Which Multiple Sclerosis Patient and Disease Factors are Associated with the Relatives' Perceptions of the Cognition of People with Multiple Sclerosis?

Katarzyna Marek

May, 2023

Research submitted in partial fulfilment of the requirements for the degree of Doctor in Clinical Psychology (DClinPsy), Royal Holloway, University of London.

Acknowledgements

I would like to thank my project supervisor, Prof Dawn Langdon, for making this project possible and for her valuable feedback throughout. Thank you to Dr Roger Mills for his support with the Rasch analysis. Thank you to Prof Carolyn Young for supervising the project externally. Thank you to the TONiC research group for sharing this dataset with me for this secondary analysis. A big thank you to all patients with multiple sclerosis and their family members who participated in the TONiC study. Without you, this project could not have happened.

Thank you to Joanna and Enida for always being there to listen to me. Thank you to Hannah who made this DClinPsy journey a little less lonely. Thank you to my parents, Ela and Robert, who always believed in me and supported my journey into psychology from the start – Dziękuję Wam! Last but not least, thanks to Jack, who supported me throughout the three years of the course in every possible way!

Table of Contents

List of Figures

List of Tables

Chapter I: Lay Summary

Empirical Study

"Which Multiple Sclerosis Patient and Disease Factors are Associated with the Relatives' Perceptions of the Cognition of People with MS?"

Multiple sclerosis (MS) is a disease of the brain and spinal cord, where the immune system attacks nerve cells of one's own body. MS can lead to physical disability and, for about 45-70% of patients, difficulties with cognition (thinking skills such as remembering, learning new things, concentrating, or making decisions) which can result in patients' lower quality of life.

Patients' quality of life can be increased by accessing early intervention for cognitive functioning improvement (DeLuca et al., 2020). Regular cognitive testing may help with identification of these patients, but it is expensive, time-consuming, and not easily available, so using shorter screening instruments may be a useful, alternative indicator of cognitive abilities. The Multiple Sclerosis Neuropsychological Questionnaire-Informant (MSNQ-I), which is completed by an MS patient's carer/relative, is a well validated example of such a measure. Which patient and disease factors might be contributing to the relatives' scores on the MSNQ-I needed further investigation.

Although doctors and researchers use the MSNQ-I to measure people with MS' (PwMS') thinking skills, we needed to study its psychometric properties (i.e., how consistently and accurately it measures PwMS' thinking skills) to make sure it is a good questionnaire. Rasch

analysis is an example of a statistical method used to measure how well a set of items in a questionnaire work together to assess a particular trait.

First, we needed to check how well the MSNQ-I worked as a measure. Using a special analysis, we could tell that three of the fifteen items did not fit. They were more about behaviour, such as inappropriate laughing and crying. The rest of the scale included twelve items about thinking, such as forgetting appointments. We removed the three "different" items, to create the "MSNQ-I-12". We also worked out how the old MSNQ-I scores could be changed to new scores which made it a better measure, so that all of the new scores indicated equal steps in the relative's reported score of the patient's cognition.

With the more accurate MSNQ-I-12, we used another analysis to see how the relative's report of the patient's cognition related to other aspects of the patient's illness and other characteristics. We found that about a quarter of the relative's reported score of patient cognition was linked to the patient's gender, physical disability, anxiety, fatigue, how their illness affects their daily life and how their health affects their quality of life.

Going forward, we now have the MSNQ-I-12, which is a better measure of a relative's report of an MS patient's cognition. This can be used in clinics where separate cognitive assessments are not available. It can also be used in research studies when a quick, inexpensive way of including patient cognition is helpful. We also now know that some other things about patients can affect relatives' reports of the patient's cognitive abilities. Healthcare professionals using the MSNQ-I-12 can take care to consider these other things when using a relative's report as part of an MS patient's assessment.

Systematic Review

"The Associations of the Multiple Sclerosis Screening Questionnaire for Informants with Patients' Scores on Neuropsychological Assessments and Depression Scales: A Systematic Review"

Many people with multiple sclerosis (PwMS) experience changes with their cognition (thinking skills such as remembering, learning new things, concentrating, or making decisions) that can affect their daily life. Standard cognitive assessments with multiple sclerosis (MS) patients are expensive and not always possible so patients' changes in their thinking skills might not be addressed in clinics. Using shorter screening questionnaires may be another way of assessing MS patients' cognition. The Multiple Sclerosis Neuropsychological Questionnaire-Informant (MSNQ-I; Benedict et al., 2003), which is completed by an MS patient's family member/friend, is an example of such a measure. The aim of the current review was to determine whether the MSNQ-I can be used in the place of cognitive assessment in MS.

A search of research databases was conducted to find studies reporting how relatives' reports of patient thinking skills relate to patients' cognitive test scores and patient depression. Twenty-two studies were relevant to our question. To check how well these studies were done, we used the Effective Public Health Practice Project quality tool.

A method called "Systematic Review" was used to bring together the findings of 22 studies related to MSNQ data from patients and relatives and explore how the MSNQ relates to the patients' scores on objective cognitive testing and to other patient characteristics. All 22 studies included in the systematic review included numerical data, collected at one time point.

The links between the different areas of thinking skills (such as remembering, learning new things, concentrating, or making decisions) and the MSNQ questionnaires were not always clear. In general, patients' scores on objective cognitive tests were more closely related to relatives' scores on the MSNQ than to patients' scores on the MSNQ. Moreover, patient depression was more closely related to patients' scores on the MSNQ than to relatives' scores on the MSNQ.

The patient version of the MSNQ may not always give an accurate picture of patients' thinking skills. The MSNQ-I completed by a patient's family member could be a helpful way for doctors to check patients' thinking skills. Healthcare professionals should also consider patient depression when they assess PwMS' thinking skills.

Integration, Impact and Dissemination

Before starting the project, I attended webinars and read case studies about PwMS to understand the impact of difficulties with thinking skills in MS on people's lives. The research study provided additional information to the systematic review by suggesting that for more precise results it is best to use the updated version of the MSNQ-I questionnaire (MSNQ-I-12). The project had some challenges, for example a lot of data could not be used because not enough participants responded to some questions. Overall, I enjoyed working with other researchers in the MS field and learned a lot from them.

The results of these two research components might impact PwMS and their families as well as healthcare professionals and future research. The MSNQ-I-12 can be used instead of long and expensive tests. It means that more MS patients can be screened for changes in their thinking skills and offered the right support early (e.g., using calendars or reminders to remember things). Other researchers are encouraged to use similar statistical methods to check how good other commonly used questionnaires are.

We will share the findings with PwMS, researchers and clinicians working in MS. We are planning to send the results to scientific journals, MS magazines and charities. We also hope to present the findings to PwMS and their families as well as the researchers and health professionals working in MS.

Chapter II: Empirical Study

"Which Multiple Sclerosis Patient and Disease Factors are Associated with the Relatives' Perceptions of the Cognition of People with MS?"

Abstract

Cognitive difficulties in multiple sclerosis (MS) are common but often difficult to detect. Objective cognitive testing is expensive and time-consuming. Relatives' reports of MS patients' cognition on the Multiple Sclerosis Neuropsychological Questionnaire-Informant (MSNQ-I; Benedict et al., 2003) have been shown to correlate with patients' objective cognitive testing (e.g., Fenu et al., 2018; Migliore et al., 2021). The aim of the current study was to assess and improve the psychometric properties of the MSNQ-I with the Rasch analysis (Rasch, 1980) approach as well as understand which patient and disease factors predict relatives' scores on the Rasch-analysed MSNQ-I. This was a secondary data analysis study of the data from the Trajectories of Outcome in Neurological Conditions (TONiC; https://tonic.thewaltoncentre.nhs.uk/) study. The MSNQ-I data from 2,039 participants was Rasch-analysed in RUMM2030. Subsequently, Rasch-analysed scores were used in multiple regression. Rasch analysis confirmed that a 12-item version of the questionnaire (MSNQ-I-12) had a unidimensional structure. The last three items of the MSNQ-I measured a separate concept and were deleted from the scale. The multiple linear regression revealed that patients' gender, level of disability, anxiety, fatigue, health-related quality of life and disadvantage experienced as a result of ill health explained 28% of the variance on the MSNQ-I-12. Patients' depression was not a significant predictor in the multiple regression model. The MSNQ-I-12 has superior

psychometric properties to the MSNQ-I. The MSNQ-I-12 can be used as a proxy for objective cognitive assessment in MS, but the score should be carefully interpreted in the context of patients' disease and psychosocial factors.

Introduction

Cognitive impairment (CI) in multiple sclerosis (MS) is prevalent and reduces patients' quality of life (QoL). Traditional cognitive assessments with MS patients are not always feasible as they incur significant resource costs for the National Health Service (NHS) and, thus patients' CI may not be addressed in clinics. Adopting shorter screening instruments may be a valuable, alternative indicator of cognitive abilities. Patients' self-reports of cognition cannot reliably replace objective testing (Akbar & Finlayson, 2021), but informants' reports of patients' CI have been shown to correlate more reliably with patients' objective cognitive testing (e.g., Charest et al., 2020; Fenu et al., 2018; Migliore et al., 2021; O'Brien et al., 2007). The Multiple Sclerosis Neuropsychological Questionnaire-Informant (MSNQ-I; Benedict et al., 2003), which is completed by an MS patient's carer/relative, is a well-validated example of a subjective measure of difficulty with cognitive functioning. Further work is needed to determine whether the MSNQ-I can be used as a proxy cognitive assessment in MS. It is important to understand how a relative's perception of an MS patient's cognitive status is correlated with other aspects of the MS patient's disease (e.g., physical disability, type of MS).

Neurology of MS

Once diagnosed, MS is a lifelong autoimmune neurological condition of a largely unpredictable course, affecting the central nervous system (McGinley et al., 2021). MS affects over 2.8 million people globally and existing treatments do not cure MS but only slow progression. There are different MS subtypes. In relapsing-remitting multiple sclerosis (RRMS), acute flare-ups of symptoms lasting at least 24 hours (relapses) are followed by remission

(lengthy periods of full or almost full recovery). Rapidly evolving MS (REMS) is a severe type of relapsing remitting MS. It is defined by two or more disabling relapses in one year as well as either one or more gadolinium-enhancing lesions or significant increase in T2 lesions on brain Magnetic Resonance Imaging (MRI; Huisman et al., 2017). Secondary progressive multiple sclerosis (SPMS) follows from initial RRMS, when disability starts to accumulate at a variable rate with or without relapses, minor remissions, and plateaus. Primary progressive multiple sclerosis (PPMS) presents with a gradual accumulation of disability from the outset. MS can lead to a significant disability and involve a range of physical and psychological symptoms, including cognitive difficulties in all MS phenotypes (McGinley et al., 2021). These categories are increasingly recognised as scientifically unsatisfactory, although they remain the clinical standard (Granziera et al., 2023).

Impact of Cognition in MS on Life

Cognitive difficulties in MS are experienced by 45-70% of patients and decrease a patient's QoL (Gil-González et al., 2020; Lakin et al., 2021). Cognitive deficits in MS impair people's social cognition (Dulau et al., 2017), driving ability and safety (Morrow et al., 2018), the performance of activities of everyday living (Goverover et al., 2007), money management skills (Tracy et al., 2017), rehabilitation outcome (Langdon & Thompson, 1999), employment (Campbell et al., 2017) and disease management, including adherence to medication (Washington & Langdon, 2022). Since most people are diagnosed with MS in their 20s-30s, many patients experience cognitive difficulties for most of their adult lives, therefore the topic of cognition in MS is important.

Profile and Impact of Cognitive Difficulties in MS

People with MS (PwMS) can experience differing cognitive profiles, but there are broad patterns. CI is experienced the most frequently (50-75%) by those with SPMS and PPMS (Ruano et al., 2017). The most frequently affected cognitive domains in MS are information processing speed (IPS), memory, visual perceptual functions, attention, and executive skills (Benedict et al., 2020). Language in MS is rarely impaired (Benedict et al., 2020) which could be why healthcare professionals (HCPs) may fail to detect CIs during routine consultations. Additionally, the priorities of HCPs are different to those of PwMS. Whilst PwMS deem CI to be one of the most important domains of living with MS which needs to be addressed (Westergaard et al., 2022), neurologists do not consider cognition an important topic in routine clinic appointments (Marin et al., 2021). Reduced function resulting from cognitive changes can be wrongly attributed to other MS symptoms (e.g., depression or sleep-related disorders; Thomas et al., 2022b); or, in older adults, age-associated mild CI (Chiang et al., 2022). However, even when cognitive issues are raised in routine clinic appointments, objective cognitive testing is often not available outside of a few specialist MS centres, hence health professionals often rely on patient self-report to identify cognitive difficulties (Langdon et al., 2022; Meca-Lallana et al., 2021).

HCPs' Perceptions of PwMS' Cognition

There are some questions about relying on the HCPs' clinical judgements of MS patients' cognition. CI, unlike physical disability, is not a visible symptom of MS. This makes CI difficult to detect (Lakin et al., 2021). The extent of neurological disability in MS is typically measured

by the Expanded Disability Status Scale (EDSS; Kurtzke, 1983). Although the EDSS includes some items related to patient cognition (different "functional systems"), it fails to capture cognition reliably. CI of PwMS with a low physical disability (EDSS ≤ 4.0, meaning that PwMS were able to walk unaided) was not detected by neurologists in 25% of patients (Saccà et al., 2017). Jackson et al. (2022) reported that clinicians have a moderate accuracy in detecting global CI in MS patients. However, even less reassuringly, neurologists were at chance when identifying cognitively impaired patients in routine consultations (Romero et al., 2015). Clinicians were more likely to report greater CI for PwMS who were older, have higher levels of disability and reported more depressive symptoms (Jackson et al., 2022).

Confounds of Self-Reports of Cognition in MS

At first glance, it might seem that simply asking MS patients about their cognition might suffice, in the absence of formal cognitive testing. However, patients' self-reported cognitive status and their objective cognitive test scores do not perfectly align. Both significant associations (e.g., Benedict & Zivadinov, 2006; Nauta et al., 2019; Thomas et al., 2022) and non-significant associations (e.g., Benedict et al., 2003; Sejbæk et al., 2018) have been reported, suggesting that patients' self-reports of cognition are unreliable. Common confounds include depressive symptoms (Sejbæk et al., 2018; Van Laethem et al., 2022), anxiety (Akbar et al., 2011), unemployment (van Wegen et al., 2022), fatigue and poor self-efficacy (Strober et al., 2016). Patients' self-report of cognition cannot therefore reliably replace objective cognitive testing (Akbar & Finlayson, 2021).

Informants' Reports of Patients' Cognition in Other Conditions

Studies from other diseases suggest that informants' reports of patients' cognitive status may not always be reliable. There is evidence from other conditions that caregivers' perceptions of patients' functional independence (e.g., driving abilities) is influenced by external factors, such as a patient's diagnosis, independent of objective cognitive status (Schmidt & Steffen, 2020). Findings from a systematic review exploring factors associated with informant-reported cognitive decline in older adults also suggest that informants' reports are not always reliable (Morrell et al., 2019). Interestingly, dementia severity was correlated with relatives' reports of patients' cognitive decline (Morrell et al., 2019). Informants' reports of cognition of people with mild CI were also found to be influenced by their relationship with the person, whether they lived with the person and their education and race/ethnicity, even after controlling for patients' demographics, cognition and depression (Hackett et al., 2020).

Informants' Reports of Patients' Cognition in MS

Importantly, informants' ratings of MS patient cognitive status match objective testing more closely. The CI of PwMS significantly contributes to caregivers' reports of high-stress levels, suggesting relatives have some calibrated awareness of their relatives' CI level, which is unsurprising (Halstead et al., 2021). Significant associations between informants' reports of MS patients' cognition and patients' objective cognitive testing have been found (e.g., Benedict et al., 2004; O'Brien et al., 2007). Therefore, the informants' perceptions of PwMS' cognition could contribute to MS clinic management. Previous small studies from around the world have

shown that informants' reports of MS patients' cognition on the MSNQ-I can be related to other patients' characteristics and these potential confounds require further exploration.

The MSNQ-I

Because traditional neuropsychological assessment is expensive, time-consuming, and not easily available, and patients' self-reports are not always reliable, informants' reports may be useful to clinics. The MSNQ-I (Benedict et al., 2003) is an example of a short screening measure which is validated in many countries. This questionnaire can be completed by an informant (e.g., carer, family member) of PwMS. Informants use a five-point scale to rate the occurrence of 15 items related to patients' cognitive functioning (e.g., distractibility, slowed processing, forgetting what is read). Higher scores on the MSNQ-I represent a bigger perceived impact of cognitive deficits on daily functioning.

How do Relatives' Reports of PwMS' Cognition on the MSNQ-I Relate to PwMS' General Profile?

Aspects of the psychosocial, demographic and disease profile of PwMS have been shown to relate to the MSNQ-I, but the data is inconsistent. Patients' depression has been reported to correlate weakly and positively with the MSNQ-I (Akbar et al., 2010, 2011; Benedict et al., 2004; Langdon et al., 2013; Migliore et al., 2021; Rosti-Otajärvi et al., 2014; Sejbæk et al., 2018; Sonder et al., 2012), but not consistently (Benedict et al., 2003; Benedict & Zivadinov, 2006; Dagenais et al., 2013; Fenu et al., 2018, 2021; Konstantinopoulou et al., 2018; O'Brien et al., 2007; Thomas et al., 2022a, 2022b; Vanotti et al., 2009). The MSNQ-I has been shown to

correlate positively with patients' anxiety (Akbar et al., 2011; Sonder et al., 2012), but again, not consistently (Fenu et al., 2018, 2021; Thomas et al., 2022b).

Patients' fatigue does not always correlate with the MSNQ-I (Langdon et al., 2013), although this study was with patients with clinically isolated syndrome (CIS; the first episode of demyelination lasting at least 24 hours which may lead to a diagnosis of MS in the future), pre-MS diagnosis, when fatigue may not be so prominent. However, in a group of patients with mean disease duration of 13.5 years, the MSNQ-I was associated with greater patients' cognitive fatigue, but not with patients' self-reported physical fatigue (Thomas et al., 2022b). Lastly, the MSNQ-I was positively associated with most subtests of a QoL measure (Rosti-Otajärvi et al., 2014), but these correlations were weak. This was not the case for a CIS (pre-diagnosis) group (Langdon et al., 2013).

In terms of MS patients' demographics, the MSNQ-I has correlated with patients' gender (Sonder et al., 2012), but not consistently (Migliore et al., 2021). Patients' age was unrelated to the MSNQ-I (Akbar et al., 2011; Benedict et al., 2003; Langdon et al., 2013; Migliore et al., 2021; O'Brien et al., 2007), except in one study (Sonder, et al., 2012). Higher scores on the MSNQ-I were associated with less years of patients' education (Migliore et al., 2021; Sonder et al., 2012), but this finding was not consistent (Akbar et al., 2011; Benedict et al., 2003; Langdon et al., 2013; O'Brien et al., 2007). The MSNQ-I did not correlate with patients' premorbid intelligence (Akbar et al., 2011; O'Brien et al., 2007). The MSNQ-I scores were found to be significantly higher for unemployed MS patients, compared to employed patients (Akbar et al., 2011; Benedict & Zivadinov, 2006).

PwMS' disease factors have been inconsistently correlated with the MSNQ-I. The MSNQ-I has not been shown to correlate with time since CIS onset (in a short study of mainy

CIS patients; Langdon et al., 2013) or MS duration (Akbar et al., 2011; Benedict et al., 2003; Konstantinopoulou et al., 2018; O'Brien et al., 2007; Vanotti et al., 2009), except in one study (Sonder et al., 2012). Relatives of patients with SPMS reported a significantly greater CI on MSNQ-I compared to relatives of patients with RRMS (Benedict & Zivadinov, 2006). The MSNQ-I has been shown to correlate positively with the EDSS (Kurtzke, 1983; Migliore et al., 2021; Sonder et al., 2012; Vanotti et al., 2009), but not consistently (Akbar et al., 2011; Benedict et al., 2003; Konstantinopoulou et al., 2018; Langdon et al., 2013; O'Brien et al., 2007; Sejbæk et al., 2018).

MRI biomarkers have also been related to the MSNQ-I. Relatives' perceptions of PwMS' CI on the MSNQ-I was correlated with patients' brain volume as measured by MRI (whole brain, gray matter, cortical gray matter, Fenu et al., 2021; T1 lesion volume, T2 lesion volume, brain parenchymal fraction, Benedict & Zivadinov, 2006), except for the white matter volume (Fenu et al., 2021).

In summary, a number of disparate studies have explored which patients' variables correlate with informants' reports of patients' cognition in MS. The findings of these studies were mixed. The majority of studies reported no association between informants' perceptions of patients' cognition and patients' age, fatigue, depressive symptoms, anxiety symptoms, premorbid IQ levels, years of education, MS duration, and disability status. The associations of the MSNQ-I with patients' gender and patients' QoL were inconsistent. These studies were small and the associations between patients' variables and relatives' perceptions of patient cognition have not been extensively explored.

Theoretical Framework

Accurate reports of cognitive functioning require "thinking about thinking", so called "metacognition" (i.e., being aware of cognitive processes). Relatives use their observations of PwMS' everyday behaviours to form a model in their mind of PwMS' cognitive status. The relatives, based on their observations, decide whether behaviours are related to the CI of their loved ones. For instance, they might attribute certain failures to PwMS' levels of fatigue or disability rather than to their CI. A theoretical metacognition framework that can be adapted to capture this has been suggested by Morris and Mograbi (2013; see Figure 1). This model has been used in research about metacognition in MS (e.g., Mazancieux et al., 2019). Relatives' metacognitive awareness of PwMS' cognitive status relies on relatives' observations of PwMS' behaviour (including conversations) and the attributions and evaluations relatives make about these observations. According to the model, metacognition involves three main components: sensory and perceptual processing (step 1), monitoring and evaluation (step 2), and control and regulation (step 3).

Step 1: Informants observe PwMS and infer any sensory or perceptual deficits that may affect PwMS' cognitive function. For example, they may notice that a person with MS has difficulty with visual acuity or fine motor control, which could impact their ability to read or write. This will moderate the extent to which PwMS' disability is attributed to CI by relatives.

Step 2: Informants assess PwMS' cognitive functions by observing their relatives. For example, they may notice that their loved one has difficulties with attention or memory, such as forgetting appointments or having trouble following a conversation. The informants may also ask PwMS questions about their cognitive function and compare PwMS' assessment to their own

observations. This is likely to involve episodic memory, for example, a person with MS used to complete a crossword in 30 minutes but is observed to take an hour currently. Autobiographical memory could be involved, for example at the relative's 40th birthday a few years ago, a person with MS organised the guest list and the invitations, booked the restaurant and generally made the arrangements. In comparison, for the relative's 50th birthday, a person with MS is needing prompting for each action and monitoring who to invite, in order for the arrangements to be made.

Step 3: Informants adjust their internal model of PwMS' cognitive abilities as they continue to observe and monitor PwMS' cognition over time. Information about PwMS' cognition is monitored by the relatives' Cognitive Comparator Mechanisms which compare the new knowledge with the existing knowledge stored in a Database (Figure 1). Informants may provide feedback and support to PwMS to help them compensate for cognitive deficits. For example, informants may suggest external aids, such as a calendar or reminders, to help PwMS with memory and executive function.

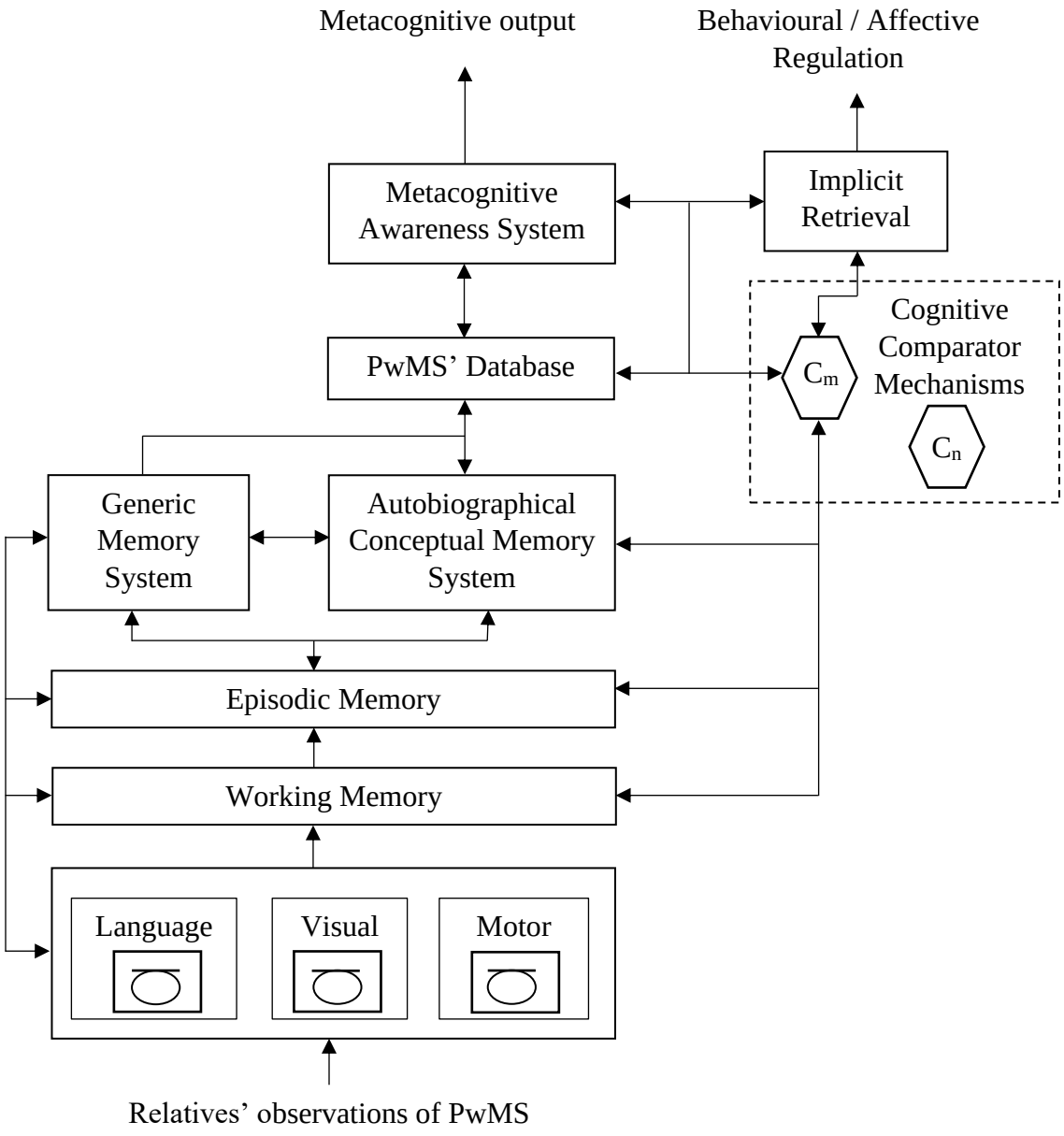


Figure 1

Model of Metacognition Adapted From Morris and Mograbi (2013)

Psychometric Properties of the MSNQ-I

If the MSNQ-I is to be considered a proxy for objective cognitive assessment in MS, the psychometric properties of the MSNQ-I need to be assessed and, if necessary, improved with a robust approach. Previous studies (e.g., Benedict et al., 2004) reviewed the MSNQ-I's internal validity, sensitivity (0.87) and specificity (0.84), but to date, no studies have been conducted on assessing or increasing the construct validity of the MSNQ-I using the more robust Rasch analysis (Rasch, 1980) approach. Rasch analysis was used in the original study (Benedict et al., 2003) only on the MSNQ-P (MSNQ-Patient version) items to reduce the number of items in the scale from 68 to 15. Only two studies have previously investigated the psychometric properties of the MSNQ-I. Both Sonder et al. (2012) and Migliore et al. (2021), based on the results of the factor analysis (principal component analysis), reported that the MSNQ-I items load on one general factor and confirmed that the MSNQ-I has a unidimensional structure. The samples in these studies were small and the results cannot be generalisable. However, another study exploring the structure of the patient version of the MSNQ, through the use of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), reported multidimensionality of MSNQ-P (Konstantinopoulou et al., 2018). These discrepant findings are unexpected for two reasons. First, both versions of the MSNQ were designed to measure one construct (i.e., patient cognition). Second, the wording of the questions on the MSNQ-P and MSNQ-I varies only slightly and only to reflect that different people answer the same questions (e.g., “Do you forget what you read?” versus “Does the person with MS forget what they read?”). The unidimensionality of the MSNQ-I needs to be confirmed.

Rasch Analysis

Rasch analysis is commonly used to create new scales and to assess psychometric properties of existing scales. It ensures that they are unidimensional, internally consistent, free of redundant items and capable of interval level measurement (Pallant & Tennant, 2007). Rasch analysis involves testing whether or not the data satisfy the requirements of the Rasch model. These requirements include unidimensionality (measurement of a single latent trait, i.e., cognition), monotonicity (increase in item responses consistent with underlying trait, e.g., item scores increase as the PwMS' CI increases), local item independence (zero correlation between items when conditioned on the score, i.e., items in a test are not related to each other), the stochastic (probabilistic) ordering of items (e.g., when comparing informants of PwMS with minimal versus profound cognitive difficulties, informants of PwMS with greater CI have a higher probability of attributing an item than informants of PwMS who are less cognitively impaired) and group invariance (no difference in response to an item by group membership, e.g., sex, age, type of MS, when at the same level of scores; Gustafsson, 1980; Teresi et al., 2000).

The Rasch model is used to measure latent traits, like attitude or ability. Rasch analysis allows researchers to use a respondent's nonlinear raw (ordinal) data and convert it to a linear (interval) scale that accounts for the unequal difficulties across all test items (Boone, 2016). This can be then evaluated through the use of parametric statistical tests, for which normally distributed interval-level data is needed (Kazis et al., 1989), and which have greater statistical power and precision than non-parametric statistical tests. Rasch analysis provides an internally valid measure that, when developed from an appropriate sample, is independent of the particular sample to which it is applied, meaning that the findings for the sample can be extended to its population. Editors are increasingly looking for psychological scales to be Rasch-analysed in

published work. In the current study Rasch analysis will document and evaluate the measurement function of the MSNQ-I, given fit of data to Rasch model expectations. The study will use the MSNQ-I data of 2,039 patient-informant dyads. No large dataset has been interrogated to determine which patient characteristics are independently linked to the MSNQ-I.

Summary

Cognitive assessment is a crucial component of MS management. Given the possibility of cognitive decline over time in MS, with subsequent negative impact on social life and QoL, a quick and reliable measure of cognition in MS is needed. Neuropsychological testing is expensive, time-consuming and in short supply, and patients' self-reports of cognition are unreliable, therefore the use of short screening instruments for relatives' reports may be a useful indicator of cognitive status. Assessing and improving the psychometric properties of the MSNQ-I with a robust approach, as well as understanding what extraneous factors impact relatives' evaluations of MS patients' cognition are needed before the MSNQ-I can be adopted as a proxy for objective cognitive assessment in MS. The objectives of this study will be to apply Rasch analysis to the MSNQ-I and, subsequently, to use the Rasch-analysed scores in the exploratory multiple linear regression to understand which patient factors are associated with informants' scores on the MSNQ-I.

Aims

The current study aims to apply Rasch analysis to increase the construct validity of the MSNQ-I and determine how good a measure the MSNQ-I is. The Rasch analysis will provide a more psychometrically robust dataset for the exploratory multiple linear regression models. Subsequently, the Rasch transformed MSNQ-I data will be used to explore how patients' characteristics relate to MSNQ-I scores, this is which patients' disease (type and duration, physical disability), demographic (age, gender) and psychosocial characteristics (depression, anxiety) are related to relatives' reports of MS patients' cognitive functioning on the MSNQ-I. This will increase understanding of relatives' reports of MS patient cognition.

First Hypothesis (Part I)

Following Rasch analysis, the MSNQ-I data will be consistent with the stochastic (probabilistic) ordering of items, monotonicity, local item independence, unidimensionality, and group invariance.

Second Hypothesis (Part II)

Scores on the Rasch-transformed MSNQ-I will not be associated with patients' age, disability, fatigue, depressive symptoms, anxiety symptoms, gender, or MS duration.

Method

Trajectories of Outcome in Neurological Conditions (TONiC) study

TONiC is a multiphase and multicentre UK national study examining quality of life of people with neurological conditions (https://tonic.thewaltoncentre.nhs.uk/). The data was from a uniquely large sample and captured a broad range of patients' experiences and stages of disease.

TONiC Sample Inclusion and Exclusion Criteria

The inclusion criteria:

1. Adults diagnosed with MS
2. Capable of answering questionnaires
3. Capable of informed consent

The exclusion criterion:

1. Suffering from a concomitant serious medical or psychiatric condition (the seriousness of which was at the clinician/researcher's discretion)

Participants were identified by the clinical care team or the research team at each centre. All participants were given information about the study and asked for informed consent once they read a written information sheet and had the opportunity to ask questions. They were provided a follow-up contact from the study team after about two weeks, by phone, email or letter, as per patient preference. Demographic and disease data from the case notes were recorded (age, gender, diagnosis, year of diagnosis, MS disability status according to EDSS bands) of those patients who consented to take part in the study. Participants who wanted to complete a

questionnaire pack signed another informed consent. A questionnaire pack consisted of MS-specific scales along with comparison generic scales used in routine care and it was administered to those participants by post, clinic recruitment or a secure website. Participants were told that they did not need to complete all questionnaires in one sitting. All documents were part-anonymised so that respondents can be identified by a study number only. Participants who were unable to attend clinics received home visits. Lay carers were also recruited, and they were asked to complete a scale on patients' cognitive functioning (MSNQ-I). Participants were informed that they had the right to decline taking part or withdraw from the study and that it would not affect their clinical care. The collection of questionnaire data started in November 2013.

Research Approvals

The project uses data from the TONiC study which received a favourable ethical opinion from a relevant local Research Ethics Committee (North West - Greater Manchester West 11/NW/0743; IRAS project ID: 88372). All subjects of the original study received information on the study and gave written, non-written or electronic remote informed consent prior to participation, including for secondary analysis. The current study was approved by the Research Ethics Committee at Royal Holloway, University of London (REC Project ID: 3153; Appendix I).

Current Study Procedure and Participants

The present study is a secondary analysis of data relating to participants who had a corresponding MSNQ-I score provided by their relatives. This study included demographic and disease information about patients and Patient Reported Outcome Measures (PROMs). Thus, this

data was extracted from the TONiC database to form a unique dataset for this study. These records were securely emailed as an Excel document and saved on the Royal Holloway drive.

Materials

1) Informants' Perceptions of PwMS' Cognitive Difficulties

The Multiple Sclerosis Neuropsychological Questionnaire-Informant (MSNQ-I; Benedict et al., 2003) is a brief test completed by PwMS' relatives or carers, consisting of 15 questions about PwMS' cognition during activities of daily living (see Appendix II). Informants are asked to read each item (e.g., Does he/she forget appointments?) and indicate, using a five-point Likert scale, how often each behaviour occurred in the past three months and how severe it was ("Very often, very disruptive", "Quite often, interferes with life", "Occasionally, seldom a problem", "Very rarely, no problem", "Never, does not occur"). Scores range from zero to 60 and higher scores indicate more cognitive complaints. The MSNQ-I has a good internal validity (sensitivity 0.87, specificity 0.84; Benedict et al., 2004) and has been validated in many countries, including Argentina (Vanotti et al., 2009), Netherlands (Sonder et al., 2012), Denmark (Sejbæk et al., 2018), Greece (Konstantinopoulou et al., 2018), and Italy (Migliore et al., 2021).

2) The Impact of Health Conditions on Functioning

The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0; Ustün et al., 2010) is a standardized assessment instrument for measuring health and disability across cultures. The WHODAS 2.0 is an easily administered PROM of the impact of health conditions on functioning in six domains (cognition, mobility, self-care, getting along, life activities, participation), during the previous 30 days. It assesses individuals' difficulties (i.e., increased

effort, discomfort or pain, slowness, changes in the way the person does the activity) that they experience in doing different activities, regardless of their medical diagnosis (e.g., "How much difficulty did you have in washing your whole body?"). A five-point Likert scale contains the following response options: "None", "Mild", "Moderate", "Severe", "Extreme or cannot do". The WHODAS 2.0 consists of 36 items which takes about 5 minutes to complete when it is self-administered. The total score is a sum of all scores from all six domains converted into a metric ranging from zero to 100 (0 = no disability, 100 = full disability). WHODAS 2.0 with 32-items was used, excluding the questions related to work (D5.5 – D5.8) as not all participants were in employment (see Appendix III). The WHODAS 2.0 has strong theoretical underpinnings, linked to the International Classification of Functioning, Disability and Health (World Health Organization, 2001). It has good psychometric properties, particularly a unidimensional structure (Rasch, 1980), a high internal consistency (Cronbach's alpha, α: 0.86) and high test-retest reliability (intra-class coefficient: 0.98; Ustün et al., 2010). In a validation study of the WHODAS 2.0 in MS it was reported that the scale is reliable and valid, with a Crobach's alpha of 0.93 and a Person Separation Index of 0.83 (Magistrale, Medori, et al., 2015; Magistrale, Pisani, et al., 2015).

3) The Effect of Chronic Disease on Functional Ability

The London Handicap Scale (LHS; Harwood et al., 1994) is a self-report measure for determining the effect of chronic disease on one's functional ability (see Appendix IV). It assesses a perceived level of disadvantage in six dimensions: mobility, orientation, occupation, physical independence, social integration, and economic self-sufficiency. Each response on a six-point Likert scale ("Not at all", "Very slightly", "Quite a lot", "Very much", "Almost completely", "Completely") is assigned a scale weight and the total scale value ranges from zero (minimum value, indicating total disability) and 1.00 (maximum value, indicating normal function). In the

current study higher scores indicate higher levels of disability. The LHS is based on the World Health Organisation (WHO) handicap framework (International Classification of Impairments, Disabilities, and Handicaps; World Health Organization, 2001). It is a valid, reliable, and acceptable measure (Pearson's correlation coefficient between predicted and measured values: 0.98; Kendall's coefficient of concordance: 1.00; Harwood et al., 1994). The LHS has been used in MS (e.g., Veillard et al., 2021).

4) Fatigue

The Neurological Fatigue Index (NFI-MS; Mills et al., 2010) is a PROM measuring fatigue severity and factors influencing fatigue. It consists of 10 items measuring the physical and cognitive areas of MS fatigue. Items are rated on a four-point Likert scale (“Strongly disagree”, “Disagree”, “Agree”, “Strongly agree”) based on the past two weeks. Scores range from zero to 30 and higher scores indicate more fatigue complaints. The NFI-MS has a good test-retest reliability (correlations coefficient above 0.7) and a good external validity (correlations of 0.7 with another comparable fatigue scale). The NFI-MS has been utilised in several studies to assess fatigue in PwMS (e.g., Ekmekyapar Fırat et al., 2021).

5) Anxiety and Depression

The Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) is a PROM consisting of 14 questions, seven about anxiety and seven about depression (see Appendix V). The HADS was designed to assess mood in physical health conditions (Wu et al., 2021). Items are rated on a four-point severity scale based on the past seven days (e.g., “I can laugh and see the funny side of things” or “Worrying thoughts go through my mind”). Scores for anxiety and depression subscales range from 0 to 21 and higher scores indicate more depression and anxiety

complaints. This scale is a reliable tool for screening for clinically significant anxiety (Cronbach's alpha, α: 0.86) and depression (Cronbach's alpha, α: 0.82) in MS and a valid measure of the severity of these mood disorders (Patel & Feinstein, 2017).

6) Disability Status

The Expanded Disability Status Scale (EDSS; Kurtzke, 1983) is a clinician-assessed disability status scale and it is the most commonly used outcome measure in MS clinical research. A neurologist examines various functions (e.g., balance, coordination, tremor, ability to move arms and legs) and assigns a score on the scale (see Appendix VI). The higher the score, the more severe the disability. Scores range from zero (everything is normal) to 10 (death due to MS) in 0.5-unit increments. If the participant was able to walk unaided for up to 500 metres, they were in the first band (i.e., had the lowest level of disability; 0 – 4.0). If they needed a walking stick or two walking sticks, they were in the second band (4.5 – 6.5). If they needed a wheelchair outside the house, they were in the third band (7.0 – 7.5). If they were totally chairbound or bedbound, they were in the fourth band (i.e., had the highest level of disability; 8.0-9.5). These bands were used in previous TONiC studies (e.g., Young et al., 2021).

7) Health-Related Quality of Life

The EuroQol five-dimension (EQ-5D; Balestroni & Bertolotti, 2012) is a self-report measure for describing and valuing health (see Appendix VII). It assesses five areas of one's health: mobility (i.e., walking), self-care (i.e., washing/dressing self), usual activities (i.e., work, study, housework, family, leisure activities), pain/discomfort and anxiety/depression. The respondent indicates their health state, based on the day they complete the questionnaire, by choosing one of the responses on a five-point severity scale (1 = "No problems", 2 = "Slight problems", 3 =

"Moderate problems", 4 = "Severe problems", 5 = "Extreme problems/Unable to"). Each response is coded as a single-digit number indicating the severity level in each dimension and the total score is written as a 5-digit code, where each digit corresponds to one health dimension. The EQ-5D includes an additional question about one's health in general on a scale from zero to 100 (0 = the worst health one can imagine, 100 = the best health one can imagine). The EQ-5D has excellent psychometric properties, including a test-retest reliability of above 0.7 (Feng et al., 2021). Its validity is established across different subgroups and it correlates with other health-related quality of life measures (Feng et al., 2021). The EQ-5D has been utilised in several studies to assess health-related quality of life in PwMS (e.g., Claflin et al., 2022; Visser et al., 2021). The EQ-5D was reported to have a good test/retest reliability (intra-class correlation coefficient: 0.81) and construct validity (correlation coefficient with other clinical measures was 0.70; Fisk et al., 2005) in a study with MS sample.

Statistical Analysis Plan

There will be two stages of data analysis:

Stage 1) Rasch analysis of the MSNQ-I scores

Stage 2) A multiple linear regression analysis assessing which patients' demographic and disease variables are associated with the Rasch-transformed MSNQ-I scores

Stage 1) Rasch Analysis of the MSNQ-I scores

The ordinal scores from the MSNQ-I will be first transformed into interval-scaled latent estimates through fit of their data to the pre-determined measurement model, called Rasch model (Rasch, 1980). The RUMM2030 Rasch analysis software will be used to apply the Rasch Partial Credit model (Andrich & Hagquist, 2015; Masters, 1982) to examine the following requirements of the Rasch model: unidimensionality, monotonicity; homogeneity; local independence and group invariance (Gustafsson, 1980; Teresi et al., 2000). Rasch analysis will be used to test the validity of the total score.

Exploratory and validation samples of 500 cases each will be created. Sample sizes ranging from 250 to 500 cases are recommended when using the RUMM2030 software to ensure accurate Type I error on the chi-square interaction fit statistic (Hagell & Westergren, 2016). Once a valid solution is found in the exploratory sample, it will be applied to the validation sample. The final scale solution will need to demonstrate satisfactory fit statistics across both samples.

Summary and Item fit Statistics

The main fit statistic will be the overall chi-square interaction; the chi-square probability (Bonferroni adjusted for number of scale items) should be non-significant indicating no deviation from the model.

If a bifactor solution is found, then a conditional chi-square fit statistic will be used since conditional inferences are less erroneous in large samples (Christensen et al., 2013; Müller, 2020). In addition, for testlet solutions, the explained common variance (‘A value’), will be

calculated; this represents the variance retained in the scale in the principal factor and should be ≥ 0.9.

Individual items will also need to display satisfactory fit residuals (<± 2.5) and non-significant chi-square (Alpha 0.01). Whilst a high positive fit residual might indicate an under-discrimination of an item (i.e., it does not discriminate between different groups and the person responds in an opposite way than expected), a high negative fit residual might be an indication of an over-discrimination of an item (e.g., the person always chooses the same response). A high negative fit residual might be also associated with either dependency or a high item total correlation (e.g., redundancy).

Reliability statistics will include extreme person scores and be based on the Person Separation Index (PSI) and Cronbach's alpha. Both should be greater than 0.7 for group use and greater than 0.85 for individual subject measurement.

Threshold Ordering

The difficulty thresholds in responses to polytomous items should progress in order.

Unidimensionality

A post-hoc test based upon the residuals, after applying the Rasch model, will be conducted to test for unidimensionality (Smith Jr., 2002). Here less than 5% of the tests between two item sets identified in the principal component analysis of the residuals should be significant. The lower bound of 95% confidence intervals of a binomial distribution, applying the Agresti-Coull method (Hagell, 2014), will be used and calculated using the online resource at https://epitools.ausvet.com.au/ciproportion.

Differential Item Functioning

To test the MSNQ-I for invariance (lack of Differential Item Functioning – DIF) an Analysis of Variance (ANOVA) of the residuals will be undertaken across the following groups: age (quartiles: < 43, 44 - 52, 53 - 60, > 61), gender (male or female), MS type (PPMS, REMS, RRMS, SPMS), self-reported EDSS bands (0 - 4, 4.5 – 6.5, 7 – 7.5, 8 – 9.5), marital status (divorced, married, single, widowed), relationship with patient (child, domestic partner, friend, other family, other friend, parent, spouse). The sample size of the current study will be sufficient (above 243 participants) to provide accurate item and person location estimates irrespective of the scale targeting (Kline, 1998).

Local Independence

The residual correlation between items should ideally be zero to meet the local item independence requirement. Simulation work has shown that a residual correlation of 0.2 above the average is indicative of a breach of this requirement, and this will be the value applied (Christensen et al., 2017).

Given satisfactory fit to the Rasch model, a conversion (nomogram) between the raw MSNQ-I score and the interval scale estimate will be provided (the person location), based on the whole sample for the highest accuracy. The logit estimates will be converted to the same range as the raw score by a further linear transformation. This can be then used in other samples to convert raw scores into linear estimates, provided that subjects respond to all items on the scale.

Strategies for Achieving fit

Given that the MSNQ-I is an existing scale, all items will be attempted to be maintained. If there are misfitting items, testlets will be used (based on conceptual groupings or known subscales or directed by local dependency). If using those testlets fails, then a bifactor solution will be used taking alternate items into two testlets. If fit is not achieved with these, it will be examined which items are misfitting and those items will be excluded from the analysis. If any items are discarded, they will be assessed for fit to Rasch model to examine if they create a separate scale.

Stage 2) A Multiple Linear Regression Analysis Assessing Which Patient Demographic and Disease Variables are Associated with the Rasch-Transformed MSNQ-I Scores

Statistics will be conducted using the statistical software IBM SPSS Statistics version 25. The data distribution will be explored to check for normality and identify outliers. Descriptive statistics will be performed for all variables. To assess which variables will be included in the multiple regression, parametric and non-parametric correlations (for PROMs, patients' age, disease duration, EDSS band), *t*-tests (for patients' gender) and ANOVA (for MS type) will be performed, using the Rasch-transformed MSNQ-I scores. Bonferroni corrections will be applied for multiple testing. The variables demonstrating significant relations to the Rasch-transformed MSNQ-I will be put forward to the multiple regression analysis. An exploratory multiple linear regression model will be built to identify significant predictors of relatives' reports of MS patients' cognitive status.

Results

The data analysed was collected between November 2013 and September 2019.

Stage 1) Rasch Analysis of the MSNQ-I Scores

Exploratory and validation samples were randomly and separately derived from the original total sample of 2,039. Data in the exploratory sample were fitted to the Rasch model (Partial Credit Model). The process involved examination of summary fit statistics, ordering of thresholds, individual item fit, residual correlations, overall item fit, dimensionality and DIF.

Initial Examination of fit in the Exploratory Sample, n=500

The 15-item MSNQ-I scale did not fit the Rasch model with a highly significant overall chi-square ($p < 0.00001$; Table 3, analysis 1). Mean fit residual for items was -0.04 with SD of 3.57 (deviating from the expected value of 1.00). Mean fit residual for persons was -0.36 with SD of 1.72. The person separation index (PSI) was 0.94 and a Cronbach's alpha of 0.958 which indicated that the scale could resolve subjects into at least five distinct groups of ability.

All item thresholds were ordered which suggested that PwMS' informants were able to discriminate between response options (Figure 2), but several items had misfit (Table 1). The worst fitting items were items 14 and 15 (laughing and crying without cause and egocentric speech).

Table 1

Item fit Statistics for all MSNQ-I Items for Exploratory Sample, n=500

Item nr	Item description*	Location	Fit residual	DF	Chi-square	DF	P
1	Distractibility	-0.503	0.147	439.4	7.70	9	0.564374
2	Problems with listening to others	-0.265	-2.158	439.4	14.54	9	0.104447
3	Slowed problem processing	-0.243	-1.650	439.4	9.34	9	0.406514
4	Forgetting appointments	0.106	1.624	439.4	14.81	9	0.096293
5	Forgetting what is read	0.128	-0.619	439.4	15.80	9	0.071185
6	Forgetting shows/programs	0.175	-0.736	439.4	9.19	9	0.41986
7	Forgetting instructions	-0.270	**-3.793**	439.4	21.28	9	0.011464
8	Needing frequent reminders	-0.187	-2.217	439.4	15.00	9	0.090939
9	Forgetting future errands	-0.074	-1.447	439.4	12.39	9	0.192411
10	Coherent question answering	0.341	**-4.073**	439.4	21.60	9	0.01024
11	Failing to track two tasks at once	-0.320	-2.507	439.4	16.90	9	0.05037
12	Failing to follow conversations	-0.224	-1.565	439.4	14.29	9	0.112239
13	Impulse control	0.448	**3.804**	439.4	9.89	9	0.359898
14	Without cause laughing/crying	0.424	**7.483**	439.4	92.89	9	**<0.000001**
15	Excessive egocentric speech	0.464	**7.066**	439.4	68.16	9	**<0.000001**

Notes. DF= Degrees of Freedom, Nr= Number, P= Probability. **Misfit is indicated in bold.**

*For full text see Appendix II.

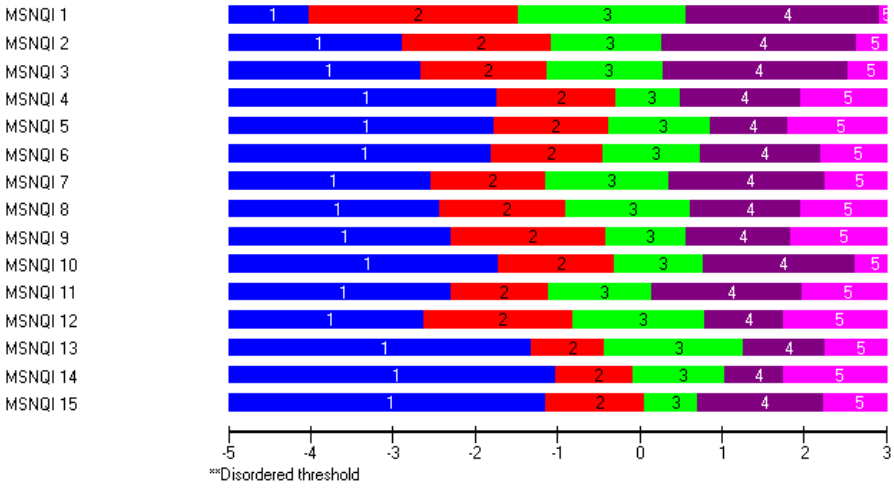


Figure 2

Category Thresholds

Examination of the residual correlations revealed local dependency between items 1 and 2 (distractibility and problems listening to others), 4 and 9 (forgetting appointments and forgetting future errands) and 8 and 9 (needing frequent reminders and forgetting future errands). The correlation matrix is given in Table 2.

Table 2

Residual Correlation Matrix for the 15-Item MSNQ-I in the Exploratory Sample, n=500

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	0.215	1	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-0.001	0.073	1	-	-	-	-	-	-	-	-	-	-	-	-
4	-0.047	-0.175	-0.101	1	-	-	-	-	-	-	-	-	-	-	-
5	-0.122	0.003	-0.034	0.183	1	-	-	-	-	-	-	-	-	-	-
6	-0.119	-0.046	0.029	-0.084	0.173	1	-	-	-	-	-	-	-	-	-
7	-0.134	-0.024	0.104	-0.125	-0.028	-0.018	1	-	-	-	-	-	-	-	-
8	0.029	-0.087	-0.140	0.129	-0.146	-0.214	0.142	1	-	-	-	-	-	-	-
9	-0.053	-0.163	-0.183	0.368	-0.043	-0.137	-0.025	0.423	1	-	-	-	-	-	-
10	-0.173	-0.022	0.127	-0.210	-0.132	0.084	0.051	-0.112	-0.100	1	-	-	-	-	-
11	-0.124	-0.064	0.096	-0.080	-0.117	-0.096	0.024	0.090	0.034	0.174	1	-	-	-	-
12	-0.116	0.022	-0.011	-0.330	-0.140	0.004	0.082	-0.205	-0.301	0.116	-0.039	1	-	-	-
13	-0.118	-0.169	-0.210	-0.191	-0.115	-0.083	-0.186	-0.218	-0.192	-0.123	-0.122	0.029	1	-	-
14	-0.076	-0.104	-0.231	-0.150	-0.149	-0.098	-0.286	-0.287	-0.216	-0.144	-0.213	-0.031	0.186	1	-
15	-0.053	-0.198	-0.221	-0.173	-0.220	-0.215	-0.177	-0.094	-0.193	-0.125	-0.218	0.021	0.160	0.159	1

The scale was multidimensional with the number of significant *t*-tests at 0.11 (lower bound of 95% confidence interval [CI] = 0.09).

The MSNQ-I was free from DIF for all six factors (age, gender, MS type, self-reported EDSS bands, marital status, relationship with patient).

The scale targeting was reasonable with a modest floor effect of 4.8% and no ceiling effect (Figure 3).

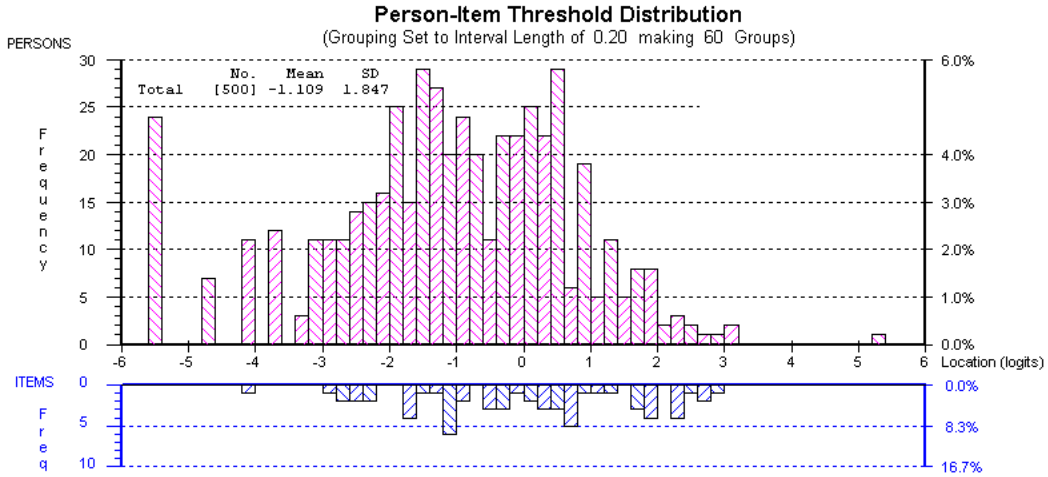


Figure 3

Targeting of the 15-item MSNQ-I in the Exploratory Sample, n=500

Further Analysis

The analysis steps are presented in Table 3. First, the misfit of items 13, 14 and 15 was addressed by deleting these three items. Multidimensionality was still present as indicated by the post-hoc *t*-test (Table 3, analysis 2). Item 4 had a high fit residual (3.067) and a high residual

correlation (0.3) with item 9 therefore item 4 was deleted (Table 3, analysis 3). The scale was still multidimensional and there was a high residual correlation between items 9 and 8 (0.389). Item 9 was deleted since it had a lower individual chi-square probability than item 8 (Table 3, analysis 4). This solution resulted in an acceptable overall fit, so it was applied to the validation sample (Table 3, analysis 5). There was a residual correlation of 0.257 between items 1 and 2 so item 1 was deleted because it had a high fit residual (2.907). This resulted in an overall acceptable fit (Table 3, analysis 6) so the solution was applied to the exploratory sample (Table 3, analysis 7). Item 5 had a misfit, and the overall chi-square value was low (0.00378) so the solution of deleting items 13, 14, 15, 4, 9 and creating a testlet (i.e., subtest/ST) for items 1 and 2 was applied (Table 3, analysis 8). When this solution was applied to the validation sample (Table 2, analysis 9) the testlet had an item residual of 4.3.

In an attempt to retain deleted items, all items were explored conceptually and grouped into three testlets (Subtest1: 1, 2, 3, 9; Subtest2: 4, 5, 6, 7, 8; Subtest3: 13, 14, 15; Table 3, analysis 10). Although the fit was acceptable (alpha = .93), Subtest2 had a residual of -6.43 and Subtest3 had a residual of 4.75. This solution was applied to the validation sample (Table 3, analysis 11) and the fit was almost identical to the exploratory sample.

Final Solution

The solution with two testlets (items 1, 2, 3,9 and items 4, 5, 6, 7, 8) was also tested following the deletion of items 13, 14, 15 since it was clear from the preceding analyses that the last three items were misfitting the scale and representing a different dimension. This solution was applied to the exploratory (Table 3, analysis 12) and validation (Table 3, analysis 13)

samples and resulted in acceptable fits. This was the final solution which was subsequently applied to the whole sample and resulted in a unidimensional scale with an acceptable fit (Table 2, analysis 14).

The discarded items 13, 14 and 15 were also separately analysed, however they did not fit the Rasch model as a separate scale. A bifactor solution using alternative items was also applied to exploratory and validation samples, however the solution with two testlets (Subtest1: 1, 2, 3, 9; Subtest2: 4, 5, 6, 7, 8) resulted in superior fit statistics.

The final scale had ordered thresholds within the testlets (Figure 4).

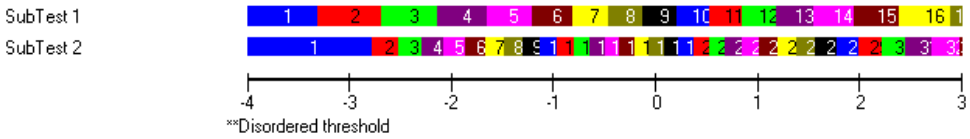


Figure 4

Thresholds Ordering of the Final Two-Testlet Solution

The final scale had an excellent spread of item thresholds and targeting very similar to the 15-item scale applied to the exploratory sample, with a modest floor effect of 4.9% and a negligible ceiling effect of 0.4% (Figure 5).

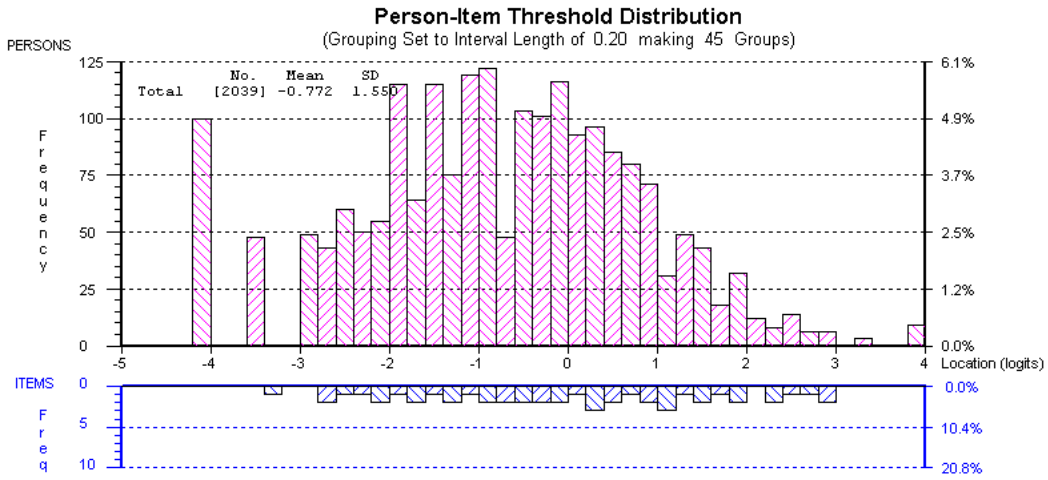


Figure 5

Person-Item Targeting of the Final Two-Testlet Solution on the Whole Sample, n=2,039

Table 3

Summary fit Statistics for Rasch Analyses

Analysis			Item Fit Residual		Person Fit Residual		Chi-square Interaction			Condi-tional				Unidimensionality	Lower
Nr	Name	Sample	M	SD	M	SD	Value	df	*p*	*p*	A	PSI	α	*t*-tests (Nr of sig. tests/out of =)	bound 95% CI
1	Initial	exp	-0.04	3.57	-0.36	1.72	343.78	135	<0.00001	-	-	0.94	0.96	56/500 = 0.11	0.09
2	Del 13,14,15	exp	-0.04	1.85	-0.48	1.65	172.28	108	0.00009	-	-	0.94	0.96	47/500 = 0.09	0.07
3	Del 13,14,15,4	exp	-0.03	1.83	-0.50	1.63	131.02	99	0.01727	-	-	0.94	0.96	38/500 = 0.08	0.06
4	Del 13,14,15,4,9	exp	-0.02	1.93	-0.49	1.58	126.39	90	0.00692	-	-	0.93	0.96	24/500 = 0.05	0.03
5	Del 13,14,15,4,9	val	-0.04	2.29	-0.40	1.36	92.44	90	0.40895	-	-	0.93	0.95	34/500 = 0.07	0.05
6	Del 13,14,15,4,9,1	val	0.10	2.21	-0.38	1.34	66.90	81	0.87014	-	-	0.93	0.95	26/500 = 0.05	0.03

Analysis			Item Fit Residual		Person Fit Residual		Chi-square Interaction			Condi-tional	A	PSI	α	Unidimensionality	Lower bound
Nr	Name	Sample	M	SD	M	SD	Value	df	*p*	*p*				*t*-tests (Nr of sig. tests/out of =)	95% CI
7	Del 13,14,15,4,9,1	exp	0.09	1.98	-0.47	1.55	119.12	81	0.00378	-	-	0.93	0.95	27/500 = 0.05	0.04
8	Del 13,14,15,4, 9 ST1:1,2	exp	0.06	1.93	-0.46	1.51	117.20	81	0.00532	-	1.00	0.93	0.95	25/500 = 0.05	0.03
9	Del 13,14,15,4,9 ST1:1,2	val	0.04	2.59	-0.39	1.34	104.93	81	0.03814	-	1.00	0.93	0.95	30/500 = 0.06	0.04
10	ST1:1,2,3,9 ST2:4,5,6,7,8 ST3:13,14,15	exp	-1.27	5.62	-0.48	0.95	30.46	27	0.29388	-	0.93	0.87	0.80	-	-
11	ST1:1,2,3,9 ST2:4,5,6,7,8 ST3:13,14,15	val	-1.11	5.08	-0.49	1.01	32.51	27	0.21377	-	0.92	0.86	0.80	-	-
12	Del 13,14,15 ST1:1,2,3,9 ST2:4,5,6,7,8	exp	-0.40	3.95	-0.52	0.87	10.00	18	0.93197	0.17	0.97	0.92	0.82	18/500 = 0.04	0.02

Analysis			Item Fit Residual		Person Fit Residual		Chi-square Interaction			Condi-tional	A	PSI	α	Unidimensionality	Lower bound
Nr	Name	Sample	M	SD	M	SD	Value	df	*p*	*p*				*t*-tests (Nr of sig. tests/out of =)	95% CI
13	Del 13,14,15 ST1:1,2,3,9 ST2:4,5,6,7,8	val	-0.37	3.91	-0.51	0.86	11.23	18	0.88456	0.03	0.96	0.90	0.80	15/500 = 0.03	0.01
14	Del 13,14,15 ST1:1,2,3,9 ST2:4,5,6,7,8	full	-0.97	8.10	-0.57	0.94	23.73	18	0.16418	0.07	0.97	0.92	0.81	70/2039 = 0.03	0.03
-	Acceptable Values	-	0	<1.4	0	<1.4	-	-	> .05*	>0.05*	>0.9	>0.85	>0.85	<0.05	<0.05

Note. α= Alpha level, A= explained common variance, CI= Confidence Interval, df= Degrees of Freedom, Del= Deleted items; exp= Exploratory sample, full= Full sample, M= Mean, Nr = Number, *p*= Probability, PSI= Person Separation Index, SD= Standard Deviation, ST= Subtest, val= Validation sample. *= Bonferroni-adjusted

The plot of the summed raw score to interval logit measure is given in Figure 6 and the nomogram of the summed raw score to a conversion metric for the MSNQ-I-12 items is provided in Table 4. These conversions are only valid when there is no missing data.

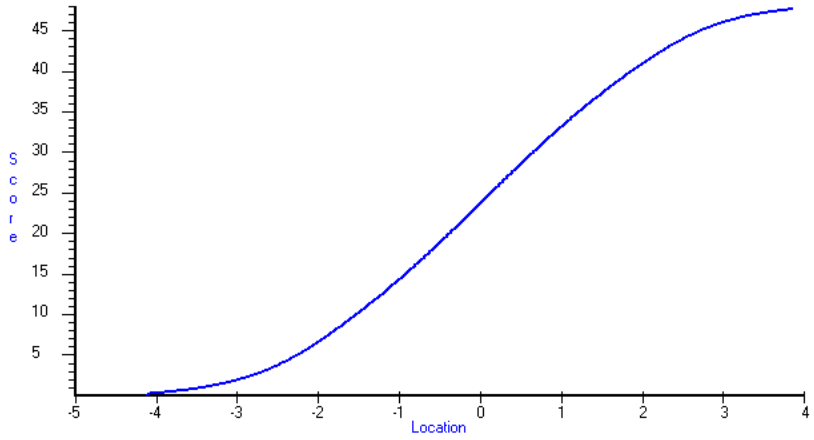


Figure 6

Plot of Summed Raw Score to Interval Logit Measure of the Final Two-Testlet Solution, n=2039

Table 4

Nomogram of Summed Raw Scores to Interval Level Conversion for the MSNQ-I-12

Raw Score	Converted Score
0	0.00
1	4.11
2	6.80
3	8.56
4	9.90
5	11.02
6	12.00
7	12.90
8	13.75
9	14.56
10	15.35
11	16.12

Raw Score	Converted Score
12	16.86
13	17.58
14	18.28
15	18.97
16	19.64
17	20.29
18	20.93
19	21.56
20	22.18
21	22.80
22	23.40
23	24.01
24	24.61
25	25.22
26	25.82
27	26.44
28	27.05
29	27.68
30	28.30
31	28.94
32	29.60
33	30.27
34	30.95
35	31.65
36	32.37
37	33.11
38	33.87
39	34.66
40	35.48
41	36.33
42	37.23
43	38.19
44	39.26
45	40.50
46	42.09
47	44.46
48	48.00

From henceforth only the MSNQ-I-12 scores will be considered.

2) Multiple Regression Analysis

Data Distribution

First, the whole sample data for EQ-5D, WHODAS32, LHS, NFI-MS, HADS Depression (HADS D), HADS Anxiety (HADS A), MSNQ-I-12 was checked for normality through the inspection of histograms. Not all variables appeared to be normally distributed, with some data being positively skewed (WHODAS32, LHS) and some negatively skewed (EQ-5D). The remaining variables appeared to be normally distributed. This was not further addressed statistically as in samples of hundreds of cases the issue of distribution of data can be ignored (Altman & Bland, 1995, p. 298), therefore parametric statistics were used. No data has been excluded from the analysis as a small number of outliers (N = 20) identified by a visual inspection of boxplots were likely to be reflective of the heterogenous MS population.

Sample Description

In total data records from 2,039 patients with complete MSNQ-I data were extracted from the TONiC database. Informants reported their relation to the corresponding patient to be spouse (71.2%), domestic partner (10.4%), parent (6.9%), child (4.9%), friend (2.5%), other family member (1.8%) or other friend (1.0%), not reported (1.5%). Participants had been recruited across 29 MS centres in the UK and 29.4% was collected in Liverpool. There was some heterogeneity in disease modifying therapies (DMT) which participants reported using: Interferon (10.8%), Natalizumab (8.6%), Glatiramer (8.2%), Dimethyl fumarate (5.4), Fingolimod (3.3%), Alemtuzumab (1.8%), Teriflunomide (1.0%), study drug (0.2%), Azathioprine (0.1%), Daclizumab (<0.1%), Laquinimod (<0.1%), Mitoxantrone (<0.1%),

Ocrelizumab (<0.1%), Stem cell transplant (<0.1%). The majority of participants reported taking no DMT (59.8%) and a few did not provide any information on DMT (0.5%; Amin & Hersh, 2023). Descriptive statistics for participant demographic and disease variables were calculated (Tables 5 and 6).

Table 5

Whole MSNQ-I-12 Sample Demographics by EDSS Band

	Total sample*	EDSS 0 – 4	EDSS 4.5 – 6.5	EDSS 7 – 7.5	EDSS 8 – 9.5
					
N	2,039	872	854	171	133
%	100	42.8	41.9	8.4	6.5
Mean age (SD)	51.3 (12.0)	45.5 (11.4)	54.4 (10.6)	59.0 (10.3)	58.8 (9.8)
Female %	73.6	76.1	71.9	76.6	64.7
Married %	80.5	79.9	79.5	84.2	85.7
Working full-time %	11.2	22.7	3.3	0.6	0.0

Note. EDSS= Expanded Disability Status Scale, *N*= Number, SD= Standard Deviation

*Nine patients had missing data for EDSS band

Table 6

Disease Variables by EDSS Band for the Whole MSNQ-I-12 Sample

	N	Total sample	EDSS 0 – 4	EDSS 4.5 – 6.5	EDSS 7 – 7.5	EDSS 8 – 9.5
						
Phenotype (%)						
PP	288	14.1	6.7	19.7	21.1	19.5
RE	72	3.5	4.4	3.5	2.3	0.0
RR	1,111	54.5	84.5	41.8	8.8	1.5
SP	559	27.4	4.5	35.0	67.8	78.9
Nr	9	0.4	-	-	-	-
Mean duration (SD)	2,011	11.5 (10.0)	7.6 (7.7)	12.6 (9.6)	18.1 (11.8)	21.5 (10.2)
Taking DMT (%)	2,029	39.7	53.0	37.4	12.3	6.0

Note. DMT= Disease Modifying Therapies, MS= Multiple Sclerosis, *N*= Number, Nr= Not reported, PPMS= Primary Progressive MS, REMS= Rapidly Evolving MS, RRMS= Relapsing-Remitting MS, SD= Standard Deviation, SPMS= Secondary Progressive MS

Relation of the Demographic and Disease Characteristics to the MSNQ-I-12 – Whole Sample

An independent samples *t*-test was used to compare the MSNQ-I-12 scores of male and female participants. Means and standard deviations of the MSNQ-I-12 for males and females for the whole sample are displayed in Table 7.

Table 7

Mean and Standard Deviation of the MSNQ-I-12 for Males and Females – Whole Sample

Gender	*N*	M	SD
Female	1,500	19.44	9.10
Male	539	21.17	9.35

Note. M= Mean, *N*= Number, SD= Standard Deviation

Levene's test of homogeneity was not significant ($F = .887$, $p = .347$), therefore equal variances were assumed. Male patients were rated as having more cognitive difficulties than female patients on the MSNQ-I-12 ($t(2037) = -3.77$, $p < .001$). Means and standard deviations of the MSNQ-I-12 for all MS types for the whole sample are displayed in Table 8.

Table 8

Means and Standard Deviations of the MSNQ-I-12 for all MS Types – Whole Sample

MS Type	*N*	M	SD
PPMS	288	19.13	9.23
REMS	72	20.77	9.37
RRMS	1,111	19.47	8.88
SPMS	559	20.93	9.65

Note. M= Mean, *N*= Number, SD= Standard Deviation

A one-way ANOVA was conducted to examine the effect of MS type (PPMS, REMS, RRMS, SPMS) on the MSNQ-I-12 score (Table 9).

Table 9

ANOVA for MS Type

	F	*p*
MS type	4.03	.007

The analysis of variance demonstrated a significant effect of MS type on the MSNQ-I-12 scores ($F(3,2026) = 4.03$, $p = .007$), therefore Bonferroni's post-hoc comparisons were carried out with adjusted alpha levels of .008 to account for multiple comparisons. Fisher's protected *t*-tests showed that the SPMS group were rated as having significantly more cognitive difficulties than the RRMS group ($t(1040.13) = -2.99$, $p = .003$) and PPMS group ($t(845) = -2.60$, $p = .009$). However, when using the corrected significance value of .008, there is no significant difference between PPMS and SPMS groups on the MSNQ-I-12 scores, and therefore this finding is inconclusive.

Exploring the Associations Between the MSNQ-I-12, Patients' age, Disease Duration, EDSS Bands and PROMs

Pearson's correlation coefficients between the MSNQ-I-12, patients' age, disease duration and PROMs are presented in Table 10. Table 10 also includes Spearman's correlation coefficient which was used to correlate the EDSS bands (ordinal data) with all other variables

(interval data). Spearman's correlation was chosen over Kendall's coefficient due to large sample size (Khamis, 2008). LHS was the most strongly correlated with the MSNQ-I-12.

Performing multiple correlation coefficients increases the risk of obtaining significant results by chance (committing Type I error; Curtin & Schulz, 1998). The level of significance for correlation coefficients were adjusted for multiple comparisons and the adjusted level was set to .001 (.05 / 45 = .001). Moreover, having a large sample could have also transformed small differences into significant differences (Faber & Fonseca, 2014).

Table 10

Pearson's and Spearman's Correlation Coefficients Between the MSNQ-I-12, Patients' age, Disease Duration, EDSS Bands and PROMs

	1	2	3	4	5	6	7	8	9	10
1. EDSS band	-	**.126[b] *p* < .001**	**.430[b] *p* < .001**	**.410[b] *p* < .001**	**-0.630[b] *p* < .001**	**.614[b] *p* < .001**	**.560[b] *p* < .001**	**.318[b] *p* < .001**	**.304[b] *p* < .001**	-0.024[b] *p* = .294
2. MSNQ-I-12[a]		-	.007 *p* = .761	.045 *p* = .041	**-.299 *p* < .001**	**.450 *p* < .001**	**.454 *p* < .001**	**.448 *p* < .001**	**.401 *p* < .001**	**.275 *p* < .001**
3. Age			-	**.474 *p* < .001**	**-.255 *p* < .001**	**.227 *p* < .001**	**.162 *p* < .001**	**.093 *p* < .001**	**.089 *p* < .001**	**-.187 *p* < .001**
4. MS duration				-	**-.246 *p* < .001**	**.247 *p* < .001**	**.200 *p* < .001**	**.079 *p* = .001**	**.083 *p* < .001**	**-.079 *p* < .001**
5. EQ-5D					-	**-.797 *p* < .001**	**-.743 *p* < .001**	**-.538 *p* < .001**	**-.539 *p* < .001**	**-.294 *p* < .001**
6. WHODAS32[a]						-	**.873 *p* < .001**	**.695 *p* < .001**	**.700 *p* < .001**	**.394 *p* < .001**
7. LHS							-	**.656 *p* < .001**	**.662 *p* < .001**	**.341 *p* < .001**
8. NFI-MS[a]								-	**.597 *p* < .001**	**.390 *p* < .001**
9. HADS D[a]									-	**.567 *p* < .001**
10. HADS A[a]										-

Notes. EDSS= Expanded Disability Status Scale, EQ-5D= Health Status Scale, HADS A= Hospital Anxiety and Depression Scale-Anxiety, HADS D= Hospital Anxiety and Depression Scale-Depression, LHS= London Handicap Scale, MS= Multiple Sclerosis, MSNQ-I-12= The Multiple Sclerosis Neuropsychological Questionnaire-Informant (Rasch-analysed), NFI-MS= The Neurological Fatigue Index, WHODAS 32= World Health Organisation Disability Assessment Schedule omitting work-related items; **Significant correlations are in bold.** The level of significance was set to .001. Pearson's correlations unless indicated ([b] = Spearman's). [a] = Rasch-transformed scores

Multiple Linear Regression

Prior to conducting multiple linear regression, the assumptions of the regression analysis were checked.

1. The visual inspection of the histogram of the MSNQ-I-12 (dependent variable) revealed that residuals approximate a normal distribution.
2. Linearity between the MSNQ-I-12 and other variables was examined earlier (Table 10). The MSNQ-I-12 was not significantly correlated with patients' age and disease duration, therefore these two variables were excluded from the multiple regression. The remaining independent variables (patients' gender, EDSS band, EQ-5D, LHS, NFI-MS, HADS D, HADS A) were assessed to have a linear relationship with the MSNQ-I-12.
3. Multicollinearity was also examined, and a very high correlation (above 0.8) was observed between LHS and WHODAS32, therefore only LHS was included in the regression as it had a higher number of responses (N = 1,983) than WHODAS 32 (N = 1,704). Variance Inflation Factor values were checked for the independent variables, and they indicated no problematic multicollinearity (i.e., values below 5).
4. The visual inspection of a plot of standardised predicted values versus standardised residual values revealed that there was an absence of heteroscedasticity.
5. Outliers (by distance and influence) were examined. Twenty-one outliers were found, and they were retained in the analysis as they were likely to be representative of the MS population and, given a small number, unlikely to significantly affect the results.
6. The dependent variable was at the interval level and predictors were at the mixture of interval, ratio, and dichotomous levels.

Disease type was excluded from multiple regression because almost 60% of the sample had RRMS type. Therefore, PwMS with low levels of disability would have been overly represented which could have introduced some bias and affected the analysis. The level of disability (EDSS band) is a clear indication of disease progression in MS (Meyer-Moock et al., 2014) and all EDSS bands were represented in this dataset more evenly than the MS types. Spearman's 'Rho' revealed a significant and moderate positive correlation between EDSS band and MS type ($r(2028) = .29$, $p < .001$), therefore only EDSS band was retained for the multiple regression. EDSS bands are also ordinal data, unlike MS types, and can be treated as interval data without the need to create "dummy" variables.

Multiple regression analysis excludes participants with missing data (see Appendix VIII for the table with the number of participants from total sample for all variables considered in multiple regression). After excluding participants with missing data on any of the included variables, the total number of participants in the multiple regression was 1,806.

Power Analysis

To detect reasonable-size effects with reasonable power, 10 or 15 cases of data for each predictor are required in the multiple regression model (Field, 2018, p. 519). However, to detect a small effect size a sample of 806 is required for the multiple regression with up to 10 predictors (Cohen, 1988). The available sample of 1,806 with seven predictors will therefore suffice for detecting small effect size and achieving a high level of power.

A multiple linear regression was performed to analyse the combined effect of patients' gender, patients' level of disability (EDSS band), self-reports of patients' anxiety (HADS A), patients' depression (HADS D), patients' fatigue (NFI-MS), health-related quality of life (EQ-5D) and patients' disadvantage experienced as a result of ill health (LHS) on the relatives' perceptions of the cognition of PwMS (MSNQ-I-12). A multiple linear regression was carried out using the ENTER method. These seven variables accounted for a significant amount of variance in the score on the MSNQ-I-12 ($F(7,1798) = 99.43$, $p < .001$; $R^2 = .28$, adjusted $R^2 = .28$). The partial regression coefficients, displayed in Table 11, showed that variables which had a significant unique relationship to the MSNQ-I-12 scores were: gender ($t(1799) = 4.35$, $p < .001$), EDSS band ($t(1799) = -5.17$, $p < .001$), HADS A ($t(1799) = 2.63$, $p = .009$), NFI-MS ($t(1799) = 7.81$, $p < .001$), LHS ($t(1799) = 10.17$, $p < .001$) and EQ-5D ($t(1799) = 2.13$, $p = .034$).

Table 11

A Multiple Linear Regression for the Predictors of the Relatives' Perceptions of the Cognition of PwMS on the MSNQ-I-12 (n = 1,806)

Model	*B*	*SE B*	*β*	*t*	*p*
Gender	1.836	.422	.089	4.346	<.001
EDSS Band	-1.474	.285	-.137	-5.170	<.001
EQ-5D	2.466	1.161	.070	2.125	.034
LHS	.711	.070	.370	10.168	<.001
NFI-MS	.316	.040	.219	7.812	<.001
HADS D	.141	.083	.053	1.701	.089
HADS A	.155	.059	.068	2.631	.009

Notes. EDSS= Expanded Disability Status Scale, EQ-5D= Health Status Scale, HADS A= Hospital Anxiety and Depression Scale-Anxiety, HADS D= Hospital Anxiety and Depression Scale-Depression, LHS= London Handicap Scale, NFI-MS= The Neurological Fatigue Index, WHODAS 32= World Health Organisation Disability Assessment Schedule omitting work-related items

Patient-reported depression on HADS was not independently associated with the MSNQ-I-12 score ($t(1799) = 1.70$, $p = .089$). The analysis suggested that LHS ($\beta = .37$) was the most influential predictor and HADS D ($\beta = .05$) was the least influential predictor in the model. There were no significant differences between the whole sample and the multiple regression sample (n = 1,806 with complete data) in terms of patient demographic (age, gender, marital status, employment) and disease (MS type, duration, DMT data) variables. Multiple regression sample demographics and disease information are displayed in Appendices IX and X.

Discussion

CI in MS is prevalent but objective cognitive assessments are not always available. Further work was needed to determine whether the MSNQ-I can be used as a proxy cognitive assessment in MS. This study aimed to utilise the Rasch analysis to assess the psychometric properties of the MSNQ-I and then identify which patient and disease factors are associated with informants' scores on the Rasch-analysed MSNQ-I.

Summary of Findings

1) Rasch-Analysed MSNQ-I

The MSNQ was developed in 2003 using standard techniques of item pool generation from qualitative interviews and then item reduction by Rasch analysis. Benedict et al. (2003) found four conceptual domains amongst the items of a) attention and speed of processing, b) memory, c) other cognitive ability, and d) personality and behaviour. Rasch analysis was performed on a sample of 102 MS patients (notably not any informants) using the BIGSTEPS software and mean square infit and outfit statistics. Rasch analysis has advanced considerably in the past 20 years and the original analysis simply would not have had the power to detect misfit or multidimensionality which might be clinically meaningful.

The current study using the latest Rasch techniques and software on much larger sample sizes of the partners and carers of PwMS, found that the MSNQ-I did not have a strictly unidimensional structure. Only after deleting the last three items (item 13: "Does he/she have difficulty controlling his/her impulses?"; item 14: "Does he/she laugh or cry with little cause?"; item 15: "Does he/she talk excessively or focus too much on his/her own interests?"), the revised

MSNQ-I (MSNQ-I-12) satisfied the psychometric requirements of the Rasch model. Items 13, 14 and 15 were combined into one smaller scale and Rasch-analysed, but they did not form a separate scale which would have been valid for measurement. These three items are related to executive function (a set of mental process needed for controlling one's behaviours, e.g., social and emotional control, stopping ongoing action; Perone et al., 2021), and although not explicitly stated in the original scale development paper, likely represented the personality and behaviour domain identified in the qualitative work. Following in-depth discussions with the experts in the fields of neuropsychology, neurology and psychometrics, it was concluded that these three items should be removed from the scale. Whether they truly represented a separate latent construct is an empiric question, but it is possible, and these items certainly did not conform mathematically to the rest of the scale in the analysis and their removal improved the measurement properties of the remaining items.

2) Multiple Regression

The MSNQ-I-12 was not significantly correlated with patients' age or MS duration. These findings were expected and in line with the findings of other studies (Akbar et al., 2011; Benedict et al., 2003; Konstantinopoulou et al., 2018; Langdon et al., 2013; Migliore et al., 2021; O'Brien et al., 2007; Vanotti et al., 2009). In the current study, male patients were rated by informants as having more cognitive difficulties than female patients on the MSNQ-I-12. This contradicted the finding from a study of Migliore et al. (2021) which reported no correlation between patient gender and the MSNQ-I score. A possible explanation for it might be that men with MS are more cognitively impaired than women with MS on objective cognitive tests of verbal learning and memory (Donaldson et al., 2019). The association between relatives' reports

of patient cognition and verbal memory tests was previously reported (i.e., the higher the reported CI, the lower the verbal memory score; e.g., Akbar et al., 2010; Benedict et al., 2004; Fenu et al., 2021; Vanotti et al., 2009).

The multiple linear regression revealed that patients' gender, patients' level of disability (EDSS band), self-reports of patients' anxiety (HADS A), patients' fatigue (NFI-MS), health-related quality of life (EQ-5D) and patients' disadvantage experienced as a result of ill health (LHS) explained 28% of the variance in the relatives' perceptions of the cognition of PwMS (MSNQ-I-12). Despite self-reports of patients' depression (HADS D) significantly correlating with the MSNQ-I-12, HADS D was not a significant predictor in the multiple regression model. This suggests that the shared variance between the MSNQ-I-12 and patients' self-reports of depression is accounted for by other variables.

Review of the Theoretical Model

Considering the results of the study (i.e., that about a quarter of the relative's reported score on the MSNQ-I was linked to the patient's gender, physical disability, anxiety, fatigue, how their illness affects their daily life and how their health affects their quality of life) and the theoretical model of Morris and Mograbi (2013) discussed in the introduction, it is possible that certain patient demographic and disease characteristics may be feeding into a relative's internal model of PwMS' cognition. Relatives take into account these variables and they get updated in their PwMS' database (Figure 1). This might contribute to a more accurate estimation of patient cognition. Mazancieux et al. (2019) suggested that patients' inacurrate reports of their cognition are associated with depression, fatigue and objective cognitive impairment as well as that they

are more reflective of PwMS' beliefs and worries. This suggests that there is a difference between informants' and patients' reports of PwMS' cognition, such that the same variables (e.g., patient fatigue, mood) simultaneously may act as predictors of informants' reports and confounders of patients' self-reports. Mazancieux et al. (2019) also proposed that patients with MS types associated with greater cognitive impairments are less likely to be aware of their cognitive impairments. In contrast, the implications of greater cognitive impairments may be more noticeable for PwMS' informants in day-to-day situations and therefore reported on the MSNQ-I. However, this study only focused on patient and disease variables, and it would be of interest to understand how PwMS' relatives' factors might be contributing to MSNQ-I scores. Particularly because Mazancieux et al. (2019) highlighted the evidence from other conditions that informants' depression and carer burden are associated with informants' ratings. Nonetheless, for the variables explored in the current study the adapted theoretical model of Morris and Mograbi (2013) offered a reasonable theoretical underpinning for relatives' metacognitive awareness of PwMS' cognitive status.

Strengths of This Study

The sample in this current study was reasonably representative of the MS population in the UK and worldwide. In the current study there were almost three times more females with MS than males. This was consistent with the gender ratio in MS (e.g., Català-Senent et al., 2023; McGinley et al., 2021). In relation to participants' age, the average patient age of 51.3 was higher than in other observational studies conducted in the UK (mean = 43.5, Dobson et al., 2021; mean = 41.8, Jick et al., 2015). This could represent sampling differences. Participants'

ethnicity data was not collected in the TONiC study. The lack of data on MS patients' ethnicity is a wider issue in England as reported by Public Health England (2020).

This dataset included data on participants' EDSS bands but not their exact EDSS scores. As these bands are not commonly used in studies, it was difficult to compare the distribution of levels of disability in this study with other studies and MS population in general. However, it appears that the more severe range of the disease was underrepresented.

The current sample had a higher percentage of married individuals (80.5%) compared with the UK MS Register (UKMSR), one of the largest repositories of PROMs in Europe (57.0%; Nicholas et al., 2020). In terms of patients' employment, in the current study the participants working full-time constituted 11.2% of the total sample. This was a lower percentage than reported in the UKMSR (42.0%), however it did not capture those PwMS who worked part-time. The UKMSR also used a different classification system ("active" vs "inactive") than the current study which might be contributing to the discrepancy in percentages.

In relation to MS types, in the current study RRMS was the most commonly reported MS type (58.0% including REMS) and this was also the most commonly reported MS type in other UK studies (e.g., 77.0%, Jick et al., 2015; 43.0%, Nicholas et al., 2020). In the current study 27.4% were SPMS, which fell between the worldwide prevalence (22.4%; Vasanthaprasad et al., 2022) and the UK prevalence of SPMS (47.7%; Vasanthaprasad et al., 2022). The proportion of patients with PPMS in the current study (14.1%) was consistent with other UK studies (e.g., 13.9%, Jick et al., 2015; 11.0%, Nicholas et al., 2020).

Another important strength of this study was a large sample size. This meant that the study had an excellent statistical power for both Rasch and multiple regression analyses. This

increases the confidence in the validity of the obtained results. Furthermore, the methodology in this study was another strength. The use of a well-established modern measurement theory method (Rasch analysis) enabled assessment of the existing scale and development of a revised version of the questionnaire, with superior psychometric properties. Since the fit to Rasch model was achieved, the ordinal-to-interval measurement conversion table (nomogram) is generalisable across samples and can be used in clinics and other studies.

Limitations

There were some sample biases in the current study. First, not everyone with MS from each centre volunteered to take part in the TONiC study and certain sociodemographic subgroups might be underrepresented (e.g., those who had caring or work responsibilities). Second, this unique dataset extracted from the TONiC database only included those participants who had a corresponding MSNQ-I score provided by their informants. In this sample over 70% of informants were spouses. Given that people with MS experience higher rates of relationship breakdown (Neate et al., 2019), PwMS who were not in a relationship were underrepresented. Third, the sample that the study used for the multiple regression excluded participants who had missing data on some variables. To address this both whole sample and multiple regression sample were compared on all variables and no significant differences were found. But they might differ on characteristics not measured. Fourth, there was a subtle overrepresentation of those with 10 or more years of disease duration and therefore of those with high levels of CI. CI may shift with disease duration, with significant differences in cognitive functioning between those who had MS for 5-10 years versus over 10 years (Freedman et al., 2023). These high levels of CI might have been more noticeable for PwMS' informants and led to higher scores on the MSNQ-

I-12. Fifth, the data has been collected across six years (2013-2019) which could be affecting the findings. For instance, the management of the disease changed over the years as new medications have been approved to treat RRMS (e.g., highly-efficacious monoclonal antibody treatments; Tillery et al., 2017), with different administration routes and timings. This could have impacted the health status of patients and therefore the comparability of the data within the sample. During that time period the 2010 McDonald MS diagnostic criteria were replaced with the revised 2017 McDonald criteria, which could have changed the accuracy and consistency of diagnosis of MS over time (Beesley et al., 2018).

Deleting items from a published and already widely-used scale in order to achieve fit to the Rasch model is not an ideal solution, however it is not uncommon in the process of scale refinement (e.g., Hadžibajramović et al., 2022; Huang et al., 2022; Pellicciari et al., 2020). All reasonable attempts were made to retain the removed items (i.e., examining if they form a separate scale, using testlets and using a bifactor solution). The 12-item scale had better measurement properties than the original 15-item MSNQ-I.

Additionally, there was no objective cognitive patient data for comparison with the MSNQ-I-12, which would have given a useful perspective on the validity of the MSNQ-I. Previous non-significant associations between relatives' reports and patients' objective cognitive tests (e.g., Langdon et al., 2013; Rosti-Otajärvi et al., 2014; Sejbæk et al., 2018) suggest that this might have introduced some bias to study results. This study was also limited in terms of the variables explored and it could only investigate variables chosen for the original TONiC study (e.g., there was no information about patients' ethnicity).

Future Directions

The MSNQ-I-12 can be used in clinics and future studies to identify and monitor CI in PwMS. It is recommended that clinicians and researchers administering the MSNQ-I exclude the final three items when calculating the total score. Retaining all 15 items for measurement would substantially increase measurement error because the original MSNQ-I measures two different concepts. Items 13, 14 and 15 might be administered to gather additional qualitative information as a part of the assessment.

Future studies could implement longitudinal designs to explore how relatives' perceptions of cognition of PwMS (MSNQ-I-12) change over time and which factors might be contributing to it. This study also prompts future studies to include objective cognitive testing to compare objective scores with subjective informants' scores on the MSNQ-I-12.

Conclusions

The MSNQ-I-12 has a unidimensional structure, and its psychometric properties are superior to those of the original MSNQ-I. The deleted three items of the MSNQ-I appear to be measuring executive functioning of PwMS and capture behavioural disturbances. The MSNQ-I-12 can be used as a proxy for objective cognitive assessment in MS, however the score should be carefully interpreted in the context of patients' gender, patients' level of disability (EDSS band), self-reports of patients' anxiety (HADS A), patients' fatigue (NFI-MS), health-related quality of life (EQ-5D) and patients' disadvantage experienced as a result of ill health (LHS). The MSNQ-I-12 raw score should be transferred to interval level conversion using the nomogram provided in this study.

Chapter III: Systematic Review

"The Associations of the Multiple Sclerosis Screening Questionnaire for Informants with Patients' Scores on Neuropsychological Assessments and Depression Scales: A Systematic Review"

Abstract

The Multiple Sclerosis Neuropsychological Questionnaire (MSNQ; Benedict et al., 2003) is a brief test about cognition in multiple sclerosis (MS) during activities of daily living. There are two versions of this questionnaire (patient's self-report: MSNQ-P and relative's report: MSNQ-I). It has been suggested that the MSNQ-I may be a useful alternative to objective cognitive testing, however how it relates to the objective cognitive testing has not been systematically assessed. A systematic literature search of PubMed and PsycINFO databases was conducted. The review included studies published between the MSNQ publications date (2003) and the date of the search (19th September 2022). Studies were included if they reported associations between the MSNQ-I and patients' objective cognitive test scores. There were twenty-two studies included in the systematic review, all cross-sectional and quantitative. The Effective Public Health Practice Project (EPHPP; Thomas et al., 2004) tool for quantitative studies was used to assess the quality of each study. This systematic review found that the correlations of different cognitive categories (information processing speed, auditory memory, visual memory, language, executive functioning, and visual perception/spatial processing) with MSNQ-I and MSNQ-P were inconsistent. All significant correlations were negative which means that higher MSNQ scores (i.e., more reports of cognitive difficulties) were associated with lower performance on

objective cognitive tests (i.e., a greater cognitive decline). The MSNQ-I correlations with cognitive tests were significant more frequently than the MSNQ-P correlations with cognitive tests. The patient's version of the MSNQ on its own may not reliably reflect patients' objective cognitive profiles. The MSNQ-I, however, may be a useful screening tool of patients' cognition for the healthcare professionals working with PwMS when the scores are considered in the context of other factors, including patients' depressive symptoms.

Introduction

Cognitive impairment (CI) in multiple sclerosis (MS) is common and it can affect patients' quality of life (QoL). Standard cognitive assessments with MS patients are not always feasible and, thus patients' cognitive difficulties might not be addressed in clinics. Objective cognitive assessments for patients incur significant resource costs. Using shorter screening instruments may be a useful, alternative indicator of cognitive abilities. The Multiple Sclerosis Neuropsychological Questionnaire-Informant (MSNQ-I; Benedict et al., 2003), which is completed by MS patients' carers/relatives, is a well-validated example of such a measure. Further work is needed to determine whether the MSNQ-I can be used as a proxy cognitive assessment in MS. It is important to understand how other aspects of MS patients' disease (e.g., duration, physical disability, type of MS) contribute to relatives' perceptions of MS patients' cognitive status. In other diseases, patients' characteristics are correlated with relatives' reports of patients' cognition. Associations between relatives' reports of patients' CI and patients' age and education were reported in a systematic review investigating factors related to informant-reported cognitive decline in older adults (Morrell et al., 2019).

MS is an autoimmune neurological condition of an unpredictable and progressive course affecting the central nervous system (McGinley et al., 2021). Different subtypes of MS include clinically isolated syndrome (CIS), relapsing-remitting multiple sclerosis (RRMS), secondary progressive multiple sclerosis (SPMS), primary progressive multiple sclerosis (PPMS) and benign multiple sclerosis (BMS). CIS is defined as the first episode of MS lasting at least 24 hours. CIS may or may not lead to development of MS. In RRMS, periods of acute flare-ups of symptoms, lasting at least 24 hours (relapses) are followed by lengthy periods of full or almost full recovery (remission). SPMS follows from initial RRMS, when disability starts to accumulate

at a variable rate with or without relapses, minor remissions, and plateaus. PPMS presents with a gradual accumulation of disability from the outset. BMS can be described as a type of RRMS with very mild attacks between long periods without any MS symptoms. It was estimated that in 2016 2,221,188 individuals were living with MS globally and that the second-highest age-standardised MS prevalence estimates per 100,000 population were in western Europe (Wallin et al., 2019). There is currently no cure for MS and the existing treatments only slow progression. MS can lead to a significant disability and involve a range of physical and psychological symptoms, including cognitive difficulties (McGinley et al., 2021).

Over the years increasingly more attention has been given to the topic of cognition in MS. Cognition is an umbrella term which encompasses many different mental processes. Although not all MS patients have the same cognitive profile, the most frequently affected domains in MS are information processing speed (IPS), memory, visual perceptual functions, attention, and executive skills (Benedict et al., 2020). Given that 45-70% of MS patients have cognitive difficulties and people are diagnosed with MS in their 20s-30s, many patients experience cognitive difficulties for the rest of their lives. Therefore, the topic of cognition in MS is important.

CI in MS is often overlooked despite its significant impact on patients and their systems. The challenges of CI can be particularly difficult for those who are at the starting a family and/or building a career stage of their lives. Cognitive deficits in MS impair people's social cognition (Dulau et al., 2017), driving ability and safety (Morrow et al., 2018), the performance of activities of everyday living (Goverover et al., 2007) and money management skills (Tracy et al., 2017). Cognitive decline in MS is also a factor adversely affecting disease management, including adherence to medication (Washington & Langdon, 2022), rehabilitation outcome

(Langdon & Thompson, 1999) and employment (Campbell et al., 2017). In a systematic review investigating the relationship between cognition and employment in working-age adults in MS, it was reported that those who were unemployed or had reduced their working hours had a higher level of CI on objective cognitive testing, when compared with those who maintained their employment status or working hours (Benedict et al., 2016). Cognitive difficulties in MS are also correlated with the economic burden of the disease, namely a high cost to the healthcare and community systems (Maltby et al., 2022). Cognitive difficulties in MS can decrease patients' QoL due to their psychosocial and physical consequences (Gil-González et al., 2020; Lakin et al., 2021), hence understanding the extent of cognitive challenges in MS is important.

Assessments of cognition in MS at all stages of the disease are important. The extent of neurological disability is typically measured by the Expanded Disability Status Scale (EDSS; Kurtzke, 1983), however, EDSS alone fails to reliably capture cognition (Saccà et al., 2017). Perceived and objective cognitive functioning in MS was found to be associated with patients' QoL (Crouch et al., 2022). Routine cognitive testing in MS may enable the identification of patients needing early intervention that can improve patients' cognitive functioning (DeLuca et al., 2020), and therefore their QoL (Gil-González et al., 2020; Lakin et al., 2021), as well as act as an additional clinical sign of disease activity (Benedict et al., 2020). Lastly, a recent qualitative study examining MS patients' opinions on patient-reported outcomes found CI to be one of the most important domains of living with MS that needs to be addressed (Westergaard et al., 2022). The assessment of cognition is therefore pivotal for both health professionals managing the disease as well as MS patients.

There are different options for assessing cognition in MS. These include objective testing, self-reports, relatives' reports, and clinic staffs' informal evaluations. There are some difficulties

with relying on the healthcare professionals' (HCPs') judgements of MS patients' cognition. CI, unlike physical disability, is not a visible symptom of MS. This makes cognitive difficulties difficult to detect (Lakin et al., 2021). Since clinic staffs' reports of cognition in MS are unreliable, additional cognitive evaluation is required to obtain an accurate picture of cognition.

Objective MS cognitive batteries assess several cognitive domains. The Brief Repeatable Battery of Neuropsychological Tests (BRB-N; Rao, 1991) takes approximately 45 minutes to complete, and it assesses IPS, working memory, verbal memory, visuospatial memory, and verbal fluency. The Minimal Assessment of Cognitive Functioning in Multiple Sclerosis (MACFIMS; Benedict et al., 2002) takes approximately 90 minutes to complete and, includes some of the BRB-N tests, and additionally assesses executive function and visuospatial perception. The Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS; Langdon et al., 2012) can be completed in 15 minutes, and it assesses IPS, verbal memory and visuospatial memory.

Completing a standard cognitive assessment with every MS patient on a regular basis (for example, annually) is not feasible in most MS services. Most of the available cognitive assessment batteries are long and expensive. Moreover, it is recommended that most of these assessments are completed by a qualified neuropsychologist (Ruet & Brochet, 2020), but not all MS patients have access to a neuropsychologist's expertise (Foley et al., 2012). There may be issues with the interpretation of the objective cognitive testing results since fatigue (Davenport et al., 2022; Tur, 2016), depression (Golan et al., 2018) and anxiety (Morrow et al., 2016) can have subtle bias effects on objective cognitive test performance. The results from standardized cognitive assessments completed in a controlled environment might not truly reflect patients' everyday functioning (Ruet & Brochet, 2020). Given that cognitive testing is expensive, time-

consuming and in short supply, short screening instruments, like the MSNQ, may be a useful indicator of cognitive status.

The MSNQ is an example of one of the most commonly used subjective measures of MS cognitive impairment in published work (Elwick et al., 2021). It is a brief test consisting of 15 questions about cognition during activities of daily living. There are two versions of this questionnaire, one self-report version for patients (MSNQ-P) and the other one for informants (e.g., carers, family members; MSNQ-I). On the MSNQ both patients and informants use a five-point scale to rate the occurrence of 15 items related to concrete examples of patients' cognitive functioning (e.g., distractibility, slowed processing, forgetting what is read). Higher scores on the MSNQ represent a bigger impact of cognitive deficits on daily functioning. There are multiple advantages of the MSNQ, for example, it can be completed at a distance, online, alone, quickly, it is widely available, and it has been validated in many countries.

However, patients' self-reports of cognition cannot reliably replace objective testing (Akbar & Finlayson, 2021). Fatigue, one of the most prevalent self-reported MS symptoms (Tur, 2016), and depression are common confounding variables of patients' subjective reports of cognition (Hughes et al., 2019). MSNQ-P scores are significantly correlated with patients' depression (Davenport et al., 2022), but only slightly with their objective cognitive performance (e.g., Akbar et al., 2011; Benedict et al., 2004; O'Brien et al., 2007). Conversely, MSNQ-I scores are usually correlated with patients' cognitive performance on objective tests (e.g., Charest et al., 2020; Fenu et al., 2018; Migliore et al., 2021; O'Brien et al., 2007), with a few exceptions (no correlation, Sejbæk et al., 2018; minimal correlation, Konstantinopoulou et al., 2018). The MSNQ-I has also been found to predict follow-up objective cognitive testing (Benedict & Zivadinov, 2006). Although a low to moderate correlation between the MSNQ-I and the MSNQ-

P has been shown (e.g., Benedict et al., 2003; Dagenais et al., 2013; Konstantinopoulou et al., 2018; Migliore et al., 2021; Sejbæk et al., 2018), this has not been a consistent finding. Significant discrepancies have been reported between scores on the MSNQ-P and the MSNQ-I (e.g., Fenu et al., 2018; Sonder et al., 2013).

The MSNQ-I has been used to measure cognitive function in isolated cognitive relapses (ICRs) in an MS scientific study (Meli et al., 2020). ICRs are transient cognitive deficits (Morrow et al., 2011). Patients with ICR were reported to have both lower MSNQ-I and objective cognitive tests scores, compared to patients without ICR (Meli et al., 2020). Although MSNQ-I data is being reported in scientific studies of cognition, properties of caregivers' perception of cognition of PwMS are not fully understood. Caring for someone can be very difficult and the CI of PwMS significantly contributes to caregivers' reports of high-stress levels (Halstead et al., 2021). Caregivers' perceptions of patients' cognition could contribute to MS clinic management and the MSNQ-I may be a useful measure, given that it correlates with scores on objective cognitive tests. Understanding what contributes to relatives' perspectives of MS patients' cognition is needed before the MSNQ-I can be adopted as a proxy for objective cognitive assessment in MS.

Previous small MSNQ-I studies from around the world have shown inconsistent correlations of the MSNQ-I with patients' characteristics. In terms of the MSNQ-I and patients' mood, patients' depression has been reported to correlate with the MSNQ-I (Akbar et al., 2010, 2011; Benedict et al., 2004; Langdon et al., 2013; Migliore et al., 2021; Sejbæk et al., 2018; Sonder et al., 2012), but not consistently (Benedict et al., 2003; Benedict & Zivadinov, 2006; Dagenais et al., 2013; Fenu et al., 2018; Fenu et al., 2021; Konstantinopoulou et al., 2018; O'Brien et al., 2007; Vanotti et al., 2009). Higher scores on patients' depression scale were also

associated with higher scores on the MSNQ-I (Carone et al., 2005). The MSNQ-I has been shown to correlate with patients' anxiety (Akbar et al., 2011; Sonder et al., 2012), but again, not consistently (Fenu et al., 2018; Fenu et al., 2021). Patients' neuropsychiatric symptoms (such as frequency and severity of delusions, hallucinations, aggression, euphoria, apathy, disinhibition) reported by an informant were also reported to correlate with the MSNQ-I (Benedict & Zivadinov, 2006; Carone et al., 2005). Lastly, patients' self-reported fatigue has not been found to correlate with the MSNQ-I (Langdon et al., 2013), although this study was with CIS patients, at a relatively early stage of the disease, when fatigue may not be so prominent.

In terms of the relationship between MSNQ-I and patient demographics, the MSNQ-I was shown to correlate with patients' gender (Sonder et al., 2012), but not consistently (Migliore et al., 2021). Patients' age was unrelated to the MSNQ-I (Akbar et al., 2011; Benedict et al., 2003; Langdon et al., 2013; Migliore et al., 2021; O'Brien et al., 2007), except in one study (Sonder, et al., 2012). Patients' higher age, however, has been significantly positively correlated with the difference between scores on the MSNQ-P and the MSNQ-I (Fenu et al., 2018). Patients' education has been shown to correlate with the MSNQ-I (Migliore et al., 2021; Sonder et al., 2012), but again not consistently (Akbar et al., 2011; Benedict et al., 2003; Langdon et al., 2013; O'Brien et al., 2007). The MSNQ-I did not correlate with patients' premorbid IQ (Akbar et al., 2011; O'Brien et al., 2007). MSNQ-I scores were found to be significantly higher for unemployed MS patients compared to employed patients (Akbar et al., 2011; Benedict & Zivadinov, 2006).

PwMS' disease factors and MRI biomarkers have also been inconsistently correlated with the MSNQ-I. The MSNQ-I has not been shown to correlate with time since CIS onset (Langdon et al., 2013) or MS duration (Akbar et al., 2011; Benedict et al., 2003; Konstantinopoulou et al.,

2018; O'Brien et al., 2007; Vanotti et al., 2009), except in one study (Sonder et al., 2012). Relatives of patients with SPMS reported a significantly greater CI on MSNQ-I compared to relatives of patients with RRMS (Benedict & Zivadinov, 2006). Relatives' perceptions of PwMS' CI on the MSNQ-I was correlated with patients' brain volume as measured by MRI (whole brain, gray matter, cortical gray matter, Fenu et al., 2021; T1 lesion volume, T2 lesion volume, brain parenchymal fraction, Benedict & Zivadinov, 2006), except for the white matter volume (Fenu et al., 2021).

Many PwMS report that their physical disability affects how their cognitive status is perceived. Although CI in MS is only mildly associated with a physical disability (Lynch et al., 2005), low levels of physical disability unduly enhance the perception of cognitive competence. CI of PwMS with a low physical disability may be undetected by neurologists during a routine examination in 25% of patients (Saccà et al., 2017). Saccà and colleagues (2017) also reported that neurologists failed to identify CI during routine clinical practice in two out of three PwMS when EDSS was not supplemented by at least a brief cognitive test. Relatives' reports of CIS cognitive status may also be less accurate because there is a low level of physical disability (Langdon et al., 2013). There is also evidence from other conditions that caregivers' perceptions of patients' functional independence are impacted by external factors, such as a patients' diagnosis, independent of objective cognitive status (Schmidt & Steffen, 2020).

The MSNQ-I has been shown to correlate positively with the EDSS (Migliore et al., 2021; Sonder et al., 2012; Vanotti et al., 2009), but not consistently (Akbar et al., 2011; Benedict et al., 2003; Konstantinopoulou et al., 2018; Langdon et al., 2013; O'Brien et al., 2007; Sejbæk et al., 2018). A positive correlation has been shown between patients' disability level as assessed by EDSS and the difference in perceptions of CI between relatives and patients, as assessed by

the MSNQ-I and MSNQ-P (Fenu et al., 2018). Thus, there is a complex and inconsistent relationship between MSNQ-I scores and MS patients' variables reported in previous small studies which highlights the need for further investigation. Although it is evident that CI in MS is often missed by both HCPs and carers, research to date has not yet yielded a definitive view of how other aspects of the disease affect how cognition is perceived by relatives.

The MSNQ-P and the MSNQ-I have been used to assess the awareness of CI in PwMS (Mazancieux et al., 2019). However, that study looked at the association between objective cognitive scores and patients' self-reports of cognition (using a range of different measures) as well as explored the discrepancies between two versions of the MSNQ as a way of measuring PwMS' awareness of their cognitive status. The current study is different in that it is a comprehensive exploration of both versions of the MSNQ and their relationships with patients' objective cognitive testing as well as each other. This study focuses only on the MSNQ questionnaires as they are free, easily available, quick to complete and is it hoped that this work will establish if they can be adopted in clinics to assess PwMS' cognition.

Summary

The aim of this systematic review is to synthesise available MSNQ-I data and explore how the MSNQ-I relates to patients' characteristics and patients' objective cognitive testing. No other published reviews exist in this area despite many authors highlighting the importance of taking informants' reports into account (e.g., Fenu et al., 2021; Migliore et al., 2021). Understanding which aspects of patients and their disease are associated with carers'/relatives' perceived cognition may help refine relatives' reports of MS patient cognition and optimise

MSNQ-I use in clinics, thus improve clinical assessment, monitoring, and management of MS cognitive difficulties.

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement was followed (Moher et al., 2009).

Systematic Literature Search Strategy

The following search terms were agreed upon between two reviewers (KM and DL) and used as keywords in titles and/or abstracts of the PubMed and PsycINFO electronic databases on the 19th September 2022:

(1) "multiple sclerosis" or "MS" or "Clinically Isolated Syndrome" or "CIS"

(2) "multiple sclerosis neuropsychological questionnaire" OR "multiple sclerosis neuropsychological screening questionnaire" OR "MSNQ" OR "MSNQ-I" OR "I-MSNQ" OR "cgMSNQ" OR "cg-MSNQ" OR "MSNQcg" OR "MSNQp" OR "MSNQ-P" OR "pMSNQ" OR "p-MSNQ" OR "MSNQ-S" OR "P-MSNQ"

Additionally, to establish any full-text publications unidentified by the database search, internet searches were conducted, reference lists of articles were scanned and an expert in the field (DL) was consulted.

Eligibility Criteria

Studies were required to meet all of the following criteria:

- Study of the MSNQ-I published since 2003 in peer-reviewed journals and written in English
- Reporting data from informants' version of the MSNQ
- Including contemporaneous objective cognitive assessment of patients
- Reporting at least one correlation of objective cognitive test/battery with the MSNQ-P and/or the MSNQ-I
- Reporting at least one patient disease, demographic, or psychosocial characteristic
- Patient participants 18 years of age and older, with any clinical subtype of MS (or subtype combination) and/or CIS

Study Selection

A citation software was used to remove duplicate records. All titles and abstracts were screened to examine the relevance of studies. The full texts of all relevant studies were accessed through online journals and Inter-Library Loan and read to assess the eligibility.

Data Extraction

The reviewers co-created a table to support the consistency of the process of extraction of relevant information from articles and determining articles' eligibility to be included in the review. One reviewer (KM) extracted the data from studies and both reviewers decided whether

studies met systematic review criteria. Following data extraction, 68 studies were excluded, and 22 studies were shortlisted.

Relevant study characteristics were extracted from the studies: which country the study was conducted in, sampling method and sample size. Patients' disease, demographic and psychosocial characteristics were extracted. These consisted of mean duration of MS, MS phenotype, physical disability, age, gender, education, and employment. The MSNQs' psychometric properties and the correlations between MSNQ scores and patients' objective cognitive tests and depression questionnaires were extracted. All data in this review were acquired from the information presented in text, tables, and graphs.

Quality Assessment

The Effective Public Health Practice Project (EPHPP; Thomas et al., 2004) tool for quantitative studies was used to assess the quality of each study. Both reviewers completed the ratings of all studies independently and any disagreements were discussed and resolved.

Results

Study Selection

A total of 22 studies were included in the current systematic review. Figure 7 presents the selection process.

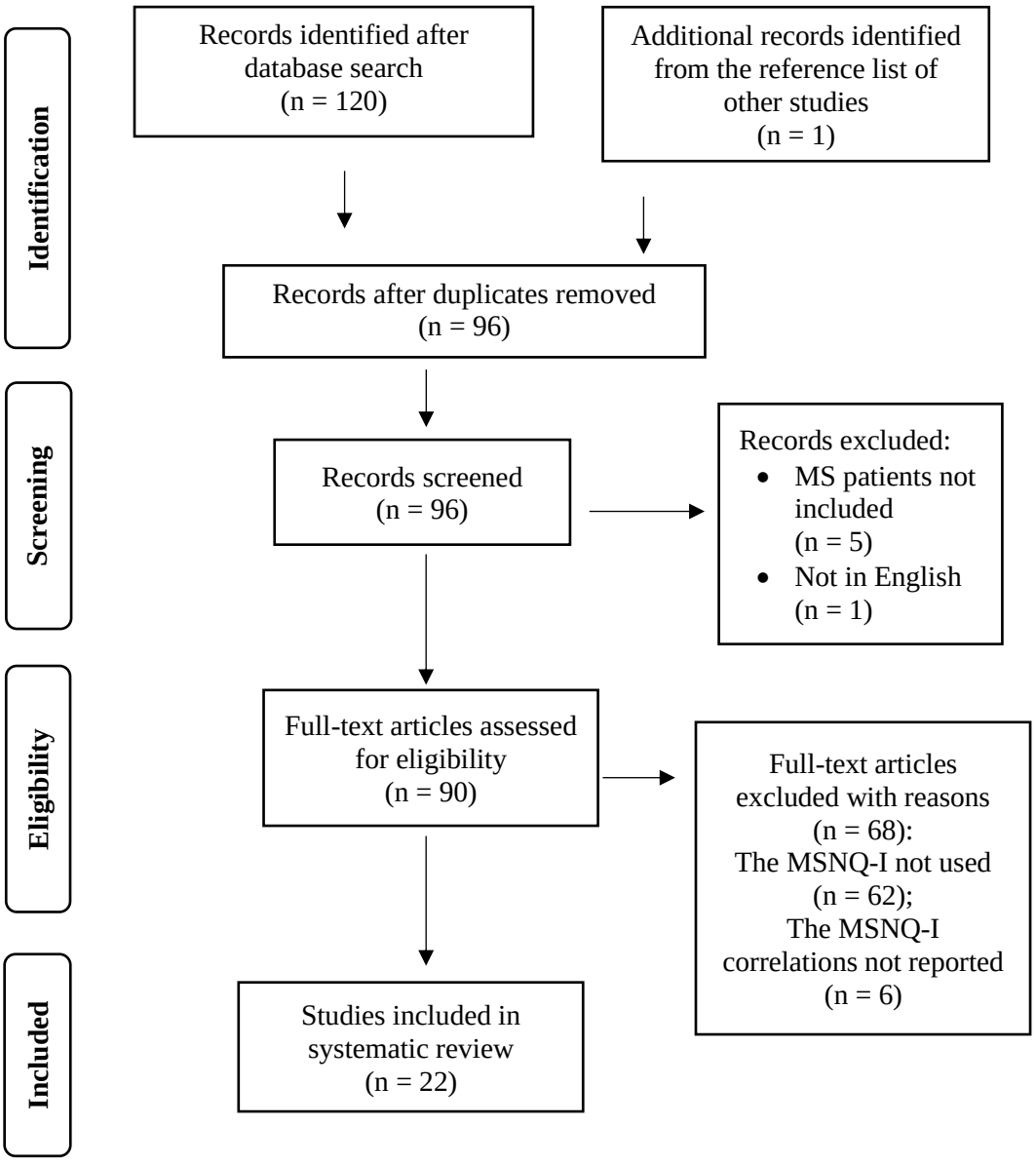


Figure 7

PRISMA Flowchart for the Process of Selecting Studies in the Systematic Review

Data Extraction

Tables 12, 13 and 14 show the relevant data extracted from the 22 included studies which met the inclusion criteria of this systematic review. The data comprised of study characteristics, sample demographic information, patients' disease information as well as MSNQs' psychometric properties and the correlations between MSNQ scores and patients' objective cognitive tests and depression questionnaires. For clarity, all verbal fluency tests in the Table 14 (Word List Generation, Controlled Oral Word Association Test, Verbal Fluency) were called VF.

Table 12

Study Characteristic, Sample Demographic and Patients' Disease Information

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Akbar et al. (2010)[a]									
MSNQ-P	Canada; Outpatient MS clinic and through advertisement	82	44.5 (8.9)	78	15.0 (2.2)	43	0/62/21/7/0	9.5 (7.4)	{2.0} [0.0–8.5]
MSNQ-I		82	Nr	Nr	Nr	Nr	-	-	-
Akbar et al. (2011) [a]									
MSNQ-P	Canada; Outpatient MS clinic and through advertisement	108	45.0 (9.0)	75	15.0 (2.0)	Nr	0/68/23/2/7	9.8 (8.1)	{2.3} [0.0-8.5]
MSNQ-I		108	Nr	Nr	Nr	Nr	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Benedict et al. (2003)[b]*									
MSNQ-P	USA; Randomly from an urban MS clinic	50	42.9 (6.7)	66	14.9 (2.2)	Nr	0/80/20/0/0	10.5 (7.0)	3.4 (2.0) [1.0-7.5]►
MSNQ-I		50	Nr	Nr	Nr	Nr	-	-	-
Benedict et al. (2004)									
MSNQ-P	USA; Patients from four MS clinics	85	42.4 (9.3)	80	14.8 (2.3)	Nr	0/80/20/0/0	Nr	2.5 (0.0-7.5) ►►
MSNQ-I		85	Nr	Nr	Nr	Nr	-	-	-
Benedict & Zivadinov (2006)									
MSNQ-P	USA; Consecutive clinical referrals/ volunteers	162	43.4 (8.6)	75	14.5 (2.3)	27	0/74/21/4/1	Nr	Nr
MSNQ-I		147	Nr	Nr	Nr	Nr	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Carone et al. (2005)									
MSNQ-P	USA;	122	44.0	72	14.5 (2.1)	Nr	0/72/24/2/2	12.0	2.5
	Clinical evaluation or		(8.8)					(8.0)	[0.0-8.0]
MSNQ-I	previously participating	122	Nr	Nr	Nr	Nr	-	-	-
	in the MSNQ studies								
Charest et al. (2020)									
MSNQ-P	Canada;	98	49.6	81	14.6 (2.8)	Nr	7/78/15/0/0	10.8	Nr
	University hospital MS		(11.4)					(7.6)	
MSNQ-I	clinic at a follow-up	98	Nr	Nr	Nr	Nr	-	-	-
Dagenais et al. (2013)									
MSNQ-P	Canada;	41	44.5	70	Nr	Nr	0/85/15/0/0	13.5	2.3
	A follow-up from the		(7.4)					(6.8)	(1.9)
MSNQ-I	MS clinic	41	Nr	Nr	Nr	Nr	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Erlanger et al. (2014)									
MSNQ-P	USA; Four MS centres	59	47.9 (7.9) [26-61]	72	Nr	Nr	0/77/23/0/0	13.2 (8.5) [1-33]	{2.5}
MSNQ-I		59	Nr	Nr	Nr	Nr	-	-	-
Fenu et al. (2018)									
MSNQ-P	Italy; Outpatients from the MS Centre	49	43.7 (11.9)	76	11.4 (4.1)	Nr	Nr	12.0 (7.8)	3.2 (2.1)
MSNQ-I		49	Nr	Nr	Nr	Nr	-	-	-
Fenu et al. (2021)									
MSNQ-P	Italy; Consecutive RRMS patients from the MS Centre	95	43.7 (11.9)	72	13.0 (3.5)	Nr	0/100/0/0/0	12.1 (7.8)	{2.0} [0.0-5.5]
MSNQ-I		95	49.5 (10.2)	63	12.3 (4.4)	Nr	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Konstantinopoulou et al. (2018)[b]									
MSNQ-P	Greece; MS centre at University Hospital	103	42.0 (9.9)	63	13.2 (2.9)	Nr	0/81/14/5/0	11.2 (7.0)	{4.5} [1.0-7.0]
MSNQ-I		60	Nr	Nr	Nr	Nr	-	-	-
Langdon et al. (2013)[c]									
MSNQ-P	France/Germany/ Russia/Spain/	130	Nr	Nr	Nr	Nr	100/0/0/0/0	0.7 (0.7)	1.3 (1.10)
MSNQ-I	Sweden/UK/ Switzerland/; Patients were part of an international observational trial	60	Nr	Nr	Nr	Nr	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Migliore et al. (2021)[b]									
MSNQ-P	Italy; Outpatient MS clinic	122	42.7 (12.9)	69.7	12.4 (3.3)	Nr	0/74/20/6/0	11.0 (7.3)	{RR = 1.0} [0.0-6.0] {SP = 6.5} [5.0-8.0] {PP = 6.0} [3.5-7.0]
MSNQ-I		122	Nr	Nr	Nr	Nr	-	-	-
O'Brien et al. (2007)									
MSNQ-P	USA; Two neurological clinics and the community through advertisements	48	45.1 (9.1)	80	14.7 (2.1)	48	0/69/21/10/0	14.6 (10.3)	3.7 (2.4)
MSNQ-I		Nr	Nr	Nr	Nr	48	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Ozakbas et al. (2018)									
MSNQ-P	Turkey; Six MS centres	462	35.3 (9.4)	69	11.3 (4.0)	Nr	0/100/0/0/0	7.8 (5.9)	1.9 (1.3)
MSNQ-I		462	Nr	Nr	Nr	Nr	-	-	-
Rosti-Otajärvi et al. (2014)									
MSNQ-P	Finland; Pool sample study from two previous studies	196	45.7 (9.1)	72	14.6 (3.3)	Nr	0/71/16/13/0	10.6 (7.3)	RRMS = 0.0-4.0 ((74.6)) 4.5-6.0 ((19.6)) ≥6.5 ((5.8)) SPMS & PPMS = 0.0-4.0 ((15.5)) 4.5-6.0 ((43.1)) ≥6.5 ((41.4))
MSNQ-I		Nr	Nr	Nr	Nr	Nr	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Sejbæk et al. (2018)[b]									
MSNQ-P	Denmark; Department of Neurology	77[ǂ]	45.6 [26-71]	57	Nr	Nr	Nr	7.8 [4.0-10.0]	2.8 [0.0-7.0]
MSNQ-I		77	Nr	Nr	Nr	Nr	-	-	-
Sonder et al. (2012)[b]									
MSNQ-P	Netherlands; University MS Centre at	121	52.8 (11.7)	62	Nr	Nr	3/36/33/28/0	16.9 (8.8)	{4.5} [3.0-6.5]
MSNQ-I		121	{55.0} [45-63]	40	Nr	Nr	-	-	-
Thomas et al. (2022a)									
MSNQ-P	USA; Archived data from another longitudinal study	87	51.8 (10.6)	75	14.8 (2.0)	Nr	0/63/30/7/0	13.6 (7.9)	4.3 (1.8)
MSNQ-I		87	Nr	Nr	Nr	Nr	-	-	-

Study	Country; Sampling method	N	Age in years Mean (SD), {median}, [range]	Female (%)	Education in years Mean (SD)	Employed (%)	MS phenotype (CIS/RR/SP/ PP/PR) %	Disease duration in years Mean (SD), {median}, [range]	EDSS Mean (SD), {median}, [range], ((% within range))
Vanotti et al. (2009)[b]									
MSNQ-P	Argentina; Participating centres in another MS study	125	42.3 (10.5)	67	13.7 (3.4)	Nr	0/86/10/3/1	8.8 (7.0)	3.3 (2.3)
MSNQ-I		125	Nr	Nr	Nr	Nr	-	-	-
Vanotti et al. (2018)									
MSNQ-P	Argentina; MS clinic	50	43.42 (10.17) [19-60]	74	14.9 (2.78)	60	0/78/18/4/0	13.1 (9.1) [1-40]	3.3 (2.6) {2.3} [0.0-8.0]
MSNQ-I		50	Nr	Nr	Nr	Nr	-	-	-

Note. CIS= Clinically Isolated Syndrome; EDSS= Expanded Disability Status Scale; MS=Multiple Sclerosis; MSNQ-I= Multiple Sclerosis informants' group; MSNQ-P= Multiple Sclerosis patients' group; Nr= Not reported; PPMS= Primary Progressive MS; PRMS= Progressive-Relapsing MS; RRMS= Relapsing-Remitting MS; SD= Standard deviation; SPMS= Secondary Progressive MS

* Benedict et al. (2003) study was the original scale validation study and only the final data is included in the table.

► this data was available for 37 patients

►► this data was available for 77 patients

[a] = Validation of Internet Version

[b] = Validation Study

[c] = Observational Study

Table 13

The MSNQs' Psychometrics

Study	Language	Mean score (SD)	Cronbach's α	Cut off score	Sensitivity (%)	Specificity (%)	Cognitively impaired (%)
Akbar et al. (2010)							
Online MSNQ-P	English	32.1 (10.0)	.91	Nr	Nr	Nr	35.4
Online MSNQ-I		26.7 (11.5)	.93	26	72	60	-
Akbar et al. (2011)							
Online MSNQ-P	English	32.9 (10.1)	Nr	≥26	Nr	Nr	36.0
Online MSNQ-I		27.3 (11.7)	Nr	Nr	Nr	Nr	-
Benedict et al. (2003)							
MSNQ-P	English	22.5 (10.2)	.93	Nr	Nr	Nr	24.0
MSNQ-I		18.4 (11.1)	.94	27	83	97	-
Benedict et al. (2004)							
MSNQ-P	English	27.4 (11.9)	.94	>23	80	68	35.0
MSNQ-I		21.3 (12.9)	.93	>22	87	84	-
Benedict & Zivadinov (2006)							
MSNQ-P	English	Nr	Nr	>23	56	94	Nr
MSNQ-I		Nr	Nr	>22	43	90	-
Carone et al. (2005)							
MSNQ-P	English	Nr	Nr	Nr	Nr	Nr	52.0
MSNQ-I		Nr	Nr	Nr	Nr	Nr	-
Charest et al. (2020)							
MSNQ-P	French	14.5 (5.9)	Nr	24	Nr	Nr	23.5
MSNQ-I		12.2 (4.8)	Nr	Nr	Nr	Nr	-
Dagenais et al. (2013)							
MSNQ-P	French	21.2 (10.1)	Nr	Nr	Nr	Nr	34.1
MSNQ-I		17.3 (11.8)	Nr	Nr	Nr	Nr	-

Study	Language	Mean score (SD)	Cronbach's α	Cut off score	Sensitivity (%)	Specificity (%)	Cognitively impaired (%)
Erlanger et al. (2014)							
MSNQ-P	English	Nr	Nr	Nr	Nr	Nr	42.0
MSNQ-I		Nr	Nr	Nr	Nr	Nr	-
Fenu et al. (2018)							
MSNQ-P	Italian	Nr	Nr	Nr	Nr	Nr	54.0
MSNQ-I		Nr	Nr	Nr	Nr	Nr	-
Fenu et al. (2021)							
MSNQ-P	Italian	Nr	Nr	Nr	Nr	Nr	53.7
MSNQ-I		Nr	Nr	Nr	Nr	Nr	-
Konstantinopoulou et al. (2018)							
MSNQ-P	Greek	16.7 (12.0)	.92	Nr	Nr	Nr	25.0 (in one cognitive domain); 3.0 (in two cognitive domain)
MSNQ-I		13.9 (11.3)	.93	Nr	Nr	Nr	-
Langdon et al. (2013)							
MSNQ-P	Various	15.5 (10.8)	.94	Nr	Nr	Nr	Nr
MSNQ-I		11.3 (9.6)	.93	Nr	Nr	Nr	-
Migliore et al. (2021)							
MSNQ-P	Italian	Nr	Nr	Nr	72	67	Nr
MSNQ-I		Nr	Nr	Nr	Nr	Nr	-
O'Brien et al. (2007)							
MSNQ-P	English	23.4 (11.2)	Nr	24	63	70	Nr
MSNQ-I		19.4 (12.7)	Nr	22	66	77	-
Ozakbas et al. (2018)							
MSNQ-P	Turkish	24.3 (12.0)	Nr	Nr	Nr	Nr	53.7
MSNQ-I		22.3 (13.1)	Nr	Nr	Nr	Nr	-
Rosti-Otajärvi et al. (2014)							
MSNQ-P	Finnish	26.2 (11.5)	Nr	≥27	Nr	Nr	≈ 50.0
MSNQ-I		20.5 (11.7)	Nr	≥27	Nr	Nr	-

Study		Language	Mean score (SD)	Cronbach's α	Cut off score	Sensitivity (%)	Specificity (%)	Cognitively impaired (%)
Sejbæk et al. (2018)								
	MSNQ-P	Danish	Nr	Nr	≥26	21	76	Nr
	MSNQ-I		Nr	Nr	≥26	33	66	-
Sonder et al. (2012)								
	MSNQ-P	Dutch	18.4 (10.3)	.93	Nr	Nr	Nr	10.0
	MSNQ-I		17.1 (11.2)	.94	21	75	71	-
Thomas et al. (2022a)								
	MSNQ-P	English	24.7 (9.7)	Nr	Nr	Nr	Nr	Nr
	MSNQ-I		20.0 (11.9)	Nr	Nr	Nr	Nr	-
Vanotti et al. (2009)								
	MSNQ-P	Spanish	18.1 (11.7)	.90	Nr	Nr	Nr	28.8
	MSNQ-I		17.2 (12.6)	.93	≥26	91	80	-
Vanotti et al. (2018)								
	MSNQ-P	Spanish	19.48 (13.6)	Nr	Nr	Nr	Nr	Nr
	MSNQ-I		16.18 (12.6)	Nr	Nr	Nr	Nr	-

Note. Nr= Not reported

Table 14

Correlations Between MSNQ Scores and Patient Objective Cognitive Tests and Depression Questionnaires

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
Akbar	Nr	**Online**	.62	.44	**SDMT**	-.10	-.27
et al. (2010)		**CES-D**	*(p < .01)*	*(p < .01)*		*(Ns)*	*(p < .05)*
					PASAT 2s	-.27	-.29
						(p < .01)	*(p < .01)*
					PASAT 3s	-.27	-.42
						(p < .01)	*(p < .01)*
					SRT-LTS	-.19	-.28
						(Ns)	*(p < .01)*
					SRT-CLTR	-.20	-.30
						(Ns)	*(p < .01)*
					SRT-D	-.23	-.34
						(p < .01)	*(p < .01)*
					10/36 SPART	-.13	-.24
						(Ns)	*(p < .05)*
					10/36 SPART-D	-.19	-.23
						(Ns)	*(p < .05)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					VF	-.16 *(Ns)*	-.34 *(p < .01)*
Akbar et al. (2011)	Nr	**HADS-D**	.45 *(p < .01)*	.27 *(p < .01)*	**SDMT**	-.09 *(Ns)*	-.14 *(Ns)*
					PASAT 2s	-.13 *(Ns)*	-.09 *(Ns)*
					PASAT 3s	-.20 *(p < .05)*	-.26 *(p < .01)*
					SRT-LTS	-.04 *(Ns)*	-.08 *(Ns)*
					SRT-CLTR	-.11 *(Ns)*	-.09 *(Ns)*
					SRT-D	-.14 *(Ns)*	-.20 *(p < .05)*
					10/36 SPART	-.10 *(Ns)*	-.20 *(p < .05)*
					10/36 SPART-D	-.09 *(Ns)*	-.23 *(p < .05)*
					VF	-.11 *(Ns)*	-.19 *(Ns)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					ANART VIQ	-.09 *(Ns)*	-.18 *(Ns)*
Benedict et al. (2003)	0.37 (*p < .01*)	**BDI**	.53 *(p < .01)*	Nr *(Ns)*	**PASAT**	Nr *(Ns)*	-.47 *(p < .001)*
					BVMT-R	Nr *(Ns)*	Nr *(Ns)*
					BVMT-R-D	Nr *(Ns)*	-.43 *(p < .001)*
		CES-D	.46 *(p < .001)*	Nr *(Ns)*	**CVLT-II**	Nr *(Ns)*	-.53 *(p < .01)*
					CVLT-II-D	Nr *(Ns)*	-.43 *(p < .001)*
					TMT-B	Nr *(Ns)*	.55 *(p < .01)*
					WCST Perseverative Responses	Nr *(Ns)*	.37 *(p < .001)*
					JLO	Nr *(Ns)*	Nr *(Ns)*
					Boston Naming Test	Nr *(Ns)*	-.45 *(p < .001)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
Benedict	Nr	**CES-D**	.61	.37	**BVMT-R**	-.46	-.57
et al. (2004)			*(p < .001)*	*(p < .001)*		*(p < .001)*	*(p < .001)*
		BDI-FS	Nr	Nr	**BVMT-R-D**	-.45	-.55
						(p < .001)	*(p < .001)*
					CVLT-II	-.37	-.45
						(p < .001)	*(p < .001)*
					CVLT-II-D	-.42	-.50
						(p < .001)	*(p < .001)*
					PASAT	-.38	-.59
						(p < .001)	*(p < .001)*
					SDMT	-.45	-.58
						(p < .001)	*(p < .001)*
					VF	-.17	-.33
						(Ns)	*(p < .01)*
					D-KEFS Sorting Correct Sorts	.30	-.38
						(Ns)	*(p < .01)*
					WCST Perseverative Responses	.29	.35
						(Ns)	*(p < .01)*
					JLO	-.44	-.36
						(p < .001)	*(p < .01)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					Composite Z	-.49 *(p < .001)*	-.64 *(p < .001)*
Benedict & Zivadinov (2006)	Nr	**BDI-FS**	.56 *(p < .001)*	Ns	**PASAT**	-.30 *(p <.001)*	-.46 *(p < .001)*
					SDMT	-.27 *(p < .01)*	-.55 *(p < .001)*
					BVMT-R	-.23 *(p < .01)*	-.46 *(p < .001)*
					BVMT-R-D	-.24 *(p < .01)*	-.48 *(p < .001)*
					CVLT-II	-.19 *(p < .05)*	-.49 *(p < .001)*
					CVLT-II-D	-.24 *(p < .01)*	-.50 *(p < .001)*
Carone et al. (2005) ▶	Nr	-	-	-	**PASAT**	Nr *(Ns)*	Nr *(Nr)**
					SDMT	Nr *(Ns)*	-.47 *(p < .001)*
					BVMT-R	Nr *(Ns)*	Nr *(Nr)**

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					CVLT-II-D	-.21 *(p = .03)*	Nr *(Nr)**
					VF	Nr *(Ns)*	-.31 *(p < .001)*
					WCST Perservative Responses	Nr *(Ns)*	Nr *(Nr)**
					JLO	Nr *(Ns)*	Nr *(Nr)**
Charest et al. (2020)	Nr	-	-	-	**MoCA**	Ns *(Nr)*	-.246 *(p = .017)*
					MACFIMS overall score	Ns *(Nr)*	-.278 *(p = .007)*
Dagenais et al. (2013)	.48 (*p = .001*)	**BDI-FS**	- .031 *(p = .847)*	.084 *(p = .601)*	**MoCA**	-.300 *(Ns)*	-.390 *(p = .012)*
					Executive/Processing Speed Factor	-.223 *(Ns)*	-.240 *(Ns)*
					Learning Factor	-.025 *(Ns)*	-.484 *(p < .001)*
					Delayed Recall Factor	-.069 *(Ns)*	-.463 *(p = .002)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
Erlanger et al. (2014)	Nr	-	-	-	**Total score of SRT, BVMT-R, PASAT and SDMT**	RRMS -.38 *(Nr)* SPMS .13 *(Nr)*	RRMS -.40 *(Nr)* SPMS -.37 *(Nr)*
Fenu et al. (2018)	Nr	**BDI**	.543 *(p = .001)*	Ns	**SDMT** **BVMT-R** **CVLT-II** **Total N of failed tests**	-.381 *(p < .004)* -.189 *(Ns)* -.269 *(p < .031)* .275 *(p < .028)*	-.521 *(p < .001)* -.423 *(p = .002)* -.338 *(p = .012)* .562 *(p < .001)*
Fenu et al. (2021)	Nr	**BDI-II**	.225 *(Ns)*	.008 *(Ns)*	**SDMT** **BVMT-R** **CVLT-II**	-.349 *(p < .001)* -.217 *(p < .01)* -.300 *(p < .001)*	-.451 *(p < .001)* -.328 *(p < .001)* -.328 *(p < .001)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					BICAMS overall score	-.317 *(p < .01)*	-.416 *(p < .001)*
					Total N of failed tests	.168 *(Ns)*	.401 *(p < .001)*
Konstantinopoulou et al. (2018)	.53 (*p < .001*)	**BDI**	.315 *(p = .025)*	Ns	**EPST**	-.233 *(Ns)*	-.049 *(Ns)*
					DS	-.173 *(Ns)*	-.388 *(p < .01)*
					DSF	-.158 *(Ns)*	-.227 *(Ns)*
					DSB	-.108 *(Ns)*	-.362 *(p < .01)*
					GVLT-D	-.007 *(Ns)*	-.303 *(p < .05)*
					VFTS	-.163 *(Ns)*	-.287 *(p < .05)*
					VFTP	-.170 *(Ns)*	-.301 *(p < .05)*
					STROOP W	-.112 *(Ns)*	-.050 *(Ns)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					STROOP C	-.036 *(Ns)*	-.175 *(Ns)*
					STROOP CW	.151 *(Ns)*	.070 *(Ns)*
Langdon et al. (2013)	.46 (*p < .01*)	**CES-D**	.35 *(p < .01)*	.37 *(p < .01)*	**PASAT 3s**	.09 *(Ns)*	-.23 *(Ns)*
					SDMT	.01 *(Ns)*	.00 *(Ns)*
					SRT-LTS	.00 *(Ns)*	-.08 *(Ns)*
					10/36 SPART	.00 *(Ns)*	.07 *(Ns)*
					VF	-.06 *(Ns)*	-.39 *(Ns)*
Migliore et al. (2021)	.36[a] (*p = .00005*)	**BDI**	.24[a] *(p* = .000001)	.25[a] *(p* = .005)	**SDMT (errors)**	.005[a] *(Ns)*	.01[a] *(Ns)*
					SDMT (correct)	-.027[a] *(p < .002)*	-.35[a] *(p < .0001)*
O'Brien et al. (2007)	Nr	**BDI**	.42 *(p = .01)*	Nr *(Ns)*	**Digit symbol**	Nr *(Ns)*	-.39 *(p < .05)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					PASAT 2s	Nr *(Ns)*	-.41 *(p < .05)*
					PASAT 3s	Nr *(Ns)*	-.40 *(p < .05)*
					SDMT	Nr *(Ns)*	-.41 *(p < .01)*
					Symbol search	Nr *(Ns)*	-.35 *(p < .05)*
					DSF	Nr *(Ns)*	Ns *(Ns)*
					DSB	Nr *(Ns)*	Ns *(Ns)*
					LM Immediate	Nr *(Ns)*	Nr *(Ns)*
					LM Delayed	Nr *(Ns)*	Nr *(Ns)*
					SRT trials to criterion	.39 *(p < .01)*	.35 *(p < .05)*
					VF corrected score	Nr *(Ns)*	Nr *(Ns)*
					STROOP CW	Nr	-.36

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
						(Ns)	*(p < .05)*
					WCST perseverative responses	Nr	Nr
						(Ns)	*(Ns)*
					WRAT-3 Reading Subtest	Nr	Ns
						(Ns)	*(Ns)*
					JLO	-.38	-.39
						(p < .05)	*(p < .05)*
					Processing Speed Index	Nr	-.39
						(Ns)	*(p < .05)*
Ozakbas et al. (2018)	Nr	**BDI**	Nr	Nr	**PASAT**	-.202[a]	-.184[a]
						(p < .001)	*(p = .006)*
					SDMT	-.220[a]	-.288[a]
						(p < .001)	*(p < .001)*
					SRT- Immediate	-.213[a]	-.138[a]
						(p < .001)	*(p = .036)*
					SRT-D	-.132[a]	-.049[a]
						(p = .004)	*(p = .464)*
					10/36 SPART- immediate	-.122[a]	-.148[a]
						(p = .008)	*(p = .025)*
					10/36 SPART-D	-.103[a]	-.127[a]
						(p = .023)	*(p = .054)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
Rosti-Otajärvi et al. (2014)	Group A: .51 (*p < .001*) Group B: .68 (*p < .001*)	**BDI-II**	Group A: .413 *(p < .001)* Group B: .538 *(p < .001)*	Group A: .249 *(p = .005)* Group B: .370 *(p = .011)*	**BRB composite Z score (Group A)** **BRB composite Z score (Group B)**	.177 *(p = .038)* .189 *(p = .166)*	.035 *(p = .699)* .247 *(p = .098)*
Sejbæk et al. (2018) ►►	0.22[b] (*p = .0001*)	**BDI**	.25[b] *(p < .0001)*	.197[b] *(p < .0001)*	**Digit-Symbol Coding** **RAVLT** **TMT-B** **WCST** **Boston Naming Test**	.000385[b] *(p < .8663)* .00713[b] *(p < .4651)* .01447[b] *(p < .2974)* .010923[b] *(p < .3689)* .002[b] *(p < .6993)*	.058501[b] *(p < .0604)* .02064[b] *(p < .2652)* .00327[b] *(p < .6588)* 2.158e-6[b] *(p < .9910)* .0085[b] *(p < .8217)*
Sonder et al. (2012)	.52[a] *(Nr)*	**HADS-D**	.49 *(Nr)*	.36 *(Nr)*	**PASAT 3s** **PASAT 2s**	-.12[a] *(Nr)* -.27[a] *(Nr)*	-.14[a] *(Nr)* -.34[a] *(Nr)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					SDMT	-.06[a] *(Nr)*	-.02[a] *(Nr)*
					SRT-LTS	-.28[a] *(Nr)*	-.29[a] *(Nr)*
					SRT-CLTS	-.22[a] *(Nr)*	-.34[a] *(Nr)*
					SRT-D	-.17[a] *(Nr)*	-.36[a] *(Nr)*
					10/36 SPART - immediate	-.21[a] *(Nr)*	-.37[a] *(Nr)*
					10/36 SPART-D	-.13[a] *(Nr)*	-.32[a] *(Nr)*
					VF	-.17[a] *(Nr)*	-.30[a] *(Nr)*
					BRB-N global (amount impaired out of 5 subtests)	.26[a] *(Nr)*	.39[a] *(Nr)*
Thomas et al. (2022a)	.60 *(Nr)*	**BDI-FS**	.39 *(Nr)**	Nr *(Ns)*	**Processing speed component** **Verbal memory component**	1.37[c] *(p = .25)* 6.39[c] *(p = .01)*	2.52[c] *(p = .12)* 2.13[c] *(p = .15)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
					Focused attention component	2.94[c]	.03[c]
						(p = .09)	*(p = .86)*
					Visual memory component	6.81[c]	9.85[c]
						(p = .01)	*(p = .002)*
					Executive functioning component	7.26[c]	9.85[c]
						(p = .01)	*(p = .002)*
Vanotti	Nr	**BDI-FS**	.30	.23	**SDMT**	-.05	-.29
et al. (2009)			*(p < .0001)*	*(Ns)*		(*Ns)*	*(p < .001)*
					PASAT 3s	-.17	-.37
						(*Ns)*	*(p < .0001)*
					SRT-LTS	-.34	-.32
						(p < .001)	*(p < .001)*
					SRT-CLTR	-.23	-.26
						(*Ns)*	*(p < .01)*
					7/24 SPART	-.09	-.31
						(*Ns)*	*(p < .001)*
					7/24 SPART-D	-.12	-.24
						(*Ns)*	*(p < .001)*
					VF	-.14	-.36
						(*Ns)*	*(p < .0001)*

Study	MSNQ-P and MSNQ-I r (*p*)	Depression scale	MSNQ-P and depression r *(p)*	MSNQ-I and depression r *(p)*	Objective cognitive test/battery	MSNQ-P and test/battery r *(p)*	MSNQ-I and test/battery r *(p)*
Vanotti et al. (2018)	Nr	**BDI-FS**	Nr	Nr	**SDMT**	-.292 *(p < .05)*	-.498 *(p < .01)*
					BVMT-R	-.158 *(Ns)*	-.227 *(Ns)*
					CVLT	-.294 *(p < .05)*	-.457 *(p < .01)*

Note. 10/36 SPART= 10/36 Spatial Recall Test; 10/36 SPART-D= 10/36 Spatial Recall Test-Delayed; 7/24 SPART= 7/24 Spatial Recall Test; 7/24 SPART-D= 7/24 Spatial Recall Test-Delayed; ANART VIQ= American National Adult Reading Test Verbal Intelligence; BDI= Beck Depression Inventory; BDI-FS= Beck Depression Inventory - Fast Screen for Medical Patients; BVMT-R= Brief Visuospatial Memory Test Revised; BVMT-R-D= Brief Visuospatial Memory Test Revised Delayed; CES-D= Centre for Epidemiological Studies Depression Scale; CIS= Clinically Isolated Syndrome; CVLT-II= California Verbal Learning Test Second Edition; CVLT-II-D= California Verbal Learning Test Second Edition Delayed; DS= Digit Span Test Total; DSB= Digit Span Test Backwards; DSF= Digit Span Test Forwards; EDSS= Expanded Disability Status Scale; HADS-D= Hospital Anxiety and Depression Scale-Depression Subscale; JLO= Judgement of Line Orientation Test; LM= Logical Memory; MACFIMS= Minimal Assessment of Cognitive Function in MS; MoCA= Montreal Cognitive Assessment; MS= Multiple Sclerosis patient group/Multiple Sclerosis; MSNQ-I= Multiple Sclerosis Neuropsychological Questionnaire-Informant Version; MSNQ-P= Multiple Sclerosis Neuropsychological Questionnaire-Patient Version; Nr= Not reported; Ns= Non-significant; PASAT= Paced Auditory Serial Attention Test; RAVLT= Rey Auditory Verbal Learning Test; RRMS= Relapsing-remitting MS; SDMT= Symbol Digit Modalities Test; SPMS= Secondary Progressive MS; SRT-CLTR= Selective Reminding Test- Continuous Long-Term Retrieval; SRT-D= Selective Reminding Test-

Delayed; SRT-LTS= Selective Reminding Test-Long-Term Storage; STROOP C= Stroop Colour-Word Test Total Score On Colour Condition; STROOP CW= Stroop Colour-Word Test Score On Colour-Word Condition; STROOP W= Stroop Colour-Word Test Total Score On Word Condition; TMT-B= Trail Making Test Part B; VF= Verbal Fluency; WCST= Wisconsin Card Sorting Test; WRAT-3= Wide Range Achievement Test-3

Information Processing Speed Tests; Verbal Memory Tests; Visual Memory Tests; Language Tests; Executive Functioning Tests; Visual Perception/Spatial Processing Tests

*p**= reported as significant, no *p* value given

[a]Spearman's 'Rho'

[b] R^2

[c] Regression coefficient

► – Objective cognitive tests data comes from *n* = 116

►► – Objective cognitive tests data comes from *n* = 77

Group A - relapsing MS, *n* = 138

Group B - progressive MS, *n* = 58

Quality Assessment

The EPHPP was used to assess the quality of each study in six categories: selection bias, study design, confounders, blinding, data collection method and withdrawals or drop out (Table 15). Each component was given a "weak", "moderate" or "strong" rating and then each study was given an overall global rating using the same scale. All studies received a "weak" overall quality rating, apart from one which received a "moderate" rating.

The selection of participants was not free of bias and the results may therefore be not generalisable to the target population. Only one study named the design adopted. Some studies considered confounders: seven controlled for at least 80% of relevant confounders and two for at least 60-79%. No studies addressed blinding. Regarding psychometric properties, seven studies reported the MSNQ to be both acceptably valid and reliable and a further nine reported only acceptable validity. The follow-up data relates to test-retest investigations and does not reflect the "attrition" of an intervention study. It is of note that the EPHPP was constructed to review public health interventions (Thomas et al., 2004) and some questions on the EPHPP assume that there is more than one group of participants in a study. As a result of the lack of suitable responses to some questions, a few studies included in this systematic review which only had one group of participants were disadvantaged and awarded lower ratings.

Table 15

Quality Assessment of Studies Included in the Systematic Review

Study	Selection bias	Study design	Confounders	Blinding	Data collection method	Withdrawals and drop out	Overall quality rating
Akbar et al. (2010)	Moderate	Weak	Strong	Moderate	Moderate	Strong	Moderate
Akbar et al. (2011)	Moderate	Weak	Strong	Moderate	Weak	Moderate	Weak
Benedict et al. (2003)	Moderate	Weak	Weak	Moderate	Moderate	Weak	Weak
Benedict et al. (2004)	Moderate	Weak	Strong	Moderate	Strong	Weak	Weak
Benedict & Zivadinov (2006)	Weak	Weak	Strong	Moderate	Strong	Weak	Weak
Carone et al. (2005)	Weak	Weak	Strong	Moderate	Moderate	Moderate	Weak
Charest et al. (2020)	Moderate	Weak	Strong	Moderate	Strong	Weak	Weak
Dagenais et al. (2013)	Moderate	Weak	Weak	Moderate	Moderate	Weak	Weak
Erlanger et al. (2014)	Moderate	Weak	Weak	Moderate	Moderate	Weak	Weak
Fenu et al. (2018)	Weak	Weak	Weak	Moderate	Moderate	Moderate	Weak
Fenu et al. (2021)	Weak	Weak	Weak	Moderate	Moderate	Weak	Weak
Konstantinopoulou et al. (2018)	Weak	Weak	Weak	Moderate	Strong	Moderate	Weak
Langdon et al. (2013)	Moderate	Moderate	Weak	Moderate	Weak	Weak	Weak
Migliore et al. (2021)	Moderate	Weak	Weak	Moderate	Strong	Moderate	Weak
O'Brien et al. (2007)	Moderate	Weak	Strong	Moderate	Strong	Weak	Weak
Ozakbas et al. (2018)	Moderate	Weak	Weak	Moderate	Weak	Moderate	Weak
Rosti-Otajärvi et al. (2014)	Moderate	Weak	Weak	Moderate	Weak	Moderate	Weak
Sejbæk et al. (2018)	Moderate	Weak	Moderate	Moderate	Weak	Strong	Weak
Sonder et al. (2012)	Weak	Weak	Weak	Moderate	Moderate	Moderate	Weak
Thomas et al. (2022a)	Weak	Weak	Weak	Moderate	Weak	Moderate	Weak
Vanotti et al. (2009)	Moderate	Weak	Moderate	Moderate	Strong	Weak	Weak
Vanotti et al. (2018)	Moderate	Weak	Weak	Moderate	Moderate	Weak	Weak

Study Characteristics

There were seven studies carried out in the USA, four in Canada, three in Italy, two in Argentina and one in each of Netherlands, Greece, Turkey, Finland, Denmark as well as one internationally (France, Germany, Russia, Spain, Sweden, Switzerland, UK). Of the 22 studies, nine studies were partially or fully funded by a biotechnology or pharmaceutical company and seven by an MS Society/Foundation/Research Association. Three studies were partially or fully funded by a health research body, two by a government organisation and one by a higher education institution. There were five studies with no external funding. All studies were quantitative and had a cross-sectional design. There were eight MSNQ validation studies, two of which validated an Internet version. The sample size ranged from 41 to 462, indicating that some samples were small. In 19 studies participants were recruited from outpatient MS clinics or MS centres. In five studies participants were, either exclusively or in addition to recruitment from MS clinics, recruited from a pool of patients who either at that time or previously took part in other MS studies. In three studies participants were recruited through advertisements in addition to recruitment from MS clinics or MS centres.

Overview of Study Parameters

There are a considerable range of parameters and reporting schedules across the studies. For example, sample selection is on the basis of differing inclusion and exclusion criteria. Most significantly, some studies only included MS patients with subjective cognitive complaints (e.g., Rosti-Otajärvi et al., 2014). Others only included MS patients with no subjective cognitive complaints (e.g., Charest et al., 2020). More generally, inclusion and exclusion criteria were

individual to each study. One study recruited patients from both arms of a randomised controlled trial (RCT) evaluating the efficacy of the beta-interferon medication Rebif, after no treatment effect was demonstrated (Charest et al., 2020). This is concerning because cognition was not assessed as an outcome in the RCT and patients are highly selected for RCT's, being a significantly healthier subset of patients than the general clinic sample.

The other variability was in evidence across cut-offs adopted for the MSNQ-P and MSNQ-I, making sensitivities and specificities difficult to compare. Two studies employed an internet version of the MSNQ (Akbar et al., 2010, 2011). Not all studies employed a validated language version of the MSNQ (e.g., Fenu et al., 2018). The neuropsychological tests used to formally evaluate cognition also varied and the criteria for designating cognitive impairment were inconsistent (e.g., Fenu et al., 2018, reported cognitive impairment in any patient failing at least one BICAMS scale; Konstantinopoulou et al., 2018, gave separate percentages impaired on one, two or three domains; Sonder et al., 2012, chose three or more impaired subtests on the BRB-N to indicate cognitive impairment). Some studies compared MS patient cognitive performance with contemporary healthy control cognition data (e.g., Akbar et al., 2010; Benedict et al., 2003). Lastly, the relationship of the informants was not routinely reported, or frequency of contact with the patient. This heterogeneity made a meta-analysis inappropriate and is considered in more detail below.

Sampling

There were various levels of selection in samples which affects the comparability of results. The majority of participants were recruited from MS clinics, which could be influenced

by demographics and health insurance in some countries; therefore, these samples might not be representative of the whole MS population. In the study of Charest et al. (2020) two arms of a drug trial were combined and those patients who agreed to additionally complete the MACFIMS as well as reporting no cognitive impairment on the MSNQ-P (score below 24) were included. One observational study funded by a pharmaceutical company only included CIS patients who were recruited within 3 years of a CIS event, with no confirmed diagnosis of MS and had mild and subtle cognitive deficits. In contrast, participants in the Benedict and Zivadinov's (2006) study were referred for cognitive assessment, suggesting that they might have had some cognitive difficulties which led to the referral for the assessment. Similarly, in the study of Carone et al. (2005) participants were included if they had at least one severe and one mild impairment, or at least three mild impairments, on cognitive testing. It is possible that whilst in some studies there was an artificially high percentage of cognitively impaired participants, in other studies cognitively impaired patients were not selected for participation. This means that the samples in those studies did not include MS patients with the full spectrum of various levels of cognition, as it would occur in the general MS population. It is also of note that patients with a BDI-FS score greater than 4 (indicating possible depression) were not included in the study of Dagenais et al. (2013), which once again, limits the confidence with which comparisons can be made with other studies that included depressed patients.

Sample Characteristics

Participants' mean age ranged from 35.4 to 52.8 years and mean years of education ranged from 11.3 to 15.0. In all included studies, apart from one which did not report participants' genders, at least 57% of MS patients were females. Only five studies reported

participants' ethnicity and in all of them at least 87% were Caucasian. Only two studies reported participants' marital status. Most commonly reported patients' demographic characteristics were Caucasian married females. In four studies which reported participants' employment status there were between 27 and 60% of patients in either full- or part-time employment. Fifteen studies reported the percentage of patients impaired on objective cognitive testing, but each study used a different classification system. Unsurprisingly, the percentage of "cognitively impaired" participants varied widely across studies (10 to 54%). The majority of participants in the studies reporting higher percentage of "cognitively impaired" individuals were diagnosed with RRMS.

Patient Disease Information

Participants' mean duration of MS ranged from 7.8 to 16.9 years, and one study (Langdon et al., 2013) included only CIS patients with a mean duration of MS of 0.7 year. In all studies, apart from one (Langdon et al., 2013) with only CIS patients included, the most common phenotype was RRMS and second most common was SPMS. The mean EDDS score reported in 12 studies ranged from 1.3 to 4.3, meaning that the average disability level was able to walk unaided, and that moderate and severe disabilities were not represented in these studies.

Control Group

Out of nine studies reporting a contemporaneous healthy control cognitive data (Akbar et al., 2010, 2011; Benedict et al., 2004; Benedict & Zivadinov, 2006; Carone et al., 2005; O'Brien et al., 2007; Ozakbas et al., 2018; Rosti-Otajärvi et al., 2014; Vanotti et al., 2009), healthy control was not reported to have completed cognitive testing in the Benedict and Zivadinov's

(2006) study. The lack of control groups in the majority of included studies limits the validity of conclusions which can be drawn from the results.

Collection and Analysis Methods

All studies used a mixture of objective cognitive testing and self-report measures, including the MSNQ-P and MSNQ-I. Some studies included other self-report psychosocial questionnaires for patients (e.g., depression scales). In all studies data was analysed using appropriate statistical tests to examine the associations between MSNQ-I, MSNQ-P, and objective cognitive tests.

MSNQ

Patients' and informants' subjective reports of patients' cognitive functioning were measured by the MSNQ-P and MSNQ-I respectively. The mean score on the MSNQ-P was numerically higher than on the MSNQ-I in all studies reporting mean MSNQ scores. Significant positive correlations between the scores on patients' and informants' MSNQ versions were reported in seven studies, with the strength of these correlations ranging from 0.22 (weak) to 0.68 (moderate). The remaining 15 studies did not report correlations between MSNQ-P and MSNQ-I. Seven studies assessed the internal consistency of the MSNQ and reported the

Cronbach's α of MSNQ, where the range was .90 to .94 for the MSNQ-P and the range for the MSNQ-I was .93 to .94.

MSNQ and Depression Correlations

Patients' depression was measured in 19 studies using three different scales (CES-D, HADS, BDI) and two studies included more than one depression scale. Fourteen studies reported a significant positive relationship between the MSNQ-P and at least one patients' depression scale and two studies a non-significant relationship. As for the relationship between the MSNQ-I and at least one patients' depression scale, nine studies reported it to be non-significant whilst seven other studies reported it to be significant and positive. Three studies did not report depression correlations with the MSNQ-P or the MSNQ-I.

Cognitive Testing

Different types of cognitive tests were used in studies and details of how cognition was measured objectively are given in Table 14. In 17 studies correlations of MSNQ were reported with individual tests or subtests. In three studies tests assessing a similar cognitive domain were grouped together into one category of neuropsychological functioning and only a total score of each category was correlated with the MSNQ. In two studies correlations of MSNQ were reported with the total score on one of the two cognitive test batteries only (MACFIMS and BRB). To summarise the findings, where it was possible, individual cognitive tests used in the studies were grouped into the following categories: information processing speed tests, auditory memory tests, visual memory tests, language tests, executive functioning tests, and visual

perception/spatial processing tests. Some studies used more than one test within each category and the summary will include the count of all reported correlations, i.e., negative significant, positive significant, non-significant, significant with no direction of the correlation reported as well as non-reported correlations.

Cognitive Batteries and Screening Tests

The Brief Repeatable Battery of Neuropsychological Tests (BRB-N; Rao, 1991) takes approximately 45 minutes to complete, and it assesses IPS, working memory, verbal memory, visuospatial memory, and verbal fluency. The Minimal Assessment of Cognitive Functioning in Multiple Sclerosis (MACFIMS; Benedict et al., 2002) takes approximately 90 minutes to complete and, includes some of the BRB-N tests, and additionally assesses executive function and visuospatial perception. The Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS; Langdon et al., 2012) can be completed in 15 minutes, and it assesses IPS, verbal memory and visuospatial memory. Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) is a screening tool for detecting cognitive decline, takes 10-12 minutes to complete and it assesses language, memory, visual and spatial thinking, reasoning, and orientation skills.

There was a significant negative correlation of both versions of the MSNQ with BICAMS overall score. Only informants' version of MSNQ was significantly negatively correlated with the total scores on MACFIMS (once) and MoCA (twice). The association between MSNQ-P and BRB total score was non-significant (once); not reported (once) and once significant (albeit positive and weak) in the group of patients with RRMS.

The relationship between MSNQ-I and BRB total score was non-significant (twice) and not reported (once). Total number of failed tests (SDMT, BVMT-R, CVLT) was significantly positively correlated with MSNQ-I (twice), but with MSNQ-P the relationship was non-significant once and once significant in a positive direction.

Information Processing Speed

The relationship between the MSNQ-P and information processing speed tests was shown to be inconsistent. Higher scores on the MSNQ-P were associated with lower performance on information processing speed tests and this was reported to be significant 13 times. In contrast, this association was reported as non-significant 18 times and not reported at all three times. It was also reported to be non-significant when correlated with processing speed category in three studies.

The relationship between the MSNQ-I and information processing speed tests was reported to be negatively significant 23 times; reported as non-significant seven times; reported as significant with no direction one time and not reported three times.

Verbal Memory

The association between MSNQ-P and verbal memory tests was reported to be significant and negative 12 times. It was reported as non-significant 17 times and not reported three times. On one occasion a significant positive correlation was reported but there was a reversed scoring

in the verbal memory test (i.e., which means that the higher the score, the worse performance). This was in line with other findings reporting negative correlations with verbal memory tests.

A negative significant correlation between the MSNQ-I and verbal memory tests was reported 19 times. On one occasion a reversed scoring was used in the cognitive test and the significant positive correlation was in line with other significant correlations in this group of cognitive tests. This correlation was reported as non-significant nine times; significant with no direction once and not reported three times.

Visual Memory

The association between MSNQ-P and visual memory tests was reported to be significant and negative twice. It was reported as non-significant eight times and not reported twice.

A negative significant correlation between the MSNQ-I and visual memory tests was reported seven times. This correlation was reported as non-significant three times and not reported two times.

Language

Across all studies the relationship between language tests and MSNQ-P was consistently reported as non-significant (nine times) and not reported on one occasion.

A negative significant correlation between the MSNQ-I and language tests was reported six times. This correlation was reported as non-significant three times and not reported on one occasion.

Executive Functioning

The association between MSNQ-P and executive functioning tests was reported to be non-significant 13 times. This is the only correlation reported consistently across all studies.

A negative significant correlation between the MSNQ-I and executive functioning tests was reported twice. On three occasions a reversed scoring was used in the cognitive test (i.e., higher scores reveal greater impairment) and the significant positive correlations were in line with other significant correlations in this group of cognitive tests. This correlation was reported as non-significant seven times and significant with no direction once.

Visual Perception/Spatial Processing

The association between MSNQ-P and visual perception/spatial processing tests was reported to be significant and negative seven times and non-significant in a further seven studies.

A negative significant correlations between the MSNQ-I and visual perception/spatial processing tests were reported nine times. This correlation was reported as non-significant three times and twice as significant with no direction.

Premorbid Cognitive Functioning

The association between MSNQ-P and premorbid cognitive functioning tests was reported to be non-significant in three studies. A negative significant correlation between the MSNQ-I and premorbid cognitive functioning tests was reported once and non-significant correlations were reported twice.

Overall Findings

The correlations of different cognitive categories with MSNQ-I and MSNQ-P were inconsistent. All significant correlations were negative (unless the reversed scoring was used in cognitive tests). This means that higher MSNQ scores (i.e., more reports of cognitive difficulties) were associated with lower performance on objective cognitive tests (i.e., a greater cognitive decline). The MSNQ-I correlations with cognitive tests were significant more frequently than the MSNQ-P correlations with cognitive tests.

Patient-Informant Discrepancies of Reported Cognitive Symptoms

One study examined the discrepancies between scores on the MSNQ-P and MSNQ-I (Carone et al., 2005). Patients were classified in relation to their informants' ratings as either "overestimators" if they overestimated their cognitive abilities, "underestimators" or "accurate estimators". It was reported that patients over-estimating their cognitive abilities performed worse than underestimators and accurate estimators on most objective cognitive tests. There were no differences found in performance on objective testing between the underestimators and accurate estimators. This classification can be potentially a helpful way of identifying those patients who would benefit from an objective cognitive testing.

Sensitivity and Specificity of the MSNQ

Five studies reported the sensitivity and specificity of the MSNQ-P (ranges 21-80% and 67-94% respectively). The range of the specificity of the MSNQ-P is narrower than the range of its sensitivity. Three studies reported that the specificity of the MSNQ-P was higher than its sensitivity and two other studies reported the opposite. Eight studies reported the sensitivity and specificity of the MSNQ-I (range 33-91% and 60-97% respectively). Fifty percent of studies said that the specificity of the MSNQ-I is higher than its sensitivity. In two studies (Benedict et al., 2003, 2004) the sensitivity and specificity values of both MSNQ-P and MSNQ-I are acceptable and comparable to those reported in a previous study using objective cognitive testing (sensitivity = 68%, specificity = 85%; Rao, 1991). There are inconclusive findings on the sensitivity and specificity of both MSNQ-P and MSNQ-I. It cannot be concluded whether either of them is better able to correctly identify people with or without cognitive impairment.

Discussion

This systematic review aimed to synthesise MSNQ data from patients and relatives and explore how the MSNQ relates to patients' performance on objective cognitive testing and to other patients' characteristics. All twenty-two studies included in the systematic review were cross-sectional and quantitative. Patient participants completed objective cognitive assessments with investigators and the MSNQs were completed by patients and their informants.

The majority of correlations across all studies showed that patients' impairments in IPS, verbal memory, visual memory, language, visual perception/spatial processing were significantly negatively related to the MSNQ-I scores. The majority of correlations across all studies indicated no significant associations between MSNQ-P and patients' scores on objective cognitive tests. Nauta et al. (2019) reported that even excluding patients with depression from the analysis did not result in a significant relation between the MSNQ-P and objective test scores. This suggests that there might be other factors related to patients' perceptions of their cognition (e.g., fatigue; Tur, 2016 and depression; Hughes et al., 2019) which are beyond the scope of this systematic review. This indicates that the MSNQ-I is more reliably related to patients' objective cognitive status than the MSNQ-P.

Seven studies reported weak to moderate significant positive correlations between the scores on patients' and informants' MSNQ versions. This indicates that higher scores on the MSNQ-P are associated with higher scores on the MSNQ-I, however this association is not strong. This is not surprising, given the closer relation of MSNQ-I scores to objective test scores. The majority of included studies reported that depression was significantly and positively correlated with the MSNQ-P and non-significantly associated with the MSNQ-I. This shows that

patients' depression was related to patients' reports of cognitive status. However, patients' depression was inconsistently related to relatives' reports of patient cognitive status.

Seven studies assessed the internal consistency of the MSNQ, and both patients' and informants' versions of the MSNQ had acceptable levels of Cronbach's alpha (MSNQ-P range: .90 - .94, MSNQ-I range: .93 - .94). Cronbach's alpha is expressed as a number between zero and one. Although there is a consensus that the closer this value is to one, the higher the internal consistency of a scale, it has been recommended that a value of alpha should not exceed 0.90 (Streiner, 2003). Alpha value is affected by the number of items on the scale, how related these items are, as well as the dimensionality of the scale (i.e., how many underlying concepts a scale measures; Cortina, 1993). High Cronbach's alpha values of the MSNQ-P and MSNQ-I might be an indication that some items are redundant which means that they measure the same thing but use different wording.

The results of this systematic review suggest that patients' and informants' versions of the MSNQ may be assessing two different concepts as there was no association between the MSNQ-I and the MSNQ-P in the majority of studies. Whilst the MSNQ-I appears to be measuring PwMS' cognitive abilities, the MSNQ-P might be revealing patients' insight into their CI. Mazancieux et al. (2019) suggested that patient-informant discrepancies can be used to evaluate patients' awareness of cognitive abilities. Mazancieux et al. (2019) also proposed that patients' lack of awareness of their memory difficulties may be explained by patients' memory impairments and therefore their inability to recall what they are able/unable to do to. Perhaps the general lack of consistent associations between two versions of the MSNQ as well as between

the MSNQ-P and objective cognitive tests confirms that the MSNQ-P is not to be used to assess CI in PwMS.

Strengths

This review has several strengths. There were variations of the MSNQ measures' name used in studies. This might have prevented the retrieval of all studies using the MSNQ-I, therefore a range of possible variations of the measure's name were included in the search in an effort to capture all relevant data. This resulted in focused search terms and allowed for discrete findings of studies using the MSNQ. Including only studies with the MSNQ-I and MSNQ-P, where they had objective testing, was a very good way of evaluating the associations between objective and subjective cognitive measures. The robustness of this review was increased by having a second reviewer's (DL's) input in the screening and selection of studies as well as cross-checking the extracted data and assessing the quality of included studies. It is the first review to synthesise data on the MSNQ-P and MSNQ-I and their correlations with patients' objective test scores and depression. This is particularly important as there is a clinical role for the MSNQ-I and the MSNQ-I is now being used to gather data about cognition for scientific studies (Meli et al., 2020). Alongside exploring how both versions of the MSNQ relate to objective testing, each other, and patient depression, the review also presents their psychometric properties.

Limitations

The current review and included studies are not free of limitations. Most available quality assessment tools were designed for RCTs. This includes the EPHPP used in this review which was not ideal for the included studies. Even high-quality validation studies performed weakly on some components (e.g., study design).

Studies took place in different countries which could have led to some biases. For instance, in some countries there is a limited access to healthcare, for example diagnosis of MS and subsequent access to disease modifying drugs (DMDs; Berger et al., 2020). Even within Europe, there is a wide range of patient access to DMDs (26% to 79%, Kobelt et al., 2017). This means that patients with no or delayed access to medication experience disease progression more quickly than patients receiving DMDs. The disease progression, and its physical, cognitive, and psychosocial consequences, is therefore different for patients with and without access to medication. This could have influenced the relations of disease variables recorded to cognitive and other variables. Different countries have varied healthcare systems which also has implications for medication adherence. For example, in the USA, whether people can access neuropsychological assessments depends on their specific health insurance. Patients who do not have access to a neurologist are less likely to be referred for neuropsychological testing. Patients with no or little health insurance are more likely to be poor, live in rural areas, belong to ethnic minority groups, have had MS for over 15 years, or have mobility difficulties (e.g., cannot walk, use a wheelchair/scooter or are bedbound; Minden et al., 2008). These variations across the world in access to healthcare limit the validity of comparing the identified studies.

There were some limitations related to the designs of the included studies. The quality assessment revealed that the selection of participants was biased, and some samples were small.

Some studies recruited patients who were referred for cognitive testing hence those with cognitive difficulties might have been represented in disproportionately large numbers (e.g., in the studies of Benedict & Zivadinov, 2006; Carone et al., 2005). Other studies only included patients self-reporting no cognitive impairment on the MSNQ-P (Charest et al., 2020). This means that samples in those studies did not capture MS patients with cognitive impairment profiles representative of the general MS population. Thus, the validity of comparing data across identified studies and broader generalisation may be limited. Two studies (Akbar et al., 2010, 2011) collected the data over the Internet which meant that researchers had no control or way of knowing whether patients and their informants completed the questionnaires without collaboration. This raises a question about the validity of data due to the bias of excluding those who were poorer, less proficient in the use of modern technology and had sensorimotor difficulties. Not only were different objective cognitive tests used across studies, but also some studies reported participants' raw scores on the objective cognitive tests whilst other studies reported an overall cognitive domain score. The wide range of tests used across studies makes it difficult to compare the findings. Not all of the measures used within the included studies were validated for MS in the language or in the country where the research took place. This means that the results from different studies may not be comparable. For instance, studies used adaptations of original standardised cognitive tests (e.g., Vanotti et al., 2018 used an adapted SDMT) and the MSNQ measures (e.g., Fenu et al., 2018 translated the MSNQ to Italian and used it in a study without a prior validation in a larger Italian population). This might have compromised the results of those studies and therefore the conclusions in this systematic review since the validity and reliability of measures assessing patient reported outcomes should be evaluated in the target population before they are used (Boateng et al., 2018). Moreover, not all studies defined cut-off

scores for the MSNQ-P and MSNQ-I for a reliable discrimination of those with and without cognitive impairments. This was often reported to be due to small MS samples, no control group, or low sensitivity of the scale.

Lastly, studies underreported some characteristics of PwMS. Of all studies included in this review, only five reported participants' ethnicity. The Caucasian population was disproportionately represented. The underreporting of data on race and ethnicity as well as underrepresentation of non-White patients is a wider issue observed in MS disease modifying treatment trial publications (Onuorah et al., 2022). Similarly, only four studies reported participants' employment status and the majority (apart from one study of O'Brien et al., 2007) used a dichotomized employment status criterion (i.e., "employed" / "unemployed" or "disabled"). This is potentially overly restrictive and may not detect finer gradations in employment status, for example that people might not work for other reasons than disability or change their employment to a less demanding role or work part-time. There are different levels of support across countries for PwMS to stay in employment (within Europe, employment rates for people with MS vary from 31% to 65%, Kobelt et al., 2017), therefore how cognition is related to employment in MS may vary depending on the cultural context of the study. Studies also underreported participants' years of education. PwMS' education has been shown to be associated with their clinical outcomes related to MS (Dobson et al., 2022), therefore it is an important characteristic to consider when exploring factors linked to cognition. Moreover, most studies included patients with different MS subtypes but reported the results as if this was a homogenous group. It cannot be concluded with certainty whether the findings for those subtypes differ from each other.

Clinical Implications

The findings of this systematic review do not support the MSNQ-P as a reliable measure of patients' objective cognitive status. Clinicians assessing patients' cognition should not therefore rely solely on patients' self-reports. Although PwMS are not the most accurate judges of their own cognition and the MSNQ-P may not be the indication of PwMS' cognitive functioning, collecting this data is useful and important because it represents the perspective of patients. Completing the MSNQ-P as an outcome measure in MS clinics may also be a way of starting the conversation about cognition, a topic which is important but often neglected. The results from a qualitative study exploring neurologists' views on patient reported outcome measures (PROMs) in MS care revealed that clinical conversation was considered to be the most fundamental source of patient-reported information (Reitzel et al., 2022). The view that PROMs should encompass cognitive functioning measures and be integrated into clinical care is also shared by PwMS (Westergaard et al., 2022). The inconsistencies in correlations between MSNQ-P and objective cognitive testing mean however that the MSNQ-P cannot be relied on as a measure of cognition.

Although the MSNQ-I was associated with patients' objective cognitive testing more consistently than the MSNQ-P, the majority of these correlations were inconsistent. It should be noted that reports of no cognitive impairments on the MSNQ-P and MSNQ-I do not always mean that there is an absence of such difficulties.

It might be better for clinicians working in MS clinics to at least ask informants to complete the MSNQ-I than not to address cognition at all, if there are not enough resources to complete formal neuropsychological assessments with patients. The MSNQ-I might be a useful screening tool for identifying those patients who would benefit from an objective cognitive

testing, where it is available. This approach accepts that informants' ratings are a reliable source of information and there is evidence from other conditions that carers' ratings can be associated with patients' diagnosis (Schmidt & Steffen, 2020). Therefore, informants' reports should be interpreted carefully. It is of note that since most studies did not include PwMS with severe psychiatric comorbidities, the accuracy of the MSNQ-I in this population is unknown. The MSNQ-I may be a useful indication of patients' cognitive status where objective cognitive testing is not available. Whilst the MSNQ-I is relatively robust to patient mood, MSNQ-P is significantly related to patient mood and other psychosocial variables, which should be taken account of in clinical practice.

Research Implications

Future research should further explore the psychometric properties of the MSNQ, namely high Cronbach's alpha values and potential redundant items. Although deleting items from existing scales is not an ideal solution (it limits usefulness of previous research findings), shortening the length of the test was suggested to obtain an internal consistency value within the recommended range (Tavakol & Dennick, 2011). It should be noted that Cronbach's alpha value increases with the number of items on the test, regardless of the unidimensionality of the scale. Compared to the MSNQ-I, the MSNQ-P correlated more consistently with patients' depression scales than with patients' objective cognitive testing, whereas, compared to the MSNQ-P, the MSNQ-I correlated more consistently with patients' objective cognitive testing. Given these discrepancies in the associations of both versions of the MSNQ with patients' depression and patients' objective cognitive testing, it may be helpful to explore the dimensionality (i.e., the

number of underlying traits a scale measures) of the MSNQ. Future studies should focus on increasing the validity through careful and adequate recruitment strategies.

Further research would benefit from utilising a standardised objective cognitive battery, to increase validity of comparisons and generalisability of results, as part of a standardised international validation protocol. Researchers should ensure that patients with different ages, MS types and durations, ethnic backgrounds, levels of disability, employment and marital status are represented and that these characteristics are reported in studies. Researchers should also further explore informants' and patients' factors which might be associated with informants' perceptions of PwMS' cognitive abilities. This is because there may be confounding variables associated with either the MSNQ or patients' cognition and therefore the findings should be carefully considered in the context of these factors. It would be beneficial to conduct longitudinal research to explore whether the MSNQ can be used to reliably measure cognitive changes across the MS course.

Conclusions

The MSNQ-P on its own may not reliably reflect patients' cognitive profile. The MSNQ-I, however, may be a useful proxy screening tool of patients' cognition for the healthcare professionals working with PwMS when the scores are considered in the context of any possible confounding variables, including patient depressive symptoms.

Chapter IV: Integration, Impact, and Dissemination

Overview

The following section provides a synthesis of important methodological aspects of the systematic review (SR) and empirical study. Both the SR and the empirical article represent the development of a single topic and this section includes the integration of the SR findings with the conceptual basis for the empirical study. The chapter also consists of a summary of how the results of these two research components might impact people with multiple sclerosis (PwMS), clinical practice and future research. This section concludes with steps that will be taken to disseminate the findings by research publications.

Integration

The Development of the Project

In the early stages of the development of my thesis, I attended an MS-UK webinar "Cognition and MS" delivered by my supervisor, Prof Dawn Langdon for MS-UK on 20th October 2021. I heard people with MS ask questions about cognition and discuss their experiences. This helped me understand how important the topic of cognition is for MS patients. It also made me realise that this topic is often neglected. Having not directly worked with or met PwMS before starting this thesis, I made an effort to continue to increase my knowledge about cognition in MS whilst completing the project. I read the case studies from Prof Anthony Feinstein's book "Mind, Mood, and Memory: The Neurobehavioral Consequences of Multiple

Sclerosis" (Thompson & Feinstein, 2022). I found that Feinstein's case studies showed me the impact of cognitive difficulties on people's lives (e.g., vocational impact). I attended "A 3D Virtual Symposium Patient Considerations in MS: Achieving Optimal Outcomes Through an Individualized, Comprehensive Approach" on the 16th February 2022. This reinforced that there is a need for more holistic approaches to MS treatment, which includes addressing patients' cognition. On the 18th March 2022 I attended the Continuing Professional Development session "Cognition in MS" delivered by Prof Dawn Langdon during the Neuropsychology Special Interest Group's meeting. This provided me with an overview of the developments and gaps in research about MS cognition.

My supervisor and I agreed on conducting Rasch analysis as a part of the statistical analysis of the empirical study. As I was not familiar with this analysis and it is not a part of the course curriculum, I attended an online course to learn it.

My learning and reflection from these experiences gave me a useful context for the project and made me think about its implications for patients with MS and their families.

Overall Project Description

The SR chapter of this thesis aimed to synthesise all available information on how closely relatives' perceptions of PwMS' cognitive status, as measured by the Multiple Sclerosis Neuropsychological Questionnaire for Informants (MSNQ-I), is related to objective cognitive testing. This was to determine whether the MSNQ-I can be used as a proxy cognitive assessment in MS. The empirical study then aimed to evaluate and improve the psychometric properties of the MSNQ-I and explore which patient and disease factors could predict relatives' scores on the Rasch-analysed MSNQ-I. The empirical study extended the findings of the SR by providing the

recommendations on how the MSNQ-I can be best used and interpreted in the context of patient and disease factors.

Integration of SR and Empirical Study

I hoped that together both chapters of this thesis would provide an understanding of whether and how to best use the MSNQ-I in clinical practice and research as a measure of PwMS' cognition. The SR was completed before the empirical study. I wanted to summarise which aspects of patients' and their disease may be predictive of the relatives' perceptions of patients' cognition. The empirical study felt like a natural extension of the SR. The conclusion of the SR that the MSNQ-I can be used as a reliable indication of PwMS' cognition, informed the empirical study which examined how the MSNQ-I should be used (i.e., which items give a reliable score). The SR summarised the sensitivity and specificity of the MSNQ-I which provided background information for the psychometric properties further explored in the empirical study.

Challenges of the Project

There were some methodological challenges when designing the SR and deciding on the studies' inclusion criteria. The vast variability in patients' characteristics explored, and the types of objective cognitive measures used across studies made comparisons between studies difficult. I initially planned to explore the associations between the MSNQ-I and patients' characteristics and objective cognitive testing. I hoped that this would have had provided context for the empirical study investigating how the MSNQ-I relates to patients' disease, demographic, and

psychosocial characteristics in a large dataset. In the end the associations of the MSNQ-I and MSNQ-P with patients' objective cognitive testing and mood measures were included as they were the two most frequently reported correlations.

There were also some challenges with the empirical study, particularly with selecting which patients' factors, disease variables and Patient Reported Outcome Measures (PROMs) to analyse. The database included several other variables, including patients' education, religion, health status (e.g., about smoking, having comorbidities, MS relapses in the past year), recent time off sick from work, how often patients and their informants saw each other and how long they have known each other for. Unfortunately, only a small number of patients had this additional information and including them in the analysis would have decreased the number of cases and therefore decreased the statistical power of the analysis.

Perhaps the biggest challenge and learning for me was the fact that I worked with a large database. Unfortunately, I was restricted in my analysis only to those variables for which TONiC study collected data. This meant that I had no qualitative or qualifying information and that I encountered the issue of incomplete data. Although the dataset had a large number of participants, the data was collected over a period of almost a decade, and this may have contributed to the cohort effects. I recognise that despite these limitations and challenges, working with already collected data made the completion of this project more efficient and allowed me to have more time for data analysis and write-up stages of this thesis. I have not anticipated that the process of familiarising myself with the data, making sense of it, editing it to fit the SPSS format would be as time-consuming as it was. Completing a secondary data analysis, like this one, maximises the use of data which patients and their families have spent time and effort to provide. Those who contributed to the TONiC database were identified by

their existing clinical care team or the research team, so they have already been in contact with MS services. This means that those who were not seen in services were also not represented in this research. Although completing web-based questionnaires requires some digital literacy and can be excluding people without Internet access, the TONiC study offered an alternative solution to help overcome these issues (i.e., posting a questionnaire pack along with a pre-paid envelope). The uniqueness of the TONiC database is that it collected data related to cognition as well as a number of self-report questionnaires (PROMs). Most database studies are concerned with the health insurance claims and therefore only include patients who have health insurance and only records medical data (e.g., Amiri et al., 2023; Ghiani et al., 2023). Databases, particularly "big data" (i.e., high-volume data), are being increasingly used in research to improve the quality of medical care (e.g., Ullah et al., 2022; Wu et al., 2021; Yang et al., 2020). Therefore, skills related to working with databases are increasingly important.

Service User Involvement

In the phase 1 of the TONiC study patients had been invited to individual interviews and discussion groups about their experiences of neurological illness, views on quality of life and the impact of their illness on their quality of life. The most important factors in determining patients' quality of life were established and used to develop a set of questionnaires for participants and their relatives/carers during the TONiC study. This project involved a secondary data analysis, and it was not possible to involve service users in the process of analysis. However, I hope to involve PwMS in the dissemination of findings to UK MS charities.

Future Research Directions

To build upon the work presented in this project future research could explore the relationship between patients' objective cognitive testing and Rasched-analysed MSNQ-I scores. Comparing the results with the findings from previous research would increase the confidence in the conclusions which can be drawn. It would be also of interest to conduct Rasch analysis on the data from the translated versions of the MSNQ-I. This could confirm whether its psychometric properties are consistent across different cultures and languages. To compare the results with the current empirical study, future research should include samples of a similar size.

Whilst using the MSNQ-I-12 going forward offers interval level data, it makes comparison with previous publications difficult. Administering a shorter, 12-item questionnaire, reduces burden for relatives.

Reflections on the Project

Whilst my empirical project did not require data collection from participants, the secondary data analysis had its own challenges and learning points. The process has taught me valuable skills and gave me insights into the collaboration with highly respected international researchers. I consider myself lucky to have had a chance to collaborate with a large research team across the Liverpool and Royal Holloway universities as well as experience the multidisciplinary working with neurologists. It has also shown me the importance of secure data handling. I had to demonstrate this skill at several different stages and obey the terms of the non-disclosure agreement to protect data confidentiality throughout the process. I became aware of the General Data Protection Regulation law when managing data concerning health and how to document it in the Data Management Plan.

I have realised that understanding how to read the secondary data is a crucial first step to avoid erroneous analyses. I welcomed and enjoyed the opportunity, made possible thanks to my supervisor, to learn Rasch analysis. Having to learn something which was outside of the remit of the Clinical Psychology course required careful planning of the future steps of the project and, at times, multi-tasking. Whilst carrying out this research project, I have been receiving regular feedback from my supervisor on how to improve my scientific writing skills. This was difficult to hear at times but fundamental for improving the quality of my written work.

Having worked during my final year placement in a neuro-rehabilitation setting, I realised how debilitating cognitive impairment can be. In my clinical worked I observed that family members often do not understand what cognition means and how cognitive changes can affect day-to-day life. This thesis project as well my clinical experience influenced me to start a family support group in the hospital I work in. These sessions consist of a mixture of education about the brain as well as peer support. I am hopeful that family members will benefit from them and that these sessions will be continued when I finish my placement.

I would also like to acknowledge the political circumstances in which I was completing this project. The war outbreak in Ukraine in February 2022 has negatively impacted my engagement with this research. Particularly in the first few weeks of the war I was scared for my family who lives in a country that boarders with Ukraine. I discussed this in supervision and was offered adequate support.

Impact

This was an important and inclusive project because it had a sample from a population which is typically underrepresented in clinical research. MS patients, based on their health status

(e.g., having a cognitive impairment, multiple health conditions, physical disabilities), belong to an under-served group across the research landscape (NIHR, 2020).

To date all studies concerning the MSNQ-I, apart from one, were completed outside of the UK and may therefore not be reflective of the UK NHS context. This was also the first study utilising the MSNQ-I with a sample size of above 500. Therefore, the findings of this SR and empirical study have potential real-world implications which might impact PwMS as well as clinicians working with PwMS.

Patients and their families could benefit from an early diagnosis of PwMS' cognitive changes. The MSNQ-I-12 could facilitate a more reliable identification of those patients who require further cognitive assessments or cognitive interventions. Given that cognitive difficulties in MS affect people's overall QoL (Gil-González et al., 2020; Lakin et al., 2021), it is hoped that improving the management of cognition in MS would also improve their QoL.

There are also advantages of using the MSNQ-I-12 by healthcare professionals (HCPs) working with PwMS. They can use the nomogram to obtain an interval level scale by transforming a person's MSNQ-I-12 raw score. This will provide a more accurate representation of a person's true score. The MSNQ-I-12 may be therefore used in MS clinical services as a reliable screening tool for cognitive impairment and as proxy for objective cognitive assessment. This could enhance HCPs' assessments and aid their decisions regarding treatment or the support for PwMS.

This project can also impact the field of cognitive assessments/psychometric assessments. The results highlighted the benefit of using Rasch analysis to improve the psychometric properties of an existing scale. Clinicians with expertise in the areas of statistics and psychometrics could be encouraged to use Rasch analysis to refine existing scales or to develop

new questionnaires. The use of Rasch analysis may improve the reliability and validity of a range of assessment tools used in healthcare and beyond. This could lead to better-informed decisions in various fields (e.g., education, social sciences, and business).

This project would not have been possible if PwMS did not contribute to TONiC database. There are several databases collecting data from patients with various diseases. I am planning to get in touch with professional networks (e.g., https://mstrust.org.uk/health-professionals/therapists-ms-tims; https://www.mssociety.org.uk/care-and-support/experts-ms-professional-network; https://www.nationalmssociety.org/For-Professionals) with ideas on how to increase patients' awareness of databases which could facilitate recruitment.

Dissemination

To increase the impact of this research the findings will be disseminated to the research and clinical communities as well as those affected by MS.

Research

An abstract of the empirical study was submitted to the European Committee for Treatment and Research in Multiple Sclerosis conference, which will take place in Milan in October (https://www.ectrims.eu/). This is the biggest scientific MS conference in the world and should the abstract be accepted, the findings will be presented in the form of a poster.

Both the SR and empirical study will be submitted for peer review to academic journals. The SR will be submitted to the "Neurology and Therapy" journal (impact factor = 4.446). This journal has a broad scope, and it covers neurological and psychiatric therapies. Research related to PROMs and cognitive assessments in MS has been previously published in this journal. The findings of the empirical study will be written up as two independent studies reflecting the two parts of the project. The first article will outline the process and results of the Rasch analysis of the MSNQ-I. This part of the empirical study is planned to be submitted to the Multiple Sclerosis Journal (impact factor = 5.855). The second article will be submitted to the Journal of Neurology (impact factor = 6.682). It will utilise the findings of the first part (i.e., scores transformed through the Rasch analysis) and include the results of the linear multiple regression. Additionally, the nomogram (the table with a conversion of raw scores for the MSNQ-I items) will be made freely available on the TONiC website (https://tonic.thewaltoncentre.nhs.uk/) for

other researchers to use. I am also planning to present the findings in person to the TONiC research team and their collaborators based in Liverpool during the Liverpool Neuroscience Day in June 2023 (abstract submitted, https://lng.org.uk/event/lnd2023/).

The findings of this project have been disseminated among fellow Trainee Clinical Psychologists as well as course staff at Royal Holloway. The SR and empirical paper will also be shared on Royal Holloway's research information system (Pure). Pure enables research outputs to be made openly available for anyone to read and download. Copyright restrictions and publisher's requirements will be taken into account.

Clinical

If my submissions to the academic journals result in any successful publications, I hope that clinicians who work with PwMS will have access to these papers. Additionally, I plan to email a summary of the findings to the MS centres that took part in the TONiC study. To reach a bigger audience, I will also share the papers on social media platforms (i.e., Twitter and LinkedIn) where I am connected with other clinical health professionals for whom these findings may be of interest. I will also ask the administration team of Clinical Psychology programme at Royal Holloway to post it on their Twitter account to increase the visibility of these findings.

PwMS

I am planning to deliver a (possibly online) presentation with the main findings of this thesis to patients and their family members who contributed their data to the TONiC study. This will take place via the Liverpool centre where the majority of data was collected.

I also hope to liaise with relevant to MS charities (e.g., Multiple Sclerosis Trust) to be able to disseminate the findings on their websites and in magazines (e.g., MS Matters). I plan to involve PwMS in the dissemination process to co-produce the summary of findings with PwMS.

References

Akbar, N., & Finlayson, M. (2021). A mixed-methods study of cognitive performance in persons with multiple sclerosis: Association between neuropsychological test performance and interviews about daily cognitive functioning. *Multiple Sclerosis and Related Disorders*, *52*(February), 102911. https://doi.org/10.1016/j.msard.2021.102911

Akbar, N., Honarmand, K., & Feinstein, A. (2011). Self-assessment of cognition in Multiple Sclerosis: the role of personality and anxiety. *Cognitive and Behavioral Neurology*, *24*(3), 115–121. https://doi.org/10.1097/WNN.0b013e31822a20ae

Akbar, N., Honarmand, K., Kou, N., Levine, B., Rector, N., & Feinstein, A. (2010). Validity of an Internet version of the Multiple Sclerosis Neuropsychological Questionnaire. *Multiple Sclerosis Journal*, *16*(12), 1500–1506. https://doi.org/10.1177/1352458510379615

Altman, D. G., & Bland, J. M. (1995). Statistics notes: the normal distribution. *BMJ*, *310*(6975), 298. https://doi.org/10.1136/bmj.310.6975.298

Amin, M., & Hersh, C. M. (2023). Updates and advances in multiple sclerosis neurotherapeutics. *Neurodegenerative Disease Management*, *13*(1), 47–70. https://doi.org/10.2217/nmt-2021-0058

Amiri, Z., Azmin, M., Amiri, S., Akbarisari, A., Sahraian, M. A., Farzadfar, F., & Soleymani, F. (2023). Prevalence of comorbidities in patients with multiple sclerosis using administrative data from 2007 to 2016 in Iran. *Multiple Sclerosis and Related Disorders*, *74*, 104693. https://doi.org/https://doi.org/10.1016/j.msard.2023.104693

Andrich, D., & Hagquist, C. (2015). Real and Artificial Differential Item Functioning in

Polytomous Items. *Educational and Psychological Measurement*, *75*(2), 185–207. https://doi.org/10.1177/0013164414534258

Balestroni, G., & Bertolotti, G. (2012). [EuroQol-5D (EQ-5D): an instrument for measuring quality of life]. *Monaldi Archives for Chest Disease*, *78*(3), 155–159. https://doi.org/10.4081/monaldi.2012.121

Beesley, R., Anderson, V., Harding, K. E., Joseph, F., Tomassini, V., Pickersgill, T. P., Robertson, N. P., & Tallantyre, E. C. (2018). Impact of the 2017 revisions to McDonald criteria on the diagnosis of multiple sclerosis. *Multiple Sclerosis Journal*, *24*(13), 1786–1787. https://doi.org/10.1177/1352458518778007

Benedict, R. H. B., Amato, M. P., DeLuca, J., & Geurts, J. J. G. (2020). Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues. *The Lancet Neurology*, *19*(10), 860–871. https://doi.org/10.1016/S1474-4422(20)30277-5

Benedict, R. H. B., 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

Benedict, R. H. B., Drake, A. S., Irwin, L. N., Frndak, S. E., Kunker, K. A., Khan, A. L., Kordovski, V. M., Motl, R. W., & Weinstock-Guttman, B. (2016). Benchmarks of meaningful impairment on the MSFC and BICAMS. *Multiple Sclerosis Journal*, *22*(14), 1874–1882. https://doi.org/10.1177/1352458516633517

Benedict, R. H. B., Fischer, J. S., Archibald, C. J., Arnett, P. A., Beatty, W. W., Bobholz, J., Chelune, G. J., Fisk, J. D., Langdon, D. W., Caruso, L., Foley, F., LaRocca, N. G., Vowels,

L., Weinstein, A., DeLuca, J., Rao, S. M., & Munschauer, F. (2002). Minimal neuropsychological assessment of MS patients: A consensus approach. *Clinical Neuropsychologist, 16*(3), 381–397. https://doi.org/10.1076/clin.16.3.381.13859

Benedict, R. H. B., 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*, *9*(1), 95–101. https://doi.org/10.1191/1352458503ms861oa

Benedict, R. H. B., & Zivadinov, R. (2006). Predicting neuropsychological abnormalities in multiple sclerosis. *Journal of the Neurological Sciences*, *245*(1–2), 67–72. https://doi.org/10.1016/j.jns.2005.05.020

Berger, T., Adamczyk-Sowa, M., Csépány, T., Fazekas, F., Fabjan, T. H., Horáková, D., Ledinek, A. H., Illes, Z., Kobelt, G., Jazbec, S. Š., Klímová, E., Leutmezer, F., Rejdak, K., Rozsa, C., Sellner, J., Selmaj, K., Štourac̆, P., Szilasiová, J., Turc̆áni, P., … Havrdová, E. K. (2020). Factors influencing daily treatment choices in multiple sclerosis: practice guidelines, biomarkers and burden of disease. *Therapeutic Advances in Neurological Disorders*, *13*, 1756286420975223. https://doi.org/10.1177/1756286420975223

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How? *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-04-0148

Campbell, J., Rashid, W., Cercignani, M., & Langdon, D. (2017). Cognitive impairment among patients with multiple sclerosis: associations with employment and quality of life. *Postgraduate Medical Journal*, *93*(1097), 143–147. https://doi.org/10.1136/postgradmedj-2016-134071

Carone, D. A., Benedict, R. H. B., Munschauer, F. E. 3rd, Fishman, I., & Weinstock-Guttman, B. (2005). Interpreting patient/informant discrepancies of reported cognitive symptoms in MS. *Journal of the International Neuropsychological Society*, *11*(5), 574–583. https://doi.org/10.1017/S135561770505068X

Català-Senent, J. F., Andreu, Z., Hidalgo, M. R., Soler-Sáez, I., Roig, F. J., Yanguas-Casás, N., Neva-Alejo, A., López-Cerdán, A., de la Iglesia-Vayá, M., Stranger, B. E., & García-García, F. (2023). A deep transcriptome meta-analysis reveals sex differences in multiple sclerosis. *Neurobiology of Disease*, *181,* 106113. https://doi.org/10.1016/j.nbd.2023.106113

Charest, K., Tremblay, A., Langlois, R., Roger, É., Duquette, P., & Rouleau, I. (2020). Detecting Subtle Cognitive Impairment in Multiple Sclerosis with the Montreal Cognitive Assessment. *The Canadian Journal of Neurological Sciences*, *47*(5), 620–626. https://doi.org/10.1017/cjn.2020.97

Chiang, H.-S., Khera, A., Stopschinski, B. E., Stuve, O., Hart, J. J., Kelley, B., & Nguyen, T. (2022). Cognitive Decline in Older People with Multiple Sclerosis-A Narrative Review of the Literature. *Geriatrics*, *7*(3). https://doi.org/10.3390/geriatrics7030061

Christensen, K. B., Kreiner, S., & Mesbah, M. (2013). *Rasch Models in Health.* John Wiley & Sons, Inc. https://doi.org/10.1002/9781118574454

Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical Values for Yen's Q(3):

Identification of Local Dependence in the Rasch Model Using Residual Correlations. *Applied Psychological Measurement*, *41*(3), 178–194. https://doi.org/10.1177/0146621616677520

Claflin, S., Campbell, J. A., Norman, R., Mason, D. F., Kalincik, T., Simpson-Yap, S., Butzkueven, H., Carroll, W. M., Palmer, A. J., Blizzard, C. L., van der Mei, I., Henson, G. J., & Taylor, B. V. (2022). Using the EQ-5D-5L to investigate quality-of-life impacts of disease-modifying therapy policies for people with multiple sclerosis (MS) in New Zealand. *The European Journal of Health Economics : HEPAC : Health Economics in Prevention and Care*. https://doi.org/10.1007/s10198-022-01518-x

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York: Academic Press.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Crouch, T. A., Reas, H. E., Quach, C. M., & Erickson, T. M. (2022). Does depression in multiple sclerosis mediate effects of cognitive functioning on quality of life? *Quality of Life Research*, *31*(2), 497–506. https://doi.org/10.1007/s11136-021-02927-w

Curtin, F., & Schulz, P. (1998). Multiple correlations and bonferroni's correction. *Biological Psychiatry*, *44*(8), 775–777. https://doi.org/https://doi.org/10.1016/S0006-3223(98)00043-2

Dagenais, E., Rouleau, I., Demers, M., Jobin, C., Roger, E., Chamelian, L., & Duquette, P. (2013). Value of the MoCA test as a screening instrument in multiple sclerosis. *The Canadian Journal of Neurological Sciences*, *40*(3), 410–415. https://doi.org/10.1017/s0317167100014384

Davenport, L., Cogley, C., Monaghan, R., Gaughan, M., Yap, M., Bramham, J., Tubridy, N., McGuigan, C., & O'Keeffe, F. (2022). Investigating the association of mood and fatigue with objective and subjective cognitive impairment in multiple sclerosis. *Journal of Neuropsychology*, *16*(3), 537–554. https://doi.org/https://doi.org/10.1111/jnp.12283

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

Dobson, R., Craner, M., Waddingham, E., Miller, A., Cavey, A., Webb, S., Hemingway, C., Hobart, J., Evangelou, N., Scolding, N., Rog, D., Nicholas, R., Marta, M., Blain, C., Young, C. A., Ford, H. L., & Matthews, P. M. (2021). OPTIMISE: MS study protocol: a pragmatic, prospective observational study to address the need for, and challenges with, real world pharmacovigilance in multiple sclerosis. *BMJ Open*, *11*(11), e050176. https://doi.org/10.1136/bmjopen-2021-050176

Dobson, R., Rice, D. R., D'hooghe, M., Horne, R., Learmonth, Y., Mateen, F. J., Marck, C. H., Reyes, S., Williams, M. J., Giovannoni, G., & Ford, H. L. (2022). Social determinants of health in multiple sclerosis. *Nature Reviews. Neurology*, *18*(12), 723–734. https://doi.org/10.1038/s41582-022-00735-5

Donaldson, E., Patel, V. P., Shammi, P., & Feinstein, A. (2019). Why Sex Matters: A Cognitive Study of People With Multiple Sclerosis. *Cognitive and Behavioral Neurology*, *32*(1), 39–45. https://doi.org/10.1097/WNN.0000000000000188

Dulau, C., Deloire, M., Diaz, H., Saubusse, A., Charre-Morin, J., Prouteau, A., & Brochet, B. (2017). Social cognition according to cognitive impairment in different clinical phenotypes

of multiple sclerosis. *Journal of Neurology*, *264*(4), 740–748. https://doi.org/10.1007/s00415-017-8417-z

Ekmekyapar Fırat, Y., Akçalı, A., Geyik, S., Çomruk, G., Kılıçparlar Cengiz, E., & Erten, M. (2021). Relationship of Smoking with Fatigue and Depression in Patients with Multiple Sclerosis. *Turkish Journal of Neurology*, *27*(3), 289–294. https://doi.org/10.4274/tnd.2021.24119

Elwick, H., Topcu, G., Allen, C. M., Drummond, A., Evangelou, N., & Nair, R. das. (2021). Cognitive measures used in adults with multiple sclerosis: A systematic review. *Neuropsychological Rehabilitation*, 1–18. https://doi.org/10.1080/09602011.2021.1936080

Erlanger, D. M., Kaushik, T., Caruso, L. S., Benedict, R. H. B., Foley, F. W., Wilken, J., Cadavid, D., & DeLuca, J. (2014). Reliability of a cognitive endpoint for use in a multiple sclerosis pharmaceutical trial. *Journal of the Neurological Sciences*, *340*(1–2), 123–129. https://doi.org/10.1016/j.jns.2014.03.009

Faber, J., & Fonseca, L. M. (2014). How sample size influences research outcomes. *Dental Press Journal of Orthodontics*, *19*(4), 27–29. https://doi.org/10.1590/2176-9451.19.4.027-029.ebo

Feng, Y.-S., Kohlmann, T., Janssen, M. F., & Buchholz, I. (2021). Psychometric properties of the EQ-5D-5L: a systematic review of the literature. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *30*(3), 647–673. https://doi.org/10.1007/s11136-020-02688-y

Fenu, G., Fronza, M., Lorefice, L., Arru, M., Coghe, G., Frau, J., Marrosu, M. G., & Cocco, E. (2018). Performance in daily activities, cognitive impairment and perception in multiple sclerosis patients and their caregivers. *BMC Neurology*, *18*(1), 212.

https://doi.org/10.1186/s12883-018-1224-z

Fenu, G., Lorefice, L., Carta, E., Arru, M., Carta, A., Fronza, M., Coghe, G., Frau, J., Contu, F., Barracciu, M. A., & Cocco, E. (2021). Brain Volume and Perception of Cognitive Impairment in People With Multiple Sclerosis and Their Caregivers. *Frontiers in Neurology*, *12*, 636463. https://doi.org/10.3389/fneur.2021.636463

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Fisk, J. D., Brown, M. G., Sketris, I. S., Metz, L. M., Murray, T. J., & Stadnyk, K. J. (2005). A comparison of health utility measures for the evaluation of multiple sclerosis treatments. *Journal of Neurology, Neurosurgery, and Psychiatry*, *76*(1), 58–63. https://doi.org/10.1136/jnnp.2003.017897

Foley, F. W., Benedict, R. H. B., Gromisch, E. S., & Deluca, J. (2012). The Need for Screening, Assessment, and Treatment for Cognitive Dysfunction in Multiple Sclerosis: Results of a Multidisciplinary CMSC Consensus Conference, September 24, 2010. *International Journal of MS Care*, *14*(2), 58–64. https://doi.org/10.7224/1537-2073-14.2.58

Freedman, D. E., Oh, J., & Feinstein, A. (2023). Neuropsychiatric Status of Patients With Multiple Sclerosis Across Disease Duration Intervals. *The Journal of Neuropsychiatry and Clinical Neurosciences*, appineuropsych20220124. https://doi.org/10.1176/appi.neuropsych.20220124

Ghiani, M., Zhuleku, E., Dillenseger, A., Maywald, U., Fuchs, A., Wilke, T., & Ziemssen, T. (2023). Data Resource Profile: The Multiple Sclerosis Documentation System 3D and AOK PLUS Linked Database (MSDS-AOK PLUS). *Journal of Clinical Medicine*, *12*(4). https://doi.org/10.3390/jcm12041441

Gil-González, I., Martín-Rodríguez, A., Conrad, R., & Pérez-San-Gregorio, M. Á. (2020). Quality of life in adults with multiple sclerosis: A systematic review. *BMJ Open*, *10*(11). https://doi.org/10.1136/bmjopen-2020-041249

Golan, D., Doniger, G. M., Wissemann, K., Zarif, M., Bumstead, B., Buhse, M., Fafard, L., Lavi, I., Wilken, J., & Gudesblatt, M. (2018). The impact of subjective cognitive fatigue and depression on cognitive function in patients with multiple sclerosis. *Multiple Sclerosis*, *24*(2), 196–204. https://doi.org/10.1177/1352458517695470

Goverover, Y., Genova, H. M., Hillary, F. G., & DeLuca, J. (2007). The relationship between neuropsychological measures and the Timed Instrumental Activities of Daily Living task in multiple sclerosis. *Multiple Sclerosis Journal*, *13*(5), 636–644. https://doi.org/10.1177/1352458506072984

Granziera, C., Derfuss, T., & Kappos, L. (2023). Time to Change the Current Clinical Classification of Multiple Sclerosis? *JAMA Neurology*, *80*(2), 128–130. https://doi.org/10.1001/jamaneurol.2022.4156

Grzegorski, T., & Losy, J. (2017). Cognitive impairment in multiple sclerosis - a review of current knowledge and recent research. *Reviews in the Neurosciences*, *28*(8), 845–860. https://doi.org/10.1515/revneuro-2017-0011

Gustafsson, J.-E. (1980). Testing and obtaining fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology*, *33*, 205–233. https://doi.org/10.1111/j.2044-8317.1980.tb00609.x

Hackett, K., Mis, R., Drabick, D. A. G., & Giovannetti, T. (2020). Informant Reporting in Mild Cognitive Impairment: Sources of Discrepancy on the Functional Activities Questionnaire.

Journal of the International Neuropsychological Society, *26*(5), 503–514. https://doi.org/10.1017/S1355617719001449

Hadžibajramović, E., Schaufeli, W., & De Witte, H. (2022). Shortening of the Burnout Assessment Tool (BAT)-from 23 to 12 items using content and Rasch analysis. *BMC Public Health*, *22*(1), 560. https://doi.org/10.1186/s12889-022-12946-y

Hagell, P. (2014). Testing Rating Scale Unidimensionality Using the Principal Component Analysis (PCA)/t-Test Protocol with the Rasch Model: The Primacy of Theory over Statistics. *Open Journal of Statistics*, *4*, 456–465. https://doi.org/10.4236/ojs.2014.46044.

Hagell, P., & Westergren, A. (2016). Sample Size and Statistical Conclusions from Tests of Fit to the Rasch Model According to the Rasch Unidimensional Measurement Model (Rumm) Program in Health Outcome Measurement. *Journal of Applied Measurement*, *17*(4), 416–431.

Halstead, E. J., Stanley, J., Fiore, D., & Mueser, K. T. (2021). Impact of Cognitive Impairment on Adults with Multiple Sclerosis and Their Family Caregivers. *International Journal of MS Care*, *23*(3), 93–100. https://doi.org/10.7224/1537-2073.2019-091

Harwood, R., Rogers, A., Dickinson, E., & Ebrahim, S. (1994). Measuring handicap: the London Handicap Scale, a new outcome measure for chronic disease. *Quality in Health Care : QHC*, *3*, 11–16. https://doi.org/10.1136/qshc.3.1.11

Huang, F., Chen, W.-T., Shiu, C.-S., Lin, S. H., Tun, M. S., Nwe, T. W., Nu Oo, Y. T., & Oo, H. N. (2022). Adaptation and Validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R) in People Living with HIV in Myanmar. *Mindfulness*, *13*(1), 188–197. https://doi.org/10.1007/s12671-021-01784-5

Hughes, A. J., Bhattarai, J. J., Paul, S., & Beier, M. (2019). Depressive symptoms and fatigue as predictors of objective-subjective discrepancies in cognitive function in multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *30*, 192–197. https://doi.org/10.1016/j.msard.2019.01.055

Huisman, E., Papadimitropoulou, K., Jarrett, J., Bending, M., Firth, Z., Allen, F., & Adlard, N. (2017). Systematic literature review and network meta-analysis in highly active relapsing-remitting multiple sclerosis and rapidly evolving severe multiple sclerosis. *BMJ Open*, *7*(3), e013430. https://doi.org/10.1136/bmjopen-2016-013430

Jackson, D. A., Nicholson, R., Bergmann, C., Wilken, J., Kaczmarek, O., Bumstead, B., Buhse, M., Zarif, M., Penner, I.-K., Hancock, L. M., Golan, D., Doniger, G. M., Bogaardt, H., Barrera, M., Covey, T. J., & Gudesblatt, M. (2022). Cognitive impairment in people with multiple sclerosis: Perception vs. performance - factors that drive perception of impairment differ for patients and clinicians. *Multiple Sclerosis and Related Disorders*, *69*, 104410. https://doi.org/10.1016/j.msard.2022.104410

Jick, S. S., Li, L., Falcone, G. J., Vassilev, Z. P., & Wallander, M.-A. (2015). Epidemiology of multiple sclerosis: results from a large observational study in the UK. *Journal of Neurology*, *262*(9), 2033–2041. https://doi.org/10.1007/s00415-015-7796-2

Kazis, L. E., Anderson, J. J., & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical Care*, *27*(3 Suppl), S178-89. https://doi.org/10.1097/00005650-198903001-00015

Khamis, H. (2008). Measures of Association: How to Choose? *Journal of Diagnostic Medical Sonography*, *24*(3), 155–162. https://doi.org/10.1177/8756479308317006

Kline, R. K. (1998). *Principles and practice of structural equation modeling*. Guildford.

Kobelt, G., Thompson, A., Berg, J., Gannedahl, M., & Eriksson, J. (2017). New insights into the burden and costs of multiple sclerosis in Europe. *Multiple Sclerosis*, *23*(8), 1123–1136. https://doi.org/10.1177/1352458517694432

Konstantinopoulou, E., Ioannidis, P., Bakirtzis, C., Giantzi, V., Afrantou, T., Parissis, D., Messinis, L., Nasios, G., & Grigoriadis, N. (2018). Estimating Everyday Neuropsychological Functioning in Multiple Sclerosis: Reliability and Validity of the Greek Multiple Sclerosis Neuropsychological Questionnaire. *Multiple Sclerosis International*, *2018*, 6301535. https://doi.org/10.1155/2018/6301535

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*, *33*(11), 1444–1452. https://doi.org/10.1212/wnl.33.11.1444

Lakin, L., Davis, B. E., Binns, C. C., Currie, K. M., & Rensel, M. R. (2021). Comprehensive Approach to Management of Multiple Sclerosis: Addressing Invisible Symptoms—A Narrative Review. *Neurology and Therapy*, *10*(1), 75–98. https://doi.org/10.1007/s40120-021-00239-2

Langdon, D., Coutts, M., McGuigan, C., Morrow, B., Peryer, G., & Young, C. A. (2022). What would improve MS clinic services for cognition? – A stakeholder panel and survey exploration. *Multiple Sclerosis and Related Disorders*, *63*, 103930. https://doi.org/https://doi.org/10.1016/j.msard.2022.103930

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. H. B. (2012).

Recommendations for a brief international cognitive assessment for multiple sclerosis (BICAMS). *Multiple Sclerosis Journal*, *18*(6), 891–898. https://doi.org/10.1177/1352458511431076

Langdon, D. W., Benedict, R. H. B., Wicklein, E. M., Beckmann, K., & Fredrikson, S. (2013). Reports of patients and relatives from the CogniCIS study about cognition in clinically isolated syndrome: what are our patients telling us? *European Neurology*, *69*(6), 346–351. https://doi.org/10.1159/000345698

Langdon, D. W., & Thompson, A. J. (1999). Multiple sclerosis: a preliminary study of selected variables affecting rehabilitation outcome. *Multiple Sclerosis Journal*, *5*(2), 94–100. https://doi.org/10.1177/135245859900500205

Lynch, S. G., Parmenter, B. A., & Denney, D. R. (2005). The association between cognitive impairment and physical disability in multiple sclerosis. *Multiple Sclerosis Journal*, *11*(4), 469–476. https://doi.org/10.1191/1352458505ms1182oa

Magistrale, G., Medori, R., Cadavid, D., Argento, O., Incerti, C., Pisani, V., Caltagirone, C., Bozzali, M., DeLuca, J., & Nocentini, U. (2015). Assessing clinical correlates of self-rated disability in patients with multiple sclerosis. *Multiple Sclerosis Journal - Experimental, Translational and Clinical*, *1*, 2055217315592425. https://doi.org/10.1177/2055217315592425

Magistrale, G., Pisani, V., Argento, O., Incerti, C. C., Bozzali, M., Cadavid, D., Caltagirone, C., Medori, R., DeLuca, J., & Nocentini, U. (2015). Validation of the World Health Organization Disability Assessment Schedule II (WHODAS-II) in patients with multiple sclerosis. *Multiple Sclerosis*, *21*(4), 448–456. https://doi.org/10.1177/1352458514543732

Maltby, V. E., Lea, R. A., Reeves, P., Saugbjerg, B., & Lechner-Scott, J. (2022). Reduced cognitive function contributes to economic burden of multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *60*, 103707. https://doi.org/10.1016/j.msard.2022.103707

Marin, C. E., Kfouri, P. P., Callegaro, D., Lana-Peixoto, M. A., Neto, A. P. G., Vasconcelos, C. C. F., d'Almeida, J. A. C., Gonçalves, M. V. M., Mendes, M. F., Parolin, M. K. F., Nascimento, O., da Gama, P. D., Dias-Carneiro, R. P. C., Dias, R. M., Damasceno, A., San Martin, G., & Becker, J. (2021). Patients and neurologists have different perceptions of multiple sclerosis symptoms, care and challenges. *Multiple Sclerosis and Related Disorders*, *50*, 102806. https://doi.org/10.1016/j.msard.2021.102806

Masters, G. (1982). A rasch model for partial credit scoring. *The Psychometric Society*, *47*(2). https://doi.org/10.1007/BF02296272

Mazancieux, A., Souchay, C., Casez, O., & Moulin, C. J. A. (2019). Metacognition and self-awareness in Multiple Sclerosis. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *111*, 238–255. https://doi.org/10.1016/j.cortex.2018.11.012

McGinley, M. P., Goldschmidt, C. H., & Rae-Grant, A. D. (2021). Diagnosis and Treatment of Multiple Sclerosis: A Review. *JAMA*, *325*(8), 765–779. https://doi.org/10.1001/jama.2020.26858

Meca-Lallana, V., Gascón-Giménez, F., Ginestal-López, R. C., Higueras, Y., Téllez-Lara, N., Carreres-Polo, J., Eichau-Madueño, S., Romero-Imbroda, J., Vidal-Jordana, Á., & Pérez-Miralles, F. (2021). Cognitive impairment in multiple sclerosis: diagnosis and monitoring. *Neurological Sciences*, *42*(12), 5183–5193. https://doi.org/10.1007/s10072-021-05165-7

Meli, R., Roccatagliata, L., Capello, E., Bruschi, N., Uccelli, A., Mancardi, G., Inglese, M., &

Pardini, M. (2020). Ecological impact of isolated cognitive relapses in MS. *Multiple Sclerosis Journal*, *26*(1), 114–117. https://doi.org/10.1177/1352458518813722

Meyer-Moock, S., Feng, Y.-S., Maeurer, M., Dippel, F.-W., & Kohlmann, T. (2014). Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. *BMC Neurology*, *14*, 58. https://doi.org/10.1186/1471-2377-14-58

Migliore, S., Landi, D., Proietti, F., D'Aurizio, G., Squitieri, F., Mataluni, G., Nicoletti, C. G., Curcio, G., & Marfia, G. A. (2021). Validity of the Italian multiple sclerosis neuropsychological screening questionnaire. *Neurological Sciences*, *42*(11), 4583–4589. https://doi.org/10.1007/s10072-021-05141-1

Mills, R. J., Young, C. A., Pallant, J. F., & Tennant, A. (2010). Development of a patient reported outcome scale for fatigue in multiple sclerosis: The Neurological Fatigue Index (NFI-MS). *Health and Quality of Life Outcomes*, *8*, 22. https://doi.org/10.1186/1477-7525-8-22

Minden, S. L., Hoaglin, D. C., Hadden, L., Frankel, D., Robbins, T., & Perloff, J. (2008). Access to and utilization of neurologists by people with multiple sclerosis. *Neurology*, *70*(13 Pt 2), 1141–1149. https://doi.org/10.1212/01.wnl.0000306411.46934.ef

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

Morrell, L., Camic, P. M., & Genis, M. (2019). Factors associated with informant-reported cognitive decline in older adults: A systemised literature review. *Dementia*, *18*(7–8), 2760–

2784. https://doi.org/10.1177/1471301218759836

Morris, R. G., & Mograbi, D. C. (2013). Anosognosia, autobiographical memory and self knowledge in Alzheimer's disease. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(6), 1553–1565. https://doi.org/10.1016/j.cortex.2012.09.006

Morrow, S. A., Classen, S., Monahan, M., Danter, T., Taylor, R., Krasniuk, S., Rosehart, H., & He, W. (2018). On-road assessment of fitness-to-drive in persons with MS with cognitive impairment: A prospective study. *Multiple Sclerosis Journal*, *24*(11), 1499–1506. https://doi.org/10.1177/1352458517723991

Morrow, S. A., Jurgensen, S., Forrestal, F., Munchauer, F. E., & Benedict, R. H. B. (2011). Effects of acute relapses on neuropsychological status in multiple sclerosis patients. *Journal of Neurology*, *258*(9), 1603–1608. https://doi.org/10.1007/s00415-011-5975-3

Morrow, S. A., Rosehart, H., & Pantazopoulos, K. (2016). Anxiety and depressive symptoms are associated with worse performance on objective cognitive tests in MS. *Journal of Neuropsychiatry and Clinical Neurosciences*, *28*(2), 118–123. https://doi.org/10.1176/appi.neuropsych.15070167

Müller, M. (2020). Item fit statistics for Rasch analysis: can we trust them? *Journal of Statistical Distributions and Applications*, *7*(1), 5. https://doi.org/10.1186/s40488-020-00108-7

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nauta, I. M., Balk, L. J., Sonder, J. M., Hulst, H. E., Uitdehaag, B. M., Fasotti, L., & de Jong, B. A. (2019). The clinical value of the patient-reported multiple sclerosis neuropsychological screening questionnaire. *Multiple Sclerosis*, *25*(11), 1543–1546. https://doi.org/10.1177/1352458518777295

Neate, S. L., Taylor, K. L., Jelinek, G. A., De Livera, A. M., Simpson, S. J., Bevens, W., & Weiland, T. J. (2019). On the path together: Experiences of partners of people with multiple sclerosis of the impact of lifestyle modification on their relationship. *Health & Social Care in the Community*, *27*(6), 1515–1524. https://doi.org/10.1111/hsc.12822

Nicholas, R. S., Heaven, M. L., Middleton, R. M., Chevli, M., Pulikottil-Jacob, R., Jones, K. H., & Ford, D. V. (2020). Personal and societal costs of multiple sclerosis in the UK: A population-based MS Registry study. *Multiple Sclerosis Journal - Experimental, Translational and Clinical*, *6*(1), 2055217320901727. https://doi.org/10.1177/2055217320901727

NIHR. (2020). *Improving inclusion of under-served groups in clinical research: Guidance from the NIHR-INCLUDE project.* www.nihr.ac.uk/documents/improving-inclusion-of-under-served-groups-in-clinical-research-guidance-from-include-project/25435

O'Brien, A., Gaudino-Goering, E., Shawaryn, M., Komaroff, E., Moore, N. B., & DeLuca, J. (2007). Relationship of the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ) to functional, emotional, and neuropsychological outcomes. *Archives of Clinical Neuropsychology*, *22*(8), 933–948. https://doi.org/10.1016/j.acn.2007.07.002

Onuorah, H.-M., Charron, O., Meltzer, E., Montague, A., Crispino, A., Largent, A., Lucas, A., & Freeman, L. (2022). Enrollment of Non-White Participants and Reporting of Race and

Ethnicity in Phase III Trials of Multiple Sclerosis DMTs: A Systematic Review. *Neurology*, *98*(9), e880–e892. https://doi.org/10.1212/WNL.0000000000013230

Ozakbas, S., Turkoglu, R., Tamam, Y., Terzi, M., Taskapilioglu, O., Yucesan, C., Baser, H. L., Gencer, M., Akil, E., Sen, S., Turan, O. F., Sorgun, M. H., Yigit, P., & Turkes, N. (2018). Prevalence of and risk factors for cognitive impairment in patients with relapsing-remitting multiple sclerosis: Multi-center, controlled trial. *Multiple Sclerosis and Related Disorders*, *22*, 70–76. https://doi.org/10.1016/j.msard.2018.03.009

Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *The British Journal of Clinical Psychology*, *46*(Pt 1), 1–18. https://doi.org/10.1348/014466506x96931

Patel, V. P., & Feinstein, A. (2017). Comparison of Two Versions of the Hospital Anxiety and Depression Scale in Assessing Depression in a Neurologic Setting. *Cognitive and Behavioral Neurology*, *30*(4), 145–149. https://doi.org/10.1097/WNN.0000000000000138

Pellicciari, L., Piscitelli, D., Basagni, B., De Tanti, A., Algeri, L., Caselli, S., Ciurli, M. P., Conforti, J., Estraneo, A., Moretta, P., Gambini, M. G., Inzaghi, M. G., Lamberti, G., Mancuso, M., Rinaldesi, M. L., Sozzi, M., Abbruzzese, L., Zettin, M., & La Porta, F. (2020). "Less is more": validation with Rasch analysis of five short-forms for the Brain Injury Rehabilitation Trust Personality Questionnaires (BIRT-PQs). *Brain Injury*, *34*(13–14), 1741–1755. https://doi.org/10.1080/02699052.2020.1836402

Perone, S., Simmering, V. R., & Buss, A. T. (2021). A Dynamical Reconceptualization of Executive-Function Development. *Perspectives on Psychological Science*, *16*(6), 1198–1208. https://doi.org/10.1177/1745691620966792

Public Health England. (2020). *Multiple sclerosis: prevalence, incidence and smoking status - data briefing*. https://www.gov.uk/government/publications/multiple-sclerosis-prevalence-incidence-and-smoking-status/multiple-sclerosis-prevalence-incidence-and-smoking-status-data-briefing

Rao, S. M. (1991). *Neuropsychological screening battery for multiple sclerosis.* New York: Society NMS.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (2nd ed.). University of Chicago Press.

Reitzel, S. B., Lynning, M., & Skovgaard, L. (2022). Neurologists' views on patient reported outcomes in multiple sclerosis care. *Heliyon*, *8*(6), e09637. https://doi.org/10.1016/j.heliyon.2022.e09637

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

Rosti-Otajärvi, E., Ruutiainen, J., Huhtala, H., & Hämäläinen, P. (2014). Relationship between subjective and objective cognitive performance in multiple sclerosis. *Acta Neurologica Scandinavica*, *130*(5), 319–327. https://doi.org/10.1111/ane.12238

Ruano, L., Portaccio, E., Goretti, B., Niccolai, C., Severo, M., Patti, F., Cilia, S., Gallo, P., Grossi, P., Ghezzi, A., Roscio, M., Mattioli, F., Stampatori, C., Trojano, M., Viterbo, R. G., & Amato, M. P. (2017). Age and disability drive cognitive impairment in multiple sclerosis across disease subtypes. *Multiple Sclerosis*, *23*(9), 1258–1267. https://doi.org/10.1177/1352458516674367

Ruet, A., & Brochet, B. (2020). Cognitive assessment in patients with multiple sclerosis: From neuropsychological batteries to ecological tools. *Annals of Physical and Rehabilitation Medicine*, *63*(2), 154–158. https://doi.org/10.1016/j.rehab.2018.01.006

Saccà, F., Costabile, T., Carotenuto, A., Lanzillo, R., Moccia, M., Pane, C., Russo, C. V., Barbarulo, A. M., Casertano, S., Rossi, F., Signoriello, E., Lus, G., & Brescia Morra, V. (2017). The EDSS integration with the Brief International Cognitive Assessment for Multiple Sclerosis and orientation tests. *Multiple Sclerosis Journal*, *23*(9), 1289–1296. https://doi.org/10.1177/1352458516677592

Schmidt, N. E., & Steffen, A. M. (2020). Neurocognitive Disorder Diagnoses Matter: A Brief Report on Caregiver Appraisal of Driving Ability. *Journal of Applied Gerontology*, *39*(9), 966–970. https://doi.org/10.1177/0733464818803006

Sejbæk, T., Blaabjerg, M., Sprogøe, P., & Ravnborg, M. (2018). Reliability and Validity of a Danish Version of the Multiple Sclerosis Neuropsychological Screening Questionnaire. *International Journal of MS Care*, *20*(1), 49–54. https://doi.org/10.7224/1537-2073.2017-011

Smith Jr., E. V. (2002). Understanding Rasch measurement: Detecting and evaluating the impact of multidimenstionality using item fit statistics and principal component analysis of residuals. In *Journal of Applied Measurement* (Vol. 3, pp. 205–231). Richard M Smith.

Sonder, J. M., Holman, R., Knol, D. L., Bosma, L. V. A. E., Polman, C. H., & Uitdehaag, B. M. J. (2013). Analyzing differences between patient and proxy on Patient Reported Outcomes in multiple sclerosis. *Journal of the Neurological Sciences*, *334*(1–2), 143–147. https://doi.org/10.1016/j.jns.2013.08.014

Sonder, J. M., Mokkink, L. B., Van Der Linden, F. A. H., Polman, C. H., & Uitdehaag, B. M. J. (2012). Validation and interpretation of the Dutch version of the Multiple Sclerosis Neuropsychological Screening Questionnaire. *Journal of the Neurological Sciences*, *320*(1–2), 91–96. https://doi.org/10.1016/j.jns.2012.06.024

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. In *Journal of Personality Assessment* (Vol. 80, pp. 99–103). Lawrence Erlbaum. https://doi.org/10.1207/S15327752JPA8001_18

Strober, L. B., Binder, A., Nikelshpur, O. M., Chiaravalloti, N., & DeLuca, J. (2016). The Perceived Deficits Questionnaire: Perception, Deficit, or Distress? *International Journal of MS Care*, *18*(4), 183–190. https://doi.org/10.7224/1537-2073.2015-028

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. In *International Journal of Medical Education* (Vol. 2, pp. 53–55). https://doi.org/10.5116/ijme.4dfb.8dfd

Teresi, J. A., Kleinman, M., & Ocepek-Welikson, K. (2000). Modern psychometric methods for detection of differential item functioning: application to cognitive assessment measures. *Statistics in Medicine*, *19*(11–12), 1651–1683. https://doi.org/10.1002/(sici)1097-0258(20000615/30)19:11/12<1651::aid-sim453>3.0.co;2-h

Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing*, *1*(3), 176–184. https://doi.org/10.1111/j.1524-475X.2004.04006.x

Thomas, G. A., Riegler, K. E., Bradson, M. L., O'Shea, D. U., & Arnett, P. A. (2022a). Relationship between subjective report and objective assessment of neurocognitive

functioning in persons with multiple sclerosis. *Journal of the International Neuropsychological Society*, 1–8. https://doi.org/10.1017/S1355617722000212

Thomas, G. A., Riegler, K. E., Bradson, M. L., O'Shea, D. U., & Arnett, P. A. (2022b). Subjective Report, Objective Neurocognitive Performance, and "Invisible Symptoms" in Multiple Sclerosis. *Archives of Clinical Neuropsychology*. https://doi.org/10.1093/arclin/acac086

Thompson, A., & Feinstein, A. (2022). *Mind, Mood, and Memory: The Neurobehavioral Consequences of Multiple Sclerosis.* Johns Hopkins University Press. https://doi.org/doi:10.1353/book.100157

Tillery, E. E., Clements, J. N., & Howard, Z. (2017). What's new in multiple sclerosis? *The Mental Health Clinician*, *7*(5), 213–220. https://doi.org/10.9740/mhc.2017.09.213

Tracy, V. L., Basso, M. R., Marson, D. C., Combs, D. R., & Whiteside, D. M. (2017). Capacity for financial decision making in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *39*(1), 46–57. https://doi.org/10.1080/13803395.2016.1201050

Tur, C. (2016). Fatigue Management in Multiple Sclerosis. *Current Treatment Options in Neurology*, *18*(6), 26. https://doi.org/10.1007/s11940-016-0411-8

Ullah, S., Rahman, W., Ullah, F., Ahmad, G., Ijaz, M., & Gao, T. (2022). DBHR: a collection of databases relevant to human research. *Future Science OA*, *8*(3), FSO780. https://doi.org/10.2144/fsoa-2021-0101

Ustün, T. B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J., Saxena, S., von Korff, M., & Pull, C. (2010). Developing the World Health Organization Disability

Assessment Schedule 2.0. *Bulletin of the World Health Organization*, *88*(11), 815–823. https://doi.org/10.2471/BLT.09.067231

Van Laethem, D., De Cock, A., Van Schependom, J., Benedict, R. H. B., Nagels, G., & D'hooghe, M. (2022). Correlates of patient-reported cognitive performance with regard to disability. *Scientific Reports*, *12*(1), 13489. https://doi.org/10.1038/s41598-022-17649-3

van Wegen, J., van Egmond, E. E. A., Benedict, R. H. B., Beenakker, E. A. C., van Eijk, J. J. J., Frequin, S. T. F. M., de Gans, K., Gerlach, O. H. H., van Gorp, D. A. M., Hengstman, G. J. D., Jongen, P. J., van der Klink, J. J. L., Reneman, M. F., Verhagen, W. I. M., Middelkoop, H. A. M., Visser, L. H., Hulst, H. E., & van der Hiele, K. (2022). Subjective cognitive impairment is related to work status in people with multiple sclerosis. *IBRO Neuroscience Reports*, *13*, 513–522. https://doi.org/10.1016/j.ibneur.2022.10.016

Vanotti, S., Benedict, R. H. B., Acion, L., & Cáceres, F. (2009). Validation of the Multiple Sclerosis Neuropsychological Screening Questionnaire in Argentina. *Multiple Sclerosis Journal*, *15*(2), 244–250. https://doi.org/10.1177/1352458508097924

Vanotti, S., Smerbeck, A., Eizaguirre, M. B., Saladino, M. L., Benedict, R. R. H., & Caceres, F. J. (2018). BICAMS in the Argentine population: Relationship with clinical and sociodemographic variables. *Applied Neuropsychology: Adult*, *25*(5), 424–433. https://doi.org/10.1080/23279095.2017.1323751

Vasanthaprasad, V., Khurana, V., Vadapalle, S., Palace, J., & Adlard, N. (2022). Systematic literature review and meta-analysis of the prevalence of secondary progressive multiple sclerosis in the USA, Europe, Canada, Australia, and Brazil. *BMC Neurology*, *22*(1), 301. https://doi.org/10.1186/s12883-022-02820-0

Veillard, D., Baumstarck, K., Edan, G., Debouverie, M., Wiertlewski, S., De Sèze, J., Clavelou, P., Pelletier, J., Verny, C., Chauvin, K., Cosson, M. E., Loundou, A., & Auquier, P. (2021). Assessing the experience of the quality of care of patients living with multiple sclerosis and their caregivers: The MusiCare questionnaire. *European Journal of Neurology*, *28*(3), 910–920. https://doi.org/10.1111/ene.14685

Visser, L. A., Louapre, C., Uyl-de Groot, C. A., & Redekop, W. K. (2021). Health-related quality of life of multiple sclerosis patients: a European multi-country study. *Archives of Public Health*, *79*(1), 39. https://doi.org/10.1186/s13690-021-00561-z

Wallin, M. T., Culpepper, W. J., Nichols, E., Bhutta, Z. A., Gebrehiwot, T. T., Hay, S. I., Khalil, I. A., Krohn, K. J., Liang, X., Naghavi, M., Mokdad, A. H., Nixon, M. R., Reiner, R. C., Sartorius, B., Smith, M., Topor-Madry, R., Werdecker, A., Vos, T., Feigin, V. L., & Murray, C. J. L. (2019). Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, *18*(3), 269–285. https://doi.org/10.1016/S1474-4422(18)30443-5

Washington, F., & Langdon, D. (2022). Factors affecting adherence to disease-modifying therapies in multiple sclerosis: systematic review. *Journal of Neurology*, *269*(4), 1861–1872. https://doi.org/10.1007/s00415-021-10850-w

Westergaard, K., Skovgaard, L., Magyari, M., & Kristiansen, M. (2022). Patient perspectives on patient-reported outcomes in multiple sclerosis treatment trajectories: A qualitative study of why, what, and how? *Multiple Sclerosis and Related Disorders*, *58*, 103475. https://doi.org/10.1016/j.msard.2021.103475

World Health Organization. (2001). *International classification of functioning, disability and*

health (ICF). World Health Organization. https://apps.who.int/iris/handle/10665/42407

Wu, W.-T., Li, Y.-J., Feng, A.-Z., Li, L., Huang, T., Xu, A.-D., & Lyu, J. (2021). Data mining in clinical big data: the frequently used databases, steps, and methodological models. *Military Medical Research*, *8*(1), 44. https://doi.org/10.1186/s40779-021-00338-z

Wu, Y., Levis, B., Sun, Y., He, C., Krishnan, A., Neupane, D., Bhandari, P. M., Negeri, Z., Benedetti, A., & Thombs, B. D. (2021). Accuracy of the Hospital Anxiety and Depression Scale Depression subscale (HADS-D) to screen for major depression: systematic review and individual participant data meta-analysis. *BMJ*, *373*, n972. https://doi.org/10.1136/bmj.n972

Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T., Zheng, S., Xu, A., & Lyu, J. (2020). Brief introduction of medical database and data mining technology in big data era. *Journal of Evidence-Based Medicine*, *13*(1), 57–69. https://doi.org/10.1111/jebm.12373

Young, C. A., Mills, R., Rog, D., Sharrack, B., Majeed, T., Constantinescu, C. S., Kalra, S., Harrower, T., Santander, H., Courtald, G., Ford, H. L., Woolmore, J., & Tennant, A. (2021). Quality of life in multiple sclerosis is dominated by fatigue, disability and self-efficacy. *Journal of the Neurological Sciences*, *426*, 117437. https://doi.org/https://doi.org/10.1016/j.jns.2021.117437

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Appendices

Appendix I:

Ethics Approval

Result of your application to the Research Ethics Committee (application ID 3153)

You forwarded this message on Tue 29/03/2022 21:18

ES Ethics Application System <ethics@rhul.ac.uk>
To: Marek, Katarzyna (2020); Langdon, D; Ethics
Mon 28/03/2022 11:40

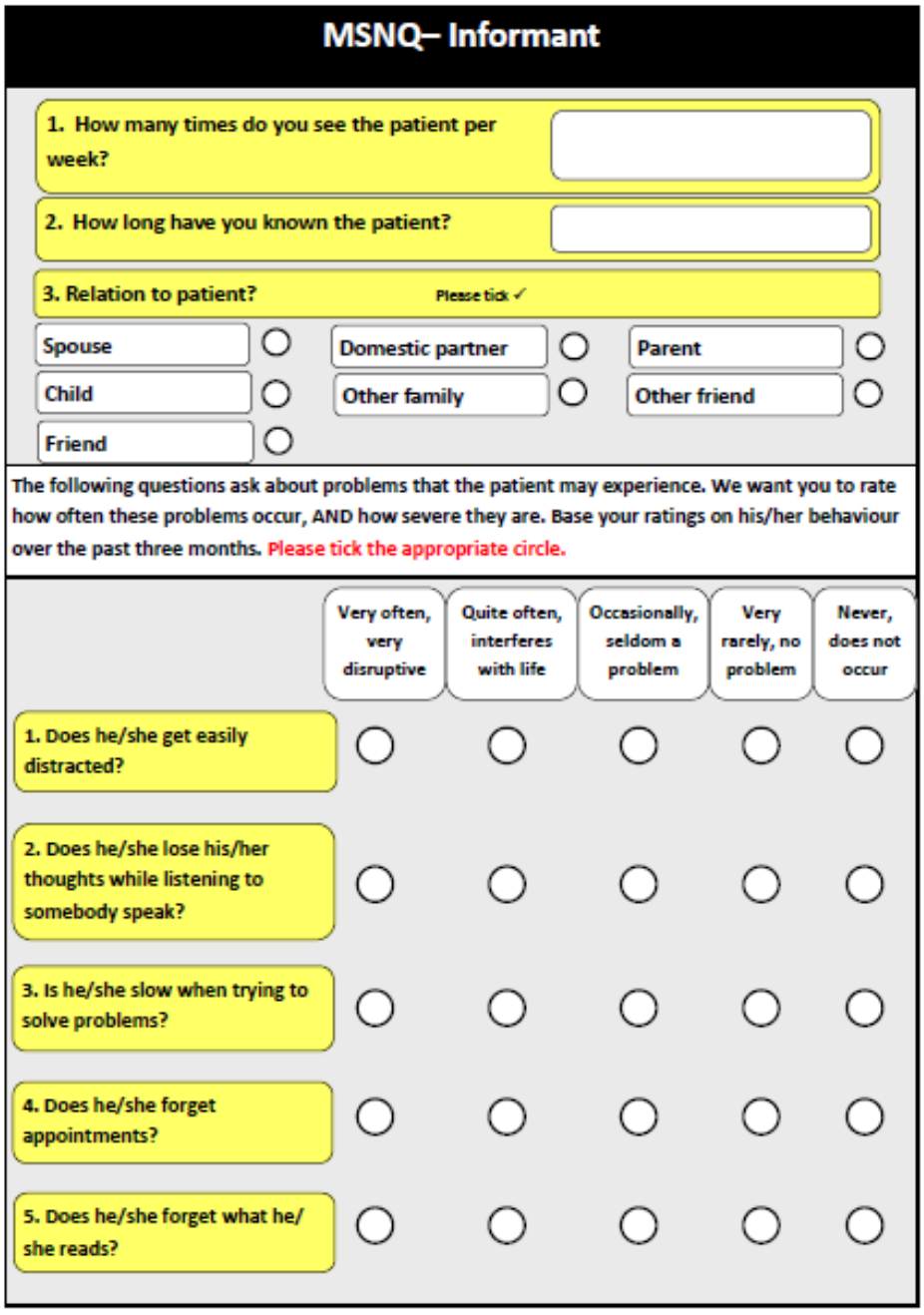
PI: Prof Dawn Langdon
Project title: What influences relatives' perceptions of the cognition of people with Multiple Sclerosis (PwMS)?

REC ProjectID: 3153

Your application has been approved by the Research Ethics Committee.
Please report any subsequent changes that affect the ethics of the project to the University Research Ethics Committee ethics@rhul.ac.uk

Appendix II:

The Multiple Sclerosis Neuropsychological Questionnaire-Informant (MSNQ-I)



MSNQ– Informant

1. How many times do you see the patient per week?

2. How long have you known the patient?

3. Relation to patient? Please tick ✓

Spouse ○	Domestic partner ○	Parent ○
Child ○	Other family ○	Other friend ○
Friend ○		

The following questions ask about problems that the patient may experience. We want you to rate how often these problems occur, AND how severe they are. Base your ratings on his/her behaviour over the past three months. Please tick the appropriate circle.

	Very often, very disruptive	Quite often, interferes with life	Occasionally, seldom a problem	Very rarely, no problem	Never, does not occur
1. Does he/she get easily distracted?	○	○	○	○	○
2. Does he/she lose his/her thoughts while listening to somebody speak?	○	○	○	○	○
3. Is he/she slow when trying to solve problems?	○	○	○	○	○
4. Does he/she forget appointments?	○	○	○	○	○
5. Does he/she forget what he/she reads?	○	○	○	○	○


Very often, very disruptive
Quite often, interferes with life
Occasionally, seldom a problem
Very rarely, no problem
Never, does not occur
6. Does he/she have trouble describing shows or programs recently watched?
7.Does he/she need to have instructions repeated?
8. Does he/she have to be reminded to do tasks?
9. Does he/she forget errands that were planned?
10. Does he/she have difficulty answering questions?
11. Does he/she have difficulty keeping track of two things at once?
12. Does he/she miss the point of what someone is trying to say?
13. Does he/she have difficulty controlling his/her impulses?
14. Does he/she laugh or cry with little cause?
15. Does he/she talk excessively or focus too much on his/her own interests?

Appendix III:

The World Health Organization Disability Assessment Schedule 2.0

(WHODAS 2.0)

WHODAS 2.0

World Health Organization Disability Assessment Schedule 2.0

36-item version, self-administered

Patient Name: _____ **Age:** _____ **Sex:** ❑ Male ❑ Female **Date:** _____

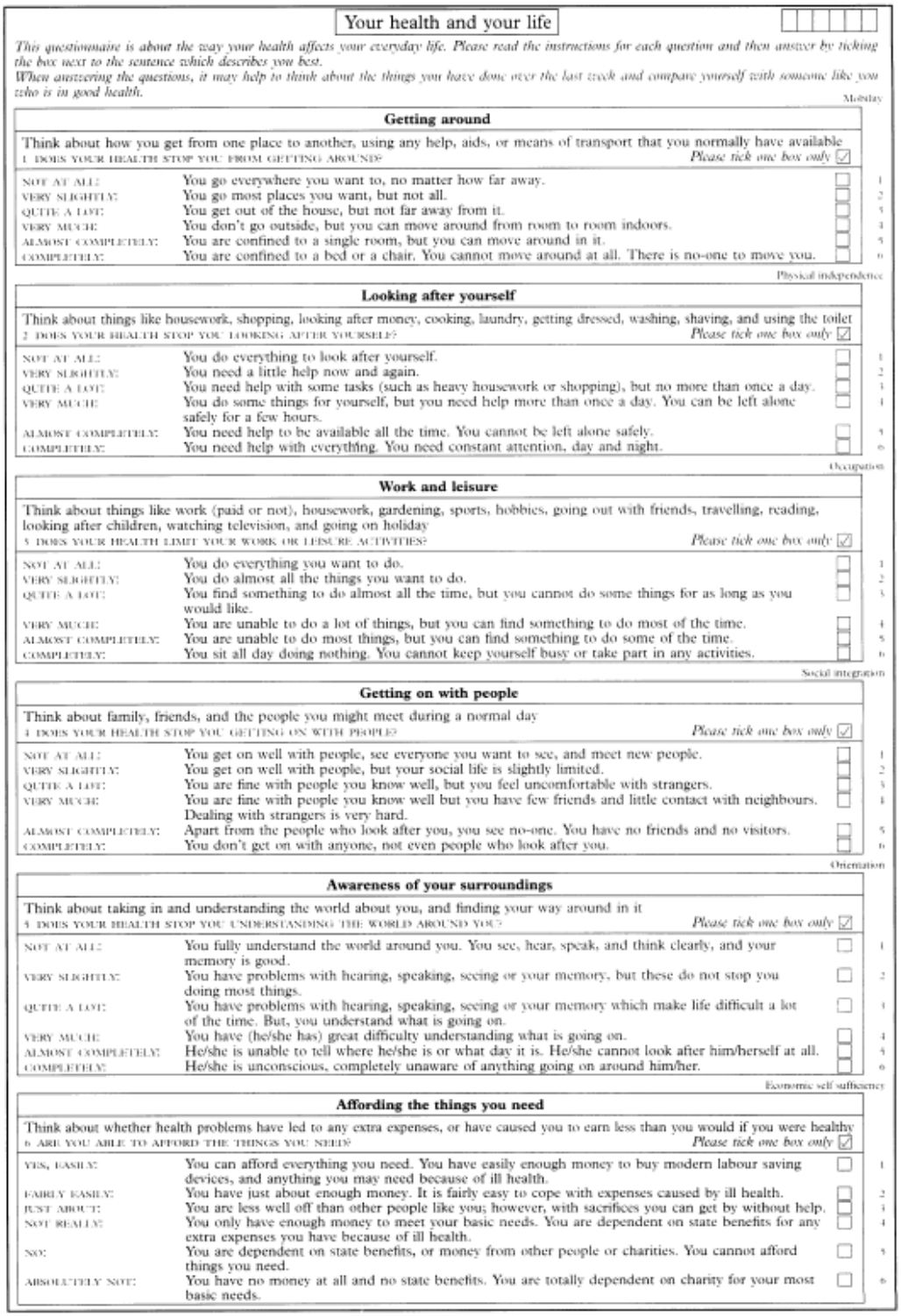
This questionnaire asks about difficulties due to health/mental health conditions. Health conditions include **diseases or illnesses, other health problems that may be short or long lasting, injuries, mental or emotional problems, and problems with alcohol or drugs.** Think back over the **past 30 days** and answer these questions thinking about how much difficulty you had doing the following activities. For each question, please circle only **one** response.

							Clinician Use Only		
Numeric scores assigned to each of the items:		1	2	3	4	5	Raw Item Score	Raw Domain Score	Average Domain Score
In the last 30 days, how much difficulty did you have in:									
Understanding and communicating									
D1.1	Concentrating on doing something for ten minutes?	None	Mild	Moderate	Severe	Extreme or cannot do			
D1.2	Remembering to do important things?	None	Mild	Moderate	Severe	Extreme or cannot do			
D1.3	Analyzing and finding solutions to problems in day-to-day life?	None	Mild	Moderate	Severe	Extreme or cannot do			
D1.4	Learning a new task, for example, learning how to get to a new place?	None	Mild	Moderate	Severe	Extreme or cannot do		___ 30	___ 5
D1.5	Generally understanding what people say?	None	Mild	Moderate	Severe	Extreme or cannot do			
D1.6	Starting and maintaining a conversation?	None	Mild	Moderate	Severe	Extreme or cannot do			
Getting around									
D2.1	Standing for long periods, such as 30 minutes?	None	Mild	Moderate	Severe	Extreme or cannot do			
D2.2	Standing up from sitting down?	None	Mild	Moderate	Severe	Extreme or cannot do			
D2.3	Moving around inside your home?	None	Mild	Moderate	Severe	Extreme or cannot do			
D2.4	Getting out of your home?	None	Mild	Moderate	Severe	Extreme or cannot do		___ 25	___ 5
D2.5	Walking a long distance, such as a kilometer (or equivalent)?	None	Mild	Moderate	Severe	Extreme or cannot do			
Self-care									
D3.1	Washing your whole body?	None	Mild	Moderate	Severe	Extreme or cannot do			
D3.2	Getting dressed?	None	Mild	Moderate	Severe	Extreme or cannot do			
D3.3	Eating?	None	Mild	Moderate	Severe	Extreme or cannot do		___ 20	___ 5
D3.4	Staying by yourself for a few days?	None	Mild	Moderate	Severe	Extreme or cannot do			
Getting along with people									
D4.1	Dealing with people you do not know?	None	Mild	Moderate	Severe	Extreme or cannot do			
D4.2	Maintaining a friendship?	None	Mild	Moderate	Severe	Extreme or cannot do			
D4.3	Getting along with people who are close to you?	None	Mild	Moderate	Severe	Extreme or cannot do			
D4.4	Making new friends?	None	Mild	Moderate	Severe	Extreme or cannot do		___ 25	___ 5
D4.5	Sexual activities?	None	Mild	Moderate	Severe	Extreme or cannot do			

							Clinician Use Only		
Numeric scores assigned to each of the items:		1	2	3	4	5	Raw Item Score	Raw Domain Score	Average Domain Score
In the last 30 days, how much difficulty did you have in:									
Life activities—Household									
D5.1	Taking care of your household responsibilities?	None	Mild	Moderate	Severe	Extreme or cannot do			
D5.2	Doing most important household tasks well?	None	Mild	Moderate	Severe	Extreme or cannot do			
D5.3	Getting all of the household work done that you needed to do?	None	Mild	Moderate	Severe	Extreme or cannot do		___ 20	___ 5
D5.4	Getting your household work done as quickly as needed?	None	Mild	Moderate	Severe	Extreme or cannot do			
Life activities—School/Work									
If you work (paid, non-paid, self-employed) or go to school, complete questions D5.5–D5.8, below. Otherwise, skip to D6.1.									
Because of your health condition, in the past 30 days, how much difficulty did you have in:									
D5.5	Your day-to-day work/school?	None	Mild	Moderate	Severe	Extreme or cannot do			
D5.6	Doing your most important work/school tasks well?	None	Mild	Moderate	Severe	Extreme or cannot do			
D5.7	Getting all of the work done that you need to do?	None	Mild	Moderate	Severe	Extreme or cannot do		___ 20	___ 5
D5.8	Getting your work done as quickly as needed?	None	Mild	Moderate	Severe	Extreme or cannot do			
Participation in society									
In the past 30 days:									
D6.1	How much of a problem did you have in joining in community activities (for example, festivities, religious, or other activities) in the same way as anyone else can?	None	Mild	Moderate	Severe	Extreme or cannot do			
D6.2	How much of a problem did you have because of barriers or hindrances around you?	None	Mild	Moderate	Severe	Extreme or cannot do			
D6.3	How much of a problem did you have living with dignity because of the attitudes and actions of others?	None	Mild	Moderate	Severe	Extreme or cannot do			
D6.4	How much time did you spend on your health condition or its consequences?	None	Some	Moderate	A Lot	Extreme or cannot do		___ 40	___ 5
D6.5	How much have you been emotionally affected by your health condition?	None	Mild	Moderate	Severe	Extreme or cannot do			
D6.6	How much has your health been a drain on the financial resources of you or your family?	None	Mild	Moderate	Severe	Extreme or cannot do			
D6.7	How much of a problem did your family have because of your health problems?	None	Mild	Moderate	Severe	Extreme or cannot do			
D6.8	How much of a problem did you have in doing things by yourself for relaxation or pleasure?	None	Mild	Moderate	Severe	Extreme or cannot do			
						General Disability Score (Total):		___ 180	___ 5

Appendix IV:

The London Handicap Scale (LHS)



Your health and your life

This questionnaire is about the way your health affects your everyday life. Please read the instructions for each question and then answer by ticking the box next to the sentence which describes you best.
When answering the questions, it may help to think about the things you have done over the last week and compare yourself with someone like you who is in good health.

Mobility

Getting around

Think about how you get from one place to another, using any help, aids, or means of transport that you normally have available
1 DOES YOUR HEALTH STOP YOU FROM GETTING AROUND? *Please tick one box only* ☑

NOT AT ALL:	You go everywhere you want to, no matter how far away.	1
VERY SLIGHTLY:	You go most places you want, but not all.	2
QUITE A LOT:	You get out of the house, but not far away from it.	3
VERY MUCH:	You don't go outside, but you can move around from room to room indoors.	4
ALMOST COMPLETELY:	You are confined to a single room, but you can move around in it.	5
COMPLETELY:	You are confined to a bed or a chair. You cannot move around at all. There is no-one to move you.	6

Physical independence

Looking after yourself

Think about things like housework, shopping, looking after money, cooking, laundry, getting dressed, washing, shaving, and using the toilet
2 DOES YOUR HEALTH STOP YOU LOOKING AFTER YOURSELF? *Please tick one box only* ☑

NOT AT ALL:	You do everything to look after yourself.	1
VERY SLIGHTLY:	You need a little help now and again.	2
QUITE A LOT:	You need help with some tasks (such as heavy housework or shopping), but no more than once a day.	3
VERY MUCH:	You do some things for yourself, but you need help more than once a day. You can be left alone safely for a few hours.	4
ALMOST COMPLETELY:	You need help to be available all the time. You cannot be left alone safely.	5
COMPLETELY:	You need help with everything. You need constant attention, day and night.	6

Occupation

Work and leisure

Think about things like work (paid or not), housework, gardening, sports, hobbies, going out with friends, travelling, reading, looking after children, watching television, and going on holiday
3 DOES YOUR HEALTH LIMIT YOUR WORK OR LEISURE ACTIVITIES? *Please tick one box only* ☑

NOT AT ALL:	You do everything you want to do.	1
VERY SLIGHTLY:	You do almost all the things you want to do.	2
QUITE A LOT:	You find something to do almost all the time, but you cannot do some things for as long as you would like.	3
VERY MUCH:	You are unable to do a lot of things, but you can find something to do most of the time.	4
ALMOST COMPLETELY:	You are unable to do most things, but you can find something to do some of the time.	5
COMPLETELY:	You sit all day doing nothing. You cannot keep yourself busy or take part in any activities.	6

Social integration

Getting on with people

Think about family, friends, and the people you might meet during a normal day
4 DOES YOUR HEALTH STOP YOU GETTING ON WITH PEOPLE? *Please tick one box only* ☑

NOT AT ALL:	You get on well with people, see everyone you want to see, and meet new people.	1
VERY SLIGHTLY:	You get on well with people, but your social life is slightly limited.	2
QUITE A LOT:	You are fine with people you know well, but you feel uncomfortable with strangers.	3
VERY MUCH:	You are fine with people you know well but you have few friends and little contact with neighbours. Dealing with strangers is very hard.	4
ALMOST COMPLETELY:	Apart from the people who look after you, you see no-one. You have no friends and no visitors.	5
COMPLETELY:	You don't get on with anyone, not even people who look after you.	6

Orientation

Awareness of your surroundings

Think about taking in and understanding the world about you, and finding your way around in it
5 DOES YOUR HEALTH STOP YOU UNDERSTANDING THE WORLD AROUND YOU? *Please tick one box only* ☑

NOT AT ALL:	You fully understand the world around you. You see, hear, speak, and think clearly, and your memory is good.	1
VERY SLIGHTLY:	You have problems with hearing, speaking, seeing or your memory, but these do not stop you doing most things.	2
QUITE A LOT:	You have problems with hearing, speaking, seeing or your memory which make life difficult a lot of the time. But, you understand what is going on.	3
VERY MUCH:	You have (he/she has) great difficulty understanding what is going on.	4
ALMOST COMPLETELY:	He/she is unable to tell where he/she is or what day it is. He/she cannot look after him/herself at all.	5
COMPLETELY:	He/she is unconscious, completely unaware of anything going on around him/her.	6

Economic self sufficiency

Affording the things you need

Think about whether health problems have led to any extra expenses, or have caused you to earn less than you would if you were healthy
6 ARE YOU ABLE TO AFFORD THE THINGS YOU NEED? *Please tick one box only* ☑

YES, EASILY:	You can afford everything you need. You have easily enough money to buy modern labour saving devices, and anything you may need because of ill health.	1
FAIRLY EASILY:	You have just about enough money. It is fairly easy to cope with expenses caused by ill health.	2
JUST ABOUT:	You are less well off than other people like you; however, with sacrifices you can get by without help.	3
NOT REALLY:	You only have enough money to meet your basic needs. You are dependent on state benefits for any extra expenses you have because of ill health.	4
NO:	You are dependent on state benefits, or money from other people or charities. You cannot afford things you need.	5
ABSOLUTELY NOT:	You have no money at all and no state benefits. You are totally dependent on charity for your most basic needs.	6

The London handicap scale questionnaire

Appendix V:

The Hospital Anxiety and Depression Scale (HADS)

Hospital Anxiety and Depression Scale (HADS)

Tick the box beside the reply that is closest to how you have been feeling in the past week.
Don't take too long over you replies: your immediate is best.

D	A		D	A	
		I feel tense or 'wound up':			**I feel as if I am slowed down:**
	3	Most of the time	3		Nearly all the time
	2	A lot of the time	2		Very often
	1	From time to time, occasionally	1		Sometimes
	0	Not at all	0		Not at all
		I still enjoy the things I used to enjoy:			**I get a sort of frightened feeling like 'butterflies' in the stomach:**
0		Definitely as much		0	Not at all
1		Not quite so much		1	Occasionally
2		Only a little		2	Quite Often
3		Hardly at all		3	Very Often
		I get a sort of frightened feeling as if something awful is about to happen:			**I have lost interest in my appearance:**
	3	Very definitely and quite badly	3		Definitely
	2	Yes, but not too badly	2		I don't take as much care as I should
	1	A little, but it doesn't worry me	1		I may not take quite as much care
	0	Not at all	0		I take just as much care as ever
		I can laugh and see the funny side of things:			**I feel restless as I have to be on the move:**
0		As much as I always could		3	Very much indeed
1		Not quite so much now		2	Quite a lot
2		Definitely not so much now		1	Not very much
3		Not at all		0	Not at all
		Worrying thoughts go through my mind:			**I look forward with enjoyment to things:**
	3	A great deal of the time	0		As much as I ever did
	2	A lot of the time	1		Rather less than I used to
	1	From time to time, but not too often	2		Definitely less than I used to
	0	Only occasionally	3		Hardly at all
		I feel cheerful:			**I get sudden feelings of panic:**
3		Not at all		3	Very often indeed
2		Not often		2	Quite often
1		Sometimes		1	Not very often
0		Most of the time		0	Not at all
		I can sit at ease and feel relaxed:			**I can enjoy a good book or radio or TV program:**
	0	Definitely	0		Often
	1	Usually	1		Sometimes
	2	Not Often	2		Not often
	3	Not at all	3		Very seldom

Appendix VI:

The Expanded Disability Status Scale (EDSS)

0	Normal neurologic exam
1.0	No disability, minimal signs in one functional system
1.5	No disability, minimal signs in more than one functional system
2.0	Minimal disability in one functional system
2.5	Minimal disability in two functional systems
3.0	Moderate disability in one functional system, or mild disability in three or four functional systems though fully ambulatory
3.5	Fully ambulatory but with moderate disability in three or four functional systems
4.0	Fully ambulatory without aid, self-sufficient, up and about some 12 hours a day despite relatively severe disability. Able to walk without aid or rest some 500 meters
4.5	Fully ambulatory without aid, up and about much of the day, able to work a full day, may otherwise have some limitation of full activity or require minimal assistance, characterized by relatively severe disability. Able to walk without aid or rest for some 300 meters
5.0	Ambulatory without aid or rest for about 200 meters; disability severe enough to preclude full daily activities (e.g. to work full day without special provisions)
5.5	Ambulatory without aid or rest for about 100 meters; disability severe enough to preclude full daily activities
6.0	Intermittent or unilateral constant assistance (cane, crutch, or brace) required to walk about 100 meters with or without resting
6.5	Constant bilateral assistance (canes, crutches, or braces) required to walk about 20 meters without resting
7.0	Unable to walk beyond about 5 meters even with aid. Essentially restricted to a wheelchair. Wheels self in standard wheelchair and transfers alone. Active in wheelchair about 12 hours a day
7.5	Unable to take more than a few steps. Restricted to wheelchair. May need aid to transfer. Wheels self but cannot carry on in standard wheelchair a full day. May require a motorized wheelchair
8.0	Unable to walk at all, essentially restricted to bed, chair or wheelchair but may be out of bed much of the day. Retains many self-care functions. Generally has effective use of the arms
8.5	Essentially restricted to bed much of the day. Has some effective use of arm(s). Retains some self-care functions
9.0	Helpless bed patient. Can communicate and eat
9.5	Totally helpless bed patient. Unable to communicate effectively or eat/ swallow
10	Death due to Multiple Sclerosis

SOURCE: Kurtzke JF. 1983. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*;33:1444-52.

Appendix VII:

The EuroQol five-dimension (EQ-5D)

Under each heading, please check the ONE box that best describes your health TODAY.

MOBILITY

- I have no problems walking ❑
- I have slight problems walking ❑
- I have moderate problems walking ❑
- I have severe problems walking ❑
- I am unable to walk ❑

SELF-CARE

- I have no problems washing or dressing myself ❑
- I have slight problems washing or dressing myself ❑
- I have moderate problems washing or dressing myself ❑
- I have severe problems washing or dressing myself ❑
- I am unable to wash or dress myself ❑

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

- I have no problems doing my usual activities ❑
- I have slight problems doing my usual activities ❑
- I have moderate problems doing my usual activities ❑
- I have severe problems doing my usual activities ❑
- I am unable to do my usual activities ❑

PAIN / DISCOMFORT

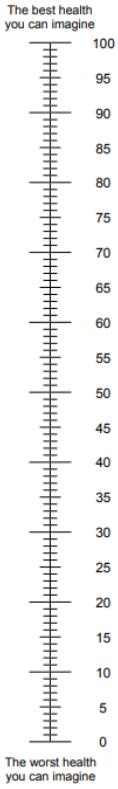
- I have no pain or discomfort ❑
- I have slight pain or discomfort ❑
- I have moderate pain or discomfort ❑
- I have severe pain or discomfort ❑
- I have extreme pain or discomfort ❑

ANXIETY / DEPRESSION

- I am not anxious or depressed ❑
- I am slightly anxious or depressed ❑
- I am moderately anxious or depressed ❑
- I am severely anxious or depressed ❑
- I am extremely anxious or depressed ❑

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



Appendix VIII:

Table of Number of Reponses for Each Variable – Whole Sample

Variable Name	*N*
Age	2,039
Gender	2,039
MS type	2,030
Disease duration	2,011
EDSS band	2,030
EQ-5D	1,995
WHODAS32	1,704
LHS	1,983
NFI-MS	1,956
HADS D	1,997
HADS A	1,983
MSNQ-I-12	2,039

Notes. EDSS= Expanded Disability Status Scale, EQ-5D= Health Status Scale, HADS A= Hospital Anxiety and Depression Scale-Anxiety, HADS D= Hospital Anxiety and Depression Scale-Depression, LHS= London Hospital Scale, MS= Multiple Sclerosis, MSNQ-I-12= The Multiple Sclerosis Neuropsychological Questionnaire-Informant (Rasch-analysed), *N*= Number, NFI-MS= The Neurological Fatigue Index, WHODAS 32= World Health Organisation Disability Assessment Schedule omitting work-related items

Appendix IX:

Table of Multiple Regression MSNQ-I-12 Sample Demographics by EDSS Band

	Regression sample	EDSS 0 – 4	EDSS 4.5 – 6.5	EDSS 7 – 7.5	EDSS 8 – 9.5
					
n	1,806	799	745	146	116
%	100	44.2	41.3	8.1	6.4
Mean age (SD)	51.0 (12.0)	45.6 (11.4)	54.3 (10.6)	58.3 (10.5)	57.9 (9.6)
Female %	73.5	76.2	71.1	76.7	65.5
Married %	80.8	80.0	80.3	84.2	85.3
Working full-time %	11.7	23.0	3.5	0.7	0.0

Note. EDSS= Expanded Disability Status Scale, *n*= Number, SD= Standard Deviation

Appendix X:

Table of Disease Variables by EDSS Band for the Multiple Regression MSNQ-I-12 Sample

	n	Regression sample	EDSS 0 – 4	EDSS 4.5 – 6.5	EDSS 7 – 7.5	EDSS 8 – 9.5
						
Phenotype (%)						
PP	243	13.5	6.1	19.1	21.9	17.2
RE	62	3.4	4.1	3.5	2.1	0.0
RR	1,007	55.8	85.0	41.9	9.6	1.7
SP	494	27.4	4.8	35.6	66.4	81.0
Mean duration (SD)	1,790	11.3 (9.8)	7.6 (7.7)	12.6 (9.5)	17.9 (11.7)	20.9 (10.0)
Taking DMT (%)	1,805	40.0	52.6	37.3	13.0	5.2

Note. DMT= Disease Modifying Therapies, MS= Multiple Sclerosis, *n*= Number, Nr= Not reported, PPMS= Primary Progressive MS, REMS= Rapidly Evolving MS, RRMS= Relapsing-Remitting MS, SD= Standard Deviation, SPMS= Secondary Progressive MS