

Genetic contributions to self-reported tiredness

Vincent Deary^{1*#}, Saskia P Hagenaars^{2,3,4*}, Sarah E Harris^{2,5}, W David Hill^{2,3}, Gail Davies^{2,3}, David CM Liewald BSc^{1,2}, International Consortium for Blood Pressure GWAS, CHARGE consortium Aging and Longevity Group, Andrew M McIntosh⁴, Catharine R Gale^{2,3,6†}, Ian J Deary PhD^{2,3†#}

1 Department of Psychology, Northumbria University, Newcastle, NE1 8ST, UK

2 Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK

3 Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK

4 Division of Psychiatry, University of Edinburgh, Edinburgh, EH10 5HF, UK

5 Medical Genetics Section, University of Edinburgh Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

6 MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, SO16 6YD, UK

*These authors contributed equally to the work.

†These authors contributed equally to the work.

Corresponding author,
Vincent Deary or Ian J. Deary

vincent.deary@northumbria.ac.uk

ian.deary@ed.ac.uk

Abstract

Self-reported tiredness and low energy, often called fatigue, is associated with poorer physical and mental health. Twin studies have indicated that this has a heritability between 6% and 50%. In the UK Biobank sample (N = 108 976) we carried out a genome-wide association study of responses to the question, “Over the last two weeks, how often have you felt tired or had little energy?” Univariate GCTA-GREML found that the proportion of variance explained by all common SNPs for this tiredness question was 8.4% (SE = 0.6%). GWAS identified one genome-wide significant hit (Affymetrix id 1:64178756_C_T; $p = 1.36 \times 10^{-11}$). LD score regression and polygenic profile analysis were used to test for pleiotropy between tiredness and up to 28 physical and mental health traits from GWAS consortia. Significant genetic correlations were identified between tiredness and BMI, HDL cholesterol, forced expiratory volume, grip strength, HbA1c, longevity, obesity, self-rated health, smoking status, triglycerides, type 2 diabetes, waist-hip ratio, ADHD, bipolar disorder, major depressive disorder, neuroticism, schizophrenia, and verbal-numerical reasoning (absolute rg effect sizes between 0.11 and 0.78). Significant associations were identified between tiredness phenotypic scores and polygenic profile scores for BMI, HDL cholesterol, LDL cholesterol, coronary artery disease, HbA1c, height, obesity, smoking status, triglycerides, type 2 diabetes, and waist-hip ratio, childhood cognitive ability, neuroticism, bipolar disorder, major depressive disorder, and schizophrenia (standardised β 's between -0.016 and 0.03). These results suggest that tiredness is a partly-heritable, heterogeneous and complex phenomenon that is phenotypically and genetically associated with affective, cognitive, personality, and physiological processes.

“Hech, sirs! But I’m wabbit, I’m back frae the toon;
I ha’ena dune pechin’—jist let me sit doon.

From Glesca’

By William Dixon Cocker (1882-1970)

Introduction

The present study examines genetic contributions to how the UK Biobank’s participants answered the question, “Over the last two weeks, how often have you felt tired or had little energy?” Ideal questionnaire items don’t have conjunctions, but the “or” is understandable here, and it may even allow capture of both peripheral and central fatigue. The first and last authors of the present study grew up in South Lanarkshire in Scotland, where fatigue was often self-reported in terms of feeling “wabbit”. The Scots word wabbit encompasses both peripheral fatigue, the muscle weakness after a long walk, and central fatigue, the reduced ability to initiate and/or sustain mental and physical activity, such as we might experience whilst having flu. Throughout the paper we refer mainly to the single English words “fatigue” and/or “tiredness” as the construct captured by the question, but the Scottish vernacular word is a good reminder that these constructs are multi-dimensional.

Fatigue is a common complaint. In a Dutch adult general population survey with 9375 respondents, 4.9% reported short term fatigue (<6 months duration), 30.5% chronic fatigue (> 6 months duration), and 1% fulfilled diagnostic criteria for Chronic Fatigue Syndrome (CFS)¹. These findings are similar to a London-based survey of general practice patients in England, aged 18-45, with 15283 respondents, where 36.7% reported substantial fatigue, 18.3% substantial fatigue of > 6 months duration and 1% fulfilled the diagnostic criteria for CFS². Two other large surveys of US workers and community dwelling adults aged 51 and over report fatigue rates of 37.9% (two week period prevalence) and 31.2% (one week period prevalence) respectively^{3,4}. In an early review of fatigue epidemiology, Lewis and Wessely⁵ argue that fatigue “is best viewed on a continuum”, and the continuous distribution of fatigue in the general population is supported by the Pawlikowska et al. study². Fatigue is also a common presentation in primary care. In a survey of 1428 consultations to 89 GPs in Ireland, fatigue prevalence was 25% and the main reason for attendance in 6.5%⁶.

Demographically, higher levels of self-reported fatigue are associated with female gender, lower socio-economic status¹ and poorer self-rated health status⁷. There are less clear

associations between age and fatigue, with some studies reporting a small but significant positive correlation between age and fatigue², whereas others report no association⁶ or a negative association^{1,7}. There is a clearer link with the Fried phenotype of frailty in older adults⁸, which has significant associations with mortality. The frailty phenotype comprises weakness (as measured by grip strength), weight loss, reduced mobility, reduced walking speed and fatigue⁹.

Fatigue is associated with a number of lifestyle-related factors and conditions. Smoking is a risk factor for fatigue¹⁰ and fatigue has strong cross-sectional associations with Type 2 Diabetes¹¹ and increased BMI¹². Fatigue is consistently associated with poorer physical and mental health status. It is one of the most common symptom complaints of cancer patients. For those undergoing treatment, prevalence estimates vary between 25% and 99%, and 25 to 30% of survivors report long term fatigue¹³. Fatigue is also a significant symptom of, to name just a few conditions, primary biliary cirrhosis (PBC)¹⁴, multiple sclerosis¹⁵, rheumatoid arthritis¹⁶, primary Sjogren's syndrome¹⁷, and Parkinson's disease¹⁸. It is associated with chronic disease in general, and there is a linear relationship between number of chronic diseases and self-reported fatigue¹⁹. Fatigue is also strongly associated with depression²⁰, with self-reported psychological distress^{2,21}, and with the personality trait of neuroticism^{22,23}.

Research into the biological mechanisms of fatigue has focussed on a few key areas. Fatigue is associated with the cytokine-mediated inflammatory response, particularly Interleukin (IL)-1Beta and IL-6. These latter have been shown, for instance, to be elevated in cancer patients²⁴, and administration of interferon-alpha produces depression and/or fatigue in the majority of patients receiving it as a treatment²⁵. HPA axis dysregulation in the form of hypocortisolaemia, blunted diurnal variation and blunted stress reactivity have been found in cross-sectional studies of CFS patients (see Tomas et al.²⁶ for a recent review). Other popular candidate aetio-pathological mechanisms for fatigue include serotonin pathways, circadian dysregulation, autonomic dysfunction¹⁷, 5HT neurotransmitter dysregulation, alterations in ATP metabolism, and vagal afferent activation²⁴. Some authors have suggested that, rather than being located with one biological system, fatigue represents a systemic dysregulation of the interaction between these systems²⁷.

At the other end of the biopsychosocial spectrum, psychosocial models of fatigue focus on the role that the individual's response to their symptom may serve in perpetuating it. For

instance, in a cross sectional study of 149 patients with multiple sclerosis, illness-related cognitions and behaviours were associated with a higher level of fatigue independent of neurological impairment²⁸. More integrative models are predicated on the notion of allostatic load, the psychophysiological work done to adjust to stress, and its impact upon the body's self-regulatory systems. As such, these models complement biological accounts of fatigue and provide potential pathways for integrating psychosocial and biological findings²⁹. Multifactorial accounts and models of fatigue exist in multiple sclerosis³⁰, primary biliary cirrhosis¹⁴, obesity³¹, diabetes³², frailty³³ and cancer¹³. These multi-factorial models postulate that fatigue is likely to be the product of physiological factors (generic, such as inflammation, and/or disease specific such as hyperglycaemia in diabetes), psychosocial factors (e.g. emotional distress), lifestyle and behavioural factors (e.g. reduced activity), illness consequences (e.g. sleep disturbance, weakness), and the interaction of these contributors.

Research into the genetics and epigenetics of fatigue has tended to focus on genes associated with the biological mechanisms described above, and done so usually within fatiguing illnesses such as those listed above. Candidate gene studies have suggested several genes to be involved in chronic fatigue syndrome, particularly genes involved in the immune system and HPA axis. Reviewing this literature Landmark-Høyvik et al.³⁴ suggest that findings are inconclusive, and are hampered by phenotypic heterogeneity, lack of power, and poor study design. Twin studies have shown heritability of fatigue to be between 6% and 50% with a higher concordance in monozygotic twins than dizygotic twins^{35, 36}. In a study of fatigue, insomnia and depression in 3758 twins (893 MZ pairs and 884 DZ pairs) the best model was a common pathway model suggesting that the high association between the symptoms (correlations of 0.35 to 0.44) was mediated by an underlying common factor whose variation was 49% genetic and 51% unique environmental. Unique specific variance in fatigue was 38% genetic and 62% unique environmental³⁷. Genome-wide association studies have shown an association between SNPs in genes associated with impaired cognitive abilities (*GRIK2*, $p = 1.26 \times 10^{-11}$)^{38, 39} and the circadian clock (*NPAS2*, not genome-wide significant)³⁹ and chronic fatigue syndrome, but this was in a sample of just 42 cases of chronic fatigue syndrome and 38 controls.

To sum up in the words of Landmark-Høyvik et al.⁴⁰: “fatigue can be conceptualized as a final common end point for psychological and biological processes. Fatigue is therefore both

heterogeneous (occurring across different conditions) and multifactorial”. The aim of the present study is to understand further the genetic contribution to self-reported tiredness and/or low energy. We conducted a genome-wide association analysis, in the UK Biobank sample, of a response to a single item question: “Over the past two weeks, how often have you felt tired or had little energy?” Based on the foregoing literature overview, we also investigated pleiotropy with physical- and mental health- related traits.

Materials and methods

Study design and participants

UK Biobank is a large resource for identifying determinants of human diseases in middle aged and older individuals (<http://www.ukbiobank.ac.uk>)⁴¹. A total of 502 655 community-dwelling individuals aged between 37 and 73 years were recruited in the United Kingdom between 2006 and 2010. Baseline assessment included cognitive testing, personality self-report, and physical and mental health measures. For the present study, genome-wide genotyping data were available for 112 151 participants (58 914 females, 53 237 males) after quality control (see below). They were aged from 40 to 73 years (mean = 56.9 years, SD = 7.9). UK Biobank received ethical approval from the Research Ethics Committee (REC reference for UK Biobank is 11/NW/0382). This study has been completed under UK Biobank application 10279. Figure 1 shows the study flow for the present report.

Procedures

Tiredness

Participants were asked the question, "Over the past two weeks, how often have you felt tired or had little energy?" Possible answers were: "Not at all/Several days/More than half the day/Nearly every day/Do not know/Prefer not to answer". This question was asked as part of the Mental Health Questionnaire, which consists of items from the Patient Health Questionnaire⁴². Participants answering with "Do not know" or "Prefer not to answer" were excluded, resulting in a four-category variable for tiredness ranging from "Not at all" to "Nearly every day". We will refer to this question in the rest of the paper as 'tiredness' but we ask the reader to bear in mind the question as it was asked, i.e. its referring to tiredness and/or low energy.

Genotyping and quality control

The interim release of UK Biobank included genotype data for 152 729 individuals, of whom 49 979 were genotyped using the UK BiLEVE array and 102 750 using the UK Biobank axiom array. These arrays have over 95% content in common. Details of the array design, genotyping procedures and quality control details have been published elsewhere^{43,44}.

Imputation

An imputed dataset was made available as part of the UK Biobank interim data release. The 1000 Genomes phase 3 and UK10K haplotype reference panels were merged and the genotype data imputed to this merged reference panel. Further details can be found at the following URL: <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=157020>. Autosomal variants with a minor allele frequency less than 0.1 % and an imputation quality score of 0.1 or less were excluded from further analysis (N ~ 17.3M SNPs).

Curation of summary results from GWAS consortia on health related variables

Published summary results from international GWAS consortia were gathered in order to perform linkage disequilibrium (LD) score regression and polygenic profile analysis between the UK Biobank tiredness variable and the genetic predisposition to multiple health-related traits. Details of the health-related variables, the consortia's websites, key references for each consortium, and number of subjects included in each consortium's GWAS are given in Supplementary Table 1.

Statistical analysis

Phenotypic correlations

Spearman's rank correlation coefficients were calculated between responses to the tiredness question, grip strength, forced expiratory volume in 1s, height, BMI, self-rated health, verbal-numerical reasoning and neuroticism, all of which were measured phenotypes in UK Biobank. Details on measurements of these phenotypes can be found in the Supplementary Material.

Genetic association analysis

111 749 individuals answered the tiredness question and had genotypic information. After visual inspection of the distribution of the UK Biobank tiredness variable no exclusions were made. Prior to analysis, tiredness was adjusted for age, gender, assessment centre, genotyping batch and array, and 10 principal components for population stratification. Genotype-phenotype association analyses were conducted using SNPTEST v2.5.1⁴⁵ and can be found at the following URL:

https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html#introduction. An additive model was specified using the 'frequentist 1' option. Genotype uncertainty was accounted for by analysing genotype dosages.

Genetic association analyses were also performed on the following UK Biobank phenotypes in order to perform further analyses: self-rated health⁴⁶, grip strength, forced expiratory volume in 1s, neuroticism⁴⁷, verbal-numerical reasoning⁴⁸. We specifically examined whether any variants associated with tiredness were also associated with grip strength, self-rated health, and neuroticism because we judged these to provide some coverage of physical and mental resilience in UK Biobank.

Estimation of SNP based heritability

To estimate the proportion of variance explained by all common SNPs in tiredness, univariate GCTA-GREML analysis was performed⁴⁹. This analysis included only unrelated individuals, using a relatedness cut-off of 0.025 in the generation of the genetic relationship matrix.

Gene-based association analysis

Gene-based associations were derived using MAGMA⁵⁰ using the summary GWAS statistics for tiredness. SNPs were assigned to 18 062 genes using the NCBI build 37.3. The gene boundary was defined as the start and stop site of each gene. In order to account for linkage disequilibrium between the SNPs used, the European panel of the 1000 Genomes data (phase 1, release 3) was used. A Bonferroni correction was used to control for 18 062 tests ($\alpha=0.05/18,062$; $P < 2.768 \times 10^{-6}$).

Partitioned heritability

The summary statistics from the GWAS on tiredness were portioned into functional categories using the same data processing pipeline as Finucane et al.⁵¹; more details on this method can be found in the Supplementary Materials.

Pleiotropy; LD score regression and polygenic profiling

Genetic associations between tiredness and health-related variables from GWAS consortia were computed using two methods, LD score regression and polygenic profile score analysis. Each provides a different metric to infer the existence of pleiotropy between pairs of traits. LD score regression was used to compute genetic correlations between two traits; this tests the degree to which the polygenic architecture of one trait overlaps with that of other traits. Polygenic profile score analysis was used to test the extent to which individual differences in the tiredness phenotype in UK Biobank could be predicted by the polygenic architecture of the health-related traits from other GWAS consortia. Both of these methods are dependent on

a trait's being highly polygenic in nature, that is, where a large number of variants of small effect contribute toward phenotypic variation^{52, 53}. LD score regression was performed between tiredness and 28 health-related traits. Polygenic profile score analysis was performed on 25 of the 28 health-related traits, as this method requires independent samples to provide the summary GWAS information from which the polygenic profile score is based.

LD score regression. This was used to quantify the extent of pleiotropy between tiredness in UK Biobank and 28 health-related traits^{53, 54}. This technique examines the correlational structure of the SNPs found across the genome. In the present study, LD score regression was used to derive genetic correlations between tiredness and health-related traits using the GWAS results of 24 large GWAS consortia and four UK Biobank phenotypes. The data processing pipeline devised by Bulik-Sullivan et al.⁵³ was followed. In order to ensure that the genetic correlation for the Alzheimer's disease phenotype was not driven by a single locus or biased the fit of the regression model, a 500kb region centred on the *APOE* locus was removed and this phenotype was re-run. This additional model is referred to in the tables and figures as 'Alzheimer's disease (500kb)'.

Polygenic profile scores. The UK Biobank genotyping data was recoded from numeric (1,2) allele coding to standard ACGT coding using a bespoke programme developed by one of the present authors (DCL)⁴⁴. Polygenic profile scores were created for 25 health-related traits in all genotyped participants using PRSice⁵⁵. Prior to creating the scores, SNPs with a minor allele frequency < 0.01 were removed and clumping was used to obtain SNPs in linkage equilibrium with an $r^2 < 0.25$ within a 200bp window. Five polygenic profile scores were created for each trait including SNPs according to their significance of association with the relevant trait at p-value thresholds of $P < 0.01$, $P < 0.05$, $P < 0.1$, $P < 0.5$ and all SNPs.

Regression models were used to examine the association between the polygenic profiles and tiredness in UK Biobank, adjusting for age at measurement, sex, genotyping batch and array, assessment centre, and the first ten principal components for population stratification. All associations were corrected for multiple testing using the false discovery rate method⁵⁶. Sensitivity analyses were performed to test whether the results were confounded by individual's neuroticism levels, their self-rated health scores, or a diagnosis of major depressive disorder. This was done by adjusting the models for the neuroticism and self-rated health scores. Individuals with a probable diagnosis of major depressive disorder were

excluded from the sensitivity analysis, based on the diagnostic method formulated by Smith et al.⁵⁷. Further details can be found in the Supplementary Material. To examine whether any association between polygenic risk for type 2 diabetes and tiredness was confounded by having had a diagnosis of type 2 diabetes, all individuals with a self-reported doctor's diagnosis of type 2 diabetes were excluded from that specific sensitivity analysis (Supplementary Material). Multivariate regression was performed using all FDR significant polygenic profile scores and earlier described covariates.

Comparison of gene-based analysis results within UK Biobank

Gene-based associations for tiredness were compared with gene-based results for other UK Biobank health-related traits that, in the present report's results, showed a statistically significant genetic correlation with tiredness, using previously described methods⁵⁰.

Results

Phenotypic correlations

108 976 individuals from UK Biobank with genotypic data answered the question “Over the past two weeks, how often have you felt tired or had little energy?”, referred to hereinafter as “tiredness”. There were 51 416 individuals who answered “not at all”, 44 208 individuals responded “several days”, 6404 individuals answered “more than half the days”, and 6948 individuals responded “nearly every day”. Correlations indicated that individuals who reported feeling more tired tended to have lower grip strength, lower lung function, poorer self-rated health, lower scores for verbal-numerical reasoning, and shorter stature (Table 1). Correlations indicated that individuals who reported feeling more tired tended to have a higher BMI and higher neuroticism scores (Table 1). Absolute effect sizes ranged from very small to moderate (Supplementary Figures 1a-f). The mean scores and distribution for each of these variables at each level of tiredness are shown in Supplementary Figure 1 and Supplementary Table 2.

Genome-wide association study

There was one genome-wide significant SNP (Affymetrix id 1:64178756_C_T; $p = 1.36 \times 10^{-11}$) on chromosome 1 (Figure 2). This SNP is not in a gene and does not have an rs id. It has both a low minor allele frequency (0.001) and a low imputation quality score (0.43). It is not in a peak with other SNPs. Therefore, this result should be treated with caution. Two suggestive peaks were identified on chromosomes 1 and 17, with the lowest p values of 5.88×10^{-8} (rs142592148; an intronic SNP in *SLC44A5*) and 6.86×10^{-8} (rs7219015; an intronic SNP in *PAFAH1B1*) for each peak, respectively. The peak on chromosome 1 contains three genes (*CRYZ*, *TYW3* and *SLC44A5*) and the peak on chromosome 17 contains one gene (*PAFAH1B1*). The *CRY/TYW3* locus has previously been associated with circulating resistin levels, a hormone associated with insulin resistance, inflammation, and risk of type 2 diabetes and cardiovascular disease⁵⁸. *SLC44A5* encodes a solute carrier protein and is important for metabolism of lipids and lipoproteins and has been associated with birth weight in cattle⁵⁹. *PAFAH1B1* encodes a subunit of an enzyme that has important roles in brain development and spermatogenesis. Mutations in this gene cause the neurological disorder lissencephaly⁶⁰.

SNP based heritability estimate

The proportion of variance in tiredness explained by all common genetic variants using GCTA-GREML was 8.4 % (SE 0.006).

Gene based association analysis

Gene-based association analysis identified five genes, *DRD2*, *PRRC2C*, *C3orf84*, *ANO10* and *ASXL3*, that attained genome-wide significance for tiredness (Table 2 and Supplementary Table 3). *DRD2* encodes a dopamine receptor and has previously been associated with psychiatric illnesses⁶¹. Alternative splicing of *PRRC2C* has been associated with lung cancer⁶². Mutations in *ANO10* cause cerebellar ataxias⁶³. *ASXL3* encodes a polycomb protein and mutations in this gene are associated with intellectual disability, feeding problems and distinctive facial features⁶⁴.

In addition, each of these genes was also nominally significant in a GWAS of neuroticism, with *DRD2* being genome wide significant in both phenotypes⁴⁷. A comparison with a UK Biobank GWAS of self-rated health²² showed that four of the five genes (*DRD2*, *PRRC2C*, *C3orf84* and *ASXL3*) were significant across phenotypes. Grip strength showed less overlap, with only *C3orf84* being nominally associated with grip strength. Of the five genes examined here, *C3orf84* was associated with each of the four phenotypes. Whereas the genetic correlations between these traits (see below) are likely to encompass multiple genic and non-genic regions, as well as unique points of overlap between pairs of phenotypes, the variants found in the *C3orf84* represent a point of the genome where the genetic architecture of these four traits converges.

Partitioned heritability

From the full baseline model using 52 annotations, only evolutionarily conserved regions were found to be enriched for tiredness (Supplementary Figure 2). This annotation contained only 2.6% of the SNPs from the summary statistics, but they collectively explained 40% of the heritability of tiredness (SE = 11%, enrichment metric = 15.34, SE = 4.05, P = 0.00039). By clustering the histone marks into tissue specific categories we found significant enrichment for variants found in the central nervous system (Supplementary Figure 3). This category contained 15% of the SNPs and explained 45% of the heritability (SE = 8%, enrichment metric = 3.02, SE = 0.54, P = 0.0002).

Genetic correlations between tiredness and physical and mental health traits

LD score regression was used to test if genetic variants associated with health-related traits also contribute towards tiredness in UK Biobank. Table 3 and Figure 3 show these genetic correlations. Positive significant (FDR corrected) genetic correlations were found between tiredness and BMI ($r_g = 0.20$), HbA1c ($r_g = 0.25$), obesity ($r_g = 0.21$), smoking status ($r_g = 0.20$), triglycerides ($r_g = 0.13$), type 2 diabetes ($r_g = 0.18$), waist-hip ratio ($r_g = 0.08$), ADHD ($r_g = 0.27$), bipolar disorder ($r_g = 0.14$), major depressive disorder ($r_g = 0.59$), neuroticism ($r_g = 0.62$), and schizophrenia ($r_g = 0.25$). Negative significant (FDR corrected) genetic correlations were found between HDL cholesterol ($r_g = -0.11$), forced expiratory volume in 1s ($r_g = -0.12$), grip strength ($r_g = -0.16$), longevity ($r_g = -0.39$), self-rated health ($r_g = -0.78$), and verbal-numerical reasoning ($r_g = 0.14$). These genetic correlations suggest that there are common genetic associations between tiredness and multiple physical- and mental health-related traits.

Polygenic prediction

The full results including all five thresholds can be found in Supplementary Table 4, as well as the number of SNPs included for the five thresholds in each trait. Table 4 shows the results for the polygenic profile scores analyses, using the most predictive threshold for each trait. Complete results for all five thresholds can be found in Supplementary Table 4. Higher polygenic profiles for nine physical health traits predicted increased tiredness (standardised β 's between 0.008 and 0.03) in UK Biobank: BMI, LDL cholesterol, coronary artery disease, HbA1c, obesity, smoking status, triglycerides, type 2 diabetes, and waist-hip ratio. Higher polygenic profiles for HDL cholesterol and height predicted lower tiredness ($\beta = -0.016$ and $\beta = -0.008$, respectively).

Of the mental health traits, higher polygenic risk for bipolar disorder, neuroticism, major depressive disorder, and schizophrenia were associated with increased tiredness (standardised β 's between 0.008 and 0.03). Polygenic profiles for childhood cognitive ability showed a negative association with tiredness ($\beta = -0.01$).

Sensitivity analysis showed that, when controlling for neuroticism, the associations between tiredness and polygenic profiles for BMI, obesity, type 2 diabetes, cholesterol (HDL and LDL), HbA1c, triglycerides, waist-hip ratio, childhood cognitive ability, and schizophrenia remained significant (after FDR correction), indicating that these associations are not confounded by scores for neuroticism. Similar analyses controlling for self-rated health

indicated that the following associations with tiredness are not confounded by self-rated health: bipolar disorder, neuroticism, major depressive disorder, and schizophrenia. Supplementary Table 5 shows the adjusted results and the percentage of attenuation in standardised β 's for the models.

When excluding individuals with a probable diagnosis of major depressive disorder (N = 7364) from all individuals with sufficient information about their mental health to make a probable depression diagnosis (N = 31 523), the associations between tiredness and polygenic profiles for BMI, HDL cholesterol, obesity, triglycerides, waist-hip ratio, neuroticism, and schizophrenia remained significant (after FDR correction), compared to BMI, HDL cholesterol, obesity, triglycerides, type 2 diabetes, waist-hip ratio, neuroticism, and schizophrenia in the full model (also N = 31 523) (Supplementary Table 6). When excluding individuals with a type 2 diabetes diagnosis (N = 725), the association between tiredness and the polygenic risk score for diabetes remained significant, indicating that this association is independent of self-reported morbidity for that disorder.

A multivariate regression model including 15 of the 16 significant polygenic profile scores (BMI, HDL cholesterol, LDL cholesterol, HbA1c, height, obesity, smoking status, triglycerides, type 2 diabetes, waist-hip ratio, bipolar disorder, childhood cognitive ability, major depressive disorder, neuroticism and schizophrenia) showed that polygenic profile scores for the following traits contributed independently to the association with tiredness: BMI, HDL cholesterol, triglycerides, waist-hip ratio, childhood cognitive ability, major depressive disorder, neuroticism, and schizophrenia. The scores together accounted for 0.25 % of the variance in tiredness (Supplementary Table 7).

Discussion

In the present study, pleiotropy was identified between tiredness and longevity, grip strength, multiple metabolic indicators, smoking status, neuroticism, childhood cognitive ability, depression, and schizophrenia. These analyses, combining data from the UK Biobank and many GWAS consortia, provide the first estimate of the overlap in the genetic variants contributing to the heritability of tiredness and these physical and mental health-related traits and disorders. Tiredness demonstrated significant SNP-based heritability of 8.4 %.

Five genes attained genome wide significance for tiredness: *DRD2*, *PRRC2C*, *ANO10*, *ASXL3* and *C3orf84*, the latter being an uncharacterised protein representing a point of genetic convergence between tiredness, neuroticism, grip strength and self-rated health. *DRD2*, *PRRC2C*, *ANO10*, *ASXL3*, and *DRD2* have previously been associated with psychiatric illnesses¹⁸, lung cancer¹⁹, cerebellar ataxias²⁰ and intellectual disability²¹ respectively. The three genes within the suggestive peak on chromosome 1 (*CRYZ*, *TYW3* and *SLC44A5*) have previously been associated with insulin resistance, inflammation, risk of type 2 diabetes, metabolism of lipids and lipoproteins and cardiovascular disease^{15,16}. These genes are consistent with the identification of pleiotropy between tiredness and metabolic irregularities, and perhaps more broadly with ‘metabolic syndrome’ and ‘allostatic load’. *PAFAH1B*, within a suggestive peak on chromosome 17 has important roles in brain development¹⁷. This is consistent with the identification of pleiotropy between tiredness and cognitive traits and the finding of significant enrichment for variants found in the central nervous system.

Evolutionarily conserved regions were found to be enriched for association with tiredness, consistent with findings for other quantitative traits including disease status⁵¹, suggesting that these are important loci where mutations result in detrimental phenotypes. The range of factors— affective, cognitive, behavioural, and physical—that are genetically associated with tiredness is in itself remarkable, and confirms the observation of Landmark-Høyvik et al³⁴, quoted in the introduction, that the related construct of fatigue is aetiologically heterogeneous and multifactorial. No overlap has been found with genes (*GRIK2*, *NPAS2*) identified in previous GWAS of fatigue^{38,39}; however, the sample size of these studies was too small to have enough power to detect statistically significant differences. The present study did not show significant associations for candidate genes previously identified³⁴.

Tiredness showed significant shared heritability with a range of factors associated with metabolic syndrome⁶⁵ including cholesterol, tri-glycerides, HbA1c, waist-hip ratio, BMI, obesity, and type 2 diabetes. Several of these factors are biomarkers of allostatic load^{61, 62}. The concept of allostatic load has been used in the context of both physical and mental ill health, including the symptom of fatigue. Conceptually, allostatic load represents the cumulative, physiological ‘wear and tear’ of a prolonged response to a stressor. Allostatic load has been shown to be a replicably-measurable multi-variate construct^{66, 67}, with first order factors comprising cardiovascular, immune, metabolic, anthropometric, and neuroendocrine markers²⁹. It is hypothesised that, in response to threats to homeostasis, the body’s self-regulatory mechanisms, such as the sympathetic-adrenal medullary axis and the hypothalamic pituitary adrenal axis, have the potential to “overcompensate and eventually collapse upon themselves”²⁹, with consequences for morbidity and mortality.

In our current analysis we included metabolic (cholesterol, HbA1c, triglycerides), anthropometric (waist hip ratio, BMI, obesity) and cardiovascular/respiratory (diastolic and systolic blood pressure, forced expiratory volume) markers of allostatic load. The results showed significant pleiotropy, as measured by both LD score regression and polygenic profile analysis, between tiredness and most of the metabolic and anthropometric markers, though not the cardiovascular/respiratory markers. This is suggestive of the fact that the pleiotropy between tiredness and these physiological factors may be due to a biological propensity to an over-compensatory physiological stress response.

Neuroticism, the tendency to experience negative affective states, and tiredness were strongly correlated, both phenotypically and genetically. This may represent a separate route to fatigue, a predominately affective one, and/or it may overlap with the concept of allostatic load. A recent paper by Gale et al.⁴⁷, also using this UK Biobank sample, supports the allostatic load link with neuroticism, showed that polygenic profile scores for several physical and mental health traits - BMI, coronary artery disease, smoking status, bipolar disorder, borderline personality, major depressive disorder, negative affect and schizophrenia - significantly predicted neuroticism. In the present study, when tiredness polygenic profile analyses were adjusted for neuroticism, the associations between tiredness and mental health disorders (bar schizophrenia) were largely attenuated, whereas most of the metabolic and anthropometric associations remained significant. This suggests that it is the propensity to neuroticism, rather than the specific propensity to these disorders, that accounts or mediates

the tiredness associated with mood disorders. Watson and Pennebaker⁶⁸, discussing competing models of how negative affectivity is related to self-reported physical and emotional well-being, found that it is associated as much with the former as with the latter, and that negative affectivity might better be conceptualised as a general tendency to experience both somatic and emotional distress. This concept of a general tendency to what they termed somatopsychic distress could explain the pleiotropy observed in the present study between neuroticism and tiredness.

That neuroticism may be a distinct route to fatigue is supported by the fact that when the polygenic profile analysis is adjusted for self-rated health, all associations between polygenic profiles for physical health and tiredness are attenuated to the point of non-significance, whereas the relationship between tiredness, and polygenic profiles for neuroticism, MDD and bipolar disorder remain significant. This is consistent with the study of Gale et al.⁶⁹, investigating pleiotropy between neuroticism and physical and mental health, where there were more and stronger genetic associations between neuroticism and mental health than between neuroticism and physical health. If we take self-rated health to be a marker, to some extent, of actual physical health (and the study by Harris et al.⁷⁰ would indicate that it is), then this would suggest that when physical health is adjusted for, polygenic profiles for neuroticism and its associated negative affective states, continue to make a unique contribution to tiredness.

These proposed affective and physiological routes to fatigue may not be mutually exclusive. The allostatic load concept postulates that individual differences in personality, cognition and behavioural responses to stress, and socio-cultural factors, affect the physiological stress response. An increased propensity to experience distress, as captured in the concept of neuroticism, would imply that there is increasing propensity to over-respond to stressors and thus to physiological dysregulation. In the study by Gale et al.⁶⁹, the only significant pleiotropy between neuroticism and physical disease state was a significant association with the polygenic risk score for coronary artery disease, which is suggestive of an overlap of affective and cardiovascular stress responses. The allostatic load model also allows us to incorporate childhood cognitive ability (reduced ability to problem solve), smoking status (at least one study has found smoking and allostatic load interaction effects²⁹) and grip strength (reduced overall system integrity/vigour) into a more general model which is suggestive of shared genetic variance between stress proneness (neuroticism, reduced cognitive ability,

reduced vigour), the physiological response to stress (biomarkers), behavioural responses (smoking), self-reported tiredness, disease and mortality⁷¹.

The large sample size of the present study is a strength of this study, providing powerful and robust tests of pleiotropy between tiredness and physical and mental health. A second strength is that all genetic samples were processed on the same platform at the same location. The use of summary data from many international GWAS consortia provided foundations for a comprehensive examination of shared genetic aetiology between tiredness and a wide range of health-related phenotypes.

A limitation of the present study is the measure of tiredness. As Wessely⁷² observed, an objective measure of fatigue is “an unattainable holy grail”. Almost all the studies cited in the introduction have used subjective self-reports. The self-report measures used vary widely, with there being several validated fatigue measures, and many of the reported studies use either double or single item questionnaires and/or single item Visual Analogue Scales. This in itself may account for some of the inconsistency in fatigue research, though the demographic studies cited at the beginning of this article, using a wide variety of measures from single questions⁷ to a well validated fatigue questionnaire², produced similar findings. However our findings of genetic associations of the “wabbit” proxy for fatigue will need replication with better validated multi-item measures. All analyses were restricted to individuals of white British ancestry, because the sample does not have enough power to generalize results for individuals with different backgrounds.

Summary

Being genetically predisposed to a range of mental and physical health complaints also predisposes individuals to report that they are more tired or lacking in energy. This study confirms that self-reported tiredness is a partly-heritable, heterogeneous and complex phenomenon that is phenotypically and genetically associated with affective, cognitive, personality, health and physiological processes. This study also served as a first step in testing some genetic hypotheses from the allostatic load model, finding suggestive links between tiredness and three genes on chromosome one associated with allostatic processes and considerable pleiotropy between tiredness and allostatic markers. We can foresee more tests of these links as more genome-wide genotyping data become available.

Acknowledgments

This research has been conducted using the UK Biobank Resource. The work was undertaken in The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC, Age UK (Disconnected Mind Project), and Medical Research Council (MRC) is gratefully acknowledged.

Conflict of Interest

IJD is a participant in UK Biobank. None of the other authors have actual or potential conflicts of interest to declare.

References

1. van't Leven M, Zielhuis GA, van der Meer JW, Verbeek AL, Bleijenberg G. Fatigue and chronic fatigue syndrome-like complaints in the general population. *The European Journal of Public Health* 2010; **20**(3): 251-257.
2. Pawlikowska T, Chalder T, Hirsch S, Wallace P, Wright D, Wessely S. Population based study of fatigue and psychological distress. *Bmj* 1994; **308**(6931): 763-766.
3. Ricci JA, Chee E, Lorandeanu AL, Berger J. Fatigue in the US workforce: prevalence and implications for lost productive work time. *Journal of Occupational and Environmental Medicine* 2007; **49**(1): 1-10.
4. Meng H, Hale L, Friedberg F. Prevalence and predictors of fatigue among middle-aged and older adults: evidence from the Health and Retirement study. *Journal of the American Geriatrics Society* 2010; **58**(10): 2033.
5. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *Journal of epidemiology and community health* 1992; **46**(2): 92.
6. Cullen W, Kearney Y, Bury G. Prevalence of fatigue in general practice. *Irish journal of medical science* 2002; **171**(1): 10-12.
7. Dolan P, Kudrna L. More years, less yawns: fresh evidence on tiredness by age and other factors. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 2013; gbt118.
8. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J *et al.* Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001; **56**(3): M146-156.
9. Bandeen-Roche K, Xue Q-L, Ferrucci L, Walston J, Guralnik JM, Chaves P *et al.* Phenotype of frailty: characterization in the women's health and aging studies. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 2006; **61**(3): 262-266.
10. Theorell-Haglöw J, Lindberg E, Janson C. What are the important risk factors for daytime sleepiness and fatigue in women. *Sleep* 2006; **29**(6): 751-757.
11. Fritschi C, Quinn L. Fatigue in patients with diabetes: a review. *Journal of psychosomatic research* 2010; **69**(1): 33-41.
12. Jarosz PA, Davis JE, Yarandi HN, Farkas R, Feingold E, Shippings SH *et al.* Obesity in urban women: associations with sleep and sleepiness, fatigue and activity. *Women's Health Issues* 2014; **24**(4): e447-e454.
13. Bower JE. Cancer-related fatigue [mdash] mechanisms, risk factors, and treatments. *Nature Reviews Clinical Oncology* 2014; **11**(10): 597-609.
14. Griffiths L, Jones DE. Pathogenesis of primary biliary cirrhosis and its fatigue. *Digestive Diseases* 2014; **32**(5): 615-625.
15. Krupp LB, Alvarez LA, LaRocca NG, Scheinberg LC. Fatigue in multiple sclerosis. *Archives of neurology* 1988; **45**(4): 435-437.

16. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D *et al.* Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. *Arthritis Care & Research* 2005; **53**(5): 697-702.
17. Ng W-F, Bowman SJ. Primary Sjögren's syndrome: too dry and too tired. *Rheumatology* 2010; keq009.
18. Alves G, Wentzel-Larsen T, Larsen J. Is fatigue an independent and persistent symptom in patients with Parkinson disease? *Neurology* 2004; **63**(10): 1908-1911.
19. Franssen PM, Bültmann U, Kant I, van Amelsvoort LG. The association between chronic diseases and fatigue in the working population. *Journal of psychosomatic research* 2003; **54**(4): 339-344.
20. Demyttenaere K, De Fruyt J, Stahl SM. The many faces of fatigue in major depressive disorder. *The International Journal of Neuropsychopharmacology* 2005; **8**(01): 93-105.
21. Bültmann U, Kant I, Schröer C, Kasl S. The relationship between psychosocial work characteristics and fatigue and psychological distress. *International archives of occupational and environmental health* 2002; **75**(4): 259-266.
22. Fernández-Muñoz JJ, Morón-Verdasco A, Cigarán-Méndez M, Muñoz-Hellín E, Pérez-de-Heredia-Torres M, Fernández-de-las-Peñas C. Disability, quality of life, personality, cognitive and psychological variables associated with fatigue in patients with multiple sclerosis. *Acta Neurol Scand* 2015; **132**(2): 118-124.
23. Kitamura H, Shindo M, Tachibana A, Honma H, Someya T. Personality and resilience associated with perceived fatigue of local government employees responding to disasters. *J Occup Health* 2013; **55**(1): 1-5.
24. Barsevick A, Frost M, Zwinderman A, Hall P, Halyard M, Consortium G. I'm so tired: biological and genetic mechanisms of cancer-related fatigue. *Quality of Life Research* 2010; **19**(10): 1419-1427.
25. Felger JC, Cole SW, Pace TW, Hu F, Woolwine BJ, Doho GH *et al.* Molecular signatures of peripheral blood mononuclear cells during chronic interferon- α treatment: relationship with depression and fatigue. *Psychological medicine* 2012; **42**(08): 1591-1603.
26. Tomas C, Newton J, Watson S. A review of hypothalamic-pituitary-adrenal axis function in chronic fatigue syndrome. *ISRN neuroscience* 2013; **2013**.
27. Klimas NG, Broderick G, Fletcher MA. Biomarkers for chronic fatigue. *Brain, behavior, and immunity* 2012; **26**(8): 1202-1210.
28. Skerrett TN, Moss-Morris R. Fatigue and social impairment in multiple sclerosis: the role of patients' cognitive and behavioral responses to their symptoms. *Journal of psychosomatic research* 2006; **61**(5): 587-593.
29. Juster R-P, McEwen BS, Lupien SJ. Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews* 2010; **35**(1): 2-16.

30. Van Kessel K, Moss-Morris R. Understanding multiple sclerosis fatigue: a synthesis of biological and psychological factors. *Journal of psychosomatic research* 2006; **61**(5): 583-585.
31. Dixon JB. The effect of obesity on health outcomes. *Molecular and cellular endocrinology* 2010; **316**(2): 104-108.
32. Lasselin J, Layé S, Dexpert S, Aubert A, Gonzalez C, Gin H *et al.* Fatigue symptoms relate to systemic inflammation in patients with type 2 diabetes. *Brain, behavior, and immunity* 2012; **26**(8): 1211-1219.
33. Brown PJ, Roose SP, Fieo R, Liu X, Rantanen T, Sneed JR *et al.* Frailty and depression in older adults: a high-risk clinical population. *The American Journal of Geriatric Psychiatry* 2014; **22**(11): 1083-1095.
34. Landmark-Høyvik H, Reinertsen KV, Loge JH, Kristensen VN, Dumeaux V, Fosså SD *et al.* The Genetics and Epigenetics of Fatigue. *PM&R* 2010; **2**(5): 456-465.
35. Schur E, Afari N, Goldberg J, Buchwald D, Sullivan PF. Twin Analyses of Fatigue. *Twin Research and Human Genetics* 2007; **10**(05): 729-733.
36. Sullivan PF, Evengard B, Jacks A, Pedersen NL. Twin analyses of chronic fatigue in a Swedish national sample. *Psychol Med* 2005; **35**(9): 1327-1336.
37. Hur Y-M, Burri A, Spector TD. The genetic and environmental structure of the covariation among the symptoms of insomnia, fatigue, and depression in adult females. *Twin Research and Human Genetics* 2012; **15**(06): 720-726.
38. Schlauch K, Khaiboullina S, De Meirleir K, Rawat S, Petereit J, Rizvanov A *et al.* Genome-wide association analysis identifies genetic variations in subjects with myalgic encephalomyelitis/chronic fatigue syndrome. *Translational psychiatry* 2016; **6**(2): e730.
39. Smith AK, Fang H, Whistler T, Unger ER, Rajeevan MS. Convergent genomic studies identify association of GRIK2 and NPAS2 with chronic fatigue syndrome. *Neuropsychobiology* 2011; **64**(4): 183-194.
40. Landmark-Høyvik H, Reinertsen K, Loge J, Fosså S, Børresen-Dale A, Dumeaux V. Alterations of gene expression in blood cells associated with chronic fatigue in breast cancer survivors. *The pharmacogenomics journal* 2009; **9**(5): 333-340.
41. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J *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**(3): e1001779.
42. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 1999; **282**(18): 1737-1744.
43. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Artigas MS *et al.* Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *The Lancet Respiratory Medicine* 2015; **3**(10): 769-781.

44. Hagenaars SP, Harris SE, Davies G, Hill WD, Liewald DC, Ritchie SJ *et al.* Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. *Mol Psychiatry* 2016.
45. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**(7): 906-913.
46. Harris SE, Hagenaars SP, Davies G, Hill WD, Liewald DC, Ritchie SJ *et al.* Molecular genetic contributions to self-rated health. *bioRxiv* 2015: 029504.
47. Gale CR, Hagenaars SP, Davies G, Hill WD, Liewald DC, Cullen B *et al.* Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. *Translational Psychiatry* In Press.
48. Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE *et al.* Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N = 112 151). *Molecular Psychiatry* In press.
49. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**(7): 565-569.
50. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *11* 2015; **4**.
51. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics* 2015; **47**(11): 1228-1235.
52. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**(7256): 748-752.
53. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**(3): 291-295.
54. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; **47**(11): 1236-1241.
55. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics* 2015; **31**(9): 1466-1468.
56. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* 1995; **57**(1): 289-300.
57. Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans J *et al.* Prevalence and Characteristics of Probable Major Depression and Bipolar Disorder within UK Biobank: Cross-Sectional Study of 172,751 Participants. *PLoS ONE* 2013; **8**(11): e75362.

58. Qi Q, Menzaghi C, Smith S, Liang L, de Rekeneire N, Garcia ME *et al.* Genome-wide association analysis identifies TYW3/CRYZ and NDST4 loci associated with circulating resistin levels. *Hum Mol Genet* 2012; **21**(21): 4774-4780.
59. Sugimoto M, Watanabe T, Sugimoto Y. The Molecular Effects of a Polymorphism in the 5'UTR of Solute Carrier Family 44, Member 5 that Is Associated with Birth Weight in Holsteins. *PLoS ONE* 2012; **7**(7): e41267.
60. Clark GD. Platelet-Activating Factor Acetylhydrolase and Brain Development. *Enzymes* 2015; **38**: 37-42.
61. Hoenicka J, aragüés M, Ponce G, Rodríguez-Jiménez R, Jiménez-Arriero MA, Palomo T. From dopaminergic genes to psychiatric disorders. *Neurotoxicity Research*; **11**(1): 61-71.
62. de Miguel FJ, Sharma RD, Pajares MJ, Montuenga LM, Rubio A, Pio R. Identification of Alternative Splicing Events Regulated by the Oncogenic Factor SRSF1 in Lung Cancer. *Cancer Research* 2014; **74**(4): 1105-1115.
63. Sailer A, Houlden H. Recent Advances in the Genetics of Cerebellar Ataxias. *Current Neurology and Neuroscience Reports* 2012; **12**(3): 227-236.
64. Russell B, Graham JM. Expanding our knowledge of conditions associated with the ASXL gene family. *Genome Medicine* 2013; **5**(2): 1-3.
65. Alberti KG, Zimmet P, Shaw J, Group IETFC. The metabolic syndrome--a new worldwide definition. *Lancet* 2005; **366**(9491): 1059-1062.
66. Booth T, Starr JM, Deary I. Modeling multisystem biological risk in later life: allostatic load in the Lothian birth cohort study 1936. *American Journal of Human Biology* 2013; **25**(4): 538-543.
67. Seeman T, Gruenewald T, Karlamangla A, Sidney S, Liu K, McEwen B *et al.* Modeling multisystem biological risk in young adults: the Coronary Artery Risk Development in Young Adults Study. *American Journal of Human Biology* 2010; **22**(4): 463-472.
68. Watson D, Pennebaker JW. Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychological review* 1989; **96**(2): 234.
69. Gale C, Hagenaars SP, Davies G, Hill WD, Liewald DC, Cullen B *et al.* Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. *bioRxiv* 2015: 031138.
70. Harris SE, Hagenaars SP, Davies G, Hill WD, Liewald DC, Ritchie SJ *et al.* Molecular genetic contributions to self-rated health. *bioRxiv* 2015.
71. Deary IJ. Looking for 'system integrity' in cognitive epidemiology. *Gerontology* 2012; **58**(6): 545-553.
72. Wessely S. Chronic fatigue: symptom and syndrome. *Annals of Internal Medicine* 2001; **134**(9_Part_2): 838-843.

Figure legends

Figure 1. Flow diagram of participant selection.

Figure 2 (A) Manhattan and (B) Q-Q plot of P-values of the SNP-based association analysis of tiredness (responses to the question, “Over the past two weeks, how often have you felt tired or had little energy?”). The red line indicates the threshold for genome-wide significance ($P < 5 \times 10^{-8}$); the grey line indicates the threshold for suggestive significance ($P < 1 \times 10^{-5}$).

Figure 3. Barplot of genetic correlations (SE) calculated using LD regression between tiredness in UK Biobank and mental and physical health measures from GWAS consortia. *, $p < 0.0281$

Figure 1. Flow diagram of participant selection.

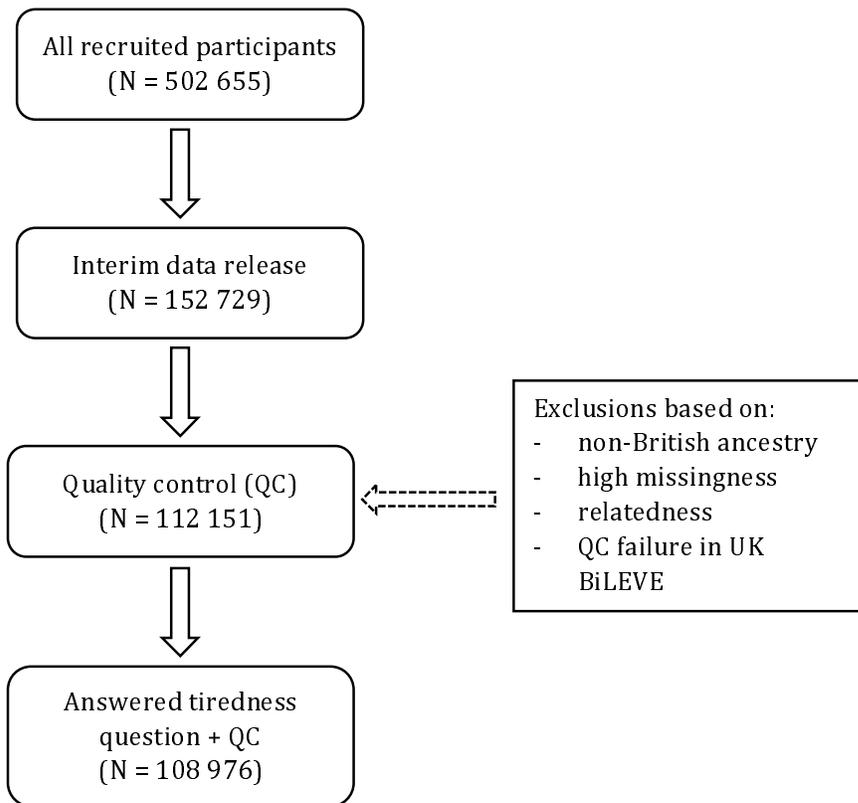


Table 1. Spearman phenotypic correlations between tiredness (responses to the question, “Over the past two weeks, how often have you felt tired or had little energy?”) and physical and mental health. All correlations had $P < 0.001$

	Tiredness
Self-rated health (N = 108 648)	-0.35
Grip strength (N = 108 573)	-0.12
Forced expiratory function in 1s (N = 101 823)	-0.08
Height (N = 108 796)	-0.09
BMI (N = 108 681)	0.10
Verbal-numerical reasoning (N = 35 101)	-0.04
Neuroticism (N = 105 456)	0.39

Figure 2 (A) Manhattan and (B) Q-Q plot of P-values of the SNP-based association analysis of tiredness (responses to the question, “Over the past two weeks, how often have you felt tired or had little energy?”).

The red line indicates the threshold for genome-wide significance ($P < 5 \times 10^{-8}$); the grey line indicates the threshold for suggestive significance ($P < 1 \times 10^{-5}$).

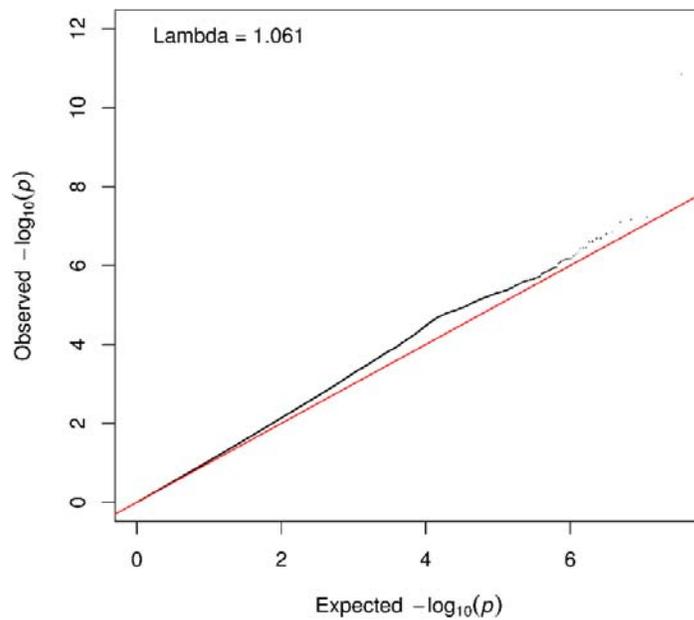
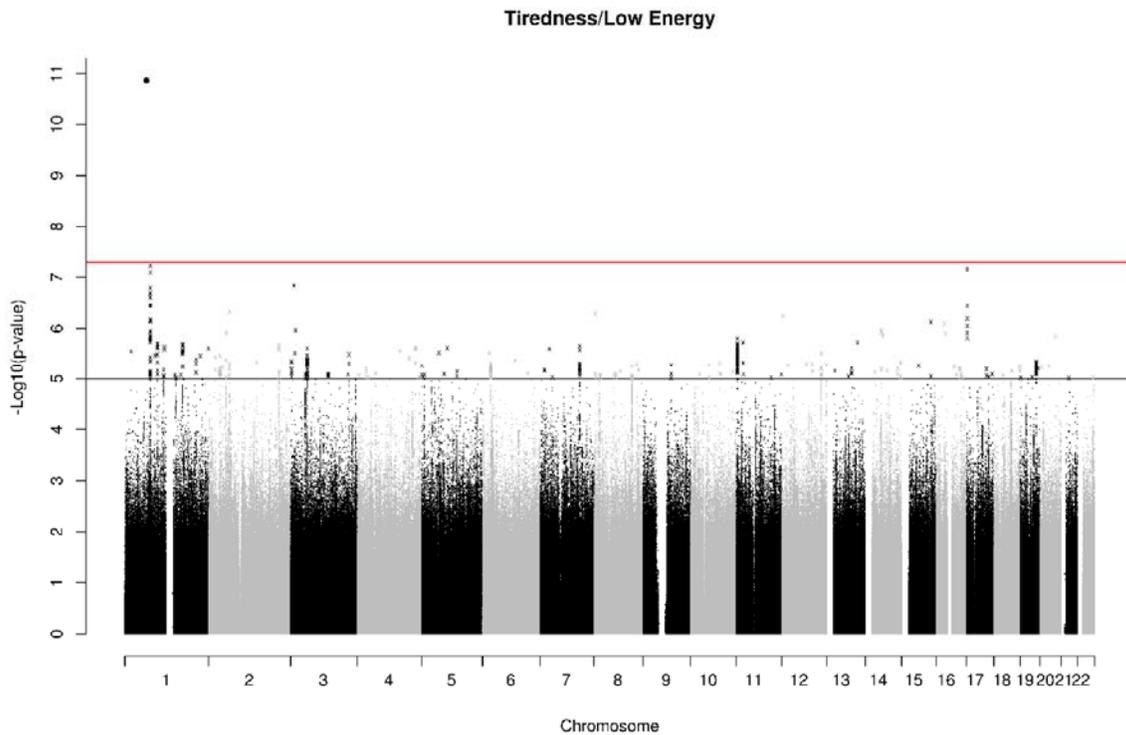


Table 2. Shows the genome wide significant genes from the UK Biobank tiredness phenotype and the significance values for the same genes using the neuroticism, self-rated health, and grip strength phenotypes, also in the UK Biobank sample. Unmodified P-values are shown for all phenotypes.

CHR	Gene	Tiredness P	SRH P	Grip P	Neuroticism P
11	<i>DRD2</i>	2.94×10^{-7}	0.012	0.156	9.69×10^{-9}
1	<i>PRRC2C</i>	1.43×10^{-6}	0.002	0.314	0.020
3	<i>C3orf84</i>	1.45×10^{-6}	7.38×10^{-5}	0.001	0.016
3	<i>ANO10</i>	1.52×10^{-6}	0.058	0.728	0.001
18	<i>ASXL3</i>	2.67×10^{-6}	1.03^{-6}	0.052	1.36×10^{-5}

Table 3. Genetic correlations between tiredness documented in the UK Biobank data set and the health-related variables collected from GWAS consortia.

Trait				
category	Traits from GWAS consortia	rg	SE	p
Physical	Blood Pressure: Diastolic	0.0332	0.0502	0.5083
Health	Blood Pressure: Systolic	-0.0698	0.0478	0.1444
	BMI	0.2024	0.0322	3.18E-10
	Cholesterol: HDL	-0.1087	0.0373	0.0036
	Cholesterol: LDL	0.0829	0.0413	0.0449
	Coronary Artery Disease	0.1338	0.067	0.0459
	Grip strength*	-0.1596	0.0482	0.0009
	HbA1c	0.2536	0.0857	0.0031
	Height	-0.0201	0.0297	0.498
	Longevity	-0.3943	0.1096	0.0003
	Forced Expiratory volume 1s*	-0.1181	0.0538	0.0281
	Obesity	0.2063	0.0381	6.31E-08
	Rheumatoid Arthritis	-0.0181	0.0674	0.7885
	Self-rated health*	-0.778	0.0349	7.30E-110
	Smoking Status	0.2009	0.0603	0.0009
	Triglycerides	0.1324	0.0332	6.62E-5
	Type 2 Diabetes	0.1784	0.0689	0.0097
Waist-hip ratio	0.2834	0.0417	1.09E-11	
Mental	ADHD	0.2694	0.1116	0.0158
Health	Alzheimer's Disease	0.0762	0.1079	0.4801
	Alzheimer's Disease (500kb)	0.0613	0.0872	0.4816
	Anorexia Nervosa	0.0192	0.0492	0.6967
	Autism	0.0129	0.0695	0.8522
	Bipolar Disorder	0.1382	0.0605	0.0223
	Childhood cognitive ability	-0.1528	0.0891	0.0864
	Major Depressive Disorder	0.5902	0.1015	6.03E-09
	Neuroticism	0.615	0.038	7.34E-59
	Schizophrenia	0.249	0.0386	1.14E-10
Verbal-numerical Reasoning*	-0.1379	0.0596	0.0206	

Statistically significant P-values (after false discovery rate correction; threshold: $P = 0.0281$) are shown in bold. *GWAS based on UK Biobank data.

Figure 3. Barplot of genetic correlations (SE) calculated using LD regression between tiredness in UK Biobank and mental and physical health measures from GWAS consortia. *, $p < 0.0281$

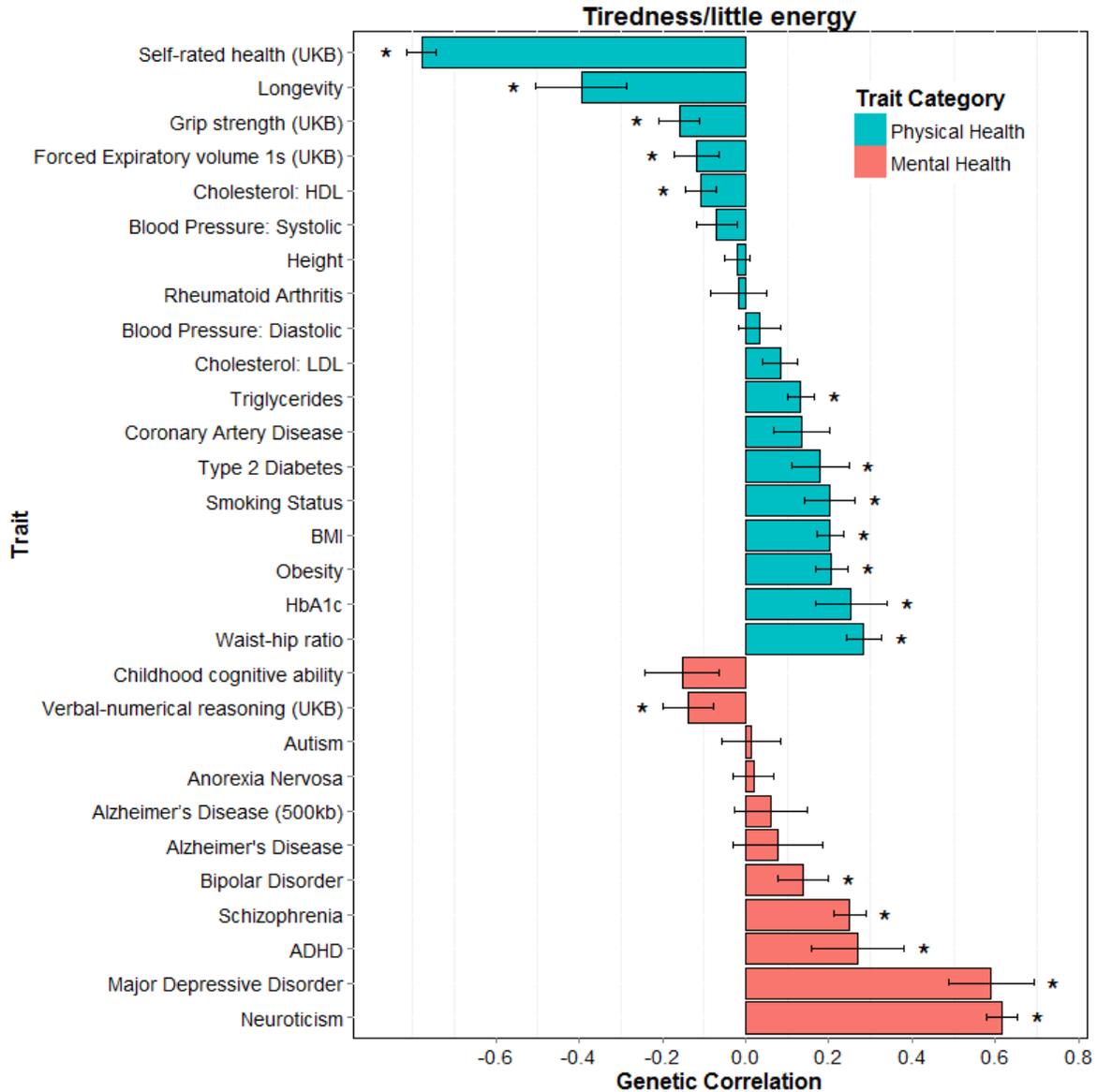


Table 4. Associations between polygenic profiles of health-related traits created from GWAS consortia summary data, and the UK Biobank tiredness phenotype controlling for age, sex, assessment centre, genotyping batch and array and 10 principal components for population structure

Trait Category	Trait	Threshold	β	P
Physical Health	Blood Pressure: Diastolic	0.1	-0.0028	0.3619
	Blood Pressure: Systolic	0.1	-0.0025	0.4077
	BMI	1	0.0280	4.90×10⁻²⁰†
	Cholesterol: HDL	0.5	-0.01634	8.49×10⁻⁸†
	Cholesterol: LDL	0.5	0.008147	0.0077†
	Coronary Artery Disease	0.5	0.0084	0.0061
	Forced Expiratory volume 1s	0.01	-0.0059	0.0529
	Longevity	0.05	-0.0067	0.0297
	HbA1c	1	0.008999	0.0033†
	Height	1	-0.0077	0.0154
	Obesity	1	0.0236	1.20×10⁻¹⁴†
	Rheumatoid Arthritis	0.1	-0.0016	0.5926
	Smoking Status	0.5	0.0091	0.0048
	Triglycerides	0.5	0.020854	1.06×10⁻¹¹†
	Type 2 Diabetes	1	0.0120	0.0001^{†a}
Waist-hip ratio	1	0.025791	7.85×10⁻¹⁷	
Mental Health	ADHD	1	0.0042	0.1647
	Alzheimer's Disease	0.05	-0.0052	0.0889
	Anorexia Nervosa	0.5	0.0048	0.1169
	Autism	1	-0.0018	0.5593
	Bipolar Disorder	0.01	0.0082	0.0076*
	Childhood cognitive ability	0.1	-0.0112	0.0002†
	Major Depressive Disorder	1	0.0183	2.00×10⁻⁹*
	Neuroticism	0.1	0.0185	2.25×10⁻⁹*
	Schizophrenia	1	0.0290	3.25×10⁻²⁰†*

FDR-corrected statistically significant values (P=0.0255) are shown in bold. The associations between the polygenic profile with the largest effect size (threshold) and tiredness are presented. Threshold is the P-value threshold with the largest effect size. †, results remain significant after controlling for neuroticism scores; *, results remain

significant after controlling for self-rated health; ^a, results remain significant after excluding individuals with type 2 diabetes ($\beta = 0.0105$, $p = 0.00076$).