Cognition in multiple sclerosis
State of the field and priorities for the future

James F. Sumowski, PhD, Ralph Benedict, PhD, Christian Enzinger, MD, Massimo Filippi, MD, Jeroen J. Geurts, PhD, Paivi Hamalainen, PhD, Hanneke Hulst, PhD, Matilde Inglese, MD, PhD, Victoria M. Leavitt, PhD, Maria A. Rocca, MD, Eija M. Rosti-Otajarvi, PhD, and Stephen Rao, PhD


Abstract

Cognitive decline is recognized as a prevalent and debilitating symptom of multiple sclerosis (MS), especially deficits in episodic memory and processing speed. The field aims to (1) incorporate cognitive assessment into standard clinical care and clinical trials, (2) utilize state-of-the-art neuroimaging to more thoroughly understand neural bases of cognitive deficits, and (3) develop effective, evidence-based, clinically feasible interventions to prevent or treat cognitive dysfunction, which are lacking. There are obstacles to these goals. Our group of MS researchers and clinicians with varied expertise took stock of the current state of the field, and we identify several important practical and theoretical challenges, including key knowledge gaps and methodologic limitations related to (1) understanding and measurement of cognitive deficits, (2) neuroimaging of neural bases and correlates of deficits, and (3) development of effective treatments. This is not a comprehensive review of the extensive literature, but instead a statement of guidelines and priorities for the field. For instance, we provide recommendations for improving the scientific basis and methodologic rigor for cognitive rehabilitation research. Toward this end, we call for multidisciplinary collaborations toward development of biologically based theoretical models of cognition capable of empirical validation and evidence-based refinement, providing the scientific context for effective treatment discovery.
**Glossary**

BVMT-R = Brief Visuospatial Memory Test–Revised; CVLT-II = California Verbal Learning Test–II; DMT = disease-modifying therapy; MS = multiple sclerosis; PST = Processing Speed Test; RCT = randomized controlled trial; SDMT = Symbol Digit Modalities Test.

It has been 140 years since Charcot described “marked enfeeblement of the memory” with “conceptions [that] are formed slowly” in persons with multiple sclerosis (MS). Such cognitive symptoms were overlooked during much of the 20th century before Rao et al. brought renewed attention to MS cognitive deficits in *Neurology®* about 25 years ago, beginning a quarter century of research on the prevalence, expression, and neural bases of MS cognitive dysfunction. Current work aims to incorporate cognitive assessment into MS clinics and clinical trials, utilize state-of-the-art neuroimaging to explicate neural bases of deficits, and develop effective symptomatic cognitive treatments. First, however, key knowledge gaps require attention and methodologic approaches need improvement to advance the field toward these goals, with the ultimate goal of effective, evidence-based, clinically feasible interventions to prevent or treat cognitive deficits. Academic articles typically emphasize what is known, but awareness of the unknown provides the catalyst for scientific discovery. As such, our international team of MS experts identified critical gaps or flaws in our knowledge and makes recommendations for future work. This is not a thorough review, but a joint statement of critical research priorities.

### Cognitive dysfunction due to MS

#### Cognitive profile

Slowed cognitive processing speed and episodic memory decline are the most common cognitive deficits in MS, with additional difficulties in executive function, verbal fluency, and visuospatial analysis. Anecdotally, patients often report difficulties with multitasking and word-finding, which are sorely underinvestigated. Cognitive decline often emerges early in disease, but impairment is more prevalent and may differ qualitatively (e.g., risk for working memory deficits) among persons with progressive vs relapsing disease. Although MS leads to deficits in multiple cognitive domains on the group level, we know little about variability in patient-level expression of cognitive deficits (e.g., patterns of isolated vs co-occurring deficits, discussed below). It is also unknown whether deficits in one cognitive domain (e.g., speed) contribute to dysfunction in other domains (e.g., memory). Although speed and memory are correlated in

### Table 1 Key priorities for understanding and measuring cognitive deficits

1. Research has identified prevalent MS cognitive deficits on the group/sample level (e.g., speed, memory). Future work should examine prevalence and expression of patient-level variability in cognitive profiles (i.e., patterns of isolated or co-occurring deficits). Also, cognitive tasks are often ascribed one function (e.g., Symbol Digit Modalities Test and processing speed), but tasks necessarily measure other functions (e.g., incidental memory). Future research should consider composite or latent variables as purer measures of targeted functions.

2. Cognitive abilities are assessed individually in optimal environments (i.e., single task performed while sitting in a quiet room), but many patients report difficulty managing multiple goals simultaneously (e.g., cooking while having a conversation). Research on cognitive–motor and cognitive–cognitive multitasking is needed to investigate real-world dual-tasking deficits. This includes validation of multitasking assessment procedures for clinical use, which may better represent patient-reported real-world cognitive deficits.

3. It is unclear whether relationships across cognitive domains (e.g., speed, memory) are dependent or independent. Although cognitive functions are correlated in MS, causal conclusions cannot be supported by correlations (e.g., slowed speed impairs learning), especially since cognition across domains is also robustly correlated in healthy persons. Caution against presuming causal links among correlated functions extends to treatment expectations: it is incorrect to assume that improvement in one function (e.g., working memory) will improve performance in a correlated function (e.g., episodic memory).

4. Substantial time and resource demands of neuropsychological assessment are obstacles to incorporation of routine cognitive monitoring into standard of care for patients with MS. We require validated cognitive monitoring tools that can be practically and seamlessly incorporated into the clinical MS center setting. Tasks must be time- and resource-efficient, with easy incorporation into medical records. These may be tablet-based.

5. Cognitive impairment is typically defined as poor performance on 2 or 3 of several diverse tasks, which leads to heterogeneous and ill-defined groups of patients with deficits in speed, memory, or other deficits. We require more precise cognitive phenotyping of patients (e.g., isolated memory impairment, speed and memory impaired). This is critical for neuroimaging research aiming to identify neural bases of functions.

6. Prospective longitudinal designs are necessary to improve understanding of disease-related cognitive decline, rather than relying on cross-sectional definitions of impairment. Clinically, brief cognitive assessments of all new patients would allow judgment of future decline relative to baseline. This is a necessary step toward using cognition as a marker of disability progression and treatment efficacy.

7. Unlike interval disability metrics (e.g., Expanded Disability Status Scale), clinical meaningfulness of quantitative cognitive test scores (or change in scores) is poorly understood. We need to identify clinically meaningful change scores to be used in treatment monitoring and clinical trials.

Abbreviation: MS = multiple sclerosis.
### Table 2: Review of cognitive tests and guidelines for cognitive assessment

<table>
<thead>
<tr>
<th>Test; cognitive domain</th>
<th>Batteries; validated outcomes</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDMT Cognitive processing speed</td>
<td>MACFIMS BRB MS-COG BICAMS MSFC</td>
<td>Very high sensitivity (mean d = 1.11)(^{13}); most sensitive task in MS Good to excellent reliability Fast and easy to administer Well-tolerated by patients Uniform across languages No floor or ceiling effects Multiple alternate forms available Preliminary evidence for 3–4 point clinically meaningful change(^{13})</td>
<td>Other functions affected by MS may contribute to performance (incidental learning of symbol-digit pairings;(^{14}) visual scanning)</td>
<td>1. Highly recommended as a cognitive monitoring tool in clinical practice 2. Highly recommended as a cognitive assessment tool in research, including cross-sectional and longitudinal designs, and as an outcome in clinical trials 3. SDMT should not be considered a pure measure of latent processing speed (due to concomitant learning and visual scanning requirements) 4. Incorporation of cognitive assessment into clinical practice with tablet- or Internet-based processing speed tasks is a goal for the future</td>
</tr>
<tr>
<td>PASAT Cognitive processing speed, working memory</td>
<td>MACFIMS BRB MS-COG</td>
<td>Moderate sensitivity (mean d = 0.63)(^{13}), but less sensitive than the SDMT as a task of cognitive processing speed/efficiency Previously the most widely used cognitive test in MS research, including clinical trials; there is therefore a large amount of research PASAT data to compare across prior studies PASAT stimuli are auditory; most cognitive efficiency tasks require visual processing, PASAT may be useful for patients with poor vision</td>
<td>Reliability limited by practice effects Susceptible to ceiling effects Poorly tolerated by patients Specificity limited by multiple factors, including math ability and test-taking anxiety</td>
<td>1. Not recommended for cognitive monitoring in clinical practice (The SDMT is more sensitive to cognitive deficits(^{15}) and has replaced PASAT as the MSFC cognitive task(^{13}) 2. Not recommended for clinical trials or designs with multiple administrations (due to practice effects) 3. The PASAT is a putative cognitive processing task, which allows results to be compared across previous studies 4. There is a need for new cognitive efficiency tasks using nonvisual stimuli for patients with visual impairment</td>
</tr>
<tr>
<td>SRT Verbal memory</td>
<td>BRB MS-COG 1. TL 2. LTS 3. CLTR 4. Delayed recall</td>
<td>High sensitivity (mean d = 0.86)(^{13}) Several alternate forms (although reliability across alternate forms needs further validation in MS) The selective reminding paradigm emphasizes retrieval of words from long-term store (secondary memory) rather than repetition from working (primary) memory</td>
<td>There is no single authoritative set of normative data, although there are different sets of published norms SRT (like other verbal memory assessments) is a poor candidate for future unsupervised assessment because tablet-based tasks assess verbal recognition, which is less sensitive than verbal recall in MS</td>
<td>1. Recommended as a verbal memory test for clinical and research use, especially when more than 2 administrations are planned 2. The SRT has 3 validated initial learning outcomes (TL, LTS, CLTR) that are highly intercorrelated; researchers are advised to identify which of these 3 measures of initial learning will be used</td>
</tr>
<tr>
<td>CVLT-II Verbal memory</td>
<td>MACFIMS BICAMS 1. TL 2. LDFR</td>
<td>High sensitivity (mean d = 0.89)(^{13}) Good age- and sex-adjusted normative data from a large standardization sample in the United States One standard and one alternate form well-validated</td>
<td>There is only one alternate form For the standard administration, additional trials between TL and LDFR (List B, Short Delay Free Recall, Short Delay Cued Recall) add time, and are required to use published normative data for LDFR CVLT-II (like other verbal memory tests) is a poor candidate for future unsupervised assessment, as stated above for the SRT</td>
<td>1. Recommended as a verbal memory test for clinical and research use, especially when robust normative data are needed (but not when more than 2 administrations are required) 2. Unlike processing speed and visuospatial memory, it will be more challenging to develop table- or Internet-based verbal memory tests appropriate for persons with MS 3. Researchers should report whether all trials of the CVLT-II (e.g., List B, SDFR, SDCR) were administered or omitted 4. Recently published CVLT-III uses the same stimuli and administration as the CVLT-II, but adds a new scoring option to sum delayed recall trials (SDFR, SDCR, LDFR, LDCR) into one composite, which is more reliable than LDFR alone; this requires validation in MS</td>
</tr>
</tbody>
</table>

Continued
### Table 2: Review of cognitive tests and guidelines for cognitive assessment (continued)

<table>
<thead>
<tr>
<th>Test; cognitive domain</th>
<th>Batteries; validated outcomes</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Recommendations</th>
</tr>
</thead>
</table>
| **BVMT-R**  
Visuospatial memory | MACFIMS BICAMS  
1. TL  
2. Delayed recall | Very high sensitivity (mean d = 1.03)\(^{1,9}\)  
Good reliability and validity  
Time-efficient for a memory test  
(much briefer than SRT and CVLT-II)  
Well-tolerated by patients  
Six well-validated alternate forms  
Good age- and sex-adjusted  
normative data published in  
a large standardization sample  
in the United States | Severe motor impairment in patients with advanced disease may complicate assessment | 1. Recommended as a good test for memory monitoring; BVMT-R is the most sensitive memory test within batteries designed for persons with MS  
2. Recommended for research, including observational cross-sectional and longitudinal designs, and as an outcome in clinical trials  
3. The drawing/construction requirement (although modest) may preclude assessment in patients with severe motor impairment; a tablet-based (or completely motor-free) visuospatial task would be ideal; however, no such task is currently validated for persons with MS |
| **10/36 SPART**  
Visuospatial memory | BRB  
1. TL  
2. Delayed recall | Construction/drawing is not required, which may be helpful in persons with severe motor impairment; however, upper extremity function is still needed (patients place markers on a grid) | Lower sensitivity (mean d = 0.48)\(^{10,15}\) than other memory tests for MS (BVMT-R, SRT, CVLT-II)  
Reliability and good normative data are lacking | 1. Due to low sensitivity, unknown reliability, and poor normative data, the SPART is not recommended as a test for memory monitoring  
2. SPART is a reasonable measure of spatial memory for research; however, BVMT-R is a more sensitive choice with established reliability and robust norms  
3. SPART may be a useful test of spatial memory for patients with severe drawing impairment, although SPART still requires placement of markers |
| **COWAT**  
Verbal fluency | BRB MACFIMS  
1. Total score (3 trials with different letters) | Moderate sensitivity (mean d = 0.54)\(^{10,15}\) but lower than other tasks in used in MS  
Validated alternate form  
Time-efficient and easy to administer | Much more related to education and vocabulary than MS disease burden  
Unclear whether task’s sensitivity is only explained by processing speed | 1. Not recommended for screening or brief monitoring in clinical practice  
2. Not recommended as an outcome for clinical trials, unless verbal fluency is the specific target of the intervention  
3. Reasonable component of a comprehensive neuropsychological assessment for research or practice |
| **D-KEFS sorting test**  
Executive function | MACFIMS  
1. Total correct sorts  
2. Description score | Alternate form is available (unlike most executive function tasks)  
Sensitivity is moderate (d = 0.67) based on one large study\(^{12}\)  
(incidence of higher-level executive dysfunction may be lower than processing speed and memory deficits in MS, at least as currently measured in the clinic) | Reliability is low (which is typical for tests of higher-level executive function)  
Administration time is long and procedures are cumbersome | 1. Not recommended for brief monitoring in clinical practice  
2. Not recommended for clinical trials, unless executive function is the specific target of the intervention  
3. May be useful in the context of a comprehensive clinical evaluation, especially when patient performance can be closely monitored for qualitative problem-solving approach |
| **JOLO**  
Visuospatial processing | MACFIMS  
1. Total correct | Good reliability and established validity as a visuospatial task  
Alternate form is available  
Easy for patients to understand | Lower sensitivity (d = 0.49)\(^{11,18}\) than other tasks used in MS (not necessarily a criticism of JOLO; visuospatial ability is less impaired than speed and memory) | 1. Not recommended for brief monitoring in clinical practice  
2. Recommended as a valid task of visuospatial function for comprehensive cognitive evaluations in research or practice  
3. Not recommended as an outcome in clinical trials unless visuospatial function is the target of intervention |

Abbreviations: BICAMS = Brief International Cognitive Assessment for MS; BRB = Brief Repeatable Battery; BVMT-R = Brief Visuospatial Memory Test, Revised; CLTR = Consistent Long-Term Retrieval; COWAT = Controlled Oral Word Association Test; CVLT-II = California Verbal Learning Test, second edition; D-KEFS = Delis-Kaplan Executive Function System; JOLO = Judgment of Line Orientation; LDFR = Long Delay Free Recall; LTS = Long-Term Storage; MACFIMS = Minimal Assessment of Cognitive Function in MS; MS = multiple sclerosis; MSFC = Multiple Sclerosis Functional Composite; PASAT = Paced Auditory Serial Addition Test; SDCR = Short-Delay Cued Recall; SDMT = Symbol Digit Modalities Test; SPART = Spatial Recall Test; SRT = Selective Reminding Test; TL = Total Learning.
MS, they are also robustly correlated in healthy persons (likely due to general ability), so conclusions about direct links between decline in speed, memory, or any function independent of premorbid ability or disease-related mediators (e.g., cerebral atrophy) are premature and potentially misleading (i.e., may encourage unfounded expectations, e.g., that treatment of one function leads to improvement in correlated functions) (key priorities in table 1).

We administer isolated cognitive tasks in rooms designed to minimize distractions; however, multitasking under ideal conditions may not capture patient-reported real-world deficits, especially in multitasking: the ubiquitous demand of young and middle adulthood to effectively manage multiple simultaneous goals (e.g., preparing dinner while having a conversation). Indeed, evidence suggests that cognition is more negatively affected in patients with MS (relative to controls) when performing cognitive tasks while walking (cognitive–motor dual task), and in the context of environmental noise (distraction). In addition to existing neuropsychological tools, the field should develop, validate, and utilize cognitive–cognitive and cognitive–motor dual-task paradigms to better address patient-reported multitasking deficits, which may be more sensitive for identifying real-world functional deficits, and also for predicting future decline.

**Cognitive assessment**

Cognitive processing speed is typically assessed as the amount of work performed within a time limit (e.g., number of items completed). Episodic memory is assessed as the amount of information learned and recalled (e.g., words, visual stimuli). Cognitive batteries developed for MS include tests of processing speed, memory, and other functions individually administered by trained professionals. We critically reviewed the most widely used tests (table 2), and identified the Symbol Digit Modalities Test (SDMT), Brief Visuospatial Memory Test-Revised (BVMT-R), and Selective Reminding Test or California Verbal Learning Test—Test-Revised (CVLT-R), and Selective Reminding Test or California Verbal Learning Test—Test-Revised (CVLT-R). Rates reflect the percentage of patients impaired on each test at one time, but not the number of patients impaired (1) on composite (or latent) measures of each function (performance on single tasks is affected by multiple cognitive processes, which reduces specificity), (2) in multiple functions (deficit co-occurrence), or (3) at any point in life (lifetime prevalence). It is also unknown whether prevalence of cognitive impairment has changed quantitatively or qualitatively since the approval of newer DMF.

Studies often characterize patients as cognitively intact or impaired based on overall performance across several tests measuring different cognitive functions (e.g., failure on 3 of 11 tasks), but this threshold can be met by failing speed or memory tasks alone, or a mix of speed, memory, and other tasks. This leads to heterogeneous “impaired” groups of patients with different isolated or co-occurring cognitive deficits, making interpretation of results challenging, and comparisons across studies troublesome, especially for imaging studies aiming to identify neural correlates of impairment (which likely differ across specific cognitive domains). Future work should better characterize groups as impaired in isolated or combined deficits (phenotypes, e.g., memory impaired but speed intact; speed and memory impaired), and also utilize purer measures of each cognitive domain (e.g., latent variables or composite domain scores).

**Cognitive decline**

When a patient reports a cognitive problem, he or she is describing a change in function from a previous level; however, the majority of cognitive research studies and clinical evaluations are cross-sectional. Clinicians and researchers examining single time points may miss decline that does not cross a given threshold for “cognitive impairment.” As shown (figure), patients with previous function above the 50th percentile but declining 1.5 SDs (red arrows) are not categorized as impaired, although such patients likely notice and report cognitive decline. Conversely, patients...
meeting criteria for impairment may experience varying degrees of decline before impairment (yellow arrows), thereby adding to heterogeneity of impaired groups. Clinically, baseline cognitive assessment in newly diagnosed patients would support accurate judgment of decline from previous function, which would be important for monitoring cognitive disability progression and potentially evaluating treatment efficacy (see discussion of regression-based norms,9 links.lww.com/WNL/A185; links.lww.com/WNL/A186). Finally, the field requires large prospective studies with combined cognitive and MRI assessment following newly diagnosed patients (relative to controls) over many years to better understand how cognitive decline in each domain progresses. A small example of such a study was previously performed with 44 patients over 7 years.4

Clinically meaningful change
The Expanded Disability Status Scale23 is an interval scale of physical disability in MS ranging from 0 (no disability) through 1 to 10 (0.5-point steps) reflecting disability milestones (e.g., 6.0 = unilateral gait assistance), with rubrics for clinically meaningful change.24 There is no interval scale for cognition in MS, and it has been challenging to validate cognitive test score changes indicative of clinical meaningfulness10; an obstacle to interpretation of cognitive outcomes in prevention and treatment research. Based on work with timed ambulation metrics,25 preliminary benchmarks of clinically meaningful cognitive test values15 or change in test values over time15 have been suggested (see reference 13 for SDMT), with employment status often used as an objective anchor. Ideally, we would also identify change in cognitive scores associated with cognitive difficulties in everyday life. One challenge is that patient-reported deficits and cognitive test performance are often discrepant,26 which may be due in part to discrepancies between laboratory cognitive tasks and real-life cognitive demands. Note also that adequate test–retest reliability with repeated measurements27,28 is an essential and prerequisite step when validating meaningful change on cognitive tasks.

Cognitive monitoring holds promise as a useful tool for disease surveillance. Indeed, cognitive decline is associated with MRI markers of MS disease burden (see below),14 and cognition can be impaired even before (or without) physical disability.29,30 Emerging evidence also supports the notion of a “cognitive relapse” whereby cognitive changes may be the only behavioral indicator of disease activity (i.e., without sensorimotor symptoms).15,61 As such, brief cognitive monitoring tools may identify disease activity that would otherwise go untreated, and early cognitive deficits may indicate a poor prognosis for later disability and cerebral atrophy.32

Neuroimaging and cognitive function

Neuroimaging research on cognition
Given the essential role of MRI in MS diagnosis and disease surveillance, the field of MS is at the forefront of novel and innovative MRI technology, which provides multiple tools for investigating cognitive deficits due to MS.14 Cognitive deficits were linked to greater lesion load in early research,16 and subsequent work shows the importance of white matter lesion location,17 microstructural injury,18 gray matter lesions,19 cortical20 and subcortical21–23 gray matter brain atrophy, and discrepant patterns of cerebral activation with fMRI.24 Advances in ultra-high-field MRI,33 myelin34 and molecular35 imaging, imaging of demyelination and remyelination,26 and nonconventional MRI techniques to assess microstructural cerebral changes27 will provide even more ways to investigate MS-related cognitive deficits. A challenging but essential next step is to integrate rich multimodality imaging data into testable and biologically informed models of disease-related cognitive deficits, utilizing unique strengths of each imaging approach. This aspirational but critical goal will require collaboration among experts in imaging modalities, neuroscience, and cognition. Such work will inform development of biologically plausible approaches to cognitive rehabilitation (key priorities in table 3).

Neural bases of cognitive deficits
Neuroanatomical correlates of cognitive deficits exist (e.g., thalamus21–23), but it is unclear whether such correlates (1) directly underlie deficits, or (2) are reliable proxies of overall (or other) cerebral damage, which mediate links to cognition. For instance, the thalamus is highly susceptible to retrograde degeneration,35 and has better scan-to-scan reliability than other structures,36 perhaps making thalamic volume a good summary measure of disease burden across patients with variable CNS damage, even if thalamic change does not directly underlie a specific deficit (e.g., memory). Although reliable correlates of cognition may be useful for prediction (discussed below), knowledge of precise neural bases is important for identifying therapeutic targets (e.g., neurotransmitter systems) for treatment of specific deficits. We need large prospective longitudinal studies with multimodality neuroimaging to carefully document temporal correlations of specific emerging cognitive deficits with changes in specific brain structures and functions, thereby informing advanced models of disease-related deficits, which will help identify therapeutic targets. For instance, longitudinal work may help substantiate cross-sectional links between memory deficits and hippocampal changes: atrophy on MRI,28 lesions on double inversion recovery,29 glutamate concentration on magnetic resonance spectroscopy,30 abnormal activation and functional connectivity on fMRI,31 and demyelination and synaptic loss on histology.32 (See links.lww.com/WNL/A185 and links.lww.com/WNL/A186 for information how segmentation of hippocampal subfields28,33,34 and thalamic nuclei35 may advance understanding of cognitive deficits.)

The search for specific neural correlates of distinct cognitive functions is encumbered by the imprecision of (1) heterogeneously impaired groups and (2) single cognitive tasks with multiple processing demands. MS research may be informed by lifespan research using factor analytic techniques to derive purer latent measures of speed, memory, and other
functions. This approach to behavioral assessment would complement recent factor analytic MRI analyses identifying nonrandom patterns of regional cortical thinning due to MS.

**Quantifying risk for future cognitive decline**

Longitudinal studies have linked baseline MRI to risk for cognitive decline, including T2 lesion volume, cerebral atrophy, microstructural damage, and cortical lesions. Although labor-intensive and expensive, we need more prospective multimodality neuroimaging studies with large representative samples to create algorithms of risk for cognitive decline. Combined with demographic, reserve, and clinical variables, such algorithms should be assessed for specificity and sensitivity in confirmation samples, thereby evaluating clinical utility. Accurate algorithms may aid in early treatment decisions (e.g., aggressiveness of DMT), and advance research and practice of early cognitive intervention. Note that clinical feasibility of MRI-informed risk algorithms is currently limited by the feasibility of employing advanced scanning sequences during clinical MRIs, as well as access to the specialty skills needed to derive quantitative MRI metrics. Indeed, even relatively basic metrics of total cerebral atrophy and T2 lesion volume are rarely available to clinicians. Different groups are working to bridge this gap by providing cerebral atrophy analysis services to clinicians, with the goal of incorporating atrophy consideration into standard clinical care.

**Functional neuroimaging**

fMRI provides a proxy of brain function, which may help explain cognitive deficits due to MS. There has been growing concern, however, about poor reproducibility of fMRI results, including contradictory findings of cross-sectional MS studies (e.g., memory deficits linked to lower and higher functional connectivity), and a lack of sufficiently large longitudinal studies. Inconsistencies may be due in part to differences in data collection and analysis approaches, statistical power, and heterogeneity across patient samples. There may be ways to improve the approach and consequent value of fMRI. First, reproducibility should be established with large collaborative multicenter studies using uniform methods, with out-of-sample replication of results. Next, following a model of “registered reports” in cognitive neuroscience and trial registration (clinicaltrials.gov), an online repository for posting specific study methods and hypotheses prior to data collection may be helpful. This would yield 2 levels of evidence: exploratory and confirmatory. Although exploratory research is important, certified confirmatory results would improve confidence in (1) individual study results, (2) the fMRI method in general, and (3) any future clinical utility of fMRI techniques.

Finally, it is integral that fMRI and behavioral findings be incorporated into working theoretical models of cognitive dysfunction, which will provide the scientific context for a priori hypothesis generation and testing, and model refinement. As discussed in the context of the functional reorganization hypothesis (see references 42 and e39), such models must move beyond overly simplistic views of large network changes as either maladaptive or compensatory, and instead create more dynamic, biologically plausible models informed by multimodality neuroimaging methods.

**Treatment and prevention of cognitive impairment**

**Cognitive rehabilitation**

Improved understanding of MS cognitive deficits will inform the nascent field of cognitive rehabilitation, which seeks to restore cognitive functioning (often through intensive cognitive training programs) or teach compensatory strategies to attenuate the deleterious effect of refractory cognitive deficits on quality of life. Efficacy for such interventions in MS is currently low, inconclusive, or preliminary, as concluded by (1) a Cochrane review of 20 randomized or quasi-randomized controlled trials (RCTs) of behavioral interventions to improve cognition in MS (data search up to July 2013), (2) a separate systematic review of 33 original intervention studies (including nonrandomized trials, search up to January 2014), and (3) a Cochrane review of 15 intervention trials specifically targeting memory (search up to June 2015). In addition to small sample sizes, quality of studies was limited by several methodologic flaws (e.g., poor blinding, unvalidated outcomes). Note also that a Cochrane review of symptomatic pharmaceutical treatment of memory deficits in MS revealed only 7 RCTs (data search up to June 2015).
Observations and implications

1. Researchers should aim to make the challenging but important distinction between (1) neuroimaging correlates and (2) neural bases of cognitive deficits, as these may not overlap. Strong correlates of cognition may ultimately be useful for prediction, but knowledge of precise neurobiological bases of deficits will be important for identifying specific therapeutic targets (e.g., neurotransmitter systems). Multimodal imaging and experimental paradigms incorporated into biologically plausible models of MS-related deficits may facilitate discovery of precise neurobiological bases/treatment targets.

2. Investigation of isolated cognitive constructs (e.g., memory) rather than heterogeneous composites of multiple cognitive domains is necessary to advance models of precise neuroanatomical and neurophysiologic bases of disease-related cognitive deficits. The use of domain-specific composite scores or latent variables would be useful.

3. Development of multivariate models (incorporating demographic, neuroimaging, and clinical variables) to better predict decline in separate cognitive domains is needed to develop clinically useful risk algorithms. The importance of risk and protective factors is not a new idea in itself; however, the development of clinically useful algorithmic tools to predict decline and inform treatment would be novel and important for the field.

4. We require clinically feasible tools or services for the quantification of disease burden from MRIs in the MS clinic, including T2 lesion volume and cerebral atrophy. With training, cortical lesions and cortical thinning may also be quantified. Once available, such metrics may improve disease surveillance and clinical decision-making. Also, quantification of lesion volume and cerebral atrophy from clinical MRIs within large representative samples would yield rich datasets for new research opportunities.

5. We need to address the reproducibility problem in fMRI work, perhaps by promoting multicenter collaborations with comparable methods, but also by registering a priori hypotheses. This would yield 2 levels of fMRI evidence: exploratory and confirmatory. Exploratory evidence is important for novel discoveries, but confirmatory evidence is needed for model building and validation of fMRI as a valid research and clinical tool.

Abbreviation: MS = multiple sclerosis.

Table 3 Key priorities for neuroimaging investigations of cognitive deficits

<table>
<thead>
<tr>
<th>Priority</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Researchers should aim to make the challenging but important distinction between (1) neuroimaging correlates and (2) neural bases of cognitive deficits, as these may not overlap. Strong correlates of cognition may ultimately be useful for prediction, but knowledge of precise neurobiological bases of deficits will be important for identifying specific therapeutic targets (e.g., neurotransmitter systems). Multimodal imaging and experimental paradigms incorporated into biologically plausible models of MS-related deficits may facilitate discovery of precise neurobiological bases/treatment targets.</td>
</tr>
<tr>
<td>2.</td>
<td>Investigation of isolated cognitive constructs (e.g., memory) rather than heterogeneous composites of multiple cognitive domains is necessary to advance models of precise neuroanatomical and neurophysiologic bases of disease-related cognitive deficits. The use of domain-specific composite scores or latent variables would be useful.</td>
</tr>
<tr>
<td>3.</td>
<td>Development of multivariate models (incorporating demographic, neuroimaging, and clinical variables) to better predict decline in separate cognitive domains is needed to develop clinically useful risk algorithms. The importance of risk and protective factors is not a new idea in itself; however, the development of clinically useful algorithmic tools to predict decline and inform treatment would be novel and important for the field.</td>
</tr>
<tr>
<td>4.</td>
<td>We require clinically feasible tools or services for the quantification of disease burden from MRIs in the MS clinic, including T2 lesion volume and cerebral atrophy. With training, cortical lesions and cortical thinning may also be quantified. Once available, such metrics may improve disease surveillance and clinical decision-making. Also, quantification of lesion volume and cerebral atrophy from clinical MRIs within large representative samples would yield rich datasets for new research opportunities.</td>
</tr>
<tr>
<td>5.</td>
<td>We need to address the reproducibility problem in fMRI work, perhaps by promoting multicenter collaborations with comparable methods, but also by registering a priori hypotheses. This would yield 2 levels of fMRI evidence: exploratory and confirmatory. Exploratory evidence is important for novel discoveries, but confirmatory evidence is needed for model building and validation of fMRI as a valid research and clinical tool.</td>
</tr>
</tbody>
</table>

2013), with no evidence for efficacy (key priorities in table 4).

We require a science of cognitive rehabilitation capable of yielding high levels of evidence. Toward this end, we must develop theoretical models of MS-related cognitive dysfunction and identify mechanisms of action to treat deficits, followed by large RCTs with validated outcomes. This rigorous pathway to high-level evidence is required by regulatory agencies before approving clinical use of new agents (e.g., DMTs), and perhaps similar regulation should be considered for cognitive rehabilitation. There is less industry funding to perform large labor-intensive RCTs of cognitive rehabilitation interventions; however, grant funding may be sought for multicenter collaborations, which may increase sample size and representativeness. Finally, standards for a priori reporting of methods must be upheld for cognitive rehabilitation RCTs, including greater transparency for outcomes (e.g., specific scores on specific tests registered on clinicaltrials.gov, rather than nonspecific references to “cognition” or “memory tests”). Cognitive rehabilitation researchers are directed to Simons and colleagues for a thorough discussion of essential guidelines for the conduct of high-quality cognitive intervention trials.

Theoretical models

As proposed for the field of rehabilitation generally, we must identify mechanisms of action and active ingredients of cognitive rehabilitation interventions. For instance, if we discover that a training program improves memory by strengthening hippocampal structure and function, we may also consider other interventions targeting hippocampal health; e.g., aerobic exercise, intellectual enrichment, glucose control, and stress management. Availability of alternative treatments is important when considering clinical feasibility, as time-consuming, expensive cognitive training programs may be feasible for some patients, whereas other patients will be more likely to engage in alternative approaches (e.g., exercise). A model should also distinguish different types of rehabilitation approaches and goals: (1) restorative interventions aiming to bolster underlying neurophysiologic bases of memory, vs (2) compensatory approaches aiming to improve memory through strategies (e.g., mnemonics) or aids (e.g., diaries; see discussion). This has implications for trial outcomes: e.g., given that compensatory strategies only work when patients use them (like a cane for walking), we should not assess treatment efficacy with standardized memory tests that prevent or encumber strategy usage (as standardized administrations may prevent compensatory strategy usage).

Secondary structural and functional neuroimaging outcomes may help identify mechanisms of action for interventions and identify markers of capacity to benefit from interventions (responders vs nonresponders). This may help hone treatments and inform new treatment development, and identify subgroups with residual capacity to respond to interventions. Like cognitive outcomes, specific neuroimaging outcomes should be stated (registered) prior to data collection, or otherwise be described as exploratory. Finally, the goal of enhancing brain structure and function in biologically plausible and lasting ways likely requires greater “doses” (duration and intensity) of interventions than is typically performed, and perhaps combined therapies with potentially synergistic effects (e.g., training plus neurostimulation).

Primary prevention

We should also promote primary prevention of cognitive decline, in part through interventions and healthy lifestyles that promote brain maintenance. Candidate modifiable lifestyle factors to build or maintain brain reserve include...
physical exercise,52 mentally active lifestyles (cognitive reserve),38,53 management of cardiovascular risk factors54 and other comorbidities,55 smoking cessation,56 and stress management.57 Research should strive to understand mechanisms of action for protective factors (e.g., moderating MS disease activity vs maintaining brain volume in a non-disease-specific way). Studies typically examine few risk or protective factors at a time, but we need larger studies of numerous factors in the same large cohort to understand (1) whether and to what extent each risk or protective factor makes independent contributions to an outcome, (2) mediating or interactive effects among different factors on an outcome, and (3) how (1) and (2) differ across outcomes (e.g., speed, memory). We also need to raise the level of evidence (e.g., RCTs) linking lifestyle factors to cognition, and explore additional variables (e.g., diet57,49). (DMTs as protective factors are discussed in links.lww.com/WNL/A185 and links.lww.com/WNL/A186.)

Holistic approach to cognitive rehabilitation

Effective intervention will require patient understanding, motivation, and compliance, as well as willingness by clinicians to consider the unique circumstances of individual patients (e.g., family/social support, comorbidities, goals). Ideal holistic rehabilitation approaches consider cognitive, emotional, and psychosocial aspects of each patient’s life.58 Patient education may promote metacognition and active participation. For instance, therapeutic feedback after neuropsychological assessments may support understanding of one’s cognitive profile, and aid patients in finding ways to maximize cognitive strengths and minimize weaknesses in daily life. Education on cognitive deficits and factors affecting cognition (e.g., sleep, medications, mood, fatigue) may promote active participation and a positive sense of agency among patients.59 Indeed, preliminary results on structured metacognitive training with peer support are encouraging.69 To advance this holistic approach, research is also needed to better understand shared neural bases for mood and cognitive dysfunction,66 which may yield new directions for cognitive treatments. Finally, treatment will likely be most effective when tailored to a patient’s specific deficit in the context of his or her degree of spared cognition and cerebral reserve. For instance, a memory deficit due to diffuse white matter lesions may require a different treatment approach than a deficit secondary to a focal hippocampal lesion. Consideration of distinct subtypes of deficits requiring distinct treatment approaches (i.e., precision medicine) is a key challenge and opportunity for the future.

Discussion

The literature on cognition in persons with MS has grown exponentially over the last 25 years, and cognitive dysfunction is now recognized as a core symptom of MS. Herein we discussed obstacles and challenges for the field and made recommendations for moving research forward. The next 25 years will bring redoubled collaboration across centers and areas of expertise, and utilize advances in neuroimaging, genetics/epigenetics, and validation of cognitive endpoints. Collaborations and advanced methods are invaluable, but the real science of cognition and cognitive rehabilitation in MS will rely on multidisciplinary collaborations toward development of biologically based theoretical models of

**Table 4: Key priorities for treatment and prevention of cognitive deficits**

1. Rigorous research designs are required to produce higher levels of evidence for cognitive rehabilitation research, including multicenter double-blind randomized controlled trials, with clear and specific a priori outcomes. Essential guidelines for the conduct of high-quality cognitive intervention trials have been discussed by Simons and colleagues.47 Adherence to these recommendations by investigators, post hoc reviewers, and journal editors will greatly improve trial quality and the science of MS cognitive rehabilitation.

2. We require theoretical frameworks to build a science of cognitive rehabilitation in MS, with biologically plausible mechanisms of action, and clear delineation of rehabilitation approaches (e.g., restorative vs compensatory). Models will be informed by greater understanding of neural bases of cognitive function. The goal of changing brain structure and function in plausible and lasting ways likely requires greater “doses” (duration and intensity) of interventions, or combined therapies with synergistic effects.

3. Structural and functional neuroimaging outcomes may explore mechanisms of action for interventions; however, when used as trial endpoints, specific hypotheses of treatment effects on neuroimaging outcomes should be stated a priori. Otherwise, such outcomes should be considered exploratory evidence requiring confirmation in a subsequent trial.

4. Observational research has identified candidate modifiable lifestyle factors that may protect against cognitive decline, including mental activity, physical exercise, and stress management. Research is needed to (1) establish causal relationships between protective factors and outcomes, including utilization of RCTs, (2) examine the unique and possibly differential contributions of each protective (or risk) factor to each individual cognitive outcome (speed, memory), as well as potential synergistic effects (interactions) among protective factors, and (3) explore mechanisms of action for different protective (or risk) factors (e.g., moderating MS disease activity itself vs building/preserving reserve).

5. Any promising cognitive interventions must be implemented in the context of a patient’s life. A holistic approach to match treatment with a patient’s unique circumstances, priorities, and abilities is necessary in future translational research and practice. That is, interventions will only be effective if they are clinically feasible for an individual patient.

6. Cognitive rehabilitation interventions may be most effective when tailored to a patient’s specific deficit, which may differ within the same cognitive domain (e.g., memory deficit secondary to diffuse white matter lesions vs focal hippocampal lesion). Consideration of distinct subtypes/etiologies of deficits requiring distinct treatment approaches (i.e., precision medicine) is both a challenge and a potential opportunity for rehabilitation.

Abbreviations: MS = multiple sclerosis; RCT = randomized controlled trial.
cognition capable of empirical validation and evidence-based refinement, providing the necessary context for effective treatment discovery.

**Author contributions**
J.F.S. conceptualized and initiated the project and coordinated efforts by coauthors. All authors made intellectual contributions, and read, revised, and approved the final manuscript.

**Study funding**
No targeted funding reported.

**Disclosure**
J. Sumowski reports no disclosures relevant to the manuscript. R. Benedict receives research support from Acorda, Novartis, Genzyme, Biogen, and Mallinckrodt Pharmaceuticals; is on the speakers’ bureau for EMD Serono; consults for Biogen, Genentech, Genzyme, Novartis, AbbVie, Roche, and Sanofi; and receives royalties from Psychological Assessment Resources. C. Enzinger has received funding for travel, speaker honoraria, or research funding from Biogen Idec, Bayer Schering Pharma, Merck Serono, Novartis, Genzyme, and Teva Pharmaceutical Industries Ltd./Sanofi-Aventis. He is serving on scientific advisory boards for Bayer Schering Pharma, Biogen Idec, Genzyme, Merck Serono, Novartis, and Teva Pharmaceutical Industries Ltd./Sanofi-Aventis; and as an academic editor for *PLoS One*. M. Filippi is Editor-in-Chief of the *Journal of Neurology*; serves on scientific advisory boards for Teva Pharmaceutical Industries Ltd.; has received funding for travel from Bayer Schering Pharma, Biogen Idec, Merck Serono, and Teva Pharmaceutical Industries Ltd.; serves as a consultant to Bayer Schering Pharma, Biogen Idec, Merck Serono, Novartis, Peggen Corporation, and Teva Pharmaceutical Industries Ltd.; serves on speakers’ bureaus for Bayer Schering Pharma, Biogen Idec, Merck Serono, and Teva Pharmaceutical Industries Ltd.; and receives research support from Bayer Schering Pharma, Biogen Idec, Novartis, Merck Serono, Teva Pharmaceutical Industries Ltd., CurePSP, and the Jacques and Gloria Gossweiler Foundation. J. Geurts serves on the editorial boards of *MS Journal*, *BMC Neurology*, *MS International*, and *Neurology*, the Scientific Advisory Board of the Dutch MS Research Foundation and MS Academy, Merck-Serono, and has served as a consultant for Merck-Serono, Biogen Idec, Novartis, Genzyme, and Teva Pharmaceuticals. P. Hamalainen reports no disclosures relevant to the manuscript. H. Hulst serves as a consultant for Genzyme, Merck-Serono, Teva Pharmaceuticals, and Novartis. M. Inglese has received research grants from Novartis and Teva Neuroscience. V. Leavitt reports no disclosures relevant to the manuscript. M. Rocca has received speakers honoraria from Biogen Idec, Excemed, and Novartis. E. Rosti-Otajari reports no disclosures relevant to the manuscript. S. Rao has received royalties from the Cleveland Clinic for licensing the tablet-based Processing Speed Test. He has received honoraria, royalties, consulting fees, or research funding from Biogen, Genzyme, Novartis, and the CHDI Foundation. Go to Neurology.org/N for full disclosures.

Received February 28, 2017. Accepted in final form October 10, 2017.

**References**