Clinical phenotyping and biomarker discovery for neurological involvement in Wilson's disease

PhD thesis (September 2021)

Dr Samuel Shribman MA (Cantab) MBBS MRCP (UK)
Reta Lila Weston Institute of Neurological Studies
UCL Queen Square Institute of Neurology

Primary supervisor: Professor Thomas Warner
Word count (excluding references and appendix): 54,455

Declaration

I, Samuel Shribman, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Wilson's disease is an autosomal-recessive disorder of copper metabolism with neurological and hepatic presentations. Chelation therapy is used to 'de-copper' patients and monitoring is based on copper indices, such as serum non-caeruloplasmin-bound copper and 24-hour urinary copper output, which do not reflect neurological involvement. Biomarkers of end-organ damage, i.e. measures of the pathophysiological process in the brain, are needed to improve outcomes and prepare for clinical trials of novel therapies. In this thesis, I examine the pattern of movement disorders, cognitive deficits and psychiatric features in 40 prospectively-recruited patients before describing work identifying wet (fluid) and imaging biomarkers for neurological involvement using cross-sectional data from this cohort.

The findings demonstrate that movement disorders are associated with cognitive deficits but not psychiatric features in chronically-treated patients and that some patients with hepatic presentations have subtle deficits in cognitive flexibility, associative learning and recognition memory for faces. I confirm previous observations that copper indices do not correlate with the severity of movement disorders and show that plasma neurofilament light is a promising monitoring biomarker. I also report preliminary data identifying serum proteins that might reflect neurological involvement using label-free proteomics.

Using a combination of whole-brain, quantitative MRI analyses, I provide evidence that grey matter volumes in subcortical regions may serve as prognostic biomarkers and diffusivity in white matter tracts may serve as monitoring biomarkers. I identify neuroimaging correlates for serum non-caeruloplasmin-bound ('free') copper and demonstrate that increasing neurological severity is associated with widespread cortical iron deposition. I also determine the neuroanatomical bases for cognitive deficits and psychiatric features in Wilson's disease, some of which relate to specific patterns of cortical volume loss or white matter pathology.

These data support the idea that there is a spectrum of neurological involvement in Wilson's disease, which I suggest can be modelled as a brain injury. I discuss next steps for validating these biomarkers and conclude by

proposing a unifying theory for the evolution of neurological involvement in Wilson's disease.

Impact statement

Novel approaches to predicting neurological outcomes and monitoring neurological involvement in Wilson's disease are described in this thesis. These require validation in independent, longitudinal cohorts but may have important implications for clinical practice and clinical trial design in the future.

I anticipate that prognostic or predictive biomarkers, such as measures of subcortical volumes, could be used to stratify patients in clinical trials or guide treatment decisions when initiating chelation therapy. Monitoring biomarkers, such as plasma neurofilament light or measures of white matter diffusivity, could be used as trial endpoints. In a clinical setting, they might indicate when patients are at risk of neurological worsening or when neurological disease activity has been suppressed and doses should be reduced in order to minimise adverse effects.

Using biomarkers for neurological involvement to optimise treatment decisions and dosing regimens may have significant financial consequences for health systems. Although Wilson's disease is rare, neurological involvement can cause significant morbidity and may be associated with substantial health and social care costs. As an example, trientine, the second most widely used chelating agent in the UK, costs between £33,671 to £101,014 per annum and is often introduced after patients taking penicillamine deteriorate neurologically. The majority of patients start lifelong treatment in early adulthood.

I identified cognitive deficits and psychiatric features in patients with hepatic presentations with implications for the classification of Wilson's disease and provision of clinical services. These findings suggest that a binary classification is an oversimplification and these symptoms may be underappreciated in patients otherwise assumed to have isolated liver disease. This highlights the importance of multi-disciplinary care with access to neurology, neuropsychology, neuropsychiatry and vocational rehabilitation services for all patients with Wilson's disease. In addition, the data suggest that neurological involvement in Wilson's disease can be modelled as a brain injury. Patients are often advised their condition is treatable and symptoms should resolve. Re-framing the disease course in this way may help patients

and their clinicians better conceptualise anticipated outcomes and the potential for recovery.

Finally, the observation that increasing neurological severity is associated with widespread cortical iron deposition and my theory on the evolution of neurological involvement has implications for basic science research on Wilson's disease. They illustrate the importance of exploring brain iron metabolism in animal and cell-based models of the disease. The demonstration that cognitive deficits and psychiatric features are associated with specific patterns of cortical pathology highlights the need for whole-brain analyses in further neuroimaging studies.

I have already begun to realise the potential impact of the work presented in thesis. I am continuing to disseminate our findings in academic journals, at medical conferences, through Wilson's Disease Support Group UK newsletters and meetings and through the British Association for the Study of the Liver Special Interest Group on rare diseases.

Table of Contents

List of tables

List of figures

Abbreviations

ACN	Acetonitrile
AD	Axial diffusivity
ADC	Apparent diffusion coefficient
ALS	Amyotrophic lateral sclerosis
AUC	Area under the curve
CPAL	Camden paired associate learning
CROWD	Cohort research on Wilson's disease
CSF	Cerebrospinal fluid
DKEFSI	Delis-Kaplan Executive Function System – Interference subtest
DSB	Digit span backwards
DSym	Digit symbol
DTE	Dithioerithritol
DTI	Diffusion tensor imaging
DWI	Diffusion-weighted imaging
EASL	European Association for the Study of the Liver
EDSS	Expanded Disability Status Scale
ELI	Electrospray ionisation
ELISA	Enzyme-linked immunosorbent assay
EXC	Exchangeable copper
FA	Fractional anisotropy
FAS	Phonemic fluency for words beginning with F, A and S
FDR	False detection rate
FLAIR	Fluid attenuated inversion recovery
FSL	Functional MRI of the Brain software library
FTD	Fronto-temporal dementia
FWE	Family-wise error
GAD7	Generalised Anxiety Disorder 7 questionnaire
GDA	Graded difficulty arithmetic
GENFI	Genetic Frontotemporal Dementia Initiative
GFAP	Glial fibrillary acidic protein
GIF	Geodesic information flow
GM	Grey matter

GNT	Graded naming test
HD	Huntington's disease
IAA	Iodoacetic acid
KFR	Kayser-Fleischer rings
LBD	Lewy body disease
LC-MS	Liquid chromatography mass spectrometry
MALDI	Matrix-assisted laser desorption/ionisation
MD	Mean diffusivity
MMSE	Mini-mental state examination
MNI	Montreal Neurological Institute
ModNPI	Modified Neuropsychiatric Inventory
MRT	Matrix reasoning test
MS	Multiple sclerosis
MS/MS	Tandem mass spectrometry
mYS	Modified Young scale
m/z	mass-to-charge ratio
NART	National Adult Reading Test
NCC	Non-caeruloplasmin-bound copper
NfL	Neurofilament light
NHNN	National Hospital for Neurology and Neurosurgery
NODDI	Neurite orientation dispersion and density imaging
NPI	Neuropsychiatric Inventory
N4PA	Neurology 4-Plex A
ODI	Orientation dispersion index
PD	Parkinson's disease
PHQ9	Patient Health Questionnaire on depression 9
RD	Radial diffusivity
RMTF	Recognition memory test for faces
RMTW	Recognition memory test for words
ROC	Receiver operating characteristics
ROI	Region-of-interest
QSM	Quantitative susceptibility mapping
SF36	36-item short form survey on quality of life
Simoa	Single molecule array

SPM	Statistical parametric mapping
SWI	Susceptibility-weighted imaging
TBSS	Tract-based spatial statistics
TE	Echo time
TMTA	Trail making test – Part A
TMTB	Trail making test – Part B
TOF	Time-of-flight
TR	Repetition time
UWDRS	Unified Wilson's disease rating scale
UWDRS-D	Disability subscore
UWDRS-N	Neurological examination subscore
UWDRS-P	Psychiatric subscore
UCH-L1	Ubiquitin carboxyl-terminal hydrolase isozyme L1
UCu	Urinary copper output
VBM	Voxel-based morphometry
Vic	Intracellular volume fraction
VOSPFL	Visual Object and Space Perceptual battery – fragmented letters
VOSPNL	Visual Object and Space Perceptual battery – number location
WAIS	Weschler Adult Intelligence Scale
WASI	Weschler Abbreviated Scale of Intelligence
WD	Wilson's disease
WDSG-UK	Wilson's Disease Support Group UK
WM	White matter

1. Introduction

Background

Wilson's disease (WD) in an autosomal-recessive disorder of copper metabolism with an estimated disease prevalence of 1 per 100,00 in Northern European populations.[1,2] Originally referring to the condition as *progressive lenticular degeneration*, Samuel Alexander Kinnier Wilson described six patients with involuntary movements who were found to have liver cirrhosis and bilateral softening of the putamen at post-mortem in 1912.[3] He commented that the disease was 'invariably fatal'. The observation that these patients accumulate copper in the liver and brain was made in 1948,[4] and copper chelating agents were introduced in 1951.[5] *ATP7B* was identified as the causative gene by three independent groups in 1993.[6-8]

We now recognise that WD has a variable clinical phenotype and can present with movement disorders, psychiatric symptoms, liver disease or any combination of these in childhood, adolescence or adulthood.[9] Chelation therapy is the mainstay of treatment and more than 90% of patients survive.[10] However, outcomes for patients with neurological and/or psychiatric symptoms remain unpredictable and around half have persisting symptoms or disability.[11,12]

Genetics and pathophysiology

ATP7B is a copper-transporting ATPase with six metal-binding and eight transmembrane domains that form a pore for ATP-dependent copper transport across membranes.[13] It is abundantly expressed in hepatocytes where it supplies copper to the trans-Golgi network for incorporation into caeruloplasmin, the main copper-carrying protein secreted into the circulation. As intracellular copper concentrations increase ATP7B translocates to late endosomal compartments and sequesters copper into vesicles for export across the apical membrane into bile.[14] Loss of function therefore leads to an inability to excrete copper in bile with subsequent accumulation in the liver and other organs, including the brain.

Over 650 pathogenic or likely pathogenic mutations in ATP7B have been described.[15] The majority are missense, nonsense or truncating

mutations but additional mechanisms including exon skipping,[16] large deletions,[17] and intronic variants,[18] have recently been reported. The most common mutation in Northern European populations is c.3207C>A, p.(His1069Gln) with an allele frequency of 19% among patients from the UK.[19] This variant leads to retention of ATP7B in the endoplasmic reticulum where it undergoes more rapid degradation.[20] Other mutations lead to ATP7B dysfunction through alterations in protein expression and subcellular localisation, copper transport or phosphorylation activity.[21]

The genetic basis for phenotypic variation in WD remains unclear. In a cohort of 1,172 European patients there was no correlation between the presence of neurological involvement and *ATP7B* genotype.[22] Attempts to identify genetic modifiers have also been unsuccessful. *APOE* genotype has been shown to influence age of onset but not phenotype.[23] A whole-exome sequencing study that focussed on 46 proteins in the ATP7B interactome found that variants in *ESD* and *INO80* were weakly associated with phenotype, although these were rare.[24]

Neurodegeneration

Several post-mortem studies have confirmed that patients with WD have markedly increased copper in the basal ganglia and cortex.[25-27] Histopathological changes in the brain correlate with cerebral copper content,[28] and include neuronal loss, spongiosis, cavitation, demyelination and astrogliosis.[3,29,30] These can be seen in the basal ganglia, thalamus, brainstem and cerebral white matter but are usually most prominent in the putamen.

The precise mechanisms through which copper toxicity leads to these neuropathological changes are unclear, in part because animal models of WD do not produce a convincing neurological phenotype.[31] However, mitochondrial dysfunction, common in other neurodegenerative diseases of the basal ganglia, appears to play a crucial role.[32] The prevailing view up until recently was that this copper-induced mitochondrial dysfunction was mediated by the generation of reactive oxygen and nitrogen species in Fenton-like reactions.[33] Copper has now been shown to have more direct effect on mitochondrial function in *ATP7B* knockout rats, independent of redox activity.[34,35]

The relationship between systemic copper balance, liver disease and neurodegeneration is complicated. Advanced liver disease is not a prerequisite to developing neurological involvement: Only half of patients with neurological presentations have clinical or radiological features of cirrhosis.[36] Moreover, in a cohort of 131 patients who underwent liver biopsy at diagnosis, Ferenci *et al* found that 8 of 34 patients with neurological presentations had normal histology or steatosis.[37] Serum copper concentrations do not correlate with neurological symptom severity,[38] and some patients without neurological or psychiatric symptoms have strikingly increased brain copper at post-mortem.[26,39] It therefore appears that a disease model where copper accumulates in the liver, overflows in to the systemic circulation when the liver is damaged and causes neurological symptoms when it accumulates in the brain is an oversimplification. Some patients appear to be more vulnerable to the toxic effects of copper deposited in the brain than others.

Clinical presentations and classifications

The majority of patients present between the age of 3 and 40 years with a subacute or insidious onset of liver disease, a movement disorder, psychiatric symptoms or a combination of these.[9] Neurological features are seen in 39-75% of cases at diagnosis,[22,40] and patients with neurological involvement have higher mean age of onset (23.6 vs 17.6 years) and are more likely to be male (56 vs 44%).[22]

The clinical phenotype, discussed in more detail in Chapter 3, is highly variable.[40,41] Movement disorders including tremor, dystonia, parkinsonism, ataxia or chorea can occur and are usually associated with dysarthria.[38,42] Cognition is often affected; mild executive dysfunction is common but impaired processing speed, memory, visuospatial function and social cognition are also reported.[43-45] Psychiatric features such as personality and behavioural changes, depression, anxiety, hypomania and psychosis, may be seen.[46] Liver disease ranges from asymptomatic steatosis, deranged liver function or cirrhosis to decompensated (symptomatic) cirrhosis or acute liver failure and can occur without overt neurological features, or vice versa.[47] Kayser-Fleischer (KF) rings, copper deposits within Descemet's membrane, occur in the majority of patients with neurological symptoms.[48]

Several classifications of WD have been proposed. The current international consensus refers to *Neurological*, *Hepatic* and *Other* presentations where the presence of any neurological or psychiatric features indicates a *Neurological* presentation, irrespective of underlying liver disease.[49] Any clinical, biochemical or radiological features of liver disease in the absence of neurological or psychiatric features denote a *Hepatic* presentation. Patients identified through family screening, sometimes referred to as 'pre-symptomatic', often have biochemical or radiological features of liver disease, if not undiagnosed neurological or psychiatric symptoms, and may be classified as *Neurological* or *Hepatic* presentations.[50] *Other* presentations are therefore relatively rare. For the purpose of this thesis, I avoid the term 'neuropsychiatric' and refer to movement disorders, cognitive deficits and psychiatric features, all of which might constitute a neurological presentation, separately.

Unified Wilson's Disease Rating Scale

The Unified Wilson's Disease Rating Scale (UWDRS) was proposed in 2008 and is divided into neurological, psychiatric and hepatic sections.[51] The neurological section includes an 18-item neurological examination (UWDRS-N) score, sometimes referred to as *Part III*, that can be used to measure the severity of movement disorders in WD. It is based on components of the Unified Parkinson's Disease Rating Scale Part III, Burke-Fahn-Marsden Dystonia Rating Scale, an ataxia rating scale and the Unified Huntington's Disease Rating Scale, in addition to a subscore for speech disturbance (see **Table 0-2** in the appendix). It has been validated in several patient populations,[52,53] and has become the gold standard for measuring neurological involvement clinically in observational and interventional research on WD. The psychiatric (UWDRS-P) score consists of 19 interview and examination-based items from the Bech-Rafalsen Mania and Melancholia Scales and the Comprehensive Psychopathology Rating Scale. Subjective items for *failing memory* and *concentration difficulties* are used to assess cognition. The UWDRS-P score is seldom reported in research studies and psychiatric features are often underrepresented in translational research on WD.

Diagnosis

WD often mimics more common conditions and the diagnosis is frequently delayed, particularly for patients with neurological presentations.[12,54] The mean interval between symptom onset and diagnosis is 21-44 months.[2,12] No single test can confirm or exclude WD and the diagnosis is usually based on a combination of investigations, discussed below. An international panel of experts proposed a diagnostic scoring system for WD, often referred to as the Leipzig criteria, in 2001.[49] This incorporates clinical, biochemical, histopathological and genetic findings to determine if a diagnosis of WD is highly likely, probable or unlikely and is widely used in clinical research (see **Table 0-1** in the appendix).

A low serum caeruloplasmin (<0.20 g/L) is seen in 90% of patients with neurological presentations but is relatively non-specific.[22,55] A high 24-hour urinary copper output (UCu, >0.64 μmol/L) and the presence of KF rings on slit lamp examination are seen in 50-80% and 90% of these patients, respectively.[55,56] These tests are more specific but inconvenient. Genetic testing with Sanger sequencing of *ATP7B* detects two mutations in 98% of patients with WD but takes several months.[19] Hepatic parenchymal copper quantification is rarely required and is normal (<250 μg/g dry weight) in 12% of patients with neurological presentations.[37] Neuroimaging abnormalities including hyperintense signal abnormality in the basal ganglia, thalami and/or brainstem on T2-weighted or fluid attenuated inversion recovery (FLAIR) sequences are present in 90% of patients with neurological or psychiatric symptoms and are helpful diagnostically.[57] Cerebral, brainstem and/or cerebellar volume loss is also common.[58]

Chelation therapy and monitoring

Chelating agents, such as penicillamine or trientine, are used to ‘de-copper’ patients by mobilising intracellular copper into the circulation and increasing the urinary excretion of copper.[59] Higher doses are often used in the first few years when the aim is to arrest the disease process and reduce symptoms. Lower doses are continued lifelong as maintenance therapy to prevent the re-accumulation of copper while minimising adverse effects. Zinc salts, which

reduce the intestinal absorption of copper, are considered third-line in the UK given reports that liver disease can progress despite monotherapy.[60]

In the absence of clinical trials, dosing regimens for chelating agents vary between clinicians. European Association for the Study of the Liver (EASL) guidelines recommend initial doses anywhere between 750-1500 mg/day for penicillamine and 900-2700 mg/day for trientine and advice on dose titration is not provided.[59] Recommendations for monitoring include a neurological examination with blood tests every six months and a UCu collected on medication and, separately, after two days of treatment cessation, every year based on level C evidence. The non-caeruloplasmin-bound copper (NCC), calculated using the serum copper and serum caeruloplasmin, is also suggested to be a useful parameter. Treatment targets for UCu in the de-coppering phase, during the initial cupriuresis, are not provided in the guidelines.[59,61] In the maintenance phase, an output while on medication 'in the vicinity' of 3-8 µmol/day or after two days of treatment cessation of less than 1.6 µmol/day are considered adequate based on expert opinion.

Adverse effects with chelating agents are relatively common and include acute or chronic immune-mediated reactions, bone marrow suppression, nephrotoxicity and dermatological complications, which may be dose-dependent.[59,62] Paradoxical neurological worsening in the first six months of treatment is also a major concern. It is reported in 23% of patients with neurological presentations and is irreversible in 33% of these.[62,63] It is associated with increasing neurological severity at presentation and thalamic/brainstem lesions on MRI and can occur with chelating agents and zinc salts. The pathophysiological basis remains unclear. A widely accepted theory is that rapid mobilisation of copper stores leads to a transient increase in copper in the brain - many clinicians therefore gradually introduce chelating agents.[64]

Prognosis

Despite being considered 'treatable', neurological outcomes in WD are unpredictable. Of 137 patients with neurological presentations from the UK, Walshe reported that 42% became symptom free, 26% had residual neurological deficits without significant disability, 18% had significant

neurological disability and 8% had no response to treatment and died.[62] The author noted that 'clinically, biochemically and in their response to cupriuresis, the group that died was indistinguishable from those that responded to treatment'. In a similar study from Germany, 20 of the 105 patients (19%) with hepatic presentations later developed progressive neurological involvement.[12]

Novel treatments

The role of tetrathiomolybdate, a chelating agent that also increases biliary excretion and prevents intestinal absorption of copper, is currently being re-evaluated. In a recent open-label phase II trial, bis-choline tetrathiomolybdate led to 16 of 28 patients achieving or maintaining target NCC concentrations over 24 weeks.[65] The authors suggest that this drug may have a more favourable safety profile, with less neurological complications. However, neurological disease was measured with the UWDRS alone and serious adverse events included psychiatric disorders (four patients), gait disturbance (one patient) and 'declining neurological function' (one patient). A phase III trial comparing bis-choline tetrathiomolybdate to standard care (NCT03403205) is ongoing.

Liver transplantation, usually reserved for patients with acute liver failure of decompensated liver disease, was recently evaluated in 18 patients with neurological worsening on chelation therapy.[66] These patients had severe neurological disability, with a median modified Rankin scale (mRS) score 5 and median UWDRS score 105, at baseline. At last follow up, these had decreased to median mRS score 1.5 and median UWDRS score 36. Eight patients had major improvement, four patients had moderate improvement, two were stable and four died. The mechanism by which liver transplantation can ameliorate neurological involvement is currently unclear, as is the role for this treatment in clinical practice.

Finally, gene therapy using an adeno-associated viral (AAV) vector has shown encouraging results in animal models. Murillo *et al* transduced the liver of *ATP7B* knockout mice with AAV8 encoding *ATP7B* cDNA under the control of a liver specific promoter. They found a dose-dependent restoration of biliary copper excretion with normalisation of urinary copper output, hepatic copper content and liver function.[67] A subsequent study has shown that AAV-based

therapy reduces cerebral copper content in knockout mice but how this might translate into neurological improvement clinically is unclear given these mice do not have a neurological phenotype.[68] An open-label phase I/II trial is due to start this year (NCT04537377), although no neurological outcome measures are included in the trial design.

Improving neurological outcomes

While there is optimism that novel treatments might improve outcomes, our means of testing their safety and effectiveness from the neurological perspective remain limited. Copper indices, such as NCC and UCu, are currently used to guide chelation therapy and NCC has been used as a primary endpoint in clinical trials. However, these measures but do not differ between patients with neurological and hepatic presentations or correlate with neurological severity.[38,69,70] They are also highly variable between individuals;[71] a given NCC may be well tolerated by some patients but associated with disease progression in others.

Using bedside neurological examination to guide treatment decisions or as a primary endpoint is also problematic. UWDRS-N scores may not improve (or deteriorate) until months or years after initiating a treatment or altering doses. In a clinical setting, the consequences of under-treating a patient may not become apparent until after a significant delay or at all if the patient is assumed to have had a poor response to treatment. Prolonged follow up periods are needed for clinical trials that use UWDRS-N scores as an outcome measure to determine neurological response to treatment.

End-organ biomarkers that directly measure the pathophysiological process in the brain in WD are needed. Monitoring biomarkers, which measure changes in neurological disease status in response to treatment, could be used to test efficacy early in the de-coppering phase of treatment and promptly identify whether dose changes are helpful or harmful for brain health in newly-diagnosed or chronically-treated patients. Predictive biomarkers, which indicate whether an individual is likely to respond well to a given treatment, could be used to guide choice of treatment and for enrichment strategies in the design and conduct of clinical trials.[72]

Biomarker discovery

There are a number of candidate biomarkers for neurological involvement, and potential approaches to identifying biomarkers, that have yet to be explored in WD. These can be divided into wet (fluid) biomarkers and imaging biomarkers and are discussed in more detail in Chapters 4 and 5, respectively.

Briefly, *exchangeable copper*, a novel copper assay which reflects the labile fraction of serum copper bound to albumin and other peptides, was found to correlate with extrahepatic manifestations of WD but these findings have not yet been replicated.[73] Proteins derived from neurons in the central nervous system, such as neurofilament light, can be detected at very low concentrations in serum or plasma samples using Single Molecule Array technology. These have been validated as non-specific markers of neurodegeneration in several diseases but have not been tested in WD.[74] Label-free peptide quantification using tandem mass spectrometry is another powerful technique for biomarker discovery that uses an unbiased approach and has not been applied in this context.[75]

Several imaging biomarkers for neurological involvement in WD have been proposed. Regional volume loss on T1-weighted imaging, hyperintense signal abnormalities on T2-weighted or FLAIR sequences and diffusion- and susceptibility-weighted imaging abnormalities are common in neurological presentations of WD.[57,58,76,77] As discussed in Chapter 5, the majority of previous studies used qualitative or semi-quantitative approaches for analysing these neuroimaging abnormalities, usually focussing on specific regions of interest. Where quantitative MRI analyses were performed, cohort stratification was limited and associations with clinical rating scales, such as UWDRS-N, and copper indices were not tested. The neuroanatomical basis for cognitive deficits and psychiatric features, which are common in WD, is also poorly understood.

Aims and objectives

The Cohort Research on Wilson's Disease (CROWD) study was set up to comprehensively characterise the clinical features in a cohort of patients with WD and then identify wet (fluid) and imaging biomarkers for neurological involvement that could be used for monitoring disease activity in the brain and

predicting neurological outcomes. Secondary aims were to use these biomarkers to derive a clearer understanding for the pathophysiological and neuroanatomical bases for movement disorders, cognitive deficits and psychiatric features in WD.

Specific objectives are discussed in relevant chapters and include determining the relationship between movement disorders, cognitive deficits, psychiatric features in WD, testing the association between clinical features and a range of copper indices (NCC, UCu and exchangeable copper) and neuronal proteins measured using Single Molecule Array technology, identifying novel wet biomarkers for neurological involvement using tandem mass spectrometry and determining the clinical significance of brain volume loss, WMHs, diffusion-weighted and susceptibility weighted imaging abnormalities in WD using quantitative, whole-brain MRI analyses. The study protocol and design are discussed in the following chapter. The clinical phenotype in our cohort and then wet and imaging biomarker studies are covered in subsequent chapters.

2. CROWD study protocol

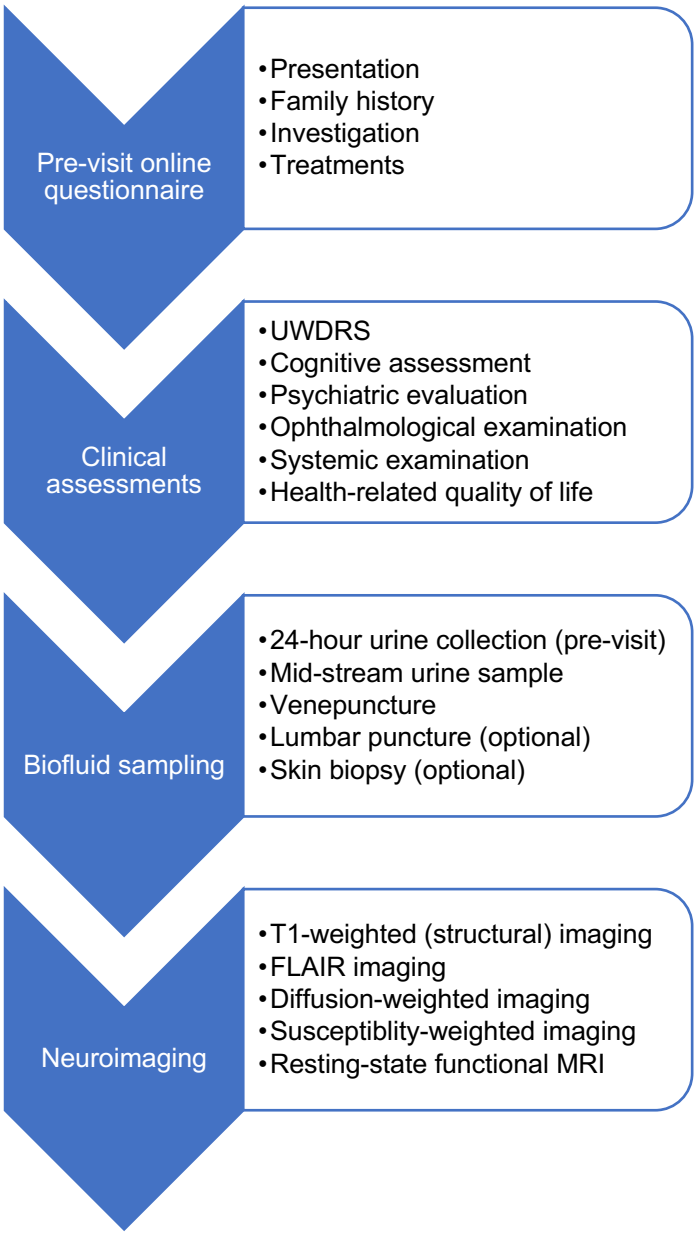
2.1 Study design

The CROWD study is an observational study of 40 patients with WD based at the UCL Queen Square Institute of Neurology and National Hospital for Neurology and Neurosurgery (NHNN) that was launched in 2018. The study is led by Professor Thomas Warner (Reta Lila Weston Institute of Neurological Studies, UCL) and supported by Dr Jonathan Rohrer (Dementia Research Centre, UCL), Professor Henrik Zetterberg (Dementia Research Institute, UCL), Dr Godfrey Gillett (NHNN and Department of Clinical Biochemistry, Northern General Hospital), Professor Emmanouil Tsochatzis (UCL Institute for Liver and Digestive Health), Mr Fion Bremner (Neuro-Ophthalmology Department, NHNN) and Professor Oliver Bandmann (Sheffield Institute of Translational Neuroscience).

Participants attended a baseline research visit for clinical assessments, biofluid sampling and neuroimaging at NHNN between January and December 2019. Research activities during visits are summarised in **Figure 2-1** and discussed in more detail below. Follow up research visits were originally planned for 12-18 months later but were not performed due to the COVID-19 pandemic. The data presented in this thesis is therefore cross-sectional data from the baseline visits.

Participants were recruited through neurology, hepatology and metabolic clinics at the NHNN and Royal Free Hospital, in addition to the Wilson's Disease Support Group UK (WDSG-UK) research register. Consecutive patients attending clinics were offered information about participating in the study. Members of the WDSG-UK research register were sent an invitation letter via post in January 2019. Patients with a confirmed diagnosis according to the Leipzig criteria (see Appendix),[49] who were aged 16 years or over and lived in the UK were eligible. Exclusion criteria were other medical or psychiatric illnesses that would interfere with completing assessments and pregnancy. I anticipated recruiting 30 patients with neurological presentations and 10 patients with hepatic presentations based on clinical caseloads in the respective clinics.

Figure 2-1 Overview of research activities



Flow diagram summarising research activities performed before and during baseline research visits. FLAIR, fluid attenuated inversion recovery; UWDRS, unified Wilson's disease rating scale.

There were no previous wet (fluid) biomarkers studies and few quantitative MRI studies to inform sample size calculations when the study was designed. The recruitment target was therefore based on a combination of samples sizes in successful pilot studies in other neurodegenerative diseases,[78-80] and previous imaging cohorts for WD,[77,81,82] in addition to the estimated number of patients attending local clinics who would be eligible to participate.

The study was funded by an Association of British Neurologists Clinical Research Training Fellowship supported by the Guarantors of Brain. Additional funding was provided by the Reta Lila Weston Trust and Wilson's Disease Support Group UK, who were also consulted on the study design and helped devise the participant information sheet through a patient focus group. The study was approved by the Newcastle and North Tyneside 2 Research ethics committee (18/NE/0279) in December 2018. All participants provided written consent.

2.2 Clinical assessments

Participants were asked to complete an online questionnaire focussing on their presentation, family history, investigations and treatment prior to their baseline research visit. These details were subsequently confirmed by interview or review of the medical records, where necessary. Neurological and psychiatric features at presentation were documented. Liver status was determined by review of previous investigations (ultrasonography, transient elastography and histopathology) and a combination of clinical assessment (systemic examination and UWDRS hepatic scale) and haematological and biochemical investigations (full blood count, coagulation profile and liver function tests) performed during research visits. The presence or absence of cirrhosis and, separately, features of decompensated liver disease (ascites, bleeding, encephalopathy and jaundice) were recorded.

Patients were classified into *Neurological*, *Hepatic* or *Other* presentations based on their clinical features at initial presentation as described in the international consensus classification of WD. Patients were also subcategorised according to their recent neurological status. I defined those patients with neurological presentations, or neurological deterioration associated with non-adherence, in the preceding six months as having active, as opposed to stable, disease.

The Unified Wilson's Disease Rating Scale (UWDRS) was performed and UWDRS-N subscores were recorded. A 16-item neuropsychological battery designed to test abstract reasoning, language, memory, executive function, processing speed, calculation, visuoperceptual, visuospatial and social cognition was performed. The battery was designed in collaboration with an expert in cognitive neurology (JDR). Tests where performance depends on motor function were avoided where possible. Testing took between 60 and 90 minutes and all participants took at least one break. Results were converted into Z scores using normative data from manuals or other published data. The cognitive domain, abbreviation and source of normative data for each test in the battery are outlined in **Table 2-1**. I refer to these neuropsychological tests as cognitive tests for the purposes of this thesis.

Psychiatric evaluation included the UWDRS psychiatric score (UWDRS-P), a modified version of the Neuropsychiatric Inventory (ModNPI), the Generalised Anxiety Disorder-7 (GAD7) questionnaire and Patient Health Questionnaire-9 (PHQ9) on depression. The UWDRS-P involves a brief interview and mental state examination of the participant, with or without a family member present to assist. The ModNPI is a 20-item third-person questionnaire that includes questions on the frequency and severity of a range of psychological symptoms and behaviours. I administered it in person or via telephone to a family-member, friend or carer. Participants and their chosen respondent were informed in advance that responses would be

Table 2-1 Cognitive test battery

Domain	Test	Abbreviation	Normative data
Abstract reasoning	Weschler Abbreviated Scale of Intelligence (WASI) - matrix reasoning test	MRT	Manual
Language	National adult reading test	NART	Journal article[83]
	Graded naming test	GNT	Journal article[83]
Memory	Recognition memory test for faces	RMTF	Manual
	Recognition memory test for words	RMTW	
	Camden paired associate learning	CPAL	
Processing speed	Trail making test – part A	TMTA	Journal article[84]
Executive function	Weschler Memory Scale Revised (WMS-R) - digit span backwards	DSB	GENFI cohort
	Phonemic verbal fluency– F-A-S	FAS	Journal article[85]
	Semantic (category) verbal fluency - animals	Animals	
	Delis-Kaplan Executive Function System - interference (stroop) test	DKEFSI	Manual
	Trail making test – part B	TMTB	Journal article[84]
	Weschler Adult Intelligence Scale (WASI) – digit symbol test	DSym	Manual
Calculation	Graded difficulty arithmetic	GDA	Journal article[86]
Visuoperceptual	Visual Object and Space Perception battery (VOSP) - fragmented letters test	VOSPFL	Manual
Visuospatial	Visual Object and Space Perception battery (VOSP) - number location test	VOSPNL	
Social	Ekman 35 faces test	Ekman	GENFI cohort

Summary of cognitive tests performed during baseline research visits. These abbreviations listed are used throughout the thesis.

confidential. The PHQ9 and GAD7 questionnaires were completed by hand and participants were offered help by reading the questions aloud or writing responses where necessary. Participants were also asked to complete the 36-item short form survey (SF36) on quality of life. Scores were corrected for age using normative data from the Omnibus Survey in Britain.[87]

Ophthalmological assessments including bedside and slit lamp examination for KF rings, corneal photography and retinal ocular coherence tomography performed by a neuro-ophthalmologist. A systemic examination was performed to assess for signs of chronic liver disease.

2.3 Biosamples

All participants were asked to collect a 24-hour urine sample, provide a mid-stream urine sample and undergo venepuncture. A subset of participants was offered lumbar puncture and/or skin biopsy. All samples were stored within the Leonard Wolfson Experimental Neurology Centre Biobank at the UCL Queen Square Institute of Neurology unless otherwise stated. Studies related to the mid-stream urine samples, lumbar punctures and skin biopsies are not covered in this thesis. These aspects of the CROWD study protocol are mentioned here for completeness.

The 24-hour urine samples were collected up to two days before the research visit and while continuing medication. Aliquots were stored at -20 °C prior to copper quantification in the Department of Clinical Biochemistry at Northern General Hospital, Sheffield. The mid-stream urine samples were collected during the research visit and aliquots were then stored at -80 °C in the Leonard Wolfson Biomarkers Laboratory.

Venepuncture was performed during the research visit using a 21 or 23 g butterfly needle. Plasma and serum samples were stored at -80 °C for biomarker studies. Whole blood and serum samples were transferred to pathology services at NHNN for routine clinical analysis, including full blood count, liver function tests and international normalised ratio, and a plasma sample was transferred overnight at 4 °C to Southampton University Hospital Trace Elements Laboratory for serum copper, serum caeruloplasmin and exchangeable copper assays.

Lumbar puncture was performed using a 22 g atraumatic needle. Aliquots of CSF were stored at -80 °C and a sample was transferred to pathology services at NHNN for routine clinical analysis. Skin biopsy was performed using a 3-6 mm punch biopsy kit. Samples were transferred in DMEM culture medium for fibroblast culture prior to transfer to the Sheffield Institute of Translational Neuroscience for further studies.

2.4 Neuroimaging

MRI data were acquired on a Siemens Prisma 3T system with a 64-channel head/neck coil (Siemens Healthcare, Erlangen, Germany). T1-weighted

(structural), fluid-attenuated inversion recovery (FLAIR), diffusion-weighted imaging (DWI) and susceptibility-weighted imaging (SWI) data were collected using the pulse sequence parameters summarised in **Table 2-2**. Data were

Table 2-2 MRI parameters

	3D T1	3D FLAIR	DWI	3D SWI	B0 fieldmap
Pulse sequence	MPRAGE	IR-SPACE	Double refocused PGSE-EPI	3D gradient echo	2D gradient echo
Voxel resolution (mm^3)	1.1 x 1.1 x 1.1	1.0 x 1.0 x 1.0	2.5 x 2.5 x 2.5	1.0 x 1.0 x 1.0	3.0 x 3.0 x 3.0
Matrix size	256 x 256 x 208	256 x 256 x 192	96 x 96 x 59	256 x 192 x 176	64 x 64 x 55
Field of view (mm)	282 x 282 x 229	256 x 256 x 192	240 x 240 x 148	256 x 192 x 176	192 x 192 x 165
Orientation	Sagittal	Sagittal	Axial	Axial	Axial
Phase-encoding direction	A >> P	A >> P	A >> P	R >> L	R >> L
Echo time, TE (ms)	2.93	403	90	4.94/9.88/14.82/19.76/24.70	4.92/7.38
Recovery time, TR (ms)	2000	4800	7300	30	688
Flip angle (degrees)	8	Variable	-	15	60
Acquisition bandwidth (Hz/Px)	240	751	1578	280/260/260/260/260	260
Parallel imaging (GRAPPA acceleration factor)	2	3	2	3	None
Total scan time	5 min 6 sec	4 min 54 sec	8 min 47 sec	4 min 9 sec	1 min 31 sec
Other sequence specific parameters	Inversion time, IR, 850 ms	Inversion time, TI, 1650 ms SPACE turbo factor 243	Twice-refocused 2D multi-slice SE-EPI readout, b = 1000 s/mm^2 for diffusion encoding along 64 orientations. Five interspersed b = 0 s/mm^3 scans.	Partial Fourier 6/8 Monopolar readout Flow compensation for first echo.	2D multi-slice

GRAPPA, generalised autocalibrating partial parallel acquisition; IR-SPACE, inversion recovery – sampling perfection with application optimised contrast using different flip angle evolutions; MPRAGE, magnetisation prepared rapid gradient echo; PGSE-EPI, pulsed-gradient spin echo – echo planar imaging.

visually inspected immediately after each acquisition to allow individual sequences to be repeated if movement artefacts were identified on gross inspection.

3. Clinical phenotyping

3.1 Introduction

Movement disorders

The neurological features of WD were described in Wilson's seminal paper in 1912 and have subsequently been confirmed in several large retrospective studies.[40,42,54,88,89] The movement disorders phenotype is variable; tremor, dystonia, parkinsonism, ataxia, chorea or a combination of these can occur, usually with early bulbar involvement. Left untreated, patients become anarthric, increasingly immobile, and develop contractures.[3]

Several different classifications for movement disorders in WD have been proposed. Denny-Brown referred to two subgroups: one with predominant tremor referred to as having *pseudosclerosis* and another with predominant dystonia referred to using Wilson's original terminology, *progressive lenticular degeneration*.[90] Marsden and Walshe separately described a third category referred to as parkinsonian.[91,92] Czlonkowska *et al* recently performed the UWDRS in 53 treatment-naïve patients with neurological presentations from Poland and classified patients according to the predominant neurological syndrome; 62% exhibited tremor, 15% were dystonic and 11% had parkinsonism.[38]

Dysarthria is the most common neurological feature of WD and can be caused by underlying dystonia, parkinsonism or cerebellar dysfunction.[38,42,93-95] Walshe and Yealland retrospectively analysed the initial symptoms and signs in 136 patients with neurologic presentations in the UK; speech disturbance was an initial symptom in 71% and sign in 79% of cases.[54] Drooling and dysphagia were noted in 21% and 14% of cases, respectively.

A postural upper limb tremor is the most common movement disorder in WD and was an initial symptom in 70% of Walshe and Yealland's cohort. In a cohort of 107 patients from Germany where 40 had postural tremor, 18 had kinetic and 12 had rest tremor.[52] Six patients had a head tremor and one patient had a wing-beating tremor. Dystonia is seen in up to 69% of neurologic presentations, [42] with facial or oromandibular involvement in 10–72%.[42,54] The frequency of parkinsonism is also variable between cohorts. In contrast to the

findings from Czlonkowska *et al*, other groups have reported that parkinsonism accounts for 45%-62% of presentations.[40,54] Machado *et al* highlighted that while bradykinesia was seen in 58% of their cohort, isolated parkinsonism occurred in less than 2% of neurologic presentations.[42] Other neurological phenotypes have been described. Cerebellar ataxia is often associated with postural tremor and occurs in 28% of cases.[42] Chorea has been reported in 9–16% of patients, but again, in the context of other concomitant movement disorders.[40,42] Spasticity is not usually associated with WD in the absence of advanced disease but hyperreflexia is seen in up to 8%.[42]

While overall outcomes for patients with neurological presentations have been reported in several studies,[12,26,62] only one study has systematically examined how specific neurological symptoms or movement disorders respond to treatment: Burke *et al* reported neurological examination subscores longitudinally, over a mean duration of 35 months, for 88 patients who were treated with ammonium tetrathiomolybdate or trientine in a clinical trial.[96] They found that dystonia was the movement disorder most resistant to treatment. The proportion of the initial examination subscores that persisted at last follow up was 86% for dystonia, 32% for speech, 19% for tremor and 12% for parkinsonism.

UWDRS-N scores are increasingly used in observational studies and clinical trials to measure neurological involvement clinically. However, UWDRS-N is a composite measure of the severity of several movement disorders. Little attention has been paid to the relative weighting for neurological endophenotypes or how the scale performs in chronically-treated patients with WD.

Cognitive deficits

A large number of studies, summarised in **Table 3-1**, have examined cognitive function in chronically-treated patients with WD. Most compared unselected patients with WD or those with neurological presentations to healthy controls or normative data.[43,44,97-103] Some compared patients with neurological presentations to those with hepatic presentations,[44,100,104,105] although definitions for the latter varied and some studies classified patients according to the presence or absence of neurological features at time of assessment.

Table 3-1 Summary of studies on cognitive deficits in WD

Study	N	Tests
Knehr *et al* [106]	7 WD	Wide Range Vocabulary, **Progressive Matrices***
Goldstein *et al* [107]	15 WD	WAIS (**Information**[¶], **Comprehension**[¶], Arithmetic, **Similarities**[¶], Digit Span, **Verbal IQ**[¶], **Block Design**[¶]), **WMS**[¶], Shipley-Hartford Tests (Vocabulary, **Concept Formation**[¶], **IQ**[¶])
Tarter et al [108]	10 hWD 10 HC	Token, **Symbol Digit**[§], TMTA, **TMTB**[§], Tapping, **Block Design**[§], Animal naming, Digit Span,
Medalia *et al* [97,98]	19 nWD 12 hWD 15 HC	WAIS-R (VIQ, **PIQ***, **FIQ***), **WMS***[†] (Paired Associate Learning, Logical Memory, Visual Reproduction), **Dementia Rating Scale***, WCST (categories, perseverations), Animal Naming Test, Boston Naming Test, **TMTA***, **TMTB***, RAVLT (**recall***, recognition), verbal fluency (**FAS***)
Littman *et al* [109]	17 nWD 17 HC	Visual scanning task (processing speed)
Lang *et al* [99]	11 nWD 6 hWD 17 HC	Multiple Choice Vocabulary Test (IQ), WAIS (Digit Span, Picture Arrangement, Arithmetic), Progressive Matrices, Benton test, Achievement Assessment System (**reasoning***, fluency, spatial imagination, figure discrimination), Intelligence Structure Test (mental figure assembly, mental rotation), SKT (recall, recognition, digit arrangement, **perceptual speed***, interference)
Rathbun *et al* [104]	22 nWD 12 hWD	WAIS-R (VIQ, PIQ, FIQ, Information, Comprehension, Arithmetic, Similarities, Vocabulary, **Digit Span**[†], Picture Completion, **Picture Arrangement**[†], **Block Design**[†], Object Assembly**, Digit Symbol**[†]) Hooper Visual Organisation Test, Raven's Progressive Matrices, Benton Visual Retention Test, Symbol Digit Modalities (**written**[†]**, oral**[†]), Peabody picture vocabulary, WMS (immediate verbal, delayed verbal, **immediate figural**[†], **delayed figural**[†])
Portala *et al* [100]	7 nWD 8 hWD	Automated Psychological Test system (**Finger Tapping and Alternation test***[†], Selective Attention, Simultaneous Capacity, **Digit Span***, Associative Learning, Long Term Memory, **Grammatical Reasoning***, **Perceptual Maze Test***, **Austin Maze Test**[†], **Impulsive errors index***)
Seniow *et al* [44]	50 nWD 17 hWD 50 HC	Weschler-Bellevue Intelligence Test (**VIQ***[†]**, PIQ***[†]**, FIQ***[†]**, Information***[†]**, Comprehension***, **Arithmetic***[†]**, Digit Span***[†]**, Similarities***[†]**, Picture Arrangement***[†]**, Picture Completion***[†]**, Object Assembly***[†]**, Block Design***[†]**, Digit Symbol***[†]), **RAVLT***[†], **Benton Visual Memory test***[†], **Raven's Progressive Matrices***[†]
Hegde *et al* [101]	12 nWD	**Digit Vigilance Test***, **Color Trails Test***, **Triads test***, **Digit Symbol***, **Finger-Tapping***, **verbal fluency (animals*)**, **Verbal N-back***, Tower of London, **WCST***, **RAVLT***, **Rey's Complex Figure Test***
Frota *et al* [43]	20 WD	MMSE, **Dementia Rating Scale***, WCST, DSB, memory test for figures, Clock design, Hooper, Cubes, verbal fluency (**FAS***), **Stroop test***, CERAD naming test, **Fontal Assessment Battery***
Ma *et al* [110]	30 WD 30 HC	MMSE, WAIS-R (**FIQ***), **verbal fluency***, **Digit span***, **Stroop test***, **WCST***, **Game of Dice Task***

Wenisch *et al* [105]	13 nWD 18 hWD	WAIS (Information, **Digit Span†**, Similarities, **Matrices†**, **Digit Symbol†**, **Symbol Search†**), **California Verbal Learning Test†**, Modified Taylor Complex Figure, Letter-number sequencing, WMS (Visual span), WCST, Stroop Test (Colours†, Words†, Interference), D^2 attention test, Baddeley's Double Task Test, TMTA, TMTB
Iwanski *et al* [103]	33 nWD 34 hWD 43 HC	Test of Everyday Attention **(sustained attention*†§**, **selective attention*†**, **divided attention*†**, **attentional switching*†**)
Peyroux *et al* [45]	10 nWD 9 hWD 20 HC	NART, WMS (total recall), VOSPNL, TMTA, D^2 attention test, Facial Emotion Recognition Test (Joy, **Anger*§**, Sadness, **Fear*†**, Disgust, Contempt), Movie for the Assessment of Social Cognition (**no theory of mind*†**), Ambiguous Intentions Hostility Questionnaire (Hostility bias, Attribution of responsibility, **Aggression bias*)**

Studies are listed in chronological order. Tests where group differences were identified are highlight in bold.
** indicates differences between patients with neurological presentation and healthy controls (or normative data).*
† indicates differences between patients with neurological presentations and patients with hepatic presentations (or 'asymptomatic' patients).
§ indicates differences between patients with hepatic presentations (or 'asymptomatic') and healthy controls.
¶ indicates where scores improved longitudinally with treatment.

Others included comparisons between patients with hepatic presentations, or those who are otherwise symptomatic, and healthy controls.[44,45,99,103,108]

Deficits in nearly all cognitive domains have been reported but language function is usually preserved. In the only longitudinal study, Goldstein *et a*l found that scores for the Weschler Adult Intelligence Scale (WAIS) subtests for Information, Comprehension, Similarities, Block Design and Verbal IQ, in addition to the Weschler Memory Scale (WMS) and Shipley Hartford Test for Concept Formation improved with treatment. Most of the improvements were seen in the first 30 months but scores were still improving across each of these tests at 60 months in some patients.

In the largest cross-sectional study, Seniow *et al* compared scores for the Weschler-Bellevue Intelligence Test, including all 10 subtests, and the Rey Auditory-Verbal Learning Test, Benton Visual Memory Test and Raven's Progressive Matrices between 50 patients with neurological symptoms (at the time of assessment), 17 patients without neurological symptoms and 50 healthy controls matched for age, sex and education.[44] Scores for all subtests were lower in patients with neurological symptoms than controls and, with the exception of the Comprehension subtest, were lower in patients with neurological symptoms than without neurological symptoms. The authors state that there is 'a tendency towards a clear, albeit mild, cognitive deterioration in the majority of patients, which does not justify the label of dementia' but acknowledge that patients with severe neurological involvement who may have more severe cognitive deficits were excluded.[44]

A similar conclusion regarding the severity of cognitive impairment in WD was reached by Lang *et al*. They performed a comprehensive battery of tests and, after correcting for multiple comparisons, identified deficits in reasoning and perceptual speed only.[99] They conclude that patients with WD 'may function on a psychometrically normal level and seem to be impaired only when compared to a closely matched control group'. Others have reported more marked cognitive impairment is common.[43] For example, Wenisch *et al* found that 13 of 18 patients with neurological symptoms and 7 of 13 patients without neurological symptoms scored more than 2 standard deviations below age- and sex-adjusted means on at least one cognitive test.

Executive function is the most commonly affected cognitive domain.[43,101,102,111,112] Deficits in working memory, assessed by digit span,[99,102,104,113] and inhibitory control, assessed by the Stroop test or DKEFSI,[43,102,113] are reported whereas set shifting, measured using the Wisconsin Card Sorting Test, appears to be spared in some studies.[43,97] Verbal fluency is often impaired, [43,98,99,101,102] however semantic fluency is preserved relative to phonemic fluency.[43] Studies on attention have suggested that sustained, divided and selective attention and attention switched are also affected.[103,114] Processing speed, usually measured by trail making test part A is often reported to be reduced.[43,97,102,113]

Memory impairment is also common in WD.[43] Reports on whether memory encoding or retrieval and whether recall or recognition are more affected are inconsistent. [43,98,101,112] One study has reported deficits in both visual and verbal memory,[101] whereas others found that visual memory is worse than verbal memory.[104,112] Isaacs-Glaberman *et al* identified deficits in delayed verbal recall but not delayed verbal recognition. There was no difference in the rate of learning or rate of forgetting between patients and controls. They argue that, similar to Huntington's disease, patients with WD have impaired retrieval with intact recognition. Visuospatial function is preserved in some but not all studies.[43,100,104] The extent to which visuoperceptual ability is impaired is not known, but difficulties with the Benton Visual Retention and Rey's Complex Figure tests, which also depend on other cognitive domains, have been reported.[112] Difficulties in recognizing facial emotions have been described as have other deficits in social cognition, including theory of mind and attributional bias.[45]

Several measures of verbal and non-verbal intelligence have been studied in WD. Verbal ability appears to be relatively preserved whereas, with the exception of one study,[104] scores for matrix reasoning tests, which measure abstract reasoning, are lower in cohorts of patients with neurological presentations.[99,106,112,113] Evidence from a few individual cases suggests that non-verbal performance improves in the first few years of treatment.[97,106,113,115] Bedside cognitive screening tests have also been used. In a recent study, Camarata *et al* found that 49% of chronically-treated patients, most of whom

had neurological symptoms, scored below 26 on a Montreal Cognitive Assessment (MoCA).[116]

A few studies have reported cognitive deficits in patients with hepatic presentations. Tarter *et al* compared 10 patients with hepatic presentations, none of whom had neurological symptom or hepatic encephalopathy at the time of assessment, to 20 matched healthy controls.[108] They identified low scores for tests including digit symbol, block design and TMTB. Iwanski *et al* found that patients without neurological symptoms had impaired sustained attention and Peyroux *et al* found that they had impaired facial emotion recognition, specifically for recognising fear.[45,103] Kirk *et al* recently reported that more than a half of patients in a Danish cohort of chronically-treated patients had deficits in either the continuous reaction time (CRT) or portosystemic encephalopathy (PSE) tests. These measure a combination of attention, processing speed, visual orientation and inhibitory control and deficits occurred independently of initial presentation.[117]

There are several limitations to the aforementioned studies. Normative data were seldom used to adjust test scores for age, sex or educational level and only a few studies included statistical corrections for multiple comparisons.[43,99,104] While several studies included measures of neurological severity,[43,97,101] their association with specific cognitive tests was not determined and the UWDRS had not been introduced when all but a few were performed. Finally, the majority of studies did not consider the impact of motor impairment, duration of disease or underlying disease activity on test performance.

Psychiatric features

Psychiatric features are common in WD and are often the presenting symptom. In a large retrospective cohort study of 195 patients, Denning and Berrios reported that 51% of patients had psychiatric features at presentation and 20% had seen a psychiatrist prior to presentation.[46]

Personality and behavioural changes, mood disorders and anxiety appear to be common. Denning and Berrios reported that the most common psychiatric features were incongruous behaviour, irritability, aggression and personality change.[46] They also found that dysarthria, dysphagia, tremor and

rigidity were more common in patient with than without psychiatric symptoms and that patients with incongruous behaviour or irritability had higher neurological symptom scores suggesting that movement disorders and psychiatric involvement are closely linked. Svetel *et al* prospectively characterised a cohort of 50 consecutive chronically-treated patients using the structured clinical interview for DSM-IV Axis I disorders and the Neuropsychiatric inventory (NPI). The latter is based on family or caregiver interviews.[118] The most common features were anxiety (62%), depression (36%), irritability (26%), disinhibition (24%) and apathy (24%). Twenty patients (40%) had more than three psychiatric features and seven (14%) had two highlighting that, similar to neurological phenotypes, psychiatric features are often mixed in WD. In a cohort of 68 patients, most of whom had hepatic presentations, Schaefer *et al* found that 56% were considered to be at risk of depression or suffering with manifest depression based on having PHQ9 scores >4.[119]

Several other psychiatric phenotypes have been reported. While rare in some retrospective studies, psychosis was the initial presentation in 13 of the 36 case reports in a systematic review.[120] Other case reports included presentations mimicking ADHD and OCD.[121-123] A case of psychosis associated with hypersexuality is also described.[124] Sleep is also disrupted; poor nocturnal sleep, excessive daytime sleepiness and cataplexy-like episodes are reported to be common.[125]

There is some evidence that psychiatric symptoms improve with chelation therapy. In a follow up study, Dening and Berrios described the persisting psychiatric features in 129 of their original cases at two follow up appointments, after mean 3.5 and 6.8 years, respectively.[126] Of the most common psychiatric features, incongruous behaviour improved whereas irritability and depression did not. Patients with persisting psychiatric features had higher neurological scores and, more specifically, dysarthria at follow up.

Importantly, several studies have found a marked discrepancy between the severity of psychiatric symptoms when reported by patients and observed by clinicians or care-givers. Portala *et al* identified differences for mood, emotional involvement, lack of appropriate emotions, panic attacks and physical discomfort.[127] In a subsequent paper, they found that male patients

reported less anxiety, verbal aggression, irritability, suspicion and guilt than controls.[128] Similarly, Seniow *et al* found that patients with neurological presentations rated themselves as having less anger-hostility and problems with interpersonal sensitivity than controls and suggest that impaired self-awareness could relate to an underlying prefrontal syndrome.[129] More recently, this group reported that methods relying on self-description are inferior to informant-based measures when diagnosing apathy in patients with WD.[130]

These studies highlight the limitations of using self-reported symptoms to assess psychopathology in WD and illustrate the importance of discussing symptoms with relative or carers in the clinical setting. They also cast doubt on the clinical utility of the UWDRS-P, which heavily relies on self-described symptoms and has not been widely used or tested in other cohorts. I therefore incorporated a modified version of the NPI that includes items on sleep,[125] hypersexuality,[124] compulsive behaviours,[123] and altered response to pain,[127] each of which been reported in WD, in our protocol.

Relationships between clinical features

Three studies, each conducted around three decades ago, investigated the inter-relationships between movement disorders, cognitive deficits and psychiatric features in WD:

Medalia *et al* assessed the association between neurological, cognitive, psychiatric features in 24 patients.[131] Neurological features were assessed as absent, mild, moderate or severe depending on the number of neurological signs, cognitive features were assessed using the WMS memory quotient alone and psychiatric features were assessed using a psychopathology index calculated as the mean score across eight subtest in the Minnesota Multiphasic Personality Inventory (MMPI). The neurological, memory quotient and psychopathology indices were higher in those patients with neurological presentations but there was no correlation between these scores leading the authors to conclude that neurological, cognitive and psychiatric features are not inter-related in WD.

Dening and Berrios performed exploratory factor and cluster analyses on the demographic, clinical and biochemical features of 400 patients from 4

separate cohorts.[89] When analysing the largest cohort (n=195),[46] with 39 input variables they identified four groups. There were two *neurological* groups, one associated with dysphagia, dystonia, rigidity and dysarthria and the other associated with the presence of KF rings, tremor and thrombocytopenia, a *psychiatric* group associated with personality change, incongruous behaviour, irritability and cognitive impairment and a *hepatic* group associated with oedema and spider naevi. However, the grouping of clinical features differed when the analysis was limited to 11 input variables and in the other cohorts. Cluster analysis also yielded different groups between cohorts. The authors conclude they have confirmed the presence of the clinically-observed subgroups statistically and that there may be more than one type of neurological syndrome but recognise the groupings derived are a function of the variables selected for analysis.

Oder *et al* subsequently used an exploratory factor analysis to identify the relationships between movement disorders, cognitive deficits, psychiatric symptoms and neuroimaging abnormalities in a prospectively-recruited cohort of 47 patients.[132] The first group was characterised by bradykinesia, rigidity, a high score on the Number Connection Test (similar to TMTA), low mood and third ventricular dilatation. The second group was associated with ataxia, tremor, reduced functional capacity (disability) and focal thalamic lesions. A third group exhibited dyskinesia, dysarthria, an organic personality syndrome and focal lesions in the putamen and pallidum. This mimics the classifications proposed by Walshe and Marsden where patients with neurological presentations are divided according to dystonia, tremor and parkinsonism phenotypes. The authors conclude that they provide evidence of three neurological subgroups of WD associated with distinct psychiatric features and neuroimaging findings and that these may be of use in determining prognosis and evaluating treatments.

Health-related quality of life

Quality of life (QoL) is a complex, subjective and multidimensional concept. The World Health Organisation defines it is as ‘individuals’ perception of their position in life in the context of the culture and value systems in which they live

and in relation to their goals, expectations, standards and concern'.[133] Several studies have assessed QoL in patients with WD:

Komal Kumar *et al* found that the Neurological Symptom Score correlated with the physical but not psychological, social relationship or environmental domains of the WHO-BREF scale.[134] They also identified a positive correlation between disease duration and the physical domain. The SF-36 has been used in two similar large studies. Patients with neurological presentations had lower SF-36 scores for physical function, role limitations due to physical function, mental health and emotional role than patients with hepatic presentations.[119,135] Svetel *et al* also reported that patients with psychiatric symptoms had lower scores for all SF-36 domains, except physical function, bodily pain and vitality. Scores for the Mini-Mental State Examination (MMSE), Hamilton Depression Rating Scale and cognitive/behaviour and motor domains from the Global Assessment Scale correlated with all SF-36 scores, except bodily pain. In a multivariate analysis, higher scores in the GAS cognition and behaviour domain and lower MMSE scores were the most significant contributors to health-related QoL in patients with WD. More recently, Camarata *et al* found that patient with WD had lower mental than physical QoL, measured by the SF-12, and patients with a diagnosis of major depressive disorder had lower mental QoL scores.[116] UWDRS-N scores correlated with physical QoL scores but they found no association between MoCA scores and physical or mental QoL.

While these studies consistently show that persistent neurological or psychiatric features are associated with lower QoL in chronically-treated patients, results on the relationship with cognitive function are conflicting and the extent to which other psychiatric disorders, beyond depression, affect QoL in WD is unknown.

Clinical phenotyping objectives

Overall, a large number of studies have characterised movement disorders, cognitive deficits or psychiatric features associated with WD separately but very few have investigated these in parallel. With the exception of single study investigating deficits in attention, none have explored the relationship between

neurological and cognitive deficits or psychiatric features since the UWDRS was introduced.

The primary purpose of characterising the demographic characteristics and clinical phenotypes in our participants in detail is to understand the structure of the cohort and performance of various clinical rating scales in order to prepare for biomarker studies. Additional objectives were to apply the UWDRS-N, including subscores for neurological endophenotypes, in an independent cohort of patients, to examine cognitive deficits in WD using a comprehensive battery of tests and with rigorous methods to control for demographic variables and multiple testing, to assess the psychiatric symptoms without relying on self-reported symptoms and, finally, to determine the relationships between specific movement disorders, cognitive deficits and psychiatric symptoms and how these influence QoL in WD.

3.2 Methods

UWDRS-N scores and subtotals for UWDRS-N items related to dystonia, parkinsonism, tremor, ataxia and chorea were calculated. Items for speech, handwriting and rigidity were considered separately given these can be influenced by multiple phenotypes. Scores were compared between patients with neurological and hepatic presentations and between patients with active and stable disease using the Mann-Whitney tests and correlated with other scores using Spearman's rank correlation coefficients. P values with and without false discovery rate (FDR) correction for multiple comparisons were recorded. Associations with age, sex, disease duration, age at onset, presence of cirrhosis and treatment were tested using Spearman's rank correlation coefficients and Mann-Whitney tests, as appropriate. Statistical analyses were performed in R (v3.6.0) and graphs were created in GraphPad Prism (v7), as for all subsequent analyses in this thesis unless otherwise stated.

Cognitive tests scores for CROWD participants were converted to Z scores using normative data previously outlined in **Table 2-1**. These were then compared between groups using the Mann-Whitney tests. P values with and without FDR correction were recorded. Spearman's rank correlation coefficients were used to determine the association between individual tests scores and UWDRS-N scores. Linear regression was used to determine the effect of handwriting scores for the dominant hand on the association between tests of processing speed and UWDRS-N.

UWDRS-P, ModNPI, PHQ9, GAD7 and SF-36 scores were compared between groups using Mann-Whitney tests. Spearman's rank correlation coefficients were used to determine associations with UWDRS-N and cognitive test scores with and without FDR correction.

3.3 Results

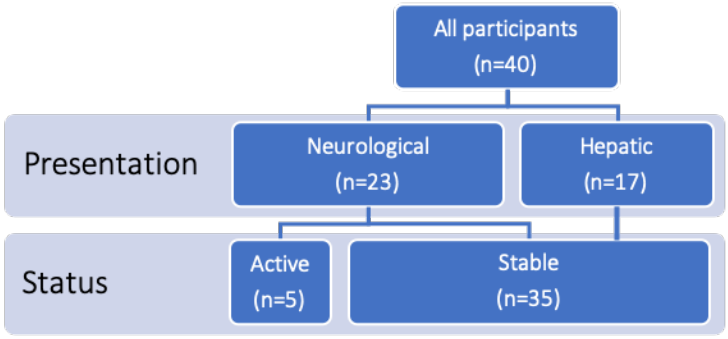
Demographics

The structure of the cohort is summarised in **Figure 3-1**. I recruited 23 patients with neurological presentations and 17 patients with hepatic presentations with an age range between 16 and 68 years. This includes six asymptomatic individuals who were diagnosed through family screening but had abnormal liver function tests or hepatic imaging and were therefore classified as having hepatic presentations. There were five patients subcategorised as having active disease, including two with recent neurological presentations and three with recent neurological deterioration associated with non-adherence, in the preceding six months. One of the patients with recent diagnosis was experiencing paradoxical worsening at the time of their assessment.

Demographic and clinical characteristics are summarised in **Table 3-2**. There were no differences in the age, sex, years of education, ethnicity, age at symptom onset, disease duration, evidence of cirrhosis, alanine transaminase, platelet counts or treatments between participants with neurological and hepatic presentations or between those with active or stable disease. One patient, subcategorised as having active disease due to non-adherence, had clinical features of decompensated cirrhosis (ascites) without evidence of hepatic encephalopathy. This was the only participant with evidence of synthetic dysfunction (INR>1.3). No other patients had symptomatic liver disease.

The mean disease duration was 23 years. *ATP7B* sequencing had been performed in 27 of 40 participants. The remaining patients were diagnosed prior the widespread use of genetic testing. Of those tested, homozygous or compound heterozygous mutations were identified in all but one patient in whom a single pathogenic mutation was identified. This patient had a typical neurological presentation associated with neuroimaging abnormalities, low serum caeruloplasmin, elevated urinary copper output and KF rings. The individual mutations have not been reported here given only one combination of pathogenic mutations occurred more than once (in participants

Figure 3-1 Flowchart of participants



Flowchart demonstrating the number of study participants classified as having neurological or hepatic presentations and subclassified as having active or stable neurological disease.

Table 3-2 Demographic and clinical characteristics

	All (n=40)	Hepatic (n=17)	Neurological (n=23)	P value	Stable (n=35)	Active (n=5)	P value
Age, mean (SD), yr	43 (14)	42 (15)	44 (14)	0.77	44 (14)	39 (17)	0.47
Sex, n (%)							
Female	20 (50%)	8 (47%)	12 (52%)	0.75	19 (54%)	1 (20%)	0.15
Male	20 (50%)	9 (53%)	11 (48%)		16 (46%)	4 (80%)	
Years of education, mean (SD)	15 (3)	15 (3)	14 (3)	0.35	15 (3)	14 (3)	0.49
Ethnicity, n							
White	28	12	16	0.07	25	3	0.84
Asian/Asian British	7	1	6		6	1	
Other ethnic group [a]	5	4	1		4	1	
ATP7B genotype							
Homozygous mutations	9	4	5		8	1	
Compound heterozygous mutations	17	10	7		14	3	
Single heterozygous mutation	1	0	1		1	0	
Not tested/results unavailable	13	3	10		12	1	
Age at symptom onset, mean (SD), yr	19 (12)	20 (17)	19 (9)	0.75	19 (13)	19 (6)	0.82
Disease duration, mean (SD), yr	23 (15)	20 (15)	25 (16)	0.37	24 (15)	20 (20)	0.56
Evidence of cirrhosis, n (%) [b]	17 (43%)	7 (41%)	10 (43%)	0.88	14 (40%)	3 (60%)	0.40
Alanine transaminase, median (IQR), IU/L	31 (17-46)	41 (27-47)	20 (17-43)	0.08	27 (17-46)	45 (32-73)	0.15
Decompensated liver disease, n (%)	1 (3)	0 (0)	1 (4)		0 (0)	1 (20)	
Treatment, n (%)							
Penicillamine	26 (65%)	9 (53%)	17 (74%)	0.39	23 (66%)	3 (60%)	0.82
Trientine	9 (23%)	5 (29%)	4 (17%)		7 (20%)	2 (40%)	
Zinc	1 (3%)	1 (6%)	0 (0%)		1 (3%)	0 (0%)	
Combination	1 (3%)	0 (0%)	1 (4%)		1 (3%)	0 (0%)	
Liver transplantation	3 (8%)	2 (12%)	1 (4%)		3 (95)	0 (0%)	

[a] *Ethnic groups other than White and Asian/Asian British were grouped to preserve the anonymity of some participants.*
[b] *Evidence of cirrhosis was based on previous imaging or histopathological results.*

who were siblings) and disclosing specific mutations may stand to unnecessarily de-anonymise certain participants within the WD community.

The majority of patients were being treated with penicillamine or trientine at the time of their research visit. Three stable patients had received liver transplants; two had initially presented with acute liver failure and one had developed decompensated cirrhosis. None of the patients were receiving dopaminergic, neuroleptic or anti-cholinergic treatments at the time of assessment.

Movement disorders

UWDRS-N scores were not associated with age (r=0.24, P=0.16), sex (13 vs 16, P=0.52), age at onset (r=0.03, P=0.86), disease duration (r=0.19, P=0.33), evidence of cirrhosis (6 vs 9, P=0.87) or treatment (penicillamine vs trientine, 12 vs 8, P=0.37) in stable patients.

Group differences in the UWDRS-N scores are summarised in **Table 3-3** and illustrated in **Figure 3-2**. UWDRS-N scores were higher and all neurological endophenotypes, except tremor and chorea, were more common in participants with neurological presentations than those with hepatic presentations. A similar pattern is seen when comparing participants with active disease to those with stable disease. Reviewing pre-visit questionnaire responses, 16 of 21 (76%) of participants with neurological presentations reported tremor at disease onset.

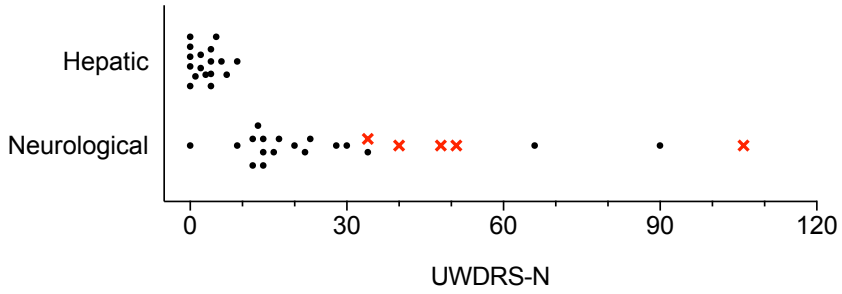
The breakdown of neurological phenotypes in individual participants is shown in **Figure 3-3**. The majority of patients with neurological findings on examination had mixed movement disorders with a combination of speech disturbance, dystonia, parkinsonism, tremor and/or ataxia. Of patients with neurological presentations, the mean breakdown of UWDRS-N scores was speech 6%, dystonia 18%, parkinsonism 35%, tremor 10%, ataxia 11%, chorea 4%, rigidity 9% and writing 7%. The frequency of scores for individual items in the UWDRS-N are depicted in **Figure 3-4**. Dystonia, parkinsonism and tremor were more frequent in the upper than lower limbs and speech disturbance and oromandibular dystonia were common. Some patients with hepatic presentations also had mild neurological signs such as ‘mildly

Table 3-3 Group differences in UWDRS-N scores

Group (maximum)	**All (n=40)** Median [IQR] (range)	**Hepatic (n=17)** Median [IQR] (range)	**Neurological (n=23)** Median [IQR] (range)	**P value**	**Stable (n=35)** Median [IQR] (range)	**Active (n=5)** Median [IQR] (range)	**P value**
Speech (4)	0 [0-2] (0-4)	0 [0-0] (0-0)	2 [0-3] (0-4)	**<0.001*** **	0 [0-2] (0-4)	3 [2-3] (1-3)	**0.003** **
Dystonia (28)	2 [0-4] (0-26)	0 [0-0] (0-1)	4 [2-5] (0-26)	**<0.001*** **	1 [0-4] (0-26)	9 [5-12] (5-18)	**0.001** **
Parkinsonism (36)	3 [1-8] (0-28)	1 [0-2] (0-7)	8 [6-13] (0-28)	**<0.001*** **	2 [1-7] (0-24)	17 [15-21] (11-28)	**0.001** **
Tremor (44)	1 [0-2] (0-17)	0 [0-1] (0-4)	1 [0-5] (0-17)	0.31	0 [0-2] (0-11)	3 [1-10] (0-17)	0.09
Ataxia (20)	1 [0-3] (0-17)	0 [0-0] (0-1)	3 [2-5] (0-17)	**<0.001*** **	0 [0-2] (0-10)	6 [5-8] (3-17)	**<0.001*** **
Chorea (24)	0 [0-1] (0-6)	0 [0-0] (0-2)	0 [0-2] (0-6)	0.47	0 [0-1] (0-6)	0 [0-0] (0-5)	0.88
Rigidity (20)	0 [0-2] (0-20)	0 [0-0] (0-2)	1 [0-3] (0-20)	**<0.001*** **	0 [0-1] (0-20)	6 [3-6] (0-15)	**0.01** *
Writing (8)	0 [0-1] (0-8)	0 [0-0] (0-2)	1 [1-4] (0-8)	**<0.001*** **	0 [0-1] (0-8)	4 [4-4] (1-8)	**0.002** **
Total (184)	12 [4-24] (0-106)	3 [0-4] (0-9)	22 [14-31] (0-106)	**<0.001*** **	9 [3-17] (0-90)	48 [40-56] (34-106)	**0.001** **

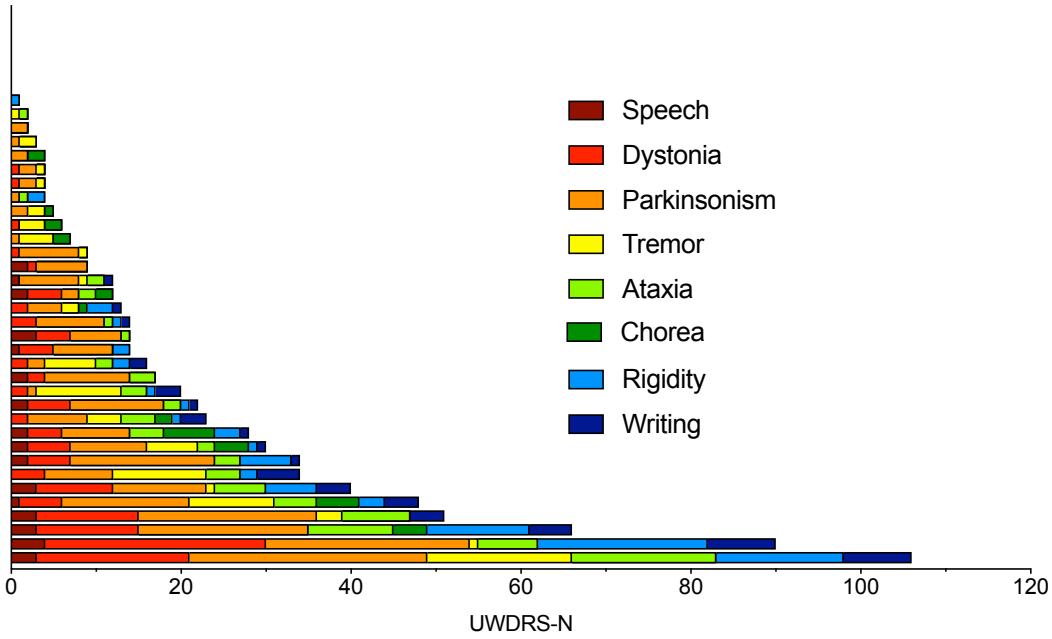
** = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. IQR, interquartile range; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.*

Figure 3-2 Group differences in UWDRS-N scores



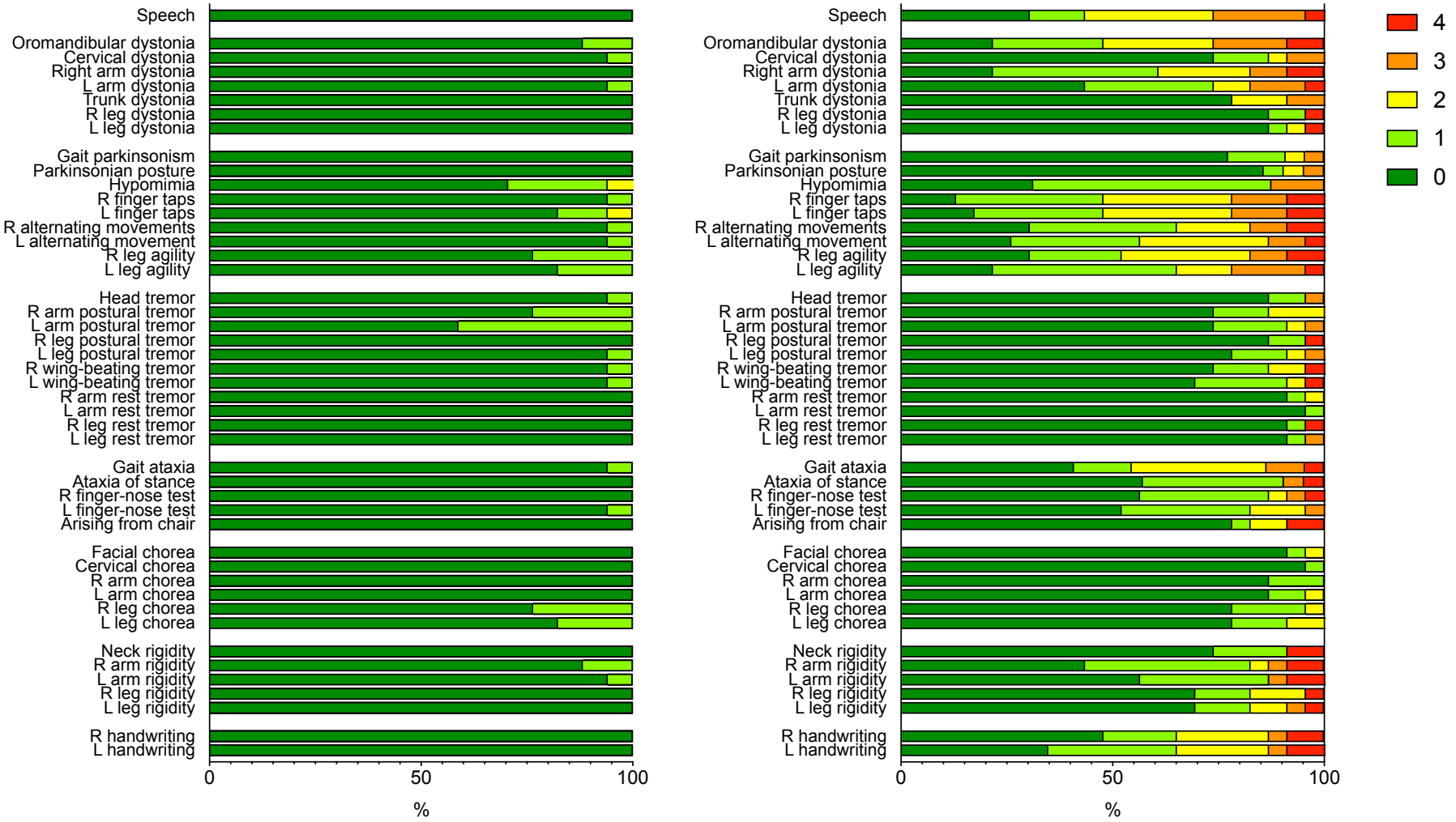
Individual scores are shown with those from active cases denoted with a red cross. UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.

Figure 3-3 Breakdown of UWDRS-N scores by phenotype



Each row represents an individual participant. Phenotypes are denoted by colours. Speech, rigidity and writing subscores relate to multiple phenotypes and are grouped separately. Six participants had UWDRS-N scores of 0. UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.

Figure 3-4 Frequency of individual UWDRS-N items



The distribution of UWDRS-N item between patients with hepatic (left) and neurological (right) presentations are shown. UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.

impaired' finger or foot taps and a 'slight or hardly perceptible' postural tremor in the upper limbs, which each count for one UWDRS-N point.

The association between specific neurological phenotypes is described in **Table 3-4**. Speech, dystonia, parkinsonism, ataxia, rigidity and writing scores are highly correlated with each other whereas scores for tremor were only correlated with ataxia. Scores for writing and chorea were not correlated with other scores.

Table 3-4 Correlation matrix of neurological phenotypes

A) Spearman's correlation coefficients

	Speech	Dystonia	Parkinsonism	Tremor	Ataxia	Chorea	Rigidity	Writing
Speech		**0.78**	**0.77**	-0.06	**0.69**	0.09	0.42	**0.49**
Dystonia	**0.78**		**0.86**	0.27	**0.79**	0.18	**0.69**	**0.76**
Parkinsonism	**0.77**	**0.86**		0.23	**0.76**	0.15	**0.61**	**0.74**
Tremor	-0.06	0.27	0.23		0.33	0.22	0.23	**0.51**
Ataxia	**0.69**	**0.79**	**0.76**	0.33		0.16	**0.65**	**0.86**
Chorea	0.09	0.18	0.15	0.22	0.16		0.17	0.16
Rigidity	**0.42**	**0.69**	**0.61**	0.23	**0.65**	0.17		**0.77**
Writing	**0.49**	**0.76**	**0.74**	**0.51**	**0.86**	0.16	**0.77**	

B) P values

	Speech	Dystonia	Parkinsonism	Tremor	Ataxia	Chorea	Rigidity	Writing
Speech		**<0.001*** **	**<0.001*** **	0.72	**<0.001*** **	0.57	0.007**	**0.001** **
Dystonia	**<0.001*** **		**<0.001*** **	0.10	**<0.001*** **	0.27	**<0.001*** **	**<0.001*** **
Parkinsonism	**<0.001*** **	**<0.001*** **		0.16	**<0.001*** **	0.36	**<0.001*** **	**<0.001*** **
Tremor	0.73	0.10	0.16		0.04*	0.17	0.16	**<0.001*** **
Ataxia	**<0.001*** **	**<0.001*** **	**<0.001*** **	0.04*		0.33	**<0.001*** **	**<0.001*** **
Chorea	0.57	0.27	0.36	0.17	0.33		0.29	0.33
Rigidity	**0.007** **	**<0.001*** **	**<0.001*** **	0.16	**<0.001*** **	0.29		**<0.001*** **
Writing	**0.001** **	**<0.001*** **	**<0.001*** **	**<0.001*** **	**<0.001*** **	0.33	**<0.001*** **	

*Correlation coefficients between UWDRS-N subscores for individual neurological phenotypes (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold.*

Cognitive deficits

Cognitive testing was performed in 39 participants. One declined to undertake any cognitive tests and several patients were unable to complete specific tests such that 32 of 780 (4%) potential test scores were excluded or missing.

Group differences in cognitive test scores are summarised in **Table 3-5** and **Figure 3-5**. The neurological group performed worse than the hepatic group in the MRT, TMTA and VOSPFL, after FDR correction. Scores for several other tests of executive function, such as the DSB, FAS, TMTB and DSym were also lower in neurological group but these differences did not persist after controlling for multiple comparisons.

The frequency of participants with Z scores below -2 for each cognitive test are shown in **Table 3-6**. Low performance was relatively common for RMTF, CPAL, TMTA and TMTB in each group and the cohort overall. For RMTF, 24% of patients with hepatic presentations and 45% of patients with neurological presentations had scores below this cut-off.

Individual participant scores for each cognitive test are shown in **Figure 3-6**. This reveals that 17 of 22 (77%) patients with neurological presentations and 9 of 17 (53%) patients with hepatic presentations scored more than two standard deviations below the mean on at least one cognitive test. Of the five patients with hepatic presentations who had UWDRS-N scores of zero, two scored more than two standard deviations below the mean in at least one cognitive test.

Associations between cognitive test scores and UWDRS-N scores in stable patients are summarised in **Table 3-7**. UWDRS-N scores were negatively correlated with the MRT, RMTF, TMTA, DSym and GDA, after controlling for multiple comparisons. Speech, dystonia and parkinsonism subscores were also negatively correlated with several further tests of abstract reasoning, processing speed and executive function, after FDR correction.

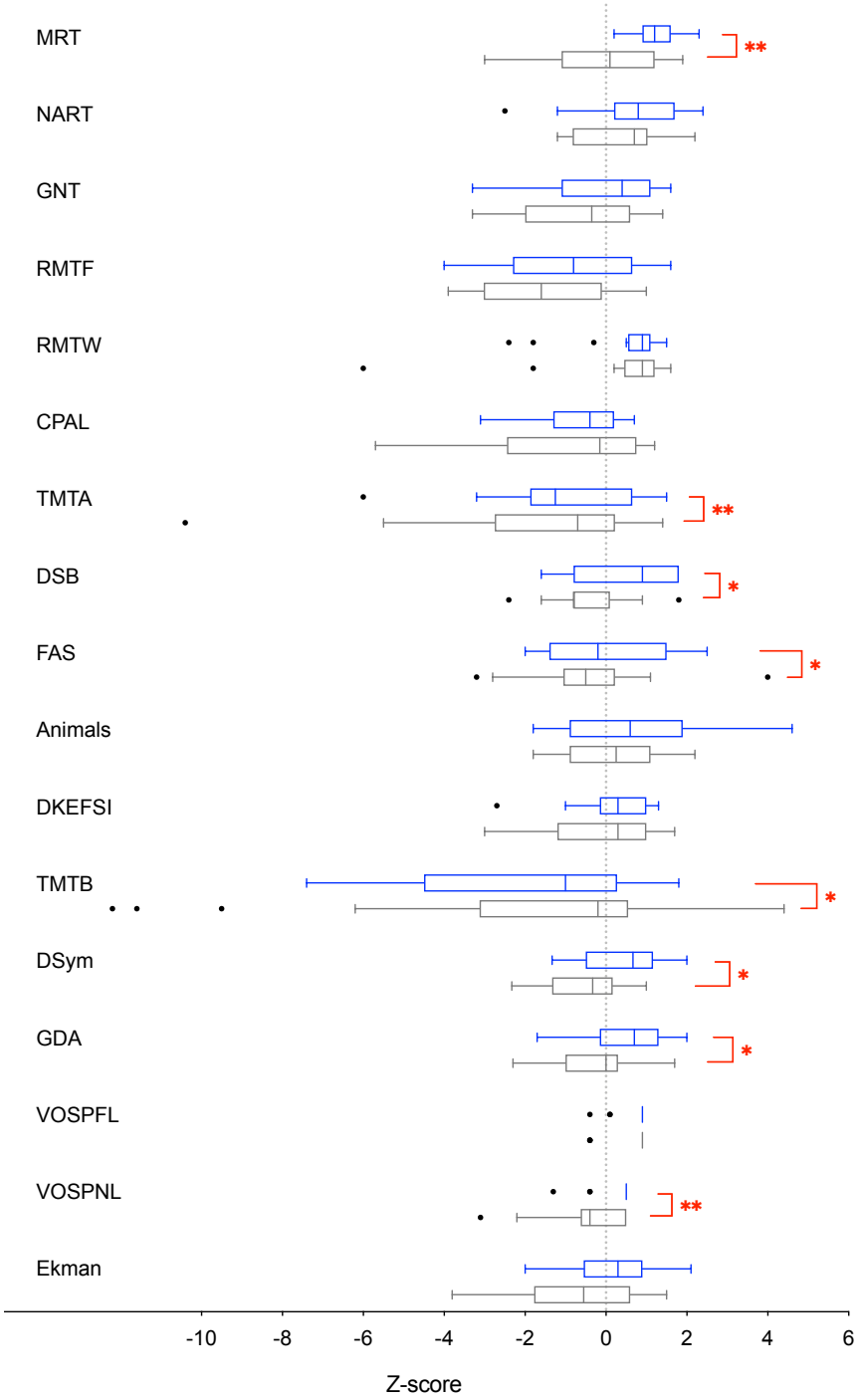
UWDRS-N handwriting scores for the dominant hand are included in **Table 3-7**. There was a negative correlation for both the TMTA and DSym tests after FDR correction and the aforementioned association between the UWDRS-N and TMTA did not persist when controlling for the handwriting

Table 3-5 Group differences in cognitive test scores

Domain	Test	All (n=39) Z score, median [IQR]	Hepatic (n=17) Z score, median [IQR]	Neurological (n=22) Z score, Median [IQR]	P value	Stable (n=35) Z score, median [IQR]	Active (n=4) Z score, Median [IQR]	P value
Abstract reasoning	MRT	1.0 [-0.1, 1.4]	1.2 [1.0, 1.6]	0.1 [-1.0, 1.2]	**0.001****	1.0 [0.2, 1.5]	-0.9 [-1.7, 0.2]	0.07
Language	NART	0.7 [-0.4, 1.2]	0.8 [0.3, 1.5]	0.7 [-0.8, 1.0]	0.33	0.8 [-0.4, 1.3]	0.7 [-0.3, 0.9]	0.57
	GNT	-0.1 [-1.6, 0.6]	0.4 [-0.9, 0.9]	-0.4 [-1.9, 0.6]	0.31	0.0 [-1.5, 0.6]	-0.1 [-1.0, 0.2]	0.71
Memory	RMTF	-1.1 [-2.9, 0.1]	-0.8 [-1.9, 0.4]	-1.6 [-3.0, -0.1]	0.11	-0.8 [-2.7, 0.1]	-3.0 [-3.2, -2.6]	0.04*
	RMTW	0.9 [0.5, 1.2]	0.9 [0.6, 1.0]	0.9 [0.5, 1.2]	0.66	0.9 [0.7, 1.2]	0.2 [-0.8, 0.7]	0.31
	CPAL	-0.4 [-1.7, 0.3]	-0.4 [-0.9, 0.1]	-0.2 [-2.0, 0.7]	0.75	-0.3 [-1.5, 0.3]	-2.7 [-3.4, -1.5]	0.10
Processing speed	TMTA	-1.0 [-1.9, 0.2]	-0.2 [-0.9, 0.8]	-1.6 [-2.7, -0.8]	**0.002****	-0.9 [-1.6, 0.3]	-3.6 [-5.8, -2.4]	0.03*
Executive function	DSB	0.1 [-0.8, 0.9]	0.9 [-0.8, 1.8]	-0.8 [-0.8, 0.1]	0.05*	0.1 [-0.8, 0.9]	-0.4 [-1.2, 0.1]	0.32
	FAS	-0.4 [-1.2, -0.2]	0.3 [-0.7, 1.1]	-1.0 [-1.5, 0.1]	0.02*	-0.1 [-1.0, 0.8]	-2.0 [-2.9, -1.2]	0.01*
	Animals	0.5 [-0.9, 1.3]	0.6 [-0.9, 1.5]	0.2 [-0.8, 1.2]	0.67	0.6 [-0.7, 1.4]	-1.0 [-1.3, -0.5]	0.07
	DKEFSI	0.3 [-0.7, 1.0]	0.3 [0.0, 1.0]	-0.3 [-1.3, 1.0]	0.21	0.3 [-0.6, 1.0]	-1.7 [-2.0, -1.0]	0.05*
	TMTB	-0.6 [-3.8, 0.4]	0.3 [-1.0, 0.7]	-1.4 [-6.2, -0.1]	0.03*	-0.4 [-2.4, 0.5]	-8.9 [-11.8, -4.7]	0.03*
	DSym	-0.2 [-0.9, 0.9]	0.7 [-0.3, 1.0]	-0.3 [-1.3, 0.0]	0.01*	0.0 [-0.6, 1.0]	-2.0 [-2.3, -1.5]	0.003**
Calculation	GDA	0.2 [-0.4, 0.7]	0.7 [0.0, 1.3]	0.0 [-0.9, 0.3]	0.02*	0.3 [-0.3, 0.9]	-0.9 [-1.3, -0.5]	0.05
Visuoperceptual	VOSPFL	0.9 [0.9, 0.9]	0.9 [0.9, 0.9]	0.9 [-0.9, 0.9]	0.86	0.9 [0.9, 0.9]	0.9 [0.9, 0.9]	0.40
Visuospatial	VOSPNL	0.5 [-0.4, 0.5]	0.5 [0.5, 0.5]	-0.4 [-0.4, 0.5]	**0.006****	0.5 [-0.4, 0.5]	-0.4 [-1.1, -0.2]	0.18
Social	Ekman	-0.1 [-0.9, 0.8]	0.3 [-0.4, 0.9]	-0.6 [-1.5, 0.5]	0.11	-0.4 [-2.3, 0.5]	-0.4 [-1.6, 0.3]	0.34

*Group differences in Z scores for each cognitive test are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

Figure 3-5 Box-and-whisker (Tukey) plot of cognitive test scores



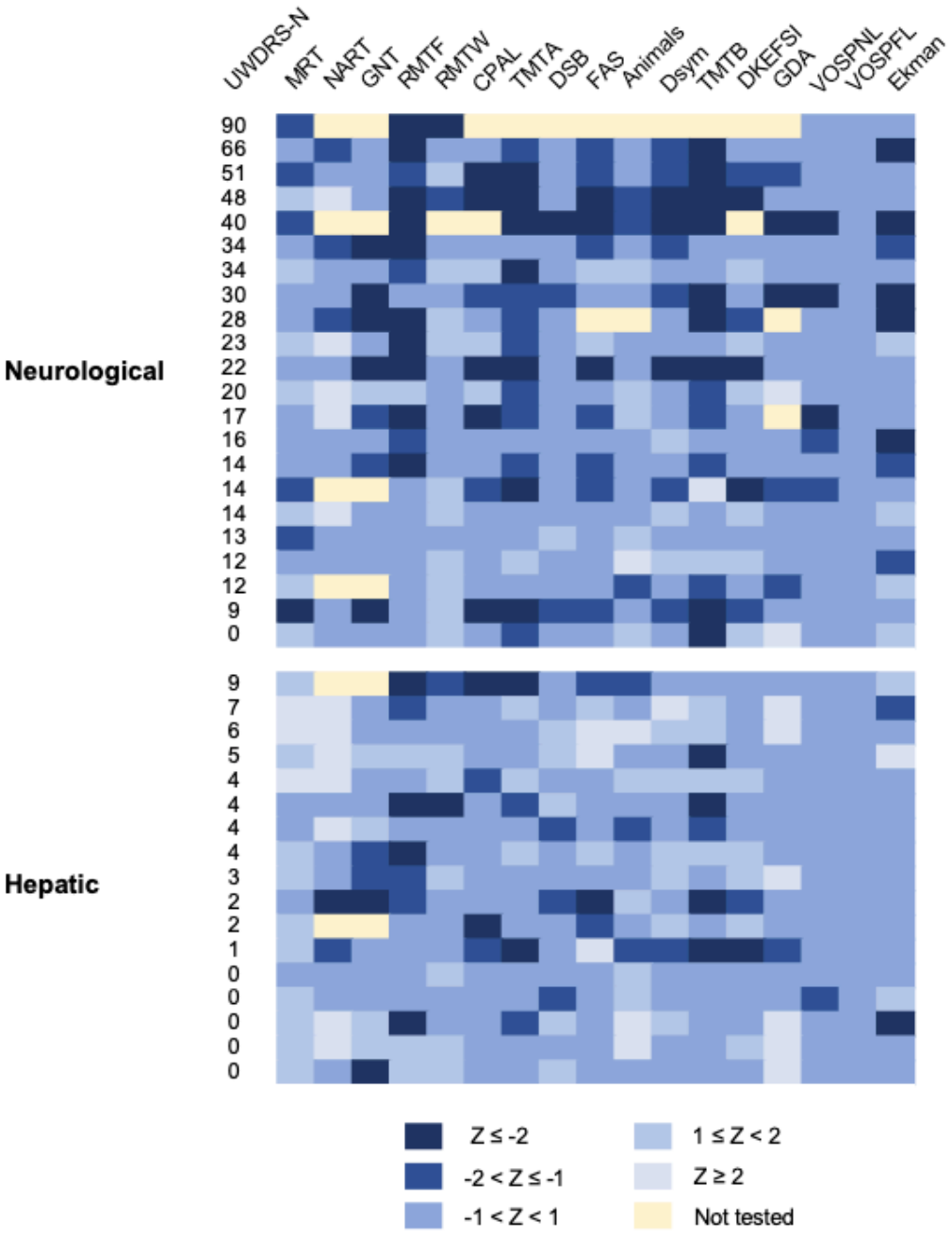
*Group differences in Z scores for cognitive test scores are shown. Scores for hepatic (blue) and neurological (grey) groups are shown. * = P value <0.05; ** = P value <0.01. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

Table 3-6 Frequency of poor performance in cognitive tests

Domain	Test	All	Hepatic	Neurological	Stable	Active
Abstract reasoning	MRT	3% (1/39)	0% (0/17)	5% (1/22)	3% (1/35)	**25% (**1/4)
Language	NART	3% (1/33)	7% (1/15)	0% (0/18)	4% (1/30)	0% (0/3)
	GNT	**15%** (5/33)	**14%** (2/15)	**17%** (3/18)	**17%** (5/30)	0% (0/3)
Memory	RMTF	**36%** (14/39)	**24%** (4/17)	**45%** (10/22)	**31%** (11/35)	**75%** (3/4)
	RMTW	5% (2/39)	6% (1/17)	5% (1/21)	3% (1/35)	**25%** (1/4)
	CPAL	**19%** (7/37)	**12%** (2/17)	**25%** (5/20)	**15%** (5/34)	**67%** (2/3)
Processing speed	TMTA	**24%** (9/38)	**12%** (2/17)	**33%** (7/21)	**18%** (6/34)	**75%** (3/4)
Executive function	DSB	3% (1/38)	0% (0/17)	5% (1/21)	0% (0/34)	**25%** (1/4)
	FAS	5% (2/38)	0% (0/17)	9% (2/21)	0% (0/34)	**50%** (2/4)
	Animals	0% (0/38)	0% (0/17)	0% (0/21)	0% (0/34)	0% (0/4)
	DKEFSI	8% (3/37)	6% (1/17)	**10%** (2/20)	6% (2/34)	**33%** (1/3)
	TMTB	**34%** (13/38)	**24%** (4/17)	**43%** (9/21)	**29%** (10/34)	**75%** (3/4)
	DSym	5% (2/38)	0% (0/17)	9% (2/21)	0% (0/34)	**50%** (2/4)
Calculation	GDA	3% (1/38)	0% (0/17)	5% (1/21)	0% (0/34)	**25%** (1/4)
Visuoperceptual	VOSPFL	0% (0/39)	0% (0/17)	0% (0/22)	0% (0/35)	0% (0/4)
Visuospatial	VOSPNL	8% (3/39)	0% (0/17)	**14%** (3/22)	6% (2/35)	**25%** (1/4)
Social	Ekman	8% (3/39)	0% (0/17)	**14%** (3/22)	6% (2/35)	**25%** (1/4)

Percentages of participants who scored more than two standard deviations below the mean are shown for each cognitive test. Percentages above 10% are highlighted in bold. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.

Figure 3-6 Individual participants scores for each cognitive test



Individual participant scores for each cognitive test are colour coded by Z score. Darker shades of blue indicate poorer performance. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.

Table 3-7 Associations between cognitive test and UWDRS-N scores

A) Spearman's correlation coefficients

Domain	Test	Speech	Dystonia	Parkins.	Tremor	Ataxia	Chorea	Rigidity	Writing	UWDRS-N
Abstract reasoning	MRT	**-0.71**	**-0.59**	**-0.56**	0.23	**-0.42**	0.10	-0.46	-	**-0.49**
Language	NART	**-0.48**	-0.25	-0.28	0.40	-0.10	0.03	-0.26	-	-0.14
	GNT	-0.35	-0.15	-0.25	0.26	0.04	0.11	-0.08	-	-0.08
Memory	RMTF	-0.38	**-0.49**	**-0.58**	-0.04	-0.37	-0.11	-0.33	-	**-0.47**
	RMTW	-0.13	-0.18	-0.22	-0.04	-0.08	0.03	-0.38	-	-0.20
	CPAL	-0.28	-0.09	-0.13	0.26	0.10	0.33	0.24	-	0.10
Processing speed	TMTA	**-0.50**	**-0.53**	**-0.56**	-0.08	**-0.43**	0.12	-0.38	**-0.47**	**-0.52**
Executive function	DSB	**-0.42**	-0.23	-0.28	0.23	-0.25	0.20	-0.24	-	-0.21
	FAS	**-0.63**	**-0.43**	**-0.54**	0.16	-0.36	0.23	-0.21	-	-0.38
	Animals	-0.35	-0.15	**-0.42**	-0.05	-0.24	-0.06	-0.27	-	-0.32
	DKEFSI	**-0.53**	-0.33	**-0.39**	0.15	-0.16	-0.05	-0.28	-	-0.31
	TMTB	**-0.40**	-0.31	**-0.45**	0.05	-0.34	0.06	-0.32	-0.35	-0.34
	DSym	**-0.63**	**-0.51**	**-0.54**	0.10	**-0.44**	0.11	-0.38	**-0.42**	**-0.46**
Calculation	GDA	**-0.53**	**-0.46**	**-0.50**	-0.06	**-0.44**	-0.02	-0.35	-	**-0.44**
Visuoperceptual	VOSPFL	-0.10	0.15	0.03	-0.03	0.10	-0.21	0.25	-	0.05
Visuospatial	VOSPNL	-0.38	**-0.41**	-0.31	0.04	-0.36	0.03	-0.34	-	-0.37
Social	Ekman	**-0.46**	-0.35	-0.21	0.16	-0.29	-0.22	-0.38	-	-0.29

B) P values

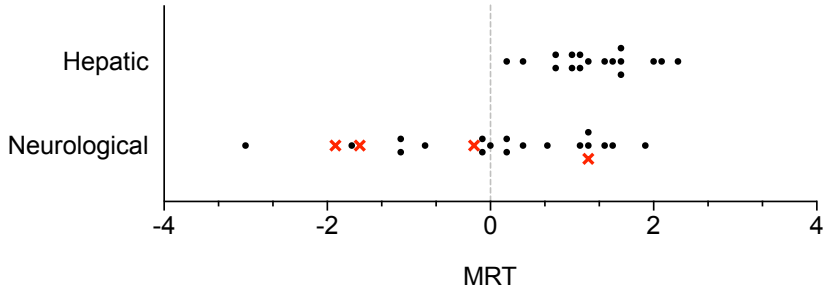
Domain	Test	Speech	Dystonia	Parkins.	Tremor	Ataxia	Chorea	Rigidity	Writing	UWDRS-N
Abstract reasoning	MRT	**<0.001*** **	**<0.001*** **	**<0.001*** **	0.15	**0.008** **	0.54	0.003**	-	**0.001** **
Language	NART	**0.005** **	0.16	0.11	0.02*	0.59	0.85	0.14	-	0.44
	GNT	0.05*	0.40	0.15	0.15	0.84	0.53	0.65	-	0.65
Memory	RMTF	**0.02***	**0.002** **	**<0.001*** **	0.80	0.02*	0.50	0.04*	-	**0.002** **
	RMTW	0.44	0.27	0.19	0.80	0.63	0.86	0.02*	-	0.23
	CPAL	0.10	0.62	0.46	0.13	0.55	0.05*	0.15	-	0.56
Processing speed	TMTA	**0.002** **	**<0.001*** **	**<0.001*** **	0.61	**0.01***	0.38	0.03*	**0.005*** **	**0.001** **
Executive function	DSB	**0.008** **	0.17	0.09	0.17	0.13	0.22	0.15	-	0.21
	FAS	**<0.001*** **	**0.01***	**<0.001*** **	0.23	0.04*	0.11	0.24	-	0.03*
	Animals	0.08	0.49	**0.02***	0.57	0.16	0.68	0.11	-	0.08
	DKEFSI	**<0.001*** **	0.05*	**0.02***	0.38	0.33	0.78	0.10	-	0.06
	TMTB	**0.01***	0.06	**0.004** **	0.78	0.04*	0.70	0.05	0.03*	0.04*
	DSym	**<0.001*** **	**<0.001*** **	**<0.001*** **	0.57	**0.005** **	0.52	0.02*	**0.009** **	**0.004** **
Calculation	GDA	**<0.001*** **	**0.004** **	**0.002** **	0.75	**0.007** **	0.89	0.04*	-	**0.008** **
Visuoperceptual	VOSPFL	0.54	0.36	0.870	0.87	0.56	0.20	0.12	-	0.76
Visuospatial	VOSPNL	0.02*	**0.009** **	0.05	0.79	0.03*	0.87	0.03*	-	0.02*
Social	Ekman	**0.004** **	0.03*	0.19	0.33	0.07	0.17	0.02*	-	0.07

*Correlation coefficients between Z scores for each cognitive test and UWDRS-N scores (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Writing scores refer to the dominant hand only in this table. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

score (ß = -0.04, P = 0.25). Individual scores for the MRT and RMTF are illustrated in **Figure 3-7** and **Figure 3-8**.

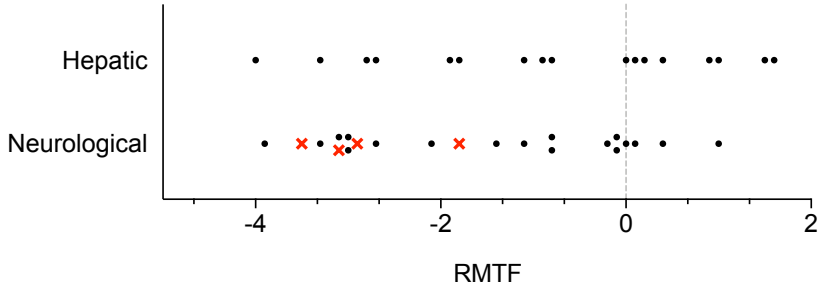
There was no association between scores for any test and disease duration in stable patients (data not shown).

Figure 3-7 Group differences in MRT scores



Individual Z scores for the MRT are shown with those from active cases denoted with a red cross. MRT, matrix reasoning test.

Figure 3-8 Group differences in RMTF scores



Individual Z scores for the RMTF are shown with those from active cases denoted with a red cross. RMTF, recognition memory test for faces.

Psychiatric features

Group differences in UWDRS-P scores, ModNPI scores and PHQ9 and GAD7 scores are summarised in **Table 3-8**, **Table 3-9** and **Table 3-10**, respectively. The distribution of UWDRS-P scores and ModNPI scores are also illustrated in **Figure 3-9** and **Figure 3-10**. Two participants did not give consent for the research team to consult a family member, friend or carer to complete the ModNPI and two participants did not complete the PHQ9 and GAD7 questionnaires.

There were no differences in any of these scores, or any individual items within these scores, between the neurological and hepatic groups, with or without FDR correction. There were differences in the UWDRS-P and ModNPI between active and stable patients: Scores for the auditory hallucinations and emotional lability items on the UWDRS-P were higher in patients with active disease, with FDR correction, as were scores for the hostile feelings, ideas of persecution, decreased verbal activity and low mood items, without correction for multiple comparisons. Scores for the delusions, agitation, depression, irritability, hypersexuality, humour, apathy and disinhibition items on the ModNPI, in addition to the total ModNPI score, were higher in patients with active disease, with FDR correction, as were the scores for the sleep and hyper-religiosity without FDR correction.

The distribution of scores for items with the UWDRS-P and ModNPI are shown in **Figure 3-11** and **Figure 3-12**. The most common UWDRS-P items were for memory loss, ideas of persecution, concentration, depressed mood, hostile feeling, anxiety and autonomic symptoms. The most common ModNPI items were irritability, depression, agitation, anxiety, sleep, apathy, disinhibition and empathy.

Associations between psychiatric rating scales and specific neurological phenotypes are summarised in **Table 3-11**. Speech scores correlated with ModNPI and GAD7 scores and parkinsonism scores correlated with PHQ9 and GAD7 scores, without FDR correction. The association between individual ModNPI items and specific neurological phenotypes are also summarised in **Table 3-12**. There were no associations with and several associations without correction for multiple testing. For example, speech

Table 3-8 Group differences in UWDRS-P subscores

Characteristic	**All (n=40)** Median [IQR] (range)	**Hepatic (n=17)** Median [IQR] (range)	**Neurological (n=23)** Median [IQR] (range)	**P *value***	**Stable (n=34)** Median [IQR] (range)	**Active (n=4)** Median [IQR] (range)	**P value**
Sleep	0 [0-1] (0-3)	0 [0-1] (0-3)	0 [0-1] (0-3)	0.75	0 [0-1] (0-3)	2 [0-2] (0-3)	0.08
Autonomic	0 [0-1] (0-4)	0 [0-1] (0-3)	0 [0-1] (0-4)	1.00	0 [0-1] (0-3)	1 [0-1] (0-4)	0.30
Sexual interest (inc)	0 [0-0] (0-3)	0 [0-0] (0-2)	0 [0-0] (0-3)	0.72	0 [0-0] (0-3)	0 [0-0] (0-3)	0.26
Sexual interest (dec)	0 [0-0] (0-4)	0 [0-0] (0-3)	0 [0-0] (0-4)	0.78	0 [0-0] (0-4)	0 [0-0] (0-3)	0.95
Memory	1 [0-1] (0-3)	1 [0-2] (0-3)	1 [0-1] (0-3)	0.83	1 [0-1] (0-3)	0 [0-1] (0-3)	0.50
Concentration	1 [0-2] (0-3)	1 [0-2] (0-3)	1 [0-2] (0-3)	0.73	0 [0-2] (0-3)	2 [2-3] (0-3)	0.07
Contact	0 [0-0] (0-3)	0 [0-2] (0-3)	0 [0-0] (0-3)	0.12	0 [0-0] (0-3)	0 [0-0] (0-2)	0.82
Hostile feelings	1 [0-2] (0-3)	0 [0-2] (0-3)	1 [0-2] (0-3)	0.35	0 [0-2] (0-3)	2 [2-3] (0-3)	0.04*
Ideas of persecution	1 [0-1] (0-4)	1 [0-1] (0-2)	1 [0-1] (0-4)	0.47	0 [0-1] (0-3)	2 [1-3] (1-4)	0.009**
Auditory hallucination	0 [0-0] (0-2)	0 [0-0] (0-1)	0 [0-0] (0-2)	0.72	0 [0-0] (0-1)	0 [0-2] (0-2)	**0.003****
Visual hallucination	0 [0-0] (0-0)	0 [0-0] (0-0)	0 [0-0] (0-0)	1.00	0 [0-0] (0-0)	0 [0-0] (0-0)	1.00
Suicidality	0 [0-0] (0-3)	0 [0-0] (0-3)	0 [0-0] (0-1)	0.39	0 [0-0] (0-3)	0 [0-1] (0-1)	0.06
Flight of thoughts	0 [0-0] (0-1)	0 [0-0] (0-1)	0 [0-0] (0-1)	0.64	0 [0-0] (0-1)	0 [0-1] (0-1)	0.11
Self esteem	0 [0-0] (0-2)	0 [0-0] (0-2)	0 [0-0] (0-2)	0.76	0 [0-0] (0-2)	0 [0-0] (0-2)	0.28
Voice level	0 [0-0] (0-1)	0 [0-0] (0-1)	0 [0-0] (0-1)	0.27	0 [0-0] (0-1)	0 [0-1] (0-1)	0.06
Verbal activity (inc)	0 [0-1] (0-2)	0 [0-1] (0-2)	0 [0-1] (0-2)	0.55	0 [0-1] (0-2)	1 [0-1] (0-1)	0.18
Verbal activity (dec)	0 [0-0] (0-3)	0 [0-0] (0-1)	0 [0-0] (0-3)	0.45	0 [0-0] (0-2)	0 [0-1] (0-3)	0.02*
Mood (elevated)	0 [0-0] (0-2)	0 [0-0] (0-1)	0 [0-0] (0-2)	0.81	0 [0-0] (0-2)	0 [0-0] (0-0)	0.39
Mood (depressed)	1 [0-2] (0-3)	1 [0-1] (0-3)	0 [0-2] (0-3)	1.00	0 [0-0] (0-2)	2 [1-2] (1-2)	0.02*
Anxiety	1 [0-1] (0-3)	0 [0-1] (0-3)	1 [0-1] (0-3)	1.00	0 [0-1] (0-3)	1 [1-2] (0-3)	0.08
Lability	0 [0-0] (0-3)	0 [0-0] (0-3)	0 [0-1] (0-3)	0.77	0 [0-0] (0-3)	2 [2-2] (0-3)	**0.002****
Disorientation	0 [0-0] (0-0)	0 [0-0] (0-0)	0 [0-0] (0-0)	1.00	0 [0-0] (0-0)	0 [0-0] (0-0)	1.00
Total	8 [3-13] (0-30)	8 [3-13] (0-27)	8 [4-15] (0-30)	0.90	8 [3-11] (0-27)	19 [16-27] (8-30)	0.01*

** = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Dec, decreased; inc, increased; IQR, interquartile range; UWDRS-P, unified Wilson's disease rating scale psychiatric subscore.*

Table 3-9 Group differences in ModNPI scores

Characteristic	**All (n=39)** Median [IQR] (range)	**Hepatic (n=16)** Median [IQR] (range)	**Neurological (n=22)** Median [IQR] (range)	**P value**	**Stable (n=34)** Median [IQR] (range)	**Active (n=4)** Median [IQR] (range)	**P value**
Delusions	0 [0-0] (0-12)	0 [0-0] (0-0)	0 [0-0] (012)	0.10	0 [0-0] (0-4)	1 [0-4] (0-12)	**0.001****
Hallucinations	0 [0-0] (0-0)	0 [0-0] (0-0)	0 [0-0] (0-0)	1.00	0 [0-0] (0-0)	0 [0-0] (0-0)	1.00
Agitation	0 [0-3] (0-9)	0 [0-0] (0-8)	1 [0-4] (0-9)	0.06	0 [0-2] (0-8)	7 [5-8] (1-9)	**0.006****
Depression	1 [0-4] (0-12)	1 [0-2] (0-8)	1 [0-4] (0-12)	0.60	1 [0-3] (0-9)	6 [3-9] (1-12)	0.03*
Anxiety	0 [0-4] (0-12)	1 [0-2] (0-12)	0 [0-4] (012)	0.80	0 [0-3] (0-12)	4 [1-8] (0-12)	0.17
Elation	0 [0-0] (0-1)	0 [0-0] (0-0)	0 [0-0] (0-1)	0.20	0 [0-0] (0-1)	0 [0-0] (0-0)	0.66
Irritability	2 [0-6] (0-12)	1 [0-2] (0-12)	2 [0-6] (0-12)	0.10	1 [0-3] (0-12)	8 [6-10] (6-12)	**0.007****
Motor behaviour	0 (0-0) (0-4)	0 [0-0] (0-0)	0 (0-0) (0-4)	0.10	0 [0-0] (0-4)	0 [0-1] (0-2)	0.20
Sleep[a]	0 [0-6] (0-12)	1 [0-7] (0-12)	0 [0-4] (0-8)	0.50	0 [0-4] (0-12)	7 [5-8] (2-8)	0.03*
Hypersexuality[a]	0 [0-0] (0-6)	0 [0-0] (0-0)	0 [0-0] (0-6)	0.40	0 [0-0] (0-0)	0 [0-2] (0-6)	**0.005****
Hyperreligiosity[a]	0 [0-0] (0-8)	0 [0-0] (0-0)	0 [0-0] (0-8)	0.40	0 [0-0] (0-0)	0 [0-2] (0-8)	**0.005****
Humour[a]	0 [0-0] (0-8)	0 [0-0] (0-8)	0 [0-0] (0-8)	0.20	0 [0-0] (0-8)	1 [0-4] (0-8)	0.04*
Apathy	0 [0-1] (0-12)	0 [0-0] (0-12)	0 [0-1] (0-12)	0.30	0 [0-0] (0-12)	12 [9-12] (0-12)	**0.01****
Disinhibition	0 [0-1] (0-12)	0 [0-0] (0-8)	0 [0-1] (0-12)	0.20	0 [0-0] (0-12)	5 [3-8] (0-12)	**0.01***
Eating habits[a]	0 [0-0] (0-12)	0 [0-0] (0-3)	0 [0-2] (0-12)	0.40	0 [0-0] (0-12)	1 [0-3] (0-6)	0.17
Empathy[a]	0 [0-0] (0-12)	0 [0-1] (0-4)	0 [0-0] (0-12)	0.40	0 [0-0] (0-4)	1 [0-5] (0-12)	0.16
Compulsive[a]	0 [0-0] (0-12)	0 [0-0] (0-12)	0 [0-0] (0-12)	0.40	0 [0-0] (0-12)	3 [0-8] (0-12)	0.06
Social/emotional[a]	0 [0-0] (0-8)	0 [0-0] (0-4)	0 [0-0] (0-8)	0.40	0 [0-0] (0-4)	0 [0-2] (0-8)	0.60
Trusting[a]	0 [0-0] (0-8)	0 [0-0] (0-1)	0 [0-2] (0-8)	0.05	0 [0-0] (0-8)	1 [0-2] (0-2)	0.15
Pain[a]	0 [0-0] (0-12)	0 [0-1] (0-12)	0 [0-0] (0-8)	0.70	0 [0-0] (0-12)	3 [0-7] (0-8)	0.14
Total	12 [3-25] (0-152)	9 [3-16] (0-76)	15 [6-38] (0-152)	0.30	10 [3-20] (0-76)	55 [38-89] (26-152)	**0.009****

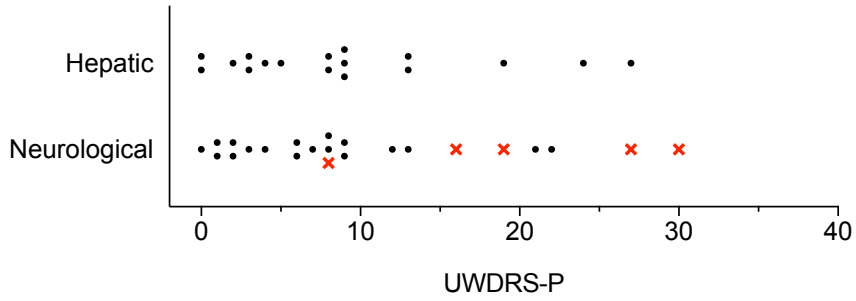
*[a] These psychiatric features are not included in the original version of the NPI. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. ModNPI, modified neuropsychiatric inventory.*

Table 3-10 Group differences in PHQ9 and GAD7 subscores

Characteristic	**All (n=39)** Median [IQR] (range)	**Hepatic (n=17)** Median [IQR] (range)	**Neurological (n=22)** Median [IQR] (range)	**P value**	**Stable (n=34)** Median [IQR] (range)	**Active (n=4)** Median [IQR] (range)	**P value**
Little interest in doing	1 [0-1] (0-3)	1 [0-2] (0-3)	1 [0-2] (0-3)	0.60	1 [0-1] (0-3)	1 [1-1] (0-1)	0.81
Feeling down	0 [0-0] (0-3)	0 [0-1] (0-3)	0 [0-1] (0-3)	0.65	0 [0-1] (0-3)	2 [1-3] (0-3)	0.09
Trouble sleeping	1 [0-2] (0-3)	0 [0-2] (0-3)	1 [0-2] (0-3)	0.69	0 [0-2] (0-3)	3 [3-3] (1-3)	0.02*
Feeling tired	1 [0-1] (0-3)	1 [0-2] (0-3)	1 [0-2] (0-3)	0.94	1 [0-1] (0-3)	3 [2-3] (1-3)	0.03*
Poor appetite/overeat	0 [0-1] (0-3)	0 [0-0] (0-3)	0 [0-1] (0-3)	0.42	0 [0-1] (0-3)	1 [0-2] (0-2)	0.34
Feel bad about self	0 [0-1] (0-3)	0 [0-1] (0-3)	0 [0-1] (0-2)	0.54	0 [0-1] (0-3)	1 [0-1] (0-2)	0.72
Difficulty concentrating	0 [0-1] (0-3)	0 [0-1] (0-3)	0 [0-1] (0-3)	0.34	0 [0-1] (0-3)	1 [0-1] (0-2)	0.74
Moving/speaking slow	0 [0-1] (0-3)	0 [0-0] (0-3)	0 [0-1] (0-3)	0.09	0 [0-0] (0-3)	3 [2-3] (0-3)	**0.005****
Think better off dead	0 [0-1] (0-3)	0 [0-0] (0-3)	0 [0-1] (0-3)	0.67	0 [0-0] (0-3)	2 [1-2] (0-3)	**0.01***
PHQ9 total	5 [2-9] (0-24)	4 [2-6] (0-24)	7 [2-9] (0-19)	0.46	4 [1-7] (0-24)	14 [10-18] (7-19)	0.02*
Feeling afraid	1 [0-2] (0-3)	1 [0-1] (0-3)	1 [0-2] (0-3)	0.77	1 [0-1] (0-3)	2 [1-2] (0-3)	0.30
Easily annoyed	0 [0-1] (0-3)	0 [0-0] (0-3)	1 [0-2] (0-3)	0.13	0 [0-1] (0-3)	2 [1-2] (0-3)	0.11
Being restless	0 [0-2] (0-3)	0 [0-1] (0-3)	1 [0-2] (0-3)	0.09	0 [0-2] (0-3)	2 [1-3] (0-3)	0.15
Trouble relaxing	1 [0-1] (0-3)	1 [0-1] (0-3)	1 [0-1] (0-3)	0.17	1 [0-1] (0-3)	1 [1-2] (0-3)	0.41
Worrying too much	1 [0-1] (0-3)	1 [0-1] (0-2)	1 [0-1] (0-3)	0.47	0 [0-1] (0-3)	1 [0-2] (0-3)	0.53
Not able to stop worry	1 [0-2] (0-3)	10 [0-1] (0-3)	1 [0-2] (0-3)	0.20	0 [0-1] (0-3)	2 [1-2] (1-3)	0.08
Nervous or anxious	0 [0-0] (0-3)	0 [0-0] (0-3)	0 [0-0] (0-3)	0.93	0 [0-0] (0-3)	1 [0-1] (0-3)	0.62
GAD7 total	4 [1-8] (0-19)	2 [1-4] (0-19)	6 [2-8] (0-18)	0.15	3 [0-6] (0-19)	9 [4-14] (2-18)	0.16

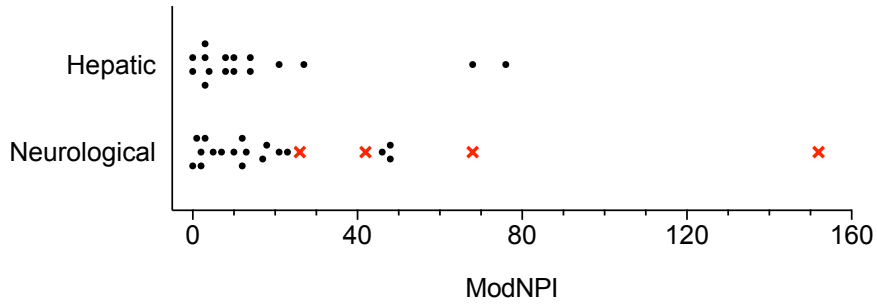
** = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. GAD7, generalised anxiety disorder assessment 7; PHQ9, patient health questionnaire 9.*

Figure 3-9 Group differences in UWDRS-P scores



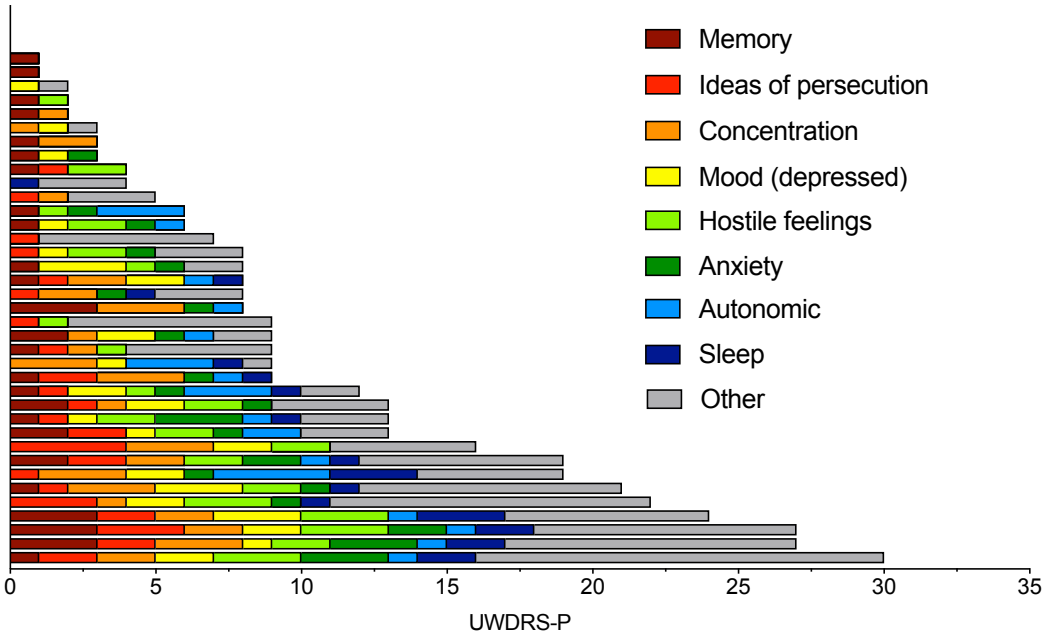
Individual UWDRS-P scores are shown with those from active cases denoted with a red cross. UWDRS-P, unified Wilson's disease rating scale psychiatric subscore.

Figure 3-10 Group differences in ModNPI scores



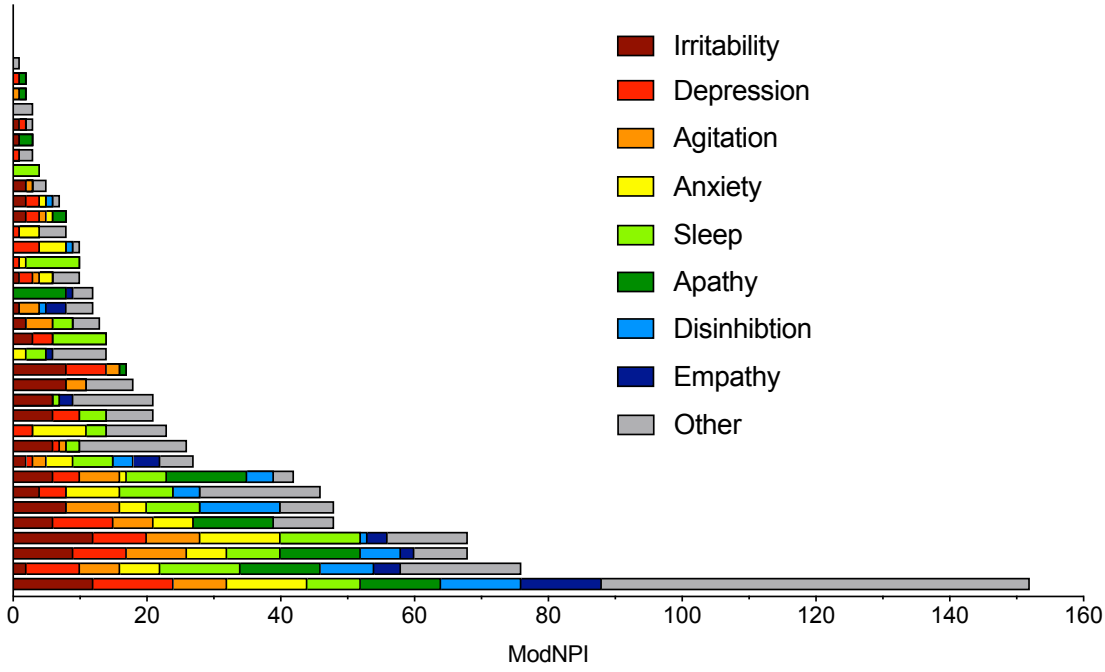
Individual ModNPI scores are shown with those from active cases denoted with a red cross. ModNPI, modified neuropsychiatric inventory.

Figure 3-11 Breakdown of UWDRS-P scores by psychiatric features



UWDRS-P scores for individual participants are shown. The legend is ordered according to the frequency of each characteristic in the cohort. Memory was the most frequent psychiatric feature. UWDRS-P, unified Wilson's disease rating scale psychiatric subscore.

Figure 3-12 Breakdown of ModNPI scores by psychiatric features



ModNPI scores for individual participants are shown. The legend is ordered according to the frequency of each characteristic in the cohort. Irritability was the most frequent psychiatric feature. ModNPI, modified neuropsychiatric inventory.

Table 3-11 Associations between psychiatric and UWDRS-N scores

A) Spearman's correlation coefficients

	UWDRS-P	ModNPI	PHQ9	GAD7
Speech	0.29	0.38	0.32	0.41
Dystonia	0.17	0.25	0.32	0.22
Parkinsonism	0.26	0.23	0.35	0.33
Tremor	0.14	-0.06	0.18	-0.02
Ataxia	0.22	0.15	0.29	0.26
Chorea	0.17	0.07	0.22	0.11
Rigidity	0.16	0.11	0.20	0.08
Writing	0.18	0.13	0.31	0.15
UWDRS-N	0.23	0.21	0.37	0.28

B) P values

	UWDRS-P	ModNPI	PHQ9	GAD7
Speech	0.07	0.02*	0.05	0.01*
Dystonia	0.28	0.13	0.05	0.19
Parkinsonism	0.11	0.16	0.03*	0.04*
Tremor	0.39	0.73	0.27	0.89
Ataxia	0.17	0.37	0.08	0.12
Chorea	0.30	0.68	0.18	0.49
Rigidity	0.33	0.51	0.22	0.65
Writing	0.27	0.43	0.06	0.36
UWDRS-N	0.14	0.20	0.02*	0.09

*Correlation coefficients between UWDRS-N scores and psychiatric rating scales (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. GAD7, generalised anxiety disorder assessment 7; ModNPI, modified neuropsychiatric inventory; PHQ9, patient health questionnaire 9; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; UWDRS-P, unified Wilson's disease rating scale psychiatric subscore.*

Table 3-12 Associations between ModNPI items and UWDRS-N scores

A) Spearman's correlation coefficients

	Speech	Dystonia	Parkinsonism	Tremor	Ataxia	Chorea	Rigidity	UWDRS-N
Delusions	0.33	0.36	0.37	0.18	0.42	0.33	0.17	0.38
Hallucinations	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Agitation	0.38	0.36	0.32	0.11	0.42	0.17	0.16	0.34
Depression	0.30	0.14	0.28	-0.04	0.08	0.01	0.04	0.19
Anxiety	0.26	0.04	0.10	-0.07	-0.06	0.11	-0.02	0.04
Elation	0.26	0.24	0.23	-0.21	0.21	0.17	0.22	0.20
Irritability	0.33	0.29	0.27	0.11	0.34	-0.04	0.14	0.29
Motor	0.03	0.04	0.15	0.35	0.24	0.07	0.16	0.19
Sleep	0.08	0.10	0.02	0.10	-0.13	-0.04	-0.07	0.00
Hypersexuality	0.27	0.26	0.27	0.18	0.28	-0.10	-0.14	0.25
Hyperreligiosity	0.27	0.26	0.27	0.18	0.28	-0.10	-0.14	0.25
Humour	0.35	0.38	0.45	-0.04	0.32	-0.27	0.18	0.33
Apathy	0.27	0.27	0.37	0.00	0.43	0.27	0.33	0.33
Disinhibition	0.19	0.08	0.20	0.16	0.05	0.14	-0.05	0.12
Eating habits	0.25	0.28	0.24	0.09	0.24	0.12	0.24	0.34
Empathy	0.01	-0.12	-0.11	-0.12	-0.06	-0.08	-0.20	-0.16
Compulsive	0.15	0.03	0.07	-0.18	-0.12	-0.20	-0.09	-0.02
Social/emotional	-0.05	-0.16	-0.08	0.04	-0.14	-0.03	-0.32	-0.07
Trusting	0.46	0.37	0.36	-0.15	0.09	0.00	0.02	0.25
Pain	0.11	0.03	-0.03	-0.14	0.02	-0.35	-0.06	-0.06
ModNPI	0.38	0.25	0.23	-0.06	0.15	0.07	0.11	0.21

B) P values

	Speech	Dystonia	Parkinsonism	Tremor	Ataxia	Chorea	Rigidity	UWDRS-N
Delusions	0.04*	0.03*	0.02*	0.28	0.009**	0.05*	0.31	0.02*
Hallucinations	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Agitation	0.02*	0.03*	0.05*	0.52	0.008**	0.31	0.35	0.03*
Depression	0.06	0.41	0.09	0.81	0.61	0.95	0.81	0.26
Anxiety	0.12	0.82	0.55	0.67	0.72	0.52	0.89	0.79
Elation	0.12	0.15	0.16	0.21	0.22	0.29	0.19	0.22
Irritability	0.04*	0.08	0.11	0.51	0.04*	0.80	0.40	0.08
Motor	0.87	0.79	0.36	0.03*	0.15	0.67	0.35	0.25
Sleep	0.63	0.53	0.88	0.53	0.42	0.82	0.66	0.99
Hypersexuality	0.10	0.11	0.11	0.28	0.09	0.54	0.40	0.13
Hyperreligiosity	0.10	0.11	0.11	0.28	0.09	0.54	0.40	0.13
Humour	0.03*	0.02*	0.005**	0.80	0.05	0.10	0.28	0.04*
Apathy	0.10	0.10	0.02*	0.99	0.007**	0.10	0.05*	0.04*
Disinhibition	0.25	0.64	0.23	0.35	0.76	0.40	0.78	0.47
Eating habits	0.13	0.09	0.15	0.60	0.15	0.48	0.15	0.04*
Empathy	0.95	0.47	0.50	0.47	0.72	0.63	0.23	0.34
Compulsive	0.37	0.84	0.67	0.28	0.46	0.22	0.58	0.90
Social/emotional	0.78	0.33	0.64	0.81	0.42	0.86	0.05*	0.66
Trusting	0.004**	0.02*	0.03	0.37	0.61	0.98	0.93	0.13
Pain	0.50	0.84	0.88	0.40	0.91	0.03*	0.70	0.71
ModNPI	0.02*	0.13	0.16	0.73	0.37	0.68	0.51	0.20

*Correlation coefficients between ModNPI items and UWDRS-N scores (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. ModNPI, modified neuropsychiatric inventory; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.*

scores were weakly associated with total ModNPI scores and item scores for delusions, irritability, agitation, altered sense of humour and trusting behaviour.

Associations between psychiatric rating scales and cognitive test scores are shown in **Table 3-13**. UWDRS-P scores were negatively correlated with DKEFSI scores only. Associations between ModNPI items and cognitive test are shown in **Table 3-14**. The depression item negatively correlated with GNT scores only. There were no other association between psychiatric rating scales or ModNPI items and cognitive test scores after correction for multiple testing.

Table 3-13 Associations between psychiatric and cognitive test scores

A) Spearman's correlation coefficients

Domain	Test	UWDRS-P	ModNPI	PHQ9	GAD7
Abstract reasoning	MRT	-0.27	-0.13	-0.26	-0.30
Language	NART	-0.23	-0.38	-0.33	-0.38
	GNT	-0.30	-0.43	-0.22	-0.34
Memory	RMTF	-0.03	-0.14	-0.10	-0.21
	RMTW	-0.25	-0.24	-0.08	-0.10
	CPAL	-0.10	-0.07	0.14	-0.02
Processing speed	TMTA	0.03	-0.04	-0.22	-0.31
Executive function	DSB	-0.19	-0.31	-0.01	-0.30
	FAS	0.09	-0.13	0.09	-0.03
	Animals	-0.10	0.13	-0.08	-0.14
	DKEFSI	**-0.54**	-0.29	-0.39	-0.39
	TMTB	-0.17	-0.19	-0.14	-0.33
	DSym	-0.35	-0.13	-0.31	-0.42
Calculation	GDA	-0.27	0.01	-0.24	-0.39
Visuoperceptual	VOSPFL	-0.20	-0.03	-0.10	-0.17
Visuospatial	VOSPNL	-0.11	-0.36	-0.31	-0.26
Social	Ekman	-0.11	-0.42	-0.14	-0.10

B) P values

Domain	Test	UWDRS-P	ModNPI	PHQ9	GAD7
Abstract reasoning	MRT	0.12	0.46	0.14	0.08
Language	NART	0.21	0.04*	0.08	0.04*
	GNT	0.10	0.02*	0.25	0.07
Memory	RMTF	0.84	0.43	0.56	0.24
	RMTW	0.14	0.17	0.65	0.59
	CPAL	0.56	0.69	0.42	0.91
Processing speed	TMTA	0.88	0.84	0.21	0.08
Executive function	DSB	0.28	0.08	0.95	0.08
	FAS	0.63	0.46	0.63	0.88
	Animals	0.59	0.49	0.66	0.43
	DKEFSI	**<0.001*** **	0.11	0.02*	0.02*
	TMTB	0.33	0.30	0.43	0.06
	DSym	0.04*	0.48	0.07	0.01
Calculation	GDA	0.13	0.94	0.18	0.03
Visuoperceptual	VOSPFL	0.26	0.85	0.58	0.33
Visuospatial	VOSPNL	0.55	0.04*	0.07	0.13
Social	Ekman	0.54	0.01*	0.44	0.58

*Correlation coefficients between Z scores for each cognitive test and psychiatric rating scale scores (A) and their corresponding P values (B) are shown * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GAD7, generalised anxiety disorder assessment 7; GDA, graded difficulty arithmetic; GNT, graded naming test; ModNPI, modified neuropsychiatric inventory; MRT, matrix reasoning test; NART, national adult reasoning test; PHQ9, patient health questionnaire 9; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; UWDRS-P, unified Wilson's disease rating scale psychiatric subscore; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

Table 3-14 Associations between ModNPI items and cognitive test scores

A) Spearman's correlation coefficients

	MRT	NART	GNT	RMTF	RMTW	CPAL	TMTA	DSB	FAS
Delusions	-0.20	-0.27	-0.29	-0.15	0.10	0.06	-0.05	0.03	NA
Hallucinations	NA	NA	NA	NA	NA	NA	NA	NA	NA
Agitation	-0.10	-0.17	-0.34	-0.12	0.03	-0.06	-0.05	-0.15	-0.09
Depression	-0.27	-0.30	**-0.53**	-0.11	-0.12	-0.06	0.07	0.01	-0.11
Anxiety	-0.14	-0.44	-0.46	-0.11	-0.18	0.00	0.00	-0.06	0.02
Elation	-0.20	-0.36	-0.38	-0.25	0.05	-0.16	-0.25	-0.07	-0.29
Irritability	-0.07	-0.12	-0.29	0.03	-0.10	-0.10	0.10	-0.22	-0.09
Motor	-0.07	-0.03	-0.06	0.10	0.01	0.29	0.04	-0.07	-0.11
Sleep	0.06	-0.10	-0.17	-0.02	-0.12	-0.10	0.19	-0.12	0.03
Hypersexuality	NA	NA	NA	NA	NA	NA	NA	NA	NA
Hyperreligiosity	NA	NA	NA	NA	NA	NA	NA	NA	NA
Humour	-0.15	-0.17	-0.23	-0.26	-0.13	-0.12	-0.08	-0.16	-0.32
Apathy	-0.37	-0.26	-0.26	-0.13	-0.19	0.16	-0.05	-0.05	-0.36
Disinhibition	0.06	-0.20	-0.29	-0.07	-0.10	0.12	0.12	-0.22	-0.06
Eating habits	0.03	-0.04	-0.03	-0.18	-0.16	0.36	-0.05	0.03	-0.09
Empathy	0.18	-0.08	0.08	-0.08	-0.05	0.11	0.10	-0.05	-0.01
Compulsive	-0.03	-0.12	-0.09	0.17	0.09	-0.20	-0.23	-0.28	-0.07
Social/emotion	0.25	0.06	0.06	-0.16	-0.14	0.23	0.00	0.03	-0.01
Trusting	-0.39	-0.49	-0.43	-0.16	-0.07	-0.20	-0.21	-0.52	-0.37
Pain	-0.17	-0.34	-0.16	-0.10	-0.03	-0.15	-0.25	-0.13	-0.12

	Animals	DKEFSI	TMTB	DSym	GDA	VOSPFL	VOSPNL	Ekman
Delusions	NA	NA	-0.19	-0.08	NA	0.08	-0.15	-0.24
Hallucinations	NA	NA	NA	NA	NA	NA	NA	NA
Agitation	0.31	0.31	-0.29	-0.27	0.09	0.01	-0.08	-0.22
Depression	-0.14	-0.14	0.03	-0.09	-0.05	-0.05	-0.11	-0.11
Anxiety	-0.15	-0.15	-0.16	-0.17	-0.06	-0.20	0.00	-0.27
Elation	-0.07	-0.07	-0.35	-0.28	-0.16	0.12	-0.01	-0.26
Irritability	0.14	0.14	-0.07	-0.08	-0.05	0.03	-0.37	-0.18
Motor	-0.26	-0.26	0.06	0.07	-0.23	-0.22	-0.10	0.00
Sleep	0.31	0.31	0.13	0.05	0.01	-0.09	-0.19	-0.26
Hypersexuality	NA	NA	NA	NA	NA	NA	NA	NA
Hyperreligiosity	NA	NA	NA	NA	NA	NA	NA	NA
Humour	0.09	0.09	-0.31	-0.18	-0.02	0.17	0.15	0.04
Apathy	-0.07	-0.07	-0.26	-0.29	0.10	0.25	0.03	-0.25
Disinhibition	-0.07	-0.07	-0.12	-0.03	-0.11	-0.47	0.00	-0.08
Eating habits	-0.15	-0.15	-0.08	0.06	0.26	0.06	-0.17	-0.34
Empathy	0.25	0.25	-0.08	0.07	0.10	0.04	0.02	-0.05
Compulsive	-0.10	-0.10	-0.26	0.01	0.00	0.19	-0.16	0.24
Social/emotion	0.10	0.10	-0.05	0.19	0.39	-0.17	0.12	-0.24
Trusting	-0.01	-0.01	-0.31	-0.34	-0.16	-0.03	-0.13	-0.40
Pain	-0.01	-0.01	-0.09	-0.29	-0.28	0.06	-0.05	-0.11

(B) P values

	MRT	NART	GNT	RMTF	RMTW	CPAL	TMTA	DSB	FAS
Delusions	0.27	0.15	0.12	0.39	0.58	0.76	0.80	0.87	NA
Hallucinations	NA	NA	NA	NA	NA	NA	NA	NA	NA
Agitation	0.58	0.38	0.07	0.51	0.86	0.75	0.78	0.42	0.62
Depression	0.12	0.10	**0.002****	0.55	0.52	0.75	0.71	0.96	0.54
Anxiety	0.42	0.01*	0.01*	0.54	0.31	0.98	0.99	0.72	0.90
Elation	0.25	0.05	0.04*	0.16	0.77	0.37	0.17	0.71	0.11
Irritability	0.69	0.52	0.12	0.86	0.57	0.60	0.59	0.22	0.64
Motor	0.70	0.87	0.73	0.59	0.94	0.11	0.83	0.69	0.55
Sleep	0.74	0.62	0.36	0.90	0.50	0.56	0.30	0.49	0.88
Hypersexuality	NA	NA	NA	NA	NA	NA	NA	NA	NA
Hyperreligiosity	NA	NA	NA	NA	NA	NA	NA	NA	NA
Humour	0.40	0.38	0.22	0.14	0.45	0.50	0.67	0.38	0.07
Apathy	0.03*	0.16	0.17	0.46	0.29	0.38	0.78	0.79	0.04*
Disinhibition	0.73	0.29	0.12	0.71	0.58	0.51	0.51	0.22	0.73
Eating habits	0.86	0.82	0.88	0.32	0.35	0.04*	0.78	0.86	0.64
Empathy	0.32	0.67	0.68	0.66	0.77	0.55	0.59	0.76	0.96
Compulsive	0.85	0.51	0.63	0.34	0.61	0.28	0.19	0.12	0.72
Social/emotion	0.16	0.75	0.75	0.37	0.43	0.20	1.00	0.85	0.94
Trusting	0.02*	0.006**	0.02*	0.37	0.68	0.27	0.23	0.00	0.04*
Pain	0.33	0.06	0.41	0.58	0.87	0.40	0.16	0.46	0.53

	Animals	DKEFSI	TMTB	DSym	GDA	VOSPFL	VOSPNL	Ekman
Delusions	NA	NA	0.30	0.64	NA	0.65	0.39	0.17
Hallucinations	NA	NA	NA	NA	NA	NA	NA	NA
Agitation	0.08	0.08	0.11	0.13	0.63	0.94	0.65	0.20
Depression	0.43	0.43	0.88	0.61	0.77	0.80	0.54	0.55
Anxiety	0.42	0.42	0.36	0.34	0.75	0.26	0.99	0.12
Elation	0.71	0.71	0.05*	0.12	0.38	0.52	0.97	0.14
Irritability	0.46	0.46	0.69	0.65	0.79	0.87	0.03*	0.31
Motor	0.15	0.15	0.74	0.69	0.22	0.21	0.56	0.99
Sleep	0.08	0.08	0.48	0.79	0.95	0.60	0.29	0.14
Hypersexuality	NA	NA	NA	NA	NA	NA	NA	NA
Hyperreligiosity	NA	NA	NA	NA	NA	NA	NA	NA
Humour	0.61	0.61	0.08	0.33	0.92	0.34	0.39	0.81
Apathy	0.71	0.71	0.14	0.11	0.59	0.15	0.87	0.16
Disinhibition	0.69	0.69	0.50	0.87	0.55	0.00	0.99	0.65
Eating habits	0.40	0.40	0.67	0.75	0.16	0.73	0.34	0.05*
Empathy	0.16	0.16	0.66	0.71	0.60	0.81	0.91	0.80
Compulsive	0.60	0.60	0.15	0.96	0.98	0.28	0.37	0.16
Social/emotion	0.58	0.58	0.76	0.29	0.03	0.34	0.50	0.17
Trusting	0.94	0.94	0.08	0.05	0.39	0.86	0.46	0.02*
Pain	0.98	0.98	0.62	0.10	0.13	0.75	0.80	0.53

*Correlation coefficients between Z scores for each cognitive test and ModNPI scores (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

Health-related quality of life

Group differences in SF-36 scores are shown in **Table 3-15** and **Figure 3-13**. There were no differences in any scores after correction for multiple testing. Scores for physical functioning were lower in patients with neurological and hepatic presentations and, for role limitations related to physical functioning, scores were lower in patients with active than stable disease without correction for multiple testing. Two outliers, with Z scores for physical functioning of -5.5 and -3.9, were noted among patients with hepatic presentations. On review, these patients had some of the lowest UWDRS-N scores (0 and 2) and highest self-reported depression scores on the PHQ9 questionnaire (24 and 24) in the cohort. Both reported bilateral knee arthralgia on review of systems during their research visit. On reassessing group differences after excluding these two values, physical functioning scores were lower in patients with neurological than hepatic presentations (-0.2 vs 0.5, P=0.004).

Associations between UWDRS-N and SF-36 scores, with and without exclusion of outliers, are shown in **Table 3-16**. UWDRS-N scores and subscores for speech, dystonia, and parkinsonism negatively correlated with physical function scores are correcting for multiple comparisons only when outliers were excluded. There were no other associations between UWDRS-N and SF-36 scores.

Associations between cognitive test scores and SF-36 scores, with and without exclusion of outliers, are shown in **Table 3-17**. NART and GNT scores correlated with social functioning scores after correction for multiple testing. There were no other associations when outliers were excluded.

Associations between psychiatric scales and SF-36 scores are shown in **Table 3-18**. UWDRS-P, ModNPI, PHQ9 and GAD7 scores were each associated with role limitations (physical), bodily pain, general health perceptions, energy/vitality, social functioning, role limitations (emotional) and mental health with correction for multiple testing with the exceptions that UWDRS-P scores did not correlate with general health perceptions and ModNPI scores did not correlate with role limitations.

Associations between SF-36 scores and individual ModNPI items are shown in **Table 3-19**. Items for depression and anxiety were negatively

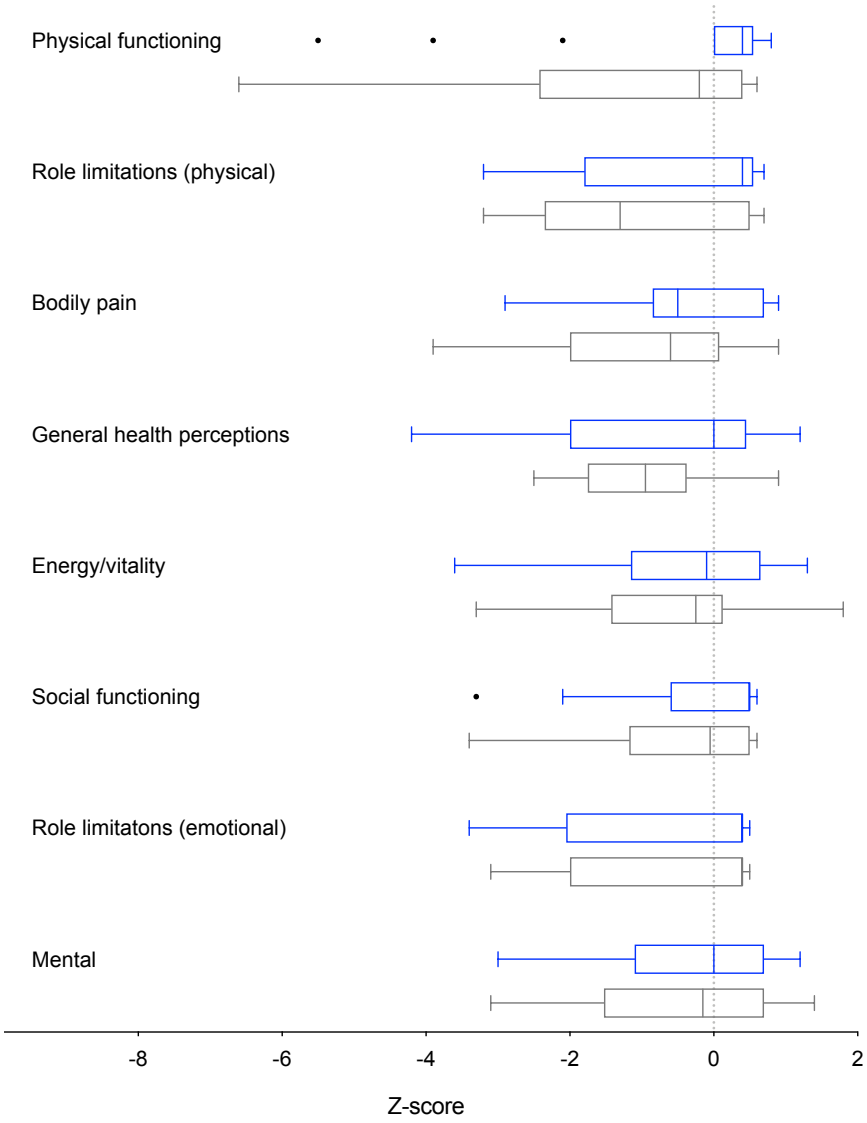
correlated with scores for social functioning, role limitations due to emotional health and mental health with correction for multiple testing and with most other SF-36 domains without correction for multiple testing.

Table 3-15 Group differences in SF-36 scores

Domain	**All (n=39)** Median [IQR]	**Hepatic (n=17)** Median [IQR]	**Neurological (n=22)** Median [IQR]	**P value**	**Stable (n=35)** Median [IQR]	**Active (n=4)** Median [IQR]	**P value**
Physical functioning	0.4 [-1.7, 0.5]	0.4 [0.0, 0.5]	-0.2 [-1.9, 0.4]	0.03*	0.4 [-1.3, 0.5]	-1.3 [-2.4, -0.5]	0.13
Role limitations (physical)	0.4 [-2.2, 0.5]	0.4 [-1.3, 0.5]	-1.3 [-2.3, 0.5]	0.29	0.4 [-1.6, 0.5]	-2.4 [-2.6, -2.1]	0.04*
Bodily pain	-0.5 [-1.1, 0.6]	-0.5 [-0.7, 0.7]	-0.6 [-1.9, -0.1]	0.33	-0.5 [-0.9, 0.7]	-1.8 [-2.2, -1.2]	0.07
General health perceptions	-0.8 [-1.8, 0.1]	0.0 [-1.8, 0.3]	-1.0 [-1.7, -0.4]	0.54	-0.7 [-1.8, 0.1]	-1.6 [-1.7, -1.1]	0.46
Energy/vitality	-0.1 [-1.2, 0.3]	-0.1 [-0.9, 0.6]	-0.3 [-1.3, 0.1]	0.31	-0.1 [-1.2, 0.5]	-0.5 [-1.2, -0.1]	0.53
Social functioning	0.1 [-1.0, 0.5]	0.5 [-0.2, 0.5]	-0.1 [-1.0, 0.5]	0.51	0.1 [-1.0, 0.5]	-1.0 [-1.6, -0.7]	0.07
Role limitations (emotional)	0.4 [-1.9, 0.4]	0.4 [-1.9, 0.4]	0.4 [-1.7, 0.4]	0.91	0.4 [-1.9, 0.4]	-1.0 [-2.5, 0.4]	0.33
Mental health	-0.1 [-1.2, 0.7]	0.0 [-1.1, 0.7]	-0.2 [-1.5, 0.7]	0.68	0.0 [-1.1, 0.7]	-1.6 [-2.0, -1.0]	0.08

*Group differences in SF-36 scores given as Z scores are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. IQR, interquartile range; SF-36, 36-item short form survey.*

Figure 3-13 Group differences in SF-36 scores



Box-and-whisker (Tukey) plots of Z scores for SF-36 scores from patients with hepatic presentations (blue) and neurological presentations (grey) are shown.

Table 3-16 Associations between SF-36 and UWDRS-N scores

A) Spearman's correlation coefficients

	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
Speech	-0.49 (-0.59)	-0.02	-0.30	-0.19	-0.10	-0.21	0.02	-0.11
Dystonia	-0.40 (-0.53)	-0.09	-0.17	-0.12	-0.26	-0.05	-0.02	-0.06
Parkinsonism	-0.36 (-0.45)	-0.19	-0.14	-0.08	-0.19	-0.13	-0.01	-0.08
Tremor	0.13 (0.08)	-0.20	0.12	-0.09	-0.17	-0.10	-0.12	-0.01
Ataxia	-0.27 (-0.26)	-0.06	-0.04	0.06	-0.09	0.04	0.17	0.13
Chorea	-0.18 (-0.15)	-0.01	0.10	-0.08	-0.19	-0.19	-0.07	-0.11
Rigidity	-0.08 (-0.23)	-0.10	0.13	0.12	-0.02	0.06	0.13	0.13
UWDRS-N	-0.28 (-0.41)	-0.14	-0.03	-0.08	-0.21	-0.12	-0.01	-0.05

B) P values

	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
Speech	**0.003**** (**<0.001***)	0.92	0.08	0.28	0.58	0.23	0.92	0.54
Dystonia	0.02* (**0.002***)	0.60	0.34	0.48	0.14	0.77	0.92	0.74
Parkinsonism	0.03* (**0.008**)	0.27	0.42	0.66	0.28	0.45	0.98	0.67
Tremor	0.45 (0.67)	0.24	0.48	0.61	0.33	0.58	0.50	0.93
Ataxia	0.11 (0.03*)	0.74	0.80	0.75	0.61	0.80	0.34	0.47
Chorea	0.30 (0.15)	0.94	0.57	0.63	0.28	0.28	0.68	0.53
Rigidity	0.64 (0.41)	0.57	0.46	0.47	0.91	0.73	0.45	0.44
UWDRS-N	0.10 (**0.02***)	0.42	0.89	0.66	0.23	0.49	0.96	0.81

*Correlation coefficients between UWDRS-N subscores and SF-36 scores (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Correlation coefficients and P values after excluding outliers are included in parentheses.*

Table 3-17 Associations between SF-36 and cognitive test scores

A) Spearman's correlation coefficients

Domain	Test	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
Abstract reasoning	MRT	**0.49** (0.48)	0.16	0.36	0.22	0.28	0.19	-0.02	0.10
Language	NART	**0.50** (0.46)	0.37	0.40	0.21	0.27	**0.53**	0.33	0.26
	GNT	0.47 (0.42)	0.50	0.40	0.19	0.33	**0.51**	0.47	0.41
Memory	RMTF	0.39 (0.43)	0.13	0.16	0.20	0.23	0.01	-0.01	0.17
	RMTW	0.13 (0.16)	0.05	0.10	0.18	0.08	0.07	-0.10	0.06
	CPAL	0.09 (0.12)	0.05	0.25	-0.12	-0.02	0.03	0.12	0.15
Processing speed	TMTA	0.20 (0.29)	0.19	0.15	0.25	0.21	0.16	0.09	0.14
Executive function	DSB	0.23 (0.21)	0.05	0.12	-0.01	0.01	0.16	0.07	0.13
	FAS	0.38 (0.38)	0.06	0.25	0.25	0.05	0.07	-0.16	-0.12
	Animals	0.00 (0.08)	0.08	-0.19	-0.11	-0.17	0.18	-0.07	-0.15
	DKEFSI	0.26 (0.22)	0.18	0.30	0.22	0.26	0.39	0.25	0.26
	TMTB	0.15 (0.12)	0.25	0.12	0.13	0.06	0.09	-0.04	0.02
	DSym	0.31 (0.24)	0.25	0.33	0.15	0.22	0.34	0.22	0.11
Calculation	GDA	0.31 (0.32)	0.33	0.24	0.05	0.12	0.36	0.26	0.05
Visuoperceptual	VOSPFL	-0.03 (0.01)	-0.07	0.09	0.01	-0.16	0.30	0.18	0.06
Visuospatial	VOSPNL	0.17 (0.19)	0.02	0.23	0.14	0.31	0.25	0.16	0.29
Social	Ekman	0.22 (0.25)	-0.17	0.12	0.11	0.07	0.09	-0.17	0.05

B) P values

Domain	Test	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
Abstract reasoning	MRT	**0.003**** (0.005)	0.36	0.04*	0.20	0.10	0.27	0.91	0.55
Language	NART	**0.005**** (0.01)	0.04*	0.03*	0.27	0.15	**0.002****	0.07	0.17
	GNT	0.009** (0.02)	0.004**	0.03*	0.32	0.08	**0.004****	0.008**	0.03*
Memory	RMTF	0.02* (0.01)	0.45	0.37	0.26	0.18	0.97	0.96	0.33
	RMTW	0.46 (0.37)	0.76	0.56	0.30	0.65	0.70	0.57	0.73
	CPAL	0.62 (0.52)	0.79	0.15	0.49	0.93	0.88	0.49	0.39
Processing speed	TMTA	0.25 (0.11)	0.27	0.40	0.15	0.24	0.37	0.60	0.43
Executive function	DSB	0.18 (0.25)	0.79	0.51	0.94	0.94	0.37	0.69	0.47
	FAS	0.03* (0.04)	0.74	0.17	0.16	0.76	0.69	0.39	0.50
	Animals	0.99 (0.67)	0.65	0.30	0.53	0.33	0.31	0.70	0.41
	DKEFSI	0.14 (0.22)	0.30	0.08	0.20	0.14	0.02*	0.15	0.14
	TMTB	0.40 (0.52)	0.16	0.51	0.45	0.73	0.62	0.81	0.93
	DSym	0.07 (0.19)	0.15	0.05	0.41	0.20	0.05*	0.21	0.54
Calculation	GDA	0.08 (0.08)	0.07	0.18	0.80	0.51	0.04*	0.15	0.80
Visuoperceptual	VOSPFL	0.85 (0.97)	0.70	0.62	0.98	0.36	0.08	0.30	0.75
Visuospatial	VOSPNL	0.33 (0.30)	0.90	0.19	0.42	0.07	0.14	0.37	0.09
Social	Ekman	0.21 (0.16)	0.33	0.50	0.52	0.69	0.61	0.33	0.78

*Correlation coefficients between Z scores for each cognitive test and SF-36 scores (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Correlation coefficients and P values after excluding outliers are included in parentheses. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

Table 3-18 Associations between SF-36 scores and psychiatric scales

A) Spearman's correlation coefficients

	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
UWDRS-P	-0.32 (-0.20)	**-0.39**	**-0.45**	-0.22	**-0.46**	**-0.53**	**-0.49**	**-0.57**
ModNPI	**-0.38** (-0.27)	-0.21	**-0.36**	**-0.44**	**-0.37**	**-0.52**	-0.31	**-0.58**
PHQ9	**-0.50** (-0.40)	**-0.51**	**-0.61**	**-0.57**	**-0.80**	**-0.78**	**-0.67**	**-0.78**
GAD7	**-0.49** (-0.39)	**-0.39**	**-0.48**	**-0.43**	**-0.56**	**-0.68**	**-0.60**	**-0.71**

B) P values

	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
UWDRS-P	0.06 (0.25)	**0.02***	**0.006****	0.20	**0.006****	**<0.001****	**0.003****	**<0.001*****
ModNPI	**0.03*** (0.13)	0.23	**0.04***	**0.009****	**0.03***	**<0.002****	0.08	**<0.001*****
PHQ9	**0.003**** (0.02*)	**0.002****	**<0.001*****	**<0.001*****	**<0.001*****	**<0.001*****	**<0.001*****	**<0.001*****
GAD7	**0.003**** (0.03*)	**0.02***	**0.004****	**0.01***	**<0.001*****	**<0.001*****	**<0.001*****	**<0.001*****

*Correlation coefficients between psychiatric rating scales and SF-36 scores (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Correlation coefficients and P values after excluding outliers are included in parentheses. GAD7, generalised anxiety disorder assessment 7; ModNPI, modified neuropsychiatric inventory; PHQ9, patient health questionnaire 9; UWDRS-P, unified Wilson's disease rating scale psychiatric subscore.*

Table 3-19 Associations between SF-36 and ModNPI items

A) Spearman's correlation coefficients

Test	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
Delusions	-0.24 (-0.26)	-0.28	-0.22	0.04	-0.26	-0.29	-0.18	-0.20
Hallucinations	NA	NA	NA	NA	NA	NA	NA	NA
Agitation	-0.45 (-0.35)	-0.27	-0.30	-0.20	-0.29	-0.18	-0.19	-0.32
Depression	-0.35 (-0.23)	-0.32	-0.42	-0.31	-0.44	**-0.57**	**-0.47**	**-0.62**
Anxiety	-0.30 (-0.18)	-0.49	-0.38	-0.40	-0.34	**-0.61**	**-0.52**	**-0.60**
Elation	-0.16 (-0.19)	-0.18	-0.07	0.22	0.03	-0.15	-0.08	0.06
Irritability	-0.23 (-0.15)	-0.06	-0.21	-0.13	-0.19	-0.19	-0.06	-0.27
Motor	-0.15 (-0.18)	-0.24	-0.22	-0.23	0.03	-0.20	0.10	0.12
Sleep	-0.14 (-0.04)	-0.19	-0.27	-0.31	-0.29	-0.37	-0.35	-0.40
Hypersexuality	NA	NA	NA	NA	NA	NA	NA	NA
Hyperreligiosity	NA	NA	NA	NA	NA	NA	NA	NA
Humour	-0.24 (-0.11)	-0.33	-0.18	0.06	0.05	0.10	-0.05	0.02
Apathy	-0.23 (-0.16)	0.00	0.01	-0.04	-0.25	0.02	0.13	-0.01
Disinhibition	-0.23 (-0.08)	-0.22	-0.17	-0.33	-0.07	-0.41	-0.32	-0.32
Eating habits	0.05 (0.01)	0.23	0.31	0.02	0.12	-0.06	0.16	0.10
Empathy	-0.15 (0.04)	0.01	-0.10	-0.29	-0.04	-0.17	-0.03	-0.19
Compulsive	0.12 (0.09)	0.09	0.22	-0.09	0.08	0.07	0.08	-0.06
Social/emotional	0.14 (0.25)	-0.01	0.03	-0.26	-0.11	-0.22	-0.16	-0.31
Trusting	-0.28 (-0.24)	-0.13	-0.16	-0.20	-0.14	-0.16	-0.12	-0.15
Pain	-0.30 (-0.14)	-0.24	-0.43	-0.27	-0.23	-0.33	-0.36	-0.30

B) P values

Test	Physical functioning	Role limitations (physical)	Bodily pain	General health perceptions	Energy/ vitality	Social functioning	Role limitations (emotional)	Mental health
Delusions	0.17 (0.14)	0.11	0.21	0.84	0.14	0.09	0.32	0.25
Hallucinations	NA	NA	NA	NA	NA	NA	NA	NA
Agitation	0.007** (0.05)	0.12	0.08	0.26	0.10	0.31	0.29	0.06
Depression	0.04 (0.22)	0.07	0.01*	0.07	0.009**	**<0.001*** **	**0.006** **	**<0.001*** **
Anxiety	0.08 (0.33)	0.003**	0.03*	0.02*	0.05	**<0.001*** **	**0.002** **	**<0.001*** **
Elation	0.36 (0.29)	0.31	0.69	0.20	0.89	0.41	0.66	0.72
Irritability	0.18 (0.42)	0.73	0.24	0.48	0.28	0.27	0.72	0.12
Motor	0.39 (0.32)	0.16	0.21	0.20	0.85	0.25	0.58	0.49
Sleep	0.44 (0.85)	0.29	0.12	0.07	0.10	0.03*	0.04*	0.02*
Hypersexuality	NA	NA	NA	NA	NA	NA	NA	NA
Hyperreligiosity	NA	NA	NA	NA	NA	NA	NA	NA
Humour	0.18 (0.54)	0.06	0.32	0.74	0.78	0.59	0.77	0.91
Apathy	0.20 (0.38)	0.98	0.96	0.82	0.15	0.90	0.47	0.95
Disinhibition	0.19 (0.67)	0.20	0.32	0.06	0.70	0.02*	0.07	0.06
Eating habits	0.79 (0.96)	0.19	0.08	0.93	0.50	0.75	0.36	0.58
Empathy	0.41 (0.82)	0.96	0.58	0.09	0.81	0.34	0.86	0.28
Compulsive	0.48 (0.62)	0.61	0.22	0.60	0.67	0.71	0.67	0.72
Social/emotional	0.43 (0.17)	0.98	0.85	0.13	0.55	0.21	0.37	0.08
Trusting	0.10 (0.19)	0.45	0.38	0.25	0.43	0.35	0.50	0.39
Pain	0.08 (0.45)	0.17	0.01	0.12	0.19	0.06	0.04*	0.09

*Correlation coefficients between ModNPI items and SF-36 scores (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Correlation coefficients and P values after excluding outliers are included in parentheses.*

3.4 Discussion

I have characterised the movement disorder, cognitive deficits and psychiatric features in a cohort of 40 prospectively-recruited patients with WD. This has allowed me to confirm previous observations related to frequency and severity of various clinical features and report novel findings, particularly with regards to the pattern of cognitive deficits in patients with hepatic presentations. I have performed the most comprehensive analysis of the relationship between movement disorders, cognitive deficits and psychiatric features of the disease to date. Our observations also have implications for the classification of WD. Here, I put these results in context and discuss the opportunities and limitations of using this cohort for biomarker studies. I defer discussing the neuroanatomical correlates of movement disorders, cognitive deficits and psychiatric symptoms until the neuroimaging analyses in Chapter 5.

Cohort structure

The cohort is broadly representative of patients with WD based on the existing literature and balanced across key demographic and clinical characteristics between groups. The proportion of patients with neurological presentations, age of onset, frequency of cirrhosis and choice of treatment is commensurate with large retrospective studies from other European countries.[22,41,88,136,137] Where previous prospective studies have focussed on a specific aspect of WD, I have comprehensively characterised movement disorders, cognitive deficits and psychiatric features in parallel. This provides a rich dataset from which to test candidate biomarkers and understand the full spectrum of brain pathology associated with WD. There are, however, a number of important limitations and considerations when using this cohort in biomarker studies.

The first limitation is the small number of active cases. This reflects the challenges in recruiting patients with rare diseases to prospective studies. There were, to my knowledge, two other new diagnoses of WD in the UK during the recruitment period, one was not eligible to participate and the other declined. The small number of active cases limits our ability to draw conclusions about differences between patients with active and stable

disease; any lack of difference between these groups should be interpreted with care.

The combination of patients presenting *de novo* and those deteriorating with non-adherence also makes the active group heterogeneous. While it is likely that these patients share the same underlying pathophysiology, i.e. accumulation or re-accumulation of copper in the brain, patients who deteriorate with non-adherence are inherently likely to be older with longer disease duration. In classifying deteriorating, non-adherent patients as having active disease I may also be introducing a selection bias for patients with cognitive impairment or specific psychiatric features. Nonetheless, I suspect distinguishing between chronically-treated patients who are stable and those who are deteriorating serves an important purpose for biomarker discovery and has not, to my knowledge, been attempted in any previous studies.

In a similar vein, I included both patients with neurological and hepatic presentations in the stable group. As discussed later in this chapter, our data supports the idea that there to be a continuum of subclinical neurological involvement in WD and classifying patients into neurological and hepatic presentations is essentially choosing a cut-off to divide patients into two easily recognisable groups at one time point early in the disease course. Most of the patients with hepatic presentations had subtle neurological or cognitive deficits. It would therefore seem counter-intuitive to exclude these patients unless there were specific concerns regarding the distribution of UWDRS-N scores with, for example, a bimodal distribution but this was not the case.

Another consideration is the age distribution of our cohort. With a mean age of 44 years and disease duration of 24 years, patients with stable disease represent a group of patients who have been receiving maintenance therapy for many years. This is not itself a problem as long as age is included as an independent variable when testing candidate biomarkers, as is usual practice, and, in some respects, is an advantage: I am confident that participants classified as having stable disease have received treatment for long enough to have been adequately ‘de-coppered’ and therefore have an arrested disease state. Associations between clinical rating scales and candidate biomarkers in stable patients can therefore be used to identify prognostic/predictive biomarkers and, with imaging, to identify

neuroanatomical correlates for specific clinical features. Importantly for our biomarker studies, I did not find any associations between age or disease duration and clinical rating scales in stable patients.

The extent to which I characterised the severity of ongoing liver disease was limited. I am aware of which patients are suspected to have compensated (asymptomatic/uncomplicated) cirrhosis on the basis of previous investigations but did not re-evaluate the severity of liver disease during research visits beyond performing clinical examination and blood tests. The gold standard for confirming the presence of cirrhosis is liver biopsy, which is clearly not justified in the context of our observational study. Transient elastography is a non-invasive method for diagnosing cirrhosis but, at a unit cost of up to £400,[138] is expensive. The FIB-4 score, which is based on a combination of age, aspartate transaminase (AST), alanine transaminase (ALT) and platelet count, has recently been shown to be a useful surrogate marker for cirrhosis in WD,[139] however I am unable to retrospectively calculate this without AST values.

There was only one patient with features of decompensated liver disease in the cohort. This limits the extent to which any findings can be generalised to newly-diagnosed patients with WD in whom these are more common.

I did not classify participants according to whether they experienced paradoxical neurological worsening. This complex and poorly understood phenomenon likely influences neurological outcomes,[140] and should ideally be incorporated into analyses in biomarker studies. I am aware of several patients from our cohort who undoubtedly experienced neurological worsening but many more who describe a mild worsening that is difficult to quantify or discern from underlying disease progression retrospectively. Indeed, Brewer *et al* reported that more than half of patients with neurological symptoms reported worsening in at least some of their symptoms after treatment initiation.[64] Prospectively measuring UWDRS-N scores in the first year of treatment would be the most effective way to characterise paradoxical neurological worsening but was not feasible in our study.

Conducting numerous statistical tests to compare the severity of movement disorders, cognitive deficits and psychiatric features and test their

associations with a range of candidate biomarkers creates issues around multiple comparisons. I correct for this using the false discovery rate here and in subsequent individual analyses. This is less stringent than other methods, such as Bonferroni correction, but reasonable given the exploratory nature of the work.

Finally, I did not recruit healthy controls to the CROWD study due to the study timeline and budget. This limits our ability to interpret the significance of some clinical findings where normative data is not available but does not impede our ability to identify whether candidate biomarkers measure neurological involvement within our cohort. As discussed in Chapter 4, biosamples collected from healthy controls from other cohorts were made available for wet (fluid) biomarkers studies.

Movement disorders

The neurological endophenotypes within our cohort are consistent with the previous literature. Speech disturbance, dystonia, parkinsonism and tremor were the most common findings on neurological examination.[38,42,54] UWDRS-N scores were higher in patients with neurological presentations, as expected, and they were also higher in those with active than stable disease. UWDRS-N scores therefore need to be included as a covariate in comparisons between active and stable patients in biomarker studies.

The breakdown of UWDRS-N scores by specific movement disorders illustrates that most patients with neurological involvement had mixed, overlapping movement disorders, as opposed to distinct phenotypes. There was also a high degree of correlation between scores for speech, dystonia, parkinsonism and ataxia but not tremor. This makes classifying our participants according to the dominant phenotype, as has previously been suggested,[38,90-92] challenging. It will also make it difficult to identify neuroimaging correlates for specific neurological signs within a relatively small cohort.

Two unexpected findings might raise questions as to whether participants have been classified accordingly or whether our cohort is representative of patients with WD and require further discussion:

Firstly, some patients with hepatic presentations had mild neurological signs when examined at their research visit. These patients were not included in the neurological group because we used the international consensus classification that determines phenotype according to the clinical symptoms and signs at initial presentation. This pragmatic approach leads to a somewhat artificial dichotomy between neurological and hepatic phenotypes. It is possible that some intermittent and subclinical neurological features at initial presentation may not have been attributed to WD by patients in our cohort, or identified by their hepatologists. An alternative explanation is that these signs represent disease progression; Merle *et al* previously reported that up to 20% of patients with hepatic presentation develop neurological signs later in the disease course.[12] Finally, some of these signs could represent normal variation; most older adults in the general population score at least two points on part III of the UPDRS.[141] This could be tested by performing the UWDRS in a cohort of age-matched healthy controls.

Secondly, there were no differences in scores for tremor, the most common movement disorder at presentation,[38] between our groups. This appears to be related to both higher scores in patients with hepatic presentations and lower scores in those with neurological presentations. However, 76% of patients with neurological presentations reported tremor at onset suggesting that tremor was present but then improved or resolved with treatment in most. This supports the findings of Burkhe *et al* that tremor is more likely to respond to chelation therapy than dystonia,[96] and highlights that the relative frequency of specific neurological phenotypes may differ between patients presenting *de novo* and those who have been chronically-treated. The lack of association between the severity of tremor and other movement disorders in chronically-treated patients with WD is also interesting. It suggests that these are driven by dysfunction or damage in distinct brain regions or networks.

These observations may have important implications for cohort stratification and the interpretation of UWDRS-N scores in clinical trials for WD. Tremor accounts for a maximum of 44 points in the UWDRS-N, compared to 28 for dystonia and 4 for speech. A disproportionate number of patients with disabling tremor in one treatment group may mean that UWDRS-N scores in

this group are more likely to improve, irrespective of the relative efficacy of the treatment.

Cognitive deficits

I assessed cognitive function in our participants using a comprehensive battery of neuropsychological tests. With the exception of the study by Seniow *et al*,[44] this is largest cohort of patients with WD to undergo detailed cognitive testing and the only study to examine associations with UWDRS-N scores across multiple tests. Unlike most previous studies I adjusted test scores for age, sex and/or educational level using normative data and corrected for multiple comparisons.

Group differences in test scores confirm previous observations that patients with neurological presentations have deficits in abstract reasoning (MRT),[44,105] processing speed (TMTA),[97] and visuospatial function (VOSNL).[101,104,107] Scores for several tests of executive function were lower in patients with neurological than hepatic presentations and lower in patients with active than stable disease without correction for multiple testing. I demonstrate that UWDRS-N scores, including subscores for speech, dystonia, parkinsonism and ataxia, are strongly correlated with measures of abstract reasoning, executive function, processing speed, memory for faces and calculation in chronically-treated patients. The lack of associations between certain neurological phenotypes and cognitive test scores is also interesting. There were no associations between tremor scores and any cognitive test. Although 19% of participants had poor performance in paired associate learning (CPAL), there was no association with UWDRS-N scores suggesting that some cognitive deficits occur independently of movement disorders in WD.

Of tests included in the battery, MRT scores were the most strongly correlated with the severity of movement disorders. Using the MRT as a marker of cognitive impairment in WD more broadly may be advantageous for several additional reasons. Firstly, deficits in abstract reasoning are consistently reported in WD and have been reported to improve with treatment.[99,112,113] Secondly, it is an untimed, multiple-choice test and so not influenced by processing speed, speech impairment or handwriting ability.

Finally, it correlates with imaging biomarkers in other neurodegenerative diseases: Schott *et al* found that deteriorations in MRT scores were more closely associated with progression of brain atrophy in Alzheimer's disease than any other cognitive test.[142] They offer the explanation that this test 'draws upon a wide range of skills supported by a highly distributed network of cognitive representations' and 'has a wide dynamic range with graded variation in item difficulty'.

Our observations on processing speed need to be carefully considered given TMTA is a timed test that is highly dependent on handwriting ability. Indeed, post hoc analysis comparing scores between patients with neurological and hepatic presentations showed that there was no difference when controlling for writing scores in the dominant hand. This is consistent with the suggestion by Littman *et al* that low scores on tests of processing speed primarily reflect motor impairment.[109] In an elegant experiment where they asked subjects to perform a scanning tasks of gradually increasing complexity, patients with WD had delayed response latencies but the rate of information processing did not differ from healthy controls as the task became more challenging.

In the absence of healthy control group, the frequency of poor performance in each test is a useful measure of cognitive deficits in patients with hepatic presentations. Scores for the RMTF and TMTB in the group are particularly interesting with 4 of 17 (24%) of patients with hepatic presentations scoring more than two standard deviations below the mean. The difference between RMTF scores (median -0.8, IQR -1.9 to 0.4) and RMTW scores (0.9, IQR 0.6 to 1.0) in patients with hepatic presentations is also striking. Moreover, RMTF and TMTB, were the tests with the highest frequency of low performers among patients with neurological presentations. Poor performance on at least one cognitive test was common in patients with hepatic presentations with minimal neurological signs. It was also seen in a few patients with hepatic presentations without any ongoing neurological signs.

Tarter *et al* previously reported that patients with hepatic presentations had low TMTB and Block Design scores but the RMTF has not previously been applied to patients with WD.[108] Several other tests that draw on visual memory including the Rey's Complex Figure Test, WMS immediate and delayed figural

recall and Benton Visual Memory Test have been studied in WD and scores were reduced in patients with neurological presentation.[44,101,104] The Benton Visual Memory test is the only test of visual memory that has been applied in patients with hepatic presentations and scores were unremarkable.[44] I suspect RMTF scores were low in our participants because visual processing of faces is likely to be distinct from visual processing of the abstract stimuli used in other the tests.[143] It is tempting to attribute these deficits to impaired emotion recognition given Peyroux *et al* have previously demonstrated that patients with hepatic presentations have impaired facial emotion recognition, particularly for recognising fear, and many of the faces included in the RMTF are made memorable by their expression. However, I found no association between RMTF scores and Ekman scores (R=0.31, P=0.07) or the subscore for fear (R=008, P=0.65), suggesting that a distinct aspect of visual processing or memory for faces is affected.

It is worth highlighting that our findings in patients with hepatic presentations seem to contradict those of Seniow *et al* who did not find any cognitive deficits in patients without neurological symptoms. Importantly, they classified patients according the presence of neurological symptoms at the time of assessment, not presentation, and none of their asymptomatic group had neuroradiological abnormalities ensuring that any patients with hepatic presentations who may have accrued subtle neurological signs were excluded. In addition, they did not include the TMTB, RMTF or similar tests in their battery. They reconcile their findings with the earlier observations of Tarter *et al* by suggesting that the tests Tarter *et al* used may be more sensitive to early neuropsychological deficits or that the lower scores may represent subclinical encephalopathy associated with cirrhosis.

In the absence of a control group consisting of patients with other liver diseases we cannot exclude the possibility that observations in patients with hepatic presentations represent subtle cognitive deficits associated with cirrhosis in general. The term *minimal hepatic encephalopathy* has been used to describe subclinical cognitive impairment in patients with cirrhosis, typically with portosystemic shunting. However, this is unlikely to be a confounding factor in our cohort because the mean disease duration among patients with hepatic presentations was 20 years and none had features of decompensated

liver disease. In addition, I cannot exclude the possibility that chelation therapy or symptomatic treatments for neurological or psychiatric symptoms contribute to cognitive deficits in patients with WD.

Overall, our findings support the idea that subtle deficits in memory for faces, cognitive flexibility and associative learning represent an early neurological phenotype that emerges prior to the development of movement disorders and deficits in abstract reasoning and other aspects of executive function. However, it is important to remember that all of our patients with hepatic presentations were chronically-treated and the extent to which other cognitive domains might be affected in newly diagnosed, untreated patients with hepatic presentations remains poorly understood. It is plausible that patients with hepatic presentations develop subtle deficits in other cognitive domains but that deficits in memory for faces and cognitive flexibility persist because they are less amenable to treatment. The theory that certain cognitive phenotypes occur independently of movement disorders and the neuroanatomical basis for cognitive deficits in WD is discussed in much more detail in Chapter 5.

Psychiatric features

I have used a combination of interviews and examinations with participants (UWDRS-P), interviews with family members, friends or carers (ModNPI) and self-reported questionnaires (PHQ9 and GAD7) to characterise the psychiatric features in our cohort. Consistent with the previous literature, irritability, depression, agitation and anxiety were among the most common psychiatric phenotypes and participants demonstrated multiple, overlapping psychiatric symptoms.

Interestingly, there were no differences in scores for any of the psychiatric rating scales, or any of their specific items, between patients with neurological and hepatic presentations, with or without FDR correction. There were also no associations between psychiatric rating scales and UWDRS-N scores and, with the exception of UWDRS-P and DKEFSI scores, no associations between psychiatric rating scales and cognitive test scores. This might lead us to conclude that, unlike most cognitive deficits, psychiatric features occur independently of neurological involvement in WD but there are

several caveats to this assertion. Firstly, I did identify weak associations between ModNPI and speech scores and between PHQ9 and parkinsonism scores, which are consistent with earlier observations by Oder *et al* and may be more apparent in a larger cohort.[132] Secondly, the UWDRS-P and ModNPI may not be as sensitive at detecting psychiatric symptoms in WD as predicted; psychiatric symptoms are complex and more detailed clinical interviews led by a psychiatrist may be required. Condensing a heterogenous group of symptoms into a single score in this way may not be appropriate. Thirdly, I was only able to test associations between neurological and psychiatric features in chronically-treated patients and any associations may be stronger in untreated patients.

Comparisons between ModNPI scores and items between patients with active and stable disease yielded interesting results. Patients with active disease had higher scores for delusions, agitation, irritability, hypersexuality, hyperreligiosity, apathy and disinhibition. As previously mentioned, there may be an issue with selection bias here whereby chronically-treated patients with these symptoms are less likely to be adherent to medications and therefore develop active disease. Nonetheless, the observation that several patients with active disease had delusions is noteworthy given psychosis has previously been reported to be rare in WD.

I considered which rating scale would be most appropriate for quantifying psychiatric involvement in biomarker studies and clinical trials. A major concern with the UWDRS-P is that self-reported memory loss and difficulty concentrating, which are better assessed using cognitive tests, were two of the three most common items on this scale in our cohort and a further nine of the 17 remaining items in the UWDRS-P are also self-reported. While our modified version of the NPI has not been widely used or validated by comparison with more comprehensive psychiatric assessment, the original version of the NPI has been studied in WD and the additional items cover psychiatric symptoms that have previously been reported in WD.[144] The ModNPI showed a greater number of group differences and associations with specific neurological phenotypes than the UWDRS-P and would appear to be a more reliable screen for psychiatric symptoms in WD.

Quality of life

I asked participants to complete SF-36 questionnaires and compared scores across the eight SF-36 domains between groups and with clinical rating scales in the chronically-treated, stable patients. Unexpectedly, there were no group differences between SF-36 scores or associations with UWDRS-N scores or subscores. Clear differences in physical functioning scores between patients with neurological and hepatic presentations and associations with UWDRS-N scores became apparent when two outliers were excluded. I did not identify clear associations between SF-36 scores and cognitive deficits, with the exception of social functioning and language function, which is usually preserved in WD. UWDRS-P, ModNPI, PHQ9 and GAD7 scores negatively correlated with all SF-36 domains, except physical functioning, highlighting clinical importance of psychiatric involvement in WD.

Our findings are largely consistent with previous observations by Svetel *et al* and Camarata *et al*.[116,145] By testing associations between multiple psychiatric rating scales, including PHQ9, GAD7 and individual ModNPI items, I can add that depression and anxiety are the main psychiatric phenotypes that appear to influence health-related QoL in WD. This is an important observation because, unlike other psychiatric symptoms common in WD, depression and anxiety may be more amenable to symptomatic treatment. It also suggests that screening tests for depression and anxiety, which have been validated in numerous other diseases, may provide clinically-meaningful measures of psychiatric involvement in clinical trials.

Unlike Svetel *et al*, who found that lower MMSE scores were the most significant contributors to SF-36 scores in a multivariate analysis,[145] I did not find associations between cognitive deficits and QoL. The reasons for this are unclear. The MMSE is a composite measure of multiple cognitive domains that is relatively insensitive to executive dysfunction.[146] In contrast to tests that focus on a single cognitive domain, scores for the MMSE may only fall when there is a more marked cognitive impairment across multiple domains, which is more likely to affect QoL.

The decision to exclude two outlying physical functioning scores requires further discussion. This was based on the marked discrepancy with

UWDRS-N scores and presence of disabling arthralgia in these patients. I chose not to exclude other SF-36 scores from these two participants on the grounds that these patients had significant psychiatric co-morbidity and removing other scores may unnecessarily introduce bias. Degenerative arthritis of the knees is common in WD,[147] and, with the benefit of hindsight, it may have been helpful to include a visual analogue scale for measuring joint pain in order to measure rheumatological involvement in our participants. Nonetheless, these observations highlight that factors other than movement disorders, cognitive deficits and psychiatric symptoms may influence measures of QoL in WD.

Concluding remarks

I have comprehensively characterised the demographic and clinical characteristics in a cohort of patients with WD. Given the structure of our cohort, I can compare candidate biomarkers between patients with hepatic and neurological presentations and between those with active and stable neurological disease. I can also investigate their association with the severity of ongoing movement disorders, cognitive deficits and psychiatric features in chronically-treated, stable patients to determine their clinical significance and, using imaging, identify neuroanatomical correlates for neurological involvement.

I found that UWDRS-N scores were higher in patients with active than stable disease, as expected, and cognitive deficits, but less so psychiatric features, were associated with neurological severity. The UWDRS-N score therefore needs to be considered as a covariate when comparing patients with active and stable disease and also when testing association with cognitive scores in biomarker studies.

Our findings so far have broader implications for our understanding of neurological involvement in WD and how it should be classified. While it is convenient to consider neurological involvement as a binary phenomenon, and this resonates with patients (and their clinicians), I have provided evidence that many patients with hepatic presentations demonstrate subtle, or subclinical, cognitive and psychiatric features. It is plausible that some of these were evident at presentation but were not identified during routine clinical

assessment and that others developed later. We should therefore recognise the possibility that there is a spectrum of neurological involvement in WD and that the initial presentation represents a single time-point in the disease course after which the phenotype may evolve.

Finally, the observation that the vast majority of chronically-treated patients have stable disease, and the lack of association between any clinical features and disease duration, illustrate a simple yet crucial point about the disease course in WD: Unlike most neurodegenerative diseases, WD mimics a protracted monophasic illness. Symptoms progress until chelation therapy is initiated and then improve, persist or, in cases of paradoxical neurological worsening, deteriorate with treatment before they reach a plateau and stabilise in the longer term. The disease process in patients with neurological involvement can therefore be modelled as a brain injury. In the context of our biomarker studies, the classification of patients into those with and without neurological symptoms at presentation can serve as a useful indicator of whether a patient has sustained a significant brain injury or not. Participants who I considered to have active disease may have ongoing, evolving brain injury with a combination of reversible and irreversible neurological damage whereas patients with stable disease are likely to have predominantly irreversible neurological damage. Carefully stratifying the cohort in this way is an essential step for successfully identifying biomarkers for neurological involvement and interpreting their clinical significance in the following chapters.

4. Wet (fluid) biomarkers

4.1 Introduction

Copper indices are currently used as the main diagnostic and monitoring biomarkers for WD. While numerous studies have investigated their diagnostic accuracy, data on their use as monitoring biomarkers is more limited. Measuring the 24-hour urinary copper output (UCu) on at least an annual basis is recommended in the European Association for the Study of the Liver (EASL) guidelines on Wilson's disease.[59] The serum non-caeruloplasmin-bound copper (NCC) is suggested to be another useful parameter for monitoring. This chapter begins by discussing the evidence base for using various copper indices for monitoring and the extent to which they inform us about neurological disease activity before discussing other candidate biomarkers for neurological involvement, including the principles behind Single Molecule Array technology and tandem mass spectrometry.

24-hour urinary copper output

Copper can be quantified in biological samples using inductively coupled mass spectrometry or atomic absorption spectroscopy.[148] Data on the reliability of spot urine and morning (first) urine samples are conflicting,[149-151] and so 24-hour urine collections remain the method of choice for measuring urinary copper excretion. These can be performed while patients continue their medication, 'on treatment', or after 48 hours of treatment cessation, 'off treatment'.[152] Specialists vary in which they prefer and guidelines recognise both approaches.[59,61,153] There is some evidence to suggest that 'off treatment' collections are better for assessing compliance with chelation therapy.[152]

Treatment targets for the initial de-coppering phase of treatment are not provided in the guidelines. During maintenance therapy, they suggest that 'on treatment' UCu 'in the vicinity' of 3-8 µmol/24 hours and 'off treatment' UCu less than 1.6 µmol/day are considered adequate.[59,61,153] These recommendations are based on expert opinion and no supporting evidence is cited. The 'off treatment' target is problematic for the 20-50% of patients with a UCu in the target range before starting treatment.[56] This target is also notably higher than reference ranges for UCu in controls: In a study of 111 healthy

adults from the UK, Sieniawska *et al* reported a mean copper output of 0.34 μmol/24 hours with 95% confidence intervals of 0.21 to 0.72 μmol/24 hours.[154] Some argue a lower 'off treatment' target below 0.77 μmol/24 hours is more appropriate,[155] but there is no consensus on this issue.

Only a few studies have investigated the relationship between UCu and neurological involvement or changes in UCu with chelation therapy in detail. Czlonkowska *et al* found no association with UWDRS-N scores in 53 newly-diagnosed patients demonstrating that UCu does not reflect neurological severity.[38] Walshe reported UCu over the first two years of treatment in 192 patients, most of whom received penicillamine, seen between 1955 and 2000.[69] Median UCu in the neurological group at baseline was 4.20 μmol/24 hours with standard deviation 2.72 μmol/24 hours. After administration of a single 500 mg dose of penicillamine, median UCu increased to 13.64 μmol/24 hours with standard deviation 10.63 μmol/24 hours. 'Off treatment' UCu decreased considerably to median 0.79 μmol/24 hours after 12 months of treatment. These findings suggest high inter-individual variation in UCu before treatment and very high inter-individual variation in UCu in the initial response to treatment.

In a similar study, Pfeiffenberger *et al* reported both 'on treatment' and 'off treatment' UCu in 321 patients treated with penicillamine, trientine or zinc salts between 2003 and 2015.[71] There were no differences in UCu between patients with predominant neurological or hepatic symptoms. Median UCu was lower in this cohort at baseline, 2.11 μmol/24 hours in patients who were subsequently treated with penicillamine. 'Off treatment' UCu was 1.41 μmol/24 hours in this group after 12 months of treatment. The more gradual restoration of copper balance relative to patients in the previous cohort likely reflects the use of lower doses of penicillamine. Long term neurological outcomes were not reported in either study and the consequences of removing copper more slowly or rapidly and the most appropriate target UCu in both the de-coppering or maintenances phases of treatment remain unclear.

Serum non-caeruloplasmin-bound copper

The majority of circulating copper is caeruloplasmin-bound in healthy individuals. Patients with WD are unable to incorporate copper into

caeruloplasmin and so typically have low serum copper (reference range 11-20 μmol/L). The non-caeruloplasmin-bound 'free' copper (NCC) can be calculated (by multiplying the serum caeruloplasmin [g/L] by 49.6 then subtracting this from the serum copper [μmol/L]) and is usually elevated in patients with WD.[59] NCC has been proposed as a diagnostic test but variation in the performance of immunological assays for caeruloplasmin between laboratories makes deciding a universal and reliable cut-off for diagnostic purposes challenging. Negative values, which are physiologically impossible, are not uncommon.[156] NCC does however have a role in monitoring when serial caeruloplasmin and copper measurements can be performed in the same laboratory. Guidelines suggest a target NCC <2.4 μmol/L, again based on expert opinion.[59]

As with UCu, Czlonkowska *et al* did not find an association between NCC and UWDRS-N scores.[38] Walshe examined changes in NCC over the first year of treatment in 80 patients, including 42 with neurological presentations.[70] The mean NCC at baseline was 4.6 μmol/L and decreased to 1.4 μmol/L after 12 months of chelation therapy. Interestingly, there was no association between NCC and UCu in these patients. The author suggests that other factor such as proteinuria, which is common in WD and as an adverse effect of penicillamine,[70] may increase variability in UCu given free copper can loosely bind to albumin, transcuprein and certain amino acids.[148] Pfeiffenberger *et al* also examined NCC in their study on UCu.[71] NCC in the penicillamine-treated group fell from 4.8 μmol/L to 1.6 μmol/L after 12 months. Interval measurements suggested that the majority of this decrease occurred in the first six months of treatment. Again, there was high degree of inter-individual variability in NCC at each time point. This led the authors to conclude that both UCu and NCC are 'less than ideal parameters by which to monitor the benefit of copper-reducing therapy'.

CSF copper

There is a small literature on CSF copper in WD and how it changes with treatment. Weisner *et al* reported the serum, urine and CSF coppers in five patients with neurological presentation starting chelation therapy.[157] CSF copper was increased to up to 3 times the upper limit of normal at presentation

(reference range 0.1-0.6 μmol/L). They fell much more slowly than serum and urine copper and in parallel with clinical improvement. This illustrates the important point that the steady-state concentration of copper is differentially regulated between serum and CSF. Stuerenburg later reported that the average duration of therapy required to normalise CSF copper was 47 months in a series of four patients,[158] much slower than restoration of systemic copper balance. The half-life of CSF copper on treatment was calculated to be 24 months.

In a case report of a patient with a good neurological recovery who stopped taking medication 15 years later, Stuerenburg observed a three-fold increase in CSF copper concentration over two years without any recurrence of neurological signs or deterioration in cognitive tests. More recently, Zhou *et al* reported CSF coppers in 30 patients with WD and 20 age-matched controls. CSF coppers were higher in patients with neurological presentations (2.0 μmol/L) than those with hepatic presentations (1.1 μmol/L) and increased in both of these groups compared to controls (0.3 μmol/L). The relationship between CSF copper and the severity of neurological involvement was not examined.

Although it is not practical to repeat lumbar punctures on a regular basis for monitoring purposes, these reports demonstrate that changes in CSF copper, and therefore presumably brain parenchymal copper, lag behind changes in systemic copper balance. This may help explain the lack of association between measures of systemic copper balance and neurological severity.

Exchangeable copper

A novel method for quantifying copper in serum samples was proposed in 2009.[159] The exchangeable copper (EXC) reflects the labile fraction of copper bound to albumin and other peptides without relying on caeruloplasmin assays. It is measured by adding ethylenediaminetetraacetic acid (EDTA) to a serum sample for one hour prior to performing ultrafiltration to remove peptides and other large biomolecules. The copper concentration in the ultrafiltrate is then measured. Expressed as percentage of total serum copper, relative exchangeable copper (REC) appears to be a useful diagnostic biomarker that

can differentiate patients with WD from controls and heterozygote carriers and patients with hepatic presentations from patients with other liver diseases with high sensitivity and specificity.[160-162] However, methodological concerns regarding the test-retest reliability of the EXC assay have prevented more widespread use of this test.[163] Data on its ability to differentiate Wilson's disease from other neurological and psychiatric disorders is also not yet available.

Poujois *et al* examined the association between EXC and neurological involvement in a cohort of 48 newly-diagnosed patients, 28 of whom were classified as having 'extra-hepatic' disease on the basis of neurological symptoms, abnormal neuroimaging or the presence of KF rings.[73] EXC was higher in the extra-hepatic than hepatic group (2.8 vs 1.3 μmol/L, $P < 0.001$). At a cut-off of 2.08 μmol/L, EXC was able to differentiate patients with hepatic and extra-hepatic involvement with sensitivity 86% and specificity 94%. Unlike NCC and UCu, EXC also correlated with UWDRS-N scores ($r=0.45$, $P=0.02$) and semi-quantitative scores for the severity of KF rings ($r=0.46$, $P=0.02$) and neuroimaging abnormalities ($r=0.38$, $P=0.05$). The authors propose that EXC may have a role as a predictive biomarker, identifying patients who may be at higher risk of neurological deterioration. The role for EXC in monitoring has been studied in *ATP7B* -/- knockout mice, but not humans; EXC did not improve after 11 weeks of treatment.[164]

End-organ biomarkers

While copper indices undoubtedly have a role in monitoring response to treatment in WD, the aforementioned studies illustrate that conventional copper indices, such as NCC and UCu, are poor biomarkers for neurological involvement. From the available data, CSF copper appears to be the most useful copper parameter for monitoring neurological response to treatment but is impractical to measure and, similar to brain parenchymal copper in neuropathological studies,[26,39] is significantly elevated in some patients without neurological involvement.[165] Despite this, the best copper indices can still only inform us about the level of copper the brain is exposed to and not the consequences for brain structure and function. End-organ biomarkers, which directly measure neurological disease activity, are needed.

There has been significant progress in identifying and validating wet (fluid) biomarkers for other neurodegenerative diseases in the last decade. This has primarily been driven by CSF proteomics. CSF communicates freely with brain interstitial fluid where proteins released or secreted from neurons and glia are relatively abundant. These proteins can be detected in CSF samples using standard enzyme-linked immunosorbent assays (ELISA). Disease-specific diagnostic biomarkers in CSF have now entered mainstream clinical practice for some neurodegenerative diseases, for example amyloid beta 1-42 (Aβ42) in Alzheimer's disease.[166]

Non-specific CSF markers of neuronal and glial injury have also been identified: Tau, a microtubule-stabilising protein predominantly expressed in short cortical unmyelinated axons,[167] and neurofilament light (NfL), a cytoskeletal filament primarily expressed in large-calibre myelinated axons,[74] are increased in the CSF of patients with Alzheimer's disease, frontotemporal dementia (FTD), atypical parkinsonian syndromes, prion disease and traumatic brain injury and higher concentrations are associated with more rapid disease progression in various neurodegenerative diseases.[168,169] CSF NfL is also increased in multiple sclerosis (MS), amyotrophic lateral sclerosis (ALS), Parkinson's disease (PD), Huntington's disease (HD) and HIV-associated dementia.[170] Glial fibrillary acidic protein (GFAP), the main cytoskeletal protein in astrocytes, is released into the interstitial fluid on glial activation or injury, and increased in the CSF of patients with traumatic brain injury and MS.[169,171] Ubiquitin C-terminal hydrolase L1 (UCH-L1) is a small cytosolic protein involved in maintaining axonal integrity that is extremely abundant in the brain and makes up an estimated 1-5% of total neuronal protein.[172] It is increased in traumatic brain and spinal cord injuries and ALS.[169,173,174]

The need to perform lumbar puncture limits the use of CSF markers in a clinical setting. However, the glycolymphatic clearance system allows CSF proteins to communicate with blood, albeit in a restricted manner, and some of these biomarkers can be detected in serum or plasma samples at extremely low concentrations. The development of ultrasensitive immunoassays over the last decade using Single Molecule Array (Simoa) technology now makes it possible to measure tau, NfL, GFAP and UCH-L1 in blood samples.

Simoa assays use the same reagents as a standard ELISA however several important steps enhance the sensitivity by more than 100-fold: The primary antibodies are attached to paramagnetic beads and, with highly diluted samples and more than 500,000 beads, the number of beads vastly outnumbers the protein of interest.[74,175] This means that when the concentration of the protein of interest is low each bead will capture a single immunocomplex or none at all. Beads are loaded into 4 x 3 μm wells that can accommodate a single bead before the wells are sealed. By digitally measuring the fluorescent signals in each well, the proportion of 'positive' wells is used to determine the concentration of the protein of interest.

Of CSF proteins that can be detected in blood samples, NfL has garnered the most interest. Serum or plasma NfL has been shown to correlate with clinical, biochemical and/or imaging markers of several neurodegenerative disease. In Alzheimer's disease, plasma NfL correlates with MMSE scores, CSF Aβ42, hippocampal volume and cortical thickness.[176] It has also been shown to have prognostic value in several diseases, including pre-symptomatic carriers of genetic neurodegenerative diseases. In a longitudinal study of genetic FTDs, serum NfL was higher in pre-symptomatic carriers that became symptomatic during the follow up period. The rate of change in serum NfL was also associated with decline in MMSE scores and atrophy in symptomatic patients.[177] In HD, plasma NfL is increased in gene carriers approximately 24 years from predicted clinical onset,[178] and correlates with subsequent decline in cognition, functional capacity and brain volume in symptomatic patients.[179]

There is particular interest in using NfL as a biomarker of disease activity in MS given disease-modifying treatment are standard of care and NfL may have a role as both a prognostic and monitoring biomarker. Serum NfL correlates with Expanded Disability Status Scale (EDSS) scores and the presence of contrast-enhancing lesions on MRI.[180] It predicts progression of disability and brain volume loss over 10 years.[181,182] Reductions in NfL have been reported with various disease-modifying treatments and with escalation from first- to second-line DMTs.[183] Reyes *et al* recently examined the use of CSF NfL in clinical practice as well as its effect on treatment strategies and outcomes in patients with MS.[184] They found that CSF NfL was the only

indicator of ongoing disease activity in a subset of patients and elevated levels were associated with more treatment or escalation of treatment, which had an impact on EDSS at one year. The concept of 'no evidence of disease activity', or NEDA, is widely used in MS and some authors have proposed that NfL measurements could be used to determine NEDA in MS.[185]

Wet (fluid) biomarkers for neurological involvement in WD had only been investigated in a single study prior to our work: Lekomtseva *et al* measured serum tau in 47 patients with WD and 30 healthy controls using an ELISA. Levels were higher in chronically-treated patients with neurological presentations than controls (221.7 vs 71 ng/L, P<0.001) but the relationship with neurological involvement was untested.[186] There are also methodological concerns about these results given the prevailing view is that tau cannot be reliable measured in serum samples using an ELISA.

Label-free quantitative proteomics

An alternative approach to biomarker discovery is to use unbiased proteomic profiling in biofluid samples to identify novel candidate biomarkers. Recent technological advances have made it possible to perform 'label-free' quantitative proteomics using mass spectrometry in relatively complex biosamples, such as plasma or serum, with high throughput.[187] Here, I introduce the principles behind a data independent approach, known as MS^E, using high-performance liquid chromatography mass spectrometry (HPLC-MS^E) and its application to biomarkers discovery.

Mass spectrometry is a powerful technique for detecting compounds ranging from small molecules, including metals, to large complex biomolecules, such as proteins. There a various types of mass spectrometry but it invariably consists of at least three key steps[188]:

> *Ionisation* – This involves converting compounds into a positively-charged or ionised state. The positively charged compounds and fragments of these compounds are referred to as ions and usually have a charge of +1 or +2. This process can be performed using several methods such as electrospray ionisation (ESI) or matrix-assisted laser

desorption/ionisation (MALDI). Cations, including copper, can be ionised using inductively-coupled plasma mass spectrometry (ICP-MS).

Mass selection – This involves separating ions according to their mass-to-charge ratio (m/z). This can be achieved using time-of-flight, quadrupole mass filters, ion trap or orbitrap analysers. Time-of-flight (TOF) analysers operate by accelerating ions through a high voltage. Ions with a smaller m/z ratio travel at higher velocity and reach the detector quickly whereas ions with a larger m/z ratio travel at lower velocity and reach the detector slowly.

Detection – This involves recording the charge as ions pass a detector over time so that the relative abundance of each ion at a given m/z ratio can be determined. The resolution of a mass spectrometer refers to its ability to differentiate two peaks in relative abundance from ions with a similar m/z ratio.

Liquid chromatography (LC) is a separation technique that can be used to enhance the ability of mass spectrometry to identify and quantify complex molecules.[189] LC-MS involves passing a sample through a column containing adsorbent material prior to performing mass spectrometry. The sample is separated according to the hydrophobic/hydrophilic interaction of its constituents with the adsorbent material. The amount of time it takes for a given compound to pass through the column is known as the retention time. LC-MS therefore enables separation of a sample by retention time prior to separation by m/z ratio. High performance liquid chromatography (HPLC) involves injecting the sample into the column under high pressure and enables better separation of constituents. Nano-LC refers to performing HPLC on columns with much smaller diameters, on a capillary scale, in order to improved sensitivity and efficiency of ionisation in the mass spectrometer.[190]

The application of mass spectrometry for quantitative proteomics can be further enhanced in several ways.[188] Ion mobility separation (IMS), where ions are separated based on their interactions with a collision gas, can be performed between ionisation and m/z separation. This increases the peak

capacity allowing a wider range of compounds to be identified in a single run when combined with mass spectrometry in IMS-MS.[75] Tandem mass spectrometry (MS/MS) involves ionisation and separation of *precursor* ions by m/z in *MS1* prior to further fragmentation and separation of selected *product* ions by m/z in *MS2*. This is, in essence, combining two mass spectrometers. Similar gains can be made by using MS^E, where data is acquired in alternating low and high collision energy modes, with a single time-of-flight or quadrupole analyser.[187] Precursor ions are detected in low collision energy acquisition and the more fragmented product ions are detected in the high collision acquisitions.

MS^E, and other label-free methods, have been used for biomarkers discovery in biofluid samples in various disorders, including a few neurological and psychiatric diseases.[75] In ALS, MS^E was used with CSF samples to identify four classifiers proteins (WD repeat-containing protein 63, amyloid-like protein 1, SPARC-like protein 1, and cell adhesion molecule 3) that could differentiate ALS from non-ALS samples with 83% sensitivity and 100% specificity.[191] Heywood *et al* used LC-MS^E to identify 38 proteins that were increased in the CSF of patients with Lewy body disease (LBD) compared to controls and were used to create a targeted MS/MS assay. They applied this assay to the CSF of 17 patients with LBD, 7 patients with PD, 16 patients with Alzheimer's disease and 15 healthy controls in a validation study and identified four novel proteins that were able to differentiate Alzheimer's disease and LBD and three further proteins that correlated with CSF Aβ42.[192] More recently, Figura *et al* used label-free LC-MS/MS to identify three proteins that are reduced in the saliva of patients with PD.[193]

Label-free quantitative proteomics has not been applied to biosamples from patients with WD. Hydrophilic interaction LC-MS has been used to examine the untargeted serum metabolome of patients with WD. Sarode *et al* identified 99 metabolites related to amino acid metabolism, the tricarboxylic acid cycle, choline metabolism and oxidative stress that differed between patients and healthy controls. There were 57 metabolites that differed between patients with neurological presentations and controls and two metabolites, urate and pyrophosphate were not altered in hepatic presentations. No

significant differences were identified between patients with neurological and hepatic presentations.

Wet (fluid) biomarker objectives

Overall, a number of studies have examined associations between copper indices and neurological involvement, although the relationship between EXC and neurological involvement in chronically-treated patients is untested. Simoa assays had not been used to measure serum neuronal or glial-derived proteins in patients with WD prior to this study and label-free MS^E has not been attempted.

We therefore aimed to characterise the biochemical profile of our participants using NCC, UCu and EXC and then perform a Simoa assay to measure tau, NfL, GFAP and UCH-L1 in plasma samples before using LC-MS^E to identify novel serum biomarkers for neurological involvement in WD. I hypothesised that copper indices would be higher in patients with active disease but not associated with neurological severity and that plasma NfL would be higher in patients with neurological presentation and active disease.

4.2 Methods

Copper indices were measured and the Neurology 4-Plex A (N4PA) Simoa assay was performed using the methods described below. These are also outlined in a previous publication.[194] Methods for LC-MS^E are discussed prior to explaining the statistical analyses for both of these methods. Other research associates or technicians performed these various methods, as outlined in the statement of contributions, and I performed the statistical analyses.

Copper indices

Serum ceruloplasmin and copper were measured using the immuno-turbidimetric test (Beckman Coulter) and ICP-MS (NexION 300, PerkinElmer), respectively, within the Supra-Regional Assay Service at University Hospital Southampton NHS Foundation Trust. The NCC was then calculated by subtracting the serum caeruloplasmin (g/L) multiplied by 49.6 from the total serum copper (µmol/L).[59] EXC was measured in the same laboratory as previously described:[73] Equal volumes of EDTA 3g/L were added to plasma for one hour prior to transfer to the Centrifree Ultrafiltration device (Millipore). Samples were centrifuged at 2000g for one hour and copper was then measured in filtrates as above. UCu was measured using ICP-MS (7700x system, Agilent) within the Departmental of Clinical Biochemistry at Northern General Hospital, Sheffield Teaching Hospitals NHS Foundation Trust. Results from participants on zinc therapy or with liver transplants, in whom UCu is inherently lower, were excluded.

Targeted proteomics – N4PA assay

Plasma samples from 38 age-matched healthy controls were provided from the Genetic Frontotemporal Dementia Initiative (GENFI) cohort for this part of the study.[13] These healthy controls were family members of patients with *MAPT*, *GRN* and *C9orf72* mutations who had all undergone genetic testing and were non-carriers. Samples had been collected, processed and stored using the same experimental procedures as in the CROWD study.

NfL, tau, UCH-L1 and GFAP concentrations were measured in plasma samples from the CROWD and GENFI cohorts. N4PA kits were used according to manufacturer instructions on a Simoa HD-X analyser (Quanterix)

at the Dementia Research Institute, UCL Queen Square Institute of Neurology. Samples were thawed at room temperature for 2 hours, centrifuged at 10,000 g for 5 minutes and then transferred to 96-well plates, which were then sealed using an X-Pierce XP-100 (Excel Scientific). Duplicate samples were analysed, as recommended, and samples from the GENFI and CROWD cohorts were inter-mixed. For each experimental run, the 96-well plate, bead reagent (beads coated with primary antibody in buffer), sample diluent (buffer with protein stabiliser acting as a heterophilic blocker), detector reagent (biotinylated secondary antibody in buffer), streptavidin-ß-galactosidase (SBG) reagent, resourfin ß-D-galactopyranoside (RGP) reagent and calibrators were loaded into the HD-X analyser. Automated measurements with the standard 4x dilution protocol were carried out on consecutive days and the operator was blinded to clinical information.

Untargeted proteomics - LC-MS^E

A total of 20 samples were analysed using nano-LC IMS-MS^E within the Biological Mass Spectrometry Centre, UCL Great Ormond Street Institute of Child Health. This included five patients with neurological presentations, all classified as having active disease, and five patients with hepatic presentation who were matched for age and sex. Serum samples from five healthy controls and five patients with SCA3 mutations were provided by the Ataxia Centre, UCL Queen Square Institute of Neurology. The samples were processed and stored under the same conditions and were also matched for age and sex with the patients with neurological presentations. Data relating to the five patients with SCA3 mutations are not reported in this thesis.

Samples were thawed and the 12 most abundant serum proteins (α1-acid glycoprotein, α1-antitrypsin, α2-macroglobulin, albumin, apolipoprotein A-I, apolipoprotein A-II, fibrinogen, haptoglobin, IgA, IgG, IgM and transferrin) were removed from each sample using Pierce Top12 Abundant Protein Depletion Spin Columns (Thermo Scientific). A volume of 10 µL of each sample was added to the resin slurry in each column. Columns were then incubated with end-over-end mixing for 60 minutes at room temperature prior to centrifugation at 1000 g for 2 minutes. The ultrafiltrates were freeze dried at -40 °C.

The dried samples were then digested. Samples were dissolved in 20 µL digest buffer (100 mM Tris, pH 7.8, 6M urea, 2 M thiourea, 2% ASB14) for 60 minutes and incubated with 1.5 µL of dithioerythritol (DTE) solution (30 mg DTE/1000 µL, 100 mM Tris, pH 7.8), to breakdown disulphide bonds, for 60 minutes. A volume of 3 µL iodoacetic acid (IAA) solution (36 mg IAA/1000 µL, 100mM Tris, pH 7.8) was added for 45 minutes in darkness to prevent re-formation of the disulphide bonds. A volume of 165.5 µL of MilliQ water was added prior to 10 µL of 0.1 g/L Trypsin Gold solution (Promega). Samples were then incubated in a 37 °C water bath for 16 hours.

Samples were simultaneous cleaned (to remove salts and other impurities) and fractionated using solid phase extraction prior to LC-MS analysis. Acetonitrile (ACN) solutions of 10 concentrations (7.4%, 10.8%, 12.6%, 14.0%, 15.3%, 16.7%, 18.3%, 20.4%, 23.5%, 60.0%) were mixed. ISOLUTE C18 cartridges (Biotage) were used to perform solid phase extraction. Samples were diluted with 200 µL NH_4OH (20 mM). Cartridges were washed with 2 x 1 mL 100% ACN and primed with 2 x 1 mL NH_4OH. Samples were loaded into cartridges and washed with 2 x 1mL NH_4OH. The first fraction was eluted into a 96-well plate using 2 x 250 µL of 7.4% ACN solution. The second fraction was eluted using the 10.8% ACN solution, and so on, until the 10 fractions were separated. Solvents were then evaporated in a SpeedVac vacuum concentrator. Quality control samples were created by pooling equal volumes from each fraction of the samples.

Peptides were separated using a NanoAcquity high-performance liquid chromatography system (Waters corporation, Manchester, UK). The chromatographic system's mobile phase was A: 0.1% formic acid, B: acetonitrile 0.1% formic acid. The peptides were loaded onto a 180 µm x 20 mm, 5 µm Symmetry C18 trap column (Waters) before entering the analytical column, a 75 µm x 150 mm, 1.7 µm Peptide BEH C18 (Waters). The column temperature was set to 45 °C. The gradient elution started at 3% B and was linearly increased to 60% B over 40 minutes after which it was increased to 85% B over 2 minutes and washed for 2 minutes before returning to the initial conditions over 2 minutes followed by 15 minutes of equilibration before the subsequent injection. The eluted peptides were detected using a Synapt-G2-Si system (Waters) equipped with a nano-electrospray ion source.

Data were acquired in positive MS^E mode from 0 to 60 minutes within the m/z range 50-2000. Capillary voltage was set to 3 kV and source temperature to 100 °C. The desolvation gas consisted of nitrogen with a flow of 50 L/h and desolvation temperature was set to 200 °C. The purge and desolvation gas consisted of nitrogen with flow rates of 600 mL/h and 600 L/h, respectively. The gas in the IMS cell was helium with a flow rate of 90 mL/h. The low energy acquisition was performed applying a constant collision energy of 4 V with a 1 s scan time. High energy acquisition was performed by applying a collision energy ramp from 15 to 40 V with a 1 s scan time. A lock spray of [glu^1]-fibrinopeptide B was delivered through the auxiliary pump at concentration 500 fmol/L with flow rate of 300 nL/min and acquired twice per minute. The doubly charged precursor ion, m/z 785.6426, was used for mass correction.

Data analysis was performed in Progenesis QI. Lock mass calibration was performed and peak threshold intensities were manually set at 200 counts for low energy and 20 counts for high energy to optimise peptide identification. Elution limits were set at 10 and 50 minutes. Automatic alignment was performed to compensate for drifts in retention time between runs. Alignment was manually reviewed and poorly aligned samples were re-aligned. Protein identifications were obtained for the ProteinLynx Global Server (PLGS) using Ion Accounting. Trypsin was set at the protease and 3 missed cleaves were allowed. Protein identification in each sample was processed using a hierarchical approach where more than two fragments ions per peptide, three fragment ions per protein and one peptide per protein had to be matched. Search tolerance parameters were set with a 4% false discovery rate. Protein identification parameters used in the database search included fixed modification with carboamidomethylation of cysteines and dynamic modifications with deamidation of asparagine and glutamine, oxidation of methionine and pyrrolidone carboxylic acid at N-termini. Relative quantitation of protein was performed using Hi-N. Proteins with >95% confidence identification were exported for further analysis and ranked (A, B or C) according the confidence of their identification.

Statistical analyses

I compared copper indices and N4PA results between the neurological, hepatic and control groups using linear regression with age as a covariate, given NfL and other neuronal and glial markers increase with age.[9] I compared results between active and stable patients using linear regression with age and UWDRS-N as covariates. I used receiver operating characteristic (ROC) curves to assess diagnostic performance in differentiating groups and calculated sensitivities and cut-off values at 70%, 80% and 90% specificities where the area-under-curve (AUC) was increased ($p<0.05$). I tested associations between biomarkers, copper indices and UWDRS subscores in stable patients using linear regression with age as a covariate and used FDR-correction for multiple testing.

For LC-MS^E data, I compared the relative abundance of each protein exported for analysis between patients with neurological presentations and healthy controls, between patients with neurological and hepatic presentations and between patients with hepatic presentations and healthy controls using t-tests in R. P values were corrected for multiple comparisons across all proteins using the FDR. Accession numbers for proteins with uncorrected $P<0.01$ were cross-referenced with the UniProtKB 2021_03 database (https://www.uniprot.org/) to ascertain their function and expression. Proteins with enhanced tissue expression in the brain were reviewed. The STRING database version 11.0 (https://string-db.org/) was used to identify predicted functional associations with other proteins.

4.3 Results

Copper indices and N4PA assay

Copper indices and N4PA results are summarised in **Table 4-1** and individual results are shown in **Figure 4-1** and **Figure 4-2**. There were no differences in NCC, EXC or UCu between patients with neurological and hepatic presentations. NfL concentrations were higher in patients with neurological presentations than those with hepatic presentations and controls. These differences persisted when patients with active disease were excluded (8.7 vs 7.0 ng/L, P =0.04; 8.7 vs 7.5 ng/L, P=0.01). UCH-L1 concentrations were higher in patients with neurological presentations compared to controls but not compared to patients with hepatic presentations. There were no differences in GFAP or tau concentrations between patients with neurological and hepatic presentations.

UCu was higher in patients with active disease than those with stable disease, as was NfL without correction for multiple testing. The highest NfL concentration, 38.5 ng/L, and UWDRS-N score, 106, was in the newly-diagnosed patient with ongoing paradoxical worsening.

The ROC curve for NfL is depicted in **Figure 4-3**. The AUC for NfL in differentiating patients with neurological and hepatic presentations was 0.71 (p=0.03). For 70%, 80% and 90% specificities, the sensitivities were 61% (41-78), 48% (29-67) and 39% (22-59) with cut-off values 8.4, 9.7 and 12.9 ng/L, respectively.

NCC was strongly correlated with EXC (r= 0.44, P = 0.005) and UCu (r= 0.54, P = 0.001). There was no association between EXC and UCu (r= 0.01, P = 0.98). There were no associations between copper indices and N4PA results after correction for multiple testing as shown in **Table 4-2**.

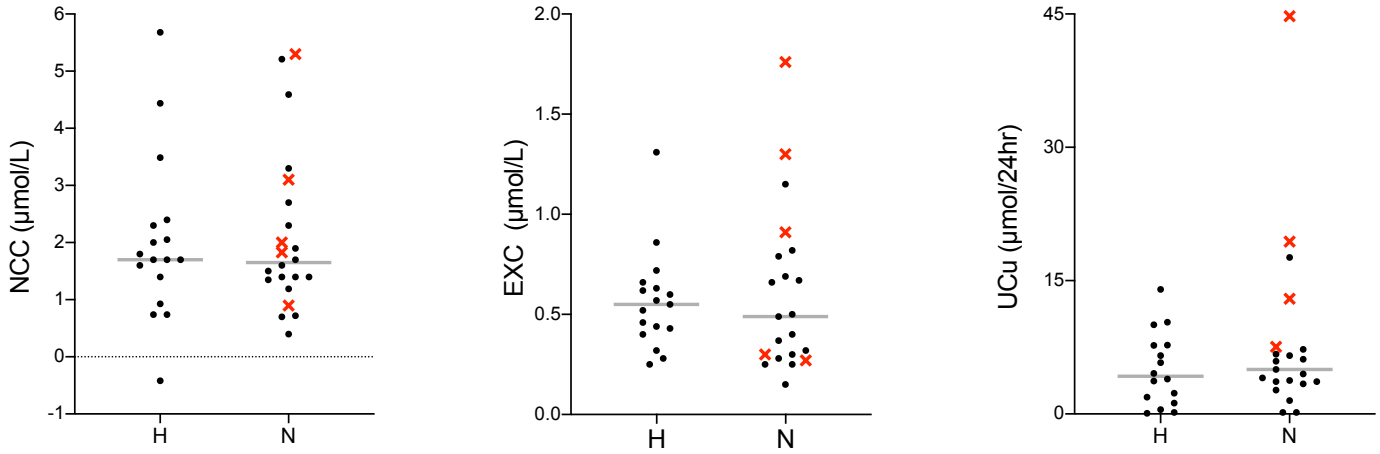
Associations between biomarker results and UWDRS-N, cognitive test and psychiatric scores in stable patients are shown in **Table 4-3**, **Table 4-4** and **Table 4-5**, respectively. Associations between N4PA results and UWDRS-N scores in stable patients without correcting for age are depicted in **Figure 4-4**. NfL, Tau and UCH-L1 were strongly associated with UWDRS-N scores when controlling for age. There was a weak association between NfL and UWDRS-N scores without controlling for age. Associations with

Table 4-1 Group differences in copper indices and N4PA results

	Controls (n=38) median [IQR]	**Hepatic (n=17)** median [IQR]	**Neurological (n=23)** median [IQR]	**P value** Neuro vs controls	**P value** Neuro vs hepatic	**Stable (n=35)** median [IQR]	**Active (n=5)** median [IQR]	**P value**
NCC (µmol/L)	-	1.7 [1.4-2.3]	1.7 [1.4-2.6]	-	0.86	1.7 [1.4-2.3]	2.0 [1.8-3.1]	0.54
EXC (µmol/L)	-	0.6 [0.4-0.6]	0.5 [0.3-0.8]	-	0.71	0.5 [0.3-0.7]	0.9 [0.3-1.3]	0.13
UCu (µmol)	-	4.2 [1.7-7.7]	5.0 [3.6-7.3]	-	0.28	4.0 [2.3-6.6]	16.2 [11.6-25.7]	**<0.001*** **
NfL (ng/L)	7.6 [5.4-9.9]	7.0 [4.9-8.8]	8.7 [6.6-16.0]	**0.005** **	**0.005** **	7.7 [5.8-11.8]	22.2 [8.6-23.2]	0.05*
Tau (ng/L)	1.4 [1.1-2.3]	1.4 [1.3-1.7]	1.8 [1.2-2.1]	0.57	0.13	1.6 [1.3-1.9]	1.2 [1.1-3.1]	0.57
GFAP (ng/L)	84 [65-136]	80 [67-90]	84 [65-136]	0.61	0.50	79 [66-106]	135 [84-155]	0.42
UCH-L1 (ng/L)	23 [14-41]	23 [17-31]	23 [14-41]	**0.01***	0.06	22 [14-32]	41 [35-45]	0.87

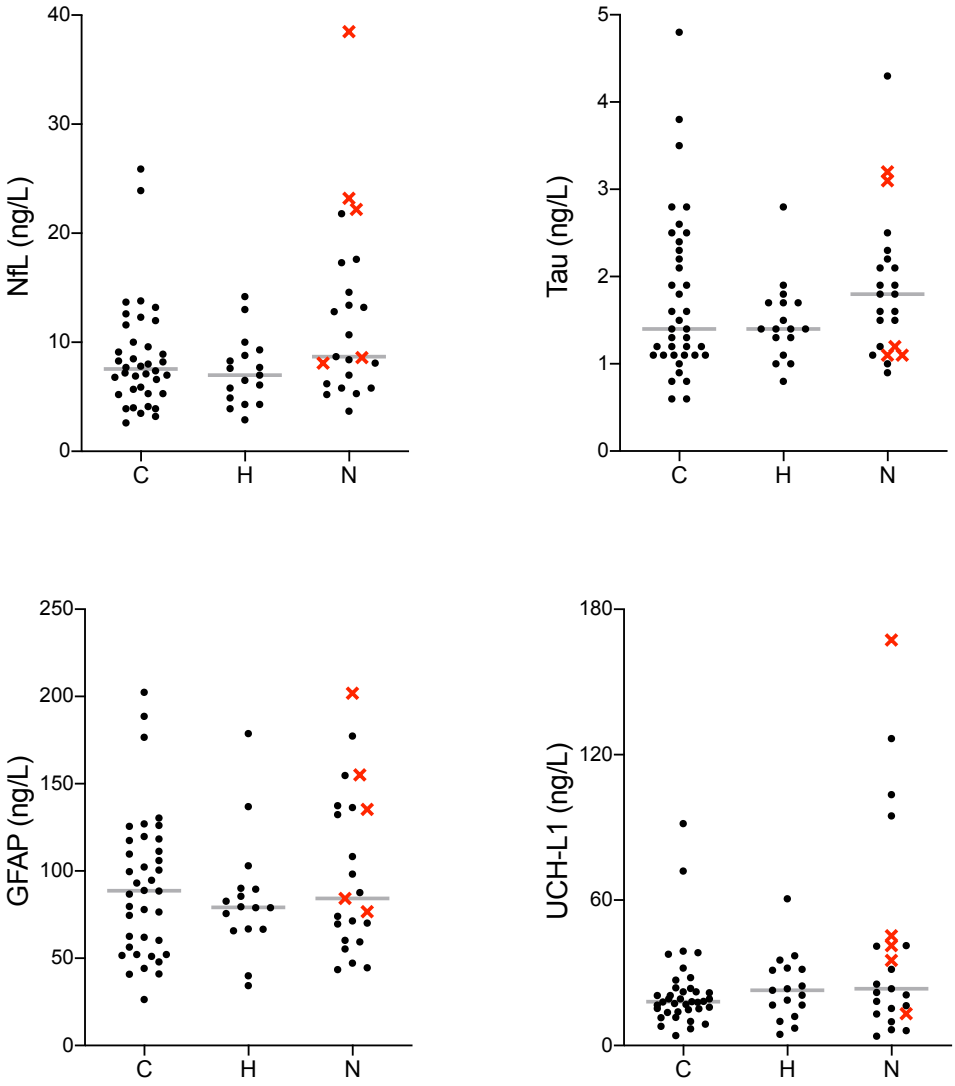
** = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. EXC, exchangeable copper; GFAP, glial fibrillary acidic protein; IQR, interquartile range; N4PA, neurology 4-plex A assay; NCC, non-caeruloplasmin-bound copper; NfL, neurofilament light; UCH-L1, ubiquitin carboxy-terminal hydrolase L1; UCu, 24-hour urinary copper output.*

Figure 4-1 Group differences in copper indices



Individual NCC, EXC and UCu results for patients with hepatic presentations (H) and neurological presentations (N) are shown. Red crosses indicate those with active disease. EXC, exchangeable copper; NCC, non-caeruloplasmin-bound copper; UCu, 24-hour urinary copper output.

Figure 4-2 Group differences in N4PA results



Individual NfL, tau, GFAP and UCH-L1 results for healthy controls (C), patients with hepatic presentations (H) and patients with neurological presentations (N) are shown. Red crosses indicate those with active disease. GFAP, glial fibrillary acidic protein; N4PA, neurology 4-plex A assay; NfL, neurofilament light; UCH-L1, ubiquitin carboxy-terminal hydrolase L1.

Figure 4-3 ROC curve for NfL in differentiating patients with hepatic and neurological presentations

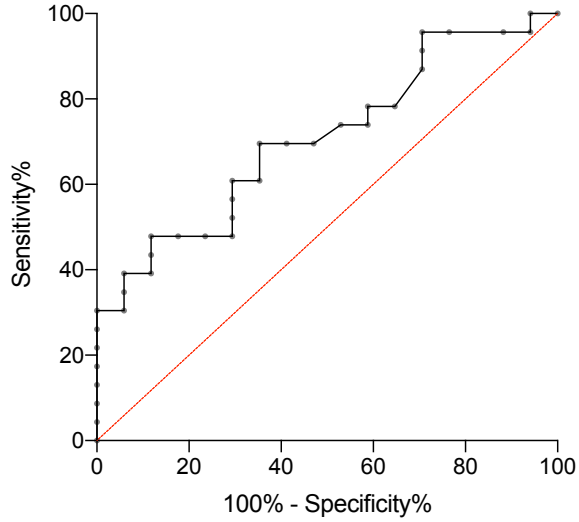


Table 4-2 Associations between N4PA results and copper indices

	NCC	EXC	UCu
NfL	0.32	0.08	0.14
Tau	0.34	0.01*	0.67
GFAP	0.65	0.46	0.81
UCH-L1	0.21	0.73	0.32

*P values for coefficients when testing associations between N4PA results and copper indices using linear regression are shown. Corresponding coefficients where P < 0.05 were negative. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. EXC, exchangeable copper; GFAP, glial fibrillary acidic protein; NCC, non-caeruloplasmin-bound copper; NfL, neurofilament light; UCH-L1, ubiquitin carboxy-terminal hydrolase L1; UCu, 24-hour urinary copper output.*

Table 4-3 Associations between wet biomarkers and UWDRS-N scores

	NCC	EXC	UCu	NfL	Tau	GFAP	UCH-L1
Speech	0.86	0.21	0.33	0.16	**0.02***	0.80	**<0.001*****
Dystonia	0.97	0.42	0.92	**0.01***	**<0.001*****	0.89	**<0.001*****
Parkinsonism	0.70	0.68	0.28	**0.009****	**0.003****	0.74	**<0.001*****
Tremor	0.71	0.21	0.73	0.20	0.52	0.17	0.80
Ataxia	0.10	0.87	0.85	**0.003****	0.07	0.71	**<0.001*****
Chorea	0.25	0.87	0.27	0.29	0.39	0.54	0.30
Rigidity	0.69	0.48	0.34	**0.02***	**<0.001*****	0.81	**<0.001*****
UWDRS-N	0.55	0.69	0.55	**0.003****	**<0.001*****	0.74	**<0.001*****

*P values for coefficients when testing associations between wet biomarkers, including N4PA results and copper indices, and UWDRS-N scores using linear regression are shown. Corresponding coefficients where P < 0.05 were positive. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. EXC, exchangeable copper; GFAP, glial fibrillary acidic protein; NCC, non-caeruloplasmin-bound copper; NfL, neurofilament light; UCH-L1, ubiquitin carboxy-terminal hydrolase L1; UCu, 24-hour urinary copper output; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.*

Table 4-4 Associations between wet biomarkers and cognitive scores

Domain	Test	NCC	EXC	UCu	NfL	Tau	GFAP	UCH-L1
Abstract reasoning	MRT	0.92	0.54	0.45	0.35	0.08	0.83	0.37
Language	NART	0.36	0.87	0.93	0.90	0.80	0.32	0.27
	GNT	0.14	0.46	0.24	0.52	0.98	0.27	0.95
Memory	RMTF	0.13	1.00	0.54	0.55	0.23	0.35	0.07
	RMTW	0.65	0.95	0.26	0.31	0.004** (-)	0.86	0.02*
	CPAL	0.46	0.95	0.75	0.90	0.85	0.87	0.36
Processing speed	TMTA	0.33	0.12	0.39	0.14	0.81	0.43	0.10
Executive function	DSB	0.89	0.11	0.83	0.34	0.77	0.77	0.75
	FAS	0.86	0.86	0.03* (+)	0.75	0.87	0.75	0.49
	Animals	0.81	0.61	0.51	0.83	0.31	**<0.001*** (+)	0.80
	DKEFSI	0.47	0.55	0.79	0.44	0.45	0.22	0.16
	TMTB	0.08	0.09	0.89	0.01* (-)	0.12	0.81	0.05*
	DSym	0.24	0.05	0.37	0.14	0.97	0.57	0.18
Calculation	GDA	0.99	0.98	0.94	0.46	0.40	0.28	0.64
Visuoperceptual	VOSPFL	0.45	0.14	0.80	0.85	0.73	0.38	0.73
Visuospatial	VOSPNL	0.33	0.35	0.11	0.60	0.83	0.68	0.46
Social	Ekman	0.02	0.84	0.53	0.21	0.33	0.81	0.08

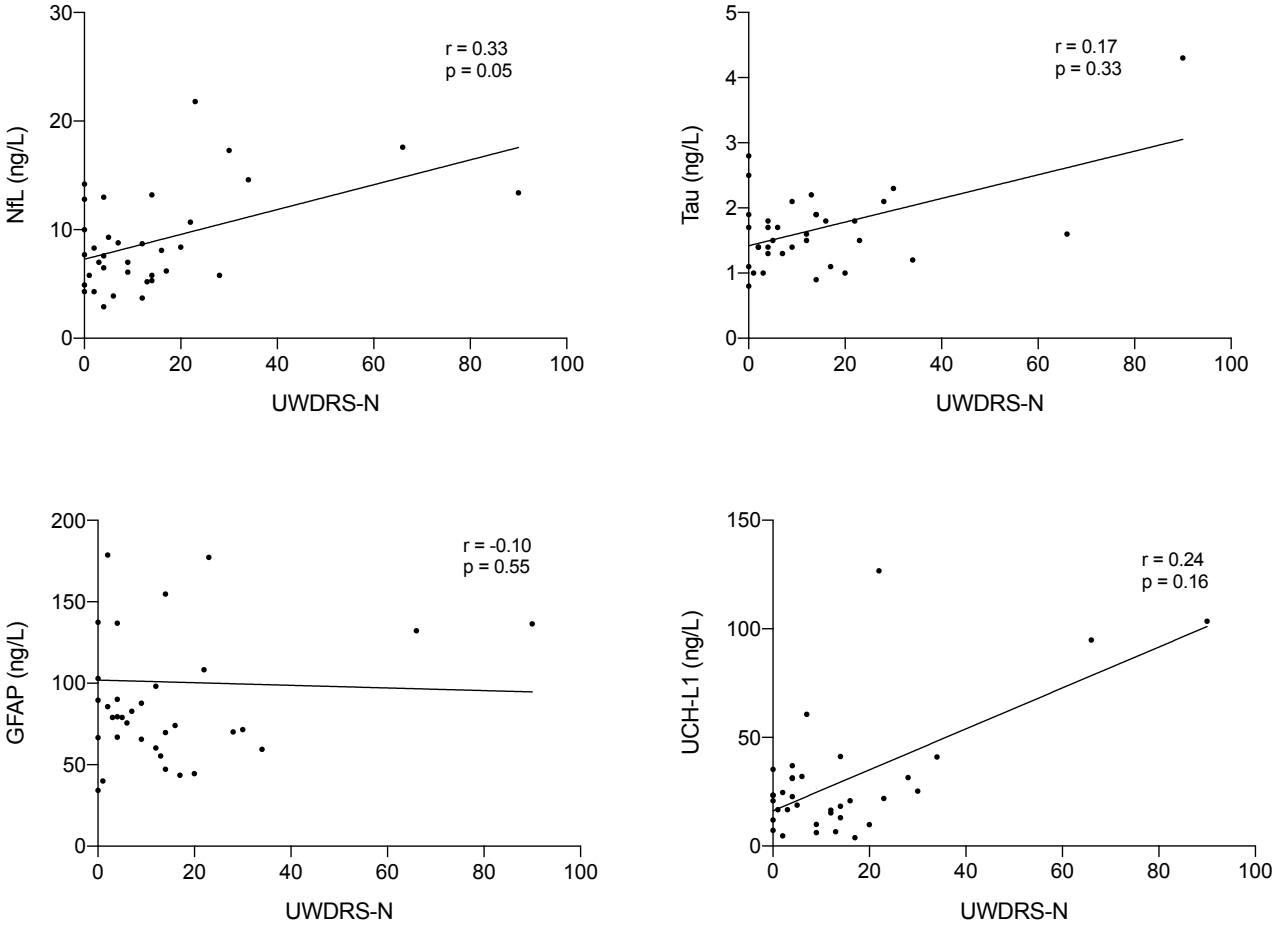
*P values for coefficients when testing associations between wet biomarkers, including N4PA results and copper indices, and cognitive test scores using linear regression are shown. Corresponding coefficients where P < 0.05 were positive (+) or negative (-). * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; EXC, exchangeable copper; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GFAP, glial fibrillary acidic protein; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; NCC, non-caeruloplasmin-bound copper; NfL, neurofilament light; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; TMTA, trail making test part A; TMTB, trail making test part B; UCH-L1, ubiquitin carboxy-terminal hydrolase L1; UCu, 24-hour urinary copper output; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

Table 4-5 Associations between wet biomarkers and psychiatric scores

	NCC	EXC	UCu	NfL	Tau	GFAP	UCH-L1
UWDRS-P	0.97	0.63	0.93	0.79	0.21	0.22	0.64
ModNPI	0.43	0.25	0.24	0.52	0.17	0.27	0.39
PHQ9	0.58	0.36	0.67	0.98	0.04*	0.32	0.46
GAD7	0.79	0.10	0.90	0.57	0.04*	0.20	0.37

*P values for coefficients when testing associations between wet biomarkers, including N4PA results and copper indices, and psychiatric rating scales using linear regression are shown. Corresponding coefficients where P < 0.05 were positive. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. EXC, exchangeable copper; GAD7, generalised anxiety disorder assessment 7; GFAP, glial fibrillary acidic protein; ModNPI, modified neuropsychiatric inventory; NCC, non-caeruloplasmin-bound copper; NfL, neurofilament light; PHQ9, patient health questionnaire 9; UCH-L1, ubiquitin carboxy-terminal hydrolase L1; UCu, 24-hour urinary copper output; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; UWDRS-P, unified Wilson's disease rating scale psychiatric subscore.*

Figure 4-4 Associations between N4PA results and UWDRS-N scores without correcting for age



Scatter plots comparing N4PA results and UWDRS-N scores are shown. Spearman correlation coefficients (r) and corresponding P values are included. GFAP, glial fibrillary acidic protein; N4PA, neurology 4-plex A assay; NfL, neurofilament light; UCH-L1, ubiquitin carboxy-terminal hydrolase L1; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.

UWDRS-N scores were driven by subscores for dystonia, parkinsonism, ataxia and rigidity and less so speech or tremor. The only association between wet biomarkers and cognitive test scores was a positive correlation between Animals scores and GFAP. There were no associations with scores for psychiatric rating scales.

Increasing age was associated with increasing NfL (r=0.74, P<0.001) and GFAP (r=0.41, P=0.009) but not Tau (r=-0.1, P=0.59) or UCH-L1 (r=-0.12, P=0.48) in controls. There was no difference in NfL between stable patients with and without cirrhosis (7.7 vs 8.3 ng/L, P = 0.19). There were no associations between disease duration and NfL (r=0.28, P=0.10), tau (r=0.1, P=0.75), GFAP (r=0.13, P=0.45) or UCH-L1 (r=0.02, P = 0.91). The mean coefficient of variation, calculated from duplicate samples, was 4.9% for NfL, 2.4% for GFAP, 6.2% for tau and 15.3% for UCH-L1.

Label-free quantitative proteomics

Due to technical issues with the Synapt-G2-Si system (Waters) that are ongoing, only the first four fractions from each sample have been analysed. In addition, the first fraction of the sample from a patient with a hepatic presentation was not inadequately processed and needs to be re-analysed. The results presented here may change when the remaining six fractions are processed.

A total of 586 proteins were identified in the serum samples. When comparing patients with neurological presentations with healthy controls, 77 proteins differed between groups with P values <0.05 of which nine had FDR-corrected P values <0.05. When comparing patients with hepatic presentations with healthy controls, 109 proteins differed between groups with P values <0.05 of which 25 had FDR-corrected P values <0.05. When comparing patients with neurological and hepatic presentations, 30 proteins differed between groups with P values <0.05, none of which had FDR-corrected P values <0.05. The proteins identified with each of these comparisons for P values <0.01 are shown in **Table 4-6**, **Table 4-7** and **Table 4-8**, respectively.

Caeruloplasmin was lower in patients with neurological presentations than controls (89,723 vs 535,108; P < 0.001 and $_{FDR}$P = 0.01) and patients

Table 4-6 Protein identifications when comparing patients with neurological and hepatic presentations

Accession	Protein	Rank	hWD	nWD	P value
Q99590	Protein SCAF11	A	2155	1528	0.003
Q14C86	GTPase-activating protein and VPS9 domain-containing protein 1	A	6498	7378	0.004
Q53F39	Metallophosphoesterase 1	B	198	687	0.005
O75771	DNA repair protein RAD51 homolog 4	C	2854	9440	0.009
Q6PKG0	La-related protein 1	A	9214	12835	0.009

Proteins where relative abundance differed between patients with neurological and hepatic presentations with uncorrected P > 0.01 are shown. Proteins where relative abundance was higher in patients with neurological presentations are shown in green and lower in patients with neurological presentations are shown in red. P values less than 0.05 after FDR correction are highlighted in bold. Rank refers to the confidence in peptide identification (A-C).

Table 4-7 Protein identifications when comparing patients with neurological presentations and healthy controls

Accession	Protein	Rank	HC	nWD	P value
P00450	Caeruloplasmin	A	535109	89723	**<0.001**
P13498	Cytochrome b-245 light chain	C	132808	37916	**<0.001**
P02788	Lactotransferrin	A	30020	12674	**<0.001**
Q68BL8	Olfactomedin-like protein 2B	B	485	217	**<0.001**
Q02224	Centromere-associated protein E	A	3968	6787	**<0.001**
O94823	Probable phospholipid-transporting ATPase VB	B	4620	2917	**<0.001**
Q6PF18	MORN repeat-containing protein 3	C	1432	480	**<0.001**
A2RRP1	Neuroblastoma-amplified sequence	A	4155	9990	**<0.001**
Q8IZT8	Heparan sulfate glucosamine 3-O-sulfotransferase 5	C	781	206	**<0.001**
Q8IWJ2	GRIP and coiled-coil domain-containing protein 2	A	2473	3775	0.001
Q70CQ1	Ubiquitin carboxyl-terminal hydrolase 49	A	3326	1013	0.001
P01023	Alpha-2-macroglobulin	A	270872	627669	0.002
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	A	10066	1494	0.002

Q17R98	Zinc finger protein 827	A	20427	14810	0.002
Q7Z3Z0	Keratin_ type I cytoskeletal 25	A	2217	5592	0.002
Q8TF21	Ankyrin repeat domain-containing protein 24	A	1086	480	0.002
Q63HN8	E3 ubiquitin-protein ligase RNF213	A	19582	35634	0.003
O14733	Dual specificity mitogen-activated protein kinase 7	A	18601	34778	0.003
P61626	Lysozyme C	A	27918	37531	0.003
Q9Y2K3	Myosin-15	A	1738	2415	0.004
P35579	Myosin-9	A	744	954	0.005
P02751	Fibronectin	A	88538	64004	0.005
P20848	Putative alpha-1-antitrypsin-related protein	B	2480	6518	0.005
P19320	Vascular cell adhesion protein 1	A	12639	23036	0.005
Q9NQZ6	Zinc finger C4H2 domain-containing protein	A	8355	16459	0.006
P0C0L4	Complement C4-A	A	172993	124425	0.008
P55056	Apolipoprotein C-IV	A	11877	8320	0.008
Q9Y6N7	Roundabout homolog 1	A	399	625	0.008
Q5JPH6	Probable glutamate--tRNA ligase_ mitochondrial	A	2530	6414	0.009
Q13609	Deoxyribonuclease gamma	A	1671	3522	0.009

Proteins where relative abundance differed between patients with neurological and healthy controls with uncorrected P > 0.01 are shown. Proteins where relative abundance was higher in patients with neurological presentations are shown in green and lower in patients with neurological presentations are shown in red. P values less than 0.05 after FDR correction are highlighted in bold. Rank refers to the confidence in peptide identification (A-C).

Table 4-8 Protein identifications when comparing patients with hepatic presentations and healthy controls

Accession	Protein	Rank	HC	hWD	P value
P13498	Cytochrome b-245 light chain	C	132808	24943	**<0.001**
P35908	Keratin_ type II cytoskeletal 2 epidermal	A	14395	6041	**<0.001**
P00450	Ceruloplasmin	A	535109	77864	**<0.001**
Q8TF21	Ankyrin repeat domain-containing protein 24	A	1086	297	**<0.001**
Q8IWJ2	GRIP and coiled-coil domain-containing protein 2	A	2473	4442	**<0.001**
P01594	Immunoglobulin kappa variable 1-33	A	1179	626	**<0.001**
Q68BL8	Olfactomedin-like protein 2B	B	485	221	**<0.001**
P02788	Lactotransferrin	A	30020	13356	**<0.001**
Q6PF18	MORN repeat-containing protein 3	C	1432	477	**<0.001**
Q70CQ1	Ubiquitin carboxyl-terminal hydrolase 49	A	3326	814	**<0.001**
P01023	Alpha-2-macroglobulin	A	270872	763929	**<0.001**
O95251	Histone acetyltransferase KAT7	A	6530	1607	**<0.001**
A4UGR9	Xin actin-binding repeat-containing protein 2	A	10788	8105	**<0.001**
A0A0C4DH25	Immunoglobulin kappa variable 3D-20	A	1453	663	**<0.001**
P01834	Immunoglobulin kappa constant	A	321875	95606	**<0.001**
P02749	Beta-2-glycoprotein 1	A	311196	376987	**0.001**
O60811	PRAME family member 2	C	507	2241	**0.001**
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	A	10066	1507	**0.001**
Q7Z3Z0	Keratin_ type I cytoskeletal 25	A	2217	7031	**0.001**
Q8WWR9	Pancreatic progenitor cell differentiation and proliferation factor-like protein	C	1416	4309	**0.001**
Q6ZN30	Zinc finger protein basonuclin-2	A	4537	1561	**0.001**
Q8IZT8	Heparan sulfate glucosamine 3-O-sulfotransferase 5	C	781	198	**0.001**
Q13740	CD166 antigen	A	2815	1382	**0.002**
Q5JPH6	Probable glutamate--tRNA ligase_ mitochondrial	A	2530	7067	**0.002**
Q63HN8	E3 ubiquitin-protein ligase RNF213	A	19582	44663	**0.002**
P20848	Putative alpha-1-antitrypsin-related protein	B	2480	7079	0.003
A0A0B4J2D9	Immunoglobulin kappa variable 1D-13	B	2524	1023	0.003
P01859	Immunoglobulin heavy constant gamma 2	A	24994	9354	0.003
Q9NQZ6	Zinc finger C4H2 domain-containing protein	A	8355	20224	0.003

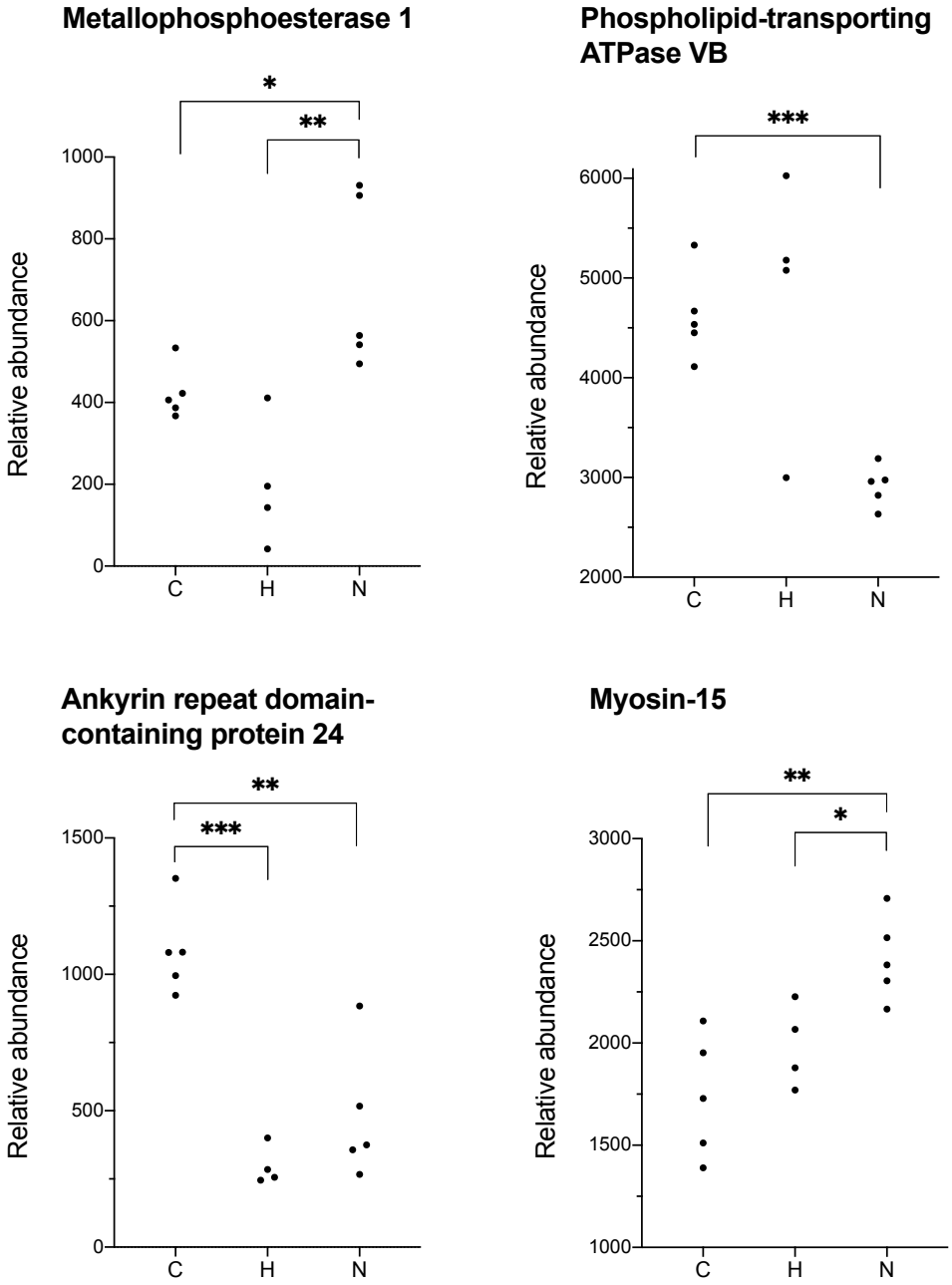
Q13609	Deoxyribonuclease gamma	A	1671	3055	0.003
Q9BXY5	Calcyphosin-2	B	896	374	0.003
A0A0B4J1V0	Immunoglobulin heavy variable 3-15	A	7528	3571	0.004
Q9UQR1	Zinc finger protein 148	B	176	30	0.004
O14733	Dual specificity mitogen-activated protein kinase 7	A	18601	35459	0.004
Q9UK55	Protein Z-dependent protease inhibitor	A	22478	17780	0.005
A0A0A0MRZ7	Immunoglobulin kappa variable 2D-26	A	9506	2720	0.005
Q02224	Centromere-associated protein E	A	3968	7656	0.005
P49798	Regulator of G-protein signaling 4	C	5822	1287	0.005
P17936	Insulin-like growth factor-binding protein 3	A	58468	75785	0.007
P15169	Carboxypeptidase N catalytic chain	A	58043	72978	0.008
P03952	Plasma kallikrein	A	65935	81548	0.008
A2RRP1	Neuroblastoma-amplified sequence	A	4155	13338	0.009
O76015	Keratin_ type I cuticular Ha8	A	2203	1526	0.009

Proteins where relative abundance differed between patients with hepatic presentations and healthy controls with uncorrected P > 0.01 are shown. Proteins where relative abundance was higher in patients with hepatic presentation are shown in green and lower in patients with hepatic presentations are shown in red. P values less than 0.05 after FDR correction are highlighted in bold. Rank refers to the confidence in peptide identification (A-C).

with hepatic presentations than controls (77,864 vs 535,108; P < 0.001 and $_{FDR}$P = 0.01). There was no difference between patients with neurological and hepatic presentations (89,723 vs 77,864; P = 0.80).

Four of the proteins identified that differed between groups were considered to be of further interest given their expression is reported to be enhanced in brain tissue on the UniProtKB 2021_03 database. These proteins all had an identification ranking of A or B. The relative abundance of these proteins in each group are shown in **Figure 4-5**. Metallophosphoesterase 1 was higher in patients with neurological presentations than patients with hepatic presentations (687 vs 198; P = 0.005 and $_{FDR}$P = 0.60) and controls (687 vs 423; P = 0.048 and $_{FDR}$P = 0.37). Phospholipid-transporting ATPase VB was lower in patients with neurological presentations than controls (2,917 vs 4,620; P = 0.003 and $_{FDR}$P = 0.03). Ankyrin repeat domain-containing protein 24 was lower in patients with neurological presentations (480 vs 1,086; P = 0.002 and $_{FDR}$P = 0.09) and hepatic presentations (297 vs 1,086; P < 0.001 and $_{FDR}$P = 0.01) than controls. Myosin-15 was higher in patients with neurological presentations than patients with hepatic presentations (2,415 vs 1,986; P = 0.02 and $_{FDR}$P = 0.60) and controls (2,415 vs 1,738; P = 0.004 and $_{FDR}$P = 0.12).

Figure 4-5 Group differences in four proteins identified by LC-MS^{E}



*Relative abundances of four proteins of interest are shown for individual participants divided into healthy controls (C), patients with hepatic presentations (H) and patients with neurological presentations (N). * = P value <0.05; ** = P value <0.01; *** = P value <0.001.*

4.4 Discussion

We have used a combination of copper indices, targeted proteomics and untargeted proteomics to identify candidate wet (fluid) biomarkers for neurological involvement in WD. I have demonstrated that plasma NfL is a promising marker of end-organ damage and have confirmed previous observations on the lack of association between urine and serum copper indices and neurological involvement. I also provide further data on EXC and its role in monitoring. Using label-free LC-MS^E on serum samples from a small subset of our cohort, we have generated preliminary data identifying several novel proteins of interest that could be taken forward for analyses with targeted proteomics in the wider cohort in the future.

Copper indices

There were no differences between NCC, EXC or UCu in patients with hepatic or neurological presentations or associations between these indices and UWDRS-N, cognitive test or psychiatric rating scale scores in stable patients. This is perhaps unsurprising given the majority of participants in the cohort are chronically-treated and, at least for NCC and UCu, there is no association between copper indices and clinical severity at initial presentation. UCu was higher in patients with active disease, as expected. I cannot draw any conclusions about the lack of difference between NCC or EXC in patients with active and stable disease given the small number of active patients.

There are several important considerations in our approach to measuring NCC, UCu and EXC:

Firstly, I calculated NCC using serum caeruloplasmin, measured with immunoturbidimetry, and serum copper, measured with ICP-MS. The limitations of using calculated NCC to measure systemic copper balance have been well described.[59] They primarily relate to variable performance of different immunoassays for caeruloplasmin, some of which cross-react with apocaeruloplasmin. [156] I feel that using calculated NCC is an acceptable approach in our cohort because all serum caeruloplasmin and copper measurements were performed using the same assays in the same laboratory.

Secondly, we used ‘on treatment’ as opposed to ‘off treatment’ urine collections. ‘Off treatment’ collections may provide a more accurate

assessment of residual copper balance and remove any confounding effects from the choice of chelating agents. However, it is our usual practice to avoid performing 'off treatment' collections in the first year of treatment when it is unlikely to alter management. Maintaining consistent treatment is important and so I chose to perform 'on treatment' collections (and exclude patients treated with liver transplantation or Zinc salts) in our analyses.

Thirdly, some authors have raised concerns about the principles underlying the EXC assay and its test-retest reliability. Solovyev *et al* recently examined the contents of the exchangeable fraction using strong-anion exchange ICP-MS.[163] They observed that it contained small amounts of caeruloplasmin and that EDTA extracted copper that was bound to caeruloplasmin. Up to 30% of caeruloplasmin-bound copper was removed from caeruloplasmin in healthy controls and 80% of caeruloplasmin-bound copper was removed from caeruloplasmin in patients with WD. The higher rate in patients with WD presumably reflects looser binding of copper to caeruloplasmin as a result of ATP7B dysfunction. They also found that the method lacks repeatability when performed on replicate serum samples, mostly because of the performance of the filters, and that the exchangeable fraction was unstable over time. The lack of differences between groups or associations with UWDRS-N scores for EXC in our cohort could therefore reflect methodological issues related to the choice of filter and sample processing times.

Targeted proteomics – N4PA assay

We have demonstrated that plasma NfL can differentiate patients with neurological and hepatic presentations, is higher in patients with active disease and, like plasma tau and UCH-L1, correlates with neurological severity in chronically-treated patients. These observations suggest blood-based biomarkers of neuro-axonal injury are promising biomarkers for non-invasively monitoring neurological involvement in WD. However, there are several important caveats to our findings and further work is required to better understand the potential role of plasma NfL measurements in a clinical setting:

Firstly, the small number of patients with active disease limits our understanding of NfL in the 'de-coppering' phase of treatment and, in the

absence of longitudinal measurements, the trajectory of NfL after treatment initiation remains unknown. Our data suggest NfL is a promising candidate biomarker that could be used for monitoring but whether it changes with time or in response to treatment is unclear. Intriguingly, the highest NfL concentration was in a participant with ongoing paradoxical neurological worsening, raising the possibility that wet biomarkers might be able to help identify this unusual phenomenon earlier.

Importantly, NfL concentrations increased with age and age-specific reference ranges are not yet available.[74] Group differences were clearer when evaluated with linear regression than with ROC curve analysis because age was a covariate, and our cohort included older adults. NfL may have greater clinical utility and detect neurological involvement at initial presentation with higher sensitivity and specificity than our results suggest given WD usually presents in adolescence or early adulthood but this remains to be tested.

There was no difference between NfL concentrations in stable patients with and without cirrhosis. However, none of these patients had features of decompensated cirrhosis and the extent to which NfL increases in WD patients with minimal or overt hepatic encephalopathy remains untested.

We prioritised sensitivity rather than specificity in our analyses because there are currently no wet (fluid) biomarkers for neurological involvement in WD. Providing proof of concept and identifying a 'signal' for a candidate biomarker was more important than refining that signal in this exploratory work.

The statistically significant group differences in UCH-L1 concentrations between patients with neurological presentations and controls despite comparable absolute values occurred because the higher UCH-L1 concentrations were in younger patients and age was included as a covariate for these analyses.

The association of NfL, tau and UCH-L1 with UWDRS-N scores implies that some stable participants had ongoing neuronal injury despite long-term maintenance therapy. One interpretation is that our definition of active disease was too restrictive and some patients classified as having stable disease may have been non-adherent or undertreated but not yet developed overt clinical worsening. While non-adherence is a common problem in chronically-treated patients,[195] this seems unlikely to be relevant given the lack of correlation

between biomarkers and copper indices and that most of the neurological deficits in these patients were longstanding. A more likely explanation is that ongoing neuroaxonal injury in chronically-treated patients does not solely depend on circulating copper levels. ATP7B is widely expressed in the brain.[13] Dysfunctional intracellular copper handling within neurons or glia or other unknown factors might also influence neuronal vulnerability in chronically-treated patients.

The observation that tau and UCH-L1 were associated with UWDRS-N scores in stable patients but did not differ between groups requires further consideration. Superficially, this might suggest that tau and UCH-L1 reflect more chronic neurological involvement. However, tau is rapidly cleared from blood and so elevated plasma tau in stable patients likely reflects ongoing neuronal injury.[196] The differential expression of these proteins within neuronal subpopulations might be relevant here: NfL is expressed in large, myelinated axons that project to deeper brain layers whereas tau is expressed in thin, unmyelinated axons of cortical interneurons.[197]

It is plausible that the pathophysiological trait that predisposes patients to more severe neurological involvement at presentation also provokes ongoing subclinical neurodegeneration in chronically-treated patients. The observation that NfL but not tau or UCH-L1 differed between groups might therefore suggest that injury to large, myelinated axons (or neuronal subpopulations with large, myelinated axons) is a more sensitive marker of ongoing, and potentially reversible, disease activity than injury to other neuronal structures (or subpopulations), which are influenced by the aforementioned pathophysiological trait. It is for this reason that we prioritise NfL over tau and UCH-L1 as the most promising biomarker for neurological involvement in WD. The lack of association between disease duration and these biomarkers is also relevant here and indicates that this pathophysiological trait is not associated with progressive neurodegeneration in chronically-treated patient, although this needs to be confirmed longitudinally.

Our observations on tau are difficult to reconcile with those of Lekomtseva *et al*. They reported that serum concentrations were higher in patients with neurological presentations than controls when measured using a

standard ELISA in a cohort of patients from the Ukraine. There are several possible explanations. Firstly, they measured tau in serum not plasma and there is some evidence that tau concentrations differ between these biofluids, even when measured with the same assays.[198] Secondly, the severity of neurological involvement in participants from the Ukranian cohort was not reported and could have been higher. Thirdly, subtle differences in sample processing affect assays for detecting proteins at such low concentrations and were not identical between cohorts.[199] It is also recognised that concentration of tau in serum is too low to reliably quantify with standard enzyme-linked immunosorbent assays.[197]

Since the publication of our findings,[200] Lin *et al* have also measured GFAP in plasma samples from 94 chronically-treated patients and 25 healthy controls using a Simoa assay.[201] Unlike in our cohort, GFAP was higher in the neurological than control group (143 vs 87 ng/L, $P < 0.001$) and higher in the neurological than hepatic group (143 vs 108 ng/L, $P = 0.005$). There were no associations between GFAP and UWDRS-N scores, UWDFRS-P scores or a semi-quantitative score for imaging abnormalities. They did not measure NfL, tau or UCH-L1. The contradictory GFAP results between our cohorts might be explained by differences in classification and age distribution: Lin *et al* classified patient according to the presence or absence of neurological features at the time of assessment and the age distribution was much lower with IQR 16-29 in their hepatic group compared to IQR 30-52 in our hepatic group.

While our observations on plasma NfL are promising, several additional steps will be required to validate it as a biomarker for neurological involvement that could be used in a clinical setting. Our results need to replicated in another, ideally larger, cohort of patients and longitudinal changes in plasma NfL in the first few years of treatment need to be examined. The latter could be achieved by including NfL as a secondary/exploratory endpoint in clinical trials. Age-specific cut-offs in healthy controls also need to be determined and these are likely to be derived from separate studies on more common neurological diseases.

I predict several potential uses for NfL in the management of WD in a clinical setting or in clinical trials. Baseline measurements might be used to

predict outcomes, i.e. as a prognostic biomarker, or guide treatment decisions around choice of agent or dose escalation plan, i.e. as a predictive biomarker. Measurements in the first few months of treatment could be used to identify patients with paradoxical neurological worsening early and as a monitoring biomarker that guides treatment escalation/de-escalation in the first year after diagnosis. They could be used to determine when a patient with WD has 'no evidence of disease activity' in the brain and therefore when doses can be safely reduced. In the longer term, they could be used to determine when patients are at risk of neurological worsening due to non-adherence or require dose adjustments. In clinical trials, plasma NfL might be used to stratify patients or as an endpoint to determine the effectiveness of treatments in newly-diagnosed patients. It could also be used to more rapidly determine whether switching treatments or amending doses improves or worsens neurological involvement in chronically-treated patients.

Untargeted proteomics

We have used label-free LC-MS^E on serum samples from a small subset of participants in the CROWD study to identify a number of proteins of interest that can be investigated further as biomarkers for neurological involvement in WD. Unfortunately, it has not been possible to complete the analysis for all fractions and the data presented in this thesis should be considered preliminary. Here, four of the proteins that we identified that have enhanced expression in brain tissue and are therefore more likely to serve as biomarkers for neurological involvement are discussed.

Metallophosphoesterase 1, also known as PGAP5, is a 45 kDa protein involved in transport of GPI-anchor proteins from the endoplasmic reticulum to the Golgi apparatus.[202] Vuoristo *et al* reported that it was selectively expressed in the brain,[203] notably the basal ganglia, although RNA sequencing data have subsequently suggested it also expressed in other organs. We observed that levels were higher in patients with neurological presentations than patients with hepatic presentations and controls. There are two other interesting characteristics to this protein. Firstly, singe nucleotide variants in *MPP1* are associated with the risk of developing bipolar affective disorder,[204] and Chen *et al* have recently reported that a missense mutation in *MPP1* is

associated with tumour recurrence in hepatocellular carcinoma.[205] Secondly, the protein has a metal binding domain for divalent cations. It is thought to bind manganese on the basis that adding EDTA abolishes enzyme activity and supplementation with manganese restores activity but copper binding has not been evaluated.[206]

Phospholipid-transporting ATPase VB is the catalytic component of the flippase complex involved in transport of glucosylceramide from the outer to the inner leaflet of lysosome membranes. It is expressed in the brain and small intestines,[207] and is thought to play an important role in maintaining lysosome membrane function in cortical neurons.[208] We observed that levels were lower in patients with neurological presentations than controls. Similar to ATP7B, this protein is a P-type ATPase, although unlike other P-type ATPases it transports phospholipids not cations. Variants in the *ATP10B* gene, which encodes this protein, have also been associated with movement disorders. In a cohort of 617 patients with PD from Belgium, 1% had loss of function mutations.[208]

Much less is known about the two remaining proteins of interest that we identified. Ankyrin repeat domain-containing protein 24 is a protein of unknown function with enhanced expression in brain tissue. There is also some expression in the testes and liver. Levels were lower in patients with neurological and hepatic presentations compared to controls. Myosin-15, or myosin heavy chain 15, has enhanced expression in brain, highest in cerebral cortex. It is also expressed in heart, skeletal muscles and retina. Levels were higher in patients with neurological presentations than those with hepatic presentations and controls.

The main limitations of the data presented here is the small numbers of samples, with only five in each group. Having identified 586 proteins in serum samples, I would expect to find group differences in 30 proteins by random chance with uncorrected P values < 0.05. Correction for multiple corrections is helpful here. Given the exploratory nature of this work, I did consider some proteins, such as metallophoshoesterase 1, to be of further interest despite FDR-corrected P values > 0.05 when comparing groups.

An additional consideration is the sensitivity of mass spectrometry to identify proteins at very low concentrations in complex samples, such as

serum or plasma. Based on our earlier observations, UCH-L1, one of the most abundant proteins in the brain, is present in serum samples at concentrations $>10^9$ times lower than caeruloplasmin. Metallophosphoesterase 1, the least abundant protein of interest that we identified had a relative abundance around 10^3 times lower than caeruloplasmin. It may therefore be extremely difficult to identify proteins that originate in the brain in blood samples using these methods. Differences in protein concentrations between our neurological and hepatic groups may reflect peripheral expression of these proteins, even if are primarily expressed in the brain, unless the protein is actively secreted from CSF into blood. Nonetheless, identifying indirect markers of neurological involvement, which reflect damage to organs other than the brain, may still be valuable given our limited understanding why only some patients develop neurological involvement.

The next steps are to complete the analysis of the remaining fractions and re-evaluate which proteins might be considered of further interest. We will then use these findings to make a high throughput, multiplexed and targeted proteomics assay through which we can quantify 20-30 of these proteins of interest in samples from the rest of the cohort using MS-MS on a triple quadrupole analyser, as previously described.[209]

5. Neuroimaging biomarkers

5.1 Introduction

WD is associated with neuroimaging abnormalities on T1-weighted, T2-weighted (or FLAIR), diffusion-weighted and susceptibility-weighted imaging. Hyperintense signal abnormalities in the basal ganglia, thalamus and/or brainstem are seen on T2-weighted sequences in 90% of cases with neurological presentations and are helpful diagnostically being uncommon among other diseases that present with movement disorders.[57] These and other neuroimaging abnormalities can also be seen in patients with hepatic presentations, without overt neurological or psychiatric symptoms.[57,210,211] The role of neuroimaging in predicting outcomes and monitoring response to treatment is less clear and not mentioned in current guidelines.[59,153,212] This chapter outlines the basic principles behind MRI and provides an overview of methods for quantitative MRI analyses. The literature on neurological involvement in WD relevant to each of the aforementioned MRI modalities is then reviewed. Finally, previous neuroimaging studies on cognitive deficits and psychiatric features in WD and our neuroimaging objectives are discussed.

Fundamentals of MRI

MRI is a powerful and versatile tool for imaging the brain. A typical scanner consists of a collection of electromagnetic coils that are used to manipulate the magnetic state of hydrogen nuclei (protons) in water molecules. The largest coil, the strength of which is referred to in units of Tesla (T), is always switched on and creates the B_0 field.[213] This aligns protons within this field so that they tend to spin around the same longitudinal axis and in the same transverse (clockwise/anti-clockwise) direction. Most align in the same longitudinal direction as the B0 field and some align in the opposite direction, creating a net magnetisation in the direction of the B_0 field.

While scanning, radiofrequency coils are used to rotate the net magnetisation, typically perpendicular to the B_0 field, by transmitting radiofrequency pulses or pairs of pulses in opposite directions. Protons ‘flip’ when exposed to a radiofrequency pulse at their resonance frequency. Radiofrequency coils then measure the immediate changes in local magnetic

fields, including the time taken for the net magnetisation to revert to its original configuration, known as the relaxation time. Gradient coils are concurrently used to change the strength of the field in a linear fashion in x, y and z directions. This means that the timing of when protons are exposed to their resonance frequency and then flip depends on their spatial location. This is used to generate a 3D image, which is divided into voxels.[213]

For most commonly used structural MRI sequences, including T1-weighted, T2-weighted, FLAIR and susceptibility-weighted sequences, the signal intensity within each voxel is derived from the relaxation time.[214] By changing timings related to the pulse sequence, such as the repetition time (TR) and echo time (TE), the image can be biased towards measuring relaxation in the longitudinal or transverse directions. In a T1-weighted sequence the image is acquired with an emphasis on the longitudinal relaxation time. This differs between grey matter (GM), white matter (WM) and CSF and so T1-weighted sequences are useful for delineating specific brain structures and changes in volume over time, i.e. atrophy. In T2-weighted and FLAIR sequences the image is acquired with an emphasis on the transverse relaxation time. Pathological changes that increase the water content of tissue or disrupt myelin lead to increased transverse relation times and areas of increased signal intensity. These are sometimes referred to as lesions or white matter hyperintensities (WMHs), even though areas of increased signal intensity can also be seen in GM. FLAIR sequences include an additional radiofrequency pulse to cancel out signal arising from free fluid, i.e. CSF, on T2-weighted images thereby improving contrast for periventricular lesions.[214]

Susceptibility is the extent to which something can become magnetised when placed in a magnetic field and can be measured using a gradient-echo (GRE) sequence. This involves applying a re-phasing gradient, in the opposite direction to initial radiofrequency pulse, while measuring the relaxation in the transverse direction.[215] This unmasks susceptibility gradients in tissues (and other magnetic field inhomogeneities) and is used to generate T2* images. T2*-weighted GRE sequences usually consist of multiple 2D images whereas more advanced susceptibility-weighted sequences are acquired in 3D and produce separate magnitude and phase images, which are then combined. The magnitude image reflects the relaxation time (T2* decay) whereas the

phase image reflects magnetic field perturbations seen by spins.[216] The latter is useful for detecting field inhomogeneities related to bone at the skull bases and air in the sinuses and allows paramagnetic effects, related to deposition of metals such as iron (or copper), and diamagnetic effects, related to myelin, to be differentiated.

Diffusion-weighted imaging is used to assess tissue microstructure by measuring diffusion of water molecules, i.e. protons, in different directions. This is achieved by acquiring numerous images and applying diffusion-encoding gradients in varying directions (using the gradient coils) during each acquisition.[213] Water molecules that are moving in the direction of the diffusion-encoding gradient experience a change in their resonance frequency and phase proportional to the amount of the diffusion in that direction. Water molecules that are not moving in the direction of the diffusion-encoding gradient are unaffected. By combining these signals from all acquisitions, the orientation and magnitude of diffusion within given voxels can be determined.

Quantitative MRI analyses

MRI data can be analysed using qualitative, semi-quantitative and quantitative methods. Qualitative approaches are descriptive and, at their most simple level, can be used to describe whether a specific abnormality is present or absent in a specific brain region for a given sequence. Semi-quantitative approaches usually involve applying a visual rating scale to derive an ordinal measure of the severity of an abnormality and may be combined from different brain regions or sequences to form an overall score. A number of quantitative methods for analysing T1-weighted and diffusion-weighted data have been developed, some of which have been incorporated into imaging analysis pipelines and widely used to study diseases of the brain over the last two decades. Approaches for quantitatively analysing WMHs and susceptibility-weighted imaging abnormalities have been developed more recently.

Several approaches can be used to quantitatively analyse changes in brain volume using T1-weighted data. The volume of specific regions-of-interest (ROI) can be calculated after segmenting the brain, either manually or using automated pipelines. The latter can be divided into atlas-based approaches that focus on aligning atlas priors, algorithmic approaches that

use signal intensity to determine ROI boundaries and learning-based approaches that depend on manually-segmented training datasets.[217-219] Voxel-based morphometry (VBM) is a type of analysis that can be used to identify regional changes in GM volume (or density) across the whole brain in an unbiased manner.[220] It involves spatially normalising T1-weighted images from all study participants in the same stereotactic space and smoothing the GM segments. Voxel-wise statistical tests are then used to compare the smoothed images between two groups or participants, or test associations with specific clinical measures, using a generalised linear model with correction for multiple testing. Other methods for quantitatively analysing T1-weighted data include cortical surface modelling, which can be used to measure cortical thickness and/or curvature.

Importantly, brain atrophy should ideally be determined by measuring serial changes in brain volume longitudinally. Regional brain volumes vary between individuals and are influenced by other factors, including age and sex, and so brain volumes are treated as a proxy for brain atrophy and age and sex included as covariates in cross-sectional neuroimaging studies.

The most widely used approach for modelling diffusion-weighted imaging data is using an analytical framework known as diffusion tensor imaging (DTI).[221] More advanced methods such as neurite orientation density and dispersion imaging (NODDI) and fixel-based analysis are increasingly used.[221,222] In DTI, diffusion within each voxel is modelled using a tensor. Three eigenvectors (ε1, ε2 and ε3), which define the direction of diffusion, and three eigenvalues (λ1, λ2 and λ3), which define the magnitude of diffusion in those three directions, are calculated. Several parameters can be derived from these values. The simplest is axial diffusivity (AD), which is the amount of diffusion (λ1) in the principal diffusion direction (ε1). The radial diffusivity (RD) is the amount of diffusion perpendicular to this axis. The mean diffusivity (MD), refers to the average amount of diffusion in all directions in a given voxel. Fractional anisotropy (FA) is the 'directionality' within each tensor, a scalar value between 0 (where diffusion occurs equally in all directions) and 1 (where diffusion occurs along one axis). A voxel in WM, which mostly contain axons, will therefore have an elongated tensor with high FA and AD and low RD. A voxel in CSF will have equal diffusion in all directions and a spherical tensor with low

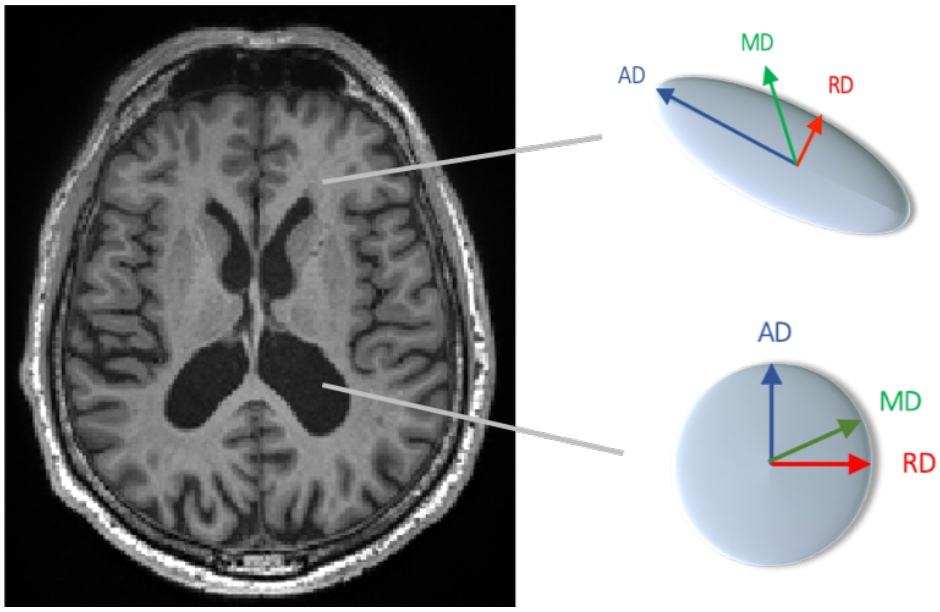
FA and equal AD, RD and MD as shown in **Figure 5-1**. This makes DTI a particularly powerful technique for studying changes in WM microstructure in disease states.

Similar to analysis of T1-weighted imaging data, an ROI-based approach can be used to determine the average FA or MD in a specific brain region. Tractography involves using DTI data to the determine the architecture of WM tracts and therefore connections between brain regions. In tract-based spatial statistics, FA maps are used to determine the spatial localisation of WM tracts and then align images across multiple subjects to enable statistically powerful comparisons of DTI parameters within WM tracts between groups of subjects.[223]

Quantitative analysis of WMHs on T2-weighted or FLAIR sequences has, up until recent years, relied on manual segmentation, which is time-consuming and has high intra- and inter-operator variability.[224] Automated lesion segmentation pipelines have now been developed,[225] mostly driven by the need to quantify inflammatory WMHs in multiple sclerosis. These can be supervised, relying on atlas-based approaches or training datasets, or unsupervised. Total WMH volumes in specific brain regions or the distribution of WMHs can then be analysed.

Finally, quantitative analysis can also be performed on susceptibility-weighted imaging. However, a major problem is that signal intensity in a given voxel on a standard SWI acquisition reflects the sum of the susceptibility from surrounding structures and other B_0 field inhomogeneities. Neither magnitude images, phase images (with or without high-pass filtering) or the relatively straightforward combination of these used to generate standard SWI images can be used alone to reliably quantify the susceptibility of local tissues.[226] Quantitative susceptibility mapping (QSM) reconstruction can be performed to overcome these issues. This typically involves accurate coil combination, to deal with issues related to the varying sensitivity of the radiofrequency coils used to generate the images, followed by phase unwrapping and background field removal steps, to disentangle the contribution of surrounding structures and isolate local field offsets.[227] These must then be deconvolved to retrieve the underlying tissue susceptibilities. A number of QSM reconstruction

Figure 5-1 A schematic diagram of two tensors



An elongated tensor, where axial diffusivity is greater than radial diffusivity, modelling diffusion in a voxel containing white matter and a spherical tensor, where axial and radial diffusivity are equal, modelling diffusion in a voxel containing cerebrospinal fluid are shown.

pipelines have been developed, most of which depend on an approach known as Morphology Enabled Dipole Inversion (MEDI).[228]

Similar to other sequences, reconstructed QSM data can be analysed using an ROI-based approach where a segmented T1-weighted image is used to determine the anatomical boundaries on QSM maps. Whole-brain approaches, similar to that used with VBM, have also recently been proposed.[229]

Brain atrophy in WD

Volume loss is common in newly-diagnosed patients with WD. Retrospective neuroimaging studies have reported cortical atrophy in 11-70%, cerebellar atrophy in 36-52% and brainstem atrophy in 66% however the proportion of patients with neurological presentations in these cohorts varied and atrophy was determined by visual inspection alone and at a single timepoint.[57,58,230]

Several MRI scoring systems that include semi-quantitative measures of brain volume have been devised.[43,230,231] Dusek *et al* recently introduced a sophisticated scale that includes an acute toxicity score, based on white matter hyperintensities on T2-weighted imaging, and a chronic damage score, based on volume loss on T1-weighted imaging and abnormal susceptibility on T2*-weighted imaging.[232] Cortical and cerebellar volume loss are visually assessed whereas central and brainstem volume loss are determined by measuring the width of the third ventricle and antero-posterior midbrain diameter, respectively. UWDRS-N scores correlated with the chronic damage score in 39 newly-diagnosed patients at baseline and after two years of treatment and post-hoc testing indicated that this association was dependent on volume loss not abnormal susceptibility on T2*-weighted imaging.

Quantitative MRI analyses have provided important insights on the pattern of regional brain volume loss in WD in recent years. Several volumetric studies using ROI-based approaches have shown that the volumes of the caudate, putamen, pallidum, thalamus, amygdala, accumbens, red nucleus and substantia nigra are lower in patients with WD than healthy controls.[233-238] The volume of some cortical ROI including the occipital fusiform, cingulate and superior frontal gyri and basal forebrain have separately been shown to be lower in patients with WD with neurological presentations than

controls.[234,237,239] Song *et al* found that the total WM volume and the volume of the lateral and third ventricles were also reduced in these patients, in addition to ROI in GM;[237] the greatest % changes in volume was seen in the superior fronto-occipital fasciculi followed by the putamen then external capsule. Tinaz *et al* and Viverios *et al* found that the subcortical ROI were lower in patients with than without neurological signs.[210,238]

Reports on the association between ROI volumes and neurological severity are conflicting: Zou *et al* and Song *et al* found that volumes of the putamen and pallidum negatively correlated with neurological severity measured using the modified Young scale and global assessment scale (GAS), respectively. Yuan *et al* found a negative correlation between UWDRS-N scores and the volume of the thalamus only and Tinaz *et al* did not find associations with any ROI volumes. Smolinksi *et al* used tissue-type segmentation to study brain volumes in 48 treatment-naïve patients.[240] UWDRS-N scores were strongly correlated ($r > 0.6$) with the overall volume of the brain parenchyma and moderately correlated ($r > 0.3$) with peripheral GM, deep GM and WM volumes. NCC was also correlated with the brain parenchymal volume.

Several whole-brain techniques for analyzing regional brain volume loss have also been applied. Using VBM, Stezin *et al* compared GM volumes in 10 adolescent patients with neurological presentations and 11 age-matched controls.[81] There were differences in the caudate nuclei bilaterally and right globus pallidus. Using deformation-based morphometry, Dusek *et al* recently showed that patients with WD have lower brain volumes than controls in a confluent region encompassing the basal ganglia, thalamus, brainstem, internal capsule and large area of the frontoparietal WM. UWDRS-N scores correlated with the volume of the putamen, pallidum and surrounding WM including in the midbrain, pons and cerebellar peduncles. Using surface-based morphometry, they also found that WD patient had lower cortical thickness in primary motor, premotor and visual cortices however there was no association between cortical thickness and UWDRS scores.

These studies suggest that several different measures of brain volumes are strongly associated with neurological severity in newly-diagnosed and chronically-treated patients. However, a direct comparison between patients

with neurological and hepatic presentations, to determine the neuroradiological sequelae of any initial brain injury, has not been performed. While a number of studies have demonstrated volume loss in subcortical ROI, none have examined the relationships between the volume of these ROI, i.e. their structural covariance. An unbiased, whole-brain approach using VBM has also only been performed in small number of patients and was not correlated with neurological severity.

White matter hyperintensities

WMHs in the basal ganglia, thalamus and brainstem are characteristic of Wilson's disease.[58,230,241,242] In the largest neuroimaging study on WD, Li *et al* reported the frequency of these lesions in specific brain regions in 535 patients with WD from China demonstrating that WMHs are common in patients with and without neurological or psychiatric symptoms.[58] Of 452 with neurological presentations and 69 patients with hepatic presentations, they were seen in the caudate in 34% and 37%, putamen in 75% and 83%, thalamus in 35% and 39%, midbrain in 40% and 50%, pons in 22% and 26% and cerebellum in 2% and 10%, respectively. Oder *et al* suggested that WMHs in the putamen were associated with dystonia whereas those in the thalamus are associated with tremor/ataxia in their factor analysis.[132,243]

Data on the relationship between the severity of neurological involvement and WMHs are conflicting. A number of semi-quantitative scales for measuring the severity of WMHs in WD have been proposed.[230,232,241] Sinha *et al* used a scoring system that combined the presence of lesions and volume loss in 10 brain regions to determine an overall score for the severity of neuroradiological findings.[244] They found scores correlated with the neurological symptoms score in 100 predominantly chronically-treated patients. Pulai *et al* subsequently graded neuroradiological abnormalities between 0 and 8 according the number of brain regions affected on T1-weighted, T2-weighted and FLAIR sequences.[245] They found this correlated with the Chu staging for neurological severity in 38 newly-diagnosed patients with WD, although the specific neuroimaging abnormalities identified were not reported.[241]In the scoring system recently proposed by Dusek *et al*, scores between 0 and 2 are given for T2-weighted or FLAIR hyperintensity in the

caudate, putamen, thalamus, midbrain, pons and 'other areas' and summed to calculate the acute toxicity score. This did not correlate with UWDRS-N scores at baseline or after 24 months of treatment in 39 newly-diagnosed patients.

Using visual assessment alone, a reduction in lesion load is seen after initiation of chelation therapy in most patients, [246-248] although lesions may persist or progress despite clinical improvement,[249-251] or resolve without any concomitant clinical improvement.[251] Dusek *et al* observed that 13 of 21 patients had an improvement in the acute toxicity score of ≥ 2 with treatment. Two patients deteriorated neurologically; one had a decrease and the other had an increase in the acute toxicity score, the latter developing periventricular and subcortical WMHs at follow up. Da Costa *et al* reported longitudinal changes over up to eight years in 18 patients and found that lesions were most likely to improve in the centrum semiovale and substantia nigra (66% of cases) and least likely to improve in the putamen (11% of cases).[252] They found that a third of patients, most of whom were taking zinc salts, had neuroradiological worsening despite clinical improvement.

Overall, T2-weighted lesions are a useful diagnostic marker but data on their associations with neurological severity in newly-diagnosed and chronically-treated patients and with clinical response to treatment are conflicting. Lesion segmentation has not, to our knowledge, been attempted in patients with WD and quantitative analysis of lesion volumes has not therefore been performed.

Diffusion-weighted imaging

The pattern of DTI abnormalities in the GM and WM of patients with WD has been explored in several studies, mostly using pre-defined ROI. However, a range of methods related to image-acquisition, pre-processing and analysis have been employed and results of some studies are conflicting. Inferences on the neuropathological basis for abnormal DTI parameters in WD, discussed later in the chapter, also vary.

FA appears to be reduced in WM and subcortical GM in patients with WD compared to controls.[82,253,254] There is also some evidence that FA in the basal ganglia and thalamus are lower in patients with neurological than hepatic

presentations and that different neurological phenotypes may be associated with low FA in specific subcortical structures.[255] While no associations between neurological rating scales and FA have been demonstrated, Song *et al* have used neurite orientation and dispersion density imaging (NODDI), which uses a more sophisticated model of tissue microstructure, to disentangle observation on FA in GM in WD.[256] They found that the intracellular volume fraction (Vic), which reflects neurite density, and orientation dispersion index (ODI), which reflects the 'sprawling' nature of dendrites in GM, in the basal ganglia were strongly negatively correlated with neurological severity.

Abnormalities in diffusivity appear to be more complex. Several studies have shown that MD is increased in the WM and subcortical GM nuclei of patients who have recently started treatment,[82,231,253,254,256] however there are several reports of drug-naïve patients with neurological presentations having low MD in subcortical GM.[257,258] In one of the earliest reports on DWI in WD, Sener described one such patient with low apparent diffusion coefficient (ADC), which is equivalent to MD, in the striatum at presentation. They went on to develop abnormally high ADC on a repeat scan after paradoxical worsening.[257] Hu *et al* recently examined compared DTI parameters between 22 chronically-treated patients with WD and 26 healthy controls using a tract-based ROI approach. They found that FA was decreased and AD, MD and RD was increased in limbic WM tracts and the superior longitudinal and uncinate fasciculi. RD and MD were also increased in the inferior longitudinal and fronto-occipital fasciculi.

Positive correlations between the neurological symptom score and MD in the anterior limb of the internal capsule and between a measure of disability and MD in the lobar WM have been reported,[253] as has a positive correlation between neurological severity and MD in the putamen and pallidum.[256] UWDRS neurological scores have not been correlated with DTI measures.[234] In one of the few prospective imaging studies on WD, Lawrence *et al* used TBSS to analyse DTI parameters within WM tracts using an unbiased, whole brain approach.[82] They showed that FA increased and MD, AD and RD decreased over 24 months in a mixed cohort of 17 drug-naïve patients and 18 treated patients who clinically improved with treatment.

On the basis of these findings, DTI parameters, particular when measured in WM tracts, appear to be a promising state biomarker for neurological involvement in WD. However, the significance of abnormally high or low MD early in the disease course and the association between diffusivity and the severity of neurological involvement is unclear. Whole-brain approaches for analysing DTI abnormalities have only been used in a single study and associations with clinical or biochemical findings were not tested.

Susceptibility-weighted imaging

Abnormal susceptibility in subcortical GM nuclei is common in WD. Tinaz *et al* recently reported that 94% of patients with neurological presentations and 29% of patients with hepatic presentations had hypointense lesions in the basal ganglia that were visible on a standard SWI sequence in a cohort of 30 chronically-treated patients.[238] Quantitative analyses have demonstrated susceptibility-related abnormalities occur in the basal ganglia, thalamus, red nucleus and substantia nigra of patients with WD using several ROI-based approaches. Some authors report this using increased (positive) susceptibility values from QSM reconstructions,[211,259-262] whereas others describe increased signal on R2* maps (the inverse of T2* maps),[260] or low phase values taken directly from standard SWI sequences.[77,263,264] All of these are consistent with increased paramagnetic mineralisation.

Several groups have reported differences between patients with neurological and hepatic presentations,[259,262,263] and between patients with hepatic presentations and healthy controls.[211,262] Fritzsch *et al* noted that conventional MRI sequences, including standard SWI sequences, appeared normal in five patients with hepatic presentations who had increased susceptibility in the substantia nigra, red nucleus, pallidum and caudate of QSM maps when acquired using 7T MRI.[211]

Data on the relationship between subcortical SWI abnormalities and neurological severity are conflicting. Dusek *et al* did not find any association between UWDRS-N scores and susceptibility values in subcortical ROI in 29 chronically-treated patients.[235] They manually segmented the ROI on co-registered T1-weighted images. In contrast, Yuan *et al* found that susceptibility values in seven subcortical nuclei did correlate with UWDRS-N values.[236] They

used QSM maps, where anatomical boundaries are less well demarcated, to manually segment ROI. This may have introduced bias and help explain the other unexpected observation that ROI volumes did not correlate with UWDRS-N scores. This group published findings from another cohort in the same year and, using similar methods, reported susceptibility values in caudate, putamen and red nucleus correlated with UWDRS-N values.[260]

Longitudinal studies using QSM in WD have not been performed with the exception of single case report: Zaino *et al* described a newly-diagnosed patient with a severe neurological presentation who was treated with Zinc salts.[265] QSM values in the basal ganglia (normalised to CSF) improved from 92 to 23 ppm over six years.

Importantly, increased susceptibility in WD, at least in chronically treated patients, appears to be driven by iron not copper. In a post-mortem 7T MRI study, Dusek *et al* acquired T2*-weighted images and, using quantitative densitometry, measured the content of both metals in brain tissue from nine WD and six control cases.[266] Copper and iron content was increased in the putamen and pallidum of WD cases but susceptibility values (and histopathological severity) correlated with iron content only. The authors suggest that cortical iron deposition might be increased in patients with neurological presentations. Cortical SWI abnormalities have, however, only been described in a case series of two patients with neurological presentations.[267]

Overall, these findings suggest that abnormal susceptibility in subcortical GM nuclei is common in WD and may be an early feature of neurological involvement. This appears to be driven by iron deposition but I cannot exclude the possibility that copper deposition also contributes in untreated patients. Susceptibility in cortical regions has not yet been examined using QSM and, although less likely to provide a useful biomarker in a clinical setting, may provide further insights on disease mechanisms.

Cognitive deficits and psychiatric features

The association between cognitive deficits and neuroimaging abnormalities in WD has been investigated in a few studies only. Patients with more widespread T2-weighted lesions and/or volume loss appear to have more

marked cognitive deficits. Seniow *et al* divided 50 patients with neurological presentations into those with neuroimaging abnormalities localised to the basal ganglia and those with more diffusive abnormalities.[44] The former performed better on only a few tests including comprehension, digit span, object assembly and block design. This led the authors to conclude that basal ganglia dysfunction is the primary cause of cognitive impairment in WD. Frota *et al* subsequently investigated the association between a semi-quantitative MRI scale and cognitive deficits in a cohort of 20 patients.[268] Their score included seven points for WMHs, six points for T2-weighted hypointensities and two points for volume loss. The number of cognitive tests affected was highly correlated with the MRI score when subscores for T2-weighted hypointensities were excluded ($r=0.72$, $P<0.001$). Associations with individual test scores were not reported.

One group has examined associations between cognitive deficits and DTI parameters in a series of recent reports. Dong *et al* performed the MMSE, and the event-based prospective memory (EBPM), digit span and verbal fluency tests in 30 patients with WD.[269] EBPM measures performance in a word selection task. The authors explain that TBSS was used but single FA values for the caudate, putamen, thalamus, brainstem and 'white matter' are reported in the results and it is unclear how these results were derived. Nonetheless, they report a negative correlation between EBPM and digit span scores and caudate, putamen and thalamus, albeit without correction for multiple testing. There was no association between cognitive tests and FA in the 'white matter' ROI.

More recently, they tested the association between DTI parameters in specific WM tracts, identified using fibre tracking, and scores for the EBPM and time-based prospective memory (TBPM) tests in 22 patients with WD.[234] TBPM measures performance in identifying specific time intervals while using a stopwatch but being distracted with another task. AD in the left inferior longitudinal fasciculus was positively correlated with EBPM scores and in the left and right uncinate fasciculi was negatively correlated with EBPM scores. FA in the right uncinate fasciculus was negatively correlated with TBPM.[234] They also performed a volumetric ROI analysis and did not find an association with MMSE, EBPM or TBPM scores. In a subsequent paper, the same group

did not find an association between the volume of the basal forebrain and these scores.[239]

The neuroanatomical correlates of psychiatric features in WD are even less clear. In a factor analysis, Oder *et al* reported that depression was grouped with dilatation of third ventricle and organic personality syndrome was grouped with lesions in the putamen and pallidum. Direct associations were not tested. Song *et al* did not find any correlation between ROI volumes and the Hamilton Rating Scale for Depression in a volumetric study.[237] The only neuroimaging correlate for psychiatric symptoms in WD that has, to our knowledge, been identified was using serotonin transporter density imaging: Eggers *et al* reported a negative correlation between Hamilton Rating Scale for Depression score and pre-synaptic serotonin transporter density in the thalamus-hypothalamus region.[270]

Neuroimaging objectives

Our understanding of imaging biomarkers for neurological involvement in WD has advanced in recent years, mostly through quantitative analyses of regional brain volume on T1-weighted images. However, the majority of quantitative studies to date have taken an ROI-based approach focussing on subcortical GM nuclei, despite cortical GM and diffuse WM abnormalities being common in WD. Few cohorts were stratified according initial presentation which, as discussed in Chapter 3, is needed to determine the neuroradiological consequences of any initial brain injury and correlations with copper indices were seldom tested. Associations with cognitive deficits and psychiatric features in WD have not been systematically investigated. The relationship between disease duration and neuroimaging abnormalities in chronically-treated patients also need to be examined.

I therefore aimed to use a combination of quantitative, whole-brain analyses to determine the neuroradiological consequences of the initial brain injury by comparing patients with neurological and hepatic presentations and then identify differences between patients with active and stable disease. This included VBM on T1-weighted acquisitions,[271] TBSS on diffusion-weighted data,[272] an analysis of the volume and distribution of WMHs measured using a lesion segmentation pipeline applied to FLAIR images,[273] and a relatively

novel pipeline for performing whole-brain QSM analyses on SWI data.[274] I also performed an ROI-based volumetric analysis on T1-weighted data. Using the same methods, I then aimed to identify neuroanatomical correlates for movement disorders, cognitive deficits and psychiatric features in chronically-treated stable patients, in addition to testing associations with copper indices and the novel wet biomarkers described in Chapter 4.

5.2 Methods

The quantitative neuroimaging analyses described below were performed on T1-weighted (structural), FLAIR, diffusion-weighted and susceptibility-weighted imaging data, as discussed in our recent publication.[275] I performed all data pre-processing/processing steps and statistical analyses, with the exception of running the geodesic information flow and Bayesian model selection pipelines, as described in the statement of contributions.

For each analysis, group differences between patients with neurological and hepatic presentations and between patients with active and stable were examined. Associations with clinical rating scales, copper indices, wet biomarkers and disease duration were then tested in stable patients. Age and sex were used as covariates in all analyses. UWDRS-N was included as a covariate when comparing patients with active and stable disease given patients with active disease had higher UWDRS-N scores. Where associations with cognitive tests or wet biomarkers were identified, the analysis was repeated after including UWDRS-N scores as a covariate.

One participant with a neurological presentation who was considered to have stable disease declined to undergo MRI during the research visit due to a relative contraindication. One participant with a hepatic presentation had T1-weighted acquisitions only after feeling claustrophobic in the MRI scanner. Two further patients with neurological presentations did not complete the DWI acquisitions.

T1-weighted (structural) imaging

T1-weighted images were bias-corrected and parcellated using the geodesic information flow (GIF) pipeline,[218] which is based on atlas propagation and label fusion. The brainstem was subsequently segmented using a customized version of a FreeSurfer module that relies on a probabilistic atlas of the brainstem and its neighbouring brainstem stuctures.[276] The volume of eight subcortical ROI, including the caudate, putamen, pallidum, thalamus, amygdala, midbrain, pons and cerebellum, were extracted. These ROI were chosen based on the previous work by Zou *et al*; the red nucleus and substantia nigra are not parcellated in the GIF pipeline and so we included the midbrain, pons and cerebellum separately. The volumes of bilateral ROI were

combined and all ROI were then expressed as a percentage of total intracranial volume (TIV). All segmentations were visually checked for quality.

ROI volumes were compared between groups and correlated with clinical rating scales, copper indices and wet biomarkers in stable patients using linear regression. P values for coefficients of interest both with and without false discovery rate (FDR) correction were calculated in R (version 3.6.0, http://www.R-project.org). Structural covariance between ROI volumes was assessed using Pearson correlation coefficients.

Voxel-based morphometry (VBM) was then performed using Statistical Parametric Mapping (SPM12, version 7771, http://www.fil.ion.ucl.ac.uk/spm), running in Matlab R2019b (Math Works, USA).[220] T1-weighted (structural) images were segmented into GM, WM and CSF standard procedures and spatially normalised using the fast-diffeomorphic image registration algorithm.[277] GM and WM segments were transformed into MNI152 space (Montreal Neurological Institute, McGill University, Canada), modulated and smoothed using a Gaussian kernel with 8 mm full-width at half maximum to create pre-processed GM tissue maps. All segmentations were visually checked for quality. A study-wise mean template based on bias-corrected T1-weighted images from study participants was created.

The pre-processed GM tissue maps were fitted to factorial design analyses to identify group differences in GM volumes and multiple regression analyses to identify associations with clinical rating scales, wet biomarkers and disease duration in stable patients. TIV, calculated in SPM, was included as a covariate, in addition to the covariates described above. Threshold-free cluster enhancement was applied using the CAT12 toolbox with statistical thresholds set at $P < 0.05$ with family-wise-error (FWE) correction after 10,000 permutations. A minimum cluster size of 20 voxels was set and statistical maps were overlaid onto the study-wise mean template. Specific gyri or cortical regions were identified using the Harvard-Oxford cortical and subcortical atlases in the Functional MRI of the Brain (FMRIB) software library (FSL, v6.0.3).

FLAIR imaging

WMHs were segmented using Bayesian model selection, an automated lesion segmentation tool applied to rigidly co-registered T1-weighted and FLAIR sequences.[278] A Gaussian mixture model with dynamically evolving number of components was fitted to the data, modelling simultaneously healthy and non-expected observations. WMH-related measures were introduced to the model through subject-specific statistical atlases obtained using the GIF pipeline. After convergence, the model was used to select candidate lesion voxels whose aggregation in connected components was automatically classified as lesion or artefact. WMH segmentations were then visually inspected and flagged if there were significant segmentation errors. This quality control stage was used to make improvements to the automated WMH segmentation, thereby maximising the number of usable segmentations.

The volume of WMHs within 40 anatomically-defined regions were calculated for each participant.[279] WM was separated into four equidistant layers between the ventricular surface and the cortical GM/WM interface. These were then divided into left and right frontal, temporal, parietal and occipital lobes using the GIF parcellation. The basal ganglia and infratentorial regions were considered separately. The volume of WMHs within each region was $\log_e$-transformed to reduce skewness. A linear regression model was used to identify group differences in the $\log_e$-transformed total volume of WMHs, $\log_e$-transformed volume of WMHs in the six brain regions (basal ganglia, frontal, temporal, parietal, occipital and infratentorial) and $\log_e$-transformed volume of WMHs in the 40 brain regions. Associations between the latter and clinical rating scales, wet biomarkers and disease duration were then tested in stable patients. Total intracranial volume (TIV), calculated in SPM, was included as a covariate, in addition to the covariates described above. P values for coefficients of interest were calculated in R. FDR-corrected P values were calculated when testing associations across 40 regions and these were summarised in bullseye plots to illustrate their anatomical distribution.[279]

Diffusion-weighted imaging

FSL was used to pre-process DWI data prior to fitting the single tensor model, resulting in volumetric diffusion tensor imaging (DTI) data. DTI datasets were then analysed using tract-based spatial statistics (TBSS).[223] Pre-processing included EDDY to correct for motion and eddy-currents with outlier replacement enabled. FUGUE was applied to correct for distortions using fieldmaps. Tensors were fitted using DTIFIT and FA, MD, AD and RD maps were generated, skeletonised and aligned using TBSS. Design matrices for identifying group differences or associations with clinical rating scales, copper indices and wet biomarkers were generated using the general linear model. Finally, RANDOMISE was used to perform non-parametric permutation analyses based on each design matrix. Covariates were mean-centred and 10,000 permutations of the data were carried out. The threshold-free cluster enhancement algorithm was used to identify clusters of voxels with FWE-corrected P values < 0.05.[280] Clusters of increased or decreased FA, MD, AD and RD were then overlaid onto a mask of the WM skeleton (created using the mean skeletonised FA map) and the MNI152 template.

Susceptibility-weighted imaging

Quantitative susceptibility maps (QSM) were reconstructed from susceptibility-weighted images using a Multi-Scale Dipole Inversion (MSDI)-based pipeline for coil-combined, multi-gradient echo data in QSMbox (https://gitlab.com/acostaj/QSMbox).[227] Pre-processing steps included unwrapping of complex 3D phase data using a discrete Laplacian method followed by background field removal using Laplacian boundary extraction and variable spherical mean filtering. All steps were applied using default settings. Whole-brain analyses were performed with the QSMexplorer pipeline (https://gitlab.com/acostaj/QSMexplorer).[281] A study-wise space was created from T1-weighted sequences using Advanced Normalisation Tools (ANTs). Bias-corrected magnitude images were then used to transform the quantitative susceptibility maps to the study-wise space. Absolute susceptibility maps smoothed with a 3 mm standard deviation 3D Gaussian kernel were used to identify group differences in susceptibility and the associations with clinical rating scales, wet biomarkers and disease duration in stable patients.

RANDOMISE was used to perform nonparametric permutation analyses based on each design matrix. Covariates were mean-centred and 10,000 permutations were performed. The GM and WM segments generated in SPM12 were combined to mask the absolute maps. Threshold-free cluster enhancement was enabled to identify clusters of voxels with FWE-corrected P values <0.05. Clusters were then overlaid onto the study-wise template for visualisation purposes.

5.3 Results

T1-weighted (structural) imaging

Structural covariance between ROI volumes is shown in **Table 5-1**. There was a positive correlation between the volume of all ROI, except the amygdala. The correlations between caudate, putamen and pallidum volumes were particularly strong with Pearson correlation coefficients >0.9.

Group differences in ROI volumes are shown in **Table 5-2**. Caudate, putamen, pallidum, thalamus, midbrain, pons and cerebellum volumes were lower in patients with neurological than hepatic presentations. VBM results showing differences in GM volumes between patients with neurological and hepatic presentations are shown in **Figure 5-2**. Patients with neurological presentations had lower GM volumes in the bilateral caudate, putamen and nucleus accumbens, left orbitofrontal and central opercular cortices and right anterior insula cortex. There were no differences between patients with active and stable disease with ROI or VBM analyses. Statistics for individual clusters, including size, P values and MNI coordinates, for these and all subsequent VBM analyses are outlined in **Table 0-3** in the appendix.

The association between UWDRS-N scores, including subscores for neurological endophenotypes, and ROI volumes in stable patients are shown in **Table 5-3**. Increasing neurological severity was associated with reduced caudate, putamen, pallidum, midbrain, pons and cerebellum volumes. Subscores for speech, dystonia, parkinsonism and ataxia but not tremor or chorea were associated caudate, putamen and pallidum volumes. In addition, subscores for speech and parkinsonism were associated with midbrain, pons and cerebellum volumes and for ataxia were associated with midbrain and cerebellar volumes. The equivalent VBM analysis is shown in **Figure 5-3**. A similar pattern of reduced GM volumes in the bilateral caudate and putamen extending to orbitofrontal and anterior insula cortices was seen.

Associations between cognitive test scores and ROI volumes are shown in **Table 5-4**. MRT, FAS and TMTB scores correlated with caudate, putamen and pallidum volumes. TMTB scores also correlated with midbrain volumes. Associations between cognitive test scores and ROI volumes after

Table 5-1 Structural covariance between ROI volumes

A) Spearman's correlation coefficient

	Caudate	Putamen	Pallidum	Thalamus	Amygdala	Midbrain	Pons	Cerebellum
Caudate		**0.92**	**0.93**	**0.79**	0.17	**0.81**	**0.83**	**0.69**
Putamen	**0.92**		**0.97**	**0.81**	0.20	**0.80**	**0.80**	**0.71**
Pallidum	**0.93**	**0.97**		**0.85**	0.20	**0.83**	**0.83**	**0.71**
Thalamus	**0.79**	**0.81**	**0.85**		0.26	**0.83**	**0.81**	**0.66**
Amygdala	0.17	0.20	0.20	0.26		0.10	0.17	**0.50**
Midbrain	**0.81**	**0.80**	**0.83**	**0.83**	0.10		**0.91**	**0.58**
Pons	**0.83**	**0.80**	**0.83**	**0.81**	0.17	**0.91**		**0.71**
Cerebellum	**0.69**	**0.71**	**0.71**	**0.66**	**0.50**	**0.58**	**0.71**	

B) P values

	Caudate	Putamen	Pallidum	Thalamus	Amygdala	Midbrain	Pons	Cerebellum
Caudate		**<0.001*** **	**<0.001*** **	**<0.001*** **	0.44	**<0.001*** **	**<0.001*** **	**<0.001*** **
Putamen	**<0.001*** **		**<0.001*** **	**<0.001*** **	0.27	**<0.001*** **	**<0.001*** **	**<0.001*** **
Pallidum	**<0.001*** **	**<0.001*** **		**<0.001*** **	0.50	**<0.001*** **	**<0.001*** **	**<0.001*** **
Thalamus	**<0.001*** **	**<0.001*** **	**<0.001*** **		0.21	**<0.001*** **	**<0.001*** **	**<0.001*** **
Amygdala	0.44	0.27	0.50	0.21		0.36	0.18	**0.003** **
Midbrain	**<0.001*** **	**<0.001*** **	**<0.001*** **	**<0.001*** **	0.36		**<0.001*** **	**<0.001*** **
Pons	**<0.001*** **	**<0.001*** **	**<0.001*** **	**<0.001*** **	0.18	**<0.001*** **		**<0.001*** **
Cerebellum	**<0.001*** **	**<0.001*** **	**<0.001*** **	**<0.001*** **	**0.001** **	**<0.001*** **	**<0.001*** **	

*Correlation coefficients between ROI volumes (A) and their corresponding P values (B) are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. ROI, region of interest.*

Table 5-2 Group differences in ROI volumes

	All (n=39) mean [SD], %	**Hepatic (n=17)** mean [SD], %	**Neurological (n=22)** Mean [SD], %	**P value**	**Stable (n=34)** mean [SD], %	**Active (n=5)** Mean [SD], %	**P value**
Caudate	0.39 [0.06]	0.45 [0.05]	0.36 [0.05]	**<0.001*** **	0.41 [0.06]	0.35 [0.08]	0.62
Putamen	0.54 [0.08]	0.61 [0.07]	0.49 [0.05]	**<0.001*** **	0.54 [0.09]	0.46 [0.07]	0.86
Pallidum	0.23 [0.04]	0.27 [0.03]	0.21 [0.03]	**<0.001*** **	0.24 [0.04]	0.20 [0.04]	0.79
Thalamus	0.71 [0.09]	0.77 [0.07]	0.67 [0.09]	**<0.001*** **	0.71 [0.09]	0.61 [0.15]	0.63
Amygdala	0.23 [0.02]	0.24 [0.02]	0.24 [0.02]	0.90	0.23 [0.02]	0.23 [0.03]	0.08
Midbrain	0.37 [0.07]	0.42 [0.04]	0.34 [0.07]	**<0.001*** **	0.37 [0.06]	0.29 [0.12]	0.74
Pons	0.88 [0.16]	0.99 [0.10]	0.80 [0.15]	**<0.001*** **	0.87 [0.14]	0.75 [0.28]	0.67
Cerebellum	9.17 [1.03]	9.61 [0.75]	8.83 [1.09]	**0.007****	9.33 [0.79]	9.16 [2.04]	0.82

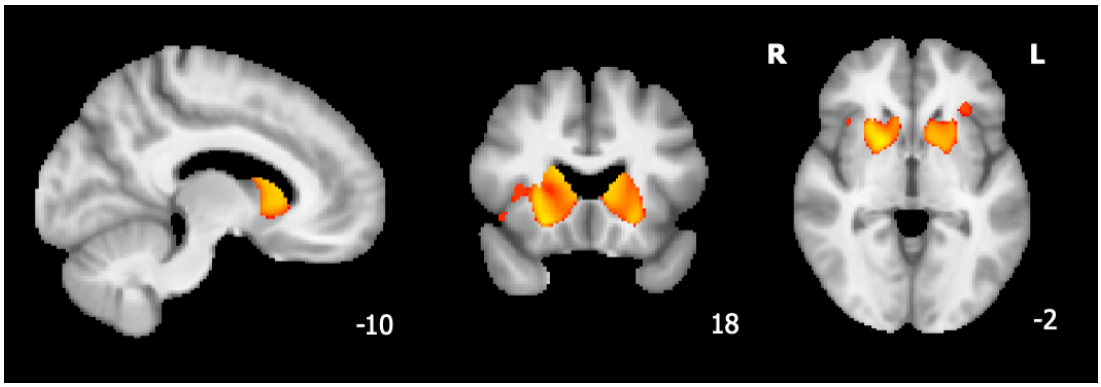
** = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. ROI, region of interest; SD, standard deviation.*

Table 5-3 Associations between ROI volumes and UWDRS-N scores

	Caudate	**Putamen**	**Pallidum**	**Thalamus**	**Amygdala**	**Midbrain**	**Pons**	**Cerebellum**
Speech	**<0.001*** **	**<0.001*** **	**<0.001*** **	0.06	0.55	**0.006****	**0.02***	0.04*
Dystonia	**0.003****	**0.005****	**0.006****	0.08	0.42	0.03*	0.04*	0.07
Parkinsonism	**<0.001*** **	**<0.001*** **	**<0.001*** **	**0.03***	0.59	**0.007****	**0.008****	**0.009****
Tremor	0.48	0.4	0.48	0.68	0.29	0.73	0.93	0.79
Ataxia	**0.02***	**0.005****	**0.01***	0.13	0.3	**0.02***	0.07	**0.03***
Chorea	0.68	0.45	0.47	0.38	0.93	0.27	0.93	0.74
UWDRS-N	**0.004****	**0.002****	**0.003****	0.05	0.32	**0.02***	**0.03***	**0.03***

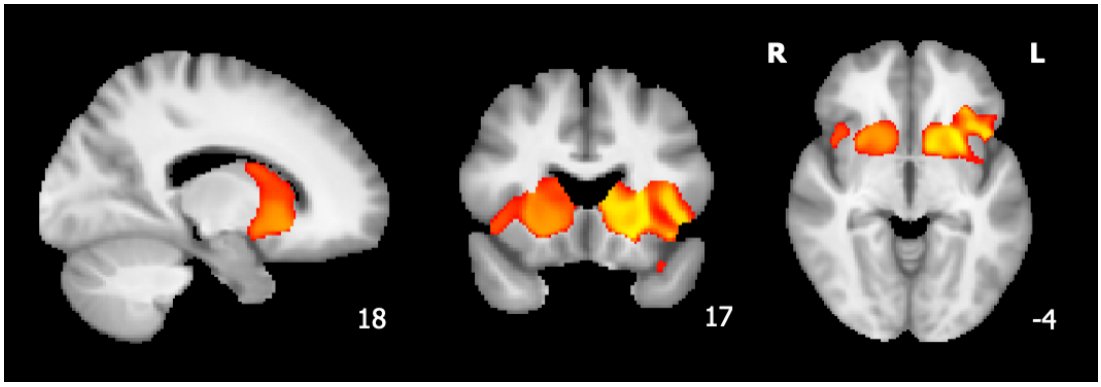
*P values for coefficients when testing associations between UWDRS-N scores and ROI volumes using linear regression are shown. Corresponding coefficients where P < 0.05 were negative. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. ROI, region of interest; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.*

Figure 5-2 VBM for comparison between hepatic and neurological presentations



Clusters where GM volume is lower in patients with neurological presentation than patients with hepatic presentations for FWE-corrected P values < 0.05 are shown in red/yellow. One slice in sagittal (x), coronal (y) and axial (z) planes was selected for visualisation purposes. MNI coordinates are provided. FWE, family-wise error; GM, grey matter; MNI, Montreal neurological institute; VBM, voxel-based morphometry.

Figure 5-3 VBM for associations with UWDRS-N scores



Clusters where GM volume is negatively correlated with UWDRS-N scores in stable patients for FWE-corrected P values < 0.05 are shown in red/yellow. One slice in sagittal (x), coronal (y) and axial (z) planes was selected for visualisation purposes. MNI coordinates are provided. MNI coordinates are provided. FWE, family-wise error; GM, grey matter; MNI, Montreal neurological institute; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; VBM, voxel-based morphometry.

Table 5-4 Associations between ROI volumes and cognitive test scores

Domain	Test	Caudate	Putamen	Pallidum	Thalamus	Amygdala	Midbrain	Pons	Cerebellum
Abstract reasoning	MRT	**0.006****	**<0.001****	**0.001****	0.03*	0.30	0.09	0.10	0.004**
Language	NART	0.06	0.04*	0.03*	0.25	0.81	0.19	0.39	0.09
	GNT	0.03*	0.04*	0.01*	0.03*	0.96	0.16	0.20	0.05
Memory	RMTF	0.03*	0.02*	0.05*	0.31	0.44	0.08	0.10	0.07
	RMTW	0.16	0.34	0.37	0.64	0.43	0.51	0.39	0.59
	CPAL	0.11	0.30	0.50	0.92	0.52	1.00	0.64	0.74
Processing speed	TMTA	0.04*	0.08	0.10	0.29	0.92	0.29	0.25	0.39
Executive function	DSB	0.06	0.02*	0.02*	0.25	0.69	0.56	0.45	0.28
	FAS	**0.009****	**<0.001*****	**0.006****	0.06	0.32	0.01*	0.01*	0.02*
	Animals	0.12	0.21	0.18	0.78	0.68	0.98	0.48	0.03*
	DKEFSI	0.16	0.07	0.13	0.41	0.38	0.37	0.46	0.16
	TMTB	**0.003****	**0.003****	**0.001***	0.004**	0.81	**0.001****	0.004**	0.35
	DSym	0.02*	0.02	0.02*	0.07	0.51	0.04*	0.06	0.05
Calculation	GDA	0.09	0.10	0.10	0.81	0.97	0.81	0.71	0.24
Visuoperceptual	VOSPFL	0.36	0.72	0.79	0.33	0.33	0.45	0.54	0.26
Visuospatial	VOSPNL	0.39	0.35	0.52	0.66	0.80	0.86	0.92	0.21
Social	Ekman	0.56	0.12	0.32	0.49	0.25	0.09	0.30	0.21

*P values for coefficients when testing associations between cognitive test scores and ROI volumes using linear regression are shown. Corresponding coefficients where P < 0.05 were positive. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; ROI, region of interest; TMTA, trail making test part A; TMTB, trail making test part B; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

including UWDRS as a covariate are shown in **Table 5-5**. MRT and FAS scores correlated with putamen volumes. VBM analyses identified associations between several cognitive test scores and distinct patterns of reduced GM volume. These findings are summarised in **Table 5-6** and VBM results for individual cognitive tests are shown in **Figure 5-4**. MRT scores were associated with reduced GM volumes in the right putamen, insula and orbitofrontal cortices and scores for semantic fluency (Animals) were associated reduced GM volumes in the left cerebellum. RMTF scores were associated with decreased GM volumes in diffuse, predominantly anterior cortical regions including the bilateral cingulate, paracingulate and insula cortices, middle frontal gyri and supplementary motor area and right superior and middle temporal gyri and subcallosal and opercular cortices. Subcortical clusters including the caudate, putamen, dorsal midbrain, and right cerebellum were also identified. Similarly, Ekman scores were also associated with decreased GM volumes in anterior cortical regions. These included the bilateral cingulate, paracingulate, insula and orbitofrontal cortices and pre-central and post-central gyri, right central opercular cortex, frontal pole and middle frontal gyrus and left superior frontal gyrus and temporal fusiform cortex. Subcortical clusters in the bilateral hippocampus and cerebellum and right putamen were also identified. Clusters in the bilateral cingulate and paracingulate cortices and right central opercular cortex persisted when including UWDRS-N score as a covariate. TMTB scores were associated with decreased volume in the left supplementary motor area and occipital fusiform gyrus and right intracalcarine cortex and these clusters also persisted when UWDRS-N scores were included as a covariate.

Associations between psychiatric scores and ROI volumes are shown in **Table 5-7**. There were no associations between UWDRS-P, ModNPI, PHQ9 or GAD7 scores and any ROI volume. Associations between scores for individual ModNPI items are shown in **Table 5-8**. There were no associations between ModNPI items and ROI volumes. Correlations between trusting behaviour and pallidum volumes and between compulsive behaviour and amygdala volumes did not persist with correction for multiple comparisons. UWDRS-P scores were associated with decreased GM volume in a small cluster in the right occipital fusiform gyrus on VBM analysis. Scores for the

Table 5-5 Associations between ROI volumes and cognitive test scores with UWDRS-N as covariate

Domain	Test	Caudate	Putamen	Pallidum	Thalamus	Amygdala	Midbrain	Pons	Cerebellum
Abstract reasoning	MRT	0.08	**0.005****	0.03*	0.15	0.51	0.43	0.44	0.03*
Language	NART	0.16	0.17	0.12	0.52	0.70	0.53	0.76	0.28
	GNT	0.05*	0.07	0.02*	0.06	0.90	0.28	0.32	0.10
Memory	RMTF	0.27	0.24	0.42	0.90	0.73	0.47	0.47	0.35
	RMTW	0.75	0.28	0.29	0.39	0.80	0.35	0.57	0.31
	CPAL	0.06	0.15	0.31	0.94	0.51	0.77	0.47	0.57
Processing speed	TMTA	0.14	0.37	0.40	0.64	0.97	0.80	0.59	0.92
Executive function	DSB	0.07	0.02*	0.02*	0.30	0.67	0.64	0.51	0.34
	FAS	0.02*	**0.004****	0.02*	0.12	0.24	0.03*	0.03*	0.07
	Animals	0.21	0.43	0.35	0.97	0.77	0.62	0.75	0.06
	DKEFSI	0.25	0.13	0.24	0.58	0.41	0.59	0.66	0.27
	TMTB	0.01*	0.01*	0.007**	0.01*	0.90	0.005**	0.01*	0.74
	DSym	0.07	0.09	0.09	0.21	0.59	0.16	0.17	0.20
Calculation	GDA	0.19	0.25	0.24	0.89	0.89	0.81	0.98	0.50
Visuoperceptual	VOSPFL	0.41	0.51	0.60	0.38	0.37	0.48	0.59	0.30
Visuospatial	VOSPNL	0.69	0.65	0.88	0.40	0.65	0.53	0.61	0.36
Social	Ekman	0.69	0.49	0.95	0.95	0.38	0.33	0.77	0.56

*P values for coefficients when testing associations between cognitive test scores and ROI volumes using linear regression with UWDRS-N as a covariate are shown. Corresponding coefficients where P < 0.05 were positive. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; ROI, region of interest; TMTA, trail making test part A; TMTB, trail making test part B; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest.*

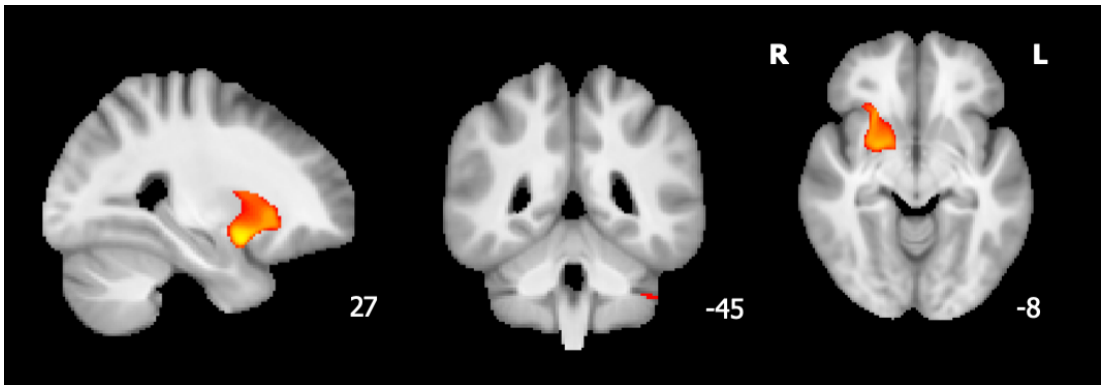
Table 5-6 Summary of VBM results for associations with cognitive test scores

Domain	Test	Anatomical regions
Abstract reasoning	MRT	Right putamen, insula and orbitofrontal cortices and temporal pole and left cerebellum.
Language	NART	Nil
	GNT	Nil
Memory	RMTF	Bilateral caudate, putamen, dorsal midbrain, insula, cingulate and paracingulate cortices and right cerebellum. Bilateral supplementary motor areas and middle frontal gyri, right subcallosal and opercular cortices and right superior and middle temporal gyri.
	RMTW	Nil
	CPAL	Nil
Processing speed	TMTA	Nil
Executive function	DSB	Nil
	FAS	Nil
	Animals	Left cerebellum
	DKEFSI	Nil
	TMTB	**Left supplementary motor area** and occipital fusiform gyrus and **right intracalcarine cortex.**
	DSym	Nil
Calculation	GDA	Nil
Visuoperceptual	VOSPFL	Nil
Visuospatial	VOSPNL	Nil
Social	Ekman	**Bilateral cingulate and paracingulate cortices**, insula, hippocampi, cerebellum and vermis and right putamen. Bilateral orbitofrontal cortices and pre-central and post-central gyri, **right central opercular cortex**, frontal pole and middle frontal gyrus and left superior frontal gyrus and temporal fusiform cortex.

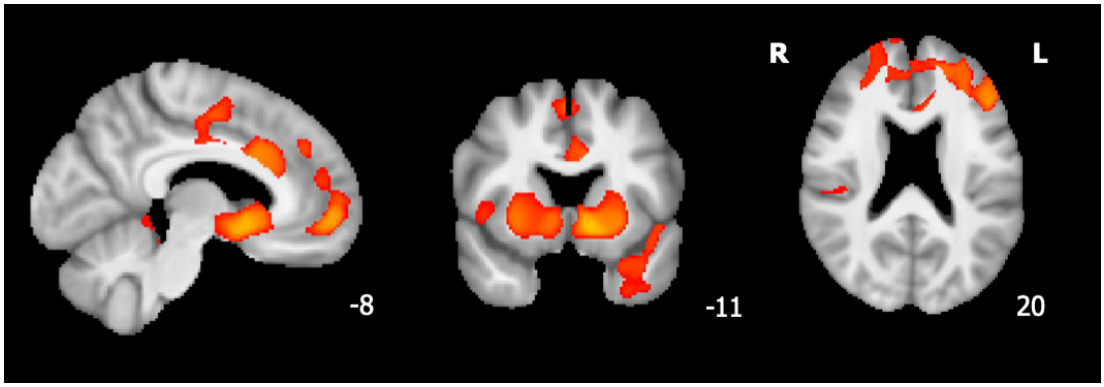
The anatomical regions of clusters where GM volume was positively correlated with cognitive test scores in stable patients for FWE-corrected P values are described for each cognitive test. Anatomical regions where clusters persist when UWDRS-N scores are included as a covariate are highlighted in bold.

Figure 5-4 VBM for associations with cognitive test scores

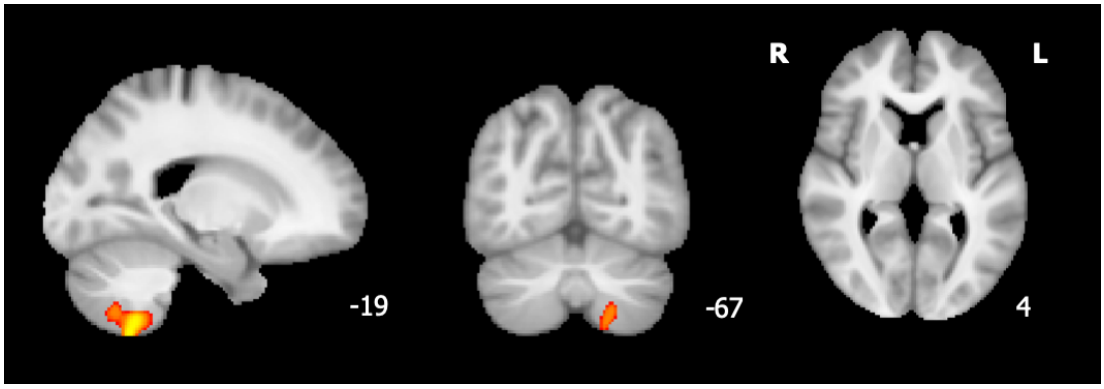
A) MRT



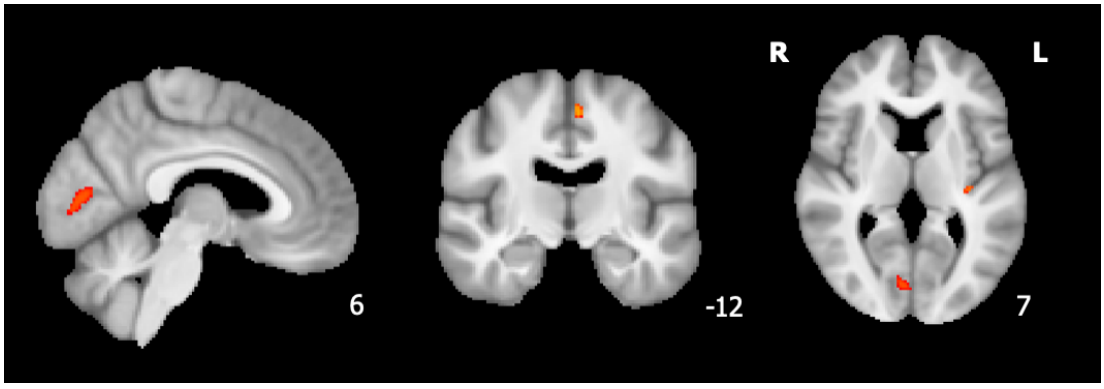
B) RMTF



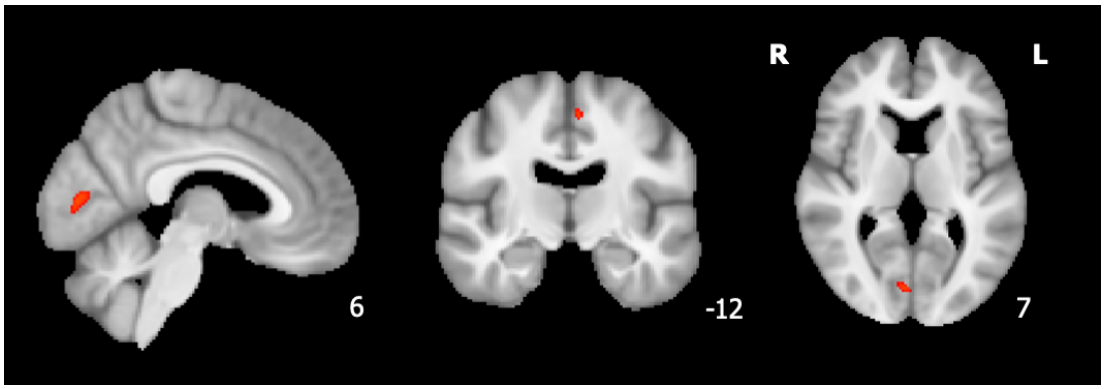
C) Animals scores



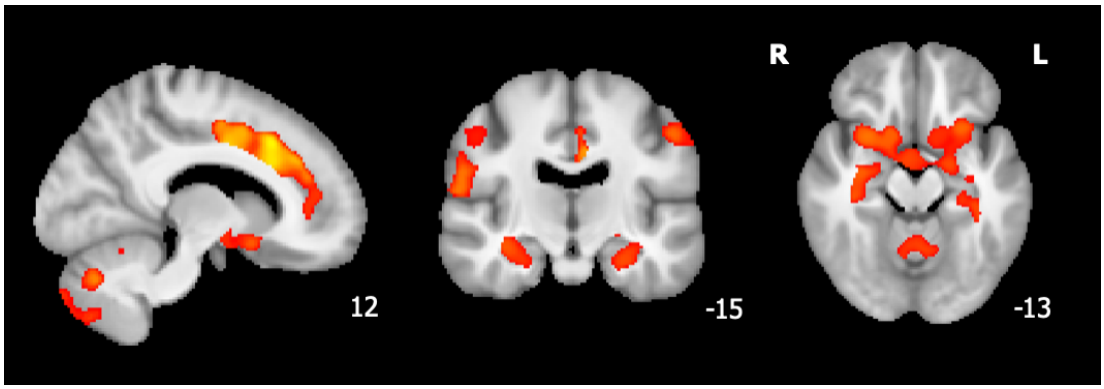
D) TMTB scores



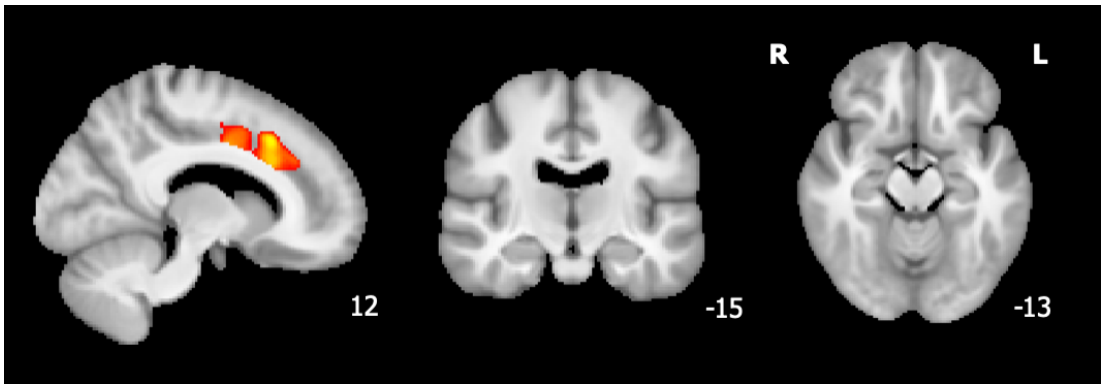
E) TMTB with UWDRS-N as covariate



F) Ekman scores



G) Ekman with UWDRS-N as covariate



Clusters where GM volume positively correlates with cognitive test scores in stable patients for FWE-corrected P values < 0.05 are shown in red/yellow for various cognitive tests. Where clusters persist after including UWDRS-N scores as a covariate these are shown in separate tissue maps. One slice in sagittal (x), coronal (y) and axial (z) planes was selected for visualisation purposes. MNI coordinates are provided. Animals, semantic fluency test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; FWE, family-wise error; GM, grey matter; MNI, Montreal neurological institute; MRT, matrix reasoning test; RMTF, recognition memory test for faces; TMTB, trail making test part B; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; VBM, voxel-based morphometry.

Table 5-7 Association between ROI volumes and psychiatric scores

	Caudate	Putamen	Pallidum	Thalamus	Amygdala	Midbrain	Pons	Cerebellum
UWDRS-P	0.74	0.76	0.79	0.91	0.21	0.95	0.74	0.75
ModNPI	0.35	0.41	0.29	0.58	0.43	0.37	0.47	0.28
PHQ9	0.96	0.93	0.89	0.54	0.86	0.75	0.38	0.69
GAD7	0.17	0.17	0.25	0.96	0.8	0.57	0.76	0.0

*P values for coefficients when testing associations between psychiatric rating scales and ROI volumes using linear regression are shown. Corresponding coefficients where P < 0.05 were positive. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. GAD7, generalised anxiety disorder assessment 7; ModNPI, modified neuropsychiatric inventory; PHQ9, patient health questionnaire 9; ROI, region of interest; UWDRS-P, unified Wilson's disease rating scale psychiatric subscore.*

Table 5-8 The association between ROI volumes and ModNPI items in stable patients

	Caudate	Putamen	Pallidum	Thalamus	Amygdala	Midbrain	Pons	Cerebellum
Delusions	0.91	0.50	0.55	0.19	0.53	0.21	0.67	0.74
Hallucinations	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Agitation	0.22	0.28	0.24	0.19	0.86	0.08	0.19	0.77
Depression	0.24	0.25	0.23	0.64	0.48	0.63	0.69	0.13
Anxiety	0.57	0.52	0.40	0.79	0.93	0.93	0.72	0.31
Elation	0.28	0.12	0.09	0.01*	0.86	0.02	0.08	0.69
Irritability	0.33	0.83	0.66	0.86	0.58	0.49	0.33	0.35
Motor	0.55	0.79	0.61	0.87	0.40	0.96	0.80	0.96
Sleep	0.79	0.53	0.79	0.51	0.43	0.73	0.82	0.80
Hypersexuality	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Hyperreligiosity	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Humour	0.05	0.09	0.09	0.17	0.84	0.27	0.39	0.12
Apathy	0.67	0.24	0.32	0.32	0.49	0.16	0.76	0.28
Disinhibition	0.49	0.36	0.23	0.49	0.16	0.44	0.49	0.33
Eating habits	0.69	0.39	0.20	0.21	0.25	0.26	0.11	0.13
Empathy	0.87	0.76	0.81	0.42	0.02	0.70	0.95	0.33
Compulsive	0.86	0.65	0.97	0.80	0.005**	0.74	0.77	0.49
Social/emotion	0.76	0.92	0.50	0.52	0.86	0.96	0.86	0.68
Trusting	0.03	0.01*	0.006**	0.01*	0.63	0.10	0.09	0.10
Pain	0.79	0.70	0.60	0.26	0.74	0.55	0.54	0.63
ModNPI	0.35	0.41	0.29	0.58	0.43	0.37	0.47	0.28

*P values for coefficients when testing associations between ModNPI items and ROI volumes using linear regression are shown. Corresponding coefficients where P < 0.05 were negative. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. ModNPI, modified neuropsychiatric inventory; ROI, region of interest.*

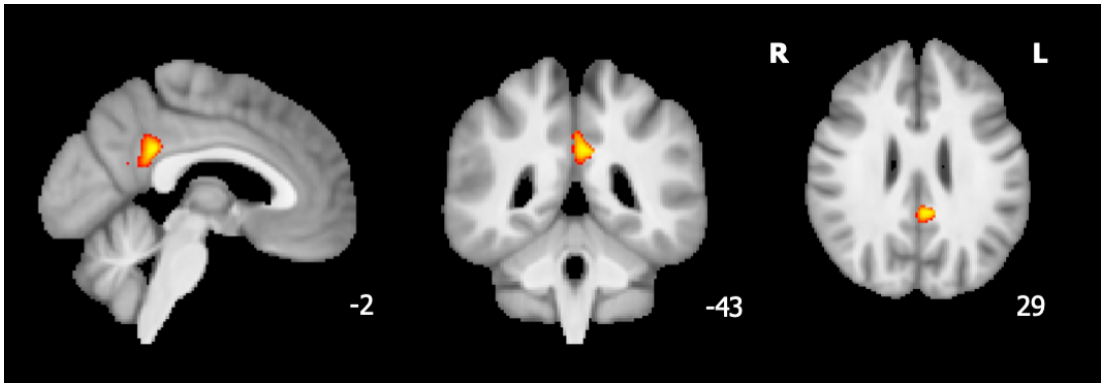
anxiety item on the ModNPI were associated with a cluster of decreased GM volume in the posterior cingulate cortex, as shown in **Figure 5-5**. Scores for the social and emotional cues item on the ModNPI were associated with diffuse regions of cortical volume loss involving the bilateral posterior cingulate cortices, hippocampi and parahippocampal gyri, amygdalae, superior parietal lobules, lingual gyri, occipital poles and cerebellar vermis, the left lateral occipital and insula cortices and the right intracalcarine cortex and temporal pole, also shown in **Figure 5-5**. There were no other associations between psychiatric rating scales and GM volume on VBM analyses.

Associations between copper indices and wet biomarkers and ROI volumes are shown in **Table 5-9**. NCC concentrations were negatively correlated with GM volumes in scattered cortical areas including the left precentral gyrus, right lateral occipital cortex and bilateral precuneus, as shown in **Figure 5-6**. There were associations between other copper indices and ROI volumes on volumetric analyses or GM volumes on VBM analyses. Tau was negatively correlated with caudate, putamen and pallidum volumes on ROI analyses and with GM volumes in the caudate and putamen on VBM analyses, as shown in **Figure 5-7**. These associations did not persist when including UWDRS-N scores as a covariate. There were no associations between NfL, GFAP or UCH-L1 and ROI volumes or GM volumes on VBM analyses.

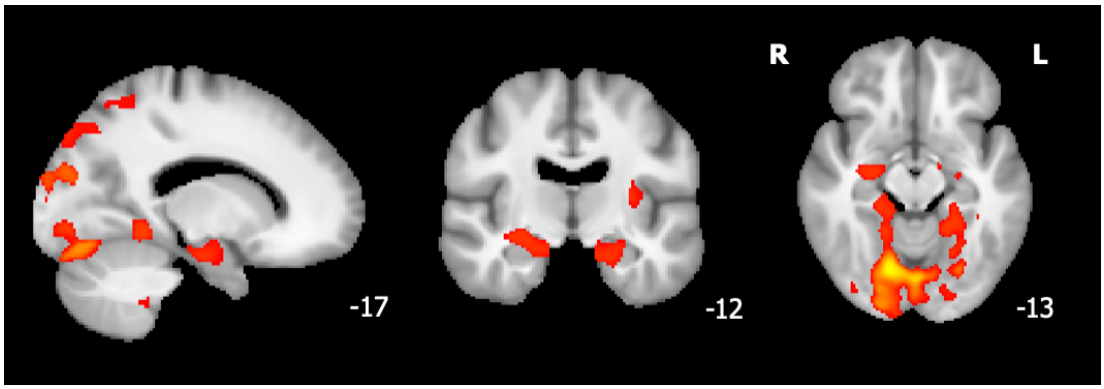
There were no associations between disease duration and ROI volumes or GM volumes on VBM.

Figure 5-5 VBM for associations with ModNPI items

A) Anxiety



B) Social and emotional cues



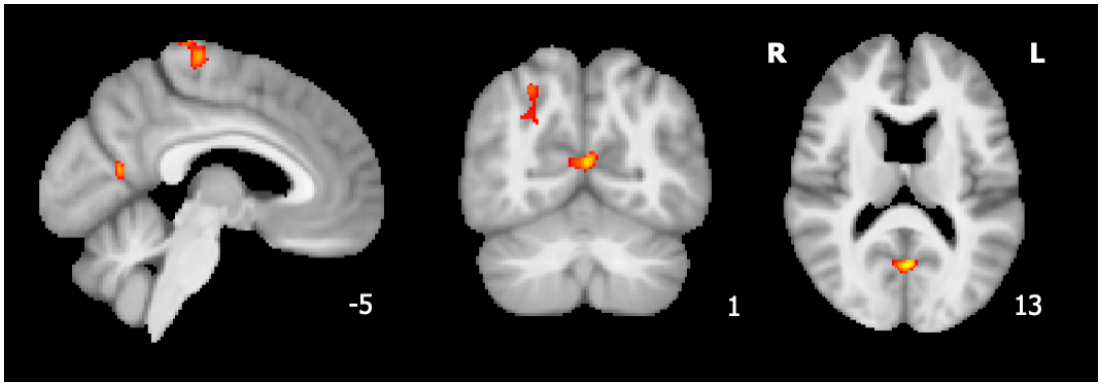
Clusters where GM volume negatively correlates with ModNPI item scores in stable patients for FWE-corrected P values < 0.05 are shown in red/yellow for the anxiety item (A) and social and emotional cues item (B). One slice in sagittal (x), coronal (y) and axial (z) planes was selected for visualisation purposes. MNI coordinates are provided. ModNPI, modified neuropsychiatric inventory; MNI, Montreal neurological institute; VBM, voxel-based morphometry.

Table 5-9 Associations between ROI volumes and wet biomarkers

	Caudate	Putamen	Pallidum	Thalamus	Amygdala	Midbrain	Pons	Cerebellum
NCC	0.38	0.82	0.39	0.41	0.35	0.55	0.79	0.94
EXC	0.30	0.23	0.37	0.55	0.35	0.66	0.62	0.67
UCu	0.88	0.72	0.76	0.83	0.43	0.60	0.49	0.09
NfL	0.37	0.24	0.28	0.29	0.83	0.26	0.48	0.88
Tau	**0.01***	**0.02***	**0.02***	0.08	0.64	0.06	0.04*	0.53
GFAP	0.15	0.65	0.47	0.56	0.31	0.37	0.16	0.01*
UCH-L1	0.33	0.13	0.11	0.38	0.91	0.09	0.14	0.68

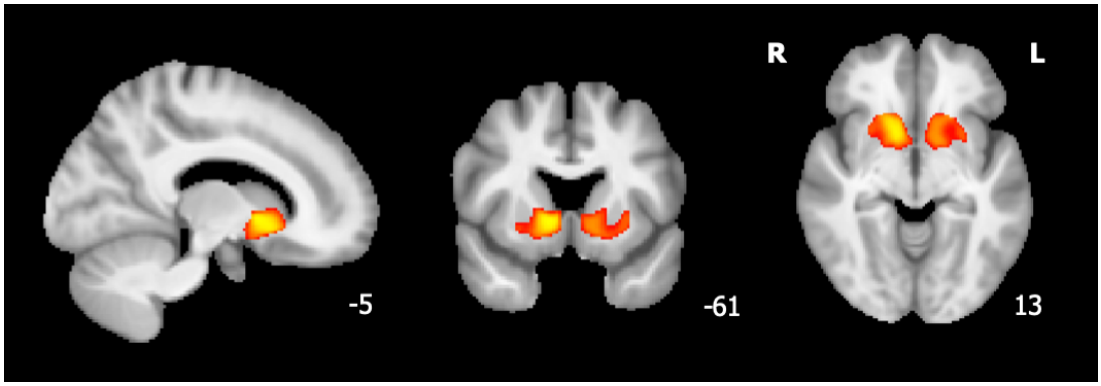
*P values for coefficients when testing associations between biomarkers, including N4PA results and copper indices, and ROI volumes using linear regression are shown. Corresponding coefficients where P < 0.05 were negative. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. EXC, exchangeable copper; GFAP, glial fibrillary acidic protein; NCC, non-caeruloplasmin-bound copper; NfL, neurofilament light; N4PA, neurology 4-plex assay; ROI, region of interest; UCH-L1, ubiquitin carboxy-terminal hydrolase L1; UCu, 24-hour urinary copper output.*

Figure 5-6 VBM for associations with NCC



Clusters where GM volume negatively correlates with NCC in stable patients for FWE-corrected P values < 0.05 are shown in red/yellow. One slice in sagittal (x), coronal (y) and axial (z) planes was selected for visualisation purposes. MNI coordinates are provided. MNI, Montreal neurological institute; NCC, non-caeruloplasmin-bound copper; VBM, voxel-based morphometry.

Figure 5-7 VBM for associations with Tau



Clusters where GM volume negatively correlates with tau in stable patients for FWE-corrected P values < 0.05 are shown in red/yellow. One slice in sagittal (x), coronal (y) and axial (z) planes was selected for visualisation purposes. MNI coordinates are provided. MNI, Montreal neurological institute; VBM, voxel-based morphometry.

FLAIR imaging

The log_e-transformed total volume of WMHs did not differ between patients with hepatic and neurological presentations (916 vs 1384 mm^3, P = 0.12) and was higher in patients with active disease than patients with stable disease (6126 vs 953 mm^3, P < 0.001). There were no associations between the log_e-transformed total volume of WMHs and scores for UWDRS-N (ß = 0.00, P = 0.63), any cognitive test, UWDRS-P (ß = -0.06, P = 0.28), ModNPI (ß = 0.01, P = 0.09) or GAD7 (ß = 0.03, P = 0.30) in stable patients. It was, however, positively correlated with PHQ scores only (ß = 0.05, P = 0.04). There were no associations between the log_e-transformed total volume of WMHs and any copper index or wet biomarkers in stable patients. This included NCC (ß = -0.06, P = 0.68), EXC (ß = -0.41, P = 0.48), UCu (ß = -0.01, P = 0.89), NfL (ß = 0.03, P = 0.57), tau (ß = -0.04, P = 0.85), GFAP (ß = 0.00, P = 0.05) or UCH-L1 (ß = 0.00, P = 0.874).

Group differences in the log_e-transformed volume of WMHs across six brain regions are shown in **Table 5-10**. Patients with neurological presentations had higher infratentorial lesion volumes than patients with hepatic presentations. This finding did not persist when excluding patients with active disease on post hoc testing. Patients with active disease had higher lesions volumes in the basal ganglia, frontal, temporal, parietal and infratentorial regions.

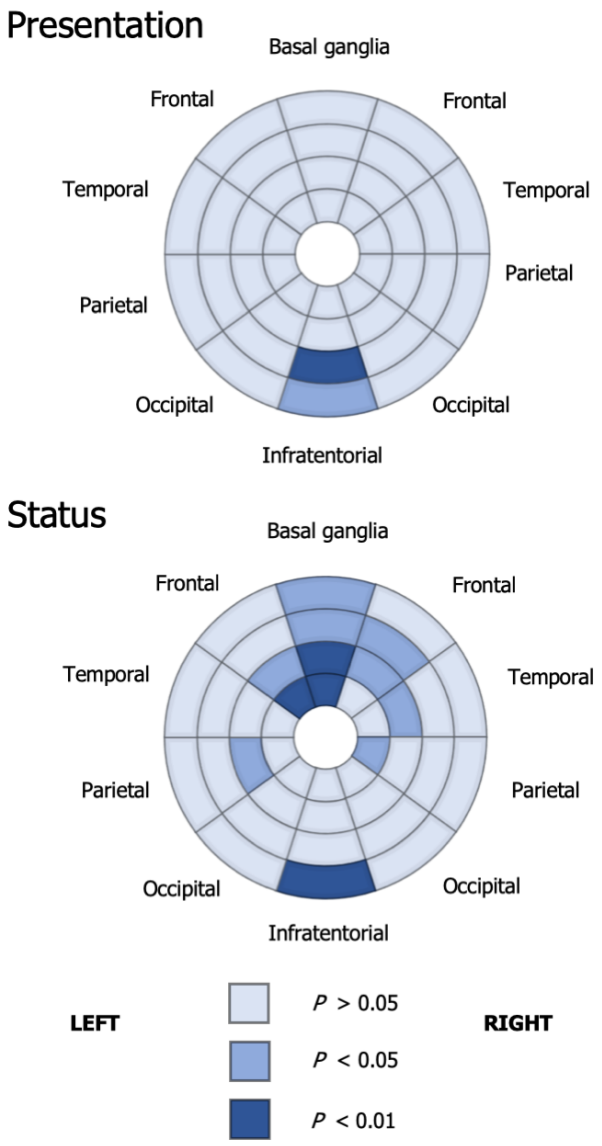
Bullseyes plots demonstrating group differences in the log_e-transformed volume of WMHs across 40 brain regions are shown in **Figure 5-8**. Patients with neurological presentations had higher infratentorial lesion volumes than patients with hepatic presentations in regions that correspond to the cerebellum and, as above, this finding did not persist when excluding patients with active disease. Patients with active disease had higher lesions volumes throughout the basal ganglia and in deeper, predominantly periventricular regions of the frontal, temporal and parietal WM. There were no associations between lesion volumes within individual regions and any of the clinical rating scales, copper indices or wet biomarkers. On testing associations with individual ModNPI items (n=20), I identified a weak association between trusting behaviour and the volume of WMHs in scattered

Table 5-10 Group differences in the volume of WMHs

	Hepatic (n=17) Median [IQR], mm^3	**Neurological (n=22)** Median [IQR], mm^3	**P value**	**Stable (n=34)** Median [IQR], mm^3	**Active (n=5)** Median [IQR], mm^3	**P value**
Basal ganglia	39 [32-81]	44 [27-112]	0.40	38 [25-69]	261 [239-1810]	**<0.001*** **
Frontal lobe	453 [169-547]	368 [119-934]	0.53	316 [106-540]	1118 [800-2194]	**0.02***
Temporal lobe	66 [33-136]	154 [75-447]	0.19	87 [41-202]	534 [467-1195]	**0.01***
Parietal lobe	41 [9-76]	147 [12-340]	0.29	41 [8-223]	298 [296-568]	**0.02***
Occipital lobe	218 [147-524]	305 [223-665]	0.30	248 [193-481]	770 [324-1302]	0.26
Infratentorial	11 [0-27]	75 [56-198]	**<0.001*** **	52 [11-81]	987 [365-1775]	**0.03***
Total	916 [525-1281]	1384 [853-3205]	0.12	953 [622-1517]	6126 [2683-6434]	**0.007****

*Group differences in the log_e-transformed volume of WMHs in six brain regions and in total are shown. * = P value <0.05; ** = P value <0.01; *** = P value <0.001. P values less than 0.05 after FDR correction are highlighted in bold. IQR, interquartile range; WMHs, white matter hyperintensities.*

Figure 5-8 Bullseye plots for group differences in the volume of WMHs

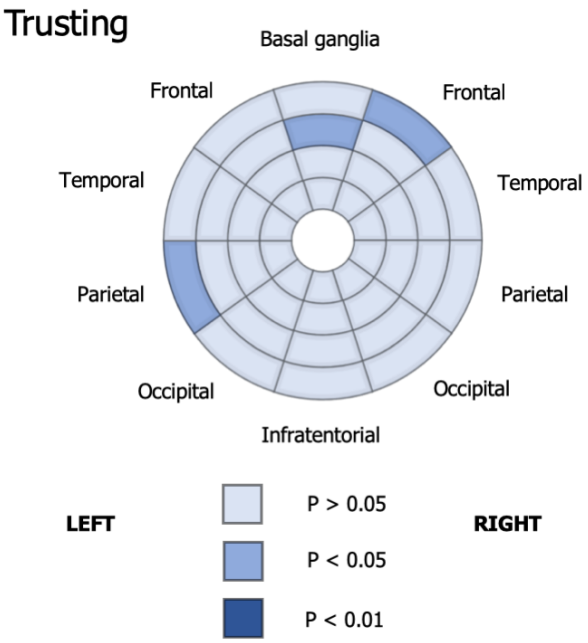


Bullseye plots consisting of 40 brain regions are shown. There are 10 sectors each divided into four layers corresponding to more central or peripheral locations. Regions where the $\log_e$-transformed volume of WMHs is higher in patients with neurological than hepatic presentation for FDR-corrected P values < 0.05 are highlighted in dark blue in the upper plot. Regions where the $\log_e$-transformed volume of WMHs is higher in patients with active than stable disease for FDR-corrected P values < 0.05 are highlighted in dark blue in the lower plot.

regions within the basal ganglia and right frontal and left parietal WM, as shown in **Figure 5-9**.

There were no associations between disease duration and the $\log_e$-transformed volume of WMHs in any brain region.

Figure 5-9 Bullseye plot for associations between trusting behaviour and the volume of WMHs



A bullseye plot consisting of 40 brain regions are shown. There are 10 sectors each divided into four layers corresponding to more central or peripheral locations. Regions where the $\log_e$-transformed volume of WMHs positively correlates with the ModNPI item score for trusting behaviour for FDR-corrected P values < 0.05 are highlighted in dark blue.

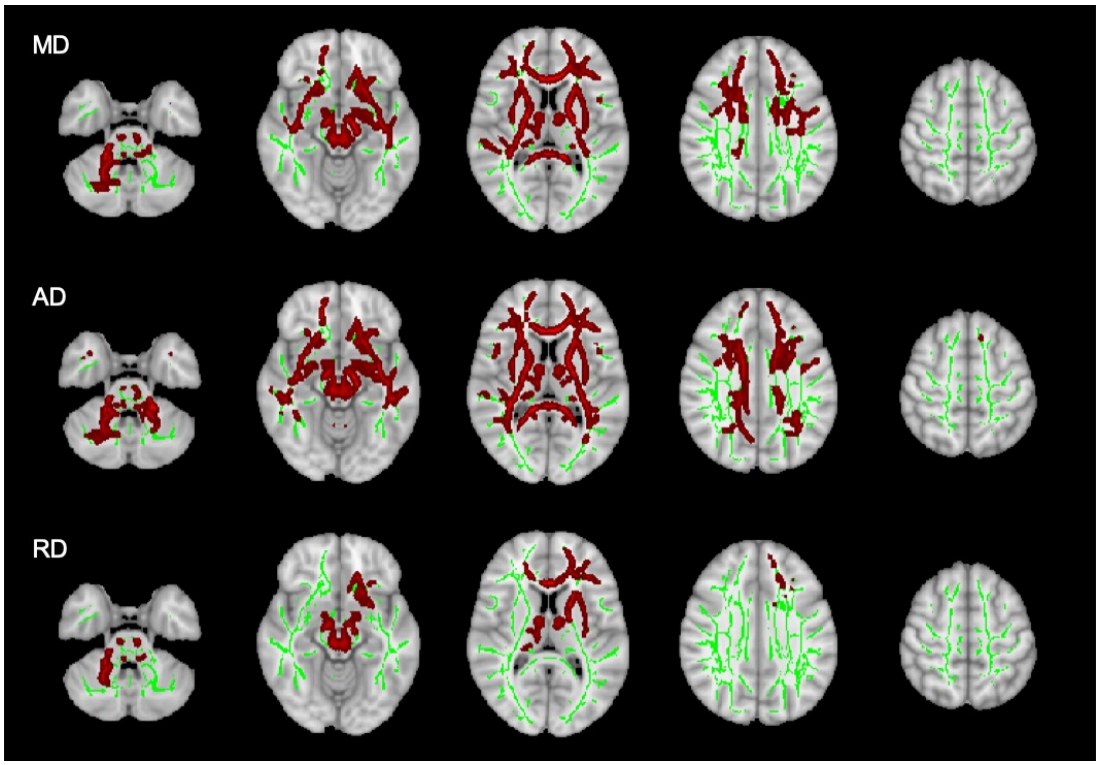
Diffusion-weighted imaging

There were no differences in FA, MD, AD or RD when comparing the WM tracts of patients with neurological and hepatic presentations using TBSS. Patients with active disease had increased MD and AD throughout the WM, with the exception of the inferior longitudinal fasciculi, compared to patients with stable disease as shown in **Figure 5-10**. RD was increased in a more restricted distribution of WM tracts including the left internal and external capsules, anterior thalamic radiation and forceps minor and the cerebral peduncle and corpus callosum bilaterally. There were no differences in FA.

Increasing UWDRS-N scores in stable patients were associated with reduced AD in left internal and external capsules, anterior thalamic radiation, uncinate fasciculus and cerebral peduncle and, bilaterally, in the corpus callosum and forceps minor as shown in **Figure 5-11**. There were no associations between UWDRS-N scores and FA, MD or RD.

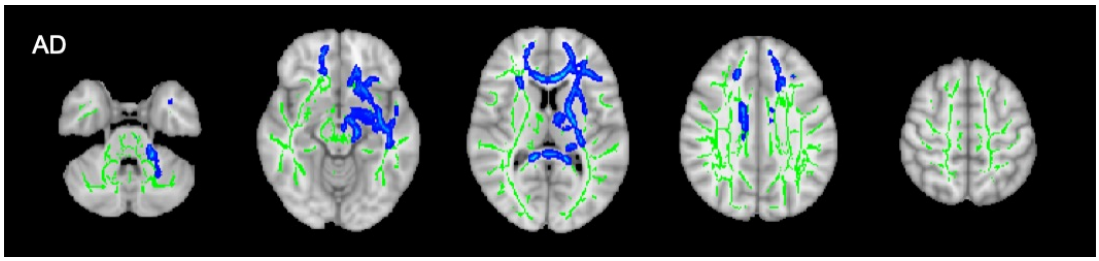
Scores for several cognitive tests were associated with DTI parameters. These findings are summarised in **Table 5-11** and associations between individual cognitive tests and DTI parameters are shown in **Figure 5-12**. Associations between individual cognitive tests and DTI parameters that persisted when including UWDRS-N as a covariate are shown in **Figure 0-1** in the appendix. MRT, RMTF, RMTW, CPAL and GDA scores were associated with widespread DTI abnormalities throughout supratentorial WM tracts: Poorer performance on the MRT and RMTW was associated with widespread decreases in AD whereas poorer performance on the RMTF and CPAL was associated with widespread increases in RD, increases in MD and decreases in FA. Poorer performance on the TMTA, TMTB and DKEFSI was associated with increases in RD in more localised WM tracts within the right corona radiata. Deficits in phonemic fluency (FAS) were associated with decreases in AD localised to WM in the cerebral peduncles, brainstem, middle cerebellar peduncles and cerebellum and deficits in semantic fluency (Animals) were associated with localised decreases in FA in WM tracts within the left corona radiata. Unlike other cognitive tests, poorer performance on the GDA was associated with widespread *decreases* in RD in addition to decreases in AD

Figure 5-10 TBSS for comparison between active and stable disease



Tracts where DTI parameters are higher in patients with active disease are shown in red. The WM skeleton is shown in green. Axial slices at z = -34, -12, 10, 32 and 54 are shown. AD, axial diffusivity; MD, mean diffusivity; RD, radial diffusivity; TBSS, tract-based spatial statistics; WM, white matter.

Figure 5-11 TBSS for associations with UWDRS-N scores



Tracts where AD decreases as UWDRS-N increases are shown in blue. The WM skeleton is shown in green. Axial slices at z = -34, -12, 10, 32 and 54 are shown. AD, axial diffusivity; TBSS, tract-based spatial statistics; WM, white matter.

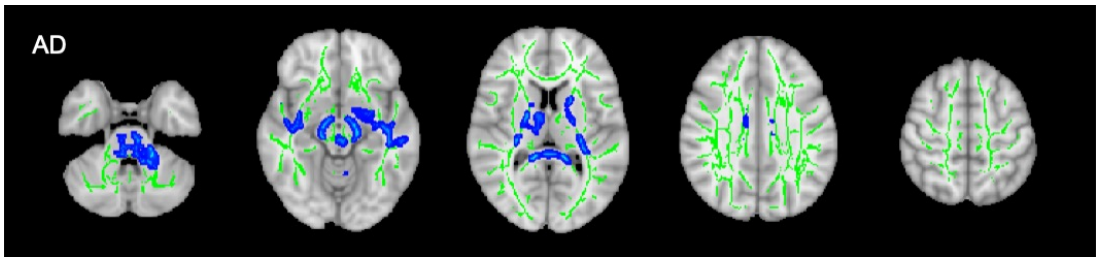
Table 5-11 Summary of TBSS for cognitive tests

Domain	Test	FA	MD	AD	RD
Abstract reasoning	MRT	-	-	Widespread	-
Language	NART	-	-	-	-
	GNT	-	-	-	-
Memory	RMTF	Widespread*	Widespread*	-	Widespread *
	RMTW	-	-	Widespread	-
	CPAL	Widespread *	Widespread *	-	Widespread *
Processing speed	TMTA	-	Localised	-	Localised
Executive function	DSB	-	-	-	-
	FAS	-	-	Localised*	-
	Animals	Localised	-	-	-
	DKEFSI	-	-	-	Localised
	TMTB	-	-	-	Localised
	DSym	-	-	-	-
Calculation	GDA	Widespread *	Widespread *	Widespread *	Widespread *
Visuoperceptual	VOSPFL	-	-	-	-
Visuospatial	VOSPNL	-	-	-	-
Social	Ekman	-	-	-	-

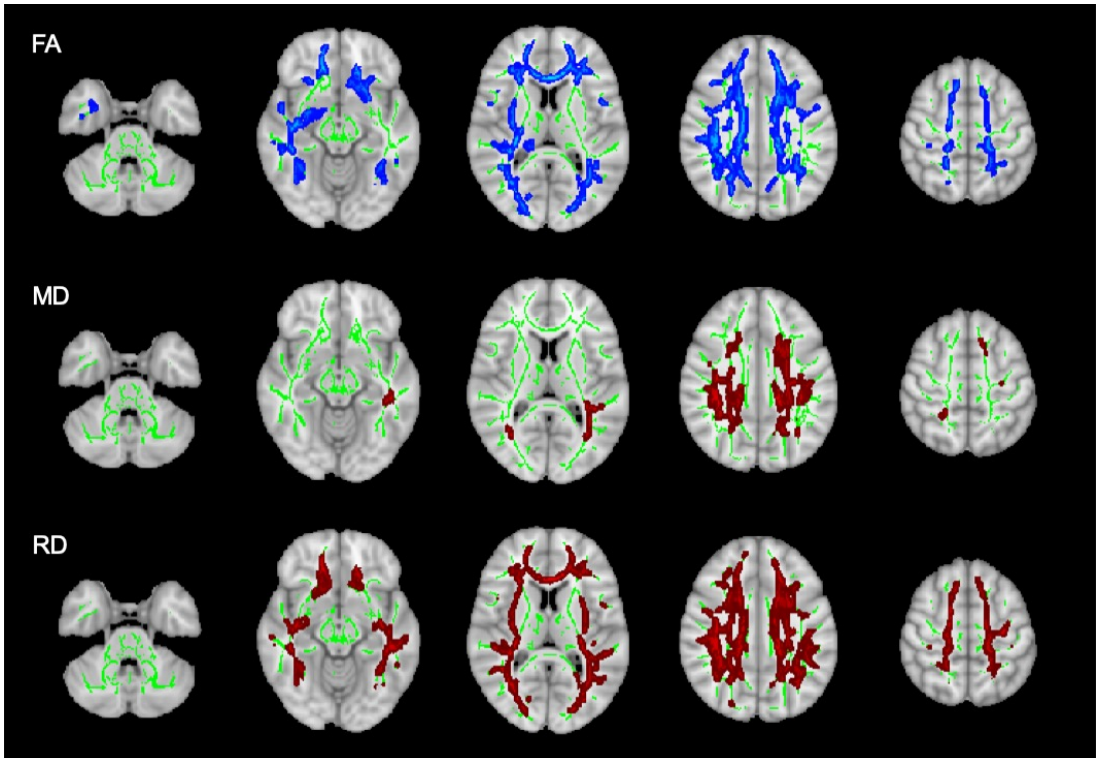
Distributions of abnormalities (widespread, localised or none) in WM tracts are shown. Associations where decreasing cognitive performance is associated with decreasing DTI parameters are shown in blue. Associations where decreasing cognitive performance is associated with increasing DTI parameters are shown in red. Associations which persist after including UWDRS-N scores as a covariate are marked with asterisks. Animals, semantic fluency test; CPAL, Camden paired associate learning test; DKEFSI, Delis-Kaplan execution function system interference subtest; DSB, digit span backwards; DSym, digit symbol test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; GNT, graded naming test; MRT, matrix reasoning test; NART, national adult reasoning test; RMTF, recognition memory test for faces; RMTW, recognition memory test for words; ROI, region of interest; TBSS, tract-based spatial statistics; TMTA, trail making test part A; TMTB, trail making test part B; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; VOSPFL, visual object and space perception battery fragmented letter subtest; VOSPNL, visual object and space perception battery number location subtest; WM, white matter.

Figure 5-12 TBSS for associations with cognitive test scores

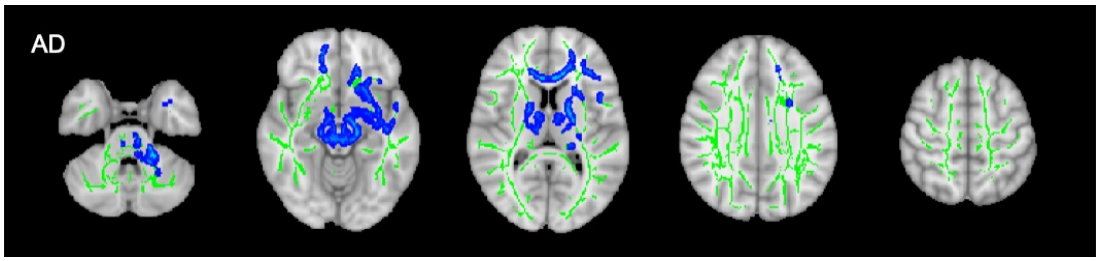
A) MRT



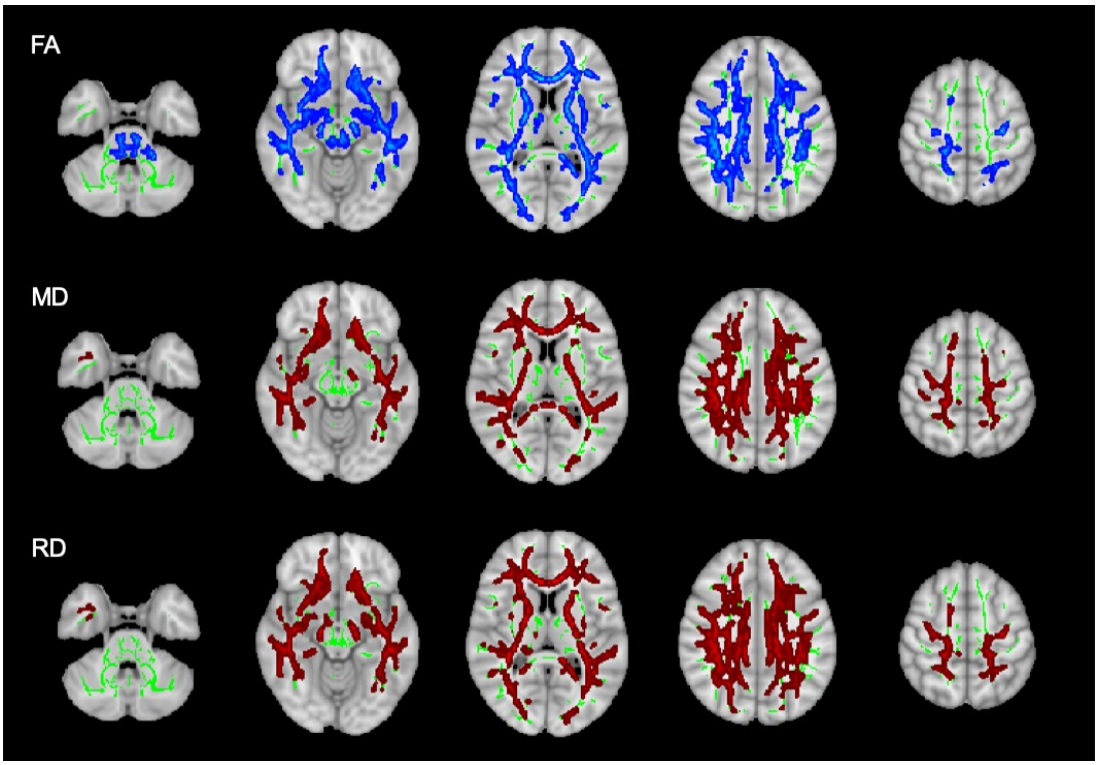
B) RMTF



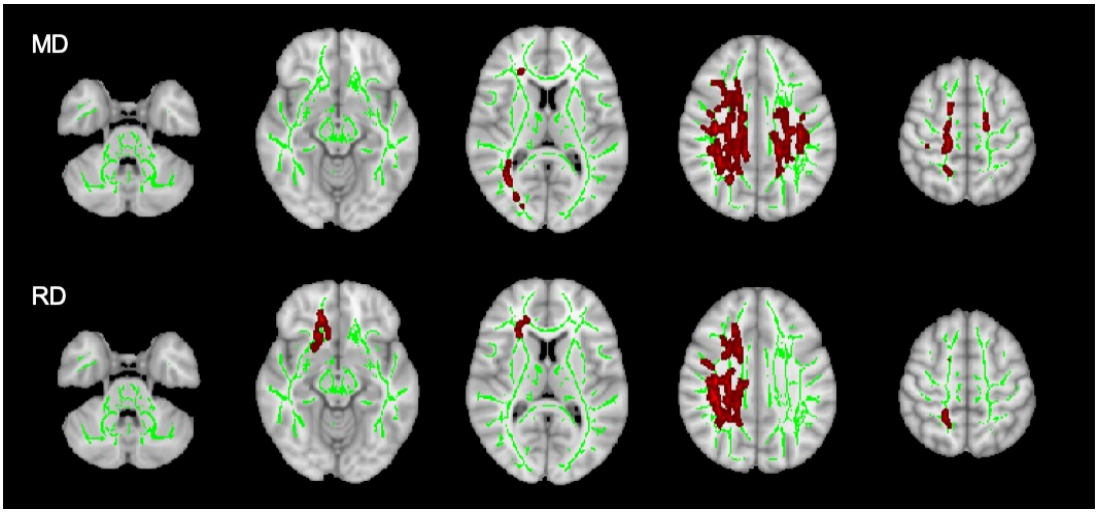
C) RMTW



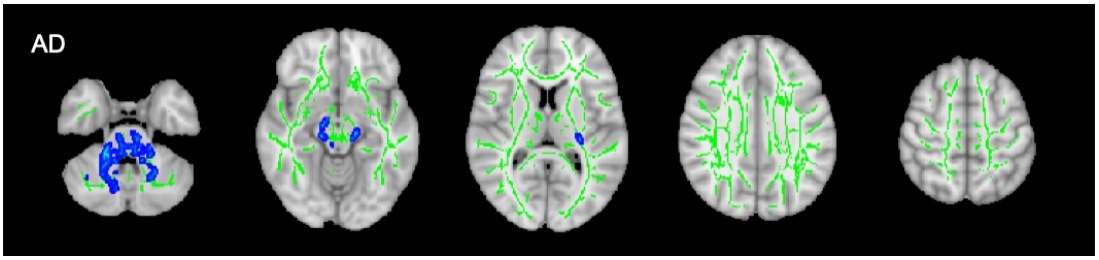
D) CPAL



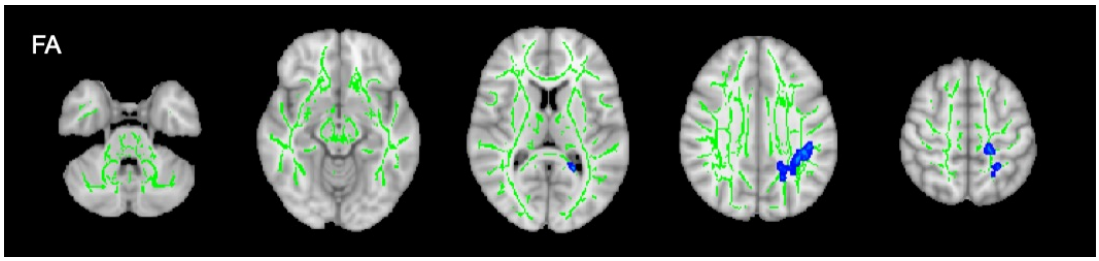
E) TMTA



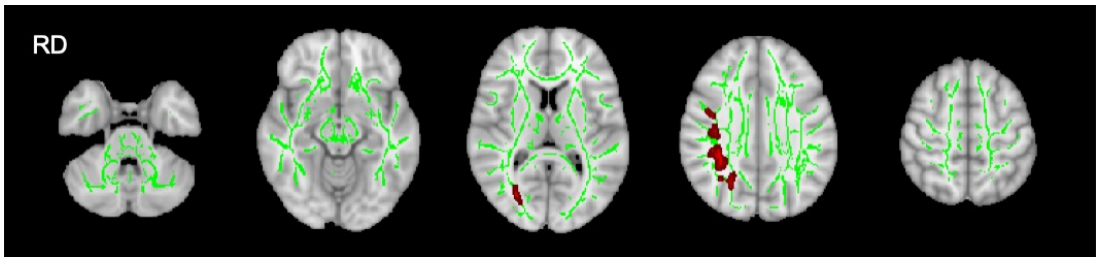
F) FAS



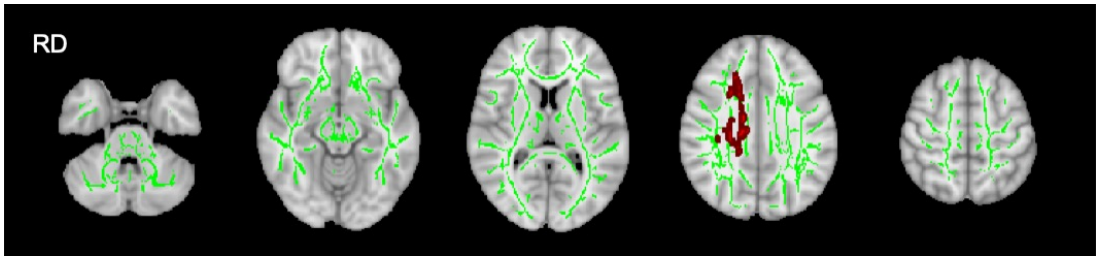
G) Animals



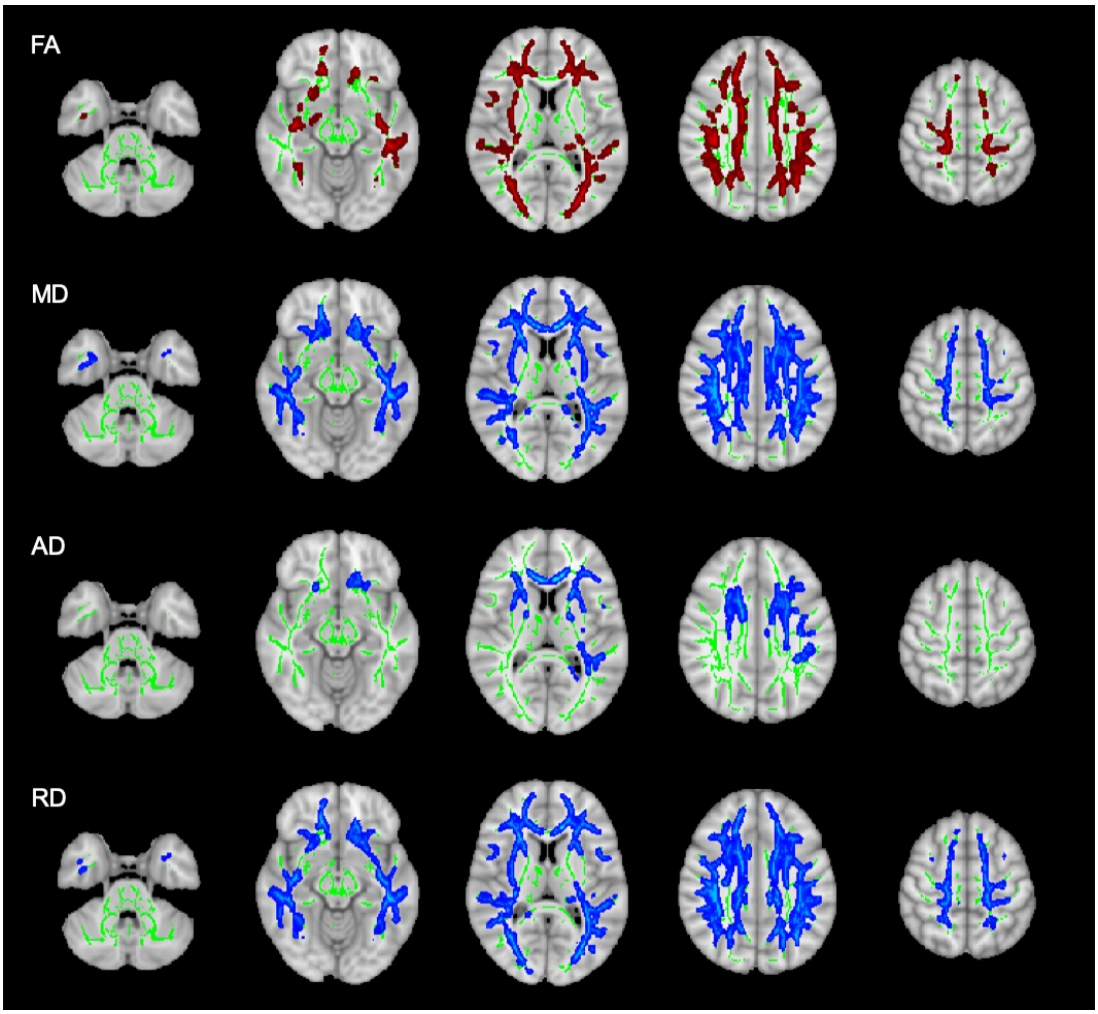
H) DKEFSI



I) TMTB



J) GDA



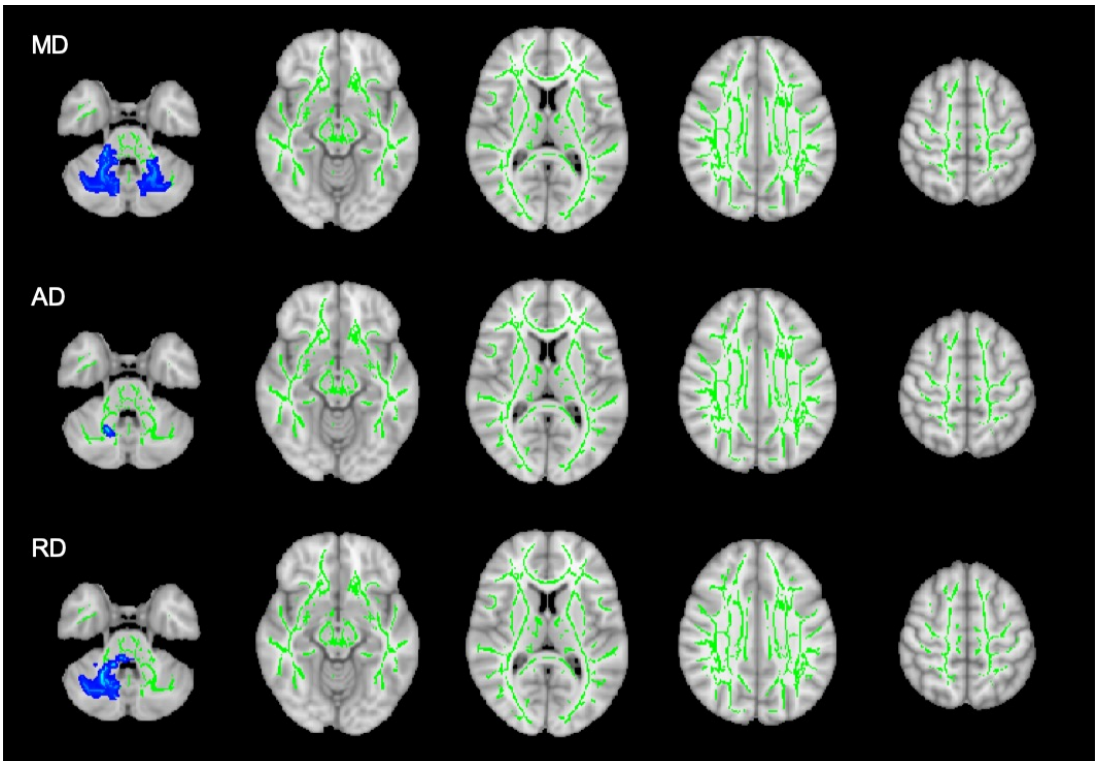
Tracts where DTI parameters increase as cognitive performance decreases are shown in red. Tracts where DTI parameters decrease as cognitive performance decreases are shown in blue. The WM skeleton is shown in green. Axial slices at z = -34, -12, 10, 32 and 54 are shown. AD, axial diffusivity; FA, fractional anisotropy; MD, mean diffusivity; RD, radial diffusivity; TBSS, tract-based spatial statistics; WM, white matter.

and, as a result, *decreases* in MD and *increases* in FA. There were no other associations between cognitive test scores and DTI parameters. Associations between ModNPI scores and DTI parameters are shown in **Figure 5-13** and associations between individual ModNPI items and DTI parameters are shown in **Figure 5-14**. Increasing ModNPI scores were associated with decreases in RD, MD and, to a lesser extent, AD in cerebellar WM tracts, including the middle cerebellar peduncle. Increasing scores for the disinhibition item were associated with increases in FA in subcortical WM tracts including the internal capsule, corpus callosum and cerebral peduncles, decreases in MD in cerebellar WM tracts and widespread decreases in RD. A similar pattern was seen for the eating habits item with widespread deceases in RD. There were no associations between DTI parameters and UWDRS-P, PHQ9 or GAD7 score or scores for other ModNPI items.

Increasing NCC concentrations in stable patients were associated with increased MD and RD throughout the WM and increased AD in the left anterior thalamic radiation, inferior longitudinal fasciculus and corpus callosum, as shown in **Figure 5-15**. Increasing tau was associated with decreases in **Figure 5-16**. There were no other associations between DTI parameters and copper indices or wet biomarkers.

There were no associations between disease duration and DTI parameters in stable patients.

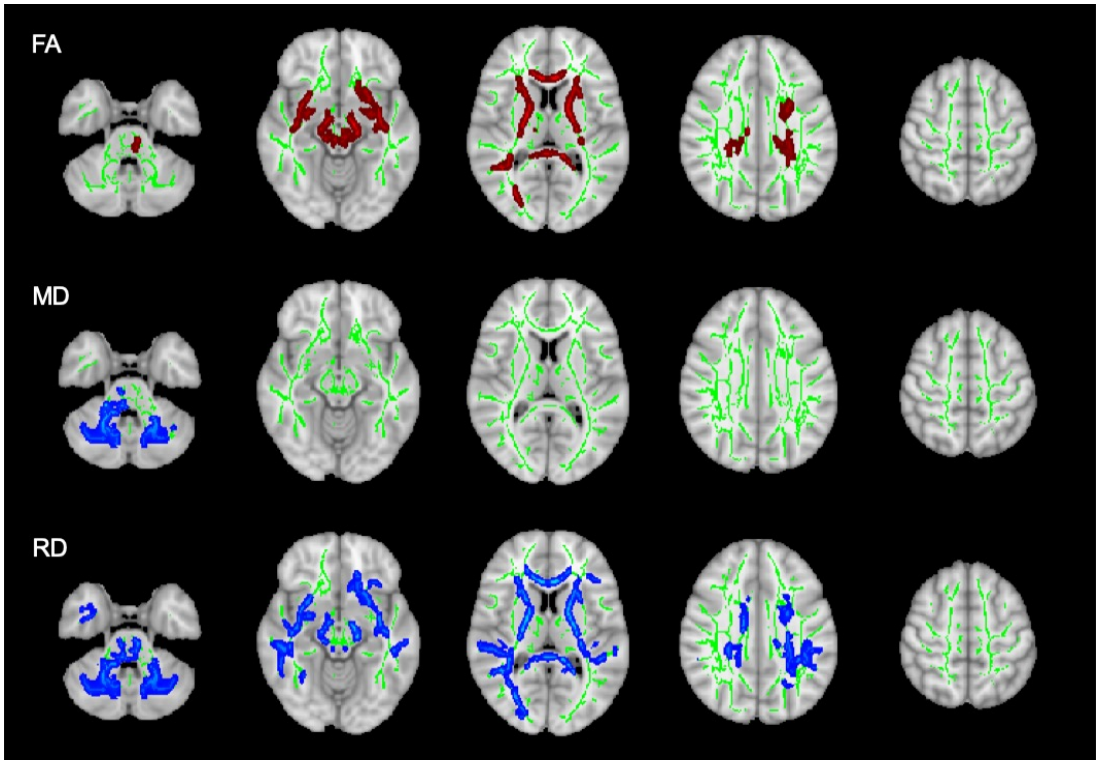
Figure 5-13 TBSS for associations with ModNPI scores



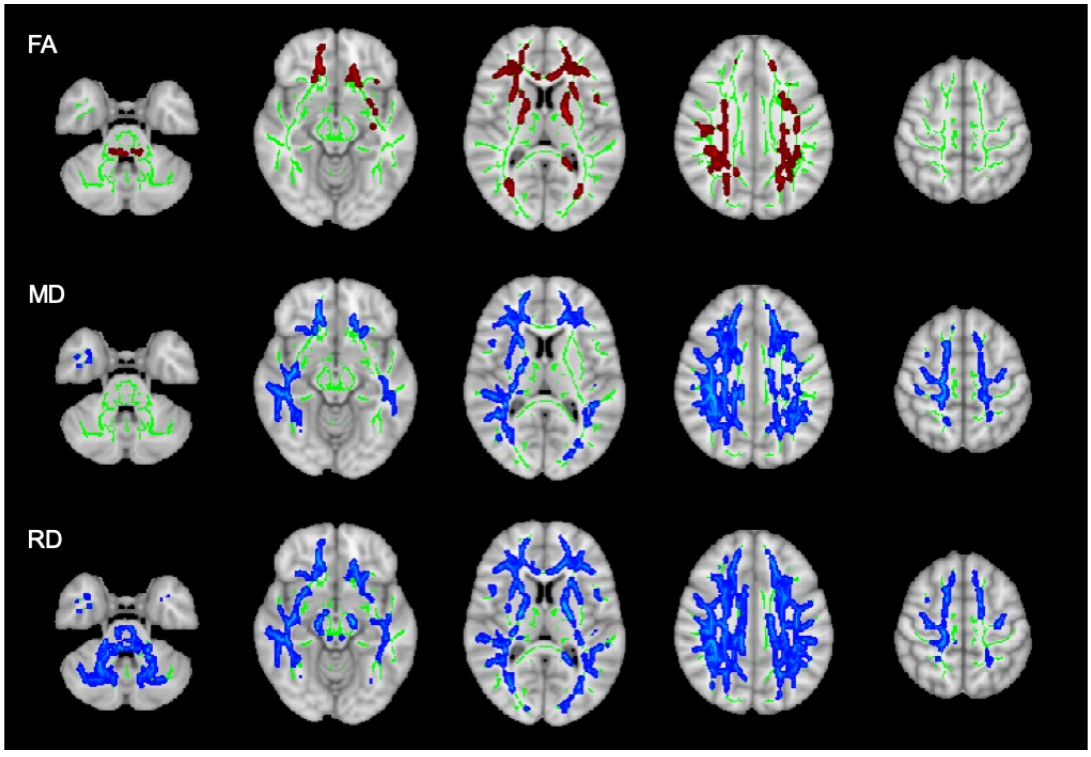
Tracts where DTI parameters increase as ModNPI decrease are shown in blue. The WM skeleton is shown in green. Axial slices at z = -34, -12, 10, 32 and 54 are shown. AD, axial diffusivity; FA, fractional anisotropy; MD, mean diffusivity; RD, radial diffusivity; TBSS, tract-based spatial statistics; WM, white matter.

Figure 5-14 TBSS for associations with ModNPI items

A) Disinhibition

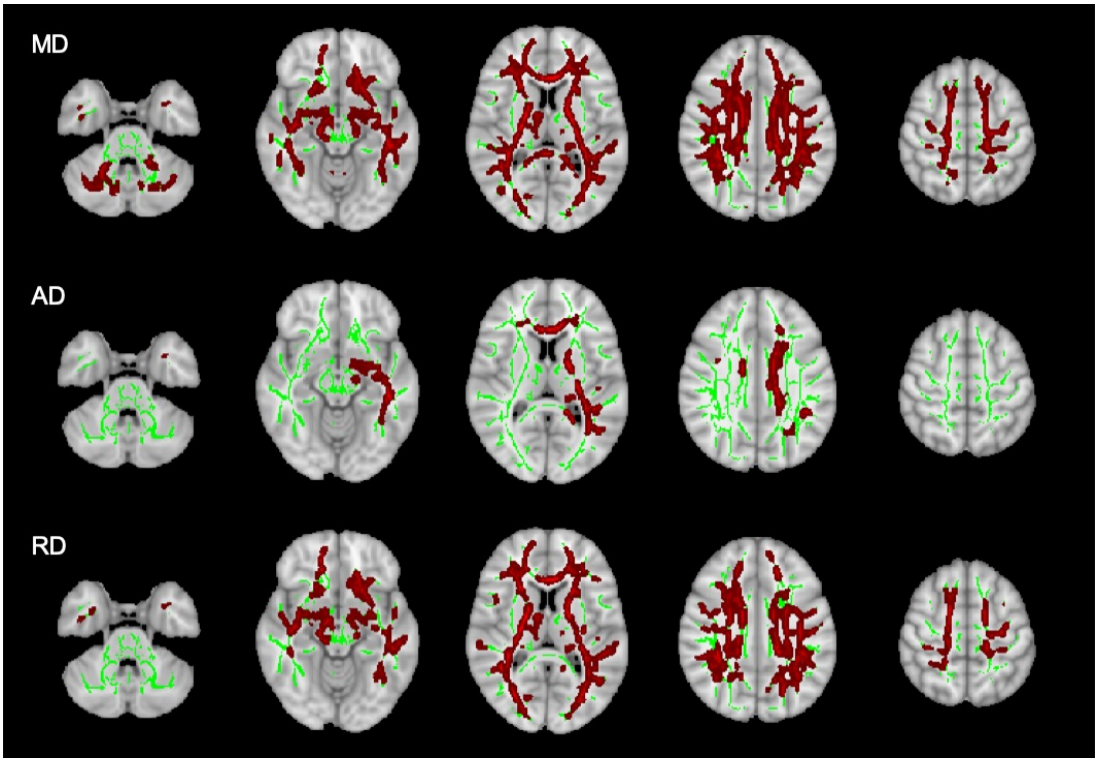


B) Eating habits



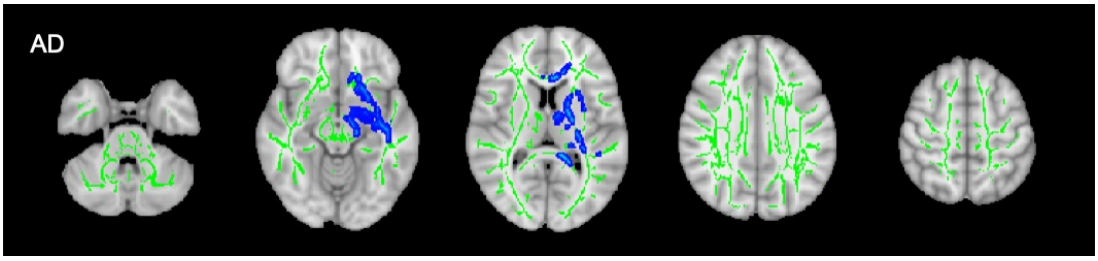
Tracts where DTI parameters increase as each ModNPI item increases are shown in red. Tracts where DTI parameters decrease as each ModNPI item increases are shown in blue. Axial slices at z = -34, -12, 10, 32 and 54 are shown. The WM skeleton is shown in green. AD, axial diffusivity; FA, fractional anisotropy; MD, mean diffusivity; RD, radial diffusivity; TBSS, tract-based spatial statistics; WM, white matter.

Figure 5-15 TBSS for associations with NCC



Tracts where DTI parameters increase as NCC increases are shown in red. The WM skeleton is shown in green. Axial slices at z = -34, -12, 10, 32 and 54 are shown. AD, axial diffusivity; MD, mean diffusivity; RD, radial diffusivity; TBSS, tract-based spatial statistics; WM, white matter.

Figure 5-16 TBSS for associations with Tau



Tracts where DTI parameters decrease as tau increases are shown in blue. The WM skeleton is shown in green. Axial slices at z = -34, -12, 10, 32 and 54 are shown. AD, axial diffusivity; TBSS, tract-based spatial statistics; WM, white matter.

Susceptibility-weighted imaging

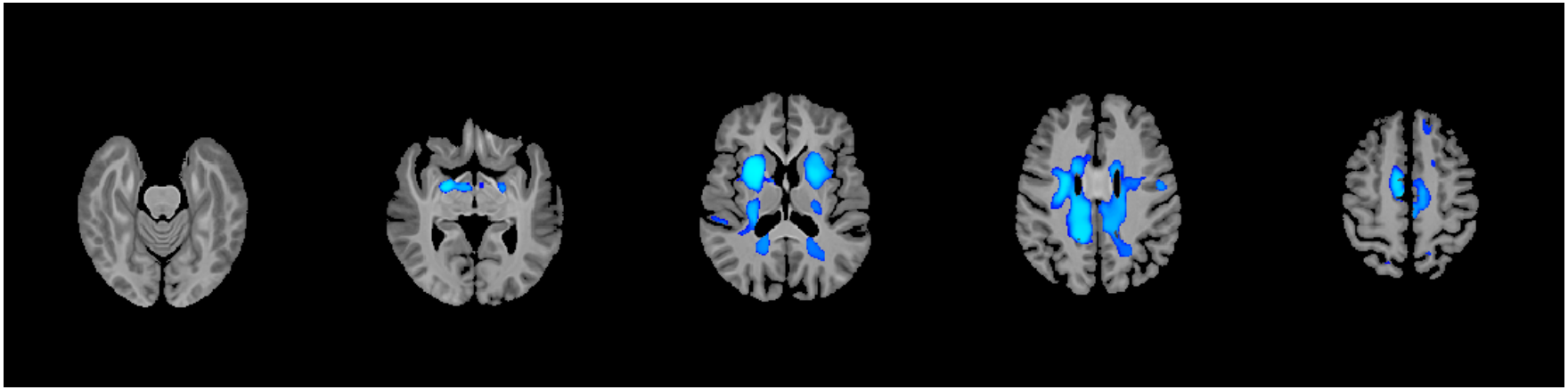
Patients with neurological presentations had increased absolute susceptibility in the bilateral putamen, cingulate and medial frontal cortices compared to patients with hepatic presentations, as shown in **Figure 5-17**. Clusters extended into regions of white matter including the internal and external capsules, corpus callosum and, to a lesser extent, corona radiata. Patients with active disease had increased susceptibility in one cluster within the left frontal pole compared with patients with stable disease, as shown in **Figure 5-18**. On post-hoc testing, this cluster appeared to be driven by the chronically-treated, non-adherent patients.

The association between UWDRS-N scores and susceptibility in stable patients is shown in *Clusters where susceptibility is higher in patients with active disease are shown in blue. QSM, quantitative susceptibility mapping.*

Figure 5-19. Increasing UWDRS-N scores were associated with increased susceptibility in scattered cortical areas within the bilateral frontal, parietal and occipital lobes, cingulate cortices, right cerebellar hemisphere but not the basal ganglia. A few of these clusters extended into adjacent white matter tracts.

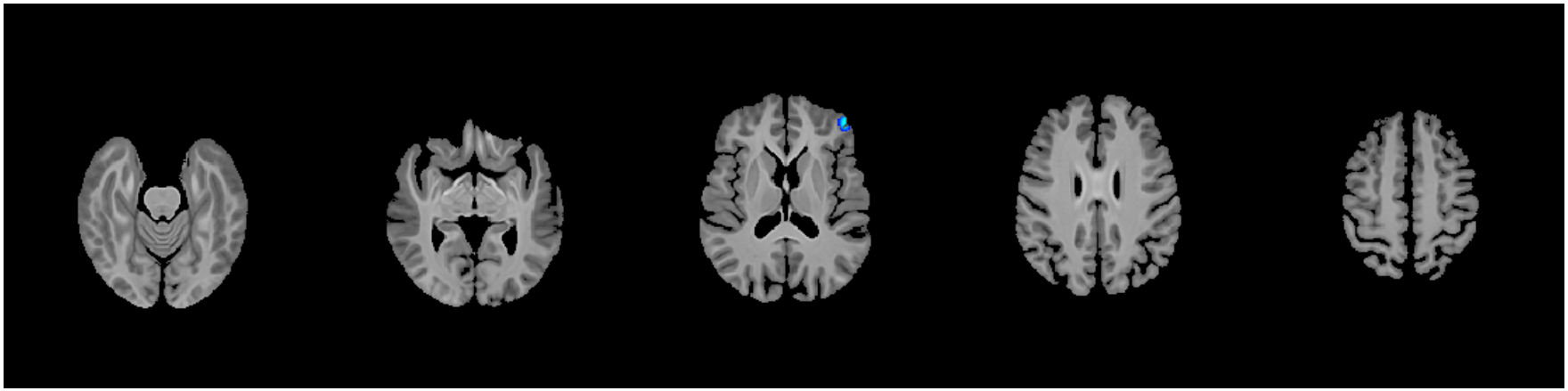
There were no associations between susceptibility and any cognitive test scores, psychiatric rating scales, copper indices, wet biomarkers or disease duration in stable patients, with one exception: A cluster in the right fronto-temporal region that was identified when test associations with tau, as shown in **Figure 5-20**.

Figure 5-17 QSM for comparison between hepatic and neurological presentations



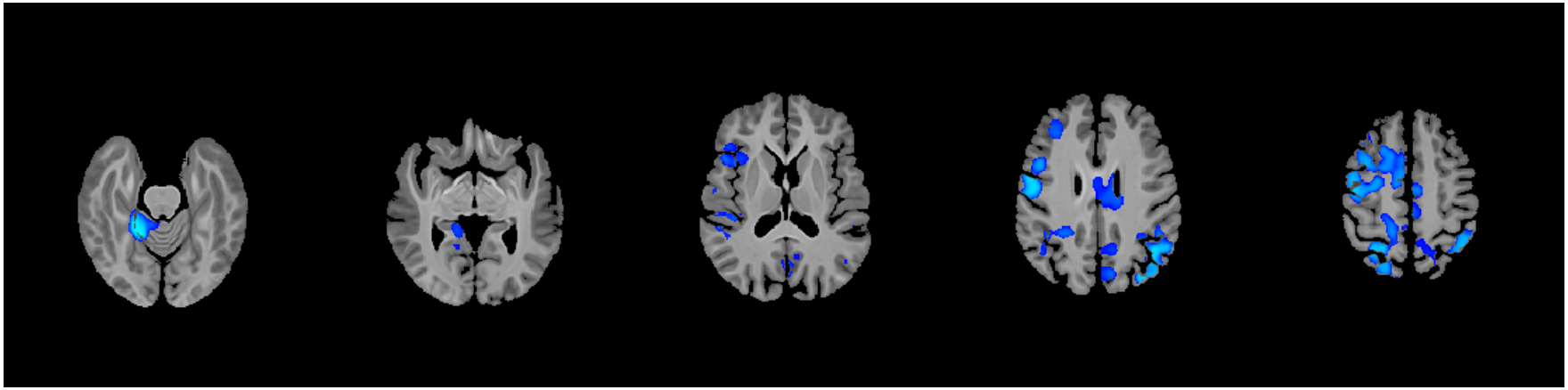
Clusters where susceptibility is higher in patients with neurological presentations are shown in blue. QSM, quantitative susceptibility mapping.

Figure 5-18 QSM for comparison between active and stable disease



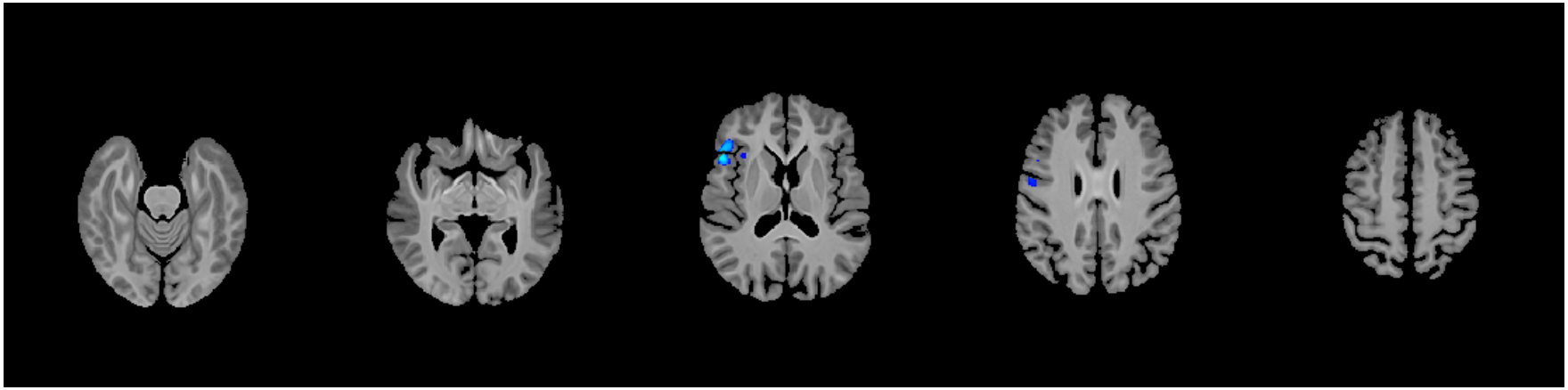
Clusters where susceptibility is higher in patients with active disease are shown in blue. QSM, quantitative susceptibility mapping.

Figure 5-19 QSM for associations with UWDRS-N scores



Clusters where increasing UWDRS-N scores are associated with increased susceptibility are shown in blue. QSM, quantitative susceptibility mapping; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore.

Figure 5-20 QSM for associations with Tau



Clusters where increasing tau is associated with increased susceptibility are shown in blue. QSM, quantitative susceptibility mapping.

5.4 Discussion

The combination of whole-brain quantitative neuroimaging methods applied here offers the most comprehensive analysis of the relationships between clinico-biochemical characteristics and neuroimaging abnormalities in WD to date. I have confirmed previous observations on associations between neurological involvement and the regional loss of brain volume and demonstrated that specific patterns of diffusion abnormalities are associated with neurological disease activity, neurological severity in chronically-treated patients and copper indices. I have identified neuroimaging correlates for a range of cognitive deficits and psychiatric features in WD using VBM and TBSS in chronically-treated patients. Using an automated lesions segmentation pipeline, I have highlighted the limitations of using WMHs as a measure of neurological involvement and with whole-brain QSM demonstrated that increasing neurological severity is associated with widespread cortical iron deposition. These findings have implications for disease mechanisms and our understanding of the progression of neuropathology in WD, in addition to validation of prognostic and monitoring biomarkers for neurological involvement.

Subcortical volume loss as a prognostic biomarker

Patients with neurological presentations had lower caudate, putamen, pallidum, thalamus, midbrain, pons and cerebellum volumes than patients with hepatic presentations. VBM analyses also identified clusters of reduced GM volume in the orbitofrontal and insula cortices. A similar pattern of predominantly subcortical volume loss was seen when testing the association with UWDRS-N scores in chronically-treated patients.

These findings are consistent with previous volumetric and deformation-based morphometry studies comparing patients with controls.[210,235,237,238,282] They indicate that these structural changes likely occur as part of the initial brain injury and represent a clinically relevant neuroimaging endpoint. Quantifying these may provide a direct measure of irreversible damage that could be used as a prognostic or predictive biomarker for neurological involvement. The high structural covariance between the

basal ganglia, thalamus, brainstem and, to a lesser extent, cerebellum is relevant here. It indicates that, despite significant phenotypic variability among patients with neurological presentations, patients with volume loss in one of these subcortical ROI are very likely to have volume loss in another.

A reductionist strategy for developing imaging biomarkers where the volume of a single ROI, such as the putamen, is calculated may therefore be as accurate as more complicated composite scores for predicting neurological outcomes. A similar approach measuring caudate volumes using automated segmentation pipelines has been proposed for tracking progression of Huntington's disease in clinical trials.[283] The chronic damage score recently proposed by Dusek *et al* includes visual ratings for cortical and cerebellar volume loss and measurements of third ventricle width and antero-posterior midbrain diameter.[284] I did not measure these in our participants but our findings suggest that measuring third ventricular width alone may be sufficient for capturing the neuroradiological sequelae of brain injury without the need for automated segmentation pipelines. Indeed, third ventricular width has previously been shown to correlate with neurological severity in transcranial sonography studies.[285,286]

I have also shown that while volume loss in these structures correlates with the severity of speech disturbance, dystonia, parkinsonism and ataxia, this was not the case for tremor. This is consistent with the emerging idea that tremor is related to disrupted neural networks,[287] and has implications for using subcortical volume as a biomarker. Measuring subcortical volume alone as a marker of irreversible neurological involvement may lead us to underappreciate the significance of tremor.

Importantly, my observations on subcortical volume loss are based on differences between patients with neurological and hepatic presentations and associations with movement disorder severity at a group level. The significance of subcortical volume loss, and other imaging biomarkers discussed below, has not yet been tested at an individual level. This will be a crucial step for demonstrating their clinical utility.

The significance of WMHs

Patients with active disease had a higher volume of WMHs in the basal ganglia and infratentorial regions than patients with stable disease. However, the lack of differences between patients with neurological and hepatic presentations and lack of association with neurological severity in chronically-treated patients are unexpected. These findings suggest that resolution of WMHs does not necessarily equate to treatment success and persistence of WMHs does not necessarily indicate treatment failure. On the basis of these results, WMHs are unlikely to be a useful monitoring biomarker in WD, despite being helpful diagnostically.

This interpretation is consistent with the recent observation that the acute toxicity score proposed by Dusek *et al*, which is based on the distribution and severity of WMHs, does not correlate with neurological severity in newly-diagnosed patients at baseline or after two years of treatment.[284] I suspect that the semiquantitative scales previously proposed by Sinha *et al* and Pulai *et al* correlated with neurological severity because they also incorporated measures of brain volume.[244,245]

Importantly, I found that patients with active disease also had WMHs in the deeper, central layers of frontal, temporal and parietal WM. These periventricular and surrounding WM regions were not included in the aforementioned semi-quantitative scales, perhaps because lesions there are less conspicuous and more difficult to characterize visually. Nonetheless, they were reported to be relatively common at initial presentation in some early qualitative MRI studies. Rubenstein *et al* noted that 5 of 22 patients, most of whom had neurological presentations, had lesions in the subcortical WM at presentation in 1987.[243] Da Costa *et al* also subsequently reported that 5 of 18 patients had WMHs in the centrum semiovale that resolved in all but one case with treatment.[252] Atypical presentations of WD with diffuse leukoencephalopathy have also been described.[288]

There are several potential mechanisms underlying WMHs in WD including demyelination, oedema, gliosis, tissue necrosis and cavitation and these may occur concurrently.[29,266,289-291] Our data support the idea that some lesions, and therefore some of the pathological processes responsible for

these lesions, are reversible whereas others are not. It is plausible that pathophysiological basis for WMHs varies in different brain regions or tissues and between newly-diagnosed and chronically-treated patients and this may account for the lack of association with clinical neurological involvement. The distribution of lesions, extending to more diffuse areas of periventricular WM, emphasises the importance of recognising WM pathology and potential consequences for network dysfunction in WD, discussed further below.

DTI as a monitoring biomarker

There were no differences in DTI parameters between patients with neurological and hepatic presentations. However, increasing neurological severity in chronically-treated patients was associated with decreasing AD in subcortical WM tracts. I suspect these decreases in AD reflect loss of axons, i.e. WM atrophy. Smolinski *et al* observed that total WM volume was inversely associated with UWDRS-N in newly-diagnosed patients,[240] and there is supporting evidence from a mouse model of Wallerian degeneration that axonal loss correlates with reductions in AD.[292]

Patients with active disease, although small in number, had higher MD, AD and, to a lesser extent, RD than patients with stable disease. This is consistent with previous observation on MD in ROI-based studies,[253,293] and the observation by Lawrence *et al* that MD, AD and RD decrease with treatment. The pathophysiological basis for these diffusion abnormalities is less clear. Increases in membrane permeability, fibre reorganisation, destructions of intracellular compartments and glial alterations might all affect diffusion of water molecules in unanticipated directions.[294]

In contrast to previous studies,[82,238,253] I did not identify any abnormalities in FA in WM. I suspect this relates to the small number of newly-diagnosed patients, who may have more pronounced changes in WM diffusivity, in our cohort. Our data does, however, highlight the importance of considering absolute measures of diffusivity, such as AD and RD, separately in WD. I have demonstrated distinct patterns of abnormalities in these parameters that may be missed or underappreciated with an overreliance on FA and MD, which are essentially functions of AD and RD. Our data also support observations in Alzheimer's disease that early neuropathological

processes can be associated with changes in WM diffusivity that are proportional along each semi-principal axis and therefore do not alter FA.[295]

In combination with the observation of Lawrence *et al*,[82] our data suggest that measure of diffusivity in WM tracts are promising monitoring biomarkers for neurological involvement in WD. The bidirectional relationship with AD in patients with active and stable disease is noteworthy and raises the possibility that DTI might help differentiate reversible from irreversible neurological involvement in WD. As discussed below, DTI parameters also correlated with NCC in chronically-treated patients. The question remains how DTI data might be applied in a clinical setting or in a clinical trial for WD. Variation in sequence parameters and pre-processing steps makes comparing images acquired from different sites and deriving universal cut-off values more challenging. However, DTI parameters at pre-specified ROI have recently been used as secondary endpoints in clinical trials for multiple sclerosis.[296,297] A similar approach has been proposed for trials for FTDs.[298] Further longitudinal studies are required to determine how DTI parameters evolve in the initial stages of treatment and with paradoxical neurological worsening.

Cortical iron deposition

Using whole-brain QSM I found that patients with neurological presentations had increased absolute susceptibility in the caudate, putamen and medial frontal cortices compared to patients with hepatic presentations. Several of the clusters I identified extended into subcortical white matter tracts, including the internal and external capsules, and the corpus callosum. Increasing UWDRS-N scores were associated with increased susceptibility in scattered cortical areas within the bilateral frontal, parietal and occipital lobes, cingulate cortices, right cerebellar hemisphere but not the basal ganglia.

Increased susceptibility in the basal ganglia was seen in previous ROI-based studies and, in the absence of any association with disease duration, suggest iron deposition occurs at the time of the initial brain injury.[235] Given the lack of correlation with the severity of movement disorders, cognitive deficits or psychiatric features in WD, susceptibility in the basal ganglia appears to have a limited role as a biomarker for neurological involvement.

Our observations on susceptibility in cortical regions are novel and intriguing. I suspect that increased absolute susceptibility in cortical regions, similar to that in the basal ganglia, represents abnormal iron deposition. However, it is important to emphasize that I used *absolute* as opposed to *signed* susceptibility maps. This involves collapsing effects from paramagnetic (positive) and diamagnetic (negative) susceptibility into an unsigned measure of susceptibility. This has been shown to reduce the spatial gradient of blooming effects and improve sensitivity for identifying changes in cortical susceptibility.[299] The disadvantage of this approach is that it is not possible to differentiate paramagnetic and diamagnetic materials.

We should consider whether changes in other causes of paramagnetic or diamagnetic susceptibility might explain our findings. The assertion that cortical SWI abnormalities reflect iron deposition is primarily based on the observation that SWI abnormalities in the basal ganglia correlated with iron and not copper deposition in post-mortem tissue from nine patients with WD only.[301] However, it seems unlikely that these observations relate to copper deposition because I did not find any association between absolute susceptibility and NCC. Myelin has a diamagnetic effect and loss of myelin has been shown to lead to increasing (less negative) susceptibility in WM and, counterintuitively, decreasing (less positive) susceptibility in GM.[300] The increased absolute susceptibility I observed in WM when comparing patients with neurological and hepatic presentations could therefore be explained by loss of myelin and/or iron deposition. Analysis of R2* maps may help differentiate these but further histopathological studies will ultimately be required.

Implications for disease mechanisms

Assuming cortical SWI abnormalities in our cohort represent abnormal iron deposition, the observations here might have important implications for our understanding of the pathophysiological basis for neurological involvement in WD. As discussed in the Chapter 1, the reason why only some patients develop movement disorders is unknown. Dusek *et al* previously demonstrated that patients with more pronounced histopathological severity have increased iron deposition in the putamen and this correlates with the

presence of iron-containing macrophages.[301] Here, I demonstrate that chronically-treated patients with more severe neurological involvement have increased cortical susceptibility. This may suggest that a tendency to mishandle brain iron in response to copper accumulation predisposes certain patients to more marked neurodegeneration, particularly in the basal ganglia. This would explain how some patients with hepatic presentations can have markedly elevated cerebral copper content but not develop neurological symptoms.[25,266]

An alternative explanation is that iron deposition occurs as an epiphenomenon to neurodegeneration or in response to ongoing copper toxicity. However, there was no association between neurological severity and cortical volume loss or between NCC and cortical susceptibility. If a tendency to mishandle iron does predispose some patients to neurodegeneration then we also need to consider why increasing neurological severity was not associated with increased susceptibility in the basal ganglia in our analyses given this has been demonstrated in a previous histopathological study. I suspect this relates to a ceiling effect for quantifying iron deposition in the basal ganglia using QSM. It might be possible to test this theory by acquiring susceptibility-weighted images at lower field strengths.

The precise mechanism through which iron deposition occurs in WD is unknown however caeruloplasmin deficiency may play an important role. In addition to being the main copper-carrying protein secreted in to the systemic circulation, caeruloplasmin also has ferroxidase activity, converting ferrous iron (Fe^{2+}) to ferric iron (Fe^{3+}). Transferrin, the main iron transport protein, can only bind ferric iron and so caeruloplasmin deficiency as a result of ATP7B dysfunction may lead to an inability to mobilise brain iron. This is evident in acaeruloplasminaemia, a rare neurodegenerative disease where loss of function mutations in the *CP* gene lead to a brain iron accumulation disorder characterised by cognitive impairment and movement disorders. Caeruloplasmin is expressed in neurons and expression in Purkinje neurons is decreased in *ATP7B* knockout mice.[302] Patients with WD also been shown to have lower CSF caeruloplasmin that healthy controls[303].

Our hypothesis on the relationship between iron deposition and neurodegeneration and the extent to which caeruloplasmin contributes to

dysregulation of brain iron metabolism needs to be tested in animal and cell-based models of WD.

Neuroanatomical correlates of cognitive deficits

I have demonstrated that a number of cognitive deficits are associated with distinct patterns of brain volume loss and diffusion abnormalities. Our observations have implications for understanding the neuroanatomical basis for cognitive impairment and, more broadly, the progression of brain pathology in WD, in addition to the role of specific neuroimaging findings as biomarkers.

While we often refer to cognitive tests representing specific cognitive domains, most draw on various cognitive functions and brain networks. Task-based functional MRI (fMRI) studies inform us as to which brain regions are activated when healthy controls perform specific cognitive tests, although the test are usually adapted to enable them to be performed in a scanner. In combination with our quantitative MRI data, these studies can be used to infer the neuroanatomical basis for specific cognitive deficits in WD with the caveat that I cannot confirm whether the associations I have identified are causal without employing more sophisticated neuroimaging methods.[304]

I found that deficits in abstract reasoning (MRT) and executive function, specifically verbal fluency (FAS), were strongly associated lower volumes in the basal ganglia, particularly the putamen. These deficits were also associated with decreasing AD in predominantly subcortical WM tracts. fMRI studies in healthy controls show that performing the MRT is associated with activation of areas involved in problem-solving including the dorsolateral prefrontal cortex (PFC), inferior frontal gyrus and anterior cingulate cortex, and areas involved in visual processing extending from the fusiform gyrus to inferior and middle occipital cortices.[305] FAS in healthy controls is associated with activation of regions in the left PFC including the superior, middle and inferior frontal gyri, and left basal ganglia and thalamus.[306] I suspect our observations on MRT and FAS therefore reflect disruption of fronto-striatal networks that originate on the PFC, project to the caudate and putamen and then pallidum before reverting back to the PFC via the thalamus.[307]

In contrast, deficits in recognition memory for faces correlated with decreased GM volumes in diffuse anterior cortical regions, including the PFC,

cingulate cortex and cerebellum, in addition to the basal ganglia. They were also associated with diffusion abnormalities in widespread WM tracts driven by increasing RD. Using [$H_2{}^{15}O$]-PET, Kim *et al* demonstrated that recognition memory for faces in healthy individuals localizes to the right lingual and fusiform gyri, which form part of the ventral stream for object recognition, and right inferior parietal lobule, which forms part of the dorsal stream for object lateralization.[308] Activation of the left cingulate cortex and cerebellum is seen when performing the recognition memory test for both faces and words. More recently, Cohen *et al* used lesion network mapping to identify connections common to lesions causing prosopagnosia.[309] To their surprise, lesions were consistently linked to a functionally connected brain network that involves not only the right fusiform gyrus but also the left PFC. I did not identify clusters in or near the dorsal or ventral streams for visual processing on VBM and so suspect that deficits in recognition memory for faces in WD could be driven by damage to this network at the left PFC, connections between the left PFC and right fusiform gyrus and/or dysfunction in the cingulate cortex. As with increases in RD seen in association with active disease and NCC, the pathophysiological basis for increases in RD are unclear but seem unlikely to be driven by axonal loss alone.

Deficits in facial emotion recognition were associated with a similar pattern of abnormalities with decreased GM volumes in diffuse anterior cortical regions. In an fMRI study, facial emotion recognition was associated with activation of the PFC, basal ganglia, thalamus, amygdala and other brain regions including the superior temporal sulcus, inferior parietal lobe and somatosensory and occipital cortices.[310] Others have highlighted the importance of the anterior cingulate cortex in detecting and discriminating affective information.[311] Using voxel-based lesion-symptom mapping, Tsuchida *et al* found that the ventromedial PFC, within which they included the medial orbitofrontal cortex, plays an important role in detecting subtle emotional signals from faces whereas the left lateral PFC plays a role in discriminating between specific emotions.[312] Wolf *et al* subsequently suggested that the ventromedial PFC mediates recognition of facial emotions by guiding visual attention to emotionally salient regions of the face.[313] Russell *et al* recently reported that impaired facial emotion recognition in genetic FTDs

is associated with networks involving the striatum and orbitofrontal and insula cortices.[314] The authors point out that the orbitofrontal cortex has a role in stimulus reinforcement learning and processing reward whereas the insula is involved in a variety of processes related to social cognition including interoception, processing of emotional experiences and awareness of positive and negative feelings. I suspect that damage to the orbitofrontal, insula and/or anterior cingulate cortices might account for deficits in social cognition in WD.

I assessed cognitive flexibility using TMTB and found that deficits were associated with decreased volume in the left supplementary motor area, left occipital fusiform gyrus and right intracalcarine cortex on VBM, in addition to the caudate, putamen and pallidum on ROI analyses. Cognitive flexibility is a complex concept that can be assessed in a variety of ways.[315] Most tests that involve set-shifting or task-shifting also draw on other cognitive functions. Performance in TMTB might be influenced by visual search, working memory, attention, inhibition and processing speed, in addition to set switching ability. Our VBM findings are consistent with fMRI studies that have shown that performing the TMTB is associated with activation in ventral and dorsal visual pathways and the supplementary motor area, particularly on the left.[316] Associations with caudate, putamen and pallidum volumes may reflect the motor aspect of these tasks and/or other executive functions recruited when performing the TMTB.

The contrasting associations between some cognitive test scores and diffusion abnormalities in WM is noteworthy. CPAL scores, like RMTF scores, were associated with widespread increases in RD whereas GDA scores were associated with widespread decreases in RD. Performing arithmetic is associated with activation of bilateral inferior frontal gyri, bilateral posterior medial frontal cortices, right insula, left intraparietal sulcus and left precentral gyrus in healthy controls.[317] However, the diffuse nature of the abnormalities I identified precludes any localisation of these cognitive functions beyond the WM, as opposed to GM. The opposing associations suggest that they might be driven by different by pathological processes both of which are present throughout WM. For example, disruptions to tissue microstructure as a result of local copper deposition or, separately, ATP7B dysfunction.

Looking back at results from Chapter 3, deficits in memory for faces, cognitive flexibility and paired associated learning were surprisingly common in patients with hepatic presentations. I proposed that these might represent an early neurological phenotype and have demonstrated here that these deficits are associated with a distinct pattern of neuroimaging abnormalities characterised by widespread increases in RD throughout WM. I previously found that deficits in abstract reasoning and verbal fluency were more closely associated with the severity of movement disorders and have shown using ROI analysis that these are associated with lower volumes in the putamen, including when controlling for neurological severity.

Basal ganglia pathology is undoubtedly the main cause of movement disorders in WD and it is tempting to assume that, similar to other neurodegenerative diseases, more widespread cortical GM and WM abnormalities reflects more advanced or severe neurological involvement in WD. However, our data suggest otherwise. Unlike the pathological protein aggregates seen in Parkinson's disease or Huntington's disease, the toxic insult in WD is ubiquitous in the brain at any early stage. The increases in brain copper, local ATP7B dysfunction or other factors that lead to damage in cortical GM and WM may provoke a subtle cognitive phenotype that precedes more localised basal ganglia pathology and the emergence of a movement disorder which is associated with more prominent executive dysfunction.

Neuroanatomical correlates of psychiatric features

I have identified neuroimaging correlates for anxiety, deficits in responding to social and emotional cues, disinhibition and altered eating habits in chronically-treated patient with WD. fMRI studies in healthy controls, who do not usually experience these symptoms, are less helpful here but parallels can be drawn from neuroimaging studies on other neurodegenerative and psychiatric disorders.

Scores for the anxiety item on the ModNPI were associated with a cluster of decreased GM volume in the posterior cingulate cortex. The function of the posterior cingulate cortex is poorly understood but it appears to be a core hub, connected with a number of different brain networks, and has been linked with various brain functions including memory, navigation and narrative

comprehension.[318] Nour *et al* recently reported that anxiety in Alzheimer's disease was associated with decreased volume in the posterior cingulate cortex, in addition to the putamen, left insula and left parahippocampal gyrus.[319] I did not identify neuroimaging correlates for GAD7 scores reinforcing the idea that family member or caregiver reports may be more relevant than subjective anxiety.

Scores for the social/emotional cues item on the ModNPI were associated with volume loss involving posterior cortical regions and limbic structures such as the posterior cingulate and insula cortices, hippocampus, and amygdala. This contrasts with regions associated with Ekman scores which included more anterior structures such as the anterior cingulate, orbitofrontal and wider pre-frontal cortices. The reasons for this are unclear. The social/emotional cues item on the ModNPI involves asking family members or care givers about a number of aspects of social and emotional behaviour including whether the participant seems to be worse at interacting socially with others, insensitive to how a normal conversation flows or have difficulty reading others emotions whereas the Ekman test specifically measures ability to recognise facial expressions. Vogt has suggested that the limbic system can be divided into at least three subsystems including an anterior 'emotional' division and posterior division that 'engages in spatial orientation to extract personally relevant objects from the visual environment'.[320] Our findings might suggest that both of these subsystems are affected in WD with the Ekman test probing the anterior subsystem and the social-emotional cues item probing the posterior subsystem.

Scores for disinhibition and eating habits were both associated with widespread decreases in RD with resulting decreases in MD and increases in FA. These scores were likely driving the association between total ModNPI scores and diffusion abnormalities in the cerebellum. It is plausible that the item on eating habits might inadvertently detect disinhibited behaviour given it asks about increased liking for sweet foods and whether participants stuff food in their mouth, in addition to more general changes in appetite, weight or eating habits. However, there was no association between scores for items on disinhibition and eating habits on post hoc testing and so these items appear to measure distinct behavioural issues in WD.

As before, it is difficult to localise the neuroanatomical basis for these symptoms beyond stating that they reflect damage in WM, as opposed to GM. In Alzheimer's disease, disinhibition measured using the NPI is associated with volume loss in bilateral cingulate and right middle frontal gyri.[321] A variety of abnormal eating behaviours including hyperphagia, indiscriminate eating and preference for sweet food have been described in behavioural variant FTD and linked to fronto-insular and anteromedial temporal regions.[322] Damage to a number of WM tracts might disrupt these networks and provoke disinhibited behaviour and abnormal eating habits in WD.

Irrespective of the neuroanatomical basis for these psychiatric symptoms, it is interesting to note that, similar to scores for GDA, they were associated with widespread decreases in RD. This supports the idea that the pathological basis for psychiatric symptoms in WD may be distinct from the disease mechanisms responsible for movement disorder and other cognitive deficits.

Copper indices and other wet biomarkers

NCC was associated with decreased GM volumes in scattered cortical areas including the left precentral gyrus, right lateral occipital cortex and bilateral precuneus and widespread decreases in RD, MD and, to a lesser extent, AD. Smolinski *et al* previously demonstrated that NCC was associated with lower total brain volume and our findings confirm that this relates to cortical volume loss.[240] Our observations on the association between NCC and diffusion parameters are novel and support the role of diffusivity as a monitoring biomarker for neurological involvement in WD. They also justify the use of NCC as a clinically-relevant copper index in our study and in clinical practice.

Importantly, I are unable to discern whether the association between NCC and diffusivity reflect dynamic changes in WM microstructure as a result of ongoing copper toxicity or more longstanding WM damage. The latter situation could plausibly arise if patients who sustain more extensive WM injury at initial presentation tend to have higher NCC when chronically-treated. However, this seems unlikely given the lack of association between NCC and neurological severity.

I also identified associations between plasma tau and decreased GM volumes in the caudate, putamen and pallidum and with decreasing AD in predominantly left-sided subcortical tracts. These abnormalities are similar to the pattern of abnormalities seen in association with UWDRS-N scores and, unsurprisingly, did not persist when including UWDRS-N scores as a covariate. I did not identify associations between plasma NfL, UCH-L1 or GFAP and neuroimaging abnormalities. I previously proposed the idea that a pathophysiological trait that predisposes patients to more severe neurological involvement at presentation might explain the association between neurological severity and plasma tau in chronically-treated patients with WD. Here, our neuroimaging data indicate that increases in plasma tau, and therefore this unknown pathophysiological trait, are related to damage localised to the basal ganglia and surrounding WM.

Disease duration

I did not identify any association between diseases duration and neuroimaging abnormalities in chronically-treated patients. Although a seemingly minor point, this has not been systematically tested in the past. It supports the widely held view that chelation therapy arrests the neurological disease process in most patients. For the purposes of our analyses, it also confirms that disease duration does not need to be included as a covariate.

However, there are some caveats here. Firstly, it is difficult to define when the disease begins. I used symptom onset, acknowledging that we had to exclude asymptomatic patients, but it seems likely that subclinical neurological involvement may precede symptom onset by years or even decades. Secondly, I excluded patients with active disease when testing association with disease duration and it is likely that disease progression despite chelation therapy would lead chronically-treated patients to be classified as having active disease. Finally, I included age as a covariate when testing associations with disease duration. This may create issues with multicollinearity given age and disease duration are correlated in chronically-treated patients ($r=0.56$, $P=0.001$). The variance inflation factor for disease duration in a linear regression model that also includes age and sex is 1.77 suggesting that the extent to which result might be influenced by

multicollinearity is low but I cannot exclude that this has influenced our observations on disease duration.

Limitations

There are some specific limitations to the neuroimaging methods applied here, beyond the more general limitations of applying these to our cohort, which are covered in Chapter 3:

VBM is susceptible to systematic misregistration errors when spatially normalising atypical brains.[323] These can introduce shape changes in some subcortical structures relevant to our study. Gitelman *et al* examined performance in segmenting and normalising images acquired from patients with herpes simplex encephalitis using SPM99 and found that caudate nucleus was particularly vulnerable to registration errors.[324] However, volumetric analyses, which I performed in parallel and yielded similar results, should avoid this issue. It is also important to highlight that while smoothing the normalised images helps overcome small misregistration errors and improve signal-to-noise ratio, it also reduces spatial resolution and introduces edge artefacts.

TBSS is a widely used technique for assessing WM microstructure using DTI however a tensor is a relatively simplistic models of diffusion. Parameters reflect average diffusion within a given voxel and this creates issues where fibre crossing occurs. For example, a voxel may contain only axons but FA will be low if there are multiple crossing fibres. I suspect that NODDI, which includes indirect measures of intracellular and extracellular water content, and fixel-based analyses, which can identify microstructural changes in individual fibre populations, might provide more insights into the significance of WM abnormalities in WD. A further limitation regarding the use of DTI in our study is that diffusion acquisitions, which require prolonged scan times, are also particularly sensitive to movement artefact.[325] The EDDY module in the pre-processing of diffusion images is used to ameliorate movement artefact but it remains a concern when imaging patients with movement disorders. I cannot exclude the possibility that head motion in patients with neurological presentations has somehow biased some of our results, although motion artefact in across other sequences was limited.

While Bayesian model selection has been shown to perform well as a tool for lesion segmentation compared to other freely-available methods,[273] I did not control for vascular risk factors when analysing the volume and distribution of WMHs. This may be important given the age distribution of our cohort. In addition, recent data published by Dezortova *et al* suggests that imaging patients with WD at 3T accentuates susceptibility-related abnormalities in the basal ganglia, and compared with imaging at 1.5 T, T2 relaxation times may be shorter (and hyperintensities less prominent).[326] By acquiring FLAIR images at 3T I may therefore be underestimating the volume of WMHs in our participants relative to other studies, which mostly used 1.5T or less.

Finally, issues with VBM related to registration and smoothing are also relevant for the whole-brain QSM pipeline we used. It is possible that QSM abnormalities in subcortical structures may have been attenuated by registration errors. The limitations of using absolute instead of signed QSM has been covered above.

6. Conclusions

WD is a complex neurodegenerative disease with a variable clinical phenotype. Chelation therapy may improve neurological symptoms but monitoring relies on copper indices, which do not reflect neurological severity, and outcomes remain unpredictable. The main aims of the CROWD study were to comprehensively characterise the clinical features in a cohort of patients with WD and then to identify wet (fluid) and imaging biomarkers for neurological involvement. Secondarily, we aimed to develop a clearer understanding of the pathophysiological and neuroanatomical bases for movement disorders, cognitive deficits and psychiatric features in WD. In this final chapter, I draw together our key findings, their wider implications and opportunities for future work before proposing a unifying theory for the evolution of neurological involvement based on our results.

I classified patients according to their initial presentation and ongoing neurological disease activity and, using a combination of clinical rating scales and assessments, examined the inter-relationships between movement disorders, cognitive deficits and psychiatric features in chronically-treated, stable patients with WD. I have shown that while movement disorders are closely associated with deficits in executive function and abstract reasoning, patients with hepatic presentations (with no or minimal ongoing neurological signs) may have deficits in recognition memory for faces, cognitive flexibility and associative learning. This supports the idea that there is a continuum of neurological involvement in WD. I also found that psychiatric features did not correlate with movement disorders or cognitive deficits, at least in chronically-treated patients, and were associated with distinct neuroimaging abnormalities. These findings raise questions about whether a binary disease classification for neurological involvement is appropriate and whether neurological and psychiatric symptoms should be combined, as they are in the current classification.

I have introduced the concept that neurological involvement in WD can be modelled as a brain injury. Supporting evidence includes observations that the majority of chronically-treated patients have stable neurological disease, specific neuroimaging findings differ between patients with neurological and

hepatic presentation (but do not correlate with persistent neurological features) and neither clinical or neuroradiological features are associated with disease duration in chronically-treated patients. This concept has important implications for clinical trial design given the extent of any initial brain injury may influence response to treatment. Cohorts may need to be stratified accordingly. Our VBM and volumetric analyses have demonstrated that measures of subcortical volume loss are promising prognostic or predictive biomarkers that likely reflect irreversible damage and may serve this purpose. Further work is required to examine to what extent neurological outcomes after several years of treatment are influenced by the severity of subcortical volume loss at presentation and whether simplified measures of subcortical volume loss, such as third ventricular width, correlate with volumetric measurements.

Using a Simoa assay, we have demonstrated that plasma NfL can differentiate between patients with neurological and hepatic presentations and those with active and stable neurological disease. Like tau and UCH-L1, it correlates with neurological severity in stable patients. This is the first description of a wet (fluid) biomarker for end-organ damage in the brain in patients with WD and supports our hypothesis that plasma NfL could be used to monitor treatment response. Further work is needed to understand how NfL changes longitudinally, particularly in the first year of treatment, and over what timeframe. Utility as a monitoring biomarker will be limited unless changes in NfL precede changes in the clinical response to treatment, measured using rating scales, by at least several weeks. We also need to test whether NfL (or other markers) can be used to identify paradoxical neurological worsening earlier or confirm NEDA in chronically-treated patients.

Furthermore, I have confirmed previous observations that serum and urine copper indices do not correlate with neurological severity, at least in chronically-treated patients. Using LC-MS^{E}, we have also generated preliminary data identifying several novel proteins of interest that could be taken forward for analyses with targeted proteomics in the wider cohort in the future.

In addition to delineating the relationship between neurological involvement and structural brain abnormalities, I have confirmed that increasing NCC is associated with cortical volume loss and used whole-brain

quantitative neuroimaging analyses to highlight the limitations of using WMHs to determine response to treatment. I have demonstrated that measures of diffusivity in WM tracts are promising monitor biomarkers that are higher in patients with active disease and correlate with both neurological severity and NCC in chronically-treated patients. The practical application of DTI parameters in measuring initial response to treatment now needs to be determined and more sophisticated NODDI or fixel-based analyses may be needed to shed light on the pathophysiological basis for these diffusion abnormalities. As with NfL, we need to better understand the interval between changes in DTI parameters and clinical response to treatment to determine their value as monitoring biomarkers.

I suspect SWI abnormalities have a more limited role as end-organ biomarkers for neurological involvement. However, our observations on cortical SWI abnormalities have implications for disease mechanisms and have led us to propose that a tendency to mishandle iron in response to the accumulation of copper in the brain may predispose some patients to more severe neurological involvement. This needs to be tested in animal and cell-based models of WD.

I identified a number of novel imaging correlates for cognitive deficits and psychiatric features in WD. I found that the cognitive deficits seen in patients with hepatic presentations, such as memory for faces, cognitive flexibility and associative learning were associated with a distinct pattern of diffusion abnormalities characterised by increasing RD. Deficits in abstract reasoning and phonemic fluency, which are more closely associated movement disorders, correlated with basal ganglia volumes and decreases in AD in subcortical WM tracts and deficits in facial emotion recognition and response to social/emotional cues were associated with cortical GM volume loss.

To conclude, I can attempt combine our findings into a unifying theory for the evolution of neurological involvement in WD. I have shown that while the severity of movement disorders and some aspects of executive function are closely associated with structural changes in the basal ganglia, other cognitive deficits and psychiatric features relate to widespread cortical GM and WM pathology and may occur independently of movement disorders. In

parallel, I found that NCC, at least in chronically-treated patients, is associated with a similar pattern of diffuse abnormalities and that brain iron deposition is linked to the severity of movement disorders but not cognitive deficits or psychiatric features. I therefore hypothesise that damage to cortical GM and WM develops early in the neurological disease course as a direct result of copper deposition in the brain and manifests with subtle cognitive deficits and psychiatric features, which may not necessarily be identified as overt neurological involvement. Damage to the basal ganglia associated with the emergence of movement disorders and more prominent executive dysfunction occurs when iron deposition in the basal ganglia, which is usually mild in patients without movement disorders, progresses and is more likely in those patients with a tendency to mishandle iron in response to copper toxicity.

This theory is consistent with previous observations that patients with hepatic presentations may have subtle cognitive deficits and develop psychiatric symptoms,[327,328] that psychiatric symptoms may precede the onset of movement disorder by up to several years,[329] and that NCC does not correlate with the presence or severity of movement disorders at initial presentation.[38] It explains how patients with can have markedly increased brain copper content without developing movement disorders,[10] and why the severity of histopathological changes in basal ganglia correlates with iron but not copper content.[301] It may also explain why attempts to identify the genetic determinants for neurological involvement, which have so far focussed on *ATP7B* genotype and the ATP7B interactome,[330,331] have been unsuccessful.

Testing this hypothesis will be challenging given all patients diagnosed with WD are started on treatment and it is not possible to conduct natural history studies in untreated patients. Nonetheless, confirming that newly-diagnosed patients without movement disorders have deficits in cognitive flexibility, recognition memory for faces and associative learning compared to healthy controls and that associations with cortical volume loss and abnormal WM microstructure persist in this patient group would provide further supporting evidence. Our hope is that this can be tested in parallel with our proposed biomarkers for neurological involvement in clinical trials for novel treatments in the coming years

Statement of contributions

I would like to thanks the following people without whom this body of work would not be possible:

Prof Thomas T Warner (Professor of Neurology)
Prof Oliver Bandmann (Professor of Neurology)
Dr Jonathan D Rohrer (Professor of Neurology)
Prof Henrik Zetterberg (Professor of Neurochemistry)
Dr Godfrey T Gillett (Consultant in Clinical Biochemistry)
Prof Emmanouil Tsochatzis (Professor of Hepatology)
Dr James Dooley (Emeritus Reader in Hepatology)
Dr David Thomas (Senior Lecturer in MRI Physics)
Miss Rhian Convery (Research Assistant)
Miss Maggie Burrows (Specialist Research Nurse)
Miss Karen Shaw (Specialist Research Nurse)
Dr Matthew Appleby (Clinical Research Physician)
Dr Elisha Chung (Research Technician)
Ms Carole Davis (MRI Radiographer)
Prof Rolf Jager (Professor of Neuroradiology)
Mr Fion Bremner (Consultant in Neuro-Ophthalmology)
Ms Lynne Jung (Clinical Biochemist)
Dr Paul Cook (Consultant in Clinical Biochemistry)
Mrs Alison Martin (Clinical Biochemist)
Dr Carolin Heller (Research Fellow)
Dr Martha S Foiani (Research Fellow)
Dr Imogen Swift (Research Fellow)
Ms Jenny Hallqvist (Research Fellow)
Prof Kevin Mills (Professor of Translational Omics)
Dr Amanda Heslegrave (Senior Research Associate)
Dr Martina Bocchetta (Senior Research Associate)
Dr Carole Sudre (Senior Research Associate)
Dr Julio Acosta-Cabronero (Senior Research Associate)

TTW, OB, JDR, HZ and I set up the CROWD study with input from GTG, ET and JD. JDR, DT and I designed the imaging protocol. JDR, RC and I devised the neuropsychological battery. TTW and I submitted the ethics and NHS sponsorship application with assistance from MB and KS. Members of the WDSG-UK assisted in the designing participant information sheets and consent forms. GTG and ET helped recruit participants. MB and I organised the study visits. I performed clinical assessments and venepuncture. MB assisted with cognitive testing and arranging self-reported questionnaires. MA assisted with lumbar punctures and EC processed and stored biosamples. CD coordinated the MRI scans and MB and I accompanied patients to the scanner. RJ reviewed MRI scans for incidental findings. FB performed slit lamp examinations. LJ and PC analysed serum/plasma samples for copper indices. AM and GTG analysed urine samples for copper output. CH performed the 4-PLEX assay with assistance from MSF, IS, AH and I. JH performed LC-MS^E on our serum samples with input from KM. I performed the VBM and TBSS analyses with assistance from MB who also ran the GIF pipeline. CS ran the Bayesian Model Selection pipeline and I ran the whole-brain QSM pipleline with support from JAC. I performed subsequent statistical analyses with input from TTW, OB, JDR and HZ.

Publications

I published the following papers related to this thesis during my fellowship:

1. Shribman S, Marjot T, Sharif A, *et al*. Investigations and management of Wilson's disease: a practical guide from the British Association for the Study of the Liver. *Lancet Gastroenterol Hepatol* (published online 13th April 2022)
2. Shribman S, Bocchetta M, Sudre CH, *et al*. Neuroimaging correlates of brain injury in Wilson's disease: A multimodal, whole-brain MRI study. *Brain* 2022;29(145):263-275.
3. Shribman S, Poujois A, Bandmann O, *et al*. Wilson's disease: Update on pathogenesis, biomarkers and treatments. *J Neurol Neurosurg Psychiatry* 2021;92(10):1053-1061.
4. Shribman S, Heller C, Burrows M, *et al*. Plasma neurofilament light as a biomarker of neurological involvement in Wilson's disease. *Mov Disord* 2021;36(2):503-508.
5. Shribman S, Warner TT, Dooley JS. Clinical presentations of Wilson disease. *Ann Transl Med* 2019;7(S2):S60.
6. Shribman S, Webb G, Taylor R, *et al*. Liver transplantation for late-onset presentations of acute liver failure in Wilson's disease: The UK experience over 2 decades. *JHEP Rep* 2020; 2(3):100096.

I published the following papers unrelated to this thesis during my fellowship:

1. Schreier D, Di Lorenzo F, Iodice F, Shribman S. Do you want to perform endovascular therapy? Perspectives from neurology trainees across Europe. *Eur J Neurol* 2020;2(3):100096.
2. Shribman S, Alexander SK, Zarkali A, *et al*. Training in neurology: lessons learnt. *Pract Neurol* 2019;19(5):431-437.
3. Shribman S, Reid E, Crosby AH, *et al*. Hereditary spastic paraplegia: from diagnosis to emerging therapeutic approaches. *Lancet Neurol* 2019;18(12):1136-1146.

References

1. Poujois A, Woimant F, Samson S, Chaine P, Girardot-Tinant N, Tuppin P. Characteristics and prevalence of Wilson's disease: A 2013 observational population-based study in France. *Clin Res Hepatol Gastroenterol* 2018; **42**: 57–63.
2. O'Brien M, Reilly M, Sweeney B, Walsh C, Hutchinson M. Epidemiology of Wilson's disease in Ireland. *Mov Disord* 2014; **29**: 1567–8.
3. Compston A. Progressive lenticular degeneration: a familial nervous disease associated with cirrhosis of the liver, by S. A. Kinnier Wilson, (From the National Hospital, and the Laboratory of the National Hospital, Queen Square, London) Brain 1912: 34; 295-509. *Brain* 2009; **132**: 1997–2001.
4. Cumings JN. The copper and iron content of brain and liver in the normal and in hepato-lenticular degeneration. *Brain* 1948; **71**: 410–5.
5. Walshe JM. The conquest of Wilson's disease. *Brain* 2009; **132**: 2289–95.
6. Tanzi RE, Petrukhin K, Chernov I*, et al.* The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene. *Nat Genet* 1993; **5**: 344–50.
7. Yamaguchi Y, Heiny ME, Gitlin JD. Isolation and characterization of a human liver cDNA as a candidate gene for Wilson disease. *Biochem Biophys Res Commun* 1993; **197**: 271–7.
8. Bull PC, Thomas GR, Rommens JM, Forbes JR, Cox DW. The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nat Genet* 1993; **5**: 327–37.
9. Shribman S, Warner TT, Dooley JS. Clinical presentations of Wilson disease. *Ann Transl Med* 2019; **7**: S60.
10. Czlonkowska A, Tarnacka B, Litwin T, Gajda J, Rodo M. Wilson's disease-cause of mortality in 164 patients during 1992-2003 observation period. *J Neurol* 2005; **252**: 698-703.
11. Walshe JM, Yealland M. Chelation treatment of neurological Wilson's disease. *Q J Med* 1993; **86**: 197–204.

12. Merle U, Schaefer M, Ferenci P, Stremmel W. Clinical presentation, diagnosis and long-term outcome of Wilson's disease: a cohort study. *Gut* 2007; **56**: 115–20.
13. Telianidis J, Hung YH, Materia S, Fontaine SL. Role of the P-Type ATPases, ATP7A and ATP7B in brain copper homeostasis. *Front Aging Neurosci* 2013; **5**: 44.
14. Roy S, McCann CJ, Ralle M*, et al.* Analysis of Wilson disease mutations revealed that interactions between different ATP7B mutants modify their properties. *Sci Rep* 2020; **10**: 13487.
15. Kumar M, Gaharwar U, Paul S*, et al.* WilsonGen a comprehensive clinically annotated genomic variant resource for Wilson's Disease. *Sci Rep* 2020; **10**: 9037.
16. Wang C, Zhou W, Huang Y*, et al.* Presumed missense and synonymous mutations in ATP7B gene cause exon skipping in Wilson disease. *Liver Int* 2018; **38**: 1504–13.
17. Liu WL, Li F, Liu L*, et al.* A novel gross deletion and breakpoint junction sequence analysis of ATP7B in a Chinese family with Wilson disease using nextgeneration sequencing and Sanger sequencing. *Mol Med Rep* 2020; **21**: 517–23.
18. Woimant F, Poujois A, Bloch A*, et al.* A novel deep intronic variant in ATP7B in five unrelated families affected by Wilson disease. *Mol Genet Genomic Med* 2020; **8**: e1428.
19. Coffey AJ, Durkie M, Hague S*, et al.* A genetic study of Wilson's disease in the United Kingdom. *Brain* 2013; **136**: 1476–87.
20. Parisi S, Polishchuk EV, Allocca S*, et al.* Characterization of the most frequent ATP7B mutation causing Wilson disease in hepatocytes from patient induced pluripotent stem cells. *Sci Rep* 2018; **8**: 6247.
21. Huster D, Kühne A, Bhattacharjee A*, et al.* Diverse functional properties of Wilson disease ATP7B variants. *Gastroenterology* 2012; **142**: 947–56.e5.
22. Ferenci P, Stremmel W, Członkowska A*, et al.* Age and Sex but Not ATP7B Genotype Effectively Influence the Clinical Phenotype of Wilson Disease. *Hepatology* 2019; **69**: 1464–76.

23. Schiefermeier M, Kollegger H, Madl C*, et al.* The impact of apolipoprotein E genotypes on age at onset of symptoms and phenotypic expression in Wilson's disease. *Brain* 2000; **123**: 585–90.
24. Kluska A, Kulecka M, Litwin T*, et al.* Whole-exome sequencing identifies novel pathogenic variants across the ATP7B gene and some modifiers of Wilson's disease phenotype. *Liver Int* 2019; **39**: 177–86.
25. Litwin T, Gromadzka G, Szpak GM, Jabłonka-Salach K, Bulska E, Członkowska A. Brain metal accumulation in Wilson's disease. *J Neurol Sci* 2013; **329**: 55–8.
26. Czlonkowska A, Tarnacka B, Litwin T, Gajda J, Rodo M. Wilson's disease-cause of mortality in 164 patients during 1992-2003 observation period. *J Neurol* 2005; **252**: 698–703.
27. Cumings. Trace metals in the brain and in Wilson's disease. *J Clin Path* 1968.
28. Horoupian DS, Sternlieb I, Scheinberg IH. Neuropathological findings in penicillamine-treated patients with Wilson's disease. *Clin Neuropathol* 1988; **7**: 62–7.
29. Meenakshi-Sundaram S, Mahadevan A, Taly AB, Arunodaya GR, Swamy HS, Shankar SK. Wilson's disease: a clinico-neuropathological autopsy study. *J Clin Neurosci* 2008; **15**: 409–17.
30. Poujois A, Mikol J, Woimant F. Wilson disease: brain pathology. *Handb Clin Neurol* 2017; **142**: 77–89.
31. Reed E, Lutsenko S, Bandmann O. Animal models of Wilson disease. *J Neurochem* 2018.
32. Sternlieb I. Mitochondrial and fatty changes in hepatocytes of patients with Wilson's disease. *Gastroenterology* 1968; **55**: 354–67.
33. Scheiber IF, Bruha R, Dusek P. Pathogenesis of Wilson disease. *Handb Clin Neurol* 2017; **142**: 43–55.
34. Zischka H, Lichtmannegger J, Schmitt S*, et al.* Liver mitochondrial membrane crosslinking and destruction in a rat model of Wilson disease. *J Clin Invest* 2011; **121**: 1508–18.
35. Borchard S, Bork F, Rieder T*, et al.* The exceptional sensitivity of brain mitochondria to copper. *Toxicol In Vitro* 2018; **51**: 11–22.

36. Przybyłkowski A, Gromadzka G, Chabik G, Wierzchowska A, Litwin T, Członkowska A. Liver cirrhosis in patients newly diagnosed with neurological phenotype of Wilson's disease. *Funct Neurol* 2014; **29**: 23–9.
37. Ferenci P, Steindl-Munda P, Vogel W*, et al.* Diagnostic value of quantitative hepatic copper determination in patients with Wilson's Disease. *Clin Gastroenterol Hepatol* 2005; **3**: 811–8.
38. Czlonkowska A, Litwin T, Dziezyc K, Karlinski M, Bring J, Bjartmar C. Characteristics of a newly diagnosed Polish cohort of patients with neurological manifestations of Wilson disease evaluated with the Unified Wilson's Disease Rating Scale. *BMC Neurol* 2018; **18**: 34.
39. Faa G, Lisci M, Caria MP*, et al.* Brain copper, iron, magnesium, zinc, calcium, sulfur and phosphorus storage in Wilson's disease. *J Trace Elem Med Biol* 2001; **15**: 155–60.
40. Taly AB, Meenakshi-Sundaram S, Sinha S, Swamy HS, Arunodaya GR. Wilson disease: description of 282 patients evaluated over 3 decades. *Medicine (Baltimore)* 2007; **86**: 112–21.
41. Beinhardt S, Leiss W, Stattermayer AF*, et al.* Long-term outcomes of patients with Wilson disease in a large Austrian cohort. *Clin Gastroenterol Hepatol* 2014; **12**: 683–9.
42. Machado A, Chien HF, Deguti MM*, et al.* Neurological manifestations in Wilson's disease: Report of 119 cases. *Mov Disord* 2006; **21**: 2192–6.
43. Frota NA, Barbosa ER, Porto CS*, et al.* Cognitive impairment and magnetic resonance imaging correlations in Wilson's disease. *Acta Neurol Scand* 2013; **127**: 391–8.
44. Seniów J, Bak T, Gajda J, Poniatowska R, Czlonkowska A. Cognitive functioning in neurologically symptomatic and asymptomatic forms of Wilson's disease. *Mov Disord* 2002; **17**: 1077–83.
45. Peyroux E, Santaella N, Broussolle E*, et al.* Social cognition in Wilson's disease: A new phenotype? *PLoS One* 2017; **12**: e0173467.
46. Dening TR, Berrios GE. Wilson's disease. Psychiatric symptoms in 195 cases. *Arch Gen Psychiatry* 1989; **46**: 1126–34.
47. Ala A, Walker AP, Ashkan K, Dooley JS, Schilsky ML. Wilson's disease. *The Lancet* 2007; **369**: 397–408.

48. Steindl P, Ferenci P, Dienes HP*, et al.* Wilson's disease in patients presenting with liver disease: a diagnostic challenge. *Gastroenterology* 1997; **113**: 212-8.
49. Ferenci P, Caca K, Loudianos G*, et al.* Diagnosis and phenotypic classification of Wilson disease. *Liver Int* 2003; **23**: 139–42.
50. Chabik G, Litwin T, Członkowska A. Concordance rates of Wilson's disease phenotype among siblings. *J Inherit Metab Dis* 2014; **37**: 131–5.
51. Czlonkowska A, Tarnacka B, Moller JC*, et al.* Unified Wilson's Disease Rating Scale - a proposal for the neurological scoring of Wilson's disease patients. *Neurol Neurochir Pol* 2007; **41**: 1–12.
52. Leinweber B, Möller JC, Scherag A*, et al.* Evaluation of the Unified Wilson's Disease Rating Scale (UWDRS) in German patients with treated Wilson's disease. *Mov Disord* 2008; **23**: 54–62.
53. Volpert HM, Pfeiffenberger J, Groner JB*, et al.* Comparative assessment of clinical rating scales in Wilson's disease. *BMC Neurol* 2017; **17**: 140.
54. Walshe JM, Yealland M. Wilson's disease: the problem of delayed diagnosis. *J Neurol Neurosurg Psychiatry* 1992; **55**: 692–6.
55. Steindl P, Ferenci P, Dienes HP*, et al.* Wilson's disease in patients presenting with liver disease: a diagnostic challenge. *Gastroenterology* 1997; **113**: 212–8.
56. Ryan A, Nevitt SJ, Tuohy O, Cook P. Biomarkers for diagnosis of Wilson's disease. *Cochrane Database Syst Rev* 2019; **2019**.
57. Litwin T, Gromadzka G, Członkowska A, Gołębiowski M, Poniatowska R. The effect of gender on brain MRI pathology in Wilson's disease. *Metab Brain Dis* 2013; **28**: 69–75.
58. Li X, Feng Z, Tang W*, et al.* Sex Differences in Clinical Characteristics and Brain MRI Change in Patients With Wilson's Disease in a Chinese Population. *Front Physiol* 2018; **9**: 1429.
59. EASL Clinical Practice Guidelines: Wilson's disease. *J Hepatol* 2012; **56**: 671–85.
60. Santiago R, Gottrand F, Debray D*, et al.* Zinc Therapy for Wilson Disease in Children in French Pediatric Centers. *J Pediatr Gastroenterol Nutr* 2015; **61**: 613–8.

61. Roberts EA, Schilsky ML, (AASLD). Diagnosis and treatment of Wilson disease: an update. *Hepatology* 2008; **47**: 2089-111.
62. Walshe JM, Yealland M. Chelation treatment of neurological Wilson's disease. *Q J Med* 1993; **86**: 197-204.
63. Litwin T, Dzieżyc K, Karliński M, Chabik G, Czepiel W, Czlonkowska A. Early neurological worsening in patients with Wilson's disease. *J Neurol Sci* 2015; **355**: 162–7.
64. Brewer GJ, Terry CA, Aisen AM, Hill GM. Worsening of neurologic syndrome in patients with Wilson's disease with initial penicillamine therapy. *Arch Neurol* 1987; **44**: 490–3.
65. Weiss KH, Askari FK, Czlonkowska A*, et al.* Bis-choline tetrathiomolybdate in patients with Wilson's disease: an open-label, multicentre, phase 2 study. *Lancet Gastroenterol Hepatol* 2017; **2**: 869–76.
66. Poujois A, Sobesky R, Meissner WG*, et al.* Liver transplantation as a rescue therapy for severe neurologic forms of Wilson disease. *Neurology* 2020; **94**: e2189–e202.
67. Murillo O, Luqui DM, Gazquez C*, et al.* Long-term metabolic correction of Wilson's disease in a murine model by gene therapy. *J Hepatol* 2016; **64**: 419–26.
68. Uerlings R, Moreno D, Murillo O*, et al.* Brain copper storage after genetic long-term correction in a mouse model of Wilson disease. *Neurol Genet* 2018; **4**: e243.
69. Walshe JM. The pattern of urinary copper excretion and its response to treatment in patients with Wilson's disease. *QJM* 2011; **104**: 775–8.
70. Walshe JM. Serum 'free' copper in Wilson disease. *QJM* 2012; **105**: 419–23.
71. Pfeiffenberger J, Lohse CM, Gotthardt D*, et al.* Long-term evaluation of urinary copper excretion and non-caeruloplasmin associated copper in Wilson disease patients under medical treatment. *J Inherit Metab Dis* 2019; **42**: 371–80.
72. Califf RM. Biomarker definitions and their applications. *Exp Biol Med (Maywood)* 2018; **243**: 213–21.

73. Poujois A, Trocello JM, Djebrani-Oussedik N*, et al.* Exchangeable copper: a reflection of the neurological severity in Wilson's disease. *Eur J Neurol* 2017; **24**: 154–60.
74. Khalil M, Teunissen CE, Otto M*, et al.* Neurofilaments as biomarkers in neurological disorders. *Nat Rev Neurol* 2018; **14**: 577–89.
75. Souza G, Guest PC, Martins-de-Souza D. LC-MS(E), Multiplex MS/MS, Ion Mobility, and Label-Free Quantitation in Clinical Proteomics. *Methods Mol Biol* 2017; **1546**: 57–73.
76. Favrole P, Chabriat H, Guichard JP, Woimant F. Clinical correlates of cerebral water diffusion in Wilson disease. *Neurology* 2006; **66**: 384–9.
77. Yang J, Li X, Yang R*, et al.* Susceptibility-Weighted Imaging Manifestations in the Brain of Wilson's Disease Patients. *PLoS One* 2015; **10**: e0125100.
78. Weston PSJ, Poole T, Ryan NS*, et al.* Serum neurofilament light in familial Alzheimer disease: A marker of early neurodegeneration. *Neurology* 2017; **89**: 2167–75.
79. Menke RA, Korner S, Filippini N*, et al.* Widespread grey matter pathology dominates the longitudinal cerebral MRI and clinical landscape of amyotrophic lateral sclerosis. *Brain* 2014; **137**: 2546–55.
80. Acosta-Cabronero J, Cardenas-Blanco A, Betts MJ*, et al.* The whole-brain pattern of magnetic susceptibility perturbations in Parkinson's disease. *Brain* 2017; **140**: 118-31.
81. Stezin A, George L, Jhunjhunwala K*, et al.* Exploring cortical atrophy and its clinical and biochemical correlates in Wilson's disease using voxel based morphometry. *Parkinsonism Relat Disord* 2016; **30**: 52–7.
82. Lawrence A, Saini J, Sinha S*, et al.* Improvement of Diffusion Tensor Imaging (DTI) Parameters with Decoppering Treatment in Wilson's Disease. *JIMD Rep* 2016; **25**: 31-7.
83. EK W. The Graded Naming Test: A Restandardisation. Neuropsychological Rehabilitation; 1997. p. 143–6.
84. Tombaugh TN. Trail Making Test A and B: normative data stratified by age and education. *Arch Clin Neuropsychol* 2004; **19**: 203–14.

85. Tombaugh TN, Kozak J, Rees L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Arch Clin Neuropsychol* 1999; **14**: 167–77.
86. Jackson M, Warrington EK. Arithmetic skills in patients with unilateral cerebral lesions. *Cortex* 1986; **22**: 611–20.
87. Bowling A, Bond M, Jenkinson C, Lamping DL. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *J Public Health Med* 1999; **21**: 255–70.
88. Litwin T, Gromadzka G, Członkowska A. Gender differences in Wilson's disease. *J Neurol Sci* 2012; **312**: 31–5.
89. Dening TR, Berrios GE. Wilson's disease: clinical groups in 400 cases. *Acta Neurol Scand* 1989; **80**: 527–34.
90. Denny-Brown D. Hepatolenticular degeneration (Wilson's disease). Two different components. *N Engl J Med* 1964; **270**: 1149–56.
91. Marsden CD. Wilson's disease. *Q J Med* 1987; **65**: 959–66.
92. JM W. Wilson's disease. Handbook of Clinical Neurology: Elsevier 1976. p. 379–414.
93. Moores A, Fox S, Lang A, Hirschfield GM. Wilson disease: Canadian perspectives on presentation and outcomes from an adult ambulatory setting. *Can J Gastroenterol* 2012; **26**: 333–9.
94. Hefter H, Arendt G, Stremmel W, Freund HJ. Motor impairment in Wilson's disease, II: Slowness of speech. *Acta Neurol Scand* 1993; **87**: 148-60.
95. Berry WR, Darley FL, Aronson AE. Dysarthria in Wilson's disease. *J Speech Hear Res* 1974; **17**: 169-83.
96. Burke JF, Dayalu P, Nan B, Askari F, Brewer GJ, Lorincz MT. Prognostic significance of neurologic examination findings in Wilson disease. *Parkinsonism Relat Disord* 2011; **17**: 551–6.
97. Medalia A, Isaacs-Glaberman K, Scheinberg IH. Neuropsychological impairment in Wilson's disease. *Arch Neurol* 1988; **45**: 502–4.
98. Isaacs-Glaberman K, Medalia A, Scheinberg IH. Verbal recall and recognition abilities in patients with Wilson's disease. *Cortex* 1989; **25**: 353–61.

99. Lang C, Müller D, Claus D, Druschky KF. Neuropsychological findings in treated Wilson's disease. *Acta Neurol Scand* 1990; **81**: 75–81.
100. Portala K, Levander S, Westermark K, Ekselius L, von Knorring L. Pattern of neuropsychological deficits in patients with treated Wilson's disease. *Eur Arch Psychiatry Clin Neurosci* 2001; **251**: 262–8.
101. Hegde S, Sinha S, Rao SL, Taly AB, Vasudev MK. Cognitive profile and structural findings in Wilson's disease: a neuropsychological and MRI-based study. *Neurol India* 2010; **58**: 708–13.
102. Ma H, Lv X, Han Y*, et al.* Decision-making impairments in patients with Wilson's disease. *J Clin Exp Neuropsychol* 2013; **35**: 472-9.
103. Iwański S, Seniów J, Leśniak M, Litwin T, Członkowska A. Diverse attention deficits in patients with neurologically symptomatic and asymptomatic Wilson's disease. *Neuropsychology* 2015; **29**: 25–30.
104. Rathbun JK. Neuropsychological aspects of Wilson's disease. *Int J Neurosci* 1996; **85**: 221–9.
105. Wenisch E, De Tassigny A, Trocello JM, Beretti J, Girardot-Tinant N, Woimant F. Cognitive profile in Wilson's disease: a case series of 31 patients. *Rev Neurol (Paris)* 2013; **169**: 944–9.
106. Knehr CA, Bearn AG. Psychological impairment in Wilson's disease. *J Nerv Ment Dis* 1956; **124**: 251–5.
107. Goldstein NP, Ewert JC, Randall RV, Gross JB. Psychiatric aspects of Wilson's disease (hepatolenticular degeneration): results of psychometric tests during long-term therapy. *Am J Psychiatry* 1968; **124**: 1555–61.
108. Tarter RE, Switala J, Carra J, Edwards N, Van Thiel DH. Neuropsychological impairment associated with hepatolenticular degeneration (Wilson's disease) in the absence of overt encephalopathy. *Int J Neurosci* 1987; **37**: 67–71.
109. Littman E, Medalia A, Senior G, Scheinberg IH. Rate of information processing in patients with Wilson's disease. *J Neuropsychiatry Clin Neurosci* 1995; **7**: 68–71.
110. Ma H, Lv X, Han Y*, et al.* Decision-making impairments in patients with Wilson's disease. *J Clin Exp Neuropsychol* 2013; **35**: 472–9.

111. Jacquelet E, Beretti J, De-Tassigny A*, et al.* [Compliance with treatment in Wilson's disease: On the interest of a multidisciplinary closer follow-up]. *Rev Med Interne* 2018; **39**: 155–60.

112. Seniów J, Bak T, Gajda J, Poniatowska R, Czlonkowska A. Cognitive functioning in neurologically symptomatic and asymptomatic forms of Wilson's disease. *Mov Disord* 2002; **17**: 1077-83.

113. Wenisch E, De Tassigny A, Trocello JM, Beretti J, Girardot-Tinant N, Woimant F. Cognitive profile in Wilson's disease: a case series of 31 patients. *Rev Neurol (Paris)* 2013; **169**: 944-9.

114. Han Y, Zhang F, Tian Y, Hu P, Li B, Wang K. Selective impairment of attentional networks of alerting in Wilson's disease. *PLoS One* 2014; **9**: e100454.

115. Rosselli M, Lorenzana P, Rosselli A, Vergara I. Wilson's disease, a reversible dementia: case report. *J Clin Exp Neuropsychol* 1987; **9**: 399–406.

116. Camarata MA, Ala A, Coskun AK*, et al.* Effect of Mental Health, Neurological and Liver Disease on Quality of Life in Patients with Wilson Disease. *J Acad Consult Liaison Psychiatry* 2021.

117. Kirk FT, Munk DE, Laursen TL*, et al.* Cognitive impairment in stable Wilson disease across phenotype. *Metab Brain Dis* 2021.

118. Svetel M, Potrebic A, Pekmezovic T*, et al.* Neuropsychiatric aspects of treated Wilson's disease. *Parkinsonism Relat Disord* 2009; **15**: 772–5.

119. Schaefer M, Gotthardt DN, Ganion N*, et al.* Wilson disease: Health-related quality of life and risk for depression. *Clin Res Hepatol Gastroenterol* 2016; **40**: 349–56.

120. Zimbrean PC, Schilsky ML. Psychiatric aspects of Wilson disease: a review. *Gen Hosp Psychiatry* 2014; **36**: 53–62.

121. Lin JJ, Lin KL, Wang HS, Wong MC. Psychological presentations without hepatic involvement in Wilson disease. *Pediatr Neurol* 2006; **35**: 284–6.

122. Silva F, Nobre S, Campos AP, Vasconcelos M, Gonçalves I. Behavioural and psychiatric disorders in paediatric Wilson's disease. *BMJ Case Rep* 2011; **2011**.

123. Kumawat BL, Sharma CM, Tripathi G, Ralot T, Dixit S. Wilson's disease presenting as isolated obsessive-compulsive disorder. *Indian J Med Sci* 2007; **61**: 607–10.

124. Volpe FM, Tavares A. Cyproterone for hypersexuality in a psychotic patient with Wilson's disease. *Aust N Z J Psychiatry* 2000; **34**: 878–9.

125. Nevsimalova S, Buskova J, Bruha R*, et al.* Sleep disorders in Wilson's disease. *Eur J Neurol* 2011; **18**: 184-90.

126. Dening TR, Berrios GE. Wilson's disease: a longitudinal study of psychiatric symptoms. *Biol Psychiatry* 1990; **28**: 255–65.

127. Portala K, Westermark K, von Knorring L, Ekselius L. Psychopathology in treated Wilson's disease determined by means of CPRS expert and self-ratings. *Acta Psychiatr Scand* 2000; **101**: 104–9.

128. Portala K, Westermark K, Ekselius L, von Knorring L. Personality traits in treated Wilson's disease determined by means of the Karolinska Scales of Personality (KSP). *Eur Psychiatry* 2001; **16**: 362–71.

129. Seniów J, Mroziak B, Członkowska A, Jedryka-Góral A. Self-rated emotional functioning of patients with neurological or asymptomatic form of Wilson's disease. *Clin Neuropsychol* 2003; **17**: 367–73.

130. Lesniak M, Roessler-Gorecka M, Czlonkowska A, Seniow J. Clinical significance of self-descriptive apathy assessment in patients with neurological form of Wilson's disease. *Neurol Sci* 2021.

131. Medalia A, Galynker I, Scheinberg IH. The interaction of motor, memory, and emotional dysfunction in Wilson's disease. *Biol Psychiatry* 1992; **31**: 823–6.

132. Oder W, Prayer L, Grimm G*, et al.* Wilson's disease: evidence of subgroups derived from clinical findings and brain lesions. *Neurology* 1993; **43**: 120–4.

133. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med* 1995; **41**: 1403–9.

134. Komal Kumar RN, Taly AB, Nair KPS*, et al.* Quality of life in Wilson's disease. *Ann Indian Acad Neurol* 2008; **11**: 37–40.

135. Svetel M, Pekmezovic T, Tomic A*, et al.* Quality of life in patients with treated and clinically stable Wilson's disease. *Mov Disord* 2011; **26**: 1503-8.

136. Bruha R, Marecek Z, Pospisilova L*, et al.* Long-term follow-up of Wilson disease: natural history, treatment, mutations analysis and phenotypic correlation. *Liver Int* 2011; **31**: 83–91.

137. Weiss KH, Thurik F, Gotthardt DN*, et al.* Efficacy and safety of oral chelators in treatment of patients with Wilson disease. *Clin Gastroenterol Hepatol* 2013; **11**: 1028–35 e1–2.

138. https://www.nice.org.uk/advice/mib216/chapter/The-technology#costs.

139. Paternostro R, Pfeiffenberger J, Ferenci P*, et al.* Non-invasive diagnosis of cirrhosis and long-term disease monitoring by transient elastography in patients with Wilson disease. *Liver Int* 2020; **40**: 894–904.

140. Prashanth LK, Taly AB, Sinha S*, et al.* Prognostic factors in patients presenting with severe neurological forms of Wilson's disease. *QJM* 2005; **98**: 557–63.

141. Keezer MR, Wolfson C, Postuma RB. Age, Gender, Comorbidity, and the MDS-UPDRS: Results from a Population-Based Study. *Neuroepidemiology* 2016; **46**: 222–7.

142. Schott JM, Crutch SJ, Frost C, Warrington EK, Rossor MN, Fox NC. Neuropsychological correlates of whole brain atrophy in Alzheimer's disease. *Neuropsychologia* 2008; **46**: 1732–7.

143. Gambarota F, Sessa P. Visual Working Memory for Faces and Facial Expressions as a Useful "Tool" for Understanding Social and Affective Cognition. *Front Psychol* 2019; **10**: 2392.

144. Svetel M, Potrebic A, Pekmezovic T*, et al.* Neuropsychiatric aspects of treated Wilson's disease. *Parkinsonism Relat Disord* 2009; **15**: 772-5.

145. Svetel M, Pekmezovic T, Tomic A*, et al.* Quality of life in patients with treated and clinically stable Wilson's disease. *Mov Disord* 2011; **26**: 1503–8.

146. Senda M, Terada S, Takenoshita S*, et al.* Diagnostic utility of the Addenbrooke's Cognitive Examination - III (ACE-III), Mini-ACE, Mini-Mental State Examination, Montreal Cognitive Assessment, and

Hasegawa Dementia Scale-Revised for detecting mild cognitive impairment and dementia. *Psychogeriatrics* 2020; **20**: 156–62.

147. Dziezyc K, Litwin T, Czlonkowska A. Other organ involvement and clinical aspects of Wilson disease. *Handb Clin Neurol* 2017; **142**: 157–69.
148. Woimant F, Djebrani-Oussedik N, Poujois A. New tools for Wilson's disease diagnosis: exchangeable copper fraction. *Ann Transl Med* 2019; **7**: S70.
149. Tang NL, Hui J, Huang D*, et al.* Reference intervals of spot urine copper excretion in preschool children and potential application in pre-symptomatic screening of Wilson disease. *Pathology* 2020; **52**: 439–46.
150. Wang YX, Feng W, Zeng Q*, et al.* Variability of Metal Levels in Spot, First Morning, and 24-Hour Urine Samples over a 3-Month Period in Healthy Adult Chinese Men. *Environ Health Perspect* 2016; **124**: 468–76.
151. Ullah A, Maksud MA, Khan SR, Quraishi SB. Morning (First) Urine Copper Concentration: a New Approach for the Diagnosis of Wilson's Disease. *Biol Trace Elem Res* 2019; **190**: 283–8.
152. Dzieżyc K, Litwin T, Chabik G, Członkowska A. Measurement of urinary copper excretion after 48-h d-penicillamine cessation as a compliance assessment in Wilson's disease. *Funct Neurol* 2015; **30**: 264–8.
153. Socha P, Janczyk W, Dhawan A*, et al.* Wilson's Disease in Children: A Position Paper by the Hepatology Committee of the European Society for Paediatric Gastroenterology, Hepatology and Nutrition. *J Pediatr Gastroenterol Nutr* 2018; **66**: 334–44.
154. Sieniawska CE, Jung LC, Olufadi R, Walker V. Twenty-four-hour urinary trace element excretion: reference intervals and interpretive issues. *Ann Clin Biochem* 2012; **49**: 341–51.
155. Dziezyc K, Litwin T, Chabik G, Czlonkowska A. Measurement of urinary copper excretion after 48-h d-penicillamine cessation as a compliance assessment in Wilson's disease. *Funct Neurol* 2015; **30**: 264–8.
156. Duncan A, Yacoubian C, Beetham R, Catchpole A, Bullock D. The role of calculated non-caeruloplasmin-bound copper in Wilson's disease. *Ann Clin Biochem* 2017; **54**: 649–54.

157. Weisner B, Hartard C, Dieu C. CSF copper concentration: a new parameter for diagnosis and monitoring therapy of Wilson's disease with cerebral manifestation. *J Neurol Sci* 1987; **79**: 229–37.
158. Stuerenburg HJ. CSF copper concentrations, blood-brain barrier function, and coeruloplasmin synthesis during the treatment of Wilson's disease. *J Neural Transm (Vienna)* 2000; **107**: 321–9.
159. El Balkhi S, Poupon J, Trocello JM*, et al.* Determination of ultrafiltrable and exchangeable copper in plasma: stability and reference values in healthy subjects. *Anal Bioanal Chem* 2009; **394**: 1477–84.
160. El Balkhi S, Trocello JM, Poupon J*, et al.* Relative exchangeable copper: a new highly sensitive and highly specific biomarker for Wilson's disease diagnosis. *Clin Chim Acta* 2011; **412**: 2254–60.
161. Trocello JM, El Balkhi S, Woimant F*, et al.* Relative exchangeable copper: a promising tool for family screening in Wilson disease. *Mov Disord* 2014; **29**: 558–62.
162. Guillaud O, Brunet AS, Mallet I*, et al.* Relative exchangeable copper: A valuable tool for the diagnosis of Wilson disease. *Liver Int* 2018; **38**: 350–7.
163. Solovyev N, Ala A, Schilsky M, Mills C, Willis K, Harrington CF. Biomedical copper speciation in relation to Wilson's disease using strong anion exchange chromatography coupled to triple quadrupole inductively coupled plasma mass spectrometry. *Anal Chim Acta* 2020; **1098**: 27–36.
164. Heissat S, Harel A, Um K*, et al.* Evaluation of the accuracy of exchangeable copper and relative exchangeable copper (REC) in a mouse model of Wilson's disease. *J Trace Elem Med Biol* 2018; **50**: 652–7.
165. Zhou XX, Qin HL, Li XH*, et al.* Characterizing brain mineral deposition in patients with Wilson disease using susceptibility-weighted imaging. *Neurol India* 2014; **62**: 362–6.
166. https://www.nice.org.uk/guidance/ng97/chapter/Recommendations#diagnosis.

167. Trojanowski JQ, Schuck T, Schmidt ML, Lee VM. Distribution of tau proteins in the normal human central and peripheral nervous system. *J Histochem Cytochem* 1989; **37**: 209–15.

168. Zetterberg H. Review: Tau in biofluids - relation to pathology, imaging and clinical features. *Neuropathol Appl Neurobiol* 2017; **43**: 194–9.

169. Agoston DV, Shutes-David A, Peskind ER. Biofluid biomarkers of traumatic brain injury. *Brain Inj* 2017; **31**: 1195–203.

170. Gaetani L, Blennow K, Calabresi P, Di Filippo M, Parnetti L, Zetterberg H. Neurofilament light chain as a biomarker in neurological disorders. *J Neurol Neurosurg Psychiatry* 2019; **90**: 870–81.

171. Sun M, Liu N, Xie Q*, et al.* A candidate biomarker of glial fibrillary acidic protein in CSF and blood in differentiating multiple sclerosis and its subtypes: A systematic review and meta-analysis. *Mult Scler Relat Disord* 2021; **51**: 102870.

172. Bishop P, Rocca D, Henley JM. Ubiquitin C-terminal hydrolase L1 (UCH-L1): structure, distribution and roles in brain function and dysfunction. *Biochem J* 2016; **473**: 2453–62.

173. Stukas S, Gill J, Cooper J*, et al.* Characterization of Cerebrospinal Fluid Ubiquitin C-Terminal Hydrolase L1 (UCH-L1) as a Biomarker of Human Acute Traumatic Spinal Cord Injury. *J Neurotrauma* 2021.

174. Li R, Wang J, Xie W, Liu J, Wang C. UCHL1 from serum and CSF is a candidate biomarker for amyotrophic lateral sclerosis. *Ann Clin Transl Neurol* 2020; **7**: 1420–8.

175. Quanterix. The Scientific Principle of Single-Molecule Array Technology. *News-Medical* 2020.

176. Mattsson N, Andreasson U, Zetterberg H, Blennow K, Alzheimer's Disease Neuroimaging I. Association of Plasma Neurofilament Light With Neurodegeneration in Patients With Alzheimer Disease. *JAMA Neurol* 2017; **74**: 557–66.

177. van der Ende EL, Meeter LH, Poos JM*, et al.* Serum neurofilament light chain in genetic frontotemporal dementia: a longitudinal, multicentre cohort study. *Lancet Neurol* 2019; **18**: 1103–11.

178. Scahill RI, Zeun P, Osborne-Crowley K*, et al.* Biological and clinical characteristics of gene carriers far from predicted onset in the

Huntington's disease Young Adult Study (HD-YAS): a cross-sectional analysis. *Lancet Neurol* 2020; **19**: 502–12.

179. Byrne LM, Rodrigues FB, Blennow K*, et al.* Neurofilament light protein in blood as a potential biomarker of neurodegeneration in Huntington's disease: a retrospective cohort analysis. *Lancet Neurol* 2017; **16**: 601–9.
180. Disanto G, Barro C, Benkert P*, et al.* Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. *Ann Neurol* 2017; **81**: 857–70.
181. Haring DA, Kropshofer H, Kappos L*, et al.* Long-term prognostic value of longitudinal measurements of blood neurofilament levels. *Neurol Neuroimmunol Neuroinflamm* 2020; **7**.
182. Kuhle J, Plavina T, Barro C*, et al.* Neurofilament light levels are associated with long-term outcomes in multiple sclerosis. *Mult Scler* 2020; **26**: 1691–9.
183. Ferreira-Atuesta C, Reyes S, Giovanonni G, Gnanapavan S. The Evolution of Neurofilament Light Chain in Multiple Sclerosis. *Front Neurosci* 2021; **15**: 642384.
184. Reyes S, Smets I, Holden D*, et al.* CSF neurofilament light chain testing as an aid to determine treatment strategies in MS. *Neurol Neuroimmunol Neuroinflamm* 2020; **7**.
185. Szilasiova J, Rosenberger J, Fedicova M*, et al.* Neurofilament Light Chain Levels Are Associated with Disease Activity Determined by No Evident Disease Activity in Multiple Sclerosis Patients. *Eur Neurol* 2021: 1–8.
186. Lekomtseva Y, Voloshyn-Gaponov I, Tatayna G. Targeting Higher Levels of Tau Protein in Ukrainian Patients with Wilson's Disease. *Neurol Ther* 2019; **8**: 59–68.
187. Helm S, Baginsky S. MSE for Label-Free Absolute Protein Quantification in Complex Proteomes. *Methods Mol Biol* 2018; **1696**: 235–47.
188. Matthiesen R, Bunkenborg J. Introduction to Mass Spectrometry-Based Proteomics. *Methods Mol Biol* 2020; **2051**: 1–58.
189. Pitt JJ. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev* 2009; **30**: 19–34.

190. Sanders KL, Edwards JL. Nano-liquid chromatography-mass spectrometry and recent applications in omics investigations. *Anal Methods* 2020; **12**: 4404–17.

191. Collins MA, An J, Hood BL, Conrads TP, Bowser RP. Label-Free LC-MS/MS Proteomic Analysis of Cerebrospinal Fluid Identifies Protein/Pathway Alterations and Candidate Biomarkers for Amyotrophic Lateral Sclerosis. *J Proteome Res* 2015; **14**: 4486–501.

192. Heywood WE, Galimberti D, Bliss E*, et al.* Identification of novel CSF biomarkers for neurodegeneration and their validation by a high-throughput multiplexed targeted proteomic assay. *Mol Neurodegener* 2015; **10**: 64.

193. Figura M, Sitkiewicz E, Swiderska B*, et al.* Proteomic Profile of Saliva in Parkinson's Disease Patients: A Proof of Concept Study. *Brain Sci* 2021; **11**.

194. Shribman S, Heller C, Burrows M*, et al.* Plasma Neurofilament Light as a Biomarker of Neurological Involvement in Wilson's Disease. *Mov Disord* 2021; **36**: 503-8.

195. Maselbas W, Chabik G, Czlonkowska A. Persistence with treatment in patients with Wilson disease. *Neurol Neurochir Pol* 2010; **44**: 260–3.

196. Simrén J, Ashton NJ, Blennow K, Zetterberg H. An update on fluid biomarkers for neurodegenerative diseases: recent success and challenges ahead. *Curr Opin Neurobiol* 2020; **61**: 29–39.

197. Zetterberg H, Smith DH, Blennow K. Biomarkers of mild traumatic brain injury in cerebrospinal fluid and blood. *Nat Rev Neurol* 2013; **9**: 201-10.

198. Sparks DL, Kryscio RJ, Sabbagh MN*, et al.* Tau is reduced in AD plasma and validation of employed ELISA methods. *Am J Neurodegener Dis* 2012; **1**: 99–106.

199. Lue LF, Guerra A, Walker DG. Amyloid Beta and Tau as Alzheimer's Disease Blood Biomarkers: Promise From New Technologies. *Neurol Ther* 2017; **6**: 25–36.

200. Shribman S, Heller C, Burrows M*, et al.* Plasma Neurofilament Light as a Biomarker of Neurological Involvement in Wilson's Disease. *Mov Disord* 2021; **36**: 503–8.

201. Lin J, Zheng Y, Liu Y*, et al.* Higher Concentration of Plasma Glial Fibrillary Acidic Protein in Wilson Disease Patients With Neurological Manifestations. *Mov Disord* 2021.

202. Hirata T, Mishra SK, Nakamura S*, et al.* Identification of a Golgi GPI-N-acetylgalactosamine transferase with tandem transmembrane regions in the catalytic domain. *Nat Commun* 2018; **9**: 405.

203. Vuoristo JT, Ala-Kokko L. cDNA cloning, genomic organization and expression of the novel human metallophosphoesterase gene MPPE1 on chromosome 18p11.2. *Cytogenet Cell Genet* 2001; **95**: 60–3.

204. Lohoff FW, Ferraro TN, Brodkin ES, Weller AE, Bloch PJ. Association between polymorphisms in the metallophosphoesterase (MPPE1) gene and bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet* 2010; **153B**: 830–6.

205. Chen X, Xu J, Wang P*, et al.* Metallophosphoesterase 1, a novel candidate gene in hepatocellular carcinoma malignancy and recurrence. *Cancer Biol Ther* 2020; **21**: 1005–13.

206. Fujita M, Maeda Y, Ra M, Yamaguchi Y, Taguchi R, Kinoshita T. GPI glycan remodeling by PGAP5 regulates transport of GPI-anchored proteins from the ER to the Golgi. *Cell* 2009; **139**: 352–65.

207. Nakanishi H, Irie K, Segawa K*, et al.* Crystal structure of a human plasma membrane phospholipid flippase. *J Biol Chem* 2020; **295**: 10180–94.

208. Martin S, Smolders S, Van den Haute C*, et al.* Mutated ATP10B increases Parkinson's disease risk by compromising lysosomal glucosylceramide export. *Acta Neuropathol* 2020; **139**: 1001–24.

209. Heywood WE, Baud A, Bliss E*, et al.* A High Throughput, Multiplexed and Targeted Proteomic CSF Assay to Quantify Neurodegenerative Biomarkers and Apolipoprotein E Isoforms Status. *J Vis Exp* 2016.

210. Viveiros A, Beliveau V, Panzer M*, et al.* Neurodegeneration in Hepatic and Neurologic Wilson Disease. *Hepatology* 2020.

211. Fritzsch D, Reiss-Zimmermann M, Trampel R, Turner R, Hoffmann KT, Schafer A. Seven-tesla magnetic resonance imaging in Wilson disease using quantitative susceptibility mapping for measurement of copper accumulation. *Invest Radiol* 2014; **49**: 299–306.

212. Roberts EA, Schilsky ML, (AASLD). Diagnosis and treatment of Wilson disease: an update. *Hepatology* 2008; **47**: 2089–111.
213. Jenkinson Ma, Chappell Ma. Introduction to neuroimaging analysis. First edition. ed.
214. MacIver CL, Ebden S, Tallantyre EC. MRI: how to understand it. *Pract Neurol* 2021; **21**: 216–24.
215. Elster A. Questions and Answers in MRI. 2015.
216. Liu C, Li W, Tong KA, Yeom KW, Kuzminski S. Susceptibility-weighted imaging and quantitative susceptibility mapping in the brain. *J Magn Reson Imaging* 2015; **42**: 23–41.
217. Desikan RS, Segonne F, Fischl B*, et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 2006; **31**: 968–80.
218. Cardoso MJ, Modat M, Wolz R*, et al.* Geodesic Information Flows: Spatially-Variant Graphs and Their Application to Segmentation and Fusion. *IEEE Trans Med Imaging* 2015; **34**: 1976–88.
219. Pagnozzi AM, Fripp J, Rose SE. Quantifying deep grey matter atrophy using automated segmentation approaches: A systematic review of structural MRI studies. *Neuroimage* 2019; **201**: 116018.
220. Ashburner J, Friston KJ. Voxel-based morphometry--the methods. *Neuroimage* 2000; **11**: 805–21.
221. Andica C, Kamagata K, Hatano T*, et al.* MR Biomarkers of Degenerative Brain Disorders Derived From Diffusion Imaging. *J Magn Reson Imaging* 2020; **52**: 1620–36.
222. Raffelt DA, Tournier JD, Smith RE*, et al.* Investigating white matter fibre density and morphology using fixel-based analysis. *Neuroimage* 2017; **144**: 58–73.
223. Smith SM, Jenkinson M, Johansen-Berg H*, et al.* Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 2006; **31**: 1487–505.
224. Ashton EA, Takahashi C, Berg MJ, Goodman A, Totterman S, Ekholm S. Accuracy and reproducibility of manual and semiautomated quantification of MS lesions by MRI. *J Magn Reson Imaging* 2003; **17**: 300–8.

225. Balakrishnan R, Valdes Hernandez MDC, Farrall AJ. Automatic segmentation of white matter hyperintensities from brain magnetic resonance images in the era of deep learning and big data - A systematic review. *Comput Med Imaging Graph* 2021; **88**: 101867.

226. Wang Y, Liu T. Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. *Magn Reson Med* 2015; **73**: 82–101.

227. Acosta-Cabronero J, Milovic C, Mattern H, Tejos C, Speck O, Callaghan MF. A robust multi-scale approach to quantitative susceptibility mapping. *Neuroimage* 2018; **183**: 7–24.

228. Langkammer C, Schweser F, Shmueli K*, et al.* Quantitative susceptibility mapping: Report from the 2016 reconstruction challenge. *Magn Reson Med* 2018; **79**: 1661–73.

229. Acosta-Cabronero J, Betts MJ, Cardenas-Blanco A, Yang S, Nestor PJ. In Vivo MRI Mapping of Brain Iron Deposition across the Adult Lifespan. *J Neurosci* 2016; **36**: 364–74.

230. Sinha S, Taly AB, Ravishankar S*, et al.* Wilson's disease: cranial MRI observations and clinical correlation. *Neuroradiology* 2006; **48**: 613–21.

231. Favrole P, Chabriat H, Guichard JP, Woimant F. Clinical correlates of cerebral water diffusion in Wilson disease. *Neurology* 2006; **66**: 384-9.

232. Dusek P, Smolinski L, Redzia-Ogrodnik B*, et al.* Semiquantitative Scale for Assessing Brain MRI Abnormalities in Wilson Disease: A Validation Study. *Mov Disord* 2020; **35**: 994–1001.

233. Zou L, Song Y, Zhou X, Chu J, Tang X. Regional morphometric abnormalities and clinical relevance in Wilson's disease. *Mov Disord* 2019; **34**: 545–54.

234. Hu S, Xu C, Dong T*, et al.* Structural and Functional Changes Are Related to Cognitive Status in Wilson's Disease. *Front Hum Neurosci* 2021; **15**: 610947.

235. Dusek P, Lescinskij A, Ruzicka F*, et al.* Associations of Brain Atrophy and Cerebral Iron Accumulation at MRI with Clinical Severity in Wilson Disease. *Radiology* 2021: 202846.

236. Yuan XZ, Li GY, Chen JL, Li JQ, Wang XP. Paramagnetic Metal Accumulation in the Deep Gray Matter Nuclei Is Associated With

Neurodegeneration in Wilson's Disease. *Front Neurosci* 2020; **14**: 573633.

237. Song Y, Zou L, Zhao J*, et al.* Whole brain volume and cortical thickness abnormalities in Wilson's disease: a clinical correlation study. *Brain Imaging Behav* 2020.

238. Tinaz S, Arora J, Nalamada K*, et al.* Structural and functional brain changes in hepatic and neurological Wilson disease. *Brain Imaging Behav* 2020.

239. Wu Y, Hu S, Wang Y*, et al.* The degeneration changes of basal forebrain are associated with prospective memory impairment in patients with Wilson's disease. *Brain Behav* 2021.

240. Smolinski L, Litwin T, Redzia-Ogrodnik B, Dziezyc K, Kurkowska-Jastrzebska I, Czlonkowska A. Brain volume is related to neurological impairment and to copper overload in Wilson's disease. *Neurol Sci* 2019; **40**: 2089-95.

241. Pulai S, Biswas A, Roy A*, et al.* Clinical features, MRI brain, and MRS abnormalities of drug-naive neurologic Wilson's disease. *Neurol India* 2014; **62**: 153–8.

242. King AD, Walshe JM, Kendall BE*, et al.* Cranial MR imaging in Wilson's disease. *AJR Am J Roentgenol* 1996; **167**: 1579-84.

243. Starosta-Rubinstein S, Young AB, Kluin K*, et al.* Clinical assessment of 31 patients with Wilson's disease. Correlations with structural changes on magnetic resonance imaging. *Arch Neurol* 1987; **44**: 365–70.

244. Sinha S, Taly AB, Ravishankar S*, et al.* Wilson's disease: cranial MRI observations and clinical correlation. *Neuroradiology* 2006; **48**: 613-21.

245. Pulai S, Biswas A, Roy A*, et al.* Clinical features, MRI brain, and MRS abnormalities of drug-naive neurologic Wilson's disease. *Neurol India* 2014; **62**: 153-8.

246. Thuomas KA, Aquilonius SM, Bergstrom K, Westermark K. Magnetic resonance imaging of the brain in Wilson's disease. *Neuroradiology* 1993; **35**: 134–41.

247. Roh JK, Lee TG, Wie BA, Lee SB, Park SH, Chang KH. Initial and follow-up brain MRI findings and correlation with the clinical course in Wilson's disease. *Neurology* 1994; **44**: 1064–8.

248. Sinha S, Taly AB, Prashanth LK, Ravishankar S, Arunodaya GR, Vasudev MK. Sequential MRI changes in Wilson's disease with de-coppering therapy: a study of 50 patients. *Br J Radiol* 2007; **80**: 744–9.

249. da Costa Mdo D, Spitz M, Bacheschi LA, Leite CC, Lucato LT, Barbosa ER. Wilson's disease: two treatment modalities. Correlations to pretreatment and posttreatment brain MRI. *Neuroradiology* 2009; **51**: 627–33.

250. Nazer H, Brismar J, al-Kawi MZ, Gunasekaran TS, Jorulf KH. Magnetic resonance imaging of the brain in Wilson's disease. *Neuroradiology* 1993; **35**: 130–3.

251. Kim TJ, Kim IO, Kim WS*, et al.* MR imaging of the brain in Wilson disease of childhood: findings before and after treatment with clinical correlation. *AJNR Am J Neuroradiol* 2006; **27**: 1373–8.

252. Costa MdDLd, Spitz M, Bacheschi LA. Wilson's disease: two treatment modalities. Correlations to pretreatment and posttreatment brain MRI. Neuroradiology; 2009. p. 627–33.

253. Jadav R, Saini J, Sinha S, Bagepally B, Rao S, Taly AB. Diffusion tensor imaging (DTI) and its clinical correlates in drug naive Wilson's disease. *Metab Brain Dis* 2013; **28**: 455-62.

254. Li G, Zhou X, Xu P, Pan X, Chen Y. Microstructure assessment of the thalamus in Wilson's disease using diffusion tensor imaging. *Clin Radiol* 2014; **69**: 294–8.

255. Zhou XX, Li XH, Qin H*, et al.* Diffusion tensor imaging of the extracorticospinal network in the brains of patients with Wilson disease. *J Neurol Sci* 2016; **362**: 292–8.

256. Song YK, Li XB, Huang XL*, et al.* A study of neurite orientation dispersion and density imaging in wilson's disease. *J Magn Reson Imaging* 2018; **48**: 423–30.

257. Sener RN. Diffusion MR imaging changes associated with Wilson disease. *AJNR Am J Neuroradiol* 2003; **24**: 965–7.

258. Xue P, Liu Y, Ma XH, Chen Y, Zhang SJ. Role of diffusion weighted magnetic resonance imaging and spectroscopy in the diagnosis and follow-up of hepatolenticular degeneration. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao* 2012; **34**: 497–502.

259. Dezortova M, Lescinskij A, Dusek P*, et al.* Multiparametric Quantitative Brain MRI in Neurological and Hepatic Forms of Wilson's Disease. *J Magn Reson Imaging* 2020; **51**: 1829–35.

260. Li G, Wu R, Tong R*, et al.* Quantitative Measurement of Metal Accumulation in Brain of Patients With Wilson's Disease. *Mov Disord* 2020; **35**: 1787–95.

261. Doganay S, Gumus K, Koc G*, et al.* Magnetic Susceptibility Changes in the Basal Ganglia and Brain Stem of Patients with Wilson's Disease: Evaluation with Quantitative Susceptibility Mapping. *Magn Reson Med Sci* 2017.

262. Saracoglu S, Gumus K, Doganay S*, et al.* Brain susceptibility changes in neurologically asymptomatic pediatric patients with Wilson's disease: evaluation with quantitative susceptibility mapping. *Acta Radiol* 2018; **59**: 1380–5.

263. Bai X, Wang G, Wu L*, et al.* Deep-gray nuclei susceptibility-weighted imaging filtered phase shift in patients with Wilson's disease. *Pediatr Res* 2014; **75**: 436–42.

264. Zhou XX, Qin HL, Li XH*, et al.* Characterizing brain mineral deposition in patients with Wilson disease using susceptibility-weighted imaging. *Neurol India* 2014; **62**: 362-6.

265. Zaino D, Chiarotti I, Battisti C, Salvatore S, Federico A, Cerase A. Six-year clinical and MRI quantitative susceptibility mapping (QSM) follow-up in neurological Wilson's disease under zinc therapy: a case report. *Neurol Sci* 2019; **40**: 199–201.

266. Dusek P, Bahn E, Litwin T*, et al.* Brain iron accumulation in Wilson disease: a post mortem 7 Tesla MRI - histopathological study. *Neuropathol Appl Neurobiol* 2017; **43**: 514–32.

267. Lee JH, Yang TI, Cho M, Yoon KT, Baik SK, Han YH. Widespread cerebral cortical mineralization in Wilson's disease detected by susceptibility-weighted imaging. *J Neurol Sci* 2012; **313**: 54–6.

268. Frota NA, Barbosa ER, Porto CS*, et al.* Cognitive impairment and magnetic resonance imaging correlations in Wilson's disease. *Acta Neurol Scand* 2013; **127**: 391-8.

269. Dong T, Yang WM, Wu MC*, et al.* Microstructure changes in whiter matter relate to cognitive impairment in Wilson's disease. *Biosci Rep* 2019; **39**.

270. Eggers B, Hermann W, Barthel H, Sabri O, Wagner A, Hesse S. The degree of depression in Hamilton rating scale is correlated with the density of presynaptic serotonin transporters in 23 patients with Wilson's disease. *J Neurol* 2003; **250**: 576-80.

271. Ashburner J, Friston KJ. Voxel-based morphometry--the methods. *Neuroimage* 2000; **11**: 805-21.

272. Smith SM, Jenkinson M, Johansen-Berg H*, et al.* Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 2006; **31**: 1487-505.

273. Sudre CH, Cardoso MJ, Bouvy WH, Biessels GJ, Barnes J, Ourselin S. Bayesian model selection for pathological neuroimaging data applied to white matter lesion segmentation. *IEEE Trans Med Imaging* 2015; **34**: 2079-102.

274. Acosta-Cabronero J, Betts MJ, Cardenas-Blanco A, Yang S, Nestor PJ. In Vivo MRI Mapping of Brain Iron Deposition across the Adult Lifespan. *J Neurosci* 2016; **36**: 364-74.

275. Shribman S, Bocchetta M, Sudre CH*, et al.* Neuroimaging correlates of brain injury in Wilson's disease: a multimodal, whole-brain MRI study. *Brain* 2021.

276. Bocchetta M, Iglesias JE, Chelban V*, et al.* Automated Brainstem Segmentation Detects Differential Involvement in Atypical Parkinsonian Syndromes. *J Mov Disord* 2020; **13**: 39–46.

277. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage* 2007; **38**: 95–113.

278. Sudre CH, Cardoso MJ, Bouvy WH, Biessels GJ, Barnes J, Ourselin S. Bayesian model selection for pathological neuroimaging data applied to white matter lesion segmentation. *IEEE Trans Med Imaging* 2015; **34**: 2079–102.

279. Sudre CH, Gomez Anson B, Davagnanam I*, et al.* Bullseye's representation of cerebral white matter hyperintensities. *J Neuroradiol* 2018; **45**: 114–22.

280. Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 2009; **44**: 83–98.

281. Acosta-Cabronero J, Cardenas-Blanco A, Betts MJ*, et al.* The whole-brain pattern of magnetic susceptibility perturbations in Parkinson's disease. *Brain* 2017; **140**: 118–31.

282. Zou L, Song Y, Zhou X, Chu J, Tang X. Regional morphometric abnormalities and clinical relevance in Wilson's disease. *Mov Disord* 2019; **34**: 545-54.

283. Mansoor NM, Vanniyasingam T, Malone I*, et al.* Validating Automated Segmentation Tools in the Assessment of Caudate Atrophy in Huntington's Disease. *Front Neurol* 2021; **12**: 616272.

284. Dusek P, Smolinski L, Redzia-Ogrodnik B*, et al.* Semiquantitative Scale for Assessing Brain MRI Abnormalities in Wilson Disease: A Validation Study. *Mov Disord* 2020; **35**: 994-1001.

285. Walter U, Krolikowski K, Tarnacka B, Benecke R, Czlonkowska A, Dressler D. Sonographic detection of basal ganglia lesions in asymptomatic and symptomatic Wilson disease. *Neurology* 2005; **64**: 1726–32.

286. Svetel M, Mijajlović M, Tomić A, Kresojević N, Pekmezović T, Kostić VS. Transcranial sonography in Wilson's disease. *Parkinsonism Relat Disord* 2012; **18**: 234-8.

287. Klaming R, Annese J. Functional anatomy of essential tremor: lessons from neuroimaging. *AJNR Am J Neuroradiol* 2014; **35**: 1450–7.

288. Patell R, Dosi R, Joshi HK, Storz D. Atypical neuroimaging in Wilson's disease. *BMJ Case Rep* 2014; **2014**.

289. Miyakawa T, Murayama E. An autopsy case of the "demyelinating type" of Wilson's disease. *Acta Neuropathol* 1976; **35**: 235–41.

290. Ma KC, Ye ZR, Fang J, Wu JV. Glial fibrillary acidic protein immunohistochemical study of Alzheimer I & II astrogliosis in Wilson's disease. *Acta Neurol Scand* 1988; **78**: 290–6.

291. Kinnier Wilson SA. Progressive lenticular degeneration: a familial nervous disease associated with cirrhosis of the liver. *Brain* 1912; **34**: 295–509.

292. Song SK, Sun SW, Ju WK, Lin SJ, Cross AH, Neufeld AH. Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *Neuroimage* 2003; **20**: 1714–22.

293. Song YK, Li XB, Huang XL*, et al.* A study of neurite orientation dispersion and density imaging in wilson's disease. *J Magn Reson Imaging* 2018; **48**: 423-30.

294. Beaulieu C. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed* 2002; **15**: 435–55.

295. Acosta-Cabronero J, Williams GB, Pengas G, Nestor PJ. Absolute diffusivities define the landscape of white matter degeneration in Alzheimer's disease. *Brain* 2010; **133**: 529–39.

296. Fox RJ, Coffey CS, Conwit R*, et al.* Phase 2 Trial of Ibudilast in Progressive Multiple Sclerosis. *N Engl J Med* 2018; **379**: 846-55.

297. Green AJ, Gelfand JM, Cree BA*, et al.* Clemastine fumarate as a remyelinating therapy for multiple sclerosis (ReBUILD): a randomised, controlled, double-blind, crossover trial. *Lancet* 2017; **390**: 2481-9.

298. Staffaroni AM, Ljubenkov PA, Kornak J*, et al.* Longitudinal multimodal imaging and clinical endpoints for frontotemporal dementia clinical trials. *Brain* 2019; **142**: 443-59.

299. Betts MJ, Acosta-Cabronero J, Cardenas-Blanco A, Nestor PJ, Düzel E. High-resolution characterisation of the aging brain using simultaneous quantitative susceptibility mapping (QSM) and R2* measurements at 7T. *Neuroimage* 2016; **138**: 43-63.

300. Hametner S, Endmayr V, Deistung A*, et al.* The influence of brain iron and myelin on magnetic susceptibility and effective transverse relaxation - A biochemical and histological validation study. *Neuroimage* 2018; **179**: 117–33.

301. Dusek P, Bahn E, Litwin T*, et al.* Brain iron accumulation in Wilson disease: a post mortem 7 Tesla MRI - histopathological study. *Neuropathol Appl Neurobiol* 2017; **43**: 514-32.

302. Barnes N, Tsivkovskii R, Tsivkovskaia N, Lutsenko S. The copper-transporting ATPases, menkes and wilson disease proteins, have distinct

roles in adult and developing cerebellum. *J Biol Chem* 2005; **280**: 9640–5.

303. Kodama H, Okabe I, Yanagisawa M*, et al.* Does CSF copper level in Wilson disease reflect copper accumulation in the brain? *Pediatr Neurol* 1988; **4**: 35-7.

304. Dijkstra N, de Bruin L. Cognitive Neuroscience and Causal Inference: Implications for Psychiatry. *Front Psychiatry* 2016; **7**: 129.

305. Allen MD, Fong AK. Clinical application of standardized cognitive assessment using fMRI. I. Matrix reasoning. *Behav Neurol* 2008; **20**: 127-40.

306. Allen MD, Fong AK. Clinical application of standardized cognitive assessment using fMRI. II. Verbal fluency. *Behav Neurol* 2008; **20**: 141-52.

307. Jobson DD, Hase Y, Clarkson AN, Kalaria RN. The role of the medial prefrontal cortex in cognition, ageing and dementia. *Brain Commun* 2021; **3**: fcab125.

308. Kim JJ, Andreasen NC, O'Leary DS*, et al.* Direct comparison of the neural substrates of recognition memory for words and faces. *Brain* 1999; **122 (Pt 6)**: 1069-83.

309. Cohen AL, Soussand L, Corrow SL, Martinaud O, Barton JJS, Fox MD. Looking beyond the face area: lesion network mapping of prosopagnosia. *Brain* 2019; **142**: 3975-90.

310. Mier D, Lis S, Neuthe K*, et al.* The involvement of emotion recognition in affective theory of mind. *Psychophysiology* 2010; **47**: 1028-39.

311. Killgore WD, Yurgelun-Todd DA. Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces. *Neuroimage* 2004; **21**: 1215-23.

312. Tsuchida A, Fellows LK. Are you upset? Distinct roles for orbitofrontal and lateral prefrontal cortex in detecting and distinguishing facial expressions of emotion. *Cereb Cortex* 2012; **22**: 2904-12.

313. Wolf RC, Pujara M, Baskaya MK, Koenigs M. Emotion recognition deficits associated with ventromedial prefrontal cortex lesions are improved by gaze manipulation. *Cortex* 2016; **82**: 255-62.

314. Russell LL, Greaves CV, Bocchetta M*, et al.* Social cognition impairment in genetic frontotemporal dementia within the GENFI cohort. *Cortex* 2020; **133**: 384-98.
315. Dajani DR, Uddin LQ. Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends Neurosci* 2015; **38**: 571-8.
316. Allen MD, Owens TE, Fong AK, Richards DR. A functional neuroimaging analysis of the Trail Making Test-B: implications for clinical application. *Behav Neurol* 2011; **24**: 159-71.
317. Brignoni-Perez E, Matejko AA, Jamal NI, Eden GF. Functional neuroanatomy of arithmetic in monolingual and bilingual adults and children. *Hum Brain Mapp* 2021.
318. Leech R, Sharp DJ. The role of the posterior cingulate cortex in cognition and disease. *Brain* 2014; **137**: 12-32.
319. Chen Y, Dang M, Zhang Z. Brain mechanisms underlying neuropsychiatric symptoms in Alzheimer's disease: a systematic review of symptom-general and -specific lesion patterns. *Mol Neurodegener* 2021; **16**: 38.
320. Vogt BA. Cingulate cortex in the three limbic subsystems. *Handb Clin Neurol* 2019; **166**: 39-51.
321. Serra L, Perri R, Cercignani M*, et al.* Are the behavioral symptoms of Alzheimer's disease directly associated with neurodegeneration? *J Alzheimers Dis* 2010; **21**: 627-39.
322. Ahmed RM, Halliday G, Hodges JR. Hypothalamic symptoms of frontotemporal dementia disorders. *Handb Clin Neurol* 2021; **182**: 269-80.
323. Mechelli A, Price CJ, Friston KJ, Ashburner J. Voxel-Based Morphometry of the Human Brain: Methods and Applications. *Curr Med Imaging Rev* 2005; **1**.
324. Gitelman DR, Ashburner J, Friston KJ, Tyler LK, Price CJ. Voxel-based morphometry of herpes simplex encephalitis. *Neuroimage* 2001; **13**: 623-31.
325. Havsteen I, Ohlhues A, Madsen KH, Nybing JD, Christensen H, Christensen A. Are Movement Artifacts in Magnetic Resonance Imaging a Real Problem?-A Narrative Review. *Front Neurol* 2017; **8**: 232.

326. Dezortova M, Lescinskij A, Dusek P*, et al.* Multiparametric Quantitative Brain MRI in Neurological and Hepatic Forms of Wilson's Disease. *J Magn Reson Imaging* 2020; **51**: 1829-35.

327. Tarter RE, Switala J, Carra J, Edwards N, Van Thiel DH. Neuropsychological impairment associated with hepatolenticular degeneration (Wilson's disease) in the absence of overt encephalopathy. *Int J Neurosci* 1987; **37**: 67-71.

328. Taly AB, Meenakshi-Sundaram S, Sinha S, Swamy HS, Arunodaya GR. Wilson disease: description of 282 patients evaluated over 3 decades. *Medicine (Baltimore)* 2007; **86**: 112-21.

329. Zimbrean PC, Schilsky ML. Psychiatric aspects of Wilson disease: a review. *Gen Hosp Psychiatry* 2014; **36**: 53-62.

330. Ferenci P, Stremmel W, Członkowska A*, et al.* Age and Sex but Not ATP7B Genotype Effectively Influence the Clinical Phenotype of Wilson Disease. *Hepatology* 2019; **69**: 1464-76.

331. Kluska A, Kulecka M, Litwin T*, et al.* Whole-exome sequencing identifies novel pathogenic variants across the ATP7B gene and some modifiers of Wilson's disease phenotype. *Liver Int* 2019; **39**: 177-86.

Appendix

Table 0-1 Leipzig criteria for the diagnosis of WD.

Score	-1	0	1	2	4
Neurologic or psychiatric features (including typical MRI abnormalities)		Absent		Present	
KF rings (slit lamp examination)		Absent		Present	
Coombs negative hemolytic anaemia		Absent		Present	
Urine copper output		Normal	1-2x ULN	>2x ULN	
> after D-penicillamine challenge				>5x ULN	
Liver copper content	Normal		1-5x ULN	>5x ULN	
Rhodanine-stained hepatocytes		Absent	Present		
Serum ceruloplasmin (mg/dL)		Normal	10-20	<10	
Pathogenic ATP7B mutations		None	One		Two

A total score of four or more indicates WD is highly likely, two to three indicates WD is probable (and more investigations required) and a score of zero to one indicates WD is unlikely.

Table 0-2 UWDRS-N scoring system

Item	Score
10. Speech	0 - Normal
	1 - Slight dysarthria or slight loss of expression, diction and/or volume
	2 - Moderate dysarthria or monotone, slurred speech. Still understandable
	3 - Marked impairment, difficult to understand
	4 - Unintelligible
11A. Facial expression - Oromandibular dystonia If >2, skip item 11B	0 - No dystonia present
	1 - Slight. Occasional grimacing or other mouth movements
	2 - Mild. Movement present less than 50% of the time
	3 - Moderate dystonic movements or contractions present most of the time
	4 - Severe dystonic movements or contractions present most of the time
11B. Facial expression - Hypomimia	0 - Normal
	1 - Minimal hypomimia, could be normal “Poker Face”
	2 - Slight but definitely abnormal diminution of facial expression
	3 - Moderate hypomimia; lips parted some of the time
	4 - Masked or fixed facies with severe or complete loss of facial expression
12. Tremor at rest Scores for RUL, LUL, RLL and LLL	0 - Absent
	1 - Slight and infrequently present
	2 - Mild in amplitude and persistent. Or moderate but intermittent
	3 - Moderate in amplitude and present most of the time
	4 - Marked in amplitude and present most of the time
13. Head tremor	0 - None
	1 - Slight or hardly perceptible tremor. May be intermittent
	2 - Moderate amplitude (< 2 cm). May be intermittent
	3 - Marked amplitude (2-4 cm)
	4 - Severe amplitude (> 4 cm)
14. Rigidity Scores for Neck, RUL, LUL, RLL and LLL	0 - Absent
	1 - Slight or detectable only when activated by mirror or other movements
	2 - Mild to moderate
	3 - Marked, but full range of motion easily achieved
	4 - Severe, range of motion achieved with difficulty
15. Finger taps Score for RUL and LUL	0 - Normal
	1 - Mild impairment
	2 - Moderate impairment
	3 - Severe impairment
	4 - Cannot perform the task
16. Rapid alternating movements of hands Score for RUL and LUL	0 - Normal
	1 - Mild impairment
	2 - Moderate impairment
	3 - Severe impairment
	4 - Cannot perform the task

17. Handwriting	0 - Normal
	1 - Slightly impaired
	2 - Moderately impaired; all words are legible
	3 - Severely impaired; few words are legible
	4 - Cannot hold a pen
18A. Tremor in arms - Postural Score for RUL and LUL	0 - None
	1 - Slight or hardly perceptible postural tremor. May be intermittent
	2 - Moderate amplitude (< 2 cm). May be intermittent
	3 - Marked amplitude (2-4 cm)
	4 - Severe amplitude (> 4 cm)
18B. Tremor in arms - Wing-beating Score for RUL and LUL	0 - None
	1 - Slight or hardly perceptible wing-beating tremor. May be intermittent
	2 - Moderate amplitude (< 2 cm). May be intermittent
	3 - Marked amplitude (2-4 cm)
	4 - Severe amplitude (> 4 cm)
19. Finger-to-nose-test Score for RUL and LUL	0 - Normal
	1 - Mild impairment
	2 - Moderate impairment
	3 - Severe impairment
	4 - Cannot perform the task
20. Leg agility Score for RLL and LLL	0 - Normal
	1 - Mild impairment
	2 - Moderate impairment
	3 - Severe impairment
	4 - Cannot perform the task
21. Postural tremor in legs Score for RLL and LLL	0 - None
	1 - Slight or hardly perceptible tremor. May be intermittent
	2 - Moderate amplitude (< 2 cm). May be intermittent
	3 - Marked amplitude (2-4 cm)
	4 - Severe amplitude (> 4 cm)
22. Cervical dystonia	0 - No dystonia present
	1 - Slight. Occasional pulling
	2 - Obvious torticollis, but mild
	3 - Moderate pulling
	4 - Extreme pulling
23. Arm and hand dystonia Score for RUL and LUL	0 - No dystonia present
	1 - Slight dystonia. Clinically insignificant
	2 - Mild. Obvious dystonia, but not disabling
	3 - Moderate. Able to grasp, with some manual function
	4 - Severe. No useful grasp

24. Arising from chair	0 - Normal
	1 - Slow; or may need more than one attempt
	2 - Needs arms of seat as support
	3 - Tends to falls back and may have to try several times
	4 - Unable to arise without help
25A. Posture - Trunk dystonia If >2, skip items 25B and 25C	0 - No dystonia present
	1 - Slight bending; clinically insignificant
	2 - Definite bending, but not interfering with standing
	3 - Moderate bending; interfering with standing
	4 - Extreme bending of trunk preventing standing
25B. Posture - Ataxia of stance	0 - Absent
	1 - Slight (swaying only present without visual feedback)
	2 - Moderate (moderate swaying; still able to stand with feet together)
	3 - Marked (marked swaying; unable to stand with feet together)
	4 - Severe to most severe (unable to stand without support or bedridden)
25C. Posture Parkinsonism	0 - Normal erect
	1 - Not quite erect, slightly stooped posture; could be normal for older person
	2 - Moderately stooped posture, definitely abnormal; can be slightly leaning
	3 - Severely stooped posture with kyphosis; can be moderately leaning
	4 - Marked flexion with extreme abnormality of posture
26A. Gait - Leg dystonia Score for RLL and LLL If >2, skip items 26B and 26C	0 - No dystonia present
	1 - Slight dystonia, but not causing impairment; clinically insignificant
	2 - Mild dystonia. Walks briskly and unaided
	3 - Moderate dystonia. Severely impairs walking or requires assistance
	4 - Severe. Unable to walk on involved leg
26B. Gait - Ataxia	0 - Absent
	1 - Slight (ataxia only visible, when walking on tandem or without visual feedback)
	2 - Moderate (ataxia visible in normal walking; difficulties when walking on tandem)
	3 - Marked (broad-based, staggering gait; unable to walk on tandem)
	4 - Severe (unable to walk without support, wheelchair-bound, bedridden)
26C. Gait - Parkinsonism	0 - Normal
	1 - Walks slowly, may shuffle with short steps, but no festination or propulsion
	2 - Walks with difficulty, requires little or no assistance; may have festination, …
	3 - Severe disturbance of gait, requiring assistance
	4 - Cannot walk at all, even with assistance.
27C. Chorea Score for face, trunk, RUL, LUL, RLL and LLL	0 - Absent
	1 - Slight/ intermittent
	2 - Mild/ common or moderate/ intermittent
	3 - Moderate/ common
	4 - Marked/ prolonged

RUL – right upper limb; LUL – left upper limb; RLL - right lower limb; LLL – left lower limb.

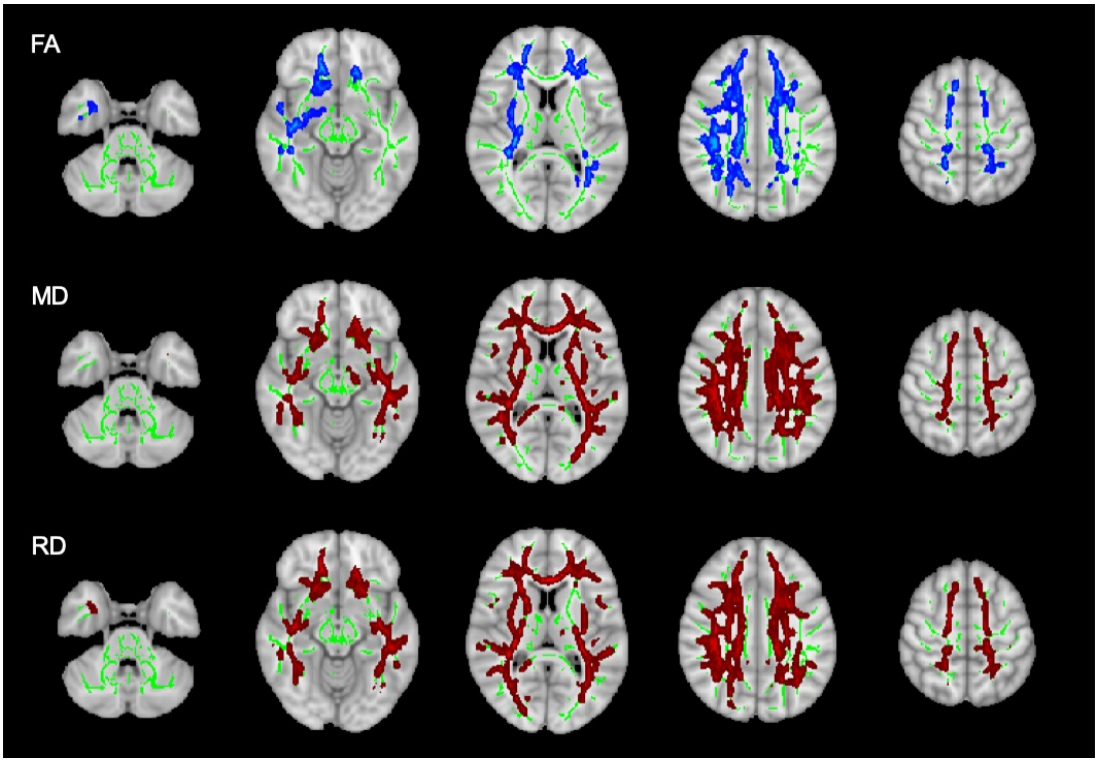
Table 0-3 Statistics for clusters identified across all VBM analyses

Analysis	Size	TFCE	P_{FWE}	P_{uncorr}	x	y	z
Presentation	2952	213390	0.009	<0.001	22	15	-10
	2180	145797	0.013	<0.001	-10	16	9
	128	53891	0.028	0.001	-32	32	-4
	21	44711	0.037	0.002	-56	-6	6
UWDRS-N	6808	2443	0.005	0.001	-27	10	-8
	3420	2040	0.011	0.001	21	20	-9
MRT	2073	2031	0.012	0.001	30	4	-12
	23	1444	0.047	0.001	-38	-44	-44
RMTF	22190	2202	0.008	<0.001	9	-6	56
	1170	1598	0.033	0.001	56	-32	-2
	374	1499	0.041	0.001	9	-56	-8
	585	1499	0.041	0.001	0	-39	-18
	215	1479	0.043	<0.001	-24	32	36
	141	1469	0.044	<0.001	21	57	4
	179	1464	0.045	0.002	44	-8	3
	24	1427	0.049	0.001	32	-18	6
Animals	601	1806	0.033	0.001	-21	-56	-60
TMTB	333	1550	0.042	0.001	-33	-22	9
	140	1527	0.044	0.001	-9	-18	60
	59	1500	0.047	0.001	-3	-66	21
	190	1499	0.047	0.001	4	-78	10
TMTB with UWDRS-N	76	1523	0.044	0.001	-22	-88	-15
	84	1482	0.048	0.001	6	-80	9
	34	1482	0.048	0.001	-3	-9	54
Ekman	5493	2008	0.012	<0.001	9	22	39
	8096	1878	0.016	<0.001	2	-74	-36
	9604	1807	0.019	<0.001	62	8	16
	734	1655	0.027	<0.001	-52	-15	45
	600	1528	0.037	<0.001	32	27	38
	763	1489	0.041	0.002	-34	-75	-34
	329	1470	0.043	<0.001	52	-16	44
	200	1457	0.044	<0.001	50	9	42
	103	1450	0.045	0.001	32	-2	56
	36	1440	0.046	<0.001	-51	8	27
	73	1427	0.048	0.001	-42	-6	2
	28	1410	0.050	<0.001	-63	-8	24
Ekman with UWDRS-N	1961	1673	0.025	<0.001	9	22	39
	168	1452	0.044	<0.001	60	8	16
Anxiety (ModNPI)	358	1522	0.036	<0.001	-4	-42	30
Social/emotional cues (ModNPI)	15206	2248	0.006	0.001	16	-69	-10
	782	1570	0.030	0.001	-3	-68	-30
	323	1551	0.031	0.001	-30	-15	10
	363	1465	0.038	<0.001	-18	-86	42
	65	1412	0.044	<0.001	24	-54	57
	281	1408	0.044	<0.001	0	-39	40
	173	1401	0.045	<0.001	-16	-57	58
	56	1371	0.048	0.001	-18	-46	-45
	33	1364	0.049	<0.001	39	-82	-14
NCC	240	1500	0.039	0.001	0	-63	14
	247	1477	0.041	0.001	-8	-26	70
	118	1443	0.045	<0.001	27	-60	48
Tau	1375	1770	0.02	<0.001	12	9	-3
	1292	1619	0.029	0.001	-12	8	-3

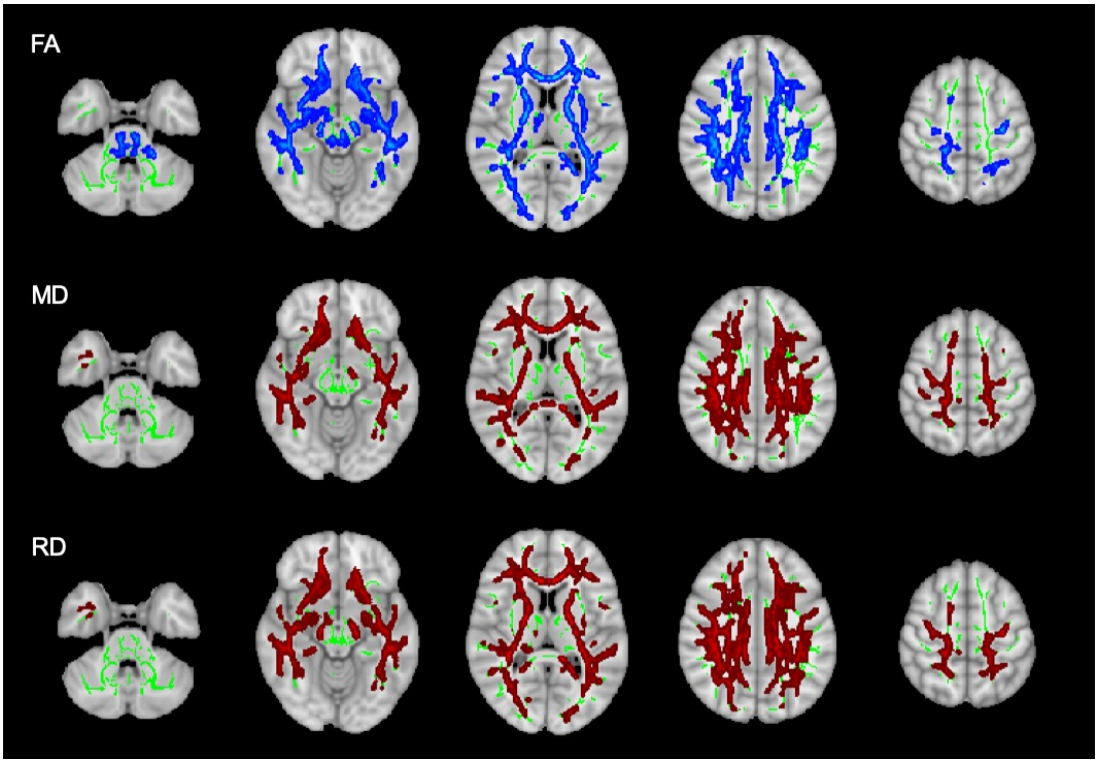
Animals, semantic fluency test; Ekman, Ekman 35-faces test; FAS, phonemic fluency test; GDA, graded difficulty arithmetic; MRT, matrix reasoning test; NCC, non-caeruloplasmin-bound copper; RMTF, recognition memory test for faces; TMTB, trail making test part B; UWDRS-N, unified Wilson's disease rating scale neurological examination subscore; VBM, voxel-based morphometry.

Figure 0-1 TBSS for associations with cognitive test scores with UWDRS-N as covariate

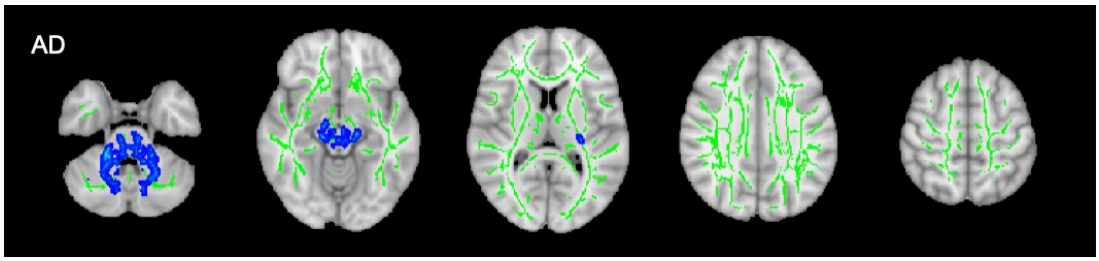
A) RMTF



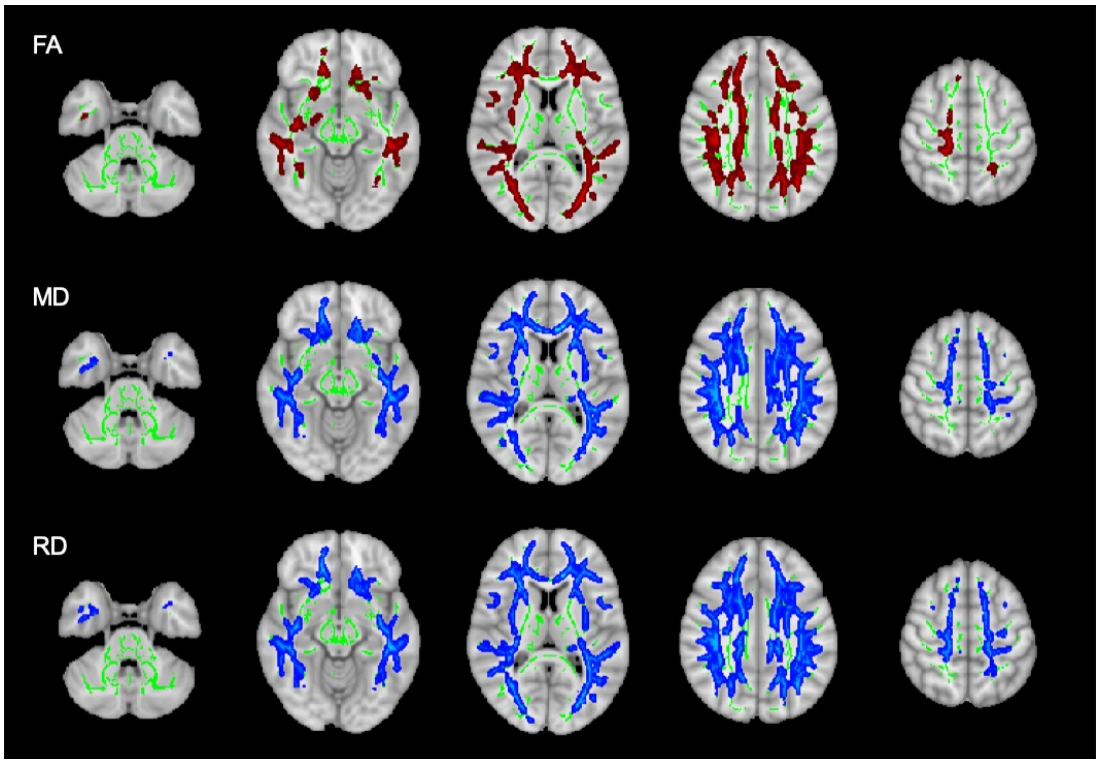
B) CPAL



C) FAS



D) GDA



Tracts where DTI parameters increase as cognitive performance decreases are shown in red. Tracts where DTI parameters decrease as cognitive performance decreases are shown in red. The WM skeleton is shown in green. Axial slices at z = -34, -12, 10, 32 and 54 are shown. AD, axial diffusivity; CPAL, Camden paired associate learning test; FA, fractional anisotropy; FAS, test for phonemic fluency; GDA, graded difficulty arithmetic; MD, mean diffusivity; RD, radial diffusivity; RMTF, recognition memory test for faces; TBSS, tract-based spatial statistics; WM, white matter.