neuroscience Campus Amsterdam, VU University Medical Center Amsterdam, Amsterdam, The Netherlands. 292Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands. 293University of Aberdeen, Institute of Medical Sciences, Aberdeen, Scotland, UK. 294Department of Psychiatry, Neurology, Neuroscience and Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA. 295Department of Psychiatry, University of Copenhagen, Copenhagen, Denmark.