Cognitive Measures Used in Adults with Multiple Sclerosis:

A Systematic Review

Hannah Elwick[1], Gogem Topcu[1], Christopher Martin Allen[1], Avril Drummond[2], Nikos Evangelou[1] and Roshan das Nair[1,3]

[1]Mental Health and Clinical Neurosciences Academic Unit, School of Medicine, University of Nottingham, UK

[2]School of Health Sciences, University of Nottingham, UK

[3]Institute of Mental Health, Nottinghamshire Healthcare NHS Foundation Trust, UK

Correspondence concerning this article should be addressed to Roshan das Nair, Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU

Contact: roshan.dasnair@nottingham.ac.uk

Word Count: 4,832

Abstract

Cognitive problems are common in people with Multiple Sclerosis (MS), and researchers and clinicians have used a vast array of measures to assess cognition. Our aim was to systematically identify cognitive measures routinely used in MS research, and outline their different uses. Previous recommendations of cognitive measures to use in MS have relied on expert consensus approaches. We believe this systematic review is a starting point for an evidence-based approach to recommending cognitive tests for use with people with MS.

We systematically searched electronic databases using relevant search terms for studies that assessed cognitive functioning in MS (last search in February 2020). From 11,854 abstracts retrieved, based on title and abstract review, 2,563 remained. Data were extracted from 1,526 studies.

Studies used 5,665 measures of cognition, with 316,053 people with MS. Substitutional style tests, serial addition tests, and word list learning tests were the most commonly used individual tests, and the Brief Repeatable Battery of Neuropsychological Tests was the most commonly used battery. Some of the most frequently used measures were potentially inappropriate due to measuring irrelevant domains of cognition, and issues with sensitivity. Further research is needed to ascertain the psychometric properties, and acceptability of measures for people with MS.

Key words: Multiple Sclerosis, cognitive impairment, systematic review, assessment, cognition

Cognitive Measures Used in Multiple Sclerosis: A Systematic Review

Multiple Sclerosis (MS) is the most common chronic disorder of the central nervous system, in which demyelination of nerve cells in the brain and spinal cord can lead to physical and psychological symptoms (Compston & Coles, 2008). In 2020, approximately 2.8 million people had MS across the world (Walton et al., 2020). There are several phenotypes of MS, including Relapsing Remitting MS (RRMS) characterised by relapses and remissions of symptoms, Primary Progressive MS (PPMS) in which there is a slow, unrelenting progression of disability from the onset, and Secondary Progressive MS (SPMS) in which the majority of people with RRMS eventually develop slowly progressively disability (Lublin et al., 2014). Although there is no cure, there are several disease modifying treatments available. In addition, symptoms of MS can be managed using medical treatments, cognitive rehabilitation, physiotherapeutic treatments, lifestyle changes, among others (Samkoff & Goodman, 2011).

One of the common symptoms of MS is cognitive impairment, which is reported to occur in 43-70% of people with MS, (DiGiuseppe et al., 2018; Rao et al., 1991). Cognitive impairment can have an adverse effect on many aspects of an individual's life, including the ability to work and socialise (Rao, Leo, Ellington, et al., 1991), treatment adherence (Bruce et al., 2010), and overall quality of life (Glanz et al., 2010). Predominantly impaired cognitive domains are memory and learning, (including immediate and delayed recall and recognition), complex attention including information processing speed, and executive functions (Chiaravalloti & DeLuca, 2008; Ferreira, 2010; Prakash et al., 2008). Recent research has also identified deficits of prospective memory in pwMS, which is associated with other impaired domains including attention and executive functions (Rouleau et al., 2018). Social cognition including Theory of Mind and emotion recognition may also be impaired, however research has been limited (Cotter et al., 2016). Cognitive impairment can worsen across the progression of MS (Achiron et al., 2005), and is unlikely to remit (Amato et al., 2006). Cognition is associated with physical and psychological symptoms in MS (Amato et al., 2019) with poor performance on cognitive assessments frequently found to be significantly associated with the common symptoms of fatigue and depression (Hansen & Lautenbacher, 2017).

Poor performance on tests of processing speed, visual and verbal memory is associated with difficulties carrying out everyday life tasks (Goverover & DeLuca, 2015), and in particular, slowed information processing speed is associated with poor outcomes in activities of daily living and quality of life (Costa et al., 2017). Executive function is considered a higher-order domain, involving subdomains of decision-making and planning, and thus impairment in this domain may have a negative impact on vocational and other activities for people with MS (Drew et al., 2008). A recent recommendation has supported domain-specific management of cognitive problems as opposed to more generalised approaches (DeLuca et al., 2020), and novel and improved methods of rehabilitation require appropriate outcome measures to determine their efficacy. Valid and reliable cognitive tests can additionally be used to identify cognitive issues at an early stage of MS, to assess progression of cognitive impairment.

Reviews by neurologists may not identify specific details of cognitive functioning and their significance (Romero et al., 2015), and appropriate, standardised cognitive measurement tools for use with people with MS are therefore required to detect changes to cognitive function accurately. A recent narrative review identified a need for incorporating neuropsychological assessment into routine care for people with MS (Sumowski et al., 2018). However, due to the high prevalence of cognitive impairment in MS and limited resources in clinics, it may not always be feasible to conduct a full neuropsychological assessment for all patients at routine clinic visits. Therefore, there is a need for both comprehensive test batteries and shorter tests that can reduce the resources needed, including the level of expertise required to use and interpret results. Tests also need to be acceptable to clinicians and patients.

A survey of North American clinical neuropsychologists (Rabin et al., 2005) reported that the most frequently used measures across different clinical populations were Wechsler Adult Intelligence Scale (WAIS), the Wechsler Memory Scale (WMS), the Trail Making Test (TMT), and the California Verbal Learning Test (CVLT). However, as this survey did not specify tests used with different neuropsychological conditions, it is not possible to determine whether this reflects what psychologists are using with people with MS in particular. A more recent UK-wide survey (Klein et al., 2018), which specifically focused on cognitive assessments for people with MS, found that the most commonly used

measures by healthcare professionals working clinically with people with MS (including neuropsychologists, neurologists, and occupational therapists) are the Montreal Cognitive Assessment (MoCA) and the Addenbrooke's Cognitive Examination (ACE-R). Generic measures may measure cognitive domains not commonly impacted in MS, such as orientation. Conversely, they might not include measurements of commonly impacted domains, such as the lack of a measure of information processing speed in the MoCA. These studies show that generic measures are prevalently used in clinical practice, and it is unclear why some measures are chosen over others.

In addition to generic measures that are not specific to MS (such as the MoCA), a number of MS-specific batteries of cognitive measures have been developed, and both comprehensive and brief screening batteries exist. The Brief Repeatable Battery of Neuropsychological Tests (BRB-N) was one of the first batteries developed specially for use with people with MS, comprising a selection of tasks from a longer battery that were shown to be sensitive to cognitive impairment (Rao et al., 1991). An expert panel then developed the Minimal Assessment of Cognitive Function in MS (MACFIMS, Benedict et al., 2003), a comprehensive battery of tests which expanded on the BRB-N to include tests with better established validity. Although the 90-minute administration time may make it impractical to use in some research and routine MS clinical practices, an abbreviated screening version, the aMACFIMS, is available. A recommendation of an assessment with shorter administration time of 15 minutes emerged from an expert committee rating individual cognitive tests (Langdon et al., 2012), however the full methodology and results have not been published. More recently, recommendations have promoted the use of the proprietary Symbol Digit Modalities Test (SDMT) as a brief, sensitive measure processing speed impairment in people with MS, which is one of the most predominantly impaired domains (Kalb et al., 2018). . The original administration of the SDMT consists of both written and verbal versions of the same task, however as dexterity problems can impact the results of the written task, the verbal administration is thought to be most appropriate, and more frequently used in MS populations (Benedict et al., 2017), and has been found to be valid and reliable (Jaywant et al., 2018).

The recommendations for cognitive measures, screens and batteries discussed above have used consensus approaches. These used expert panels comprising neuropsychologists, neurologists, clinical psychologists, researchers and Patient and Public Involvement to make recommendations, including that of full batteries (MACFIMS, Benedict et al., 2003), and more brief assessment (Kalb et al., 2018; Langdon et al., 2011). All approaches stated the need for full neuropsychological batteries to be administered for further insights if needed but highlighted the necessity for screening tools at a minimum, to promote routine assessment.

Following on from consensus recommendations, it would be useful to know which cognitive assessments are being used in research and clinical practice with people with MS. This might influence future consensus recommendations. There are ongoing concerns whether MS-specific batteries and tests (i) are able to evaluate the range of cognitive domains impaired in people with MS, particularly executive function, processing speed and memory (Hansen & Lautenbacher, 2017), and (ii) may lack ecological validity (Korakas & Tsolaki, 2016). Furthermore, in some healthcare and research settings, the use of proprietary measures may not be practical due to the financial implications of using such measures.

The aim of this review was to systematically compile and describe the neuropsychological measures that have been used in research with people with MS, and to examine the modality of administration and contexts in which they have been used. This review is the first step in determining an evidence-based approach to selecting measures for use in cognitive assessment with people with MS and complements existing consensus-based approaches.

Methods

The study protocol was prospectively registered with PROSPERO International Prospective Register of Systematic Reviews (Registration number CRD42018103384)

Search Strategy and Selection Criteria

The systematic review followed the guidelines for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, Moher, 2009). We searched the following databases: MEDLINE, EMBASE, CINAHL, PsycINFO and Web of Science, for studies published in English from inception to February 2020 and using the relevant key words including, but not limited to: Multiple Sclerosis, Cogniti*, Neuropsych*, Screen, Assessment, Test, Outcome. The search strategy was developed for all databases by one reviewer (HE) and independently reviewed by four members of the team (RdN, GT, AD and NE). The database searches were limited to adult, human studies, and texts in the English language.

Studies were included if they used psychometric measures to assess cognitive functioning in MS or screen participants into a study. We included measures that were performed face-to-face, and measures which were administered electronically or via telephone. The target population for studies was people (aged 18 and over) who had been diagnosed with any form of MS. We included individual measures, screening tests, and batteries of cognition, measures of single and multiple domains of cognition, and both subjective and objective measures of cognition. There were no constraints for the setting in which the cognitive assessment took place.

Initial title and abstract screening were completed by one reviewer (HE), and full-texts were divided between two reviewers for data extraction given the volume of the literature (HE & CA). Any disagreements were resolved by discussion and arbitration with a third reviewer (GT, RdN, AD or NE). Studies were excluded based on the hierarchy shown in Figure 1.

A formal quality assessment was not conducted due to the absence of a single appropriate assessment for all studies, as well as the impracticality of undertaking this given the considerable number of studies included in this review. As we aimed to capture the breadth of measures being used as well as their frequency of r use in a variety of contexts, including a range of research methodologies, the quality of the studies would have been difficult to compare, and this would therefore been of limited use. However, we

did capture pertinent issues, including whether studies specified the names of measures used, or how measures were administered.

Despite restricting the searches to adult and English language papers, some papers that did not meet these criteria were still present in citations extracted and were therefore removed through screening in the iterative steps of the title and abstract review, and the full-text review. In cases where two or more publications reported results of the same dataset, only the most recent study was included for data extraction.

Data Extraction and Analysis

Our main interest was to identify the measures of cognition being used with people with MS, including how the measure was used (e.g., as a screening measure, or as an independent or dependent variable), and the modality the measure was administered in (e.g., face-to-face, electronically). We also extracted information about the participants' type of MS, the country in which the study took place if a single centre study, or if it was a multicentre study, how many respondents completed the measure, and if it was a patient or proxy-reported measure. For the purpose of analysis, the different measures were characterised into a number of classes for tests so that we could understand what measures were used most often, despite the names and ascriptions given. Examples of the classes used are serial addition tests in which respondents add up numbers as they are presented (e.g., the Paced Auditory Serial Addition Test (PASAT)), substitutional tests, where respondents are presented with an array of stimuli (usually digits, symbols or letters) and asked to match them correctly with similar stimuli using a key (e.g., SDMT). The classes, named examples of these, and a description of the class are described in Table 1. These classes were determined through an assessment of how the assessments were described conceptually in relevant literature, or as described within the study. Once all individual assessments used in studies were classified, total numbers and frequency of use was calculated for each class, or multi-domain test including batteries.

Results

The searches produced a total of 11,854 papers to be screened, from which data from 1,526 studies were extracted for analysis. The PRISMA flowchart is shown in Figure 1. Full-text review was conducted by two reviewers on a subsample of randomly chosen (n=50) manuscripts, to ensure uniformity of data extraction. We found a high level of agreement qualitatively.

The studies were conducted in 52 countries (50% conducted in Europe, and 39% from North America), and included 29 multicentre studies. Included studies collected data from 316,053 people with MS. All phenotypes of MS were reflected in the studies; most of the studies (66%) used a mixed sample of people with MS, and 29% of studies enrolled people with Relapsing Remitting MS (RRMS). Only 4% of the studies specifically enrolled people with progressive forms of MS.

There was a total of 5,665 cognitive measures used. Of these measures, 93% (n=5,260) were 'objective' measures (for example, memory or processing speed tests administered by a researcher or clinician), and 7% (n=370) were 'subjective' (for example, self-report or proxy questionnaires). Thirty-five studies did not specify the names or provide sufficient description of the measures or procedures they used to classify them. However, some of these unnamed measures did report whether they were objective or subjective and were included in the overall counts. Approximately 18% (n=912) were named batteries or screening tests of cognitive functions consisting or two or more individual assessments (for example, the MoCA is a brief screening test but includes a trail making task, a verbal fluency task, word list learning among others).

Individual Tests of Cognition

There were 504 different individual measures used, however many of these overlapped considerably. For example, verbal fluency tests that used a variety of stimuli, such as different letters of the alphabet, were given different names, including the Controlled Oral Word Association Test (COWAT), F-A-S test, or letter fluency tests.

The most frequently used measures were the substitution-based tests (12% of all individual tests) such as the SDMT, and serial addition tests (11%) such as the PASAT, followed by tests which require participants to learn a list of words and recall them (immediately and/or after a delay; 7.8%) and verbal fluency tests (7.5%), such as the COWAT, in which participants are asked to name as many words that belong to a certain category (such as animals), or words beginning with a specific letter of the alphabet. The domains of cognitive impairment measured by the most frequently used measures relate to those experienced in MS, including information processing speed, attention, executive function, and aspects of memory including working and short-term memory.

Multi-Domain Tests and Batteries

Measures were operationally defined as a named collection of tests if they were a named screening test (e.g., MoCA) or a battery (e.g., BRB-N) that included one or more measure that is considered an individual test. Sixty-one different multi-domain neuropsychological tests or batteries were reported in the included studies, and these were used 913 times in included studies. The most frequently used multi-domain test used was the BRB-N, followed by the Mini–Mental State Examination (MMSE) and the WAIS (See Table 2). The Multiple Sclerosis Functional Composite (MSFC) only includes one neuropsychological measure (the PASAT) and was therefore included as an individual measure rather than a battery. The MoCA was frequently used to screen participants into studies (37%), or to report on cognitive functioning or between group differences (38%). However, 25% of the studies using the MMSE used it as a dependent or independent variable.

Consensus recommendations have advised the use of the MACFIMS and BICAMS as cognitive batteries for use with people with MS. The MACFIMS was recommended by consensus in 2003 and the BICAMS in 2011. The BRB-N was developed in 1990 and was used 10 times before the recommendation of the MACFIMS in 2003. Since then, the BRB-N has continued to be popular, with 217 studies using it, compared to 72 studies using MACFIMS. Even since the publication of BICAMS, the BRB-N has remained popular, being used in 151 studies, whereas the BICAMS was used in 62 studies.

Subjective Measures of Cognition

A breakdown of the subjective, patient (or proxy) reported measures and their frequency of use is shown in Table 3. There was a total of 370 uses of named subjective measures in the studies, 87% were completed by the people with MS, and 11% were proxy reported (8% were carer/family reported, and 3% were clinician reported). Eighteen studies did not specify who completed the subjective measure of cognition. The most frequently used subjective measures were the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ), the Perceived Deficits Questionnaire (PDQ), and Fatigue Scale for Motor and Cognitive Functions (FSMC). There were 29 studies that used a self-report measure that was not reported by name, or not a formally named measure (for example, a Likert scale or visual analogue scale of cognitive functioning not attributed to a specific measure).

Uses of Cognitive Measures

Cognitive measures were used for myriad reasons, and these were grouped into four categories (see Table 4). Many of the measures were used for two or more purposes in the study (for example to screen participants into the study, and also as a dependent variable). Cognitive measures were predominantly used to report on cognitive functioning or as a between-groups measure in people with MS, and as a dependent variable.

There were 27 studies that used measures in ways that did not fall into these predefined categories and these were recorded separately. Twelve were recorded as using the cognitive measures in a training programme for cognitive rehabilitation for people with MS, and other uses included matching healthy controls with people with MS, power calculations, and to derive the minimal clinically important difference. Some studies did not report on the cognitive measures they administered.

Administration of Cognitive Measures

Where available, we extracted data on how the cognitive measures were administered. Approximately half of the 5,670 measures used in the studies were not described in enough detail to extrapolate how they were administered. Of the remaining 2,680 measures, the majority were administered face-to-face (87%) by a healthcare professional, technician or researcher, and 11% used a computerised

administration. Remaining modalities consisted of postal and telephone administrations. The 53 measures completed by post were all subjective, self or proxy-reported measures, however, the 12 telephone administered measures were a mix of subjective and objective measures, including verbal fluency tests, word list learning and digit span tests.

Discussion

We identified a wide range of different individual cognitive tasks, screens and batteries that cover a number of cognitive domains, and despite recommendations for the use of specific measures in research with people with MS, their use is still far from being ubiquitous. Previous recommendations have focused on minimum cognitive screening for all people with MS, given the high prevalence of impairment (Kalb et al., 2018; Korakas & Tsolaki, 2016). Assessment should be able to identify impairment at an early stage, as well as becoming a part of routine monitoring. It is therefore important for assessments to be relevant, available and accessible, as well as acceptable to both people with MS and their clinicians.

Only 34% of the studies that used a neuropsychological multi-domain test or battery used one specifically developed for people with MS. Two of the most frequently used of these are screening assessments which are not specific to MS, the MMSE and the MoCA, which have been criticised for floor and ceiling effects, and therefore not being sensitive or specific enough to fully identify and reflect changes in cognitive impairment in people with MS (Beatty & Goodkin, 1990). The MMSE does not include measures of processing speed or executive functioning and is not considered an adequate assessment for use in MS clinical practice (Tobin, 2019), particularly due to inability to detect mild cognitive impairment.

Consensus-based recommendations for use of cognitive tests in MS have stated the need for both brief screening tests with the availability of comprehensive neuropsychological assessment in MS where needed. Quick screening tests of cognition (MMSE, BICAMS and MoCA, which take less than 30 minutes to complete), and longer assessments were reflected in the results, with eight batteries taking between 30 and 100 minutes to administer. While longer assessment may not be practical due to the high prevalence of

fatigue in MS (Hadjimichael et al., 2008) and the limited time available in clinics, the MS-specific batteries are important to capture a full neuropsychological evaluation of cognitive impairment.

The BICAMS has been translated into a number of languages since being recommended, however the extent of the validations is limited due to its recency, and a lack of development of translations and linguistic validation could mean that uptake of this battery has been limited. Conversely, the availability of the BRB-N in European countries, shown by the range of published translations and validations, is likely to make it more appealing to researchers and perhaps explain its high use in Europe (76% of uses of the BRB-N).

An expert consensus recently recommended the use of the SDMT as a minimum in people with MS (Kalb et al., 2018), and the high volume of use of this measure suggests that it is both acceptable and an adequate measure of processing speed for people with MS. The SDMT has been shown to have good validity and reliability (Benedict et al., 2017), as well as measuring one of the most frequently affected cognitive domains, that of information processing speed. There is some evidence that the results of the SDMT are related to factors impacted by cognitive impairment, including employment status and activities of daily living (ADLs) in people with MS (Benedict et al., 2005). However, the extent of this association has yet to be systematically reviewed. In the studies included in our review, there were several substitution style tests used with different names (for example, the Symbol Search Test; the Digit Symbol Coding Test, Coding Test), which may have been chosen over the SDMT due to cost or licencing requirements, over the psychometric properties. The omission of information about the administration of substitution style tasks in many of the included studies also demonstrates the importance of including pertinent issues about the outcome measures chosen. The use of variants and different administrations of these measures could mean that the generalisability of individual study findings is brought into question unless cross-validation has been considered. Norming data for the SDMT would be of limited value for similar tasks or different administrations, and so researchers and clinicians may use invalid comparisons to a normative sample.

The high usage of serial addition measures in research is unsurprising, given the inclusion of the PASAT in the MSFC. However more recent endorsements have moved away from using this measure, owing to its potential for inducing anxiety (Aupperle et al 2002), stress and frustration (Locke et al., 2011). Results of this task can be impacted by IQ, and mathematical ability which cannot be corrected without further and extensive data collection (Wills & Leathem, 2004). It is also likely that using the PASAT would lead to a high proportion of missing data due to participants refusing to undergo or complete the task. Participants who refuse to complete this task have been shown to be associated with higher levels of cognitive impairment (Cortés-Martínez et al., 2019), so using the PASAT could lead to significant missing data from some of the most cognitively impaired. These studies affirm the importance of choosing measures that are acceptable and accessible to people with MS.

Subjective measures of cognition in the review made up approximately 7% of the measures. Research has found that self-report measures of cognitive impairment in people with MS are more strongly related to depression than performance on objective cognitive measures, however proxy-reported measures tends to correlate higher with cognitive measures and may be valuable in assessing cognitive functioning (Benedict et al., 2004). Despite this, only 11% of subjective measures were proxy-reported, and it is possible that researchers and clinicians do not trust subjective reports of cognition. However, subjective measures can overcome many of the disadvantages of objective measures. Self- and proxy-reported measures can be less onerous on the participant and administrator, have fewer variables to control, may be more acceptable to the participant, and could reflect a global aspect of cognition that is distinct from that assessed in other ways.

Just under half of the studies had no administrative details of the assessments used (for example, verbal, written, or electronic), and thirty-five of the cognitive measures were not described well enough to ascertain the name, or what class of measures was used. Replication of these studies would be challenging. There was also considerable difficulty in classifying many of the measures due to the lack of information given in the papers, and this may have had an impact on the frequency of the classes and batteries. Both visual acuity and motor dexterity can be impaired in MS (Lau et al., 1998), and cognitive assessments

often involve the visual processing of stimuli, with participants giving motor (for example, written) or verbal responses. The written administrations of assessments can be affected by manual dexterity issues in MS, and while the verbal administrations are often used to overcome this, such as the oral administration of the SDMT, studies often did not specify this. Whilst some studies excluded participants with problems with vision and dexterity, many did not report how such issues would be overcome for the assessments. The missing information would make replication of these studies difficult.

Our finding associated with the frequency of use of tests should be treated with caution. This is based on simple counts, and the selection criteria for this review, which excluded translation and validation papers, may have meant that some papers were not counted here. Another limitation of this review is the missing information on the measures that were extracted for each study. There were some aspects of the use of cognitive measures in research that were outside of the scope for this review, such as the domains purported to be measured in the studies. Future search strategies could include specific cognitive domains and the names of measures, but this may inflate the number of articles that will need to be searched.

A formal quality assessment was not conducted as we wanted to capture the breadth and context of cognitive measures. An unexpected finding was that, for many studies, we were unable to ascertain what measures were used, or how the measures were administered, for example, data regarding the types of electronic devices used, or where the respondents completed the measure was missing in many primary studies. Devising a quality assessment checklist for future similar reviews may add further insight into this replicability issues. A further limitation of this review was that we only included papers that were published in English. Research conducted and published in other languages may have provided pertinent information regarding measures that are popular globally.

It was clear from the results of the systematic review that there are many cognitive measures being used in research with people with MS, however some, for example substitution and serial addition tasks, are prominent. The measures that were used ranged from short, individual screening tests to lengthy batteries that require trained staff to administer, as well as subjective measures that could be completed by

the people with MS, or a proxy. Cognitive measures were administered in a variety of modalities. One of the impacts of COVID-19 has been a shift towards telemedicine, and for neuropsychology to respond to this shift, research is needed on how cognitive assessments perform in remote administrations. We believe that this shift will be long lasting, beyond the COVID-19 pandemic. Evidence of validity and reliability for computerised cognitive tests and batteries for people with MS is emerging (Wojcik et al., 2019), but these do not guarantee equivalence between all potential modalities (for example, self-administered vs. technician administered; keyboard and mouse vs touchscreen), and new normative data may be required (Bauer et al., 2012). It is recommended that researchers and clinicians ensure that the measures used in all settings are appropriate to measuring cognitive impairment in MS, as well as being valid and reliable.

More research is needed to determine the acceptability of MS-specific tests from clinicians' and patients' perspectives, which may increase the use of and response to specific tests. Computerised adaptations especially require further evidence of their psychometric properties, as well as acceptability to patients. Systematic reviews are also needed to examine the association between the cognitive tests and functional outcomes, to improve clinicians' and patients' engagement with these tools.

References

Adler, G., & Lembach, Y. (2015). Memory and selective attention in multiple sclerosis: Cross-sectional computer-based assessment in a large outpatient sample. *European Archives of Psychiatry and Clinical Neuroscience*, *265*(5), 439–443.

Amato, M. P., Prestipino, E., & Bellinvia, A. (2019). Identifying risk factors for cognitive issues in multiple sclerosis. *Expert Review of Neurotherapeutics*, *19*(4), 333–347. https://doi.org/10.1080/14737175.2019.1590199

Amato, M. P., Zipoli, V., & Portaccio, E. (2006). Multiple sclerosis-related cognitive changes: A review of cross-sectional and longitudinal studies. *Journal of the Neurological Sciences*, *245*(1), 41–46. https://doi.org/10.1016/j.jns.2005.08.019

Aupperle, R. L., Beatty, W. W., deNAP Shelton, F., & Gontkovsky, S. T. (2002). Three screening batteries to detect cognitive impairment in multiple sclerosis. *Multiple Sclerosis Journal*, *8*(5), 382–389. https://doi.org/10.1191/1352458502ms832oa

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized Neuropsychological Assessment Devices: Joint Position Paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Archives of Clinical Neuropsychology*, *27*(3), 362–373. https://doi.org/10.1093/arclin/acs027

Beatty, W. W., & Goodkin, D. E. (1990). Screening for cognitive impairment in multiple sclerosis. An evaluation of the Mini-Mental State Examination. *Archives of Neurology*, *47*(3), 297–301.

Benedict, R. B. H., Munschauer, F., Linn, R., Miller, C., Murphy, E., Foley, F., & Jacobs, L. (2003). Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. *Multiple Sclerosis Journal*, *9*(1), 95–101. https://doi.org/10.1191/1352458503ms861oa

Benedict, R. H. B., Wahlig, E., Bakshi, R., Fishman, I., Munschauer, F., Zivadinov, R., & Weinstock-Guttman, B. (2005). Predicting quality of life in multiple sclerosis: Accounting for physical disability, fatigue, cognition, mood disorder, personality, and behavior change. *Journal of the Neurological Sciences*, *231*(1–2), 29–34. https://doi.org/10.1016/j.jns.2004.12.009

Benedict, R. H., DeLuca, J., Phillips, G., LaRocca, N., Hudson, L. D., & Rudick, R. (2017). Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *23*(5), 721–733. https://doi.org/10.1177/1352458517690821

Bruce, J., Hancock, L., Arnett, P., & Lynch, S. (2010). Treatment adherence in multiple sclerosis: Association with emotional status, personality, and cognition. *Journal of Behavioral Medicine*, *33*, 219–227. https://doi.org/10.1007/s10865-010-9247-y

Chiaravalloti, N. D., & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet. Neurology*, *7*(12), 1139–1151. https://doi.org/10.1016/S1474-4422(08)70259-X

Compston, A., & Coles, A. (2008). Multiple sclerosis. *Lancet*, *372*(9648), 1502–1517. https://doi.org/10.1016/S0140-6736(08)61620-7

Cortés-Martínez, A., Matias-Guiu, J. A., Pytel, V., Montero, P., Moreno-Ramos, T., & Matías-Guiu, J. (2019). What is the meaning of PASAT rejection in multiple sclerosis? *Acta Neurologica Scandinavica*, *139*(6), 559–562. https://doi.org/10.1111/ane.13090

Costa, S. L., Genova, H. M., DeLuca, J., & Chiaravalloti, N. D. (2017). Information processing speed in multiple sclerosis: Past, present, and future. *Multiple Sclerosis Journal*, *23*(6), 772–789. https://doi.org/10.1177/1352458516645869

Cotter, J., Firth, J., Enzinger, C., Kontopantelis, E., Yung, A. R., Elliott, R., & Drake, R. J. (2016). Social cognition in multiple sclerosis: A systematic review and meta-analysis. *Neurology*, *87*(16), 1727–1736. https://doi.org/10.1212/WNL.0000000000003236

Cox, D., Thompson, L. L., Foley, F., Weinstock-Guttman, B., & Munschauer, F. (2004). Reliable screening for neuropsychological impairment in multiple sclerosis. *Multiple Sclerosis Journal*, *10*(6), 675–678. https://doi.org/10.1191/1352458504ms1098oa

DeLuca, J., Chiaravalloti, N. D., & Sandroff, B. M. (2020). Treatment and management of cognitive dysfunction in patients with multiple sclerosis. *Nature Reviews Neurology*, *16*(6), 319–332. https://doi.org/10.1038/s41582-020-0355-1

DiGiuseppe, G., Blair, M., & Morrow, S. A. (2018). Short Report: Prevalence of Cognitive Impairment in Newly Diagnosed Relapsing-Remitting Multiple Sclerosis. *International Journal of MS Care*, *20*(4), 153–157. https://doi.org/10.7224/1537-2073.2017-029

Drew, M., Tippett, L. J., Starkey, N. J., & Isler, R. B. (2008). Executive dysfunction and cognitive impairment in a large community-based sample with Multiple Sclerosis from New Zealand: A descriptive study. *Archives of Clinical Neuropsychology*, *23*(1), 1–19. https://doi.org/10.1016/j.acn.2007.09.005

Ferreira, M. (2010). Cognitive deficits in multiple sclerosis A systematic review. *Arquivos de Neuro-Psiquiatria*, *68*, 632–641. https://doi.org/10.1590/S0004-282X2010000400029

Glanz, B. I., Healy, B. C., Rintell, D. J., Jaffin, S. K., Bakshi, R., & Weiner, H. L. (2010). The association between cognitive impairment and quality of life in patients with early multiple sclerosis. *Journal of the Neurological Sciences*, *290*(1–2), 75–79. https://doi.org/10.1016/j.jns.2009.11.004

Goverover, Y., & DeLuca, J. (2015). Actual reality: Using the Internet to assess everyday functioning after traumatic brain injury. *Brain Injury*, *29*(6), 715–721. https://doi.org/10.3109/02699052.2015.1004744

Hadjimichael, O., Vollmer, T., Oleen-Burkey, M., & North American Research Committee on Multiple Sclerosis. (2008). Fatigue characteristics in multiple sclerosis: The North American Research Committee on Multiple Sclerosis (NARCOMS) survey. *Health and Quality of Life Outcomes*, *6*, 100. https://doi.org/10.1186/1477-7525-6-100

Hansen, S., & Lautenbacher, S. (2017). Neuropsychological Assessment in Multiple Sclerosis. *Zeitschrift Für Neuropsychologie*, *28*(2), 117–148. https://doi.org/10.1024/1016-264X/a000197

Jaywant, A., Barredo, J., Ahern, D. C., & Resnik, L. (2018). Neuropsychological assessment without upper limb involvement: A systematic review of oral versions of the Trail Making Test and Symbol-Digit Modalities Test. *Neuropsychological Rehabilitation*, *28*(7), 1055–1077. https://doi.org/10.1080/09602011.2016.1240699

Kalb, R., Beier, M., Benedict, R., Charvet, L., Costello, K., Feinstein, A., Gingold, J., Goverover, Y., Halper, J., Harris, C., Kostich, L., Krupp, L., Lathi, E., LaRocca, N., Thrower, B., & DeLuca, J. (2018). Recommendations for cognitive screening and management in multiple sclerosis care. *Multiple Sclerosis Journal*, *24*(13), 1665–1680. https://doi.org/10.1177/1352458518803785

Klein, O. A., das Nair, R., Ablewhite, J., & Drummond, A. (2018). Assessment and management of cognitive problems in people with multiple sclerosis: A National Survey of Clinical Practice. *International Journal of Clinical Practice*, e13300. https://doi.org/10.1111/ijcp.13300

Korakas, N., & Tsolaki, M. (2016). Cognitive Impairment in Multiple Sclerosis: A Review of Neuropsychological Assessments. *Cognitive and Behavioral Neurology*, *29*(2), 55–67. https://doi.org/10.1097/wnn.0000000000000097

Langdon, D. W., Amato, M. P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., Hamalainen, P., Hartung, H. P., Krupp, L., Penner, I. K., Reder, A. T., & Benedict, R. H. (2011). Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). *Mult Scler*, *18*(6), 891–898. https://doi.org/10.1177/1352458511431076

Langdon, D. W., Amato, M. P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., Hämäläinen, P., Hartung, H.-P., Krupp, L., Penner, I. K., Reder, A. T., & Benedict, R. (2012). Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). *Multiple Sclerosis*, *18*(6), 891–898. https://doi.org/10.1177/1352458511431076

Lau, A., Chan, C. C. H., & Keung, Y. H. (1998). Visual perception and hand function in persons with multiple sclerosis. *Occupational Therapy International*, *5*(3), 194–205. https://doi.org/10.1002/oti.75

Locke, D. E. C., Stonnington, C. M., Thomas, M. L., & Caselli, R. J. (2011). Correlates of quitting the Paced Auditory Serial Addition Test in cognitively normal older adults participating in a study of normal cognitive aging. *Journal of Clinical and Experimental Neuropsychology*, *33*(8), 937–943. https://doi.org/10.1080/13803395.2011.578571

Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., Wolinsky, J. S., Balcer, L. J., Banwell, B., Barkhof, F., Bebo, B., Calabresi, P. A., Clanet, M., Comi, G., Fox, R. J., Freedman, M. S., Goodman, A. D., Inglese, M., Kappos, L., … Polman, C. H. (2014). Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology*, *83*(3), 278–286. https://doi.org/10.1212/WNL.0000000000000560

Moher, D. ; L., A. ;. Tetzlaff, J. ;. Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Prakash, R. S., Snook, E. M., Lewis, J. M., Motl, R. W., & Kramer, A. F. (2008). Cognitive impairments in relapsing-remitting multiple sclerosis: A meta-analysis. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *14*(9), 1250–1261. https://doi.org/10.1177/1352458508095004

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*(1), 33–65. https://doi.org/10.1016/j.acn.2004.02.005

Rao, S. M., Leo, G. J., Bernardin, L., & Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology*, *41*(5), 685–691.

Rao, S. M., Leo, G. J., Ellington, L., Nauertz, T., Bernardin, L., & Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning. *Neurology*, *41*(5), 692–696.

Romero, K., Shammi, P., & Feinstein, A. (2015). Neurologists accuracy in predicting cognitive impairment in multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *4*(4), 291–295. https://doi.org/10.1016/j.msard.2015.05.009

Rouleau, I., Dagenais, E., Tremblay, A., Demers, M., Roger, É., Jobin, C., & Duquette, P. (2018). Prospective memory impairment in multiple sclerosis: A review. *The Clinical Neuropsychologist*, *32*(5), 922–936. https://doi.org/10.1080/13854046.2017.1361473

Samkoff, L. M., & Goodman, A. D. (2011). Symptomatic management in multiple sclerosis. *Neurologic Clinics*, *29*(2), 449–463. https://doi.org/10.1016/j.ncl.2011.01.008

Sumowski, J. F., Benedict, R., Enzinger, C., Filippi, M., Geurts, J. J., Hamalainen, P., Hulst, H., Inglese, M., Leavitt, V. M., Rocca, M. A., Rosti-Otajarvi, E. M., & Rao, S. (2018). Cognition in multiple sclerosis. *Neurology*, *90*(6), 278–288. https://doi.org/10.1212/WNL.0000000000004977

Tobin, W. O. (2019). Management of Multiple Sclerosis Symptoms and Comorbidities. *Continuum*, *25*(3), 753–772. https://doi.org/10.1212/CON.0000000000000732

Walton, C., King, R., Rechtman, L., Kaye, W., Leray, E., Marrie, R. A., Robertson, N., La Rocca, N., Uitdehaag, B., van der Mei, I., Wallin, M., Helme, A., Angood Napier, C., Rijke, N., & Baneke, P. (2020). Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *26*(14), 1816–1821. https://doi.org/10.1177/1352458520970841

Wills, S., & Leathem, J. (2004). The Effects of Test Anxiety, Age, Intelligence Level, and Arithmetic Ability on Paced Auditory Serial Addition Test Performance. *Applied Neuropsychology*, *11*(4), 178–185. https://doi.org/10.1207/s15324826an1104_2

Wojcik, C. M., Beier, M., Costello, K., DeLuca, J., Feinstein, A., Goverover, Y., Gudesblatt, M., Jaworski, M., Kalb, R., Kostich, L., LaRocca, N. G., Rodgers, J. D., Benedict, R., & National MS Society Cognition Work Team. (2019). Computerized neuropsychological assessment devices in multiple sclerosis: A systematic review. *Multiple Sclerosis*, *25*(14), 1848–1869. https://doi.org/10.1177/1352458519879094

```
┌─────────────────────────┐                    ┌─────────────────────────────────────┐
│ Records identified      │                    │ Duplicates removed (n=2,847)        │
│ (n=11,854)              │───────────────────>│                                     │
│                         │                    │                                     │
└─────────────────────────┘                    └─────────────────────────────────────┘
            │
            v
┌─────────────────────────┐                    ┌─────────────────────────────────────┐
│                         │                    │ Title and Abstract Review (n=9,007) │
│ Title and               │                    │                                     │
│ Abstract screened       │───────────────────>│  • Not in English (n=3)             │
│ (n=9,007)               │                    │  • Not a human study (n=131)        │
│                         │                    │  • Review/validation/qualitative/case│
│                         │                    │    study (n=4,363)                  │
│                         │                    │  • People with MS are less than 50% │
│                         │                    │    of population under research     │
│                         │                    │    (n=1,944)                        │
│                         │                    │  • Sample is under 18 (n=3)         │
└─────────────────────────┘                    └─────────────────────────────────────┘
            │
            v
┌─────────────────────────┐                    ┌─────────────────────────────────────┐
│                         │                    │ Full text Review (n=2,563)          │
│ Full-text reviewed      │                    │                                     │
│ (n=2,563)               │                    │  • Not in English (n=33)            │
│                         │                    │  • Review/validation/qualitative/case│
│                         │                    │    study (n=71)                     │
│                         │                    │  • People with MS are less than 50% │
│                         │───────────────────>│    of sample under research (n=10)  │
│                         │                    │  • A measure of cognition is not used│
│                         │                    │    (n=854)                          │
│                         │                    │  • Duplicate data analysis (n=25)   │
└─────────────────────────┘                    │  • Unobtainable papers (n=44)       │
            │                                   └─────────────────────────────────────┘
            v
┌─────────────────────────┐
│ Studies extracted and   │
│ used in synthesis       │
│ (n=1,526)               │
│                         │
└─────────────────────────┘
```
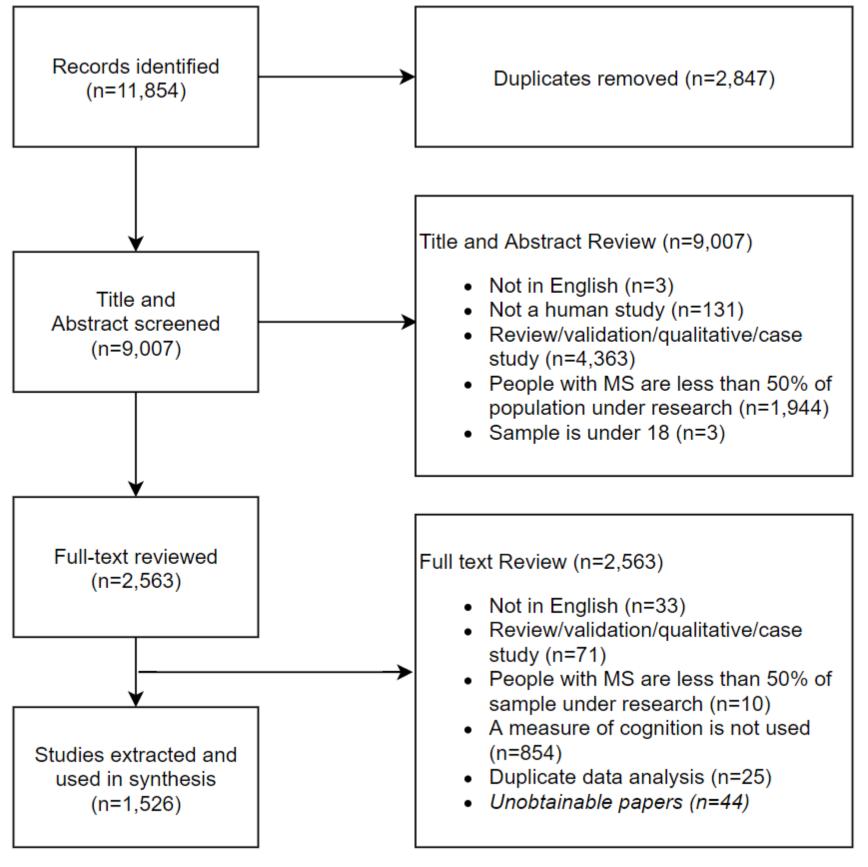
Table 1: Frequency of Individual Cognitive Assessments used with People with MS

Class	Examples of Tests	Number of Uses	% of individual test uses
ST- Substitution tests	Symbol Digit Modalities Test (SDMT; Smith, 1982), Digit Symbol (DS), Letter Digit Substitution Test	511	11.80%
SAT – Serial Addition Test	Paced Auditory Serial Addition Test (PASAT), Paced Visual Serial Addition Test (PVSAT)	476	11.00%
WLL - Word List Learning	California Verbal Learning Test (CVLT; Delis et al., 2004), Hopkins Verbal Learning Test (HVLT; Brandt & Benedict, 2001), Rey's Auditory Verbal Learning Test (RAVLT; Rey, 1941)	438	10.20%
VFT - Verbal Fluency Test	Controlled Oral Word Association Test (COWAT), or Word List Generation (WLG), FAS test	325	7.50%
ST - Stroop Test (Stroop, 1935)	Color and Word Test	242	5.60%
TMT - Trail Making Test	Colour Trails, Oral Trails, Making. Trails Test, Letter Number Sequencing	232	5.30%
DS - Digit Span	Digit Span, Forward Span, Backwards Span,	207	4.80%
ST - Sorting Test	Wisconsin Card Sorting Task (WCST) Modified Card Sorting, Weigl Test	164	3.80%
VR – Visual Reproduction Test	ROCFT - Rey-Osterrieth Complex Figure Test (ROCFT (Rey, 1941), Brief Visuospatial Memory Test (BVMT; Benedict, 1997), Figural Learning Tests, Taylor Complex Figure (TCF)	120	2.80%

Table 2: Frequency of Multidomain Batteries and Screens used with People with MS

Abbrev	Name	Type	Number of uses	% of uses	Therapeutic classification/population?
BRB-N	Brief Repeatable Battery of Neuropsychological Tests	Battery	210	23.00%	Multiple Sclerosis
MMSE	Mini-Mental State Examination	Screen	148	16.20%	Developed for geriatric population
WAIS	Wechsler Adult Intelligence Scale	Battery	86	9.40%	N/A
MACFIMS	Minimal Assessment of Cognitive Function in MS	Battery	72	7.90%	Multiple Sclerosis
WMS, WMS-III, WMS-IV	Wechsler Memory Scale, Wechsler Memory Scale Fourth Edition	Battery	52	5.90%	N/a
BICAMS	Brief International Cognitive Assessment for Multiple Sclerosis	Screen	62	6.80%	Multiple Sclerosis
MOCA	Montreal Cognitive Assessment	Screen	30	3.30%	Developed for detecting mild cognitive impairment

Table 3: Frequency of Subjective Cognitive Assessments used with People with MS

Subjective Measure Name	Frequency of uses	Percent of subjective measures
MSNQ - Multiple Sclerosis Neuropsychological Questionnaire	77	20.5%
PDQ - Perceived Deficits Questionnaire	34	9.0%
Self-report questions – not specified	*29*	*7.7%*
FSMC - Fatigue Scale for Motor and Cognitive Functions	21	5.6%
NeuroQoL	16	4.3%
FAMS - Functional Assessment of MS	14	3.7%
Global Impression Scale - Cognition	12	302%
EMQ - Everyday Memory Questionnaire	10	2.7%
MUSIC - Multiple Sclerosis Inventory of Cognition	9	2.4%
DEX - Dysexecutive Questionnaire	8	2.1%

Table 4: Frequency of Uses of Cognitive Assessments with People with MS

Use of cognitive measure	Frequency of uses*	% of uses
As a screening tool for inclusion/exclusion into a study	415	5.20%
To report on cognitive functioning in people with MS, for example as a between groups analysis	3757	47.03%
As a dependent variable	2305	28.86%
As a predictor, independent or correlational variable.	1484	18.58%
Other	27	0.34%

*Some measures were used for two or more purposes in a study