The association between the Multiple Sclerosis Screening Questionnaire and objective measures of cognition: A systematic literature review and meta-analysis.

Jessica Tunmore[a], Eirini Kontou[b]*, Nima Moghaddam[a], Laurie Hufton[a], Roshan das Nair[c,d]

[a]*School of Psychology, University of Lincoln, Lincoln, UK;* [b]*School of Medicine, University of Nottingham, Nottingham, UK;* [c]*Health Services Research, SINTEF, Trondheim, Norway;* [d]*Mental Health & Clinical Neurosciences, University of Nottingham, Nottingham, UK.*

*Corresponding author: Eirini Kontou - Eirini.Kontou@nottingham.ac.uk

Word count: 6574

Abstract

Introduction: The Multiple Sclerosis Screening Questionnaire (MSNQ) is a self-report measure used to assess cognitive difficulties in people with Multiple Sclerosis (PwMS). The aim of this systematic review was to determine the associations between the MSNQ and: objective measures of cognition, measures of mood, and quality of life measures. Method: A comprehensive search was done across three databases (PsycINFO, MEDLINE, and CINAHL). A total of 15 studies, including 1992 participants, were selected for final inclusion. Meta-analyses were conducted to determine the pooled effect size of associations. Where data were not available for meta-analyses, a narrative synthesis approach was taken. Results: Significant, but small ($r = -0.17$), associations were found between the MSNQ and objective measures of cognition. Significant, moderate associations ($r = 0.47$) were found between the MSNQ and measures of mood. Conclusions: The small association between the MSNQ and objective measures of cognition shows that the measures do not converge well. However, their divergence may be important to map the broad construct of 'cognitive ability' more fully. Limitations include a lack of reporting of non-significant effect sizes in individual studies. Clinical implications include the potential for the MSNQ to be used beyond being solely a proxy measure for objective cognition. Future research should investigate the associations between the informant version of the MSNQ and objective measures.

Keywords: Multiple Sclerosis, Cognition, MSNQ, Measurement, Mood

Introduction

Multiple Sclerosis (MS) is a disease that impacts the central nervous system by causing disruption of communication between nerve signals. Whilst the exact cause of MS remains unknown it is suggested that the leading cause of MS is infection with the Epstein-Barr Virus (He et al., 2022). MS affects people differently and can cause symptoms such as tingling, vision and motor difficulties, fatigue, numbness, and changes in mood (National MS Society, 2022). People with MS (PwMS) also experience cognitive difficulties and these have been reported in between 40 and 80% of the MS population worldwide (DiGiuseppe et al., 2018; Fischer et al., 2014; Vanotti & Caceres, 2017). Common cognitive difficulties lie within the domains of information processing (Van Geest et al., 2018), and memory and learning (DeLuca et al., 2013). Cognitive difficulties have an impact on PwMS in different ways: fear of losing employment, fear of not being understood, difficulty managing social events, changes in daily living, quality of relationships, communication, and coping with MS – although this is not an exhaustive list (Mc Auliffe & Hynes, 2019; Halstead et al., 2020).

Cognition can be assessed using neuropsychological tests (NPT). Sometimes known as objective measures, NPT assess and detect cognitive impairment in different domains of cognition, for example, processing speed or memory. NPT are usually administered by a professional who has had training and they enable quantification of cognitive ability that is purported to be free of bias (Meca-Lallana et al., 2021). Whilst there is no agreement as to the most suitable instrument to assess cognition in PwMS (Oreja-Guevara et al., 2019), some of the most used NPT are the Brief Repeatable Battery of Neuropsychological Tests (BRBN-T) (Rao, 1990), Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) (Benedict et al., 2012), and the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS) (Benedict et al., 2002). There are also new digital technologies that enable PwMS to complete objective tests themselves on mobile devices (e.g.,

www.neuroms.org; Smith et al., 2021; van Dongen et al., 2019), however, there are contextual limitations to these including distractibility, misinterpretation, and individual variability in technology literacy and skills (Bauer et al., 2012).

Self-report measures, also known as subjective or patient-reported outcomes (PROMS), are also used to measure cognition in PwMS. Subjective measures rely on self-report of cognitive difficulties and, unlike NPT, do not test cognitive ability but rather ask questions to determine difficulties the individual might be having in different cognitive domains (Meca-Lallana et al., 2021). Self-report measures have advantages: they are less time-consuming, less expensive, less resource intensive (do not require a professional to administer), and arguably less strenuous for the people completing them (Elwick et al., 2021). However, it is important to consider that subjective reports of cognition can be influenced by mood and level of insight (Goverover et al., 2005), and fatigue (Kinsinger et al., 2010). Therefore, like most assessments, subjective reports of cognitive difficulties are not without their limitations.

Peoples' perceptions of their cognitive ability, as assessed by self-report measures, have been found to be poorly correlated with performance on objective assessments of cognition (Becker et al., 2012; Bruce et al., 2010; Maor et al., 2011; Middleton et al., 2006). This has important implications for how well performance on subjective measures might predict performance on objective measures. For example, self-reported positive evaluations of cognitive ability would not rule out objective cognitive impairment, which is important to consider for clinicians using these measures in their practice. Likewise, subjective measures could give information about someone's cognitive ability that would not be readily gauged by NPT.

As well as cognitive difficulties, mood complaints are common in people with MS. Mood disorders are reported in up to 54% of the MS population (Pérez et al., 2015). Depression can negatively impact performance on objective measures (Arnett et al., 2002; Nunnari et al.,

2014); however, depression has been found to be more strongly correlated with subjective measures of cognition than with objective measures (Benedict et al., 2004; Kinsinger et al., 2010; Yigit et al., 2021). Depression affects people's ability to appraise their cognitive difficulties accurately: the relationship between objective measures of cognition and subjective measures of cognition has been found to be mediated through depression (Julian et al., 2007; Kinsinger et al., 2010; Maor et al., 2001; Shilyansky et al., 2016). However, research has demonstrated a bidirectional relationship between mood and cognition (Vissicchio et al., 2019) and it is important to consider how emotion impacts cognition, and vice versa (Tyng et al., 2017). Further, other aspects of emotional distress such as fatigue and anxiety have also been found to influence the relationship between objective and subjective cognition in PwMS (Davenport et al., 2022).

Objective and subjective assessment results are also related to quality of life (QoL) outcomes. Research has found that better subjective cognitive functioning predicts higher QoL and lower depressive symptoms (Crouch et al., 2022). A recent systematic review identified depression as a risk factor for lower QoL in PwMS (Gil-González et al., 2020), and individuals with greater subjective cognitive difficulties are more likely to have lower QoL; depression provides one explanation as to why that is (Crouch et al., 2022). This highlights an important consideration; perceived cognitive difficulties in people with MS are clinically important, regardless of whether they correlate with objective measures. Perceived cognitive difficulties can negatively impact; QoL (Samartzis et al., 2014); employment status (van Wegen et al., 2022); sleep quality, stress, and depression levels (Lamis et al., 2018). Given the associations between perceived cognitive difficulties it is important to consider the different intervention methods for PwMS. In the absence of objective deficits in cognitive functioning, that would lend itself to referral to an occupational therapist or neuropsychologist (NICE, 2022), other psychological interventions targeting the areas of difficulty associated with perceived cognitive impairments,

that may be missed by objective assessment alone, can have positive outcomes for overall wellbeing for PwMS (Dayapoğlu & Tan, 2012; Gottberg et al., 2016).

In the context of MS, a population-specific measure of subjective cognitive difficulties has been developed. The Multiple Sclerosis Neuropsychological Questionnaire (MSNQ) is a 15-item self-report measure of cognition that assesses the cognitive domains of processing speed, attention, memory, emotional control, and social skills.[1] Each item is ranked from 0-4 (0= does not occur, 4 = occurs very often, very disruptive) (Benedict et al., 2003).

The MSNQ is widely used as an assessment tool in MS care and research. There are two versions of the MSNQ: the MSNQ-Self Report (MSNQ-S) and the MSNQ- Informant report (MSNQ-I) A recent systematic review found 5665 measures of cognition used in PwMS, of which the MSNQ was the most frequently used subjective measure (Elwick et al., 2021). However, a systematic review of everyday measures for MS found that, whilst again the MSNQ was the most widely used, there were inconsistencies across the reporting of reliability and validity of the measure (das Nair et al., 2019). Out of 14 papers identified in the review that used the MSNQ, only four reported on the internal consistency and only two reported on the interrater reliability. This is problematic as missing information around reliability and validity estimates makes it difficult to ascertain how well the measure performs across samples and contexts. Although the MSNQ performed better than other self-report measures, the reporting was inconsistent; the argument could be made that whilst the performance of the MSNQ is questionable (as outlined by validity concerns), in relative terms, it outperforms other candidate measures and therefore its relative popularity is arguably justified. Others have

---

[1] The MSNQ has two forms: MSNQ – S (self-report) and MSNQ- I (informant report). The focus of the review is on self-report of cognitive ability.

argued that the MSNQ is a clinically reliable and valid screening tool that should be routinely used in MS care (Rogers & Panegyres, 2007).

Benedict et al. (2003) in their initial study found that the MSNQ-S was significantly correlated with depression, but not with objective measures of cognition. In a later, larger, study the MSNQ-S was significantly correlated with objective measures (r=-.49), although still to a lesser degree than with self-reported depression (r=-0.61) (Benedict et al., 2004). The negative relationship here reflects that those greater subjective difficulties (higher MSNQ score) were associated with weaker performance on cognitive tests. Therefore, it could be suggested that the MSNQ provides complementary information in cognitive assessments.

Detection of cognitive impairment is clinically relevant (Meca-Lallana et al., 2021); however, subjective measures of cognition may be adjunctive to NPT and have relative validity strengths as outlined previously. The aim of this review was to explore the associations between the MSNQ and outcomes of interest for PwMS: i) objective cognition ii) mood, and iii) quality of life. It is important to understand the degree of convergence, or indeed divergence, across these measures to map the broad construct of cognitive ability more fully. Understanding the degree to which these measures converge will support clinicians to use these measures in the most clinically meaningful and helpful way. Meaningful assessment of cognitive difficulties will enable cognitive interventions and support to be delivered in a person-centred way which will improve overall care for PwMS (Kalb et al., 2018).

Objectives

The review aims to address the following objectives, based on the available literature in PwMS to determine:

i)   What is the association between MSNQ-assessed subjective cognitive functioning and objective cognitive functioning?

ii)   What is the association between the MSNQ and mood, and the MSNQ and QoL?

iii)   What is the quality of the MS research that examines these associations?

Methods

Protocol and registration

This review study was registered and approved by the University of Lincoln Ethics Committee (reference: 2022_9803). The study was also registered on PROSPERO (ID: 342395).

Eligibility criteria

Papers that i) had an MS sample and ii) reported on the association between MSNQ-S and objective measures of cognition were included. Only studies that use the original variation of the MSNQ were included.

Exclusion criteria

All study designs were included apart from review or meta-review/analysis papers, which were excluded. Papers that were not published in peer-reviewed publications were excluded. Validation studies for translated or other versions of the MSNQ were excluded.

Search strategy

A search strategy was developed, in consultation with an information specialist, as presented in Table 1 below.

[Table 1 here]

A systematic search for research articles was undertaken using EBSCO host; specifically searching the PsycINFO, MEDLINE, and CINAHL databases from the date of inception to 16/07/2022. These databases were selected as the scoping exercise identified them as most

likely to contain the journals where the studies of interest would be published. These databases were searched separately and then the results combined using EndNote reference manager. Duplicates were removed using the duplication tool and any missed duplicates were removed manually. The reference lists of included papers were manually searched for additional papers.

Data extraction

Data were extracted by the researcher and a researcher external to the research team. Extracted data were cross-checked for any errors and discrepancies were resolved through discussion with the research team.

Data were aggregated and/or transformed where appropriate. To harmonise the direction of the relationship across the measures, MSNQ scores were treated such that higher scores indicated subjective cognitive difficulties and lower scores on objective measures indicated lower cognitive functioning; thus, negative correlations indicated a negative relationship between the MSNQ and objective measures of cognition, such that the higher the subjective difficulties, the lower the cognitive functioning. Measures of mood were treated as such that higher scores indicated worse mood therefore positive correlations between measures of mood and the MSNQ indicate a positive relationship between mood and subjective cognitive functioning. Scores were uniformly reported in this direction across studies. Where available, the strength of associations between the following were extracted from each paper (*r*).

1) MSNQ and objective measures of cognition

2) MSNQ and mood measures

3) MSNQ and measures of QoL

This review does not include MSNQ-I data as the focus is on patient reported data on cognition. Therefore, studies including MSNQ-I were included, however only MSNQ-S data were extracted.

Where the data were not reported, emails were sent to authors to request data.

When two or more associations with the MSNQ-S were reported within a domain (objective cognition, mood, quality of life), the *r* values were averaged within that domain as a single aggregate.

Where a battery of neuropsychological tests (objective measures of cognition) was administered, the composite score was used (if given). When two or more domains or sub-score *r*-values were reported for a single measure, and no total *r*-value given, the *r*-values were averaged and presented as a single aggregate.

Additional clinical and demographic data were collected. Where available, age (mean and standard deviation), gender, ethnicity, years of education, and MS course were collected. Level of education is important to consider and has been found to moderate outcomes on cognitive tests.

Quality assessment and study specific risk of bias.

The Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018) was used to assess the quality of the included studies. To assess the inter-rater reliability of quality appraisal, 50% of studies were randomly selected for independent coding by a second reviewer, with any discrepancies resolved through discussion. The MMAT is a critical appraisal tool that can be used to assess quantitative, qualitative, and mixed methods studies (Hong et al., 2018). All the studies within this review are quantitative descriptive designs; the MMAT was chosen due to its criteria on the assessment of these study designs. The tool uses five questions covering: sampling strategy, sample representativeness, measurement methods, non-response bias, and statistical analysis (Appendix A). The MMAT discourages quantification of risk of bias assessment and instead advocates a narrative synthesis of each criterion.

Data synthesis

Meta-Essentials (Suurmond et al., 2017) was used to conduct the meta-analyses. The associations of interest were between the MSNQ and objective measures of cognition, and the MSNQ and measures of mood and QoL. However, not all studies provided adequate data for meta-analyses and these were synthesised narratively.

Random effects models were used throughout the meta-analyses due to the variability in objective measures of cognition, and to account for within-study estimation error variance and between-studies variance. The random effects model allows us to generalise our conclusions beyond saying there is a true effect size in each study (Borenstein et al., 2010).

The variability between studies was examined and represented as I^2 statistic. The I^2 statistic is used to quantify heterogeneity within meta-analysis studies and describes the percentage of variability in the point estimates that is a result of heterogeneity as opposed to the variability being from sampling error (Higgins & Thompson, 2002). It is advised to consider the I^2 statistic within the context of clinical and methodological aspects of the studies and a strict categorisation based on the numerical value is advised against. However, the I^2 can be broadly categorised into low (25%), moderate (50%) and high (75%) (Higgins et al., 2003).

The prediction intervals (PI) were to be determined. PI demonstrates the range of the predicted effect sizes in future studies or studies not included in the meta-analysis (Nagashima et al., 2019).

Planned sensitivity analyses removing studies with outlier effect sizes, as determined by funnel plots, were carried out to determine the robustness of the pooled effect size.

Risk of Bias

Risk of publication bias was determined using funnel plots. The results of the funnel plots informed further sensitivity analyses.

The primary summary measure was the correlation, as indicated by Pearson's *r*, between the MSNQ and objective measures of cognition. Cohen's thresholds for interpreting Pearson r indicate effect sizes in the following categories: small ≥ .10, medium ≥ .30, large ≥ .50 (Cohen, 1988).

Results

Study selection

A total of 15 studies were included in the final review, with 715 screened at the abstract level, and 27 full texts examined. Figure 1 demonstrates the search strategy employed within the PRISMA flow chart.

[Figure 1 here]

Characteristics of included studies

The characteristics of included studies are presented in Table 2. A total of 15 studies were included in the final review. In total there were 1992 participants, with 1419 females, 571 males, and three who did not state their gender. The mean age (SD) was 45.33 (9.8). This excludes age data from one study (8) that reported their age as confidence intervals.

[Table 2 here]

Of these 15 studies, 12 (n=1745) had adequate data for the first meta-analysis examining the association between MSNQ and objective measures. The remaining studies did not report effect sizes or the equivalent beta values for transformation (1, 5, 6); data were requested from study

authors, but no response was received and so these studies were not included in the meta-analysis (though retained in our narrative synthesis).

To assess the associations between the MSNQ and measures of mood, a total of nine studies (n=905) had adequate data to be included in the second meta-analysis (2, 3, 6, 12, 14, 15).

Only three out of the 15 studies included a QoL measure (5, 13, 14), and only one of these studies (14) included estimate effects for the relationship between the MSNQ and the QoL measure. Therefore, the association between the MSNQ and QoL was not addressable via meta-analysis and is therefore examined narratively.

All studies were conducted from 2003 onwards. Eight studies were conducted in the United States of America (1, 2, 3, 4, 10, 11, 12, 15) and the rest were conducted in: Canada (5, 6), Ireland (7), Belgium (8), Italy (9), Turkey (13), Finland (14). Three studies had the same primary author (1, 2, 3), who was also second author on study 4. All studies were set in outpatient clinics or MS centres apart from one study that did not give the setting (15).

The average years of education was 14.3. This excludes data from four studies which did not report mean years in education (6, 7, 8, 14). All studies except one reported data on MS course (1). As reported by the studies, the most common course of MS was relapsing remitting (n=1613), followed by primary progressive (n=152), secondary progressive (n=113), clinically isolated syndrome (n=7) and progressive relapsing (n=1). Data were missing for 11 participants (10).

Nine studies described using a full NPT battery (3, 5, 7, 9, 10, 11, 13, 14, 15). The most common measure of mood was the BDI (1, 2, 3, 4, 9, 11, 12, 13,14, 15).

Quality appraisal and risk of bias within studies

Overall, the quality of studies was mixed as can be seen in Table 3.

[Table 3 here]

The level of agreement between reviewers was assessed and, prior to resolving discrepancies, overall free-marginal kappa = .74 ('substantial' agreement).

Most studies used a sampling strategy that was relevant to address the research question. Whilst it was not explicitly stated in studies, convenience sampling was the most used strategy with studies mainly recruiting from MS centres and outpatient clinics. Three studies used secondary data from other studies and did not detail the sampling strategy (5, 10, 15).

Overall, the representativeness of study samples was mixed. None of the studies reported full data on ethnicity and only five reported incomplete data on ethnicity. Four studies only reported the percentage of "Caucasian" participants (1, 2, 3, 4) and one study only categorised ethnicity as "Hispanic" or "non-Hispanic") (10). The MS course was fairly well represented across studies; however, the predominant course was relapsing remitting. There are higher prevalence rates of MS in women than in men (Dunn et al., 2015) and the included studies reflect that in their participant numbers.

All but three studies used validated measures suitable to measure either cognition, mood, or QoL. One study used a two-question screening tool to measure depression (8) and another (5) did not give the name of the depression measure used. One study stated within their paper that the Executive Functioning Performance Test (EFPT) was developed and validated for use in patients with dementia, not PwMS (12).

For most studies, it was difficult to assess the risk of non-response bias as non-response was not reported. One could assume that, as it was not explicitly stated that there were people that did not complete all measures, all participants did indeed complete all measures. One study reported that 17 patients withdrew from the initial examination and then 85 withdrew from the study after the second assessment (3).

The statistical analysis in all studies was appropriate to answer the research questions, however many studies only gave correlational data for statistically significant findings (1, 3, 4, 5, 10, 12).

Synthesis of results

MSNQ and objective measures of cognition

To assess the associations between the MSNQ and objective measures of cognition, a total of 12 studies (n=2198) had adequate data to be included in the meta-analysis (2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 15). With these 12 studies included in a random effects model, a small significant effect size was seen.

Table 4 and Figure 2 presents the results from the meta-analysis of the associations between MSNQ and objective measures of cognition.

[Table 4 here]

[Figure 2 here]

Combining the estimates from studies that assessed the association between MSNQ performance and performance on objective measures of cognition, a significant small, pooled effect size was found (r = -0.17; CI [-0.28 – -0.05]; Z= -3.06; p=.0001). Correlations ranged from r=-0.49 (2) to r=0.19 (14). Substantial heterogeneity was observed (I^2=72.21.), demonstrating between-study variability. The PI demonstrates that effect size is between -0.47 and 0.17, which a very broad range.

The central r estimate was converted to r^2 (r^2= 0.029) indicating the degree of overlap. This demonstrates <5% (r^2= 0.029) of the variance in objective performance is explainable in terms of assessed subjective difficulties, and vice versa.

The observed strength of association between the MSNQ and objective measures of cognition ranged from -0.49 (2) to 0.19 (14). Only two studies (12, 14), reported positive correlations between the MSNQ and an objective measure. However, for both studies, the effect size was very weak (r= <0.2).

Risk of bias across studies

Funnel plots were used to investigate outliers and potential risk of bias across studies. The funnel plot (Figure 3) shows three outliers when examining the associations between the MSNQ and objective measures of cognition.

[Figure 3 here]

The asymmetrical funnel plot indicates there is publication bias: most studies are clustered in the upper left quadrant. There seems to be an under-representation of the smaller associations than would be expected given the combined estimate. Given that there are 12 studies included in this analysis, it is important to note the publication bias here; this may play a role in the results and should therefore be interpreted with a degree of caution (Suurmond et al., 2016)

Secondary sensitivity analysis

When only the studies that used the MACFIMS or BICAMS were included, the heterogeneity significantly decreased ($I^2 = 0\%$) and there was a slight increase in the pooled effect size ($r = -0.25$; CI [-0.33 - -0.17]; Z= -8.63; p<.001).

The outlying effect sizes from studies (2, 14), as identified in Figure 3, were removed. The removal of these outliers considerably reduced the heterogeneity ($I^2 = 0\%$), had small impact (-0.1) on the pooled estimate, and reduced the size of the confidence intervals (r= 0.18; CI [0.24 - = 0.13]; Z= -7.24; p<.001). The prediction interval also became narrower (-0.24 - -0.13).

MSNQ & Mood

To examine the associations between the MSNQ and measures of mood, a total of 11 studies (n=1054) had adequate data to be included in the meta-analysis (1,2,3,6,7,9,10,11,12,14,15). Higher scores on measures of mood indicated greater difficulties with mood and greater distress. Table 5 and Figure 4 demonstrates the results of this meta-analysis

[Table 5 here]

[Figure 4 here]

With these 11 studies, included in a random effects model, a moderate, significant effect size was seen (r = 0.47; CI [0.35 - 0.57]; Z= 8.10; p<.0001). Correlations ranged from -0.03 (6) to 0.61 (2). Moderate heterogeneity was observed (I^2 = 69.42 %), demonstrating between-study variability. The PI demonstrates an effect size between 0.12 – 0.71, which is a broad range.

Figure 5 shows the funnel plot for studies when looking at the association between the MSNQ and measures of mood.

[Figure 5 here]

Three studies are outliers (2, 6, 9). The funnel plot is fairly symmetrical which is indicative of a moderate risk of publication bias.

Secondary sensitivity analyses

Removal of the three outlier studies, as identified by the funnel plot, reduces the heterogeneity (I^2 = 39. 97 %). However, the pooled estimate only very slightly changed, as did the confidence interval (r=0.50; CI [0.42 – 0.58]; Z= 12.27, p<.001). The prediction interval demonstrates an effect size between 0.32 – 0.65, which is a broad range.

Narrative syntheses

Three studies did not provide adequate data to be included in the meta-analysis looking at the association between objective measures of cognition and the MSNQ (1, 5, 10).

Four studies (4, 5, 8, 13) did not provide adequate data to analyse the relationship between the MSNQ and measures of mood. Only three studies (5, 13, 14) included QoL as a construct to measure alongside the MSNQ and an objective measure of cognition. Of these, only one study (14) provided correlational data between QoL and the MSNQ. Therefore, the results of these studies required narrative synthesisation.

Two studies (1, 5) state that the MSNQ scores were not significantly correlated with cognitive measures however they did not give the effect size or direction, instead just reported them as "non-significant". One study did not provide correlational data for the MSNQ and an objective measure, however, regression analyses found that in the case of depression, self-reported cognitive measures are unlikely to provide accurate estimates of objective cognitive functioning (10).

Four studies provided no or inadequate correlational data between the MSNQ and measures of mood (4 5, 813,). Two studies (5, 13) did not report on the associations between the MSNQ and measures of mood. One study (4) found that depression scores were higher in people who underestimated their cognitive ability on the MSNQ, and the inverse was also true. One study (8) used only a 2-item screening tool for depression as the measure of mood. They reported no other information on the mood measure and categorise responses into either "depressed" or "non-depressed". They presented correlational data for "depressed" and MSNQ rho= 0.30 ($p<0.05$).

Two studies did not provide data on the association between the MSNQ and QoL (5, 13). Only one study provided correlational data between the MSNQ and a measure of QoL (14). For both relapsing and progressive groups, quality of life was strongly correlated with the MSNQ and

the measures of mood. The WHOQoL was used and correlation data for both relapsing and progressive groups were provided on each of the four sub scores of the WHOQoL as follows; for progressive group: physical health $r= 0.165$ ($p=0.054$), psychological total score $r= 0.445$ ($p <0.001$), social relationship $r=0.213$ ($p=0.012$), environment $r=0.294$ ($p<0.001$); for the relapsing group, physical health $r=0.59$ ($p<0.001$) psychological $r=-.467$ ($p<0.001$) social $r=0.362$ ($p=0.007$), environment $r=0.367$ ($p=0.006$).

Discussion

There are many individual studies in the literature that compare subjective measures of cognition with objective measures of cognition. This review chose to focus on the MSNQ as the self-reported measure of cognition as it is the most widely used in clinical practice and research (Elwick et al., 2021). This is the first systematic review and meta-analysis that takes into consideration the magnitude, and indeed the quality, of the available evidence on the association between the MSNQ and objective measures of cognition; measures of mood; and measures of QoL.

The findings of this review indicate (i) small, but significant, negative associations between the MSNQ and objective measures of cognition, (ii) moderate, significant, positive associations between the MSNQ and measures of mood, (iii) variability in the quality of evidence examining these associations.

The MSNQ and objective cognition were found to be negatively correlated, as expected, although this association was weak. This indicates that scores on objective measures of cognition do not converge with scores on the MSNQ. Although they are associated, it is clear they are largely independent, with <5% ($r^2= 0.029$) of the variance in objective performance being explainable in terms of assessed subjective difficulties, and vice versa. This would indicate that patient-reported cognitive difficulties, as measured by the MSNQ, do not reflect

objective cognitive difficulties as determined by neuropsychological assessment. Likewise, the reverse is also true, and objective performance does not reflect the subjective difficulties that people with MS are reporting to affect their daily lives.

Discordance between subjective (self-report) and objective (observed behaviour) measures could reflect methodological issues (e.g., low reliability in one or both measures [and reliability information was often unreported at the level of individual studies]), but may also evidence a more fundamental distinction, in their dependence on differential response processes (Dang et al., 2020). Specifically, objective measures (1) gauge behaviour in a specific controlled situation, (2) draw on performance metrics like speed and accuracy, and (3) tend to capture 'best possible' performance. In contrast, subjective measures (1) elicit reflections on behaviours across multiple uncontrolled real-life situations, (2) draw on subjective perceptions about performance, and (3) tend to capture 'everyday' performance. Considering differences in underpinning response processes, divergence between objective and subjective measures is perhaps not surprising – but does problematise their use as indicators of a shared 'cognition' construct. Treated as discrete constructs, they may offer incrementally useful information for understanding individual functioning (e.g., in terms of both ability and typical daily performance).

The findings of this review also indicate that there is a positive, moderate, and significant association between the MSNQ and measures of mood. As discussed, mood plays a role in the self-report of cognitive difficulties, and there may be various directions of influence or overlap. Perceived cognitive difficulties may impact mood, and reports of mood could also be secondary to other factors such as symptom burden or response style. Mood has been found to influence scores on subjective measures of cognition, such that PwMS who have worse subjective functioning than objective functioning and have higher depression scores (Davenport et al., 2021; Hughes et al., 2019). This lends itself to important considerations in the interpretation of

MSNQ-S scores. The presence of low mood can impact a person's self-report of cognitive difficulties, and they are more likely to underestimate their ability (Davenport et al 2021; Hughes et al., 2019; Rosti - Ostajavri et al., 2014).

Dagenais et al. (2013) was the only study to report a negligible negative correlation with mood; however, this correlation was not significant (p=.847). The authors suggest this was due to the absence of depression within the study sample. As discussed, we know that in MS populations there are high rates of mood disorders and so it could be argued that this study was not representative of the MS population more broadly.

Whilst there was some variability in how objective cognition was assessed, most studies used a neuropsychological test battery that had been validated for use in PwMS. The MSNQ captures information on neuropsychological abilities in activities of daily living (Benedict et al., 2004). However, it is important to note that the MSNQ is not directly comparable to all objective measures. For example, many studies in this review used either the full BICAMS battery or individual tests from the BICAMS battery. The BICAMS assesses information processing speed, and verbal and visual memory, and these are not directly assessed on the MSNQ. Therefore, it is important to consider this when interpreting the associations between the MSNQ and objective measures of cognition (Davenport et al., 2021).

The overall quality of studies was mixed, and this limits the strength of conclusions that can be drawn from the reviewed evidence. The most significant issue across studies was missing data: multiple studies only reported the significant associations. This is an issue that is widespread across research in general, but particularly important when conducting a meta-analysis. The publishing of only significant effect sizes limits the extent to which the meta-analysis can draw conclusions. A meta-analysis allows us to go beyond significant results in individual studies, to get a precise and well-powered overall estimate of the effect size. However, if the non-

significant effects sizes are not reported in individual studies, it has the potential to bias the results towards larger estimates of effect.

Future research, limitations, and implications

This review has demonstrated that the MSNQ does not correlate well with objective measures of cognition. This highlights potential difficulties as it demonstrates that objective measures of cognition do not capture the perceived cognitive difficulties that PwMS are living with. Thus, subjective and objective measures of cognitive cannot be used interchangeably, as indicators of a unitary construct. Rather, these measures may offer complementary information, enabling clinicians to gain a synthesised understanding of the everyday functional impact of MS for their patients (Weber et al., 2019).

There were inconsistencies across studies providing domain-specific data and therefore the current review focuses on global cognition, as measured by objective measures, rather than domain-specific cognition. Future research should explore correlations at the domain specific level as this would allow for a more nuanced meta-analysis which would make important contributions to understanding the relationship between objective and subjective measures of cognition.

The Cognition in Multiple Sclerosis Scale (CoMSS), a newly developed self-report measure of cognition, is the first of its kind and was co-developed with PwMS (Cogger-Ward et al., 2022). Co-development of measures ensures they are grounded within the everyday experiences of people who use them (Phillips & Morgan, 2014). Future research should examine the relationship between this new measure and existing measures of cognition to determine its utility in clinical and research settings.

Whilst the strength of this review is that we were able to examine the available evidence of the association between the MSNQ-S and measures of mood, and measures of cognition, it is

important to acknowledge its limitations. This review only focused on MSNQ-S reports of cognition. The results of many studies indicate that MSNQ-I is better correlated with objective measures of cognition than with the MSNQ-S (Benedict et al., 2003; Benedict et al., 2004; Dagenais et al., 2013). Future reviews could investigate the pooled effect size of this association. This would be important in MS care as, if informant scores on MSNQ are better correlated with objective performance, this would be an important consideration when including carers and informants in appointments/clinic visits in routine care. Furthermore, this review did not examine whether level of education had any impact on the association between MSNQ and objective measures of cognition. It would have been important to consider here, as interestingly, the mean years of education were higher in this population than was the average years of education in their respective countries. Whilst this review focused on mood, other aspects of emotional distress such as fatigue and anxiety which have been found to influence the relationship between objective and subjective cognition in PwMS (Davenport et al., 2022), were not included in this review and therefore further research should explore these associations.

The measures of mood and the subjective measures of cognition both rely on self-report. Common method variance is a type of bias born from using the same method of measurement (i.e., self-report) and can compromise the relationship between the constructs being measured (Williams & McGonagle, 2016). This type of bias is inherent in measurements of mood and future research should employ objective measurements of mood to explore this issue.

Lastly, a possible limitation of the current study is that 50% of the studies were included in the interrater reliability (IRR) assessment for the quality appraisal of included studies. Whilst it is not uncommon practice to use a sample of works when assessing IRR in reviews, it is acknowledged that this could be seen as a limitation.

The limited association between the MSNQ and objective measures of cognition shows that the measures do not converge. However, their divergence may be important to help map the broad construct of 'cognitive ability' more holistically. A recent literature review informed a set of clinical recommendations that advocate for the use of cognitive evaluation for all patients diagnosed with MS, throughout their care, including follow up. It is suggested that periodic evaluation of cognitive function in MS would have many positive impacts and would provide information that could support cognitive rehabilitation and enhancement of the cognitive reserve (Meca-Lallana et al., 2021). The gap between subjective and objective measures of cognition suggests that subjective measures can be helpful as more than just proxies for objective assessment; they can provide complementary information that captures the perceived difficulties PwMS are living with, resulting in a cognitive assessment that is person-centred, something that this review would indicate is difficult to achieve with objective measurement alone. The MSNQ captures different information and is perhaps more useful in supplying incremental information on people's cognitive ability, as opposed to being used to predict performance on neuropsychological tests.

The correlation between the MSNQ and objective measures of cognition up to this point has not been analysed beyond the individual study level. This review highlighted that peoples' perceptions of their cognitive ability are poorly associated with their performance on objective assessments of cognitive ability. This has important implications for how subjective measures should be used to predict performance on objective measures. A subjective measure may pick up on everyday difficulties that are not readily gauged under conventional testing conditions, but positive self-evaluations on a subjective measure would not rule out objective cognitive impairment. It is instead suggested that subjective measures should be used as an adjunct to objective measures of cognition. The suggestion that a 'good' subjective measure of cognition is one which correlates with objective measures gives primacy to performance on objective

measures, such that objective measurement of cognition is the 'true' way to measure cognition. However, it could be argued that the two forms of measures provide equally clinically important information. As discussed, detection of cognitive difficulties in people with MS is important in treatment outcomes for people with MS (Kalb et al., 2018), and ultimately improving the life and care of PwMS is the goal of services responsible for the care of PwMS and their loved ones.

References

Arnett, P. A., Higginson, C. I., Voss, W. D., Randolph, J. J., & Grandey, A. A. (2002). Relationship between coping, cognitive dysfunction and depression in multiple sclerosis. *The Clinical Neuropsychologist*, *16*(3), 341-355. https://doi.org/10.1076/clin.16.3.341.13852

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *The Clinical Neuropsychologist*, *26*(2), 177-196. https://doi.org/10.1080/13854046.2012.663001

Becker, H., Stuifbergen, A., & Morrison, J. (2012). Promising new approaches to assess cognitive functioning in people with multiple sclerosis. *International Journal of MS Care*, *14*(2), 71-76. https://doi.org/10.7224/1537-2073-14.2.71

Benedict, R. H., Fischer, J. S., Archibald, C. J., Arnett, P. A., Beatty, W. W., Bobholz, J., ... & Munschauer, F. (2002). Minimal neuropsychological assessment of MS patients: a consensus approach. *The Clinical Neuropsychologist*, *16*(3), 381-397. https://doi.org/10.1076/clin.16.3.381.13859

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

Benedict, R. H., 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., Duquin, J. A., Jurgensen, S., Rudick, R. A., Feitcher, J., Munschauer, F. E., ... &
Weinstock-Guttman, B. (2008). Repeated assessment of neuropsychological deficits in
multiple sclerosis using the Symbol Digit Modalities Test and the MS Neuropsychological
Screening Questionnaire. *Multiple Sclerosis Journal*, *14*(7), 940-946.
https://doi.org/10.1177/1352458508090923

Benedict, R. H., Amato, M. P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., ... & Langdon, D.
(2012). Brief International Cognitive Assessment for MS (BICAMS): international standards
for validation. *BMC Neurology*, *12*(1), 1-7. https://doi.org/10.1186/1471-2377-12-55

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to
fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2),
97-111. https://doi.org/10.1002/jrsm.12

Bruce, J. M., et al. (2010). "Self-reported memory problems in multiple sclerosis: Influence of
psychiatric status and normative dissociative experiences." *Archives of Clinical
Neuropsychology* 25(1), 39-48. https://doi.org/10.1177/1352458510362819

Carone, D. A., Benedict, R. H. B., MUNSCHAUER III, F. E., 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

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. *Canadian Journal of Neurological Sciences*, *47*(5), 620-626.
https://doi.org/10.1017/cjn.2020.97

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Routledge.

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

Dagenais, E., Rouleau, I., Demers, M., Jobin, C., Roger, É., Chamelian, L., & Duquette, P. (2013).

    Value of the MoCA test as a screening instrument in multiple sclerosis. *Canadian Journal of*

    Neurological Sciences, *40*(3), 410-415. https://doi.org/10.1017/s0317167100014384

dasNair, R., Griffiths, H., Clarke, S., Methley, A., Kneebone, I., & Topcu, G. (2019). Everyday

    memory measures in multiple sclerosis: a systematic review. *Neuropsychological*

    Rehabilitation, 29(10), 1543–1568. https://doi.org/10.1080/09602011.2018.1434081

Cogger - Ward, H., Moghaddam, N., Evangelou, N., & dasNair, R. (2022). *Cognition in Multiple*

    Sclerosis Scale (CoMSS): Development and Preliminary Testing [Unpublished doctoral

    dissertation]. School of Psychology, University of Lincoln.

Dang, J., King, K. M., & Inzlicht, M. (2020). Why are self-report and behavioral measures weakly

    correlated?. *Trends in Cognitive Sciences, 24*(4), 267-269.

    https://doi.org/10.1016/j.tics.2020.01.007

Davenport, L., Cogley, C., Monaghan, R., Gaughan, M., Yap, M., Bramham, J., ... & O'Keeffe, F.

    (2022). Investigating the association of mood and fatigue with objective and subjective

    cognitive impairment in multiple sclerosis. *Journal of Neuropsychology, 00,* 1–18.

    https://doi.org/10.1111/jnp.12283

Dayapoğlu, N., & Tan, M. (2012). Evaluation of the effect of progressive relaxation exercises on

    fatigue and sleep quality in patients with multiple sclerosis. *The Journal of Alternative and*

    Complementary Medicine, 18(10), 983-987. https://doi.org/10.1089/acm.2011.0390

DeLuca, J., Leavitt, V. M., Chiaravalloti, N., & Wylie, G., (2013). Memory impairment in multiple sclerosis is due to a core deficit in initial learning. *Journal of Neurology*, *260*(10), 2491-2496. https://doi.org/10.1007/s00415-013-6990-3

D'hooghe, M. B., De Cock, A., Van Remoortel, A., Benedict, R. H., Eelen, P., Peeters, E., ... & Nagels, G. (2020). Correlations of health status indicators with perceived neuropsychological impairment and cognitive processing speed in multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *39*, 101904. https://doi.org/10.1016/j.msard.2019.101904

DiGiuseppe, G., Blair, M., & Morrow, S. A. (2018). Prevalence of cognitive impairment in newly diagnosed relapsing-remitting multiple sclerosis. *International Journal of MS Care*, *20*(4), 153-157. https://doi.org/10.7224/1537-2073.2017-029

Dunn, S. E., Gunde, E., & Lee, H. (2015). Sex-based differences in multiple sclerosis (MS): part II: rising incidence of multiple sclerosis in women and the vulnerability of men to progression of this disease. *Emerging and Evolving Topics in Multiple Sclerosis Pathogenesis and Treatments*, 57-86. https://doi.org/10.1007/7854_2015_370

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

Fenu, G., Lorefice, L., Carta, E., Arru, M., Carta, A., Fronza, M., ... & Cocco, E. (2021). Brain volume and perception of cognitive impairment in people with multiple sclerosis and their caregivers. *Frontiers in Neurology*, 570. https://doi.org/10.3389/fneur.2021.636463

Fischer, M., Kunkel, A., Bublak, P., Faiss, J. H., Hoffmann, F., Sailer, M., ... Köhler, W. (2014). How reliable is the classification of cognitive impairment across different criteria in early and late stages of multiple sclerosis? *Journal of the Neurological Sciences, 343*(1–2), 91-99. https://doi.org/10.1016/j.jns.2014.05.042

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), Article e041249. http://dx.doi.org/10.1136/bmjopen-2020-041249

Goverover, Y., Chiaravalloti, N., & DeLuca, J. (2005). The relationship between self-awareness of neurobehavioral symptoms, cognitive functioning, and emotional symptoms in multiple sclerosis. *Multiple Sclerosis Journal*, *11*(2), 203-212. https://doi.org/10.1191/1352458505ms1153oa

Gottberg, K., Chruzander, C., Backenroth, G., Johansson, S., Ahlström, G., & Ytterberg, C. (2016). Individual face-to-face cognitive behavioural therapy in multiple sclerosis: a qualitative study. *Journal of Clinical Psychology, 72*(7), 651-662. https://doi.org/10.1002/jclp.22288

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

He, R., Du, Y., & Wang, C. (2022). Epstein-Barr virus infection: the leading cause of multiple sclerosis. *Signal Transduction and Targeted Therapy*, *7*(1), 239. https://doi.org/10.1038/s41392-022-01100-0

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, *327*(7414), 557-560. https://doi.org/10.1136/bmj.327.7414.557

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M. C., Vedel, I., & Pluye, P. (2018). The

Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285-291. https://doi.org/10.3233/efi-180221

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

Julian, L., Merluzzi, N. M., & Mohr, D. C. (2007). The relationship among depression, subjective cognitive impairment, and neuropsychological performance in multiple sclerosis. *Multiple Sclerosis Journal, 13(*1), 81-86. https://doi.org/10.1177/1352458506070255

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

Kim, S., Zemon, V., Rath, J. F., Picone, M., Gromisch, E. S., Glubo, H., ... & Foley, F. W. (2017). Screening instruments for the early detection of cognitive impairment in patients with multiple sclerosis. *International Journal of MS care*, *19*(1), 1-10. https://doi.org/10.7224/1537-2073.2015-001

Kinsinger, S. W., Lattie, E., & Mohr, D. C. (2010). Relationship between depression, fatigue, subjective cognitive impairment, and objective neuropsychological functioning in patients with multiple sclerosis. *Neuropsychology, 24*(5), 573. https://doi.org/10.1037/a0019222

Lamis, D. A., Hirsch, J. K., Pugh, K. C., Topciu, R., Nsamenang, S. A., Goodman, A., & Duberstein, P. R. (2018). Perceived cognitive deficits and depressive symptoms in patients with multiple

sclerosis: Perceived stress and sleep quality as mediators. *Multiple Sclerosis and Related Disorders*, *25*, 150-155. https://doi.org/10.1016/j.msard.2018.07.019

Maor, Y., Olmer, L., & Mozes, B. (2001). The relation between objective and subjective impairment in cognitive function among multiple sclerosis patients–the role of depression. *Multiple Sclerosis Journal 7*(2) 131–135. https://doi.org/10.1177/135245850100700209

Meca-Lallana, V., Gascón-Giménez, F., Ginestal-López, R. C., Higueras, Y., Téllez-Lara, N., Carreres-Polo, J., ... & 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

Mc Auliffe, A., & Hynes, S. M. (2019). The impact of cognitive functioning on daily occupations for people with multiple sclerosis: a qualitative study. *The Open Journal of Occupational Therapy*, *7*(3), 1-12. https://doi.org/10.15453/2168-6408.1579

Middleton, L. S., Denney, D. R., Lynch, S. G., & Parmenter, B. (2006). The relationship between perceived and objective cognitive functioning in multiple sclerosis. *Archives of Clinical Neuropsychology*, *21*(5), 487-494. https://doi.org/10.1016/j.acn.2006.06.008

Nagashima, K., Noma, H., & Furukawa, T. A. (2019). Prediction intervals for random-effects meta-analysis: a confidence distribution approach. *Statistical Methods in Medical Research*, *28*(6), 1689-1702. https://doi.org/10.1177/0962280218773520

National Multiple Sclerosis Society. (2022). *Understanding MS: Discover more about Multiple Sclerosis.* National Multiple Sclerosis Society. https://www.nationalmssociety.org/What-is-MS/MS-FAQ-s

NICE. (2022). *Multiple sclerosis in adults: management.*

https://www.nice.org.uk/guidance/ng220/resources/multiple-sclerosis-in-adults-management-pdf-66143828948677

Nunnari, D., De Cola, M. C., D'Aleo, G., Rifici, C., Russo, M., Sessa, E., ... & Marino, S. (2015). Impact of depression, fatigue, and global measure of cortical volume on cognitive impairment in multiple sclerosis. *BioMed Research International*, *2015*, Article 519785. https://doi.org/10.1155/2015/519785

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

Oreja-Guevara, C., Ayuso Blanco, T., Brieva Ruiz, L., Hernández Pérez, M. Á., Meca-Lallana, V., & Ramió-Torrentà, L. (2019). Cognitive dysfunctions and assessments in multiple sclerosis. *Frontiers in Neurology*, *10*, Article 581. https://doi.org/10.3389/fneur.2019.00581

Ozakbas, S., Turkoglu, R., Tamam, Y., Terzi, M., Taskapilioglu, O., Yucesan, C., ... & 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

Pérez, L. P., González, R. S., & Lázaro, E. B. (2015). Treatment of mood disorders in multiple sclerosis. *Current Treatment Options in Neurology*, *17*(1), 1-11. https://doi.org/10.1007/s11940-014-0323-4

Phillips, A., & Morgan, G. (2014). Co-production within health and social care–the implications for Wales?. *Quality in Ageing and Older Adults*, 15(1), 10-20. https://doi.org/10.1108/QAOA-06-2013-0014

Rao, S. M. (1990). *A manual for the brief repeatable battery of neuropsychological tests in multiple sclerosis*. Medical College of Wisconsin.

Rogers, J. M., & Panegyres, P. K. (2007). Cognitive impairment in multiple sclerosis: evidence-based analysis and recommendations. *Journal of Clinical Neuroscience*, *14*(10), 919-927. https://doi.org/10.1016/j.jocn.2007.02.006

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

Samartzis, L., Gavala, E., Zoukos, Y., Aspiotis, A., & Thomaides, T. (2014). Perceived cognitive decline in multiple sclerosis impacts quality of life independently of depression. *Rehabilitation Research and Practice*, *2014*, 128751. https://doi.org/10.1155/2014/128751

Shilyansky, C., Williams, L. M., Gyurak, A., Harris, A., Usherwood, T., & Etkin, A. (2016). Effect of antidepressant treatment on cognitive impairments associated with depression: a randomised longitudinal study. *The Lancet Psychiatry*, *3*(5), 425-435. https://doi.org/10.1016/S2215-0366(16)00012-2

Smith, L., Elwick, H., Mhizha-Murira, J. R., Topcu, G., Bale, C., Evangelou, N., ... & Das Nair, R. (2021). Developing a clinical pathway to identify and manage cognitive problems in Multiple Sclerosis: Qualitative findings from patients, family members, charity volunteers, clinicians and healthcare commissioners. *Multiple Sclerosis and Related Disorders*, *49*, Article 102563. https://doi.org/10.1016/j.msard.2020.102563.

Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison, and validation of Meta-Essentials: a free and simple tool for meta-analysis. *Research Synthesis Methods*, *8*(4), 537-553. https://doi.org/10.1002/jrsm.1260

Thomas, G. A., Riegler, K. E., Bradson, M. L., O'Shea, D. U., & Arnett, P. A. (2022). 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

Tyng, C. M., Amin, H. U., Saad, M. N., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, *8*, 1454. https://doi.org/10.3389/fpsyg.2017.01454

Vanotti, S., & Caceres, F. J. (2017). Cognitive and neuropsychiatric disorders among MS patients from Latin America. *Multiple Sclerosis Journal–Experimental, Translational and Clinical*, *3*(3), 2055217317717508. Article 2055217317717508. https://doi.org/10.1177/2055217317717508

Van Dongen, L., Westerik, B., van der Hiele, K., Visser, L. H., Schoonheim, M. M., Douw, L., ... & Hulst, H. E. (2020). Introducing Multiple Screener: An unsupervised digital screening tool for cognitive deficits in MS. *Multiple Sclerosis and Related Disorders*, *38*, Article 101479. https://doi.org/10.1016/j.msard.2019.101479

Van Geest, Q., Douw, L., van't Klooster, S., Leurs, C. E., Genova, H. M., Wylie, G. R., ... & Hulst, H. E. (2018). Information processing speed in multiple sclerosis: Relevance of default mode network dynamics. *NeuroImage: Clinical*, *19*, 507-515. https://doi.org/10.1016/j.nicl.2018.05.015

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., ... & 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

Vissicchio, N. A., Altaras, C., Parker, A., Schneider, S., Portnoy, J. G., Archetti, R., ... & Foley, F. W. (2019). Relationship between anxiety and cognition in multiple sclerosis: Implications for treatment. *International Journal of MS Care*, *21*(4), 151-156. https://doi.org/10.7224/1537-2073.2018-027

Weber, E., Goverover, Y., & DeLuca, J. (2019). Beyond cognitive dysfunction: Relevance of ecological validity of neuropsychological tests in multiple sclerosis. *Multiple Sclerosis Journal*, *25*(10), 1412-1419. https://doi.org/10.1177/1352458519860318

Williams, L. J., & McGonagle, A. K. (2016). Four research designs and a comprehensive analysis strategy for investigating common method variance with self-report measures using latent variables. *Journal of Business and Psychology*, *31*, 339-359. https://doi.org/10.1007/s10869-015-9422-9

Yigit, P., Acikgoz, A., Mehdiyev, Z., Dayi, A., & Ozakbas, S. (2021). The relationship between cognition, depression, fatigue, and disability in patients with multiple sclerosis. *Irish Journal of Medical Science, 190*(3), 1129-1136. https://doi.org/10.1007/s11845-020-02377-2

Appendices

Appendix A. Mixed Methods Appraisal Tool

4. Quantitative descriptive studies	Methodological quality criteria
Quantitative descriptive studies are "concerned with and designed only to describe the existing distribution of variables without much regard to causal relationships or other hypotheses" (Porta et al., 2014, p. 72). They are used to monitoring the population, planning, and generating hypothesis (Grimes and Schulz, 2002).	4.1. Is the sampling strategy relevant to address the research question? Explanations Sampling strategy refers to the way the sample was selected. There are two main categories of sampling strategies: probability sampling (involve random selection) and non-probability sampling. Depending on the research question, probability sampling might be preferable. Non-probability sampling does not provide equal chance of being selected. To judge this criterion, consider whether the source of sample is relevant to the target population; a clear justification of the sample frame used is provided; or the sampling procedure is adequate.
Common designs include the following single-group studies (this list if not exhaustive): **Incidence or prevalence study without comparison group** In a defined population at one particular time, what is happening in a population, e.g., frequencies of factors (importance of problems), is described (portrayed).	4.2. Is the sample representative of the target population? Explanations There should be a match between respondents and the target population. Indicators of representativeness include: clear description of the target population and of the sample (such as respective sizes and inclusion and exclusion criteria), reasons why certain eligible individuals chose not to participate, and any attempts to achieve a sample of participants that represents the target population.
Survey "Research method by which information is gathered by asking people questions on a specific topic and the data collection procedure is standardized and well defined." (Bennett et al., 2011, p. 3).	4.3. Are the measurements appropriate? Explanations Indicators of appropriate measurements include: the variables are clearly defined and accurately measured, the measurements are justified and appropriate for answering the research question; the measurements reflect what they are supposed to measure; validated and reliability tested measures of the outcome of interest are used, variables are measured using 'gold standard', or questionnaires are pre-tested prior to data collection.
Case series A collection of individuals with similar characteristics are used to describe an outcome. **Case report** An individual or a group with a unique/unusual outcome is described in detail.	4.4. Is the risk of nonresponse bias low? Explanations Nonresponse bias consists of "an error of nonobservation reflecting an unsuccessful attempt to obtain the desired information from an eligible unit." (Federal Committee on Statistical Methodology, 2001, p. 6). To judge this criterion, consider whether the respondents and non-respondents are different on the variable of interest. This information might not always be reported in a paper. Some indicators of low nonresponse bias can be considered such as a low nonresponse rate, reasons for nonresponse (e.g., noncontacts vs. refusals), and statistical compensation for nonresponse (e.g., imputation). The nonresponse bias is might not be pertinent for case series and case report. This criterion could be adapted. For instance, complete data on the cases might be important to consider in these designs.
Key references: Critical Appraisal Skills Programme (2017); Draugalis et al. (2008)	4.5. Is the statistical analysis appropriate to answer the research question? Explanations The statistical analyses used should be clearly stated and justified in order to judge if they are appropriate for the design and research question, and if any problems with data analysis limited the interpretation of the results.

Appendix B. PRISMA Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Pg 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Pg 2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Pg 6 -7
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Pg 7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Pg 7
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Pg 8
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Pg 8
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 8
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Pg 8-9
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Pg 10
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Pg 9
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 9
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Pg 11
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Pg 8
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Pg 8
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Pg 14 - 18

Section and Topic	Item #	Checklist item	Location where item is reported
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Pg 14-18
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Pg 10
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Pg 10
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Pg 11
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Pg 11
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Pg 11
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Pg 11
Study characteristics	17	Cite each included study and present its characteristics.	Pg 11
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Pg 13
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Pg 11
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Pg 13
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Pg 14 - 16
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Pg 14 - 16
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Pg 14-16
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Pg 14 - 16
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Pg 14 - 16
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Pg 18 - 20
	23b	Discuss any limitations of the evidence included in the review.	Pg 20
	23c	Discuss any limitations of the review processes used.	Pg 20
	23d	Discuss implications of the results for practice, policy, and future research.	Pg 20 -22

Section and Topic	Item #	Checklist item	Location where item is reported
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Pg 7
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Pg 7
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Pg 23
Competing interests	26	Declare any competing interests of review authors.	Pg 23
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	n/a

Table 1. Search Strategy

Concept	Search terms	Limiters	Expanders	Search modes
MS	multiple sclerosis or ms	Full text online only	Apply equivalent subjects	Boolean/Phrase
	AND			
Objective cognition	(DE "Cognition" or DE "cognitive processes" or DE "cognitive aging" or DE "cognitive impairment" or DE "Cognitive linguistics") or (DE "cognitive ability" or DE "cognitive assessment") OR DE "cognitive function" or DE "Cognitive Functioning" or DE "cognitive performance" or "neuropsychological impairment" or "neuropsychological assessment" or "objective assessment" or "objective"	Full text online only	Apply equivalent subjects	Boolean/Phrase
	AND			
MSNQ	Multiple Sclerosis Neuropsychological Questionnaire or MSNQ	Full text online only	Also search within the full text of the articles; Apply equivalent subjects	Find all my search terms

Table 2. Characteristics of included studies

Study No.	Study characteristics			Sample characteristics				Measure	
	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)	Objective cognition	Mood and/or QoL*
1	Benedict et al (2003)	Outpatients MS Clinic/ USA	50	M=17, W =33	42.5 (7.2)	15 (2.3)	ns	BNT; JLO; CVLT-II Total Recall Trials 1 /5; CVLT-II Delayed Recall; BVMT-R Total Recall Trials 1 /3; BVMT-R Delayed Recall; TMT, Part B; PASAT; WCST	BDI; CES-D

Study No.	Study characteristics		Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (*n*)	Gender (*n*) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (*n*)	Objective cognition	Mood and/or QoL*
								Perseveration Responses	
2	Benedict et al (2004)	MS Clinic/ USA	85	W = 68 M = 17	42.4 (9.3)	14.8 (2.3)	RR =68 PP = 17	COWAT; JLO; CVLT-II; BVMT-R; PASAT; SDMT; WCST; and/or D-KEFS Sorting Test;	CES-D-10; BDI-FS

Study No.	Study characteristics		Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)	Objective cognition	Mood and/or QoL*
3	Benedict et al (2008)	Outpatients MS Care centre /USA	76	W = 56, M= 20 ,	47.6 (8.4)	14.7 (2.2)	RR=63, SP =11 PP =2	MACFIMS battery: COWAT; JLO; CVLT – II; BVMT-R; PASAT; SDMT; DKEFS Sorting Test.	BDI - FS
4	Carone et al (2005)	Outpatients/ USA	122	W= 88 M= 34	44 (8.8)	14.5 (2.1)	RR = 88; SP =30, PP	COWAT; JLO; CVLT – II; BVMT-R;	BDI

Study No.	Study characteristics		Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)	Objective cognition	Mood and/or QoL*
							= 2 RP=2	PASAT; SDMT; WCST	
5	Charest et al (2020)	MS Clinic /Canada	98	M= 19, W = 79	"Cognitive ly intact" 49.4 (11.8) "cognitive ly	Cognitivel y intact 14.8 (2.3); Cognitivel y impaired 14.3 (3.8)	CIS =7 RR=76 SP =15	MoCA, MACFIMS: COWAT; JLO; SDMT, PASAT; D-KEFS; BVMT-R; CVLT	MSQOL-54*

Study	Study characteristics			Sample characteristics				Measure	
No.	Author (date)	Setting/ Country	Sample size *(n)*	Gender (*n*) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (*n*)	Objective cognition	Mood and/or QoL*
					impaired" 50 (10.4)				
6	Dagenais et al (2013)	Outpatients/ Canada	41	W = 29 M = 12	44.51 (7.43)	High school n=12, college n=8,	RR = 35; SP = 6	MoCA	HADS; BDI - FS

Study No.	Study characteristics		Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)	Objective cognition	Mood and/or QoL*
						university n=21			
7	Davenport et al (2021)	Ireland/ Outpatient	177	W = 124 M =53	42.5 (11.4)	ns	RR = 152	BICAMS: SDMT; CVLT II; BVMT	HADS
8	D'hooghe et al., 2019	Belgium/ Outpatient	231	W= 147 M = 84	52 [confidenc	< 12 years n=20; 12-15 years	RR = 177; PP = 54	SDMT	2 item screening tool

Study No.	Study characteristics		Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)	Objective cognition	Mood and/or QoL*
					e interval 50.8-53.1]	n= 105; > 15 years n= 106			
9	Fenu et al 2021	Outpatients/ Italy	95	W = 68 M = 27	43.65 (11.9)	13 (3.5)	RR = 95	BICAMS: SDMT, CVLT, BVMT.	BDI-II; ZAS

Study No.	Study characteristics		Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)	Objective cognition	Mood and/or QoL*
10	Hughes et al 2019	MS Centre/ USA	99	W = 73; M =24; ns = 2	46.20 (12.96)	ns	RR =76; SP =8; PP =3; R = 1; ns =11	BICAMS: SDMT, CVLT-II, BVMT	PHQ-9

Study No.	Study characteristics			Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)		Objective cognition	Mood and/or QoL*
11	Kim et al., 2017	MS clinic/ USA	100	W = 75 M = 25	46.61 (11.72)	15.51 (2.34)	RR = 88 PP =5 SP =7		MACFIMS: COWAT; JLO Test; CVLT-II BVMT- R; PASAT; SDMT; D-KEFS	BDI-II
12	O'Brien et al., 2007	Community / USA	48	W=38 M= 10	45.1 (9.1)	14.7 (2.1)	RR =33 PP =5 SP=10		WRAT; WAIS-R DSS; SDMT; Logical Memory I and II of	BDI

Study	Study characteristics			Sample characteristics					Measure	
No.	Author (date)	Setting/ Country	Sample size *(n)*	Gender (*n*) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (*n*)		Objective cognition	Mood and/or QoL*
									the WMS-III, PASAT WAIS-R Digit Span Backwards; mSRT, Stroop Color–Word test, WCST COWAT; JOLO	

Study	Study characteristics			Sample characteristics					Measure	
No.	Author (date)	Setting/ Country	Sample size (*n*)	Gender (*n*) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (*n*)		Objective cognition	Mood and/or QoL*
13	Ozakbas et al., 2018	Multicentre MS clinics	487	W = 334; M =153	35.3 (9.39)	11.29 (4.04)	RR = 487		BRB-N; SRT; SPART;SDM; PASAT - 3, WLG; Quality of life Sort form - 36	BDI; QoL SF – 36*
14	Rosti et al., 2014	ns/Finland	196	W = 142 M =54	RR = 43.9 (8.8), PP = 49.9 (8.7)	ns	RR=13 8 , PP=58		BRB-NT;	BDI-II; WHOQO L-BREF *

Study	Study characteristics		Sample characteristics					Measure	
No.	Author (date)	Setting/ Country	Sample size (n)	Gender (n) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (n)	Objective cognition	Mood and/or QoL*
								BSRT; SPART; SDMT; PASAT 2-3); COWAT	
15	Thomas et al., 2022	Ns/USA	87	M = 22 W= 65	51.83 (10.57)	14.84 (1.96)	RR = 55; SP = 26; PP = 6	MACFIM: DS; Oral SDMT, COWAT, Animal Naming, PASAT, DSST; D-	BDI-FS

Study No.	Study characteristics			Sample characteristics					Measure	
	Author (date)	Setting/ Country	Sample size (*n*)	Gender (*n*) Men (M) Women (W) Other O	Age Mean ± SD (years)	Education Mean ± SD (years)	MS Course (*n*)		Objective cognition	Mood and/or QoL*
									KEFS , BVMT-R, CVLT-II	

Note: * indicates QoL measure

ns – not stated; SD- standard deviation; R = Progressive relapsing; RR – Relapsing remitting; SP – Secondary progressive; PP – Primary progressive; CIS – Clinically Isolated Syndrome

BNT – Boston Naming Test; JLO - Judgement of Line Orientation; CVLT-II - California Verbal Learning Test - Second Edition; BVMT- R - Brief Visuospatial Memory Test-Revised; TMT – Trial Making Test; PASAT - Paced Auditory Serial Addition Test; WCST - Wisconsin Card Sorting Test; CES-D - Centre for Epidemiologic Studies Depression Scale; COWAT - Controlled Oral Word Association Test; SDMT – Symbol Digit Modalities Test; D-KEFS - Delis-Kaplan Executive Function System; D-KEFS ST FSD - Delis-Kaplan Executive Function System Free

Sorting Description; D-KEFS ST FSC - Delis-Kaplan Executive Function System Free Sorting Categorization; D-KEFS ST SR - Delis-Kaplan Executive Function System Sort Recognition; BSRT- Buschke Selective Reminding Test; DSST - Digit Symbol Substitution Test

MACFIMS - Minimal Assessment of Cognitive Function in Multiple Sclerosis; MoCA - Montreal Cognitive Assessment; WAIS DSSS – Digit Span and Symbol Search; DS – Digit Span; WRAT - Wide Range Achievement Test; MSDRT - modified Selective Reminding Test; EPFT - Executive Functioning Performance Test; SPART - 10/36 Spatial Recall Test; WLG – Word List Generation (WGL); QoL SF – 36 – Quality of Life Short Form – 36;

BDI- Beck Depression Inventory; BDI – FS – Beck Depression Inventory Fast Screen; HADS – Hospital Anxiety and Depression Scale; ZAS - Zung Anxiety Scale; (WHOQOL-BREF) World Health Organisation Quality of Life; MSQOL-54 - Multiple Sclerosis Quality of Life

Table 3 – Quality appraisal of included studies

Study number	Is the sampling strategy relevant to address the research question?	Is the sample representative of the target population?	Are the measurements appropriate?	Is the risk of non-response bias low?	Is the statistical analysis appropriate to answer the research question?
1	Yes	Can't tell	Yes	Can't tell	Yes
2	Can't tell	Yes	Yes	Yes	Yes
3	Yes	Yes	Yes	Can't tell	Yes
4	Yes	No	Yes	Can't tell	Yes
5	Can't tell	Yes	No	Can't tell	Yes
6	Yes	No	Yes	Yes	Yes
7	Yes	Yes	Yes	Can't tell	Yes
8	Yes	Yes	No	No	Yes
9	Can't tell	No	Yes	Can't tell	Yes
10	Can't tell	Yes	Yes	Can't tell	Yes
11	Yes	Yes	Yes	No	Yes
12	Yes	Yes	No	Can't tell	Yes
13	Yes	Yes	Yes	Can't tell	Yes
14	Can't tell	Yes	Yes	Can't tell	Yes
15	Can't tell	Yes	Yes	Can't tell	Yes

Table 4 – Associations between MSNQ and objective measure of cognition

Study Number	Sample size	Correlation (r.)	Confidence Interval	Weight %

2	85	-0.49	-0.64 − - 0.31	7.32
3	76	-0.33	-0.52 − -0.11	7.00
4	122	-0.21	-0.38 − -0.03	8.26
6	41	-0.15	-0.45 − 0.17	5.14
7	177	-0.19	-0.33 − - 0.04	9.07
8	231	-0.09	-0.22 − 0.04	9.56
9	95	-0.32	-0.49 − -0.12	7.62
11	100	-0.22	-0.40 − -0.03	7.76
12	48	-0.01	-0.29 − 0.30	5.63
13	487	-0.17	-0.25 − -0.08	10.51
14 RR	138	0.18	0.01 − 0.34	8.55
14 SP	58	0.19	-0.08 − 0.43	6.20
15	87	-0.26	-0.45 − -0.05	7.38

Table 5 – Associations between the MSNQ and measures of mood

Study Number	Sample size	Correlation (r.)	Confidence Interval		Weight %
2	85	0.61	0.45 − 0.73	8.58
3	76	0.58	0.40 − 0.71	8.27
6	41	-0.03	-0.34 − 0.29	6.32
7	177	0.42	0.28 − 0.53	10.26
9	95	0.23	0.02 − 0.41	8.88
10	99	0.58	0.43 − 0.70	8.99
11	100	0.65	0.52 − 0.75	9.01

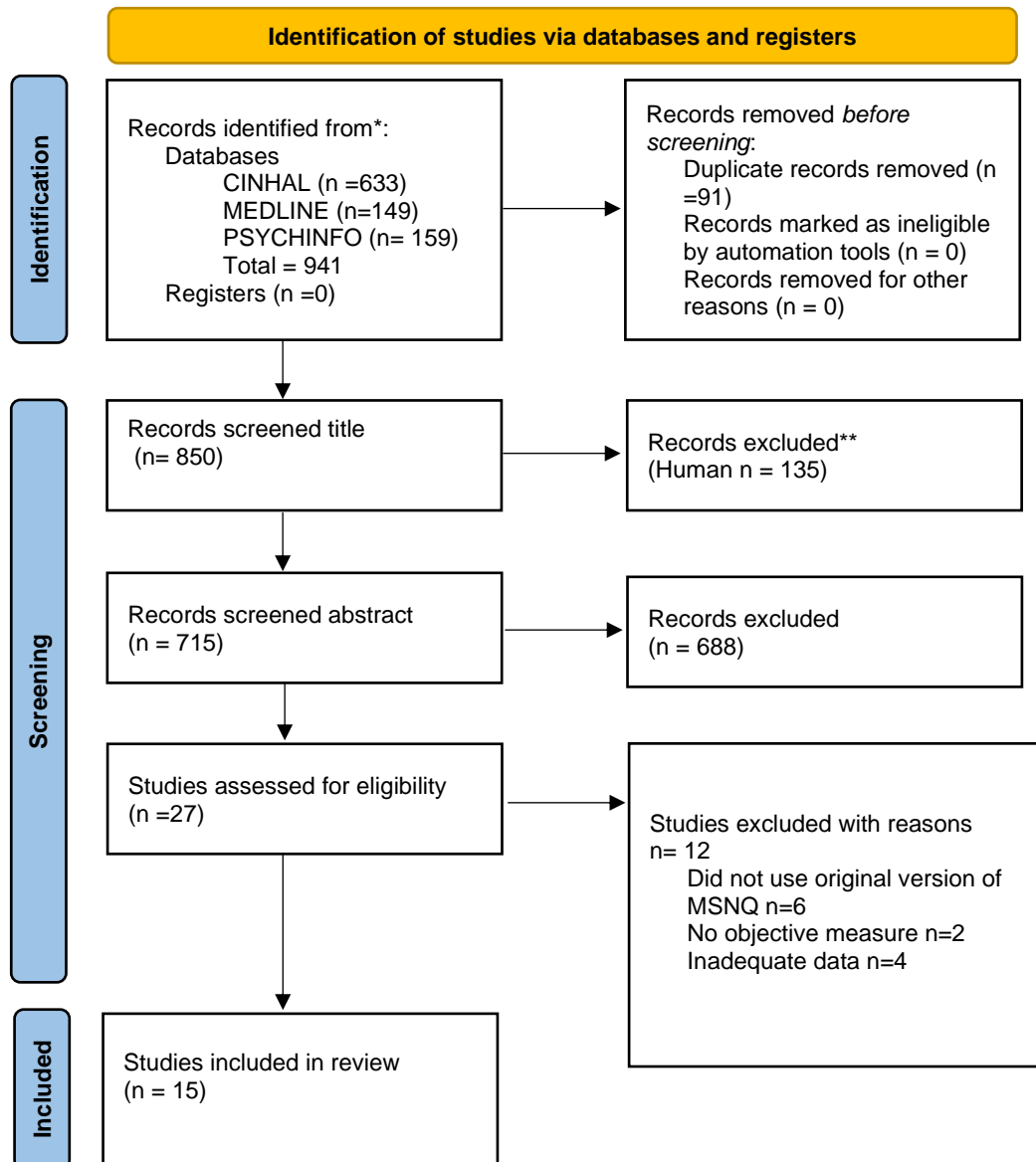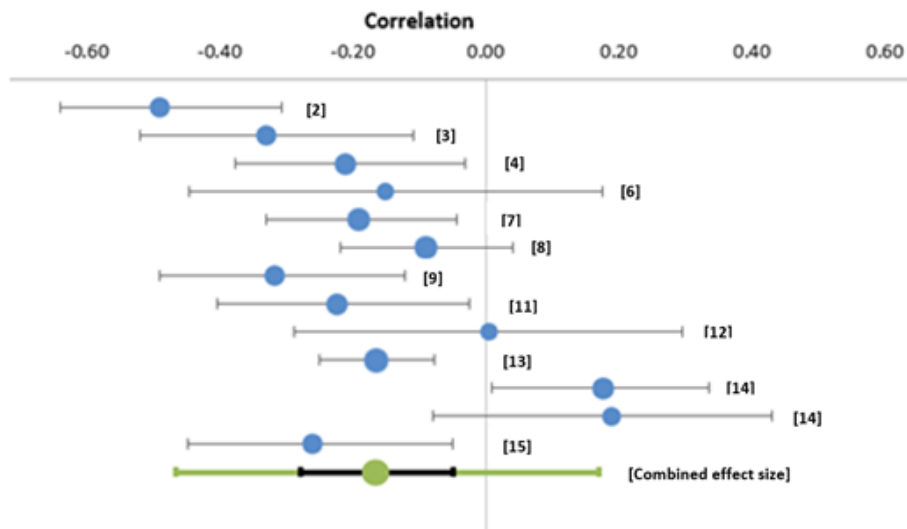| 12 | 48 | 0.42 | 0.15 | − | 0.63 | 6.84 |
14 RR	138	0.41	0.26	−	0.54	9.77
14 SP	58	0.54	0.32	−	0.70	7.45
15	87	0.39	0.19	−	0.56	8.65

Figures

Figure 1. PRISMA diagram

Figure 2

Forest plots showing associations between MSNQ and objective measure of cognition.

Note: Forest plots shows correlations between the MSNQ and objective measures of cognition. Numbers in parentheses denote study numbers. The green bar that can be seen in the figure is a prediction interval.   The prediction interval demonstrates the range of the predicted effect sizes in future studies or studies not included in this meta- analysis.

Figure 3

Funnel plots depicting publication bias in studies examining associations between the MSNQ and objective measures of cognition.
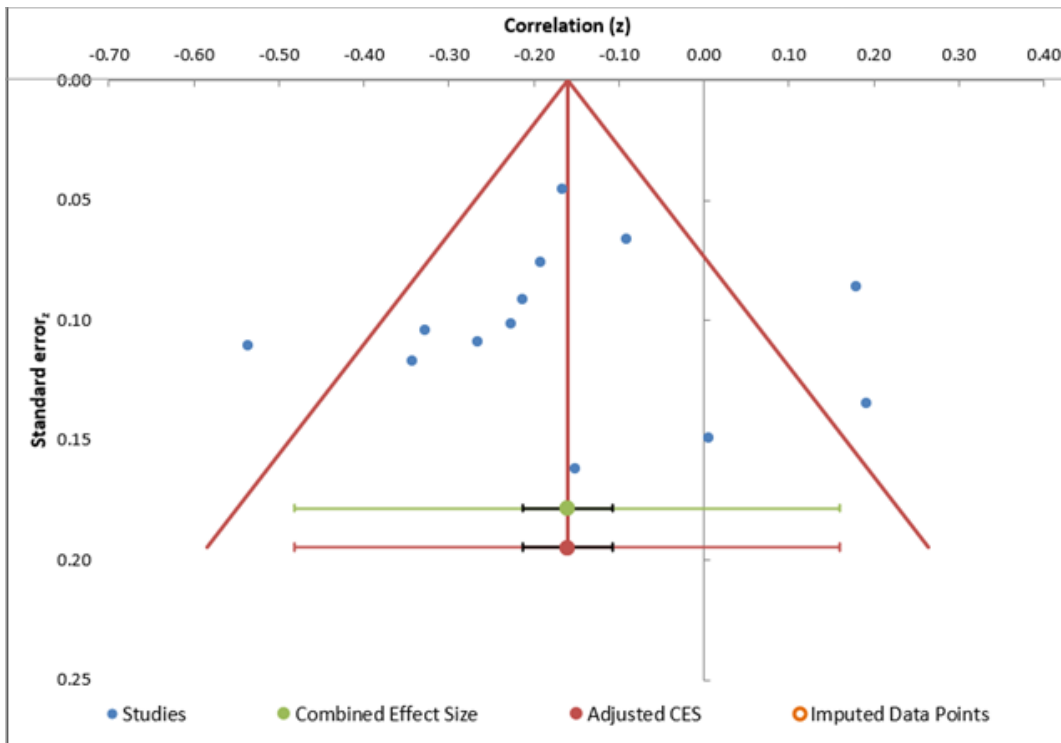
Figure 4

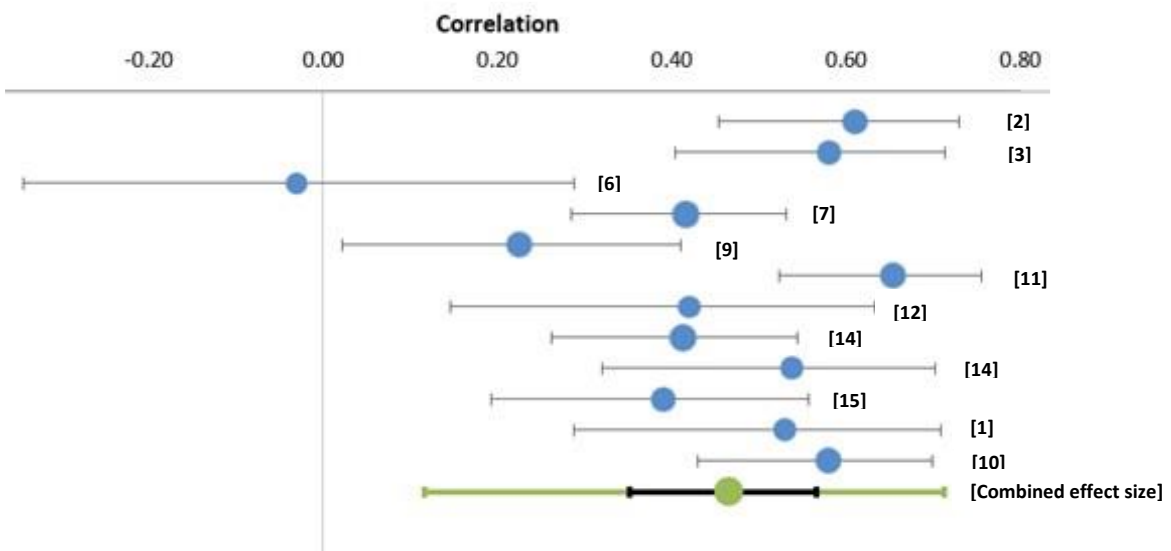Forest plots showing associations between MSNQ and measures of mood.

Figure 5

Funnel plot depicting publication bias in studies examining the association between the
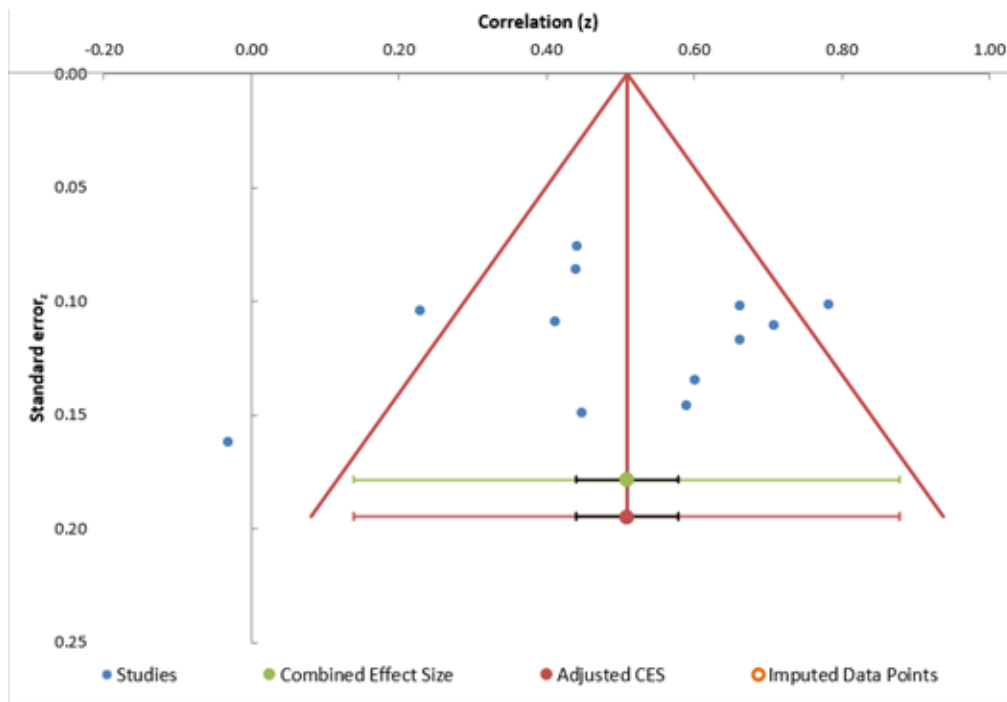
MSNQ and measures of mood.

Figure captions

Figure 1: PRISMA Digaram.

Figure 2: Forest plots showing associations between MSNQ and objective measure of cognition.

Figure 3: Funnel plots depicting publication bias in studies examining associations between the MSNQ and objective measures of cognition.

Figure 4: Forest plots showing associations between MSNQ and measures of mood.

Figure 5: Funnel plot depicting publication bias in studies examining the association between the MSNQ and measures of mood.