Semiquantitative Scale for Assessing Brain MRI Abnormalities in Wilson Disease: A Validation Study

Petr Dusek, MD, PhD,^{1,2*} ^(D) Lukasz Smolinski, MD,³ Barbara Redzia-Ogrodnik, MD,⁴ Marek Golebiowski, MD, PhD,⁵ Marta Skowronska, MD, PhD,³ Aurelia Poujois, MD, PhD,⁶ Chloe Laurencin, MD,⁷ Iwona Jastrzebska-Kurkowska, MD, PhD,³ Tomasz Litwin, MD, PhD,³ and Anna Członkowska, MD, PhD³

¹Department of Neurology and Center of Clinical Neuroscience, Charles University in Prague, First Faculty of Medicine and General University Hospital, Prague, Czech Republic

²Department of Radiology, Charles University in Prague, First Faculty of Medicine and General University Hospital, Prague, Czech Republic ³2nd Department of Neurology, Institute of Psychiatry and Neurology, Warsaw, Poland

⁴Department of Radiology, Institute of Psychiatry and Neurology, Warsaw, Poland

⁵Department of Clinical Radiology, Medical University of Warsaw, Warsaw, Poland

⁶Neurology Department, French National Reference Centre for Wilson's Disease, Fondation Ophtalmologique Adolphe de Rothschild,

Paris, France

⁷Neurology Department, French National Reference Centre for Wilson's Disease, University Hospital of Lyon, Lyon, France

ABSTRACT: Background: MRI is a sensitive method for the assessment of brain abnormalities in Wilson disease, that is, T_2 hyperintensities, T_2 hypointensities, and atrophy, but a validated scoring system for the classification of radiological severity is lacking. The objective of this study was to develop and validate a brain MRI visual rating scale for Wilson disease.

Methods: The proposed Wilson disease brain MRI severity scale consists of acute toxicity and chronic damage subscores from predefined structures. The former, calculated by summing scores of T_2 hyperintensities (excluding cavitation), is likely to be partially reversible with treatment. The latter, representing the sum of scores of T_2 hypointensities and brain atrophy, reflects pathology that is not readily reversible. Validation was performed on MRI scans acquired using 1.5T system from 39 Wilson disease patients examined at baseline and after 24 months on anticopper treatment. Intraclass correlation coefficients of 5 ratings from 3 raters were calculated. Temporal evolution of the MRI severity score and its

association with clinical severity, assessed using the Unified Wilson Disease Rating Scale part III, was calculated. **Results:** Intrarater and interrater agreement were good (r > 0.93; P < 0.001; and r > 0.74; P < 0.001, respectively). In neurologic Wilson disease patients, the total MRI severity score improved over 2 years (P = 0.032), mainly because of reduced acute toxicity (P = 0.0015), whereas the chronic damage score deteriorated (P = 0.035). Unified Wilson Disease Rating Scale part III score was positively associated with chronic damage and total score at baseline (P = 0.005 and P = 0.003, respectively) and in month 24 (P < 0.001 and P = 0.001, respectively).

Conclusions: The Wilson disease brain MRI severity scale is a simple, reliable, and valid instrument that allows semiquantitative assessment of radiological Wilson disease severity. © 2020 International Parkinson and Movement Disorder Society

Key Words: clinical scales; MRI; neuroradiology; Wilson disease

*Correspondence to: Petr Dusek, Department of Neurology, First Faculty of Medicine and General University Hospital in Prague, Katerinska 30, 120 00, Praha 2, Czech Republic; pdusek@gmail.com

Petr Dusek and Lukasz Smolinski contributed equally to this work.

Relevant conflicts of interest/financial disclosures: Petr Dusek received funding from Czech Ministry of Health, grant No. 15-25602A and No. NV19-04-00120, Czech Science Foundation, grant No. 16-07879S, European Union's Horizon2020 research and innovation programme, grant No. 633190, advisory boardpayment from Alexion pharmaceuticals, and honoraria for clinical trials from BenevolentAlBio Limited and Retrophin. Lukasz Smolinski received honoraria for clinicaltrials from Alexion Pharmaceuticals, Novartis, and GMP-Orphan. MarekGolebiowski, Barbara Redzia-Ogrodnik, Iwona Jastrzebska-Kurkowska, Marta Skowronska, Aurelia Poujois, and Tomasz Litwin have nothing to disclose. Chloé Laurencin received grant from Ellivie. Anna Czlonkowska received advisory board payments from Wilson Therapeutics, Alexion Pharmaceuticals, VIVET Therapeutics, GMP-Orphan, UNIVAR and honoraria as principal investigator of clinical trials from Wilson Therapeutics and GMP-Orphan. None of the funding bodies had any input into the study design or analysis.

Funding agencies: None.

Received: 14 November 2019; Revised: 8 February 2020; Accepted: 10 February 2020

Published online 00 Month 2020 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.28018

Brain magnetic resonance imaging (MRI) is the most sensitive neuroimaging method in the diagnosis of neurologic Wilson disease (WD).¹ More than 90% WD patients with neurologic disease and approximately 40%-70% with hepatic symptoms have abnormalities on brain MRI.² The most prominent MR findings in WD are symmetric hyperintensities in T₂-weighted or fluidattenuation inversion recovery (FLAIR) images in the deep gray matter (DGM) nuclei and white matter predominantly in the brain stem, which presumably reflect edema, demyelination, and gliosis.³ In addition, signs of diffuse tissue atrophy^{4,5} and hypointensities in T_2/T_2^* and susceptibility-weighted images (SWIs) in the DGM caused by abnormal iron accumulation^{6,7} are frequently present. It can be assumed that although T₂/FLAIR hyperintensities represent changes largely reversible with anticopper treatment, atrophy and T₂/T₂*/SWI hypointensities are much less reversible over time.⁸

It has been suggested that brain MRI could be helpful not only in the WD diagnosis but also in treatment monitoring and outcome prediction.⁹ For this, it is necessary to have a robust and validated scoring system that enables the correct classification of radiological severity. In addition, validated neuroimaging marker would be very helpful as a surrogate outcome measure in clinical trials. To quantify the degree of brain parenchyma damage, several scales were developed; they are typically represented by a crude score based on the sum of grades of radiological severity in individual brain structures.^{4,9-13} However, construct validity and reliability of these scales have never been assessed precluding their wider use in clinical and research settings.

A brain MRI severity scale was developed, taking into account clinical relevance and temporal evolution of specific MRI abnormalities present in WD patients, and the internal consistency, intra- and interrater variability, and construct validity of the scale were assessed.

Methods

Patient Characteristics

MR scans from 39 WD patients diagnosed and treated at the Institute of Neurology and Psychiatry (Warsaw, Poland) were analyzed. WD was diagnosed according to established criteria, that is, low serum ceruloplasmin, presence of Kayser-Fleischer rings, and high 24-hour urine copper excretion, with genetic confirmation (detection of 2 pathogenic *ATP7B* mutations) in 37 patients. As brain MRI abnormalities in WD differs between patients with hepatic and neurologic symptoms,¹⁴ patients were divided into 2 subgroups: neurologic presentation, defined as the presence of neurologic symptoms typical for WD regardless of concomitant hepatic abnormalities, and hepatic presentation, defined as the absence of neurologic symptoms and the presence of hepatic abnormalities, usually increased liver function test or symptoms of cirrhosis, at baseline. Patients were examined using the Unified Wilson Disease Rating Scale (UWDRS) part II (activities of daily living) and part III (neurologic examination)¹⁵ at baseline (before treatment initiation) and after 24 months on anticopper treatment. None of the patients had signs of decompensated cirrhosis or hepatic encephalopathy. The study was approved by the local ethics committee, and participants gave written informed consent prior to the study.

Brain MRI Severity Scale

The rating scale builds on previously suggested nonvalidated scoring systems and includes grading of 3 types of radiological abnormalities: T₂/FLAIR hyperintensities, T₂/T₂*/SWI hypointensities, and atrophy. The total WD brain MRI severity score consists of 2 subscores: acute toxicity score and chronic damage score (Table 1).

The acute toxicity score is calculated by summing the individual scores of T₂/FLAIR hyperintensities in all structures and is presumed to be reversible with treatment to a large extent. T₂/FLAIR hyperintensities are rated as absent, mild, or severe separately in 5 predefined regions (putamen, caudate nucleus, thalamus, mesencephalon, pons) and also in other areas, if present. In the case of asymmetric T₂/FLAIR hyperintensities, the more severe side shall be assessed. Notably, T₂ hyperintense lesions, which are at the same time clearly hypointense on T₁-weighted and

TABLE 1. Brain magnetic resonance imaging severity scale for Wilson disease

	Normal/absent	Mild/moderate	Severe
	Acute toxicity score		
	T ₂ /FLAIR hyperintensi	ity	
Putamen ^a	0	1	2
Caudate nucleus ^a	0	1	2
Thalamus ^a	0	1	2
Mesencephalon	0	1	2
Pons	0	1	2
Other area(s) — specify ^a	0	1	2
	Chronic damage scor	re	
·	T ₂ /T ₂ */SWI hypointens	sity	
Globus pallidus ^a	0		1
Putamen ^a	0		1
Caudate nucleus ^a	0		1
Thalamus ^a	0		1
Dentate nucleus ^a	0		1
I	Atrophy (assessed on	T ₁)	
Cortical	0	1	2
Central	0	1	2
Midbrain	0	1	2
Cerebellar	0	1	2

FLAIR, fluid-attenuation inversion recovery; SWI, susceptibility-weighted imaging.

^aIn case of asymmetric lesions, the side with higher severity is rated.

FLAIR images, that is, suggestive of cavities or enlarged perivascular spaces, are not regarded as T_2 hyper-intensities in this scale.

Chronic damage score reflects the sum of individual scores of brain atrophy and $T_2/T_2*/SWI$ hypointensities from predefined structures; both of these MR abnormalities likely reflect changes that are not readily reversible after anticopper treatment initiation. $T_2/T_2*/SWI$ hypointensities are rated as absent (score 0) or present (score 1) separately in 5 predefined bilateral regions (globus pallidus, putamen, caudate nucleus, thalamus, dentate nucleus). Importantly, a region may score positively for T_2 hyperintensity and T_2 hypointensity when both are present simultaneously.

Atrophy is rated as absent, moderate, or severe in predefined regions (cortex, central, midbrain, cerebellum). Visual assessment of cortical and cerebellar atrophy is based on the global cortical atrophy (GCA) score.¹⁶ Score 0 is given in the case of no atrophy or mild widening of the sulci (equivalent to GCA 0 or GCA 1). Score 1 is given in the case of clearly widened sulci and mild or moderate volume loss of the gyri (equivalent to GCA 2). Score 2 is given in the case of profoundly widened cortical sulci and "knife-blade" appearance of gyri (equivalent to GCA 3). Central atrophy is rated according to the width of third ventricle in millimeters as previously suggested¹⁷; if ≤ 6 mm, then score is 0; if >6 mm and < 10 mm, then score is 1; if ≥ 10 mm, then score is 2. Midbrain atrophy is rated according to the anteroposterior midbrain diameter in millimeters¹⁸; if ≥ 18 mm, then score is 0; if <18 mm and > 14 mm, then score is 1; if \leq 14 mm, then score is 2. Lower weighting of T₂/T₂*/SWI hypointensities (ie, range of scores 0-1) was chosen because visual assessment of MR images does not allow reliable distinction of 2 grades of T₂ hypointensities in the DGM and because of their presumably lower impact on CNS function compared with atrophy.

MRI Assessment

MRI was performed using a Philips Achieva 1.5T system (Phillips Healthcare, Eindhoven, Netherlands). The MRI protocol included following routine clinical images: T₁-weighted (spin-echo [SE], repetition time [TR], 596 milliseconds; echo time [TE], 15 milliseconds; voxel resolution, $0.9 \times 0.9 \times 5 \text{ mm}^3$), T₂-weighted (SE, TR, 6783 milliseconds; TE, 140 milliseconds; voxel resolution, $0.4 \times 0.4 \times 5 \text{ mm}^3$), FLAIR (TR, 11,000 milliseconds; TE, 140 milliseconds; inversion time, 2800 milliseconds; voxel resolution, $0.9 \times 0.9 \times 5 \text{ mm}^3$), T_2^* -weighted (gradient-echo [GRE], TR, 693 milliseconds; TE, 23 milliseconds; flip angle, 20°; voxel resolution, $0.9 \times 0.9 \times 5$ mm³), and VEN_BOLD (equivalent to SWI magnitude, GRE, TR, 49.7 milliseconds; TE, 34.7 milliseconds; flip angle, 15° ; voxel resolution, $0.4 \times 0.4 \times 2 \text{ mm}^3$). All images covered the entire brain and were acquired in the axial plane; the plane of the T_1 -weighted image used for atrophy assessment was perpendicular to the dorsal edge of the brain stem.

Thirty-eight patients were examined twice: at baseline, before treatment initiation, and after 24 months on anticopper treatment. In 1 patient who died, we used the most recent follow-up examination, which was done 18 months after baseline. All MRI examinations were anonymized by assigning random numbers. Then, 3 experienced neuroradiologists (P.D., B.O., M.G.), blinded to clinical data and examination date, independently rated the MRI examinations. Two raters (P.D., B.O.) assessed all (reanonymized) images for a second time with a 6-month delay. Specific instructions on how to grade the changes according to the WD brain MRI severity scale were provided before scoring. In addition, all raters were provided with a scoring manual with image examples (Supplementary Material 2).

Statistical Analysis

Internal consistency of the MRI severity scale was assessed with Cronbach's alpha. To determine agreement between raters, the intraclass correlation coefficient (ICC) was calculated for the acute toxicity, chronic damage, and total scores. In addition, the ICC between the first and second ratings was calculated to determine intrarater reliability. The median of scores from the first ratings from the 3 raters was used to analyze associations between the MR severity score and clinical status. Patients with the hepatic or neurologic phenotype were compared with the t test, Mann-Whitney U test, and chi-square test, as appropriate. Pairwise comparisons for clinical and imaging parameters between baseline and month 24 were performed using the Wilcoxon signed rank test. Associations among imaging and clinical scores were calculated using Spearman's rank correlation coefficient. A multiple regression model was used to test whether baseline imaging severity predicts clinical outcome. Statistical analyses were conducted using R software (version 3.5.3).

Results

Thirty-nine patients (26 men; mean age \pm standard deviation [SD], 35.1 ± 11.8 years) were examined. Eighteen patients (5 men; mean age, 31.6 ± 9.6 years) had hepatic presentation; 6 were treated by D-penicillamine and 12 by zinc salts. Twenty-one patients (13 men; mean age, 38.1 ± 12.8 years) had neurologic presentation; 15 were treated by D-penicillamine and 6 by zinc salts (Supplementary Table 1).

Reliability of the Scale

The internal consistency of the scale was good: Cronbach's alpha (95% confidence interval [CI]) was

DUSEK ET AL

0.79 (0.75–0.83) for the total MRI score, 0.78 (0.74–0.82) for the acute toxicity score, and 0.75 (0.71–0.79) for the chronic damage score. Intrarater agreement (ICC) was excellent for all examined indices, that is for acute toxicity score (0.94 [0.93–0.96]; P < 0.001), chronic damage score (0.93 [0.91–0.95]; P < 0.001), and total score (0.95 [0.93–0.96]; P < 0.001). Interrater agreement (ICC) was also good for the acute toxicity score (0.88 [0.82–0.92]; P < 0.001), and total score (0.82]; P < 0.001), and total score (0.84 [0.73–0.90]; P < 0.001).

Clinical Severity

In neurologic WD patients, UWDRS part II and part III scores generally dropped from baseline to month 24, but this change was not statistically significant (Fig. 1, Supplementary Table 2). Two neurologic WD patients (9.5%) developed irreversible neurological worsening on anticopper treatment (both were treated by D-penicillamine); the first worsened by 32/101 points in UWDRS part II/III and ultimately died, and the second worsened by 33/69 points in UWDRS part II/III, resulting in severe disability. When excluding these 2 patients with neurological worsening, UWDRS part II (P = 0.011) and part III (P = 0.0037) in neurologic patients improved significantly in month 24. UWDRS part II and part III scores did not change significantly in patients with the hepatic phenotype (Supplementary Table 2).

Imaging Severity

Analyzing median values from 3 raters, the total MRI severity score improved significantly in neurologic WD patients (P = 0.032; Fig. 2A); improvement of ≥ 2 points (which we arbitrarily consider as radiologically meaningful change) was observed in 10 patients (48%) and worsening of ≥ 2 points in 2 patients (10%). This improvement was mainly caused by a significant drop of the acute toxicity score (P = 0.0015; Fig. 2B),



FIG. 1. Change in Unified Wilson's Disease Rating Scale over time in neurologic patients with Wilson disease. Two patients with neurologic worsening on anticopper treatment are depicted by a dashed line. [Color figure can be viewed at wileyonlinelibrary.com]

whereas the chronic damage score increased over 2 years (P = 0.035, Figs. 2C and 3). An acute toxicity score improvement of ≥ 2 points was observed in



FIG. 2. (A) Total, (B) acute toxicity, and (C) chronic damage MRI score changes in patients with neurologic and hepatic phenotypes of WD. Two patients with neurologic worsening on anticopper treatment are depicted by dashed lines. Please note that the low dynamic range of values in hepatic patients leads to overlapping lines, particularly for acute toxicity score. [Color figure can be viewed at wileyonlinelibrary.com]



FIG. 3. Comparison of T₂-weighted MR images at baseline and after 24 months of anticopper treatment in an illustrative neurologic WD patient. Almost complete resolution of T₂ hyperintensities in the pons (black arrowhead), mesencephalon (black arrow), putamen (white arrowhead), and thalamus (white arrow) along with mild progression of central atrophy, that is, widening of the third ventricle, can be seen (empty arrow). Note the mixed-signal abnormality in the putamen containing both T₂ hyperintensities and T₂ hypointensities. UWDRS II/III improved from 1/10 at baseline to 0/0 in month 24, whereas the acute toxicity MR score dropped from 9 to 5 and chronic damage score increased from 3 to 5 points on treatment.

13 patients (62%) and worsening of ≥ 2 points in 1 neurologic WD patient who also experienced neurologic worsening (5%); the total MRI severity score in this patient worsened by 5 points because of the development of extensive periventricular and frontal corticosubcortical white-matter lesions (Supplementary Fig. 1A). For the chronic damage score, improvement of ≥ 2 points was observed in 2 patients (10%) and worsening of ≥ 2 points in 8 neurologic WD patients (38%). In the second patient with neurologic worsening, improvement of the acute toxicity score was offset by worsening of the chronic damage score because of progression of central atrophy and development of thalamic cavitation, resulting in an unchanged total MRI score (Supplementary Fig. 1B). The temporal evolution of acute toxicity, chronic damage, and total scores on treatment followed the same pattern in all individual raters, although the changes were less significant compared with the analysis with median values from all 3 raters (Supplementary Fig. 2, Supplementary Table 3).

To examine the contribution of T_2/T_2^* /SWI hypointensities and atrophy to the chronic damage score, we performed post hoc tests analyzing these subscores separately. Both subscores worsened with treatment, but the change was significant only for atrophy (*P* = 0.0078; Supplementary Fig. 3, Supplementary Table 2).

MRI severity scores did not change significantly between baseline and month 24 in hepatic WD patients (Fig. 2, Supplementary Table 2).

Association Between Clinical and Imaging Severity

At baseline, UWDRS part III score was positively associated with the total MR severity score (r = 0.62, P = 0.0028) and chronic damage score (r = 0.59, P = 0.0052) in neurologic WD patients (Fig. 4). In month 24, a similar pattern with an even stronger association between UWDRS part III and total MR severity score (r = 0.65, P = 0.0013) and chronic damage score (r = 0.68, P < 0.00062) was observed. In contrast, the association between the acute toxicity score and UWDRS part III was significant in month 24 only (r = 0.44, P = 0.046); however, the significance of this association was lost (r = 0.34, P = 0.13) after excluding the data point from the outlier patient with neurologic and acute toxicity score worsening on treatment (Fig. 4). Post hoc tests analyzing the subscores of the chronic damage score showed that only the atrophy subscore was significantly associated with UWDRS part III at baseline (r = 0.55, P = 0.0094) and in month 24 (r = 0.7, P = 0.00037). There was only a trend for a significant association between T2/T2*/SWI hypointensities and UWDRS part III at baseline (r = 0.37, P = 0.097; Supplementary Fig. 4).



FIG. 4. Associations between MRI severity and UWDRS part III scores (*R* is Spearman correlation coefficient). *r = 0.34, P = 0.13 after excluding an outlier patient who developed white-matter lesions accompanied by neurologic worsening. [Color figure can be viewed at wileyonlinelibrary.com]

Individual correlation coefficients between MRI and UWDRS part III scores for each rater were minimally worse but followed the same pattern as the median scores from all 3 raters (Supplementary Table 4).

To test whether baseline imaging severity predicts clinical outcome in month 24, a multiple regression model with UWDRS part III in month 24 as the dependent variable and acute toxicity and chronic damage scores at baseline as predictor variables was performed. Results showed that an increase in the chronic damage score of 1 point at baseline was associated with a greater UWDRS part III score of nearly 5 points in month 24, but there was only a statistical trend for this relationship (P = 0.062; Supplementary Table 5). Adding baseline UWDRS part III score and disease duration as predicting factors did not improve the multiple regression model (Supplementary Table 6). Baseline UWDRS part III score alone also did not predict UWDRS part III score in month 24 (Supplementary Table 7).

Discussion

We developed a novel scale for quantifying the extent of pathological changes on brain MRI in WD patients and good intrarater and interrater agreement was verified for scores of acute toxicity and chronic damage as well as for the combined total score. In addition, acute toxicity score improved on anticopper treatment, but it was not related to neurologic severity at baseline or in month 24. On the other hand, the chronic damage score significantly deteriorated on anticopper treatment, whereas it was positively associated with neurologic severity at baseline and even stronger in month 24.

The range of total scores confirms that included cases represent a wide spectrum of clinical and radiological severity. Consequently, the proposed scale is reliable across a broad range of radiological scenarios. Quantitative results confirmed previous clinical experience showing that radiological severity is higher in neurologic compared with hepatic WD patients and that it improves on anticopper treatment, mainly because of the drop in acute toxicity score.⁹ However, the acute toxicity score was not associated with clinical severity at baseline, and although its drop paralleled the improvement of neurological symptoms, it was also not associated with clinical severity after 24 months of anticopper treatment. This was corroborated by previous findings showing that T₂ hyperintensities in the basal ganglia may be present in hepatic WD patients despite the absence of neurologic symptoms¹⁹; their presence may, however, be a risk factor for the development of neurologic symptoms after treatment initiation.²⁰ There are several possible explanations for this clinicalradiologic mismatch. First, pathology not visible on routine MRI scans may significantly contribute to clinical disability. Interestingly, quantitative analysis of diffusion tensor imaging parameters showed altered tissue microstructure, even in normal-appearing thalamus and lobar white matter in WD.^{21,22} Second, the pathological basis of T₂/FLAIR hyperintensities in WD is likely heterogenous; it may theoretically reflect edema, demyelination, rarefaction, or gliosis. These pathologies likely induce distinct CNS dysfunctions with varied severity. The former 2 pathologies are reversible, and their resolution likely underlies the improvement in acute toxicity score, whereas the latter 2 pathologies are mostly irreversible and may potentially contribute by a greater deal to neurologic disability. Last, we did not include the rating of psychiatric symptoms into clinical assessment. Because organic psychiatric symptoms and cognitive impairment are frequent consequences of CNS damage in WD,²³ the addition of a psychiatric and cognitive scale may improve the correlation between clinical and radiological severity.

The acute toxicity scores meaningfully improved in 62% of neurological WD patients on treatment. Paradoxical worsening of the acute toxicity score was observed only in 1 patient in whom extensive corticosubcortical T_2 /FLAIR hyperintensities appeared. Such abnormalities were described previously in WD.²⁴⁻²⁸ Although the pathophysiological underpinnings and relation to copper toxicity are unknown, their appearance is considered a failure of anticopper treatment. Overall, these results confirm the validity of the construct that T₂/FLAIR hyperintensities are caused by acute copper toxicity and are potentially reversible with anticopper treatment.

On the other hand, chronic damage score, particularly its atrophy subscore, was moderately associated with clinical severity, indicating that tissue atrophy and, to a lesser degree, iron accumulation are important factors determining the function of the CNS. A smaller contribution of the T₂ hypointensity subscore to clinical severity justifies its lower weighting in the chronic damage score. A similar pattern was observed in multiple sclerosis in which atrophy and iron deposits in the DGM structures contributed to the disability to a degree similar to the T_2 -lesion load in the white matter.²⁹ It has been recently been shown that total brain volume and volumes of gray and white matter⁵ as well as the degree of volume loss in the putamen and globus pallidus³⁰ are associated with disability in WD. To our knowledge, there are no studies quantitatively examining the relation between cerebral iron levels and disability in WD patients. However, a positive association between iron concentration and neuropathologic severity in the putamen in WD brains was shown in a histopathological study.⁶ Although simplicity and usage of routine clinical MR scans are important advantages of this scale,

quantitative measurement of iron concentration in DGM nuclei and brain volumetric analysis are parameters that would likely further increase its accuracy.

Chronic damage score meaningfully deteriorated in 38% of neurologic WD patients on treatment. Surprisingly, improvement in chronic damage score by 2 points was observed in 2 neurologic WD patients (10%). However, post hoc cross-check revealed that it was a consequence of biased rating rather than resolution of atrophy or iron deposits. Notably, lower interrater agreement in the assessment of the chronic damage score compared with the acute toxicity score corroborates the higher complexity of the former score. Overall, the assumption that the chronic damage score reflects changes that mostly do not improve but, conversely, may even worsen during first years of anticopper treatment was confirmed. It remains unclear whether progression of atrophy and iron accumulation after anticopper treatment initiation is a consequence of mobilization of tissue copper, natural progression of disease, or other factors.³ Importantly, except for the statistical trend for the effect of chronic damage score, radiologic and clinical severity as well as disease duration at baseline were not significantly associated with clinical severity in month 24. Therefore, high total brain MRI severity or UWDRS part III scores at diagnosis do not necessarily imply unfavorable clinical outcome. The prognostic value of the chronic damage score for clinical outcome needs to be confirmed in larger studies.

Several limitations of the WD brain MR severity scale are worth mentioning. In this study, we used a 1.5T scanner for the scale validation. As iron accumulation may be present concomitantly with pathology causing T₂/FLAIR hyperintensities in tissue, T₂-hypointense lesions may become prominent at a higher magnetic field with increased sensitivity to paramagnetic effects of iron and overshadow T₂/FLAIR hyperintensities.¹⁴ Additional validation thus would be necessary before it is used on 3T scanners. Furthermore, hyperintense lesions on T₁-weighted images seen in patients with portosystemic shunt are not rated in this scale. This scale is thus not suitable for monitoring of WD patients with hepatic encephalopathy. Last, clinical symptoms and MRI abnormalities may further change beyond 24 months of anticopper treatment. Therefore, the UWDRS and MRI severity scores in month 24 may not reflect the ultimate treatment outcome.

In conclusion, the WD brain MRI rating scale is a simple and reliable visual rating scale that allows for semiquantitative assessment of radiological WD severity. We demonstrated that the theoretical assumptions for independent rating of the reversible acute toxicity and irreversible chronic damage scores are valid. This scale can be used for standardized monitoring of anticopper therapy in clinical practice and therapeutic studies. In the latter, improvement of acute toxicity score and nonworsening of chronic damage score may be employed as surrogate outcome measures.

References

- Czlonkowska A, Litwin T, Dusek P, et al. Wilson disease. Nat Rev Dis Primers 2018;4(1):21.
- Litwin T, Gromadzka G, Czlonkowska A, Golebiowski M, Poniatowska R. The effect of gender on brain MRI pathology in Wilson's disease. Metab Brain Dis 2013;28(1):69–75.
- van Wassenaer-van Hall HN, van den Heuvel AG, Jansen GH, Hoogenraad TU, Mali WP. Cranial MR in Wilson disease: abnormal white matter in extrapyramidal and pyramidal tracts. AJNR Am J Neuroradiol 1995;16(10):2021–2027.
- Sinha S, Taly AB, Ravishankar S, et al. Wilson's disease: cranial MRI observations and clinical correlation. Neuroradiology 2006;48 (9):613–621.
- 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(10):2089–2095.
- 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(6):514–532.
- Skowronska M, Litwin T, Dziezyc K, Wierzchowska A, Czlonkowska A. Does brain degeneration in Wilson disease involve not only copper but also iron accumulation? Neurol Neurochir Pol 2013;47(6):542–546.
- Dusek P, Litwin T, Czlonkowska A. Neurologic impairment in Wilson disease. Ann Transl Med 2019;7(Suppl 2):S64.
- Sinha S, Taly AB, Prashanth LK, Ravishankar S, Arunodaya GR, Vasudev MK. Sequential MRI changes in Wilson's disease with decoppering therapy: a study of 50 patients. Br J Radiol 2007;80(957): 744–749.
- Frota NA, Barbosa ER, Porto CS, et al. Cognitive impairment and magnetic resonance imaging correlations in Wilson's disease. Acta Neurol Scand 2013;127(6):391–398.
- 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(1):154–160.
- King AD, Walshe JM, Kendall BE, et al. Cranial MR imaging in Wilson's disease. AJR Am J Roentgenol 1996;167(6): 1579–1584.
- Maskova J, Skoloudik D, Burgetova A, et al. Comparison of transcranial sonography-magnetic resonance fusion imaging in Wilson's and early-onset Parkinson's diseases. Parkinsonism Relat Disord 2016;28:87–93.
- 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 2019; DOI: https://doi.org/10.1002/jmri.26984
- 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):1–12.
- Pasquier F, Leys D, Weerts JG, Mounier-Vehier F, Barkhof F, Scheltens P. Inter- and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts. Eur Neurol 1996;36(5):268–272.
- 17. Benedict RH, Weinstock-Guttman B, Fishman I, Sharma J, Tjoa CW, Bakshi R. Prediction of neuropsychological impairment in

multiple sclerosis: comparison of conventional magnetic resonance imaging measures of atrophy and lesion burden. Arch Neurol 2004; 61(2):226–230.

- Warmuth-Metz M, Naumann M, Csoti I, Solymosi L. Measurement of the midbrain diameter on routine magnetic resonance imaging: a simple and accurate method of differentiating between Parkinson disease and progressive supranuclear palsy. Arch Neurol 2001;58 (7):1076–1079.
- Kozic D, Svetel M, Petrovic B, Dragasevic N, Semnic R, Kostic VS. MR imaging of the brain in patients with hepatic form of Wilson's disease. Eur J Neurol 2003;10(5):587–592.
- Czlonkowska A, Dziezyc-Jaworska K, Klysz B, Barbara RO, Litwin T. Difficulties in diagnosis and treatment of Wilson disease-acase series of five patients. Ann Transl Med 2019;7(Suppl 2):S73.
- 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;25:31–37.
- Favrole P, Chabriat H, Guichard JP, Woimant F. Clinical correlates of cerebral water diffusion in Wilson disease. Neurology 2006;66 (3):384–389.
- Litwin T, Dusek P, Szafranski T, Dziezyc K, Czlonkowska A, Rybakowski JK. Psychiatric manifestations in Wilson's disease: possibilities and difficulties for treatment. Ther Adv Psychopharmacol 2018;8(7):199–211.
- 24. Trocello JM, Woimant F, El Balkhi S, et al. Extensive striatal, cortical, and white matter brain MRI abnormalities in Wilson disease. Neurology 2013;81(17):1557.
- Shetty AS, Kashikar R, Nagral A, Wadia PM. Cortical Cystic Necrosis in Wilson Disease. JAMA Neurol 2016;73(3):350–351.
- Aikath D, Gupta A, Chattopadhyay I, et al. Subcortical white matter abnormalities related to drug resistance in Wilson disease. Neurology 2006;67(5):878–880.
- Mikol J, Vital C, Wassef M, et al. Extensive cortico-subcortical lesions in Wilson's disease: clinico-pathological study of two cases. Acta Neuropathol 2005;110(5):451–458.
- Woimant F, Djebrani-Oussedik N, Collet C, Girardot N, Poujois A. The hidden face of Wilson's disease. Rev Neurol (Paris) 2018;174 (9):589–596.
- Burgetova A, Dusek P, Vaneckova M, et al. Thalamic Iron Differentiates Primary-Progressive and Relapsing-Remitting Multiple Sclerosis. AJNR Am J Neuroradiol 2017;38(6):1079–1086.
- Zou L, Song Y, Zhou X, Chu J, Tang X. Regional morphometric abnormalities and clinical relevance in Wilson's disease. Mov Disord 2019;34(4):545–554.
- Dusek P, Skoloudik D, Maskova J, et al. Brain iron accumulation in Wilson's disease: A longitudinal imaging case study during anticopper treatment using 7.0T MRI and transcranial sonography. J Magn Reson Imaging 2018;47(1):282–285.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

SGML and CITI Use Only DO NOT PRINT

Authors' Roles

Research project: A. Conception, B. Organization, C. Execution; 2. Statistical Analysis: A. Design, B. Execution,
 Review and Critique; 3. Manuscript Preparation: A. Writing of the first draft, B. Review and Critique
 Petr Dusek 1A,B,C; 2A,C; 3A.
 Lukasz Smolinski 1A,B,C; 2C, 3A.
 Barbara Redzia-Ogrodnik 1A,B,C; 2C; 3A.
 Marek Golebiowski 1C; 2C; 3A.
 Marta Skowronska 1B,C; 2C; 3A.
 Chloe Laurencin 1C; 2C; 3A.
 Iwona Jastrzebska-Kurkowska 1C; 2C; 3A.
 Tomasz Litwin 1B,C; 2C; 3A.
 Anna Członkowska 1A,B,C; 2C; 3B.

Financial Disclosures

Petr Dusek received funding from the Czech Ministry of Health, grants 15-25602A and NV19-04-00120, Czech Science Foundation, grant 16-07879S, European Union's Horizon 2020 research and innovation programme, grant 633190, advisory board payment from Alexion pharmaceuticals, and honoraria for clinical trials from BenevolentAI Bio Limited and Retrophin. Lukasz Smolinski received honoraria for clinical trials from Alexion Pharmaceuticals, Novartis, and GMP-Orphan. Marek Golebiowski, Barbara Redzia-Ogrodnik, Iwona Jastrzebska-Kurkowska, Marta Skowronska, Aurelia Poujois, and Tomasz Litwin have nothing to disclose. Chloé Laurencin received grant from Ellivie. Anna Czlonkowska received advisory board payments from Wilson Therapeutics, Alexion Pharmaceuticals, Therapeutics, GMP-Orphan, UNIVAR, and honoraria as principal investigator of clinical trials from Wilson Therapeutics and GMP-Orphan. None of the funding bodies had any input into the study design or analysis.