Cochrane Library

Cochrane Database of Systematic Reviews

Memory rehabilitation for people with multiple sclerosis (Review)

Taylor LA, Mhizha-Murira JR, Smith L, Potter KJ, Wong D, Evangelou N, Lincoln NB, das Nair R

Taylor LA, Mhizha-Murira JR, Smith L, Potter K-J, Wong D, Evangelou N, Lincoln NB, das Nair R.
Memory rehabilitation for people with multiple sclerosis.
Cochrane Database of Systematic Reviews 2021, Issue 10. Art. No.: CD008754.
DOI: 10.1002/14651858.CD008754.pub4.

www.cochranelibrary.com

TABLE OF CONTENTS

[Intervention Review]

Memory rehabilitation for people with multiple sclerosis

Lauren A Taylor[1], Jacqueline R Mhizha-Murira[2], Laura Smith[3], Kristy-Jane Potter[4], Dana Wong[5], Nikos Evangelou[6], Nadina B Lincoln[7], Roshan das Nair[4]

[1]Division of Psychiatry and Applied Psychology, School of Medicine, University of Nottingham, Nottingham, UK. [2]Division of Primary Care, University of Nottingham, Nottingham, UK. [3]Division of Psychiatry and Applied Psychology, University of Nottingham, Nottingham, UK. [4]Division of Psychiatry & Applied Psychology, University of Nottingham, Nottingham, UK. [5]School of Psychology and Public Health, La Trobe University, Melbourne, Australia. [6]Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham, UK. [7]Division of Rehabilitation and Ageing, University of Nottingham, Nottingham, UK

Contact address: Roshan das Nair, roshan.dasnair@nottingham.ac.uk.

Editorial group: Cochrane Multiple Sclerosis and Rare Diseases of the CNS Group.
Publication status and date: New search for studies and content updated (conclusions changed), published in Issue 10, 2021.

Citation: Taylor LA, Mhizha-Murira JR, Smith L, Potter K-J, Wong D, Evangelou N, Lincoln NB, das Nair R. Memory rehabilitation for people with multiple sclerosis. *Cochrane Database of Systematic Reviews* 2021, Issue 10. Art. No.: CD008754. DOI: 10.1002/14651858.CD008754.pub4.

ABSTRACT

Background

Problems with cognition, particularly memory, are common in people with multiple sclerosis (MS) and can affect their ability to complete daily activities and can negatively affect quality of life. Over the last few years, there has been considerable growth in the number of randomised controlled trials (RCTs) of memory rehabilitation in MS. To guide clinicians and researchers, this review provides an overview of the effectiveness of memory rehabilitation for people with MS.

Objectives

To determine whether people with MS who received memory rehabilitation compared to those who received no treatment, or an active control showed better immediate, intermediate, or longer-term outcomes in their:

1. memory functions,

2. other cognitive abilities, and

3. functional abilities, in terms of activities of daily living, mood, and quality of life.

Search methods

We searched CENTRAL which includes Clinicaltrials.gov, World Health Organization (The Whoqol) International Clinical Trials Registry Portal, Embase and PubMed (MEDLINE), and the following electronic databases (6 September 2020): CINAHL, LILACS, the NIHR Clinical Research Network Portfolio database, The Allied and Complementary Medicine Database, PsycINFO, and CAB Abstracts.

Selection criteria

We selected RCTs or quasi-RCTs of memory rehabilitation or cognitive rehabilitation for people with MS in which a memory rehabilitation treatment group was compared with a control group. Selection was conducted independently first and then confirmed through group discussion. We excluded studies that included participants whose memory deficits were the result of conditions other than MS, unless we could identify a subgroup of participants with MS with separate results.

Data collection and analysis

Eight review authors were involved in this update in terms of study selection, quality assessment, data extraction and manuscript review. We contacted investigators of primary studies for further information where required. We conducted data analysis and synthesis in accordance with Cochrane methods. We performed a 'best evidence' synthesis based on the methodological quality of the primary studies included. Outcomes were considered separately for ‘immediate’ (within the first month after completion of intervention), ‘intermediate’ (one to six months), and ‘longer-term’ (more than six months) time points.

Main results

We added 29 studies during this update, bringing the total to 44 studies, involving 2714 participants. The interventions involved various memory retraining techniques, such as computerised programmes and training on using internal and external memory aids. Control groups varied in format from assessment-only groups, discussion and games, non-specific cognitive retraining, and attention or visuospatial training. The risk of bias amongst the included studies was generally low, but we found eight studies to have high risk of bias related to certain aspects of their methodology.

In this abstract, we are only reporting outcomes at the intermediate timepoint (i.e., between one and six months). We found a slight difference between groups for subjective memory (SMD 0.23, 95% CI 0.11 to 0.35; 11 studies; 1045 participants; high-quality evidence) and quality of life (SMD 0.30, 95% CI 0.02 to 0.58; 6 studies; 683 participants; high-quality evidence) favoring the memory rehabilitation group. There was a small difference between groups for verbal memory (SMD 0.25, 95% CI 0.11 to 0.40; 6 studies; 753 participants; low-quality evidence) and information processing (SMD 0.27, 95% CI 0.00 to 0.54; 8 studies; 933 participants; low-quality evidence), favoring the memory rehabilitation group.

We found little to no difference between groups for visual memory (SMD 0.20, 95% CI -0.11 to 0.50; 6 studies; 751 participants; moderate-quality evidence), working memory (SMD 0.16, 95% CI -0.09 to 0.40; 8 studies; 821 participants; moderate-quality evidence), or activities of daily living (SMD 0.06, 95% CI -0.36 to 0.24; 4 studies; 400 participants; high-quality evidence).

Authors' conclusions

There is evidence to support the effectiveness of memory rehabilitation on some outcomes assessed in this review at intermediate follow-up. The evidence suggests that memory rehabilitation results in between-group differences favoring the memory rehabilitation group at the intermediate time point for subjective memory, verbal memory, information processing, and quality of life outcomes, suggesting that memory rehabilitation is beneficial and meaningful to people with MS. There are differential effects of memory rehabilitation based on the quality of the trials, with studies of high risk of bias inflating (positive) outcomes. Further robust, large-scale, multi-centre RCTs, with better quality reporting, using ecologically valid outcome assessments (including health economic outcomes) assessed at longer-term time points are still needed to be certain about the effectiveness of memory rehabilitation in people with MS.

PLAIN LANGUAGE SUMMARY

Memory rehabilitation in multiple sclerosis

Review question

Do people with multiple sclerosis (MS) who received memory rehabilitation compared to those who received no treatment, or a placebo show better immediate-, intermediate-, or longer-term outcomes in their:

1. memory functions,

2. other cognitive abilities, and

3. functional abilities, in terms of activities of daily living, mood, and quality of life?

Background

People with multiple sclerosis (MS) often struggle with memory problems, which can lead to difficulties in everyday life. Memory rehabilitation is offered to help people cope with memory problems, enhance their ability to perform everyday activities, and to increase independence by reducing forgetting. Such rehabilitation can involve the use of specific techniques and strategies to change the way a person tries to remember, store, or retrieve memories. However, it is unclear whether memory rehabilitation is effective in reducing forgetting or improving performance of daily activities. Historically, there were few good-quality studies that investigated the effectiveness of memory rehabilitation in people with MS, but lately there have been some larger studies. Therefore, we wanted to know whether the evidence of the effectiveness of memory rehabilitation has changed since the previous version of our review.

Study characteristics

This review included 44 studies with 2714 participants who received various types of memory retraining techniques, some using restorative techniques (e.g. computerised programmes) and others using compensatory approaches (e.g. memory aids such as diaries or calendars).

Key results and quality of the evidence

Substantial progress has been made since the last update of this review, and the results from this review suggest that there is now evidence to support the use of memory rehabilitation in people with MS. Participants who had memory rehabilitation reported better memory functioning and quality of life compared to those who did not receive memory rehabilitation, and these differences were found immediately after the intervention was completed and for some time thereafter. However, those who received memory rehabilitation did not appear to improve in terms of their anxiety symptoms or daily activities. This update has added large, good-quality studies on which to base our findings, so the evidence to support the effectiveness of memory rehabilitation is stronger than in the previous update. However, we still need large, good quality studies that examine the longer-term impact of memory rehabilitation and studies that evaluate the cost-effectiveness of memory rehabilitation in people with MS.

SUMMARY OF FINDINGS

Summary of findings 1. Memory rehabilitation for people with multiple sclerosis

Memory rehabilitation for people with multiple sclerosis

Patient or population: people with multiple sclerosis
Settings: clinic and home-based
Intervention: memory rehabilitation

Comparison: active control or no treatment

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Memory rehabilitation				
Subjective memory measures - intermediate EMQ, MSNQ, CFQ, MFQ[a] Follow-up: 1 to 6 months	-	The mean subjective memory measures - immediate in the intervention groups was **0.23 standard deviations higher** (0.11 to 0.35 higher)	-	1045 (11 studies)	⊕⊕⊕⊕ **high**	Immediate follow-up: SMD 0.32 (0.05 to 0.58) Longer-term follow-up: SMD 0.16 (0.02 to 0.30)
Objective verbal memory measures - intermediate CVLT, AVLT, HVLT, VLT, SRT, MUSIC[a] Follow-up: 1 to 6 months	-	The mean objective verbal memory measures - intermediate in the intervention groups was **0.25 standard deviations higher** (0.11 to 0.4 higher)	-	753 (6 studies)	⊕⊕⊝⊝ **low** [b,c]	Immediate follow-up: SMD 0.40 (0.22 to 0.58) Longer-term follow-up: SMD 0.13 (-0.03 to 0.29)

Objective visual memory measures - intermediate BVMT-R, SPART, CMT, ROCF Follow-up: 1 to 6 months	-	The mean objective visual memory measures - intermediate in the intervention groups was **0.2 standard deviations higher** (0.11 lower to 0.5 higher)	-	751 (6 studies)	⊕⊕⊕⊖ **moderate** [e]	Immediate follow-up: SMD 0.42 (0.25 to 0.60) Longer-term follow-up: SMD 0.12 (-0.13 to 0.37)
Objective working memory measures - intermediate PASAT, WAIS Follow-up: 1 to 6 months	-	The mean objective working memory measures - intermediate in the intervention groups was **0.16 standard deviations higher** (0.09 lower to 0.40 higher)	-	821 (8 studies)	⊕⊕⊕⊖ **moderate** [f]	Immediate follow-up: SMD 0.45 (0.18 to 0.72) Longer-term follow-up: SMD 0.04 (-0.11 to 0.2)
Informating processing - intermediate SDMT Follow-up: 1 to 6 months	-	The mean information processing measures - intermediate in the intervention groups was **0.27 standard deviations higher** (0.00 to 0.54 higher)	-	933 (8 studies)	⊕⊕⊖⊖ **low** [g,h]	Immediate follow-up: SMD 0.51 (0.19 to 0.82) Longer-term follow-up: SMD 0.21 (-0.03 to 0.45)

Quality of life - intermediate MSIS, MSQOL, SF-36, SF-12, SWLS, EQ-5D-5L[a] Follow-up: 1 to 6 months	-	The mean quality of life measures - intermediate in the intervention groups was **0.30 standard deviations higher** (0.02 to 0.58 higher)	-	683 (6 studies)	⊕⊕⊕⊕ **high**	Immediate follow-up: SMD 0.42 (0.15 to 0.68) Longer-term follow-up: SMD 0.17 (0.02 to 0.32)
Acitivities of daily living - intermediate EADL[a] Follow-up: 1 to 6 months	-	The mean activities of daily living measures - intermediate in the intervention groups was **0.06 standard deviations lower** (0.36 lower to 0.24 higher)	-	400 (4 studies)	⊕⊕⊕⊕ **high**	Immediate follow-up: SMD 0.02 (-0.26 to 0.29) Longer-term follow-up: SMD -0.11 (-0.49 to 0.27)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; SMD: standardised mean difference

GRADE Working Group grades of evidence
High quality ⊕⊕⊕⊕**:** further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality ⊕⊕⊕⊖**:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality ⊕⊕⊖⊖**:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality ⊕⊖⊖⊖**:** we are very uncertain about the estimate.

Please note: As per Cochrane guidelines, we only report seven outcomes here. Details of our other outcomes can be found in Table 1.

[a] **CMT:** Contextual Memory Text; **EAQ:** Emotional awareness questionnaire;**EMQ:** Everyday Memory Questionnaire;**HADS:** Hospital Anxiety and Depression Scale; **STAI:** State Trait Anxiety Inventory; **MSNQ:** Multiple Sclerosis Neuropsychological Screening Questionnaire;**MFQ:** Memory Functioning Questionnaire;**RBMT:** Rivermead Behavioural Memory Test;**CVLT:** California Verbal Learning Test; **AVLT:** Auditory Verbal Learning Test; **HVLT:** Hopkins Verbal Learning Test; **VLT:** Verbal Learning Test; LNNB: Luria-Nebraska Neuropsychological Battery; **BRBNT:** Brief Repeatable Battery of Neuropsychological Tests; **GHQ:** General Health Questionnaire; **BDI:** Beck Depression Inventory; **BDI-FS:** Beck Depression Inventory-Fast Screen; **EADL:** Extended Activities of Daily Living; **MSIS:** Multiple Sclerosis Impact Scale; **FAMS:** Functional Assessment of Multiple Sclerosis;**MSQOL:** Multiple Sclerosis Quality of Life; **PASAT:** Paced auditory serial addition test; **SF-36:** 36-Item Short Form Health Survey; **SF-12:** 12-Item Short Form Health Survey.

[b] 1 of 10 studies had possible risk of bias related to random sequence generation, and in 2 of the 10 it was unclear. Allocation concealment was possible in 1 study, and unclear in 3 of the 10 studies. Blinding was a potential source of bias in 2 studies, and unclear in 1 of the 10 studies. Incomplete outcome data may have been biased in 1 study, an unclear in 3 of the 10 studies. Selective reporting may have been biased in 1 study.
Downgraded by 1 due to 95% confidence intervals including no effect, and the upper or lower confidence intervals limit crosses an effect size of 0.5 in either direction.

[b]All or nearly all of the studies used a list-learning task as an objective measure of verbal memory, which has poor ecological validity.

[c]2 of the 6 studies showed unclear risk of bias relating to random sequence generation. 1 study had unclear potential risk of allocation concealment bias. 4 studies had potential risk of bias related to blinding. 3 studies had unclear risk of bias due to incomplete outcome data. 1 study had unclear risk of other bias.

[e]2 of 6 studies showed unclear potential risk of bias related to random sequence generation. 1 study showed unclear potential risk of bias related to allocation concealment. 4 of 6 studies showed potential risk of bias related to blinding. 3 of 6 studies showed unclear risk of bias related to incomplete outcome data.

[f]5 of 12 studies showed unclear potential risk of bias related to random sequence generation. 6 of 12 studies showed unclear risk of bias related to allocation concealment. 7 of 12 studies showed possible risk of bias related to blinding procedures. 1 study showed potential risk of bias related to incomplete data, and 3 of 12 studies were unclear risk of bias. 1 study had potential risk of bias related to selective reporting.

[g]3 of 8 studies showed unclear risk of bias related to random sequence generation. 1 study showed potential risk of bias related to allocation concealment, 2 of 8 studies showed unclear risk of bias. 4 of 8 studies showed potential risk of bias related to blinding procedures, 1 study showed unclear risk of bias. 3 of 8 studies showed unclear risk of bias related to incomplete data.

[h]Inconsistency with results, statistical heterogeneity>50%

BACKGROUND

Description of the condition

Multiple sclerosis (MS) is an inflammatory disease of the central nervous system that can cause physical and cognitive disturbances. The prevalence of these cognitive problems, which include dysfunctions in memory, attention, speed of information processing, and executive functions, varies up to 70% (Julian 2011). Rao 1993 reported that impaired memory functions were evident in 40% to 60% of people with MS. Impairments in cognitive functions are also related to low mood (Chiaravalloti 2008; Gilchrist 1994), and have the potential to hamper functions related to activities of daily living (ADL) (Kalmar 2008; Langdon 1996).

Description of the intervention

Cognitive rehabilitation is a specialised facet of neuropsychological rehabilitation that assists in the development of functional independence and adjustment of individuals with cognitive problems through targeted intervention or focused stimulation (Robertson 1993). Robertson 2008 defined cognitive rehabilitation as a "structured, planned experience derived from an understanding of brain function which ameliorates dysfunctional cognitive and brain processes caused by disease or injury and improves everyday life function" p565. Memory rehabilitation is a major component of the management of people with memory problems and is either implemented as part of a cognitive rehabilitation programme or as a stand-alone intervention, depending on the needs and neuropsychological profile of the patient, or clinical services available.

How the intervention might work

There is uncertainty about the precise mechanisms by which memory rehabilitation interventions work. However, it is widely believed that they provide people with the knowledge of, and information about their memory problems, by teaching them to use internal and external memory aids, different strategies to pay attention, and alternative ways of encoding, storing, and retrieving information. Targeted, repeated stimulation of certain brain areas using drill and practice cognitive exercises are thought to trigger the activation of neural networks. For group-based interventions, the therapeutic effects of being with others with similar problems may also help (Carr 2014; das Nair 2013; Klein 2019). Some of these behavioural strategies (referred to as 'restitution' or 'compensation') are believed to map onto the neural networks engaged in performing memory functions.

Why it is important to do this review

Studies have examined the effectiveness of memory rehabilitation using different methods. Single-case and small-group studies have reported positive results of memory rehabilitation, but the results obtained from randomised controlled trials (RCTs) and some systematic reviews have been less positive and reported inconclusive evidence. Some reviews (for example Cicerone 2005; Cicerone 2011; Cicerone 2019) have concluded that there is compelling evidence for memory strategy training with participants with mild memory problems, that the use of external memory aids may be beneficial for people with moderate to severe memory problems, and that errorless learning may be effective for those with severe memory impairments (albeit with limited generalisability to new tasks or overall memory problems). Cicerone 2019 also suggests that group-based interventions may be considered as part of a comprehensive neuropsychological rehabilitation of memory deficits. However, these reviews focused mainly on people with traumatic brain injury. Cochrane Reviews by Majid 2000 and das Nair 2016a found insufficient evidence to support or refute the effectiveness of memory rehabilitation following stroke. Some reviews have focused on generic psychological interventions for people with MS (Thomas 2006), or neuropsychological interventions for people with MS (Rosti-Otajärvi 2014), however these were not specific to memory rehabilitation. The Thomas 2006 review did not consider grey literature and was unable to draw any "definite conclusions". The Rosti-Otajärvi 2014 review focused on neuropsychological rehabilitation across multiple cognitive domains, as well as associated health-related factors and emotional well-being. The Goverover 2018b review was similar in that it focused on cognitive rehabilitation in six cognitive domains: attention, learning and memory, processing speed and working memory, executive functioning, metacognition, or nonspecific/combined. The current systematic review is focused solely on the effectiveness of memory rehabilitation for people with MS; databases were searched that were not searched as part of the Rosti-Otajärvi 2014 or Goverover 2018b reviews, and studies are included that were not in these reviews. This is an update of the Cochrane Review 'Memory rehabilitation for people with multiple sclerosis' (first published in the Cochrane Library 2012, Issue 3; das Nair 2012).

OBJECTIVES

To determine whether people with multiple sclerosis (MS) who received memory rehabilitation compared to those who received no treatment, or an active control, showed better immediate, intermediate, or longer-term outcomes in their:

1. memory functions,
2. other cognitive abilities, and
3. functional abilities, in terms of activities of daily living, mood, and quality of life.

METHODS

Criteria for considering studies for this review

Types of studies

For inclusion in the review, we sought randomised and quasi-randomised controlled trials, as defined by the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019), and the pre-cross-over component of randomised cross-over trials with people with MS, in which a memory treatment was compared to a control. Where papers were based on the same sample, or subset of a larger sample, we included only the study with the full sample to avoid double counting. If a study was available through both grey literature (for example conference abstract) and a peer-reviewed publication, we used the peer-reviewed publication.

Types of participants

Trials included in this review were limited to those with people with MS (including relapsing remitting, secondary progressive, and primary progressive MS). We excluded trials with participants whose memory deficits were the result of traumatic brain injury, brain tumour, stroke, epilepsy, or any other neurological condition, unless we could define a subgroup of people with MS of at

least 75% for which there were separate data. Included studies based a diagnosis of MS on well-established diagnostic criteria, for example Paty 1988 and Poser 1983 (and revised versions of the McDonald criteria (Polman 2005; Polman 2011; Thompson 2017). We did not define the type of memory deficits participants needed to have in advance, because we assumed that those people with MS who were given treatment for impaired memory had memory deficits. We placed no restrictions on the type of memory deficits participants reported .

Types of interventions

We included trials in which there was a comparison between a treatment group that received memory rehabilitation strategies, and a control group that received either a comparable standard of treatment (active control) or no memory intervention. We considered rehabilitation to take place over more than a single session; therefore, we did not consider laboratory-based experiments (such as single-session list-recall or mnemonic strategy training) as rehabilitation. Control groups needed to have people with MS, or a subgroup of people with MS amongst those with other diagnoses, for whom separate data were available. We considered memory rehabilitation to be any attempt to modify memory function by means of drill-and-practice, or by the use of internal and/or external memory aids, or by teaching people with MS strategies to cope with their memory problems. We did not include pharmacological studies.

Types of outcome measures

We included trials in which the intervention group either received memory rehabilitation or comprehensive cognitive rehabilitation with a memory component. We considered all trials that met the listed inclusion criteria and did not discriminate based on the type of memory outcomes or other cognitive outcomes they used. We considered memory outcomes to be any questionnaire or test that measures general memory or a specific domain such as verbal memory. The nine outcomes listed below were decided before the analysis was conducted to avoid bias.

Primary outcomes

Primary outcomes were measures of the extent of memory problems in everyday life. There are several ways in which this is assessed in clinical practice and research, but we only included measures that directly assessed this construct. Where multiple tests were used to assess the same construct, we followed a hierarchy that we developed prior to data analysis. We included the following commonly used tests.

1. For subjective reports of memory: we considered Everyday Memory Questionnaire (EMQ) (Sunderland 1983), over the Cognitive Failures Questionnaire (Broadbent 1982), over the Subjective Memory Questionnaire (Davis 1995), over the Memory Assessment Clinics Questionnaire (Crook 1992). If more than one questionnaire was used, we used the following hierarchy: memory problems in daily life, over general forgetting, over domain-specific questions. If a questionnaire was used that was not in this hierarchy, we arrived at a consensus through discussion prior to data extraction to avoid bias.
2. For objective verbal measure of memory: we considered California Verbal Learning Test (CVLT-II) (Delis 2000) over Selective Reminding Test (SRT) (Buschke 1973), over Doors and People Test (Baddeley 1994). For neuropsychological test batteries, we used verbal domain-specific scores over composite scores.
3. For objective visual measures of memory: we considered Brief Visuospatial Memory Test – Revised (BVMT-R) (Benedict 1996), 10/36 Spatial Recall Test (SPART) (Rao 1990), Contextual Memory Text (CMT) (Toglia 2004), Rey-Osterrieth Complex Figure text (ROCF) (Rey 1941 and Osterrieth 1944). For neuropsychological test batteries, we used visual domain-specific scores over composite battery scores.
4. For objective working measures of memory: we considered Rivermead Behavioural Memory Test (RBMT) (Wilson 1985 or newer versions of this test), over Wechsler Memory Scale (WMS) (Wechsler 1997 or newer versions of this test), over Cambridge Test of Prospective Memory (Wilson 2005), over Doors and People Test (Baddeley 1994).
5. For information processing measures: we considered the Symbol Digit Modalities Test (SDMT) (Smith 1973) over other measures, because we were aware that this is one of the most frequently used tests of information processing in MS research (Benedict 2017).

Where studies included more than one test for each outcome group, we used a hierarchy based on the tests' degree of sensitivity to assess everyday memory problems and the tests' ecological validity. If we were unsure about which outcome measure to consider in the analysis, we arrived at a consensus following a discussion with review authors which measure to consider as the primary outcome, *before* the statistical analyses were conducted to minimise bias.

Secondary outcomes

1. Mood - depression, such as the General Health Questionnaire (GHQ) (Goldberg 1988), Hospital Anxiety and Depression Scale (HADS) (Zigmond 1983); Beck Depression Inventory-Fast Screen (Beck 2003). General mood outcomes such as the GHQ were included in both the depression and anxiety scales of the mood outcome.
2. Mood - anxiety, such as the General Health Questionnaire (GHQ) (Goldberg 1988), Hospital Anxiety and Depression Scale (HADS) (Zigmond 1983), State Trait Anxiety Inventory (STAI) (Spielberger 1983), HADS.
3. Functional abilities, such as the Functional Independence Measure (FIM) (Uniform Data System for Medical Rehab 1993), Functional Assessment Measure (FAM) (Hal 1997), Nottingham Extended Activities of Daily Living (EADL) (Nouri 1987).
4. Quality of life, such as the Multiple Sclerosis Impact Scale (MSIS) (Hobart 2001) World Health Organization Quality of Life assessment (WHO-QoL) (The WHOQOL Group 1993), 36-item Short Form Health Survey (SF-36) (Ware 2001).

We also considered non-standardised measures, such as return to work and goal attainment, if studies had included these as a measure of outcome. If more than one of these scales was reported for each domain, we used the first scale in the list.

We categorised all outcomes into three separate time-points: "immediate", "intermediate", and "longer-term" and conducted separate analyses for each of these. We defined immediate as assessments conducted within the first month after completing the intervention, intermediate as assessments conducted between one

to six months later, and longer-term as any assessments conducted more than six months later.

Search methods for identification of studies

We conducted an electronic search with no restriction, and two review authors (LT, RdN) identified all potential studies.

Electronic searches

The Information Specialist used in the previous update was not available, so all studies were searched by the review authors (LT, RdN). We searched the Cochrane Central Register of Controlled Trials (CENTRAL) (2 June 2015 to 6 September 2020) which contains records from the following databases.

- MEDLINE (PubMed).
- Embase (Embase.com).
- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (EBSCO host).
- ClinicalTrials.gov.
- World Health Organization (WHO) International Clinical Trials Registry Portal (http://apps.who.int/trialsearch/).

The keywords used to search for studies for this review are listed in Appendix 1.

We also searched the following databases.

- The NIHR Clinical Research Network database (2 June 2015 to 6 September 2020).
- PsycINFO (2 June 2015 to 6 September 2020).
- Allied and Complementary Medicine Database (AMED) (2 June 2015 to 6 September 2020).
- Latin American and Caribbean Health Science Information Database (LILACS) (Bireme) (2 June 2015 to 6 September 2020).
- CAB Abstracts (2 June 2015 to 6 September 2020).

Searching other resources

We citation tracked all primary study articles and scanned reference lists from book chapters and review articles. We also examined studies identified by the Rosti-Otajärvi 2014 and Thomas 2006 MS reviews for inclusion. We only handsearched the reference lists of identified studies, not the full scientific journals, as until the early 1990s cognitive impairments were not universally recognised as a common complaint in MS (Rao 1991), and most RCTs have been reported (or updated) on electronic databases or journals. Furthermore, we would have found relevant trials from the search of the CENTRAL database, for which handsearching is carried out periodically, and we did not wish to duplicate this effort. Where necessary, we contacted authors of relevant trials to enquire whether their registered trials had been published, and to solicit more data where data required for the meta-analysis were not presented in the published paper in a format that could be used.

We accessed grey literature by searching (http://www.greynet.org/) and the British Library's EThOS database (http://ethos.bl.uk/Home.do). Grey literature is "a field in Library and Information Science that deals with the production, distribution, and access to multiple document types produced on all levels of government, academics, business, and organisation in electronic and print formats not controlled by commercial publishing i.e. where publishing is not the primary activity of the producing body" (GreyNet 2011).

Data collection and analysis

Selection of studies

One review author, (RdN), developed the search strategy in consultation with a senior librarian and the Cochrane Multiple Sclerosis and Rare Diseases of the CNS Group. Another review author (LT) evaluated abstracts of the studies obtained by this search strategy and identified trials for inclusion in the review using four inclusion criteria (types of trials, participants, interventions, and outcome measures). Five review authors (RdN, NE, DW, JMM, LS) cross-checked the search strategy, independently appraised the protocol, and confirmed the inclusion and exclusion of studies.

We eliminated articles based on the following exclusion criteria hierarchy,

- not MS, or a mixed-aetiology group without at least 75% of the sample being people with MS,
- not an RCT or quasi-RCT,
- not an adult population;
- not a memory rehabilitation study, or did not have a separate memory component if within the context of a larger "cognitive rehabilitation" (or "cognitive retraining" or "neuropsychological rehabilitation") study, or
- not a rehabilitation intervention study (not more than one session).

Data extraction and management

LT and another review author (NE, DW, JMM, or LS) independently assessed the methodological quality of each of the selected trials and rated them according to the guidelines of The Cochrane Collaboration. In case of disagreement, a third review author (RdN) arbitrated, and a verdict was reached. Our main considerations were whether participant allocation had been random and adequately concealed, and whether outcomes were performed blind to group allocation. We conducted the review using the Cochrane Review Manager software version 5.4.1 (RevMan 2020). The data extraction tool employed by the das Nair and Lincoln Cochrane review (das Nair 2016b) was used in this study and is therefore not replicated here.

Assessment of risk of bias in included studies

Review author LT and another review author (NE, DW, JMM, or LS) independently graded the included trials and completed the risk of bias table as described in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019).

The table includes the following domains.

- Random sequence generation
- Allocation concealment
- Blinding (of participants, personnel, and outcome assessors)
- Incomplete outcome data
- Selective outcome reporting
- Other sources of bias

On the basis of the information provided in the studies or by the authors of the primary studies, five review authors (LT, NE, DW,

JMM, LS) independently judged each of these domains as being low or high risk of bias, or unclear if information was insufficient. Any disagreements were arbitrated by another review author (RdN). As review authors working in the field of memory rehabilitation, we are familiar with the studies published in this area, and thus we could not be blinded to the names of the authors, institutions, or the publishing journal of the included trials. We made an evaluation of the overall risk of bias, based on the relative importance of the various domains listed. In addition to the risk of bias table, we used the GRADE approach to assessing quality of studies (GRADE Working Group 2004). This was completed across outcomes and is found in the summary of findings table. This approach allows for judgements to be made about the quality of the studies included in each outcome.

Measures of treatment effect

We planned to use odds ratio (OR) with 95% confidence intervals (CIs) for binary outcomes if reported. We used standardised mean difference (SMD) with 95% CIs for the continuous outcomes.

Unit of analysis issues

We included parallel-group, cluster-randomised, cross-over RCTs, and quasi-RCTs, and included the data from all these types of studies for the meta-analysis. For cross-over studies (as mentioned under Types of studies section), we only included the pre-cross-over phase of these trials. We did not combine the first and second phases of the cross-over studies because of uncertainty about the carry-over effects in such trials, given that they are psychological interventions, where the washout period is difficult to determine.

We included trials with more than two intervention groups and analysed them by pooling together the data on all the treatment groups (if appropriate) and compared them with the control group. If there was more than one control group, the results from the control groups were pooled together and compared with treatment.

We conducted separate analyses for the various outcomes and for the three different time-points (i.e. immediate, intermediate, and longer-term).

Dealing with missing data

Where data were not available from or unclear in the reports, we contacted the corresponding author of the studies in question for further information. We assessed the rates of attrition and missing data from the included studies (where available) and explored how these may have affected the results of the studies. If following several attempts to contact the study author we had not received a response, the missing data were not included in the analysis. Furthermore, if standard deviations (SDs) were not available from the papers, these values were inputted using methods specified in section 16.1.3.1 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019).

Assessment of heterogeneity

We considered heterogeneity by comparing the distribution of important participant factors between trials (age, gender, type of MS), and trial factors (sequence generation, allocation concealment, blinding, losses to follow-up). We employed the I^2 statistic to statistically assess heterogeneity (Higgins 2019; Huedo-Medina 2006). We further scrutinised the studies to explore the reasons for the heterogeneity if the I^2 statistic was significant at >= 50%.

Assessment of reporting biases

We considered reporting bias by conducting an exhaustive search of the literature that included but was not limited to the CENTRAL database, Embase, PsycINFO, LILACS, grey literature, reference lists of included studies and relevant reviews. We also considered reporting bias by deciding what outcomes would be assessed and reported before the meta-analysis was conducted.

Data synthesis

We consulted the Cochrane Handbook for Systematic Reviews of Interventions to plan the data synthesis (Higgins 2019), and followed the procedures outlined therein. As most psychological and neuropsychological outcome measures in memory rehabilitation tend to be ordinal-level measures, we treated these as continuous data (as recommended by Higgins 2019). the SMD was used as a summary statistic, using a random-effects model, because we predicted that multiple trials would use various outcome measures to assess memory and because of the heterogeneity of sampling.

If low scores represented a better outcome, the valence of the score was changed from positive to negative. In situations where studies combined scores from scales in which high scores are in some instances good outcomes and in some instances poor outcomes, the signs of the discrepant scores were reversed to keep them consistent. We only considered data that we deemed to be similar or comparable enough to meaningfully pool based on of the outcome measures used for the meta-analysis.

Subgroup analysis and investigation of heterogeneity

We planned subgroup analyses where at least two trials had separate data available for people with different subtypes of MS. Where significant heterogeneity was observed, we attempted to determine the causes of heterogeneity and explain this in our discussion. We did not plan on conducting subgroup analyses based on heterogeneity.

Sensitivity analysis

We considered sensitivity analyses to assess the impact of study quality (whether there was a difference between studies using an intention-to-treat analysis and an on-treatment analysis) where data needed to perform such analyses were available from the included papers. We also considered a sensitivity analysis to assess the influence of methodological quality on the intervention effect for each outcome by comparing the outcomes of those trials with low risk of bias with the outcomes of all the included studies. Following the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019), we made only informal comparisons (see Table 2), and did not conduct individual forest plots for each sensitivity analysis, but provided a summary table. Sensitivity analysis was also conducted to assess the impact of inputting the SD values as advised in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019).

Summary of findings and assessment of the quality of the evidence

We used the GRADE approach to interpret findings and present them in a summary of findings table, as advised in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2019). We considered seven key outcomes at one specifc time point they were measured to be important in the table, and thus present them in the summary of findings table.

The GRADE approach allows the quality of the evidence to be assessed clearly and without bias using four possible ratings: high, moderate, low and very low (Schünemann 2013). This rating system measures the degree of confidence that the true effect is close to that of the estimate of the effect, with high indicating very confident and very low indicating little confidence in the effect estimate. There are several factors that can lead to the downgrading of evidence such as risk of bias in included studies, inconsistency in results, and imprecision of effect estimates. If an outcome was downgraded, the reasons for this are detailed in the footnotes below the summary of findings table.

Summary of findings and assessment of the certainty of the evidence

We used the GRADE approach to interpret findings and present them in a 'Summary of findings' table, as advised in the Cochrane Handbook (Higgins 2019). We considered all of our outcomes at each time point they were measured to be important in the table, and thus present them in the 'Summary of findings' table.

The GRADE approach allows the quality of the evidence to be assessed clearly and without bias using four possible ratings: high, moderate, low and very low (Schünemann 2013). This rating system measures the degree of confidence that the true effect is close to that of the estimate of the effect, with high indicating very confident and very low indicating little confidence in the effect estimate. There are several factors that can lead to the downgrading of evidence such as risk of bias in included studies, inconsistency in results, and imprecision of effect estimates. If an outcome is downgraded, the reasons for this are detailed in the footnotes below the 'Summary of findings' table.

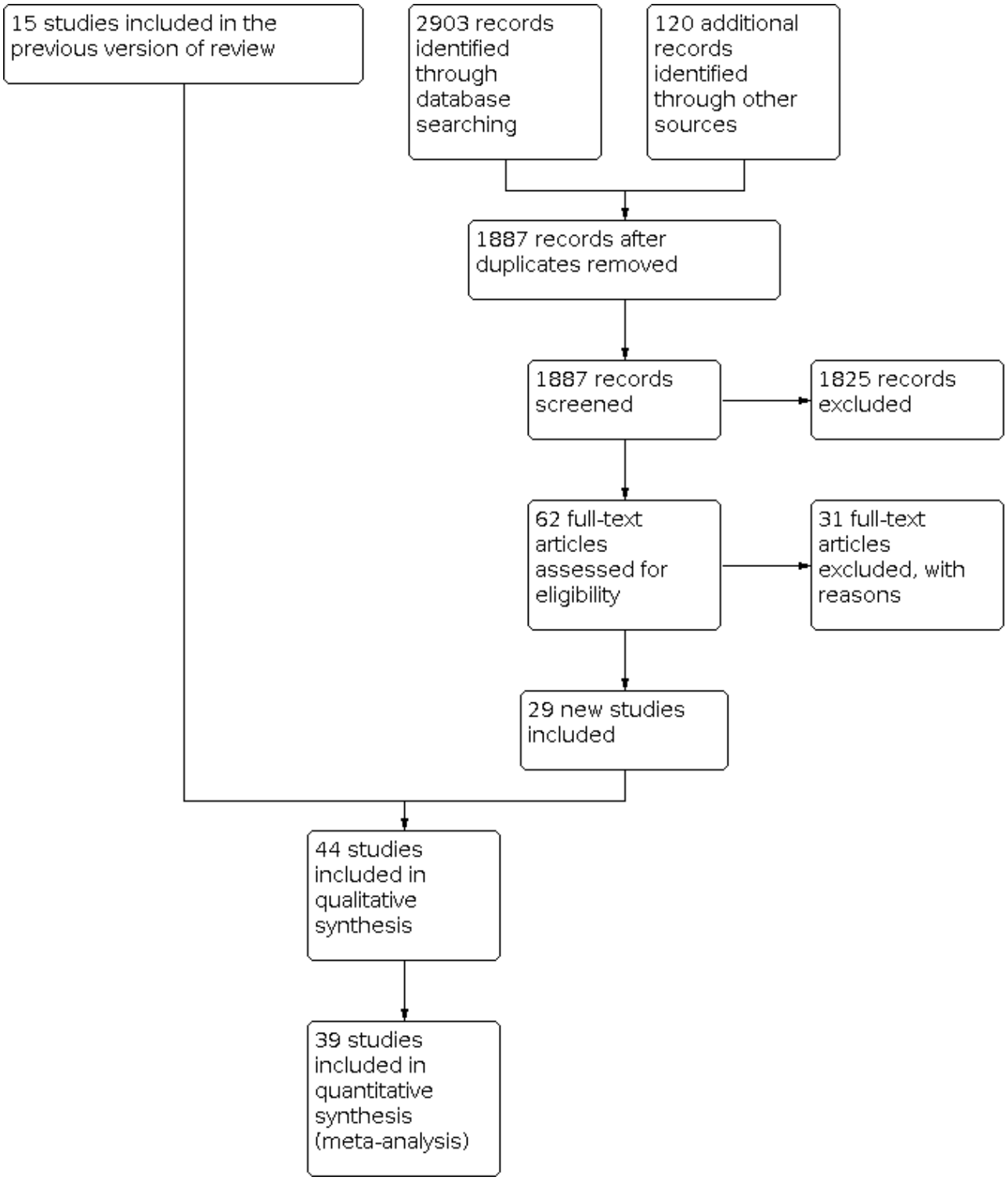
RESULTS

Description of studies

Results of the search

We identified a total of 29 studies using the above-mentioned search strategy. Fifteen studies from the previous review were added to the 29 new studies in the final analysis. Please see Figure 1.

Figure 1. Flow diagram showing article screening process



Included studies

Forty-four studies, comprising 2714 participants in total, met the inclusion criteria for this review (Campbell 2016; Carr 2014; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; Chmelařová 2020; Arian Darestani 2020; das Nair 2012; Naeeni Davarani 2020; De Luca 2019; Ernst 2015; Ernst 2018; Gich 2015; Goodwin 2020; Goverover 2018a; Hancock 2015; Hanssen 2015; Hildebrandt 2007; Huiskamp 2016; Impellizzeri 2020; Lincoln 2002; Lincoln 2020; Maggio 2020; Mani 2018; Mattioli 2016; Mendozzi 1998; Messinis 2017; Messinis 2020; Mousavi 2018a; Mousavi 2018b; Pedulla 2016; Perez-Martin 2017; Pusswald 2014; Rahmani 2020; Rilo 2018; Shahpouri 2019; Solari 2004; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005; Vilou 2020), and the Charvet

2017, De Luca 2019, and Jønsson 1993 studies were included in the review, but excluded from the meta-analysis because raw data were unattainable.

Twenty-eight of the included studies were undertaken in Europe (Austria, Denmark, Germany, Italy, Norway, Spain, Greece, Czech Republic, the UK), seven were from Iran, and nine were from the USA. All the European studies were recruited at hospital clinics or rehabilitation centres, with seven of these European studies recruiting from multiple centres (Goodwin 2020; Goverover 2018a; Lincoln 2020; Mattioli 2016; Messinis 2017; Perez-Martin 2017; Solari 2004). The maximum number of recruitment sites used was 10 (Mattioli 2016). Seven of the USA studies recruited participants from both clinic and community settings, with two of these USA studies recruiting from multiple centres (Chiaravalloti 2019a; Stuifbergen 2018).

There were nine multicentre trials (Goodwin 2020; Lincoln 2020; Hancock 2015; Mattioli 2016; Messinis 2017; Messinis 2020; Perez-Martin 2017; Solari 2004; Stuifbergen 2018). In terms of randomisation and stratification by site, Solari 2004, Messinis 2017 and Messinis 2020 used a site-stratified schedule. Lincoln 2020 had a 6:5 randomisation ratio, stratified by site and minimised by type of MS. Stuifbergen 2018 used a closed envelope method but did not specify stratification. Chmelařová 2020; Perez-Martin 2017; Rahmani 2020 did not specify their method of randomisation. Chiaravalloti 2019a; Goodwin 2020; Mattioli 2016 used random number generators, but did not specify stratification, and Stuifbergen 2018 used the closed envelope method. Hancock 2015 used a block-stratified randomisation procedure to ensure that equal types of each MS subtype were included in the intervention and control groups.

There were 35 single-centre trials (Arian Darestani 2020; Carr 2014; Campbell 2016; Charvet 2017; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; Chmelařová 2020; Naeeni Davarani 2020; De Luca 2019; Ernst 2015; Ernst 2018; Gich 2015; Goverover 2018a; Hanssen 2015; Hildebrandt 2007; Huiskamp 2016; Impellizzeri 2020; Jønsson 1993; Lincoln 2002; Maggio 2020; Mani 2018; Mattioli 2016; Mendozzi 1998; Mousavi 2018a; Mousavi 2018b; Pedulla 2016; Pusswald 2014; Rahmani 2020; Rilo 2018; Shahpouri 2019; Stuifbergen 2012; Tesar 2005; Vilou 2020). Five studies did not mention the method of generating the random schedule (Arian Darestani 2020; Ernst 2015; Mendozzi 1998; Pedulla 2016; Tesar 2005). One study reported that randomisation was quote: “performed by a lottery by the director of the rehabilitation centre” (Hanssen 2015). Four studies used quasi-randomisation: Chiaravalloti 2005 used odd-even random allocation, Hildebrandt 2007 and Pusswald 2014 allocated by alternating between intervention and control, and Arian Darestani 2020 quote: “divided [the participants] into control (n = 30) and experimental (n = 30) groups”. Six trials reported independent randomisation (Carr 2014; Chiaravalloti 2013; das Nair 2012; Lincoln 2002; Solari 2004; Tesar 2005), and Jønsson 1993 and Stuifbergen 2012 used a closed-envelope method. Mendozzi 1998 randomised the first 30 participants, and purposefully assigned the last 30 to balance age, gender, and education between groups; all data were included in our analysis. Gich 2015 stratified by level of cognitive impairment.

Participants were diagnosed with MS using the Poser criteria (Poser 1983), in seven studies, using the McDonald criteria, (McDonald 2001 in 18 studies, and the Schumacher criteria (Schumacher 1965). in one study (Jønsson 1993). Eighteen studies did not report the criteria used to diagnose MS, but merely stated that participants had clinically-definite MS. Twenty-six studies included participants with mixed types of MS (relapsing remitting MS (RRMS) and secondary progressive MS (SPMS) in Campbell 2016; das Nair 2012; Gich 2015; Lincoln 2002; Maggio 2020; Mendozzi 1998; Mousavi 2018b; Pedulla 2016; Tesar 2005; and RRMS, SPMS, and primary progressive MS (PPMS) in Carr 2014; Charvet 2017; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019b; Goverover 2018a; Hancock 2015; Hanssen 2015; Jønsson 1993; Impellizzeri 2020; Lincoln 2020; Perez-Martin 2017; Rilo 2018; Shahpouri 2019). Eight studies included participants with RRMS only (Ernst 2015; Ernst 2018; Hildebrandt 2007; Mani 2018; Mattioli 2016; Messinis 2017; Rahmani 2020; Vilou 2020). Two studies included participants with RRMS, SPMS, PPMS and progressive-relapsing MS (Chiaravalloti 2019a; Huiskamp 2016) with one study including these participants and participants with benign MS and “unknown” types (Goodwin 2020). The type of MS was not reported in eight studies (Chmelařová 2020; Arian Darestani 2020; Naeeni Davarani 2020; De Luca 2019; Mousavi 2018a; Pusswald 2014; Solari 2004; Stuifbergen 2012).

The number of participants in the studies ranged from 16, inHuiskamp 2016, to 449, in Lincoln 2020, and the number of participants in treatment or control groups ranged from seven, in das Nair 2012 and Huiskamp 2016, to 245, in Lincoln 2020. Most participants were in their 40s. Varied gender ratios were reported, with the percentage of women ranging from 36.7%, in Impellizzeri 2020, to 100%, in Mani 2018 and Rahmani 2020. The participants had a minimum of elementary education in most studies, with the participants in the Chiaravalloti 2019b having the highest number of years of education (16.07 in intervention, 16.46 in control). De Luca 2019 and Perez-Martin 2017 had the lowest number of years of education (10.8 in intervention, 11.3 in control and 10.2 in intervention, 11.6 in control, respectively). Six studies did not report education (Chmelařová 2020; Mousavi 2018a; Mousavi 2018b; Shahpouri 2019; Tesar 2005; Vilou 2020).

The groups were comparable on assessed baseline characteristics in 32 studies, and in the other studies where differences were observed, they were statistically corrected (Chiaravalloti 2005; das Nair 2012; Hancock 2015; Hildebrandt 2007; Huiskamp 2016; Jønsson 1993; Solari 2004), with the exception of Mendozzi 1998 and Stuifbergen 2012. Two studies appeared to be matched for baseline characteristics, but no statistics were reported (Arian Darestani 2020; Rahmani 2020), and one study had unequal groups due to stratification requirements but overall was well-matched (Charvet 2017).

Thirty-seven studies used two-group comparisons (treatment versus control), and six studies used three-group comparisons (das Nair 2012; Ernst 2015; Lincoln 2002; Mendozzi 1998; Mousavi 2018a; Mousavi 2018b). Lincoln 2002 used assessment versus assessment plus feedback versus assessment plus feedback and treatment; Mendozzi 1998 examined specific cognitive retraining versus non-specific cognitive retraining versus control; and das Nair 2012 investigated restitution versus compensation versus self-help control. Rahmani 2020 used computer-based versus manual-based versus mixed cognitive rehabilitation versus placebo versus control, five groups in total.

Twenty-nine studies used individual treatment, including clinic-based and home-based interventions (Campbell 2016; Charvet 2017; Chiaravalloti 2019a; Chiaravalloti 2019b; Chmelařová 2020;

Arian Darestani 2020; Naeeni Davarani 2020; De Luca 2019; Ernst 2015; Ernst 2018; Gich 2015; Goodwin 2020; Goverover 2018a; Hancock 2015; Hildebrandt 2007; Huiskamp 2016; Jønsson 1993; Lincoln 2002; Maggio 2020; Mattioli 2016; Mendozzi 1998; Messinis 2017; Messinis 2020; Pedulla 2016; Pusswald 2014; Rahmani 2020; Solari 2004; Stuifbergen 2018; Vilou 2020), and 13 had group interventions (Carr 2014; Chiaravalloti 2005; Chiaravalloti 2013; das Nair 2012; Impellizzeri 2020; Lincoln 2020; Mani 2018; Mousavi 2018a; Mousavi 2018b; Rilo 2018; Shahpouri 2019; Stuifbergen 2012; Tesar 2005). One study used a mix of both group and individual sessions (Hanssen 2015), and another used both group sessions and individual computerised sessions (Stuifbergen 2012).

The structure and content of the treatment programmes varied. Most interventions were of four to eight weeks duration (Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; Chmelařová 2020; Arian Darestani 2020; Naeeni Davarani 2020; Goodwin 2020; Hancock 2015; Hanssen 2015; Hildebrandt 2007; Huiskamp 2016; Impellizzeri 2020; Jønsson 1993; Lincoln 2002; Maggio 2020; Mani 2018; Mendozzi 1998; Messinis 2020; Mousavi 2018a; Mousavi 2018b; Pedulla 2016; Pusswald 2014; Solari 2004; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005; Vilou 2020). Carr 2014, das Nair 2012, Lincoln 2020, Messinis 2017 had 10-week programmes, Charvet 2017 and Perez-Martin 2017 had 12-week programmes, Mattioli 2016 had a 15-week programme, Rahmani 2020 had a 21-week programme, and Gich 2015 used a six-month programme. Four studies did not specify set durations of their treatment but either selected a number of sessions to be completed when the participants were available (Ernst 2015; Ernst 2018; Shahpouri 2019) or specified a timeframe for the sessions to be completed in (Lincoln 2002).

Sessions ranged from 30 minutes, (Hildebrandt 2007, Pedulla 2016, and Pusswald 2014), and two hours (Hanssen 2015, Mani 2018, and Shahpouri 2019), and participants met one to six times a week in all studies except Mendozzi 1998, where the treatment was bi-weekly. The Goodwin 2020 study lasted for two months and the session frequency varied as it was dependent on the types of text message reminders each participant required. Similarly, the Lincoln 2002 study specified a six-month timeframe in which the sessions had to be completed, but the frequency of sessions depended on individual need. The lowest number of total sessions was six (Ernst 2015; Ernst 2018; Goverover 2018a) and the highest number of total sessions was 60 (Charvet 2017). Fourteen studies had between eight and 10 sessions (Carr 2014; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; Arian Darestani 2020; Naeeni Davarani 2020; Hanssen 2015; Huiskamp 2016; Lincoln 2020; Mani 2018; Mousavi 2018a; Mousavi 2018b; Shahpouri 2019). Nine studies had between 12 and 18 sessions (Campbell 2016; das Nair 2012; Jønsson 1993; Mendozzi 1998; Perez-Martin 2017; Pusswald 2014; Solari 2004; Tesar 2005; Vilou 2020). Seven studies had between 20 and 30 sessions (De Luca 2019; Hildebrandt 2007; Maggio 2020; Messinis 2017; Messinis 2020; Rahmani 2020; Stuifbergen 2018). Six studies had between 32 and 48 sessions (Chmelařová 2020; Gich 2015; Hancock 2015; Impellizzeri 2020; Rilo 2018; Stuifbergen 2012). For two studies the frequency of sessions depended on individual need (Goodwin 2020; Lincoln 2002).

In three studies, the contents of the treatment programmes were individualised (Goodwin 2020; Hanssen 2015;Lincoln 2002), depending on the needs of the participant. Seven studies used comprehensive memory rehabilitation programmes (including teaching participants to use internal and external memory aids) (Carr 2014; das Nair 2012; Jønsson 1993; Lincoln 2002; Lincoln 2020; Pusswald 2014; Tesar 2005). Sixteen studies used computerised memory- and attention-retraining packages (Campbell 2016; Charvet 2017; Chmelařová 2020; Arian Darestani 2020; Naeeni Davarani 2020; Gich 2015; Hancock 2015; Hildebrandt 2007; Mendozzi 1998; Messinis 2020; Pedulla 2016; Pusswald 2014; Solari 2004; Stuifbergen 2012; Stuifbergen 2018; Vilou 2020), and Chiaravalloti 2005, Chiaravalloti 2013, and Chiaravalloti 2019b used the Story Memory Technique, which involved the use of imagery and story generation. De Luca 2019 used both computerised and paper-and-pencil training strategies, but did not explain the specifics of what this entailed. Goodwin 2020 used mobile phones to deliver reminders throughout the day, and the number of messages delivered varied depending on each person's needs.

Studies that had a sham or attention control group reported having ensured that these groups had minimal memory content, thereby reducing contamination (Chiaravalloti 2005; Chiaravalloti 2013; das Nair 2012; Ernst 2018; Hancock 2015; Jønsson 1993; Mousavi 2018a; Mousavi 2018b; Rahmani 2020; Solari 2004).

Lincoln 2020 assessed fidelity of intervention using three methods: firstly, a the cognitive rehabilitation followed a manual that was developed and tested in a pilot study (Carr 2014), secondly; the training was delivered by psychology graduates with clinical experience and they received training from a clinical psychologist as well as monthly teleconferences to discuss specific challenges; and thirdly, the intervention sessions were recorded and coded by an independent researcher using the time-sampling procedure and found that the intervention was delivered as intended. However, only three other studies assessed fidelity of intervention.

The 44 included studies used a range of outcome measures. All studies included at least one measure of learning or memory, with the exception of Hanssen 2015, where outcomes were related to psychological functioning and impact of disease.

Seventeen studies used subjective measures of memory. Six studies (Carr 2014; das Nair 2012; Goodwin 2020; Lincoln 2002; Lincoln 2020; Shahpouri 2019) used the Everyday Memory Questionnaire (EMQ) (Sunderland 1983), and das Nair 2012 used the Internal and External Memory Aids Questionnaires based on the Memory Aids Questionnaire (Wilson 1984). Four studies (Chiaravalloti 2005; Chiaravalloti 2019a; Goverover 2018a; Mani 2018) used the Memory Failures Questionnaire (MFQ) (Gilewski 1990); and three studies (Mousavi 2018b; Perez-Martin 2017; Stuifbergen 2012) used the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ) (Benedict 2004); one study (Chiaravalloti 2019a) used the Awareness Questionnaire (AQ) (Sherer 2004); one study (Chmelařová 2020) used the Cognitive Failures Questionnaire (CFQ) (Broadbent 1982); one study (Stuifbergen 2018) used the strategy subscale of the Multifactorial Memory Questionnaire (MMQ) (Troyer 2017).

Twenty-five trials used list-learning tasks: Hopkins Verbal Learning Task-Revised (HVLT-R) (Benedict 1998) (Chiaravalloti 2005; Rilo 2018); Verbal Learning Test (VLT) (Sturm 1999a) (Tesar 2005); California Verbal Learning Task-II (CVLT) (Delis 2000) (Arian Darestani 2020; Campbell 2016; Chiaravalloti 2013; Chiaravalloti

2019a; Chiaravalloti 2019b; Goverover 2018a; Hildebrandt 2007; Stuifbergen 2012; Stuifbergen 2018; Vilou 2020); Greek Verbal Learning Trial (GVLT) (Messinis 2020); Auditory Verbal Learning Test (AVLT) (Lezak 2004) (Hancock 2015); Selective Reminding Task (De Luca 2019; Gich 2015; Impellizzeri 2020; Lincoln 2020; Mattioli 2016; Messinis 2017; Pedulla 2016; Perez-Martin 2017; Rao 1993); and the list-learning task used by one study was not specified (Jønsson 1993). Seven studies used neuropsychological test batteries or subtests of these. One study, Mendozzi 1998, used the memory scale of the Luria-Nebraska Neuropsychological Battery (LNNB) consisting of 13 items (Golden 1980). Subtests from other test batteries included Buschke Selective Reminding Test from an Italian version of the Brief Repeatable Battery of Neuropsychological Tests (BRBNT) (Solari 2002), unspecified tests from the Rivermead Behavioural Memory Test (RBMT-E) (Wilson 1999), and the Doors and People Test (Baddeley 1994). Pusswald 2014 used the MUSIC assessment (Calabrese 2004), and Jønsson 1993 used an unspecified battery. Non-verbal memory was assessed using individual tests or part of a battery. Individual tests included the Noverbaler Lerntest (NVLT) (Sturm 1999b) (Tesar 2005), and an unspecified 50-faces recognition test (Jønsson 1993).

Seventeen trials used visual objective memory measures: Brief Visuospatial Memory Test (BVMT-R) (Benedict 1996) (Campbell 2016; Chiaravalloti 2019a; Chiaravalloti 2019b; Messinis 2017; Messinis 2020; Stuifbergen 2012; Stuifbergen 2018; Vilou 2020); 10/36 Spatial Recall Test (SPART) (Rao 1990) (De Luca 2019; Impellizzeri 2020; Lincoln 2020; Maggio 2020; Mattioli 2016; Pedulla 2016; Perez-Martin 2017); Contextual memory test (CMT) (Toglia 2004) (Goverover 2018a); Rey-Osterrieth complex figure (ROCF) (Rey 1941) (Maggio 2020).

Fourteen trials used working memory measures: Paced auditory serial addition test (PASAT) (Rao 1990) (Naeeni Davarani 2020; De Luca 2019; Impellizzeri 2020; Lincoln 2020; Maggio 2020; Mattioli 2016; Pedulla 2016; Perez-Martin 2017; Rahmani 2020; Stuifbergen 2018); N-back test (Huiskamp 2016; Pedulla 2016); Digit span WAIS subtest (Rilo 2018; Shahpouri 2019).

In terms of 'other cognitive outcomes', the most frequently assessed cognitive domain was information processing. Nineteen studies included information processing measures: Symbol Digit Modalities Test (SDMT) (Campbell 2016; De Luca 2019; Hancock 2015; Impellizzeri 2020; Lincoln 2020; Mattioli 2016; Messinis 2017; Messinis 2020; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Rilo 2018; Solari 2004; Stuifbergen 2012; Stuifbergen 2018; Vilou 2020); Stroop colour test (SCWT) (Stroop 1938) (Rahmani 2020); Trail Making Test (TMT) (Chmelařová 2020); Behaviour Rating Inventory of Executive Function – Adult version (BRIEF-A) (Hanssen 2015).

The most frequently used mood measure was the Beck Depression Inventory (BDI) (Beck 1987), used in 11 studies (Chiaravalloti 2005; Chmelařová 2020; De Luca 2019; Hancock 2015; Hildebrandt 2007; Impellizzeri 2020; Maggio 2020; Messinis 2020; Mousavi 2018a; Shahpouri 2019; Tesar 2005). Six studies, Carr 2014, das Nair 2012, Goodwin 2020, Lincoln 2002, Lincoln 2020, and Mousavi 2018a, used the General Health Questionnaire (GHQ-28) (Goldberg 1988); three, Chiaravalloti 2013, Chiaravalloti 2019b, and Goverover 2018a, used the Chicago Mood Depression Inventory (CMDI) (Nyenhuis 1998), and two of these studies also used the STAI (depression and anxiety subscale); Chiaravalloti 2019b and Goverover 2018a, and another, Solari 2004, used the Italian version of the CMDI (Solari 2003). One study used the Montgomery-Asberg Depression Rating Scale (MADRS, depression and anxiety subscale) (Mattioli 2016), one used the Centre for Epidemiological Studies Depression (CES-D) (Stuifbergen 2018).

Nine studies (De Luca 2019; Hancock 2015; Impellizzeri 2020; Lincoln 2002; Maggio 2020; Mattioli 2016; Perez-Martin 2017; Shahpouri 2019; Solari 2004) assessed quality of life using the Multiple Sclerosis Quality of Life (MSQOL-54; Vickrey 1995), three studies, Carr 2014, Hanssen 2015, and Lincoln 2020 used the Multiple Sclerosis Impact Scale (MSIS-29) (Hobart 2001), two studies, Goodwin 2020 and Messinis 2020, used the EQ-5D-5L, one, Chiaravalloti 2019a, used the SF-36 and one, Goverover 2018a, used the Satisfaction with Life Scales (SWLS).

Only two studies examined whether their rehabilitation programme affected instrumental ADL (das Nair 2012 and Lincoln 2002), by using the Extended Activities of Daily Living scale (EADL) (Nouri 1987). Four studies (Campbell 2016; Chiaravalloti 2013; Chiaravalloti 2019a; Goverover 2018a) assessed functional independence with the Functional Assessment of Multiple Sclerosis (FAMS) (Cella 1996). One study, Stuifbergen 2018, used the Instrumental Activities of Daily Living (IADL).

Eighteen studies were observer-blinded RCTs or quasi-randomised trials (Carr 2014; das Nair 2012; Gich 2015; Hildebrandt 2007; Impellizzeri 2020; Jønsson 1993; Goodwin 2020; Lincoln 2002; Lincoln 2020; Maggio 2020; Mani 2018; Mendozzi 1998; Messinis 2020; Mousavi 2018b; Perez-Martin 2017; Stuifbergen 2012; Tesar 2005; Vilou 2020), and 14 stated that they were observer- and participant-blinded (Charvet 2017; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; De Luca 2019; Goverover 2018a; Hancock 2015; Huiskamp 2016; Mousavi 2018a; Naeeni Davarani 2020; Shahpouri 2019; Solari 2004; Stuifbergen 2018). One study reported that blinding of participants was not possible due to the nature of the intervention, and there was no mention of observer blinding (Hanssen 2015). However, all outcomes were self-report questionnaire-based, therefore blinding was not deemed necessary. Twelve studies either did not use blinding procedures or were unclear in their methodology (Arian Darestani 2020; Campbell 2016; Chmelařová 2020; Ernst 2015; Jønsson 1993; Mattioli 2016; Messinis 2017; Pedulla 2016; Pusswald 2014; Rahmani 2020; Rilo 2018; Tesar 2005), therefore we determined these studies to be at high risk of bias. One study reported that while the main scorer was not blinded, a blinded rater verified the scoring accuracy for 20% of memories randomly chosen with a reliability of.95 when assessed with intraclass correlations, therefore we determined this to be a low risk of bias (Ernst 2018).

Excluded studies

We excluded 64 studies based on the exclusion criteria specified for this review. Two were studies of Alzheimer's disease, i.e. not MS (Akhtar 2006; Loewenstein 2004); four were unrelated to memory (comparative study of Barthel Index and Functional Independence Measure in van der Putten 1999, and falls in Aisen 1994, Canellopoulou 1998, and Flavia 2010); and one was a systematic review, not an intervention study (Thomas 2006). Sixteen studies were not specific to memory, but general neuropsychological rehabilitation, attention, or information processing (Amato 2014; Bhargav 2016; Cabrera-Gomez 2010; Canellopoulou 1998; Chiaravalloti 2018; De Giglio 2014; Flavia 2010; Goreover 2011; Grasso 2017; Hanssen 2016; Mattioli 2012; Mäntynen 2014; Rosti-

Otajärvi 2013a; Rosti-Otajärvi 2013b; Veldkamp 2019; Zimmer 2018). Seven studies used healthy controls instead of an MS control group (Aguirre 2019; Aldrich 1995; Chiaravalloti 2003; Ernst 2013; Lamargue 2020; Vogt 2009; Wilson 2001), and Wilson 2001 also did not distinguish between results for people with MS and others with acquired progressive brain injury. Eleven studies were not RCTs or quasi-RCTs (one quasi-experimental waiting-list control: Rodgers 1996; one small group study: Allen 1998; six without random allocation: Barker 2019; Barbarulo 2018; Brenk 2008; Brissart 2013; Pineau 2019; Shatil 2010 three with no control group: Bove 2019; Brissart 2010; Güçlü Altun 2015). One study was a brain imaging study and had an active control group (Bonavita 2015). One study used a "music intervention" (Thaut 2014). One study was not considered to be a rehabilitation study according to our inclusion criteria because it only involved one hour-long session of memory retraining (Moore 2008). Three studies used the same sample, or a subgroup of the sample, of Chiaravalloti 2013 (Chiaravalloti 2012; Dobryakova 2014; Leavitt 2014), and another, Martin 2014, was a subgroup analysis of das Nair 2012, and was therefore not included. Two studies had abstracts in English but no full-text in English available (Fiorotto 2015; Jimenez-Morales 2017). Four studies were study protocols and therefore had no results attached to them (Guijarro-Castro 2017; Harand 2019; Lincoln 2015; Nauta 2017). Finally, 13 studies were conference poster presentations, and/or no full texts could be found (Bove 2019; Campbell 2015; das Nair 2017; Harand 2017; Iaffaldano 2015; Kavaklioglu 2017; Messinis 2015; Penner 2018; Perez-Martin 2016; Rilo 2015; Rilo 2016; Rilo 2017; Nurova 2014).

Risk of bias in included studies

The risk of bias in the 44 included studies was generally low (see Figure 2 and Figure 3 – for individual study risk of bias assessments). However, high risk of selection and detection bias was found in the following: random sequence generation in four studies, allocation concealment in two studies, blinding procedures in 14 studies, incomplete outcome data in four studies, and possible selective reporting in four studies. Furthermore, we judged the risk of bias to be unclear in some instances due to insufficient reporting of methods for: randomised sequence generation in 11 studies, allocation concealment in 14 studies, blinding procedures in two studies, and incomplete outcome data in 10 studies.

Figure 2. Methodological quality graph: review authors' judgements about each methodological quality item presented as percentages across all included studies.

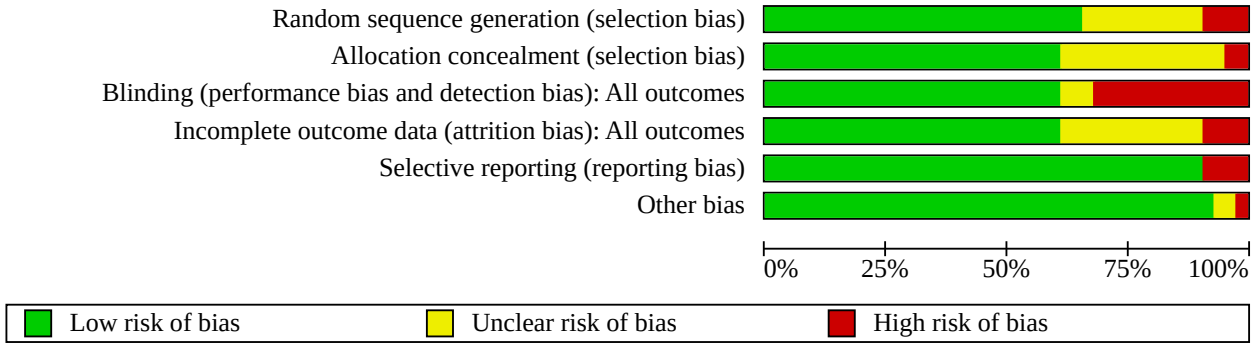


Figure 3. Methodological quality summary: review authors' judgements about each methodological quality item for each included study.

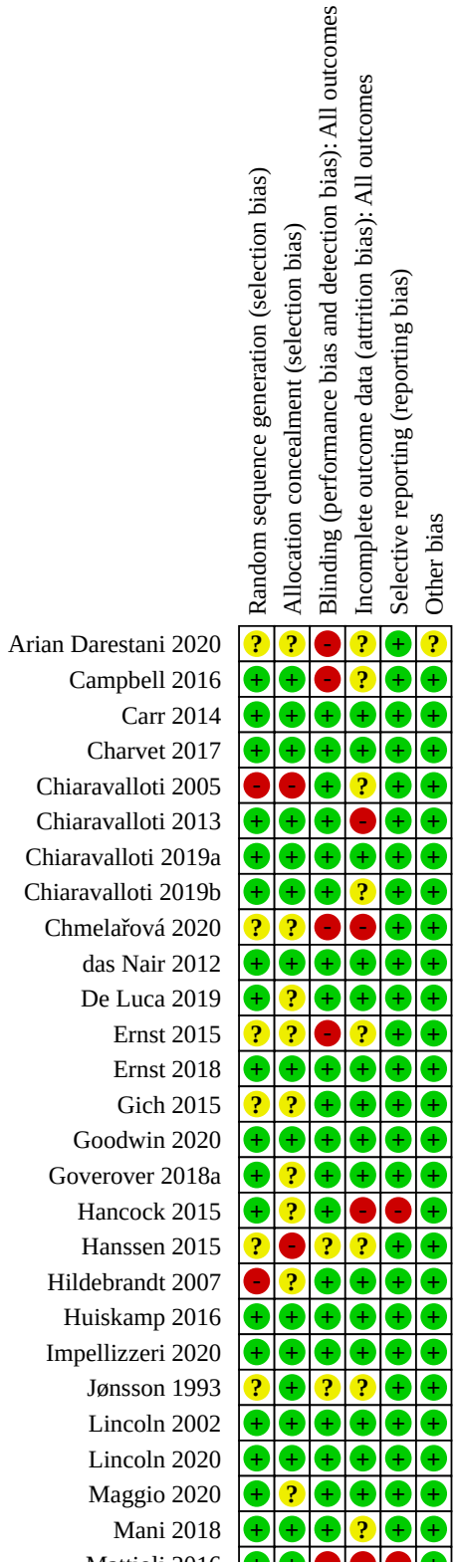
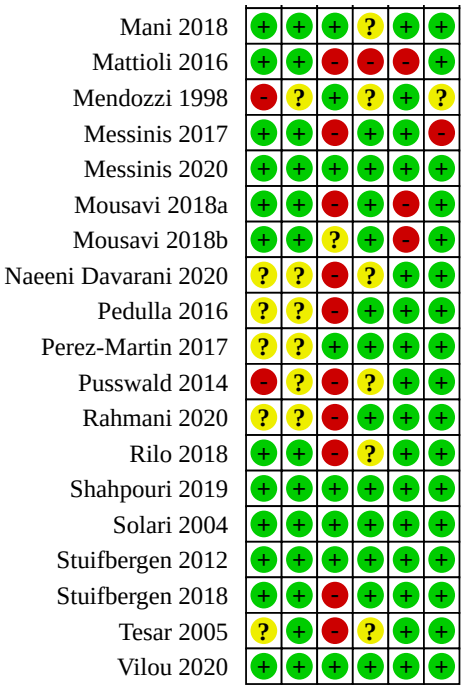


Figure 3. (Continued)



Random sequence generation

Seventeen studies were judged to have a low risk of selection bias due to having adequate random sequence generation, having used a computerised random number generator by an independent unit (Campbell 2016; Carr 2014; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; das Nair 2012; Ernst 2018; Huiskamp 2016; Impellizzeri 2020; Lincoln 2002; Mani 2018; Mattioli 2016; Messinis 2017; Messinis 2020; Rilo 2018; Shahpouri 2019; Solari 2004), two used a random number generator from the study data analyst that was created prior to recruitment and kept in sealed envelopes (Stuifbergen 2012; Stuifbergen 2018) and six used a block randomisation generated by a blind statistician (Charvet 2017; De Luca 2019; Goodwin 2020; Goverover 2018a; Lincoln 2020; Maggio 2020). Three studies used “randomised software” to randomly assign participants to three groups and therefore was determined to have low risk of bias (Mousavi 2018a; Mousavi 2018b; Vilou 2020). Four studies were judged not to have adequate sequence generation and therefore a high risk of bias, as three of these studies involved quasi-random 'odd-even' or alternating allocation (Chiaravalloti 2005; Hildebrandt 2007; Pusswald 2014), and one of these studies only randomised half the sample with no generation method stated (Mendozzi 1998). The method used for random sequence generation and the risk of bias in 11 other studies was unclear (Arian Darestani 2020; Chmelařová 2020; Ernst 2015; Gich 2015; Hanssen 2015; Jønsson 1993; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Rahmani 2020; Tesar 2005).

Allocation

We judged 19 studies to have a low risk of selection bias due to effectively concealing allocation into groups using a computerised random number generator by an independent unit (Campbell 2016; Carr 2014; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; das Nair 2012; Ernst 2018; Huiskamp 2016; Impellizzeri 2020; Lincoln 2002; Mani 2018; Mattioli 2016; Messinis 2017; Messinis 2020; Mousavi 2018a; Mousavi 2018b; Rilo 2018; Shahpouri 2019; Solari 2004), a closed envelope system (Jønsson 1993; Stuifbergen 2012; Stuifbergen 2018; Vilou 2020), or having a separate staff member who was not otherwise involved in the study complete allocation (Charvet 2017; Goodwin 2020; Goverover 2018a; Lincoln 2020; Tesar 2005). We judged two studies as not having concealed allocation to groups, suggesting a high risk of bias: one having used "odd-even" allocation completed by the principal investigator (Chiaravalloti 2005), and one stating that allocation concealment was not possible (Hanssen 2015). Fourteen studies were unclear in their explanation of allocation concealment: one informing participants whether they were to receive the intervention or assessment only (Hildebrandt 2007); one in which the principal investigator allocated groups and what other involvement they had in the study was not clearly explained (Mendozzi 1998); and 12 studies not mentioning allocation concealment (Arian Darestani 2020; Chmelařová 2020; De Luca 2019; Ernst 2015; Gich 2015; Hancock 2015; Maggio 2020; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Pusswald 2014; Rahmani 2020).

Blinding

Sixteen studies were observer blinded (Carr 2014; das Nair 2012; De Luca 2019; Gich 2015; Goodwin 2020; Hildebrandt 2007; Impellizzeri 2020; Lincoln 2002; Lincoln 2020; Maggio 2020; Mani 2018; Mendozzi 1998; Messinis 2020; Perez-Martin 2017; Stuifbergen 2012; Vilou 2020). 14 studies stated they were “double blind” (Charvet 2017; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; Goverover 2018a; Hancock 2015; Huiskamp 2016; Mani 2018; Mousavi 2018a; Naeeni Davarani 2020; Shahpouri 2019; Solari 2004; Stuifbergen 2018), however, three of these

studies (Mousavi 2018a; Naeeni Davarani 2020; Stuifbergen 2018) were judged to be high risk of bias due to lack of evidence of how they blinded the personnel and/or participants. One study reported that while the main scorer was not blinded, a blinded rater verified the scoring accuracy for 20% of memories randomly chosen with a reliability of.95 when assessed with intraclass correlations, therefore we determined this to be a low risk of bias (Ernst 2018). One study reported that blinding of participants was not possible due to the nature of the intervention (Hanssen 2015), and there was no mention of observer blinding, but because the outcomes were self-report questionnaires, we deemed this study to have an unclear risk of bias. One study was rated unclear bias due to discrpenaices in the blinding procedures found when the study stated that "Healthcare providers were not told of patients' allocation, but a few words would have given it away" (Jønsson 1993). It was not clear whether this occured or whether the authors made any attempt to prevent it by asking patients not to discuss their experience with the assessors (Jønsson 1993). Eleven studies either did not use any blinding procedures or were unclear in their methodology, suggesting a high risk of bias (Arian Darestani 2020; Campbell 2016; Chmelařová 2020; Ernst 2015; Mattioli 2016; Messinis 2017; Pusswald 2014; Pedulla 2016; Rahmani 2020; Rilo 2018; Tesar 2005). One study states the patients and statisticians were blind to group allocation, but it is unclear whether the assessor was blind, therefore suggesting an unclear risk of bias (Mousavi 2018b).

Incomplete outcome data

We deemed four studies to be at high risk of attrition bias: in three studies (Chiaravalloti 2013; Chmelařová 2020; Mattioli 2016), there was a post-randomisation attrition rate of 12%, 25% and 21%, respectively and/or no discussion of how missing data were dealt with, and the study did not use intention-to-treat analysis; in the other study, the post-randomisation attrition level was 44% (Hancock 2015). Eight studies did not address incomplete outcome data and did not use intention-to-treat analysis, which we deemed to be at unclear risk of bias: one study reported one dropout (Chiaravalloti 2005), two studies reported two dropouts (Hanssen 2015; Rilo 2018), three studies reported three dropouts (Campbell 2016; Ernst 2015; Mani 2018), two studies reported six dropouts (Chiaravalloti 2019b; Naeeni Davarani 2020); one study reported seven dropouts (Arian Darestani 2020); in another, participant outcome data were replaced with mid-trial data if a participant dropped out (Mendozzi 1998); and two studies did not explain how dropout data were handled (Jønsson 1993; Tesar 2005). One study conducted analyses on data for those participants who completed the outcome assessments (Lincoln 2002), one used list-wise deletion and baseline data imputed for any missing follow-up data (das Nair 2012), and in two studies (Solari 2004; Stuifbergen 2012), missing values were imputed according to the last observation carried forward method. In one study, where less than 10% of items were missed on a questionnaire, these were replaced with the mean for the questionnaire (Carr 2014).

Selective reporting

We deemed four studies to have a high risk of reporting bias (Hancock 2015; Mattioli 2016; Mousavi 2018a; Mousavi 2018b). One study only reported on the memory outcomes, despite other outcomes having been assessed at follow-up, and data were only reported for "good adherers" to the intervention (Hancock 2015). One study did not report outcome comparisons for control group, only the intervention group (Pedulla 2016). Three studies did not report several of their outcomes (Mattioli 2016; Mousavi 2018a; Mousavi 2018b).

Other potential sources of bias

We judged 39 studies to have a low risk of other potential sources of bias (Carr 2014; Campbell 2016; Charvet 2017; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; Chmelařová 2020; das Nair 2012; De Luca 2019; Ernst 2015; Ernst 2018; Gich 2015; Goodwin 2020; Goverover 2018a; Hancock 2015; Hanssen 2015; Hildebrandt 2007; Huiskamp 2016; Impellizzeri 2020; Jønsson 1993; Lincoln 2002; Lincoln 2020; Maggio 2020; Messinis 2020; Mousavi 2018a; Mousavi 2018b; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Pusswald 2014; Rahmani 2020; Rilo 2018; Shahpouri 2019; Solari 2004; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005; Vilou 2020). One study had a potential source of bias, as one participant in the treatment group discontinued cognitive retraining and was replaced by a new entry without further explanation (Mendozzi 1998). One study had a potential source of bias as it was unclear what the control group were told about the study (Arian Darestani 2020). One study did not collect six-month follow-up data for the control group, therefore we determined this to be high risk of potential bias (Messinis 2017).

Effects of interventions

See: **Summary of findings 1** Memory rehabilitation for people with multiple sclerosis

In this section, we first present study-specific information regarding intervention effect on memory outcomes, and then present the meta-analysis, synthesising results on various domains.

Nine studies concluded that there were no significant differences between the treatment and control groups on measures of memory, particularly after adjustments were made for multiple testing (Campbell 2016; Carr 2014; Chiaravalloti 2005; Chiaravalloti 2019a; das Nair 2012; Hancock 2015; Jønsson 1993; Lincoln 2002; Solari 2004), and Goodwin 2020 reported no significant within group improvements for the intervention group. Twenty-nine studies reported significant differences on memory measures favouring the treatment groups (Arian Darestani 2020; Chiaravalloti 2013; Chiaravalloti 2019b; Chmelařová 2020; De Luca 2019; Gich 2015; Goverover 2018a; Hildebrandt 2007; Impellizzeri 2020; Lincoln 2020; Maggio 2020; Mani 2018; Mattioli 2016; Mendozzi 1998; Messinis 2017; Messinis 2020; Mousavi 2018a; Mousavi 2018b; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Pusswald 2014; Rahmani 2020; Rilo 2018; Shahpouri 2019; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005; Vilou 2020). One study did not use memory outcomes (Hanssen 2015). Gich 2015 reported significant differences favouring treatment on some subtests of the Battery of Neuropsychological Tests (BRBN) (Rao 1993), although no significant differences were reported on the list-learning task of the BRBN used in this meta-analysis. Campbell 2016 showed no significant improvement on the California Verbal Learning Test (CVLT-II) or Brief Visuospatial Memory Test (BVMT). Chiaravalloti 2019b showed significant improvements for the intervention group in the CVLT-II at immediate follow-up but not in the Rivermead Behavioural Memory Test (RBMT). Chmelařová 2020 showed significant improvement for the intervention group in the immediate memory component of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)

(Randolph 1998), but showed no improvement for Cognitive Failures Questionnaire (CFQ). Goverover 2018a showed significant improvement for the intervention group in CVLT-II and: Contextual Memory Text (CMT), but not on the Memory Failures Questionnaire (MFQ). Hildebrandt 2007 reported improvements for the treatment group in the Learning Trials and Long Delay Free Recall subtests of the CVLT (Niemann 2003). Stuifbergen 2012 reported improvements in the CVLT total both over time and by group, and showed significantly more use of memory strategies in the intervention compared with control. Chiaravalloti 2013 and Chiaravalloti 2019a showed a greater learning slope for the treatment group compared to the control on the CVLT-II (Delis 2000). Lincoln 2020 showed significant group differences in Everyday Memory Questionnaire (EMQ-p) at both six and 12 months, in the Selective Reminding Test (SRT) total recall at six months and delayed recall at 12 months with no other significant group differences. Maggio 2020 showed significant group improvement in both groups with greater improvement in the intervention group for Spatial Recall Test (SPART) and a significant Group*Time interaction. Messinis 2017 showed significant improvement for the intervention group in SRT and BVMT-R. Messinis 2020 showed significant improvement in everyday memory at immediate follow-up but this was not sustained at longer-term follow-up. Pedulla 2016 showed significant Group*Time interaction for six out of 10 subtests of the BRBN. Tesar 2005 reported improvements on the computer-aided card-sorting test (CKV), Drühe-Wienholt 1998, and the Mosaic Test of the Hamburg Wechsler Intelligence Test (HAWIE-R), Tewes 1991, for the treatment group. Chiaravalloti 2005 observed no significant difference between the treatment and control groups on their list-learning task (HVLT-R) (Benedict 1998), but on subgroup analysis, we observed significant improvement on this task for the moderate-to-severe memory-impaired subgroup, but not for other groups. However, this subgroup analysis was carried out only on the treatment group, which had 14 participants. Mendozzi 1998 reported improvement in the specific cognitive-retraining group on seven measures of memory (Spatial Span from the Corsi, Digit Span Forward and Backward, Visual Reproduction, and Paired Associates-Hard from the Italian Weschler Memory Scale (WMS), Wechsler 1945, and the LNNB, Golden 1980. There was an improvement in Digit Span Forward only in the non-specific cognitive rehabilitation group.

Outcome 1: Subjective memory measures

Fifteen studies included subjective measures of participants' memory functioning. Ten of these studies provided immediate outcomes (Chiaravalloti 2005; Chiaravalloti 2019b; Chmelařová 2020; Goodwin 2020; Goverover 2018a; Mani 2018; Mousavi 2018b; Perez-Martin 2017; Stuifbergen 2012; Stuifbergen 2018); 11 of these studies provided intermediate outcomes (Carr 2014; Chiaravalloti 2005; Chiaravalloti 2019b; das Nair 2012; Lincoln 2002; Lincoln 2020; Mani 2018; Mousavi 2018b; Shahpouri 2019; Stuifbergen 2012; Stuifbergen 2018); and five of these studies provided longer-term outcomes (Carr 2014; das Nair 2012; Lincoln 2002; Lincoln 2020; Stuifbergen 2018). We found small to moderate differences between groups for subjective reports of memory for immediate, intermediate and longer-term follow ups: (standardised mean difference (SMD) 0.32, 95% confidence interval (CI) 0.05 to 0.58; 568 participants, moderate-quality evidence) Analysis 1.1; (SMD 0.23, 95% CI 0.11 to 0.35; 1045 participants, high-quality evidence) Analysis 1.2; and (SMD 0.16, 95% CI 0.02 to 0.30; 775 participants, high-quality evidence) Analysis 1.3, respectively. The intervention group performed better than the control group at each follow-up.

Outcome 2: Objective verbal memory measures

Twenty-one studies included objective verbal memory measures of participants' memory functioning. Nineteen of these studies provided immediate outcomes (Arian Darestani 2020; Campbell 2016; Chiaravalloti 2005; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; Gich 2015; Goverover 2018a; Hancock 2015; Impellizzeri 2020; Messinis 2017; Messinis 2020; Pedulla 2016; Perez-Martin 2017; Rilo 2018; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005; Vilou 2020), six of these studies provided intermediate outcomes (Arian Darestani 2020; Campbell 2016; Lincoln 2020; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005), and four of these studies provided longer-term outcomes (Chiaravalloti 2019b; Lincoln 2020; Mattioli 2016; Stuifbergen 2018). We found small to moderate differences between groups for objective verbal reports of memory at immediate (SMD 0.40, 95% CI 0.22 to 0.58; 922 participants, low-quality evidence) Analysis 2.1 and intermediate follow-up (SMD 0.25, 95% CI 0.11 to 0.40; 753 participants, low-quality evidence) Analysis 2.2, but no little to no difference at longer-term follow-up (SMD 0.13, 95% CI -0.03 to 0.29; 619 participants, moderate-quality evidence) Analysis 2.3. The intervention group performed better than the control group at immediate and intermediate follow-up.

Outcome 3: Objective visual memory measures

Nineteen studies included objective visual measures of participants' memory functioning. Sixteen of these studies provided immediate outcomes (Campbell 2016; Chiaravalloti 2019a; Chiaravalloti 2019b; Chmelařová 2020; Goverover 2018a; Impellizzeri 2020; Maggio 2020; Messinis 2017; Messinis 2020; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005; Vilou 2020), six of these studies provided intermediate outcomes (Campbell 2016; Lincoln 2020; Naeeni Davarani 2020; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005), four of these studies provided longer-term outcomes (Chiaravalloti 2019b; Lincoln 2020; Mattioli 2016; Stuifbergen 2018). We found a moderate difference between groups for objective reports of visual memory at immediate follow-up (SMD 0.42, 95% CI 0.25 to 0.60; 799 participants, moderate-quality evidence) Analysis 3.1, but little to no between group differences at intermediate (SMD 0.20, 95% CI -0.11 to 0.50;751 participants, moderate-quality evidence) Analysis 3.2, and longer-term follow-up (SMD 0.12, 95% CI -0.13 to 0.37; 619 participants, high-quality evidence) Analysis 3.3. The intervention group performed better than the control group at immediate follow-up.

Outcome 4: Objective working memory measures

Thirteen studies included objective working measures of participants' memory functioning. Twelve of these studies provided immediate outcomes (Chiaravalloti 2019b; Chmelařová 2020; Impellizzeri 2020; Maggio 2020; Mousavi 2018a; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Rahmani 2020; Rilo 2018; Stuifbergen 2012; Stuifbergen 2018), eight of these studies provided intermediate outcomes (das Nair 2012;Huiskamp 2016; Lincoln 2020; Mousavi 2018a; Naeeni Davarani 2020; Rahmani 2020; Stuifbergen 2012; Stuifbergen 2018), five of these studies provided longer-term outcomes (Chiaravalloti 2019b; das Nair

2012; Lincoln 2020; Mattioli 2016; Stuifbergen 2018). We found a moderate difference between groups for objective reports of working memory at immediate follow-up (SMD 0.45, 95% CI 0.18 to 0.72; 655 participants, low-quality evidence) Analysis 4.1, but little to no between group differences at intermediate (SMD -0.16, 95% CI -0.09 to 0.40; 821 participants, moderate-quality evidence) Analysis 4.2, or longer-term follow-up (SMD 0.04, 95% CI -0.11 to 0.20; 665 participants, moderate-quality evidence) Analysis 4.3. The intervention group performed better than the control group at immediate follow-up.

Outcome 5: Information processing

In terms of 'other cognitive outcomes', the most frequently assessed cognitive domain was information processing. Nineteen studies included information processing measures. Fifteen studies reported immediate outcomes (Campbell 2016; Chiaravalloti 2019b; Chmelařová 2020; Hancock 2015; Messinis 2017; Messinis 2020; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Rahmani 2020; Rilo 2018; Solari 2004; Stuifbergen 2012; Stuifbergen 2018; Vilou 2020), eight studies reported intermediate outcomes (Campbell 2016; Hanssen 2015; Lincoln 2020; Naeeni Davarani 2020; Rahmani 2020; Solari 2004; Stuifbergen 2012; Stuifbergen 2018), five studies reported longer-term outcomes (Hanssen 2015; Lincoln 2020; Mattioli 2016; Pedulla 2016; Stuifbergen 2018). We found moderate between group differences for information processing measures at immediate (SMD 0.51, 95% CI 0.19 to 0.82; 808 participants, low-quality evidence) Analysis 5.1, and intermediate follow-up (SMD 0.27, 95% CI 0.00 to 0.54; 933 participants) Analysis 5.2, but little to no difference at longer-term follow-up (SMD 0.21, 95% CI -0.03 to 0.45; 723 participants, moderate-quality evidence) Analysis 5.3. The intervention group performed better than the control group at immediate and intermediate follow-up.

Outcome 6: Mood - Depression

Twenty-two studies included measures of depression. Sixteen of these studies provided immediate outcomes (Campbell 2016; Chiaravalloti 2005; Chiaravalloti 2013; Chmelařová 2020; Goodwin 2020; Goverover 2018a; Hancock 2015; Hildebrandt 2007; Impellizzeri 2020; Maggio 2020; Messinis 2017; Messinis 2020; Perez-Martin 2017; Solari 2004; Stuifbergen 2018; Tesar 2005), 10 of these studies provided intermediate outcomes (Campbell 2016; Carr 2014; Chiaravalloti 2005; das Nair 2012; Lincoln 2002; Lincoln 2020; Shahpouri 2019; Solari 2004; Stuifbergen 2018; Tesar 2005), and seven of these studies provided longer-term outcomes (Carr 2014; Chiaravalloti 2013; das Nair 2012; Lincoln 2002; Lincoln 2020; Mattioli 2016; Stuifbergen 2018). We found a moderate difference between groups for mood measures of depression at immediate (SMD 0.34, 95% CI 0.15 to 0.53; 853 participants, moderate-quality evidence) Analysis 6.1, but little to no difference at intermediate (SMD 0.20, 95% CI -0.06 to 0.45; 1003 participants, moderate-quality evidence) Analysis 6.2, or longer-term follow-up (SMD 0.15, 95% CI -0.04 to 0.34; 891 participants, high-quality evidence) Analysis 6.3. The intervention group performed better than the control group at immediate follow-up.

Outcome 7: Mood - Anxiety

Seven studies included measures of anxiety. Four of these studies provided immediate outcomes (Campbell 2016; Goodwin 2020; Goverover 2018a; Perez-Martin 2017), four of these studies provided intermediate outcomes (Campbell 2016; Carr 2014; das Nair 2012; Lincoln 2020), and three of these studies provided longer-term outcomes (Carr 2014; das Nair 2012; Lincoln 2020). We found little to no between group differences for mood measures of anxiety at immediate (SMD 0.29, 95% CI -0.01 to 0.59; 178 participants, high-quality evidence) Analysis 7.1, intermediate (SMD 0.16, 95% CI -0.15 to 0.46; 502 participants, high-quality evidence) Analysis 7.2, or longer-term follow-up (SMD 0.27, 95% CI -0.12 to 0.65; 448 participants. high-quality evidence) Analysis 7.3.

Outcome 8: Quality of life (QoL)

Eleven studies included QoL measures. Eight of these studies provided immediate outcomes (Goodwin 2020; Goverover 2018a; Hancock 2015; Maggio 2020; Messinis 2020; Perez-Martin 2017;Shahpouri 2019; Solari 2004), six of these studies provided intermediate outcomes (Carr 2014; Hanssen 2015; Lincoln 2002; Lincoln 2020; Shahpouri 2019; Solari 2004), five of these studies provided longer-term outcomes (Carr 2014; Hanssen 2015; Lincoln 2002; Lincoln 2020; Mattioli 2016). We found small to moderate between group differences for quality of life measures at immediate, intermediate, and longer-term follow ups: (SMD 0.42, 95% CI 0.15 to 0.68; 371 participants, high-quality evidence) Analysis 8.1, (SMD 0.30, 95% CI 0.02 to 0.58; 683 participants, high-quality evidence) Analysis 8.2, and (SMD 0.17, 95% CI 0.02 to 0.32; 687 participants, high-quality evidence) Analysis 8.3, respectively. The intervention group performed better than the control group at every follow-up.

Outcome 9: Functional abilities / Activities of daily living (ADL)

Six studies included ADL measures of participants' daily functioning. Four of these studies provided immediate outcomes (Campbell 2016; Chiaravalloti 2019a; Goverover 2018a; Stuifbergen 2018), four of these studies provided intermediate outcomes (Campbell 2016; das Nair 2012; Goverover 2018a; Stuifbergen 2018),and three of these studies provided longer-term outcomes (das Nair 2012; Lincoln 2002; Stuifbergen 2018). We found little to no between group differences for ADL at immediate, intermediate, and longer-term follow-ups: (SMD 0.02, 95% CI -0.26 to 0.29; 265 participants, high-quality evidence) Analysis 9.1, (SMD -0.06, 95% CI -0.36 to 0.24; 400 participants, high-quality evidence) Analysis 9.2, and (SMD -0.11, 95% CI -0.49 to 0.27; 369 participants, high-quality evidence) Analysis 9.3, respectively.

DISCUSSION

Summary of main results

In the last two decades, research groups globally have begun to address memory problems associated with multiple sclerosis (MS). However, the literature base examining the effectiveness of memory rehabilitation for MS has been weak. While single-case and uncontrolled studies have found memory rehabilitation to be effective in reducing memory or psychological problems, these results had not been consistently replicated in randomised controlled trials (RCTs). However, more recently, we have seen larger, more methodologically-robust trials published in this area.

We included 44 RCTs or quasi-randomised trials in this review. These studies were either memory rehabilitation studies or cognitive rehabilitation trials with a specific memory component that included a memory intervention. These trials were mostly of

relatively moderate quality, with many still not adhering to the Consolidated Standards of Reporting Trials (CONSORT) guidelines (Moher 2001). Descriptions of the randomisation protocol, blinding, and content of treatment and control groups were poor in approximately 50% of studies. Studies generally had modest sample sizes and used impairment-level outcome assessments to determine the effectiveness of the intervention.

Twenty-nine individual studies reported positive results on memory outcomes from their memory rehabilitation groups (Arian Darestani 2020; Chiaravalloti 2013; Chiaravalloti 2019b; Chmelařová 2020; De Luca 2019; Gich 2015; Goverover 2018a; Hildebrandt 2007; Impellizzeri 2020; Lincoln 2020; Maggio 2020; Mani 2018; Mattioli 2016; Mendozzi 1998; Messinis 2017; Messinis 2020; Mousavi 2018a; Mousavi 2018b; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Pusswald 2014; Rahmani 2020; Rilo 2018; Shahpouri 2019; Stuifbergen 2012; Stuifbergen 2018; Tesar 2005; Vilou 2020). However, these results need to be interpreted in the context of the methodological limitations and the measures used to assess effectiveness, which may have influenced the outcome. In fact, most of the studies that reported a positive memory outcome for the participants in the intervention group were also rated as having a high or unclear risk of bias in at least one area, with the exceptions of four studies (Lincoln 2020; Shahpouri 2019; Stuifbergen 2012; Vilou 2020). One well-designed large study with larger sample size (Lincoln 2020) found a significant effect for memory outcomes at six months follow-up but did not find evidence of this effect at 12 months follow-up, suggesting that the longevity or sustainability of the positive effects of the intervention cannot necessarily be expected in the long run.

Between group differences were found for quality of life outcomes in favour of the intervention group compared to the control group at each follow-up point, suggesting that memory rehabilitation can lead to positive change in the overall perception of quality of life of people with MS. It should be noted that this positive finding has not been observed in many other cognitive rehabilitation reviews for people with MS (e.g. Rosti-Otajärvi 2014), or reviews investigating cognitive rehabilitation in other cohorts such as post-stroke patients (e.g. das Nair 2016a). This could be due to more recent trials having a broader focus on ‘impact’ of cognitive problems on MS, and therefore, likely to affect quality of life, while older trials mainly focused solely on memory impairments. It could also be that memory problems are more detrimental to the quality of life of people with MS compared to people with other neurological conditions and therefore, the tools to cope and self-manage their problems (that they are taught during memory and attention rehabilitation) lead to a greater improvement in their quality of life compared to other patient groups.

The results of this review suggest there is substantial evidence to support the effectiveness of memory rehabilitation on subjective memory measures at immediate follow-up in favour of the intervention condition, and this result is sustained at intermediate and longer-term follow-ups of up to one year. This is a significant change compared to the previous version of this review, which found no evidence to support that memory rehabilitation had a positive effect on subjective memory measures. In the current review, between group difference favouring the intervention group were also seen in the following outcomes: objective measures of verbal memory, both immediate and intermediate follow-ups; objective measures of visual memory at immediate follow-up; objective reports of working memory at immediate follow-up; information processing at immediate and intermediate follow-ups; mood measures of depressive symptoms at immediate follow-up; quality of life measures at immediate, intermediate and longer-term follow-ups. Little to no between group differences were found in activities of daily living measures or measures of anxiety.

One well-designed large study with a large sample size (Lincoln 2020) found a significant effect for memory outcomes at six-month follow-up but did not find evidence of this effect at 12-month follow-up, suggesting that the longevity or sustainability of the positive effects of the intervention cannot necessarily be expected in the long run. This supports the overall trend of these results in that, there were only two outcomes that maintained their significant effects at longer-term follow-up, suggesting that once the core intervention has been completed, maintenance plans (such as booster sessions) should be put in place to ensure the techniques learnt during the intervention are not forgotten or inconsistently used over time.

High heterogeneity (I^2>=50%) was seen in four statistically significant outcomes (working memory at immediate (I^2 = 62%) follow-up, quality of life at intermediate (I^2 = 55%) follow-up, and information processing at immediate (I^2 = 77%) and intermediate (I^2 = 69%) follow-up) and thus, these findings need to be treated with caution and explored further. Firstly, there does not appear to be one or two primary studies contributing to the increased heterogeneity for working memory measures at immediate follow-up, therefore, this could be due to the wide variation in the type of intervention used by each study. Two studies (Stuifbergen 2012; Stuifbergen 2018) used the same Memory Attention and Problem-Solving Skills in Multiple Sclerosis (MAPSS-MS) intervention, whereas the other 10 studies all used different interventions from each other. We found a large variation in both the type of methods used, e.g. computerised versus face-to-face, and the frequency at which the intervention was delivered, e.g. ranging from four to 12 weeks in duration and from once to six times per week. These results suggest that variation in both type of intervention and frequency of delivery contributed to the high heterogeneity. This theory is supported by the sensitivity analysis which shows that the heterogeneity drops to 0% when all but the two studies that used the same intervention methods are removed.

Secondly, for the high heterogeneity for information processing outcomes measured at immediate follow-up, it appears that there are three main studies contributing towards this (Campbell 2016; Messinis 2020; Naeeni Davarani 2020). All three studies used RehaCom software for their interventions which took place at home and there was some variation in the frequency of sessions between each study. The type of intervention used in these studies may have contributed to the high heterogeneity however, without a meta-regression we cannot be certain of the cause and thus these results should be interpreted with caution.

Lastly, for the high heterogeneity seen in quality of life outcomes measured at intermediate follow-up, one study appeared to contribute towards this (Shahpouri 2019). One possible cause could be that the outcomes were measured quote: "within 3 months after cognitive rehabilitation therapy" (p. 113). This suggests that some participants may have had their outcomes assessments immediately after treatment and others three months later, which could account for the large clinical variance in outcome scores.

However, these results must be viewed in relation to the quality of the evidence for this outcome, with the GRADE rating showing as low for information processing at both immediate and intermediate follow-up, low for working memory at immediate follow-up but high for quality of life at intermediate follow-up (GRADE Working Group 2004). Furthermore, improvements in outcomes are only maintained at follow-up for subjective memory measures and quality of life, which suggests that regular booster sessions of cognitive rehabilitation are necessary to maintain the improvements made and without them participants appear to revert to where they started. The degree to which this has the potential to generalise to everyday life, given the varying ecological validity of these tests, is questionable. However, it is important to note that the methodological quality of studies included in this review has improved compared to the previous review.

Overall completeness and applicability of evidence

The size of the literature-base examined in this review allowed us to address the research questions in as much depth as possible. The variety of outcomes in the trials enabled us to investigate domain-specific memory such as visual, verbal and working, thus this review not only shows the positive effects for general memory but also identified which domains are being improved by the rehabilitation intervention programmes and which of these improvements are maintained. This review fully investigated all types of studies, participants, interventions and outcome measures as stated in the methods. The positive results in trials using computerised interventions have important implications for clinical practice in the current COVID-19 pandemic, as cognitive rehabilitation may have to be delivered virtually for the foreseeable future. This review examined the evidence from RCTs and quasi-RCTs and found evidence to suggest that memory rehabilitation is effective in improving memory performance on subjective, objective (verbal, visual and working memory) assessments across immediate and intermediate follow-ups, and quality of life in the immediate, intermediate, and longer-term, and reducing depression (but only immediately after the intervention). However, this evidence should be interpreted in the context of the methodological quality as reported in the Summary of findings 1 before it is applied to a clinical setting.

Quality of the evidence

We identified 44 RCTs of memory rehabilitation for people with MS, and all but five had small sample sizes (Charvet 2017; Hanssen 2015; Lincoln 2002; Lincoln 2020; Stuifbergen 2018). However, studies included in this review were more methodologically sound than the memory rehabilitation RCTs included in systematic reviews of stroke or traumatic brain injury literature (das Nair 2007). Despite this, the CONSORT statement and guidelines were not always followed in these trials.

The randomisation protocol was inadequate and was poorly reported for 15 studies (Arian Darestani 2020; Chiaravalloti 2005; Chmelařová 2020; Ernst 2015; Gich 2015; Hanssen 2015; Hildebrandt 2007; Jønsson 1993; Mendozzi 1998; Naeeni Davarani 2020; Pedulla 2016; Perez-Martin 2017; Pusswald 2014 Rahmani 2020; Tesar 2005). Gich 2015, Hanssen 2015, and Tesar 2005 did not clearly mention how the randomisation list was created or what procedures were undertaken; Jønsson 1993 used closed envelopes but did not mention who created the random lists; Chiaravalloti 2005 employed odd-even random allocation; and Hildebrandt 2007 and Pusswald 2014 used alternating allocation. These two latter forms of allocation are not always considered acceptable in RCTs (Glanville 2006), but are classed by Cochrane as a quasi-randomised trial (Higgins 2019) and were therefore included in this review. Mendozzi 1998 randomised only half the sample, with no stated random generation method. Twenty-seven studies reported their randomisation protocols adequately (Campbell 2016; Carr 2014; Charvet 2017; Chiaravalloti 2013; Chiaravalloti 2019a; Chiaravalloti 2019b; das Nair 2012; De Luca 2019; Ernst 2018; Goodwin 2020; Goverover 2018a; Hancock 2015; Huiskamp 2016; Impellizzeri 2020; Lincoln 2002; Lincoln 2020; Maggio 2020; Messinis 2017; Messinis 2020; Mousavi 2018a; Mousavi 2018b; Rilo 2018; Shahpouri 2019; Solari 2004; Stuifbergen 2012; Stuifbergen 2018; Vilou 2020). The 29 studies we have added in this update have improved in terms of quality of reporting of trials however, more work is needed to ensure that trialists follow the CONSORT statement (Moher 2001).

Furthermore, given that memory rehabilitation is a complex intervention (Craig 2008), much more detail is required about what participants experience in both the intervention and the control arms of the trial. Indeed, the description of the interventions was adequate in most studies, however the control groups were much less well-described. Recently published guidelines such as the Template for Intervention Description and Replication (TIDieR) and the Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline (Hoffman 2014; Möhler 2015), alongside more specific guidance for memory rehabilitation (Martin 2015), may help improve the quality of reporting of trials of complex interventions.

Given the complex nature of the intervention, it is important to determine whether the intervention was delivered as intended. Only four studies (Carr 2014; Lincoln 2020; Stuifbergen 2012; Stuifbergen 2018) out of 44 reported whether a fidelity assessment was completed. Where it was assessed, authors found that the intervention was delivered with fidelity. Future trials should consider including fidelity assessments.

Inclusion and exclusion criteria were relatively well-defined. While most studies described the flow of participants through the trial, one did not (Tesar 2005), and only 14 of the 44 studies had flowcharts (Carr 2014; Campbell 2016; Chiaravalloti 2005; Chiaravalloti 2013; das Nair 2012; Gich 2015; Goodwin 2020; Hancock 2015; Hanssen 2015; Lincoln 2002; Lincoln 2020; Pusswald 2014; Solari 2004; Stuifbergen 2012).

Because MS is found in demographically diverse populations, we expected to see better description of the samples. Only five out of 44 papers described the ethnicity of the sample, while 38 out of 44 papers described the level of education. No studies reported whether or not participants were drawn from economically-disadvantaged groups or whether they had co-morbid conditions. While these factors could be balanced out through randomisation, we need to know whether the effects of the intervention are the same for these groups. Future trials should collect and report these details. Furthermore, while many studies recruited samples with people with different types of MS, we note that several studies only included people with relapsing remitting multiple sclerosis (RRMS). We recommend that future trials consider including people with other MS subtypes also, and outcomes described separately for each subtype.

Most trials opted to use impairment-level measures or tests with modest ecological validity and minimal chance of generalisation of treatment effects to activities of daily living. Fifteen studies employed subjective measures of memory (Carr 2014; Chiaravalloti 2005; Chiaravalloti 2019a; Chiaravalloti 2019b; Chmelařová 2020; das Nair 2012; Goodwin 2020; Goverover 2018a; Lincoln 2002; Lincoln 2020; Mani 2018; Mousavi 2018b; Perez-Martin 2017; Shahpouri 2019; Stuifbergen 2012), which is a big improvement from the five studies in the last update as these measures have some degree of ecological validity and were activity-level measures. However, these are prone to subjective reporting biases common to most Patient-Reported Outcome Measures (PROMs). Furthermore, the cultural appropriateness of outcomes has improved since the previous review, with more studies including translated and adapted assessment tools such as the GVLT which is the Greek adaptation of the California Verbal Learning Test (CVLT-II).

There has also been a shift in focus in some of the more recent trials from assessing only cognitive outcomes to including other outcomes such as mood and quality of life. This, we believe, is a positive step forward in memory and cognitive rehabilitation. This review highlights the importance of not only including these quality of life measures as key outcomes, as in the Lincoln 2020 trial. Only three studies assessed adverse events following memory rehabilitation (Chiaravalloti 2013; Chiaravalloti 2019a; Lincoln 2020). While, the likelihood of such adverse events is remote, trials should assess them to be certain of this.

Both parametric and nonparametric statistical tests were used to compare groups. Change scores were compared in six studies (Campbell 2016; Chiaravalloti 2005; Chiaravalloti 2013; Gich 2015; Hanssen 2015; Stuifbergen 2012), and all studies were concerned with significance testing. Contrary to the previous update of this review, the majority of the newly included studies included P values in their reporting of outcomes as opposed to the seven that included them previously (Carr 2014; das Nair 2012; Gich 2015; Hancock 2015; Lincoln 2002; Pusswald 2014; Solari 2004), with many trials providing all P values in tables that were readily accessible in the papers and online as supplementary information (Campbell 2016; Chiaravalloti 2013). Most studies also mentioned confidence intervals and often reported the post-hoc tests or statistical corrections or adjustments performed on their data. Eight studies used intention-to-treat analysis (Carr 2014; das Nair 2012; Goodwin 2020; Hildebrandt 2007; Lincoln 2002; Lincoln 2020; Solari 2004; Stuifbergen 2012).

During risk of bias assessment, we observed that some studies stated that they were “double-blind” studies without justifying how they were in fact double blind. This resulted in these studies being rated as high risk of bias. Such double-blind studies were typically those where computerised memory rehabilitation was the intervention being tested. Even in these situations where participants could potentially be blinded, it was not clear how different the computerised rehabilitation was from the computerised control group. Therefore, it was difficult to determine whether the participants were truly blinded. In some instances, the study authors reported that either participants or therapists delivering the intervention were blinded to group allocation, but from the study description, it was not always clear how this could have been the case. In future, we would strongly encourage authors to be more explicit in describing the blinding procedures used.

One limitation of this review was that we could only obtain information on whether the studies used intention-to-treat or per-protocol analyses for eight studies (Carr 2014; das Nair 2012; Hildebrandt 2007; Lincoln 2002; Lincoln 2020; Solari 2004; Stuifbergen 2012; Stuifbergen 2018), therefore we could not complete a sensitivity analysis of intention-to-treat in comparison with per-protocol analysis. We were able to conduct a sensitivity analysis comparing studies judged to be at low risk of bias to all included studies, however, we were unable to run this analysis for four outcomes due to a lack of studies with low risk of bias in every area (see Table 2). This suggests that there could be a correlation between trials that measure their outcomes immediately post-intervention and high risk of bias within the methodology. Our interpretation of the sensitivity analysis suggests that while the quality of the trials did not affect most outcomes, some differences were observed at immediate follow-up, with studies with higher risk of bias inflating the overall effect size estimates for these outcomes, and the test of overall effect changing from being statistically significant to not significant when studies at high risk of bias were excluded. This suggests that lower-quality studies may have positively influenced the outcomes, however, this could also be because only a few studies that measured immediate outcomes had low risk of bias in every area, and therefore, these results should be interpreted with caution. Furthermore, removing the studies with high risk of bias during this analysis often led to a reduction in heterogeneity. This could suggest an association between studies that have high risk of bias and increased heterogeneity. However, it is more likely that the heterogeneity was caused by wide variation in both type and frequency of intervention. There also appeared to be an association between studies that measure longer-term outcomes and low risk of bias.

We conducted a separate sensitivity analysis for the studies where standard deviations were inputted and found no clinical differences between the sensitivity analysis and the primary analysis, suggesting that inputting the standard deviations had no significant effect. Only one study had a large sample size and sufficient data available to complete a subgroup analysis (Lincoln 2002). A subgroup meta-analysis on the basis of type of MS will therefore need to be completed in a future review update when more studies become available.

Potential biases in the review process

Two of the review authors were lead investigators for three of the included studies (das Nair 2012, Lincoln 2002, Lincoln 2020), and named authors on another included study (Carr 2014), but to mitigate bias, we had multiple review authors who independently appraised the methodological quality of these studies. We only searched for papers in English, and we could only include mixed-diagnosis studies where separate data for those participants with MS were provided. Therefore, there may be more data available that we did not have access to. There were also potential overlaps between attention and memory retraining, where an intervention could be described as attention when it actually addressed memory, so we may have missed some trials. To mitigate this, we checked papers at full-text review to ensure that they were not excluded if a memory component was presented as part of the treatment. Finally, we searched GreyNet and the EThOS databases; however, we are not sure of the comprehensiveness of these, thus

creating the possibility of further relevant grey literature that was not obtained via the searches.

Agreements and disagreements with other studies or reviews

This review complements the '*Psychological interventions for multiple sclerosis*' intervention review (Thomas 2006). In one of their mini-reviews, Thomas 2006 found quote: "some evidence of effectiveness of cognitive rehabilitation on cognitive outcomes, although this was difficult to interpret because of the large number of outcome measures used". Their interpretations have therefore been based on a narrative review of results from individual studies. The Thomas 2006 review covered interventions that were not specific to 'memory rehabilitation', however, their findings related to effectiveness of interventions to help people with cognitive impairments were inconclusive.

Similarly, the Rosti-Otajärvi 2014 review found evidence that memory span, working memory, and delayed memory were significantly improved for the intervention compared with the control group. However, their review found no significant differences between intervention and control for emotional functions, whereas this review has found some significant differences, notably improved mood on depression scales and quality of life. Any discrepancies are likely due to the differences in inclusion criteria, as this review was specific to memory rehabilitation, or a cognitive rehabilitation with a memory component, whereas the Rosti-Otajärvi 2014 review evaluated a much larger breadth of neuropsychological interventions and outcomes.

The Goverover 2018b review found promising results to support cognitive rehabilitation for improving memory function and stated that there had been substantial progress made in increasing the number of cognitive rehabilitation trials to allow for practise recommendations. However, they also suggest, like we do, there is still much work to be done to optimise cognitive rehabilitation potential by applying the most rigorous methodology to ensure the quality of evidence is as high as possible.

AUTHORS' CONCLUSIONS

Implications for practice

In the last two decades increasing attention has been given to memory problems as a frequent complaint for people with multiple sclerosis (MS). Memory rehabilitation programmes are offered to some people with MS, but their effectiveness has been questionable. Small studies using a mixture of internal and external memory aids, errorless learning, and environmental manipulation have yielded positive results with many these using computer-delivered interventions. Large randomised controlled trials (RCTs) use mostly group-based and computer-delivered interventions and have also yielded positive results with improvements in outcomes seen in these trials often being maintained at follow-up. The positive results in trials using computerised interventions have important implications for clinical practice in the current COVID-19 pandemic, as cognitive rehabilitation may have to be delivered virtually for the foreseeable future. This review examined the evidence from RCTs and quasi-RCTs and found evidence to suggest that memory rehabilitation is effective in improving memory performance on subjective, objective (verbal, visual and working memory) assessments across immediate and intermediate follow-ups, and quality of life in the immediate, intermediate, and longer-term, and reducing depression (but only immediately after the intervention). Memory rehabilitation did not have an effect, at any time point, on activities of daily living or anxiety. There appeared to be no indication of harm caused by the interventions, but several studies did not routinely report adverse effects.

Implications for research

The research base from which to draw inferences for clinical practice regarding the effectiveness of memory rehabilitation for MS has improved since the previous review (das Nair 2016b). RCTs tended to be of modest sample size, and mostly used impairment-level outcome measures, which have limited value in assessing the functional effects of neurorehabilitation. These trials did not always adhere to the CONSORT guidelines (Moher 2001), which makes it difficult to get a full and true picture of the studies, and therefore limits the reader from making an informed decision regarding the fidelity of their conclusions. Missing information from such reports also make collating information for a meta-analysis difficult. Furthermore, results from 'positive' trials may be difficult to implement in clinical practice if sufficient details about the actual intervention are not clearly spelt out. The TiDieR checklist and other more specific guidance for reporting of memory rehabilitation trials may help improve the quality of reporting trials of complex interventions (Hoffman 2014; Martin 2015). Given that memory rehabilitation is a complex intervention and four studies assessed the fidelity of the intervention, we would encourage trialists to consider intervention fidelity assessments in future memory rehabilitation trials. The results of this review indicate that more research is required to arrive at a definitive answer as to whether or not memory rehabilitation for MS is effective in reducing activity limitations or restrictions to participation. It also highlights the need for more well-conceptualised, executed, and reported RCTs of memory rehabilitation that take into consideration the issues raised in this review.

ACKNOWLEDGEMENTS

We would like to thank Wendy Stanton, Faculty Team leader, Greenfield Medical Library at the University of Nottingham who helped us develop the search strategy for this review.

REFERENCES

References to studies included in this review

Arian Darestani 2020 *{published data only}*

Arian Darestani A, Naeeni Davarani M, Hassani-Abharian P, Zarrindast MR, Nasehi M. The therapeutic effect of treatment with RehaCom software on verbal performance in patients with multiple sclerosis. *Journal of Clinical Neuroscience: official journal of the Neurosurgical Society of Australasia* 2020;**72**:93-7. [PMID: 31937503]

Campbell 2016 *{published data only}*

Campbell J, Langdon D, Cercigiani M, Rashid W. A randomised controlled trial of efficacy of cognitive rehabilitation in multiple sclerosis: a cognitive, behavioural, and MRI study. *Neural Plasticity* 2016;**2016**:1-9.

Carr 2014 *{published data only}*

Carr SE, das Nair R, Schwartz AF, Lincoln NB. Group memory rehabilitation for people with multiple sclerosis: a feasibility randomized controlled trial. *Clinical Rehabilitation* 2014;**28**(6):552-61.

Charvet 2017 *{published data only}*

Charvet LE, Yang J, Shaw MT, Sherman K, Haider L, Xu J, et al. Cognitive function in multiple sclerosis improves with telerehabilitation: results from a randomized controlled trial. *PLOS One* 2017;**12**(5):e0177177. [PMID: 28493924]

Chiaravalloti 2005 *{published and unpublished data}*

Chiaravalloti ND, DeLuca J, Moore NB, Ricker JH. Treating learning impairments improves memory performance in multiple sclerosis: a randomized clinical trial. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2005;**11**:58-68.

Chiaravalloti 2013 *{published data only}*

Chiaravalloti ND, Moore NB, Nikelshpur OM, DeLuca J. An RCT to treat learning impairment in multiple sclerosis. *Neurology* 2013;**81**(24):2066-72.

Chiaravalloti 2019a *{published and unpublished data}*

Chiaravalloti ND, Moore NB, Weber E, DeLuca J. The application of Strategy-based Training to Enhance Memory (STEM) in multiple sclerosis: a pilot RCT. *Neuropsychological Rehabilitation* 2019;**29**:1-24. [PMID: 31752604]

Chiaravalloti 2019b *{published and unpublished data}*

Chiaravalloti ND, Moore NB, DeLuca J. The efficacy of the modified Story Memory Technique in progressive MS. *Multiple Sclerosis* 2019;**26**(3):354-62. [PMID: 30741103]

Chmelařová 2020 *{published and unpublished data}*

Chmelařová D, Fiala L, Dostál M, Lenz J. Intensive computer-assisted cognitive rehabilitation in persons with multiple sclerosis - results of a 12-week randomized study. *Czech and Slovak Neurology and Neurosurgery* 2020;**83/116**(4):408-15.

das Nair 2012 *{published and unpublished data}*

das Nair R, Lincoln NB. Evaluation of Rehabilitation of Memory in people with Neurological Disabilities (ReMiND): a randomized controlled trial. *Clinical Rehabilitation* 2012;**26**(10):894-903.

De Luca 2019 *{published data only}*

De Luca R, Russo M, Gasparini S, Leonardi S, Foti Cuzzola M, Sciarrone F, et al. Do people with multiple sclerosis benefit from PC-based neurorehabilitation? A pilot study. *Applied Neuropsychology. Adult* 2019;**26**:1-9. [PMID: 31414887]

Ernst 2015 *{published data only}*

Ernst A, Blanc F, De Seze J, Manning L. Using mental visual imagery to improve autobiographical memory and episodic future thinking in relapsing-remitting multiple sclerosis patients: a randomised-controlled trial study. *Restorative Neurology and Neuroscience* 2015;**33**(5):621-38. [PMID: 26410206]

Ernst 2018 *{published data only}*

Ernst A, Sourty M, Roquet D, Noblet V, Gounot D, Blanc F, et al. Benefits from an autobiographical memory facilitation programme in relapsing-remitting multiple sclerosis patients: a clinical and neuroimaging study. *Neuropsychological Rehabilitation* 2018;**28**(7):1110-30. [PMID: 27718890]

Gich 2015 *{published data only}*

Gich J, Freixanet J, García R, Vilanova J, Genís D, Silva Y, et al. A randomized, controlled, single-blind, 6-month pilot study to evaluate the efficacy of MS-Line!: a cognitive rehabilitation programme for patients with multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2015;**21**:1-12.

Goodwin 2020 *{published data only}*

Goodwin RA, Lincoln NB, das Nair R, Bateman A. Evaluation of NeuroPage as a memory aid for people with multiple sclerosis: a randomised controlled trial. *Neuropsychological Rehabilitation* 2020;**30**(1):15-31. [PMID: 29557731]

Goverover 2018a *{published data only}*

Goverover Y, Chiaravalloti N, Genova H, DeLuca J. A randomized controlled trial to treat impaired learning and memory in multiple sclerosis: the self-GEN trial. *Multiple sclerosis (Houndmills, Basingstoke, England)* 2018;**24**(8):1096-104. [PMID: 28485659]

Hancock 2015 *{published data only}*

Hancock LM, Bruce JM, Bruce AS, Lynch SG. Processing speed and working memory training in multiple sclerosis: a double-blinded randomized controlled pilot study. *Journal of Clinical and Experimental Neuropsychology* 2015;**37**(2):113-27.

Hanssen 2015 *{published data only}*

Hanssen K, Beiske A, Landrø N, Hofoss D, Hessen E. Cognitive rehabilitation in multiple sclerosis: a randomised controlled trial. *Acta Neurologica Scandinavica* 2015;**8**:1-11.

Hildebrandt 2007 *{published and unpublished data}*

Hildebrandt H, Lanz M, Hahn HK, Hoffmann E, Schwarze B, Schwendemann G, et al. Cognitive training in MS: Effects and relation to brain atrophy. *Restorative Neurology and Neuroscience* 2007;**25**:33-43.

Huiskamp 2016 *{published data only}*

Huiskamp M, Dobryakova E, Wylie GD, DeLuca J, Chiaravalloti ND. A pilot study of changes in functional brain activity during a working memory task after mSMT treatment: the MEMREHAB trial. *Multiple Sclerosis and Related Disorders* 2016;**7**:76-82. [PMID: 27237764]

Impellizzeri 2020 *{published and unpublished data}*

Impellizzeri F, Leonardi S, Latella D, Maggio MG, Foti Cuzzola M, Russo M, et al. An integrative cognitive rehabilitation using neurologic music therapy in multiple sclerosis: a pilot study. *Medicine* 2020;**99**(4):e18866. [PMID: 31977888]

Jønsson 1993 *{published data only}*

Jønsson A, Korfitzen EM, Heltberg A, Ravnborg MH, Byskov-Ottosen E. Effects of neuropsychological treatment in patients with multiple sclerosis. *Acta Neurologica Scandinavica* 1993;**131**:394-400.

Lincoln 2002 *{published and unpublished data}*

Lincoln NB, Dent A, Harding J, Weyman N, Nicholl C, Blumhardt LD, et al. Evaluation of cognitive assessment and cognitive intervention for people with multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* 2002;**72**:93-8.

Lincoln 2020 *{published data only}*

Lincoln NB, Bradshaw LE, Constantinescu CS, Day F, Drummond AE, Fitzsimmons D, et al. Group cognitive rehabilitation to reduce the psychological impact of multiple sclerosis on quality of life: the CRAMMS RCT. Health Technology Assessment (Winchester, England) 2020;**24**(4):1-182. [PMID: 31934845]

Maggio 2020 *{published data only}*

Maggio MG, De Luca R, Manuli A, Buda A, Foti Cuzzola M, Leonardi S, et al. Do patients with multiple sclerosis benefit from semi-immersive virtual reality? A randomized clinical trial on cognitive and motor outcomes. *Applied Neuropsychology. Adult* 2020;**30**:1-7. [PMID: 31920097]

Mani 2018 *{published data only}*

Mani A, Chohedri E, Ravanfar P, Mowla A, Nikseresht A. Efficacy of group cognitive rehabilitation therapy in multiple sclerosis. *Acta neurologica Scandinavica* 2018;**137**(6):589-97. [PMID: 29411360]

Mattioli 2016 *{published and unpublished data}*

Mattioli F, Bellomi F, Stampatori C, Provinciali L, Compagnucci L, Uccelli A, et al. Two years follow up of domain specific Cognitive training in relapsing remitting multiple sclerosis: a randomized clinical trial. *Frontiers in Behavioral Neuroscience* 2016;**10**:28. [PMID: 26941630]

Mendozzi 1998 *{published and unpublished data}*

Mendozzi L, Pugnetti L, Motta A, Barbieri E, Gambini E, Cazzullo CL. Computer-assisted memory retraining of patients with multiple sclerosis. *Italian Journal of Neurological Sciences* 1998;**19**:S431-8.

Messinis 2017 *{published data only}*

Messinis L, Nasios G, Kosmidis MH, Zampakis P, Malefaki S, Ntoskou K, et al. Efficacy of a computer-assisted cognitive rehabilitation intervention in relapsing-remitting multiple sclerosis patients: a multicenter randomized controlled trial. *Behavioural Neurology* 2017;**2017**:5919841. [PMID: 29463950]

Messinis 2020 *{published data only}*

Messinis L, Kosmidis MH, Nasios G, Konitsiotis S, Ntoskou A, Bakirtzis C, et al. Do Secondary Progressive Multiple Sclerosis patients benefit from Computer- based cognitive neurorehabilitation? A randomized sham controlled trial. *Multiple sclerosis and related disorders* 2020;**39**:101932. [PMID: 31927200]

Mousavi 2018a *{published data only}*

Mousavi S, Zare H, Etemadifar M, Taher Neshatdoost H. Memory rehabilitation for the working memory of patients with multiple sclerosis (MS). *Journal of Clinical and Experimental Neuropsychology* 2018;**40**(4):405-10. [PMID: 28737452]

Mousavi 2018b *{published data only}*

Mousavi S, Zare H, Etemadifar M. Evaluating the effectiveness of cognitive rehabilitation on everyday memory in multiple sclerosis patients. *Neuropsychological Rehabilitation* 2018;**30**(6):1013-23. [PMID: 30353773]

Naeeni Davarani 2020 *{published data only}*

Naeeni Davarani M, Arian Darestani A, Hassani-Abharian P, Vaseghi S, Zarrindast MR, Nasehi M. RehaCom rehabilitation training improves a wide-range of cognitive functions in multiple sclerosis patients. *Applied Neuropsychology. Adult* 2020;**30**:1-11. [PMID: 32368936]

Pedulla 2016 *{published data only}*

Pedulla L, Brichetto G, Tacchino A, Vassallo C, Zaratin P, Battaglia MA, et al. Adaptive vs. non-adaptive cognitive training by means of a personalized App: a randomized trial in people with multiple sclerosis. *Journal of Neuroengineering and Rehabilitation* 2016;**13**(1):88. [PMID: 27716336]

Perez-Martin 2017 *{published data only}*

Perez-Martin MY, González-Platas M, Eguía-Del Río P, Croissier-Elías C, Jiménez Sosa A. Efficacy of a short cognitive training program in patients with multiple sclerosis. *Neuropsychiatric Disease and Treatment* 2017;**13**:245-52. [PMID: 28223806]

Pusswald 2014 *{published data only}*

Pusswald G, Mildner C, Zebenholzer K, Auff E, Lehrner J. A neuropsychological rehabilitation program for patients with multiple sclerosis based on the model of the ICF. *NeuroRehabilitation* 2014;**35**:519-27.

Rahmani 2020 *{published data only}*

Rahmani M, Rahimian Boogar I, Talepasand S, Nokani M. Comparing the effectiveness of computer-based, manual-based, and combined cognitive rehabilitation on cognitive functions in relapsing-remitting multiple sclerosis patients. *Basic and Clinical Neuroscience* 2020;**11**(1):99-110. [PMID: 32483480]

Rilo 2018 *{published data only}*

Rilo O, Peña J, Ojeda N, Rodríguez-Antigüedad A, Mendibe-Bilbao M, Gómez-Gastiasoro A, et al. Integrative group-based cognitive rehabilitation efficacy in multiple sclerosis: a randomized clinical trial. *Disability and Rehabilitation* 2018;**40**(2):208-16. [PMID: 27927036]

Shahpouri 2019 *{published data only}*

Shahpouri MM, Barekatain M, Tavakoli M, Sanaei S, Shaygannejad V. Evaluation of cognitive rehabilitation on the cognitive performance in multiple sclerosis: a randomized controlled trial. *Journal of Research in Medical Sciences : the official journal of Isfahan University of Medical Sciences* 2019;**24**:110. [PMID: 31949461]

Solari 2004 *{published and unpublished data}*

Solari A, Motta A, Mendozzi L, Pucci E, Forni M, Mancardi G, et al. Computer-aided retraining of memory and attention in people with multiple sclerosis: a randomized, double-blind controlled trial. *Journal of the Neurological Sciences* 2004;**222**:99-104.

Stuifbergen 2012 *{published data only}*

Stuifbergen AK, Becker H, Perez F, Morison J, Kullberg V, Todd A. A randomized controlled trial of a cognitive rehabilitation intervention for persons with multiple sclerosis. *Clinical Rehabilitation* 2012;**26**(10):882-93.

Stuifbergen 2018 *{published data only}*

Stuifbergen AK, Becker H, Perez F, Morrison J, Brown A, Kullberg V, et al. Computer-assisted cognitive rehabilitation in persons with multiple sclerosis: results of a multi-site randomized controlled trial with six month follow-up. *Disability and Health Journal* 2018;**11**(3):427-34. [PMID: 29477372]

Tesar 2005 *{published and unpublished data}*

Tesar N, Bandion K, Baumhackl U. Efficacy of a neuropsychological training programme for patients with multiple sclerosis, a randomised controlled trial. *Wiener Klinische Wochenschrift* 2005;**117**(21-2):747-54.

Vilou 2020 *{published data only}*

Vilou I, Bakirtzis C, Artemiadis A, Ioannidis P, Papadimitriou M, Konstantinopoulou E, et al. Computerized cognitive rehabilitation for treatment of cognitive impairment in multiple sclerosis: an explorative study. *Journal of Integrative Neuroscience* 2020;**19**(2):341-7. [PMID: 32706198]

References to studies excluded from this review

Aguirre 2019 *{published data only}*

Aguirre N, Cruz-Gómez ÁJ, Miró-Padilla A, Bueichekú E, Broseta Torres R, Ávila C, et al. Repeated working memory training improves task performance and neural efficiency in multiple sclerosis patients and healthy controls. *Multiple Sclerosis International* 2019;**2019**:2657902. [PMID: 31139470]

Aisen 1994 *{published data only}*

Aisen ML, Iverson D, Schwalbe C, Weaver B, Aisen PS. Falls on a neurorehabilitation unit: reassessment of a prevention program. *The Journal of the American Paraplegia Society* 1994;**17**(4):179-82.

Akhtar 2006 *{published data only}*

Akhtar S, Moulin C, Bowie P. Are people with mild cognitive impairment aware of the benefits of errorless learning? *Neuropsychological Rehabilitation* 2006;**16**(3):329-46.

Aldrich 1995 *{published data only}*

Aldrich FK. Can Implicit Memory Be Exploited to Facilitate the Learning of Novel Associations?. Brighton, UK: University of Sussex, 1995.

Allen 1998 *{published data only}*

Allen DN, Goldstein G, Heyman RA, Rondinelli T. Teaching memory strategies to persons with multiple sclerosis. *Journal of Rehabilitation Research and Development* 1998;**35**(4):405-10.

Amato 2014 *{published data only}*

Amato M, Goretti B, Viterbo R, Portaccio E, Niccolai C, Hakiki B et al. Computer-assisted rehabilitation of attention in patients with multiple sclerosis: results of a randomized, double-blind trial. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2014;**20**(1):91-8.

Barbarulo 2018 *{published data only}*

Barbarulo AM, Lus G, Signoriello E, Trojano L, Grossi D, Esposito M, et al. Integrated cognitive and neuromotor rehabilitation in multiple sclerosis: a pragmatic study. Frontiers in Behavioral Neuroscience 2018;**12**:196.

Barker 2019 *{published data only}*

Barker L, Healy BC, Chan E, Leclaire K, Glanz BI. A pilot study to assess at-home speed of processing training for individuals with multiple sclerosis. *Multiple Sclerosis Iinternational* 2019;**2019**:3584259. [PMID: 31275652]

Bhargav 2016 *{published data only}*

Bhargav P, Bhargav H, Raghuram N, Garner C. Immediate effect of two yoga-based relaxation techniques on cognitive functions in patients suffering from relapsing remitting multiple sclerosis: A comparative study. *International Review of Psychiatry* 2016, 2016;**28**(3):299-308.

Bonavita 2015 *{published data only}*

Bonavita S, Sacco R, Della C, Esposito S, Sparaco M, D'Ambrosio, A et al. Computer-aided cognitive rehabilitation improves cognitive performances and induces brain functional connectivity changes in relapsing remitting multiple sclerosis patients: an exploratory study. *Journal of Neurology* 2015;**262**(1):91-100.

Bove 2019 *{published data only}*

Bove RM, Rush G, Zhao C, Rowles W, Garcha P, Morrissey J, et al. A videogame-based digital therapeutic to improve processing speed in people with multiple sclerosis: a feasibility study. *Neurology and Therapy* 2019;**8**(1):135-45. [PMID: 30506301]

Brenk 2008 *{published data only}*

Brenk A, Laun K, Haase CG. Short-term cognitive training improves mental efficiency and mood in patients with multiple sclerosis. *European Neurology* 2008;**60**:304-9.

Brissart 2010 *{published data only}*

Brissart H, Leroy M, Debouverie M. Cognitive rehabilitation in multiple sclerosis: preliminary results and presentation of a new program, PROCOG-SEP. *Revue Neurologique* 2010;**166**:406-11.

Brissart 2013 *{published data only}*

Brissart H, Leroy M, Morele E, Baumann C, Spitz E, Debouverie M. Cognitive rehabilitation in multiple sclerosis. *Neurocase* 2013;**19**(6):553-65.

Cabrera-Gomez 2010 *{published data only}*

Cabrera-Gómez J, Real-González Y, Díaz-Márquez R, del Valle AE, González-Quevedo A, Martinez-Aching G, et al. A short-term, multidisciplinary, intensive inpatient neurorehabilitation program for multiple sclerosis patients with moderate disability: a pilot study. *International Journal of MS Care* 2010;**12**(2):66-70.

Campbell 2015 *{published data only}*

Campbell J, Langdon D, Cercignani M, Rashid D. Cognitive dysfunction in multiple sclerosis. The potential for neuroplasticity and rehabilitation. *ECTRIMS online library* 2015;**2015**:1.

Canellopoulou 1998 *{published data only}*

Canellopoulou M, Richardson JT. The role of executive function in imagery mnemonics: Evidence from multiple sclerosis. *Neuropsychologia* 1998;**36**(11):1181-8.

Chiaravalloti 2003 *{published data only}*

Chiaravalloti ND, Demaree H, Gaudino EA, DeLuca J. Can the repetition effect maximize learning in multiple sclerosis? *Clinical Rehabilitation* 2003;**17**:58-68.

Chiaravalloti 2012 *{published data only}*

Chiaravalloti ND, Wylie G, Leavitt V, DeLuca J. Increased cerebral activation after behavioral treatment for memory deficits in MS. *Journal of Neurology* 2012;**259**(7):1337-46.

Chiaravalloti 2018 *{published data only}*

Chiaravalloti ND, Goverover Y, Costa SL, DeLuca J. A pilot study examining speed of processing training (spt) to improve processing speed in persons with multiple sclerosis. *Frontiers in Neurology* 2018;**9**:685. [PMID: 30210426]

das Nair 2017 *{published data only}*

das Nair R. Cognitive rehabilitation for attention and memory in people with multiple sclerosis: the CRAMMS trial. *Multiple Sclerosis journal* 2017;**23**(S3):464.

De Giglio 2014 *{published data only}*

De Giglio L, De Luca F, Prosperini L, Borriello G, Bianchi V, Pantano P, et al. A low-cost cognitive rehabilitation with a commercial video game improves sustained attention and executive functions in multiple sclerosis: a pilot study. *Neurorehabilitation and Neural Repair* 2014;**29**(5):453-61.

Dobryakova 2014 *{published data only}*

Dobryakova A, Wylie G, DeLuca J, Chiaravalloti N. A pilot study examining functional brain activity 6 months after memory retraining in MS: the MEMREHAB trial. *Brain Imaging and Behavior* 2014;**8**(3):403-6.

Ernst 2013 *{published data only}*

Ernst A, Blanc F, Voltzenlogel V, de Seze J, Chauvin B, Manning L. Autobiographical memory in multiple sclerosis patients: assessment and cognitive facilitation. *Neuropsychological Rehabilitation* 2013;**23**(2):161-81.

Fiorotto 2015 *{published data only}*

Fiorotto SM, Barroso SM. Cognitive monitoring of a patient with multiple sclerosis [Relato de experiência em acompanhamento cognitivo com um paciente com esclerose múltipla]. *Psicologia: Ciência e Profissão* 2015;**35**(3):740-53.

Flavia 2010 *{published data only}*

Flavia M, Stampatori C, Zanotti D, Parrinello G, Capra R. Efficacy and specificity of intensive cognitive rehabilitation of attention and executive functions in multiple sclerosis. *Journal of the Neurological Sciences* 2010;**288**:101-5.

Goreover 2011 *{published data only}*

Goreover Y, Basso M, Wood H, Chiaravalloti N, DeLuca J. Examining the benefits of combining two learning strategies on recall of functional information in persons with multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2011;**17**(12):1488-97.

Grasso 2017 *{published data only}*

Grasso MG, Broccoli M, Casillo P, Catani S, Pace L, Pompa A, et al. Evaluation of the impact of cognitive training on quality of life in patients with multiple sclerosis. *European Neurology* 2017;**78**(1-2):111-7. [PMID: 28738376]

Güçlü Altun 2015 *{published data only}*

Güçlü Altun I, Kirbaş D, Altun DU, Soysal A, Sütlaş PN, Yandim Kuşçu D, et al. The effects of cognitive rehabilitation on relapsing remitting multiple sclerosis patients. *Noro Psikiyatri Arsivi* 2015;**52**(2):174-9. [PMID: 28360700]

Guijarro-Castro 2017 *{published data only}*

Guijarro-Castro C, Aladro-Benito Y, Sánchez-Musulim A, Belen-Caminero A, Pérez Molina I, Gómez-Moreno I, et al. Face-to-face or telematic cognitive stimulation in patients with multiple sclerosis and cognitive impairment: why not both? *Behavioural Neurology* 2017;**2017**:5713934. [PMID: 29386749]

Hanssen 2016 *{published data only}*

Hanssen KT, Beiske AG, LandrÃ¸ NI, Hofoss D, Hessen E. Cognitive rehabilitation in multiple sclerosis: a

randomized controlled trial. *Acta Neurologica Scandinavica* 2016;**133**(1):30-40. [PMID: 25952561]

Harand 2017 *{published data only}*

Harand C, Daniel F, Mondou A, Chevanne D, Branger P, Derache N, et al. Neuropsychological management of multiple sclerosis: evaluation of the added value of a supervised and customized cognitive rehabilitation program for self-used at home. *Multiple Sclerosis Journal* 2017;**23**(S3):831.

Harand 2019 *{published data only}*

Harand C, Daniel F, Mondou A, Chevanne D, Creveuil C, Defer G. Neuropsychological management of multiple sclerosis: evaluation of a supervised and customized cognitive rehabilitation program for self-used at home (SEPIA): protocol for a randomized controlled trial. *Trials* 2019;**20**(1):614. [PMID: 31665039]

Iaffaldano 2015 *{published data only}*

Iaffaldano P, Viterbo RG, Fazio L, Taurisano P, Torterella C, Romano R, et al. Computer-assisted rehabilitation of attention in patients with multiple sclerosis increases functional activity in the left prefrontal cortex. *Multiple Sclerosis Journal* 2015;**21**(S1):649.

Jimenez-Morales 2017 *{published data only}*

Jimenez-Morales RM, Herrera-Jimenez LF, Macias-Delgado Y, Perez-Medinilla YT, Diaz-Diaz SM, Forn C. Cognitive training combined with aerobic exercises in multiple sclerosis patients: a pilot study [Entrenamiento cognitivo combinado con ejercicios aerobicos en pacientes con esclerosis multiple: estudio piloto]. *Revista de Neurologia* 2017;**64**(11):489-95. [PMID: 28555454]

Kavaklioglu 2017 *{published data only}*

Kavaklioglu BC, Ozerden M, Bitnel MK, Ozudogru A, Soysal A. Investigation of the effects of computer assisted cognitive rehabilitation in multiple sclerosis patients: a randomised controlled study. *Multiple Sclerosis Journal* 2017;**23**(S1):962.

Lamargue 2020 *{published data only}*

Lamargue D, Koubiyr I, Deloire M, Saubusse A, Charre-Morin J, Moroso A, et al. Effect of cognitive rehabilitation on neuropsychological and semiecological testing and on daily cognitive functioning in multiple sclerosis: the REACTIV randomized controlled study. *Journal of the Neurological Sciences* 2020;**415**:116929. [PMID: 32460145]

Leavitt 2014 *{published data only}*

Leavitt VM, Wylie GR, Girgis PA, DeLuca J, Chiaravalloti ND. Increased functional connectivity within memory networks following memory rehabilitation in multiple sclerosis. *Brain Imaging and Behaviour* 2014;**8**(3):394-402.

Lincoln 2015 *{published data only}*

Lincoln NB, das Nair R, Bradshaw L, Constantinescu CS, Drummond AE, Erven A, et al. Cognitive Rehabilitation for Attention and Memory in people with Multiple Sclerosis: study protocol for a randomised controlled trial (CRAMMS). *Trials* 2015;**16**:556. [PMID: 26643818]

Loewenstein 2004 *{published data only}*

Loewenstein DA, Acevedo A, Czaja SJ, Duara R. Cognitive rehabilitation of mildly impaired Alzheimer disease patients on cholinesterase inhibitors. *American Journal of Geriatric Psychiatry* 2004;**12**(4):395-402.

Mäntynen 2014 *{published data only}*

Mäntynen A, Rosti-Otajärvi E, Koivisto K, Lilja A, Huhtala H, Hämäläinen P. Neuropsychological rehabilitation does not improve cognitive performance but reduces perceived cognitive deficits in patients with multiple sclerosis: a randomised, controlled, multi-centre trial. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2014;**20**(1):99-107.

Martin 2014 *{published data only}*

Martin KJ, Lincoln N, dasNair R. Group-based memory rehabilitation for people with multiple sclerosis: Subgroup analysis of the ReMiND trial. *International Journal of Therapy and Rehabilitation* 2014;**21**(12):422-8.

Mattioli 2012 *{published data only}*

Mattioli F, Stampatori C, Scarpazza C, Parrinello G, Capra R. Persistence of the effects of attention and executive functions intensive rehabilitation in relapsing remitting multiple sclerosis. *Multiple Sclerosis and Related Disorders* 2012;**1**(4):168-73.

Messinis 2015 *{published data only}*

Messinis L, Nousia A, Kosmidis MH, Nasios G, Papathanaspoulos P. Efficacy of a computer-assisted neuropsychological training programme in cognitive performance of patients with relapsing remitting multiple sclerosis. *Journal of the Neurological Sciences* 2015;**357**(S1):354-55.

Moore 2008 *{published data only}*

Moore KS, Peterson DA, O-Shea G, McIntosh GC, Thaut MH. The effectiveness of music as a mnemonic device on recognition memory for people with multiple sclerosis. *Journal of Music Therapy* 2008;**3**:307-29.

Nauta 2017 *{published data only}*

Nauta IM, Speckens AE, Kessels RP, Geurts JJ, de Groot V, Uitdehaag BM, et al. Cognitive rehabilitation and mindfulness in multiple sclerosis (REMIND-MS): a study protocol for a randomised controlled trial. *BMC Neurology* 2017;**17**(1):201. [PMID: 29162058]

Nurova 2014 *{published data only}*

Nurova B, Kurtuncu M, Coban A, Birday E, Eraksoy M. Computer assisted cognitive rehabilitation in patients with multiple sclerosis and parenchymal neuro-Behçet's disease. *European Journal of Neurology* 2014;**21**:367.

Penner 2018 *{published data only}*

Penner IK. Potential of exercise and cognitive training. *Multiple Sclerosis Journal* 2018;**24**(S2):38.

Perez-Martin 2016 *{published data only}*

Perez-Martin MY, Gonzalez Platas M, Eguia Del Rio P, Jimenez Sosa A. Randomized, blinded, controlled study to assess the

efficacy of a cognitive training program in patients with multiple sclerosis (MS). *Multiple SclerosisJournal* 2016;**22**(S3):701-2.

Pineau 2019 *{published data only}*

Pineau F, Socha J, Corvol JC, Louapre C, Assouad R, Maillart E, et al. Impact of an adaptive program for cognitive and emotional deficits (ADACOG program) in multiple sclerosis patients with cognitive impairments. *Revue Neurologique* 2019;**175**(5):305-12. [PMID: 30910222]

Rilo 2015 *{published data only}*

Rilo O, Pena J, Ojeda N, Rodriguez-Antiguedad A, Mendibe-Bilbao M, Gomez-Gastiasoro A, et al. Integrative group-based cognitive rehabilitation efficacy in multiple sclerosis. *Multiple Sclerosis Journal* 2015;**11**(S1):230-1.

Rilo 2016 *{published data only}*

Rilo O, Pena J, Natalia O, Rodriguez-Antiguedad A, Mendibe-Bilbao M. Cognitive improvement and functional connectivity changes after integrative group-based cognitive rehabilitation in multiple sclerosis. *Multiple Sclerosis Journal* 2016;**22**(S3):41.

Rilo 2017 *{published data only}*

Rilo O, Ojeda N, Pena J, Cabrera A, Rodriguez-Antiguedad A, Mendibe-Bilbao M, et al. Maintenance of the default mode network functional connectivity through an integrative group-based cognitive rehabilitation in multiple sclerosis. *Multiple Sclerosis Journal* 2017;**23**(6):891.

Rodgers 1996 *{published data only}*

Rodgers D, Khoo K, MacEachen M, Oven M, Beatty WW. Cognitive therapy for multiple sclerosis: a preliminary study. *Alternative Therapies in Health and Medicine* 1996;**2**(5):70-4.

Rosti-Otajärvi 2013a *{published data only}*

Rosti-Otajärvi E, Mäntynen A, Koivisto K, Huhtala H, Hämäläinen P. Patient-related factors may affect the outcome of neuropsychological rehabilitation in multiple sclerosis. *Journal of Neurological Science* 2013;**334**(1-2):106-11.

Rosti-Otajärvi 2013b *{published data only}*

Rosti-Otajärvi E, Mäntynen A, Koivisto K, Huhtala H, Hämäläinen P. Neuropsychological rehabilitation has beneficial effects on perceived cognitive deficits in multiple sclerosis during nine-month follow-up. *Journal of Neurological Science* 2013;**334**(1-2):154-60.

Shatil 2010 *{published data only}*

Shatil E, Metzer A, Horvitz O, Miller A. Home-based personalized cognitive training in MS patients: a study of adherence and cognitive performance. *NeuroRehabilitation* 2010;**26**:143-53.

Thaut 2014 *{published data only}*

Thaut M, Peterson D, McIntosh G, Hoemberg V. Music mnemonics aid verbal memory and induce learning - related brain plasticity in multiple sclerosis. *Frontiers in Human Neuroscience* 2014;**13**(8):1-10.

Thomas 2006 *{published data only}*

Thomas PW, Thomas S, Hillier C, Galvin K, Baker R. Psychological interventions for multiple sclerosis. *Cochrane Database of Systematic Reviews* 2006, Issue 1. Art. No: CD004431. [DOI: 10.1002/14651858.CD004431]

van der Putten 1999 *{published data only}*

van der Putten JJ, Hobart JC, Freeman JA, Thompson AJ. Measuring change in disability after inpatient rehabilitation: comparison of the responsiveness of the Barthel index and the Functional Independence Measure. *Journal of Neurology, Neurosurgery & Psychiatry* 1999;**66**:480-4.

Veldkamp 2019 *{published data only}*

Veldkamp R, Baert I, Kalron A, Tacchino A, D'hooge M, Vanzeir E, et al. Structured cognitive-motor dual task training compared to single mobility training in persons with multiple sclerosis, a multicenter RCT. *Journal of Clinical Medicine* 2019;**8**(12):1-16. [PMID: 31835502]

Vogt 2009 *{published data only}*

Vogt A, Kappos L, Calabrese P, Stöcklin M, Gschwind L, Opwis K, et al. Working memory training in patients with multiple sclerosis - comparison of two different training schedules. *Restorative Neurology and Neuroscience* 2009;**27**:225-35.

Wilson 2001 *{published data only}*

Wilson BA, Emslie HC, Quirk K, Evans JJ. Reducing everyday memory and planning problems by means of a paging system: a randomised control crossover study. *Journal of Neurology, Neurosurgery & Psychiatry* 2001;**70**:477-82.

Zimmer 2018 *{published data only}*

Zimmer P, Bloch W, Schenk A, Oberste M, Riedel S, Kool J, et al. High-intensity interval exercise improves cognitive performance and reduces matrix metalloproteinases-2 serum levels in persons with multiple sclerosis: a randomized controlled trial. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2018;**24**(12):1635-44. [PMID: 28825348]

References to ongoing studies

ISRCTN54901925 *{published data only}***10.1186/ISRCTN54901925**

ISRCTN54901925. A randomised study of cognitive rehabilitation in multiple sclerosis. http://isrctn.com/ISRCTN54901925 (accessed 31 July 2015).

NCT03471338 *{published data only}***10.1186/s13063-019-3715-7**

Neuropsychological management of multiple sclerosis: benefits of a computerised semi-autonomous at-home cognitive rehabilitation programme. Ongoing study. March 2018. Contact author for more information.

Additional references

Baddeley 1994

Baddeley A, Emslie H, Nimmo-Smith I. The Doors and People Test: a Test of Visual and Verbal Rrecall and Recognition. Bury St Edmunds: Thames Valley Test Company, 1994.

Beck 1987

Beck AT. Beck Depression Inventory. San Antonio: The Psychological Corporation, 1987.

Beck 2003

Beck AT, Steer R, Brown G. BDI - Fastscreen for Medical Patients. San Antonio, TX: Psychological Corporation, 2003.

Benedict 1996

Benedict RH, Schretlen D, Groninger, Dobraski M, Shpritz B. Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. *Psychological Assessment* 1996;**8**(2):145.

Benedict 1998

Benedict RH, Schretlen D, Groninger L, Brandt J. Verbal Learning Test-Revised: Normative data and analysis of inter-form and test-retest reliability. *Clinical Neuropsychologist* 1998;**12**:43-55.

Benedict 2004

Benedict RH, Munschauer F, Linn R, Miller C, Murphy E, Foley F, et al. Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2003;**9**(1):95-101.

Benedict 2017

Benedict RH, DeLuca J, Phillips G, LaRocca N, Hudson LD, Rudick R, et al. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *MultIple Sclerosis* 2017;**23**(5):721-33.

Broadbent 1982

Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology* 1982;**21**:1-16.

Buschke 1973

Buschke H. Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior* 1973;**12**(5):543-50.

Calabrese 2004

Calabrese P, Klabe E, Kessler J. A neuropsychological screening test to detect cognitive impairment in MS patients - The Multiple Sclerosis Cognition Inventory Music [Ein neuropsychologisches screening zur Erfassung kognitiver Störungen bei MS-Patienten - Das Multiple Sklerose Invenatrium Cognition MUSIC]. *Psychoneurologie* 2004;**30**:384-8.

Cella 1996

Cella DF, Dineen K, Arnason B, Reder A, Webster KA, Karabatsos G, et al. Validation of the functional assessment of multiple sclerosis quality of life instrument. *Neurology* 1996;**47**(1):129-39.

Chiaravalloti 2008

Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. *Lancet Neurology* 2008;**7**(12):1139-51.

Cicerone 2005

Cicerone KD, Dahlberg C, Malec JF, Langenbahn DM, Felicetti T, Kneipp S et al. Evidence-based cognitive rehabilitation: updated review of the literature from 1998 through 2002. *Archives of Physical Medicine and Rehabilitation* 2005;**86**(8):1681-92.

Cicerone 2011

Cicerone KD, Langenbahn DM, Braden C, Malec JF, Kalmar K, Fraas M, et al. Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008. *Archives of Physical Medicine and Rehabilitation* 2011;**92**:519-30.

Cicerone 2019

Cicerone KD, Goldin Y, Ganci K, Rosenbaum A, Wethe JV, Langenbahn DM, et al. Evidence-based cognitive rehabilitation: systematic review of the literature from 2009 through 2014. *Archives of Physical Medicine and Rehabilitation* 2019;**100**(8):1515-33. [DOI: https://doi.org/10.1016/j.apmr.2019.02.011]

Craig 2008

Craig P, Dippe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655.

Crook 1992

Crook TH, Feher EP, Larrabee GJ. Assessment of memory complaint in age-associated memory impairment: the MAC-Q. *International Psychogeriatrics* 1992;**4**:165-76.

das Nair 2007

das Nair R. Effectiveness of Memory Rehabilitation Following Brain Damage. Nottingham, UK: Univ. of Nottingham, 2007.

das Nair 2013

das Nair R, Lincoln NB. The effectiveness of memory rehabilitation following neurological disabilities: a qualitative inquiry of patient perspectives. *Neuropsychological Rehabilitation* 2013;**23**(4):528-45. [PMID: 23651146]

das Nair 2016a

das Nair R, Cogger H, Worthington E, Lincoln NB. Cognitive rehabilitation for memory deficits after stroke. *Cochrane Database of Systematic Reviews* 2016, Issue 9. Art. No: CD002293. [DOI: 10.1002/14651858.CD002293]

Davis 1995

Davis AM, Cockburn AM, Wade DT, Smith PT. A subjective memory assessment questionnaire for use with elderly people after stroke. *Clinical Rehabilitation* 1995;**9**:238-44.

Delis 2000

Delis DC, Kramer E, Kaplan BA, Ober B. California Verbal Learning Test - Second Eedition (CVLT® -II). Psychological Corporation 2000.

Drühe-Wienholt 1998

Drühe-Wienholt CM, Wienholt W. Computer-Aided Card-Sorting Procedure. Frankfurt: Swets Test Services, 1998.

EuroQoL Group 1990

EuroQoL Group. EuroQoL: a new facility for the measurement of heath-related quality of life. *Health Policy* 1990;**16**:199–208.

Gilchrist 1994

Gilchrist AC, Creed FH. Depression, cognitive impairment and social stress in multiple sclerosis. *Journal of Psychosomatic Research* 1994;**38**(3):193-201.

Gilewski 1990

Gilewski MJ, Zelinski EM, Schaie KW. The memory functioning questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging* 1990;**5**:482-90.

Glanville 2006

Glanville JM, Lefebvre C, Miles JN, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: ten years on. *Journal of the Medical Library Association* 2006;**94**(2):130-6.

Goldberg 1988

Goldberg D, Williams P. A User's Guide to the General Health Questionnaire. Windsor: NFER-Nelson, 1988.

Golden 1980

Golden CJ, Hammeke TA, Purish AD. The Luria-Nebraska Neuropsychological Battery: Manual (Revised). Los Angeles: Western Psychological Service, 1980.

Goverover 2018b

Goverover Y, Chiaravalloti ND, O'Brien Amanda R, DeLuca J. Evidenced-based cognitive rehabilitation for persons with multiple sclerosis: an updated review of the literature from 2007 to 2016. *Archives of Physical Medicine and Rehabilitation* 2018;**99**(2):390-407.

GRADE Working Group 2004

GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;**328**:1490-4.

GreyNet 2011

GreyNet International: Grey Literature Network Service. Grey Literature. www.greynet.org/greynethome/aboutgreynet.html (accessed 8 July 2011).

Hal 1997

Hal KM. The functional assessment measure (FAM). *Journal of Rehabilitation Outcomes Measurement* 1997;**1**(3):63-5.

Higgins 2019

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 [updated July 2019]. Cochrane, 2019. Available from www.training.cochrane.org/handbook.

Hobart 2001

Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. *Brain* 2001;**124**(5):962-73.

Hoffman 2014

Hoffman T, Glasziou P, Boutron I, Milne R, Perera R, Moher D et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**:1-12.

Huedo-Medina 2006

Heudo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 Index? *Psychological Methods* 2006;**11**:193-206.

Julian 2011

Julian LJ. Cognitive functioning in multiple sclerosis. *Neurologic Clinics* 2011;**29**(2):507-25. [PMID: 21439456]

Kalmar 2008

Kalmar JH, Gaudino EA, Moore NB, Halper J, DeLuca J. The relationship between cognitive deficits and everyday functional activities in multiple sclerosis. *Neuropsychology* 2008;**22**(4):442-9.

Klein 2019

Klein OA, Drummond A, Mhizha-Murira JR, Mansford L, dasNair R. Effectiveness of cognitive rehabilitation for people with multiple sclerosis: a meta-synthesis of patient perspectives. *Neuropsychological Rehabilitation* 2019;**29**(4):491-512. [PMID: 28457198]

Langdon 1996

Langdon DW, Thompson AJ. Cognitive problems in multiple sclerosis. *MS Management* 1996;**3**:5-9.

Langdon 2012

Langdon DW, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson P, et al. Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). *Multiple Sclerosis* 2012;**18**(6):891-8.

Lezak 2004

Lezak M, Howieson D, Loring D, Hannay H, Fisher J. Neuropsychological Assessment. New York: Oxford University Press, 2004.

Majid 2000

Majid MJ, Lincoln NB, Weyman N. Cognitive rehabilitation for memory deficits following stroke. *Cochrane Database of Systematic Reviews* 2000, Issue 3. Art. No: CD002293. [DOI: 10.1002/14651858.CD002293]

Martin 2015

Martin KJ, Sinclair E, dasNair R. Descriptions of memory rehabilitation group interventions for neurological conditions: a systematic review. *Clinical Rehabilitation* 2015;**29**:1-9. [DOI: 10.1177/0269215515595273]

McDonald 2001

McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. *Annals of Neurology* 2001;**50**:121-7.

Moher 2001

Moher D, Schultz KF, Altman D, CONSORT Group (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;**285**:1987-91.

Möhler 2015

Möhler R, Köpke S, Meyer G. Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline (CReDECI 2). *Trials* 2015;**16**:204.

Niemann 2003

Niemann H, Sturm W, Thone-Otto AI, Willmes K. Deutsche Normierung des California Verbal Learning Tests. Leipzig: Testzentrale, 2003.

Nouri 1987

Nouri FM, Lincoln NB. An extended activities of daily living index for stroke patients. *Clinical Rehabilitation* 1987;**1**:301-5.

Nyenhuis 1998

Nyenhuis DL, Luchetta T, Yamamoto C, Terrien A, Bernardin L, Rao S, et al. The development, standardization, and initial validation of the Chicago Multiscale Depression Inventory. *Journal of Personality Assessment* 1998;**70**(2):386-401.

Osterrieth 1944

Osterrieth PA. Le test de copie d'une figure complexe; contribution a l'etude de la perception et de la memoire. Archives de Psychologie 1944.

Paty 1988

Paty DW, Oger JJ, Kastrukoff LF, Hashimoto SA, Hooge JP, Eisen AA, et al. MRI in the diagnosis of MS: a prospective study with comparison of clinical evaluation, evoked potentials, oligoclonal banding, and CT. *Neurology* 1988;**38**:180-5.

Polman 2005

Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "MacDonald Criteria". *Annals of Neurology* 2005;**58**(6):840-6.

Polman 2011

Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the MacDonald criteria. *Annals of Neurology* 2011;**69**(2):292-302.

Poser 1983

Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Annals of Neurology* 1983;**13**(3):227-31.

Randolph 1998

Randolph C, Tierney MC, Mohr E, Chase TN. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *Journal of Clinical and Experimental Neuropsycology* 1998;**20**(3):310-9.

Rao 1990

Rao SM. A manual for the brief repeatable battery of neuropsychological tests in multiple sclerosis. Milwaukee: Medical College of Wisconsin 1990.

Rao 1991

Rao SM, Leo GJ, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology* 1991;**41**(5):685-91.

Rao 1993

Rao SM, Grafman J, DiGiulio D, Mittenberg W, Bernardin L, Leo GJ, et al. Memory dysfunction in multiple sclerosis: Its relation to working memory, semantic encoding, and implicit learning. *Neuropsychology* 1993;**7**(3):364-74.

RevMan 2020 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration Review Manager (RevMan). Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2020. Available at: revman.cochrane.org.

Rey 1941

Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique.(Les problems.). *Archives de Psychologie* 1941;**28**:215-85.

Robertson 1993

Robertson IH. Cognitive rehabilitation in neurologic disease. *Current Opinion in Neurology* 1993;**6**(5):756-60.

Robertson 2008

Robertson I, Fitzpatrick S. The future of cognitive neurorehabilitation. In: Stuss D, Winocur G, Robertson I, editors(s). Cognitive Neurorehabilitation: Evidence and Application. 2nd edition. Cambridge: Cambridge University Press, 2008.

Rosti-Otajärvi 2014

Rosti-Otajärvi EM, Hämäläinen P. Neuropsychological rehabilitation for multiple sclerosis. *Cochrane Database of Systematic Reviews* 2014, Issue 2. Art. No: CD009131. [DOI: 10.1002/14651858.CD009131]

Schumacher 1965

Schumacher G, Beebe G, Kibler R, Kurland L, Kurtzke J, McDowell F. Problems of experimental trials of therapy in multiple sclerosis. *Annals of the New York Academy of Sciences* 1965;**122**:552–68.

Schünemann 2013

Schünemann H, Brożek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013.. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook..

Sherer 2004

Sherer M. The awareness questionnaire. The Center for Outcome Measurement in Brain Injury. 2004;**http://www.tbims.org/combi/aq (accessed January 8, 2021)**..

Smith 1973

Smith A. Symbol Digit Modalities Test. Western Psychological Services Los Angeles, 1973.

Solari 2002

Solari A, Mancuso L, Motta A, Mendozzi L, Serrati C. Comparison of two brief neuropsychological batteries in people with multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)* 2002;**8**:171-8.

Solari 2003

Solari A, Motta A, Mendozzi L, Aridon P, Bergamaschi R, Ghezzi A, et al. The Italian version of the Chicago Multiscale Depression Inventory: translation, adaptation and testing in people with multiple sclerosis. *Neurological Sciences* 2003;**24**:375-83.

Spielberger 1983

Spielberger CD, Gorsuch R, Lushene R, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory (STAI). Consulting Psychologists Press, 1983.

Stroop 1938

Stroop JR. Factors affecting speed in serial verbal reactions. *Psychological Monographs* 1938;**50**(5):38.

Sturm 1999a

Sturm W, Wilmes K. Verbaler Lerntest (VLT). Mödling: Schuhfried, 1999.

Sturm 1999b

Sturm W, Willmes K. Nonverbaler Lerntest (NVLT). Mödling: Schuhfried, 1999.

Sunderland 1983

Sunderland A, Harris JE, Baddeley AD. Do laboratory tests predict everyday memory? A neuropsychological study. *Journal of Verbal Learning and Verbal Behavior* 1983;**22**:341-57.

Tewes 1991

Tewes U. Hamburg-Wechsler Intelligence Test for Adults. 2nd corrected edition. Göttingen: Hans Huber, 1991.

The WHOQOL Group 1993

The WHOQOL Group. Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL). *Quality of Life Research* 1993;**2**:153-9.

Thompson 2017

Thompson AJ, Banwel BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald Criteria. Lancet Neurology 2017;**17**(2):162-73. [DOI: http://dx.doi.org/10.1016/S1474-4422(17)30470-2]

Toglia 2004

Toglia JP. Contextual Memory Test. Ecole des sciences de la rÃ©adaptation, Sciences de la santÃ©, UniversitÃ© d'Ottawa, 2004.

Troyer 2017

Troyer AK, Rich JB. Multifactorial Memory Questionnaire. *Journal of Gerontology: Psychological Sciences* 2017;**57**(1):19-27.

Uniform Data System for Medical Rehab 1993

Uniform Data System for Medical Rehabilitation. Guide for the Uniform Data Set for Medical Rehabilitation (Adult FIM), Version 4.0. Buffalo: State Univ. of New York, 1993.

Vickrey 1995

Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health related quality of life measure for multiple sclerosis. *Quality of Life Research* 1995;**4**:187–206.

Ware 2001

Ware JE, Kosinski M. SF-36 Physical and Mental Health Summary Scales: a Manual for Users of Version 1, Second Edition. Lincoln: QualityMetric Incorporated, 2001.

Wechsler 1945

Wechsler DA. A standardized memory scale for clinical use. *Journal of Psychology* 1945;**19**:87-95.

Wechsler 1997

Wechsler D. Wechsler Memory Scale-Third edition (WMS-III). San Antonio: Harcourt Assessment, 1997.

Wilson 1984

Wilson BA, Moffat N. Clinical Management of Memory Problems. Bury St Edmunds: Chapman & Hall, 1984.

Wilson 1985

Wilson BA, Cockburn J, Baddeley AD. The Rivermead Behavioural Memory Test. London: Pearson Assessment, 1985.

Wilson 1999

Wilson BA, Clare L, Baddeley AD, Cockburn J, Watson P, Tate R. The Rivermead Behavioural Memory Test-Extended version. Bury St. Edmunds: Thames Valley Test Company, 1999.

Wilson 2005

Wilson BA, Shiel A, Foley J, Emslie H, Groot Y, Hawkins K, et al. Cambridge Test of Prospective Memory (CAMPROMPT). San Antonio: Harcourt Assessment, 2005.

Zigmond 1983

Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica* 1983;**67**:361-70.

References to other published versions of this review

das Nair 2012

das Nair R, Ferguson H, Stark DL, Lincoln NB. Memory rehabilitation for people with multiple sclerosis. *Cochrane Database of Systematic Reviews* 2012, Issue 3. Art. No: CD008754. [DOI: 10.1002/14651858.CD008754.pub2]

das Nair 2016b

das Nair R, Martin KJ, Lincoln NB. Memory rehabilitation for people with multiple sclerosis. *Cochrane Database of*

Systematic Reviews 2016, Issue 3. Art. No: CD008754. [DOI: 10.1002/14651858.CD008754.pub3]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Arian Darestani 2020

Study characteristics	
Methods	Non-blinded RCT, randomisation method not specified.
Participants	n = 60 Randomised: (E: 30, C: 30) Completed: (E: 27, C: 26) Mean age: (years) (E: 37.11, C: 39.23) Mean years of education: (E: 14.9, C:14.8)
Interventions	Computerised individual intervention, 10 sessions each 60 minutes long, location not specified. Sessions used RehaCom which is an autoadaptive comprehensive software to rehabilitate cognitive impairment.
Outcomes	Quote "significant efficacy of treatment with RehaCom for verbal learning and memory [CVLT-II], and verbal fluency [COWAT] within two groups combined". Significant differences between groups for both CVLT-II and COWAT.
Notes	RCT: randomised controlled trial, E: Experimental, C: Control, CVLT-II: California verbal learning test-second UK edition, COWAT: Controlled oral word association test

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote “60 patients were chosen and randomly divided into control (n=30) and experimental (n=30) groups” Method not specified.
Allocation concealment (selection bias)	Unclear risk	See above.
Blinding (performance bias and detection bias) All outcomes	High risk	No information given.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Intention to treat not used. Seven dropouts.
Selective reporting (reporting bias)	Low risk	All outcomes reported.
Other bias	Unclear risk	Unclear what the control group participants were told about the study.

Campbell 2016

Study characteristics	
Methods	Open-design, RCT was conducted
Participants	n = 38 Randomised (E: 19, C: 19) Completed (E: 17, C: 18) Mean age (years) (E: 46.21, C: 48.52) Mean years of education (E: 14.05, C: 13.63) Treatment group had higher BICAMS baseline scores but not significant, well-matched on all other characteristics
Interventions	Computerised, home-based rehabilitation, 18 sessions over 6 weeks, 45 minutes long Intervention group: sessions used RehaCom which is an autoadaptive comprehensive software to rehabilitate cognitive impairment. Control group: quote "[participants were] asked to watch a series of natural history DVDs of corresponding duration and frequency to the rehabilitation sessions performed by the treatment group for six weeks"
Outcomes	Significant improvements found within group between time 1 and time 2 for intervention group on SDMT, but not BVMT-R or CVLT-II, no significant improvements at any other time point
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, BICAMS: Brief international cognitive test for MS, SDMT: symbol digit modalities test, BVMT-R: Brief visuospatial memory test-revised, CVLT-II: California verbal learning test-second UK edition

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "Randomisation was performed using a random number generator and allocations were placed inside sealed folders. Folders were opening following the baseline MRI"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	High risk	Quote "it was not possible for the cognitive assessment to be completed by a blind assessor"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Intention to treat not used. 3 dropouts at time 2.
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Carr 2014

Study characteristics	
Methods	Single-blind RCT, randomisation by off-site independent randomisation centre, computer-generated random number sequence
Participants	n = 48 (E: 24, C: 24) Mean age (years) E: 55.8, C: 52.9 Mean years of education E: 15.7, C: 13.5
Interventions	Group format, 10 sessions, each 1.5 hours long. Sessions included both compensation and restitution, including memory education, strategies to help focus attention, internal memory strategies, use of external aids
Outcomes	Intention-to-treat analysis used No significant differences between groups at 4 or 8 months on EMQ, MSIS-29. Experimental group scored better than control on GHQ-28 at 8 months' follow-up, no difference at 4 months
Notes	RCT: randomised controlled trial, E: Experimental, C: Control, EMQ: Everyday Memory Questionnaire, MSIS-29: Multiple Sclerosis Impact Scale, GHQ-28: General Health Questionnaire

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "computer generated list prepared in advance of the study and held by an independent researcher at the University of Nottingham"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Postal outcomes that were quote: "scored by a researcher blind to group allocation"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intention to treat used. If less than 10% of items missed, these were replaced with mean for questionnaire
Selective reporting (reporting bias)	Low risk	All outcomes analysed and reported
Other bias	Low risk	None identified

Charvet 2017

Study characteristics	
Methods	Double-blinded, randomised, active placebo-controlled trial Block randomisation generated by blinded statistician
Participants	n = 135 Mean age (years): (E: 50, C: 52)

Charvet 2017 *(Continued)*

	Mean years of education: (E: 14.82, C: 15.05) Groups all well-matched apart from there being more men assigned to ACR versus active control condition
Interventions	Home-based intervention, 60 sessions over 12 weeks, 60 minutes each ACR condition: online adaptive training programme Active control condition: computer-based game-playing
Outcomes	Intent-to-treat analysis used Significant improvements within ACR condition, very few significant between group differences
Notes	E: Experimental, C: Control, ACR: Adaptive cognitive remediation

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "Eligible and consented participants were randomly assigned to the ACR or active control condition using stratified, permuted, block randomization generated by the study statistician" strata were based on age, WRAT-3 scores and SDMT age-normative scores
Allocation concealment (selection bias)	Low risk	Quote "As any training could be potentially beneficial, participants were told they would be randomly assigned to one of two training programs that were being compared"
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote “Both the participant and study psychometricians were blinded to treatment condition” Quote “The study technician that assigned a participant’s condition was not involved in the collection of data at baseline or study end visits”
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intention-to-treat used. Quote "Missing primary outcome values were imputed via Markov Chain Monte Carlo method"
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Chiaravalloti 2005

Study characteristics

Methods	Odd-even random allocation Participants kept blind to treatment
Participants	n = 29 Randomised (E: 15, C: 14) Completed: (E: 14, C: 14)

Chiaravalloti 2005 *(Continued)*

	Age: 45 to 46 years Education: 14 to 15 years Groups comparable on all but duration of illness variable (E: group longer disease duration)
Interventions	Group format 8 sessions (45-minute sessions, 2/week) E: SMT (imagery and story) C: reading story and recall without SMT
Outcomes	Intention-to-treat analysis not used Non-significant results of group or time on HVLT-R, STAI, BDI Significant difference on MFQ (E > C); but subgroup analysis: significant difference in learning ability (HVLT-R) at follow-up 1 and 2 for moderate-severe group (E > C)
Notes	E: Experimental; C: Control; SMT: Story Memory Technique; HVLT-R: Hopkins Verbal Learning Test-Revised; STAI: State Trait Anxiety Inventory; BDI: Beck Depression Inventory; MFQ: Memory Functioning Questionnaire

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quasi-random "odd-even" allocation
Allocation concealment (selection bias)	High risk	Odd/even allocation by primary investigator
Blinding (performance bias and detection bias) All outcomes	Low risk	Participants and assessors had no knowledge of group allocation
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Intention to treat not used. 1 dropout
Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Low risk	None identified

Chiaravalloti 2013

Study characteristics

Methods	Double-blind, placebo-controlled, randomised controlled trial
Participants	n = 86 (E: 41, C: 45) Groups similar in demographic and disease characteristics, disease-modifying therapy, pretreatment cognition, and emotional symptomatology

Chiaravalloti 2013 *(Continued)*

Interventions	mSMT, 10 sessions over 5 weeks (2 per week) Session length 45 minutes to 1 hour, focused on imagery and context 2 sessions on applying mSMT to real-life scenarios
Outcomes	Intention-to-treat analysis not used. Immediate follow-up: E group showed greater learning slope on CVLT (P = 0.007), E also showed significant improvement from baseline to follow-up on CVLT slope (P = 0.009). Significant differences (E > C) on RBMT story, FAMS general contentment, FrSBe. Long-term follow-up: Decline in CVLT slope from immediate to 6 months' follow-up. Significant difference (E > C) on FAMS general contentment.
Notes	E: Experimental; C: Control; mSMT: modified Story Memory Technique; CVLT: California Verbal Learning Task; RBMT: Rivermead Behavioural Memory Test; FAMS: Functional Assessment of Multiple Sclerosis; FrSBe: Frontal Systems Behaviour Scale

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "computerised random number generator ... the individual responsible for group assignment was not otherwise involved in data collection and group assignment was verified by a second individual via duplicate copy of the randomization table generated before initiation of data collection"
Allocation concealment (selection bias)	Low risk	Quote: "treatment allocation was concealed"
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote:"All RAs conducting assessments were blinded to group membership". Masking details given. Participants also blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	High risk	88 were randomised (E = 46, C = 42), but immediate outcomes were for E = 45, C = 41, and long term outcomes were for E = 40 and C = 38. No intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Tables given as supplement to all outcomes and statistical analyses
Other bias	Low risk	None identified

Chiaravalloti 2019a

Study characteristics

Methods	Quote “This RCT used a 4-week, double-blind, parallel-groups design”
Participants	n = 20 Mean age (years) (E: 49.67, C: 45.45) Mean years of education (E: 14.33, C: 16.00)
Interventions	Multi centre study, 8 sessions over 4 weeks, 30-45 minutes long.

Chiaravalloti 2019a *(Continued)*

	STEM is designed to teach the concepts of SG, SL, and RP and the application of these techniques in everyday life.
Outcomes	No significant differences for the primary outcomes. Positive significant impact on QoL and ADL.
Notes	RCT: randomised controlled trial, E: Experimental, C: Control, STEM: Strategy based training to enhance memory, SG: Self-generation, SL: Spaced learning, RP: Retrieval practice, QoL: Quality of life, ADL: Activities of daily living.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "The groups were assigned via 1:1 randomization using a computerized random number generator"
Allocation concealment (selection bias)	Low risk	Quote "Treatment allocation was concealed"
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote: "Participants were also blind to group assignment and consented to participate in a study examining the impact of mental exercises on memory."
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Chiaravalloti 2019b

Study characteristics	
Methods	Quote "This RCT used a 5-week, double-blind, parallel groups design"
Participants	n = 30 28 participants included in immediate follow-up, 24 participants included in intermediate follow-up Mean age (years) (E: 55.2, C: 53.31) Mean years of education (E: 16.07, C: 16.46)
Interventions	Individual rehabilitation, 10 sessions over 5 weeks, 60 minutes long mSMT trains two related skills: imagery and context, early sessions encourage participants to create visual imagery to aid memory of presented stories and the final sessions focus on applying mSMT to daily life Control group engaged in non-training-specific tasks to control for professional contact and disease alterations
Outcomes	Within group: significant improvement in CVLT-II but not RBMT

Chiaravalloti 2019b *(Continued)*

	No between-group differences No significant improvements in CDMT or STAI
Notes	RCT: randomised controlled trial, E: Experimental, C: Control, mSMT: modified Story memory technique, CVLT: California Verbal Learning Task; RBMT: Rivermead Behavioural Memory Test, CMDI: Chicago multiscale depression scale, STAI: State trait anxiety inventory

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "Using a computerized random number generator, 48 randomized assignments to TX or CTL were created before data collection"
Allocation concealment (selection bias)	Low risk	Quote “Treatment allocation was concealed”
Blinding (performance bias and detection bias) All outcomes	Low risk	Double-blind study
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Intention to treat not used. Two participants dropped out after baseline due to time commitments and a further four were lost to follow-up
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Chmelařová 2020

Study characteristics	
Methods	Non-blinded, RCT, unspecified randomisation method
Participants	n = 58 Randomised (E: 35, C: 23) Completed (E: 26, C: 17) Mean age (years) (E: 41.3, C: 42.4)
Interventions	Computerised home-based intervention using the Happy Neuron Brain Jogging programme, 8 sessions over 4 weeks, 30 minutes long. The programme contains 20 tasks related to different areas of cognition, including memory and the levels vary depending on ability. Control group received no training but were contacted periodically to discuss their psychological status
Outcomes	Within group: significant improvement in RBANS total and TMT scores, no significant improvement for subjective measures

Chmelařová 2020 *(Continued)*

	No significant differences for BDI, HADS or FAMS measures.
Notes	RCT: randomised controlled trial, E: Experimental, C: Control, RBANS: Rivermead behavioural memory test, TMT: Trail making test, BDI: Beck depressive inventory, HADS: Hospital anxiety and depression scale, FAMS: Functional assessment of multiple sclerosis

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote:"Fifty-eight patients were randomized into the experimental and control groups." - method not specified
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding (performance bias and detection bias) All outcomes	High risk	No mention of blinding
Incomplete outcome data (attrition bias) All outcomes	High risk	9 experimental participants and 6 control group participants dropped out between baseline and follow-up, no mention of how they handled the missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

das Nair 2012

Study characteristics

Methods	Single-blind RCT, randomisation by off-site independent randomisation centre, computer-generated random number sequence
Participants	n = 39 with MS Randomised (A: 17, B: 12, C: 10) Mean age: 47.2 years Education years: 14.1 years
Interventions	Groups: A: Restitution - encoding and retrieval strategies, attention retraining B: Compensation - external memory aids C: Attention placebo - relaxation techniques 10 weekly sessions, 90 minutes each
Outcomes	Intention-to-treat analysis used

das Nair 2012 *(Continued)*

	Non-significant differences between groups on RBMT-E, EMQ, EMAQ, GHQ, MATBD, and EADL; significant differences in IMAQ between groups; significant main-effect on RBMT-E and MATBD over time but across all 3 groups
Notes	Analysis used in this review: A + B versus C RCT: randomised controlled trial; RBMT-E: Rivermead Behavioural Memory Test-Extended; EMQ: Everyday Memory Questionnaire; EMAQ: External Memory Aids Questionnaire; GHQ: General Health Questionnaire; MATBD: Mental Adjustment to Brain Damage; EADL: Extended Activities of Daily Living; IMAQ: Internal Memory Aids Questionnaire

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computerised random number generation by independent agency
Allocation concealment (selection bias)	Low risk	Allocation was not known by intervention provider until all 4 participants were allocated to a group
Blinding (performance bias and detection bias) All outcomes	Low risk	Outcome assessors were blind to the random allocation and the intervention participants received. Participants were requested not to disclose any information about intervention at follow-up
Incomplete outcome data (attrition bias) All outcomes	Low risk	List-wise deletion utilised and baseline data were imputed for missing follow-up data
Selective reporting (reporting bias)	Low risk	All data were analysed and results disclosed
Other bias	Low risk	None identified

De Luca 2019

Study characteristics	
Methods	Pilot study, assessor-blinded, RCT, block randomisation
Participants	n = 40 Mean age (years) (E: 52.7, C: 57.0) Mean years of education (E: 10.8, C: 11.3)
Interventions	Sessions involved computerised rehabilitation using ERICA, 24 sessions over 8 weeks, 45 minutes long Control group: sessions involved traditional CR in a face-to-face approach between patient and therapist
Outcomes	Significant treatment effect found within group for MoCA, SDMT and SRT
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, ERICA: Individual and Collective Regulation of Learning Scale, CR: Cognitive rehabilitation, MoCA: Montreal cognitive assessment, SDMT: symbol modalities digit test, SRT: selective reminding test

De Luca 2019 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "All the patients were randomized (using a block randomization with a block size of 22) into either traditional CR group (TCRG: 20 patients, 12 men and 8 women) or the computer-assisted CR group (CCRG: 20 patients, 11 men and 9 women"
Allocation concealment (selection bias)	Unclear risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Moreover, both the raters and the patients were blinded to the patient's allocation"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote "All the patients showed a mild-to-moderate cognitive impairment and none withdrew from the study"
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Ernst 2015

Study characteristics	
Methods	Sinble-blind, RCT
Participants	n = 40 Randomised (E: 17, C: 10, S: 13) Mean age (years) (E: 42.00, C: 37.40, S: 40.00) Mean years of education (E: 13.29, C: 12.20, S: 13.77)
Interventions	Sessions involved an MVI programme which is based on the ability to mentally construct mental images and pay attention to details, 6 sessions, 2 hours long C: observed the same clinical characteristics and interactions with patients as E S controls for learning effects due to repeated AM/EFT assessments
Outcomes	Significant improvement in a number of details recalled post intervention It is important to note that only 15 participants were reassessed at 6 months follow-up
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, S: Stability group, AM: Autobiographical memory, EFT: episodic future thinking

Risk of bias

Bias	Authors' judgement	Support for judgement

Ernst 2015 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Quote:"The final 40 MS patients were randomly assigned in 457 the three following groups: (i) the experimental group, 458 who followed the MVI facilitation programme; (ii) the 459 verbal control group, who underwent the verbal control 460 programme and aimed to verify the absence of a nursing effect; and (iii) the stability group, whose inclusion 462 was thought to control for learning effects due to 463 repeated AM/EFT assessments."- method not specified
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding (performance bias and detection bias) All outcomes	High risk	Quote “Patients were blind to their allocation group and, importantly, they had never before participated in similar studies” - this is not justified Quote “Neuropsychologist not blind second AI scorer was blind to the group membership in every case” Quote “the neuropsychologist was not blind to the patient’s allocation group.”
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Three patients dropped out of the stability group
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Ernst 2018

Study characteristics	
Methods	Single-blind, RCT
Participants	n = 20 Mean age (years) (E: 38.40, C: 37.40) Mean years of education (E: 13.30, C: 12.20)
Interventions	Sessions involved an MVI programme which is based on the ability to mentally construct mental images and pay attention to details, 6 sessions, 2 hours long Control group: quote "The control programme followed the same procedure but focused on the narrative structure, which plays a minor role in AM relative to MVI"
Outcomes	Significant improvement in a number of details recalled post intervention
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, AM: Autobiographical memory
Risk of bias	
Bias	**Authors' judgement** **Support for judgement**

Ernst 2018 *(Continued)*

Random sequence generation (selection bias)	Low risk	Quote “Patients were randomly assigned to two groups using a computerised random number generator and were blind to their allocation group (experimental or control)”
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote “While the main scorer (AE) was not blind to the patient’s allocation group, a blind rater verified the scoring accuracy for 20% of memories randomly chosen.” “The reliability between the two scorers was assessed with intraclass correlations and indicated a high agreement for both composites (internal details: .95; external details: .94).”
Incomplete outcome data (attrition bias) All outcomes	Low risk	All participants assessed
Selective reporting (reporting bias)	Low risk	AI, semi-structured interview and post-scan results presented
Other bias	Low risk	None identified

Gich 2015

Study characteristics	
Methods	Randomised, controlled, single-blind pilot study
Participants	n = 43 (only 41 analysed), RRMS and SPMS E: 22 (21 analysed), C: 21 (20 analysed)
Interventions	Experimental group received 2 x 75-minute sessions per week for 6 months, included written (crosswords, word searches), manipulative (origami, spatial games) and computerised tasks (working memory games, log and reasoning games), additionally participants completed 5-minute daily cognitive activities at home. Control group received no treatment.
Outcomes	BRBNT: significant differences favouring experimental on 10/36 spatial task and word list generation. No significant differences on list-learning task (selective reminding task) - used in the meta-analysis
Notes	BRBNT: Brief Repeatable Battery of Neuropsychological Tests; RRMS: Relapsing remitting multiple sclerosis; SPMS: Secondary progressive multiple sclerosis

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote:"Participants were randomly assigned to one of the two arms in a 1:1 ratio. The randomization was stratified to avoid possible confounding variables, using the level of cognitive impairment as strata". No mention of how the random sequence was generated

Gich 2015 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	No mention of allocation concealment
Blinding (performance bias and detection bias) All outcomes	Low risk	Outcome assessors blind to treatment allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Analysed only those who completed outcomes, only 1 withdrew from each group
Selective reporting (reporting bias)	Low risk	All outcomes recorded
Other bias	Low risk	None identified

Goodwin 2020

Study characteristics	
Methods	Single-blind, cross-over, RCT
Participants	n = 38 Mean age (years) (E: 48.8,C: 46.7) Mean years of education (E: 14.3, C: 14.3) This is a cross-over study so we only included the results from the first round of the trial, as shown in Figure 3
Interventions	Intervention participants received NeuroPage text messages for 2 months, the messages were based on problems identified at baseline and the prompts sent at pre-arranged times Control group: participants received non-memory text messages for two months, sent at same time and frequency as NeuroPage messages would have been Intent-to-treat used
Outcomes	No improvement in memory found Significant between group differences found in GHQ and EQ-5D-5L (only anxiety and depression subscales)
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, GHQ: General health questionnaire, EQ-5D-5L: EuroQoL five dimension five levels

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "Participants were randomly allocated to the intervention or the control group on a 1:1 ratio"
Allocation concealment (selection bias)	Low risk	Quote "Allocation was determined by an independent research assistant, using a randomisation sequence prepared in advance of the study"

Goodwin 2020 *(Continued)*

		Quote: “The independent research assistant disclosed the group allocation of the participant to the researcher delivering the intervention only after the allocation was recorded"
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Due to the nature of the intervention, both the treating researcher and participants were aware of which group they had been allocated to" Quote: “Outcome measures were scored and entered into a password protected database by a researcher blind to group allocation”
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote:“No participants dropped out from the intervention phase. One participant withdrew part way through the control condition” Intent-to-treat analysis used
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Goverover 2018a

Study characteristics		
Methods	Double-blinde, placebo-controlled, RCT	
Participants	n = 35 Mean age (years) (E: 50.15, C: 48.50) Mean years of education (E: 16, C: 15.2)	
Interventions	Sessoins involved the self-GEN trial which is based on research demonstrating that items are self-generated by an individual are better remembered than items read or heard, 6 sessions over 3 weeks, 60 minutes long Control group: met with the researcher for the same frequency and time as the intervention group and were simply asked to remember things	
Outcomes	No significant improvements for MFQ, but significant improvements found for CVLT-II and CMT No significant improvements found on SWLS and only significant differences found on depression scales of mood measures	
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, MFQ: Memory failures questionnaire, CVLT-II: California verbal learning test- second edition, SWLS: Satisfaction with life scale	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "Participants then completed baseline testing after which, using a randomized number table, participants were randomized to either the treatment or control group”
Allocation concealment (selection bias)	Unclear risk	No mention of allocation concealment

Goverover 2018a *(Continued)*

Blinding (performance bias and detection bias) All outcomes	Low risk	Quote: “Note that the individual responsible for group assignment was not involved in data collection” “The same research assistant (RA) conducted baseline and the follow-up evaluations and was blinded to group membership” “Study participants were also blinded to group assignment. Participants consented to participate in a study examining the impact of memory treatment in which they had a 50/50 chance of being in the treatment or the control group. All participants completed the study”
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Hancock 2015

Study characteristics	
Methods	Blinded, placebo-controlled design, block-stratified randomisation method
Participants	n = 40 (n = 30 analysed) Mean age: 48.8 years Mean education: 15.45 years
Interventions	Active training group: completed a computerised cognitive training programme that specifically aimed to improve information-processing speed and working memory. Completed 30-minute intervals, 6 days per week for 6 weeks. Control group: completed a computerised cognitive training programme that is almost identical to the active training group, but this programme is not intended to improve information-processing speed or working memory. This programme employed the same tasks as the former, but it did not increase in difficulty in order to challenge participants to improve. Same time intervals and length as active training group.
Outcomes	Completed immediately after the 6-week training programme. No significant differences between groups on AVLT, BDI-FS, MSIS, MSQOL
Notes	AVLT: Auditory Verbal Learning Test; BDI-FS: Beck Depression Inventory-Fast Screen; MSIS: Multiple Sclerosis Impact Scale; MSQOL: Multiple Sclerosis Quality of Life

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "A block stratified randomization method was employed"

Hancock 2015 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	No mention of allocation concealment
Blinding (performance bias and detection bias) All outcomes	Low risk	Invesigator who conducted assessment was blind to allocation, as were participants
Incomplete outcome data (attrition bias) All outcomes	High risk	71 were randomised and 31 quote:"either withdrew from the study or were lost to follow up", however, no statistical differences were observed for those who completed compared to those who withdrew/lost to follow-up. Analysis on only those who completed the trial, and were "good adherers" to intervention (at least 80% sessions attended)
Selective reporting (reporting bias)	High risk	Analysis only on those who were "good adherers" and completed trial. Not all outcomes reported in published article (BDI and MSQOL not reported), unpublished data (received from author) used in this meta-analysis
Other bias	Low risk	None identified

Hanssen 2015

Study characteristics	
Methods	Prospective, randomised controlled design
Participants	n = 120, E: 60, C: 60 Inpatients at multidisciplinary rehabilitation centre
Interventions	Inpatient cognitive rehabilitation. All participants given baseline neuropsychological testing, control offered no feedback. Experimental group offered feedback, used to develop individualised plan. Mix of individual and group sessions, focused around goal attainment. Sessions included psychoeducation, learning strategies for quote: "keeping track of appointments and belongings". After discharge, those in experimental group had 6 bi-weekly telephone sessions focused on the goals they had set during the intervention.
Outcomes	No memory outcomes. MSIS-29. Significant effect of group at 7 months' follow-up (experimental less distressed than control).
Notes	MSIS-29: Multiple Sclerosis Impact Scale; Analysis only on those completing outcomes

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomization was performed by a lottery controlled by the director of the rehabilitation center"
Allocation concealment (selection bias)	High risk	Quote: "Concealment of treatment allocation was not possible due to the nature of the intervention"
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No mention of blinding, however self reported questionnaires used as follow-ups

Hanssen 2015 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Analysis only on those who completed follow-up assessments
Selective reporting (reporting bias)	Low risk	None apparent
Other bias	Low risk	None apparent

Hildebrandt 2007

Study characteristics	
Methods	Alternating allocation. Participants informed of intervention or assessment. Outcome assessor blind
Participants	n = 42; RRMS only Randomised: E: 17, C: 25 Mean age (years) E: 42 ; C: 36.5
Interventions	E: Memory and working memory rehab tasks. 30 minutes/day, 5 days/week, for 6 weeks C: Assessments only
Outcomes	Intention-to-treat analysis used Non-significant results of CVLT - Short Delay Free/Cued Recall or CVLT - Long Delay Cued Recall Significant differences on CVLT long delay free recall Non-significant results of BDI, SF-12, EDSS
Notes	E: Experimental; C: Control; RRMS: Relapsing remitting multiple sclerosis; CVLT: California Verbal Learning Test; BDI: Beck Depression Inventory; SF-12: 12-Item Short Form Health Survey; EDSS: Expanded Disability Status Scale

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Quote: "Before the patients' assessment, randomisation was done by alternating between intervention and control group"
Allocation concealment (selection bias)	Unclear risk	Quote: "Assignment and enrolment was done by randomisation according to groups before the patients were contacted". Participants were informed of whether they would receive intervention, or assessment only
Blinding (performance bias and detection bias) All outcomes	Low risk	Participants not blinded, healthcare providers not blinded. Outcome assessors reportedly blinded: quite: "done by colleagues, who were not involved in the study"
Incomplete outcome data (attrition bias) All outcomes	Low risk	No incomplete data

Hildebrandt 2007 *(Continued)*

Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Low risk	None identified

Huiskamp 2016

Study characteristics		
Methods	Double-blind, placebo-controlled, RCT	
Participants	n = 16 Mean age (years) (E: 48.33, C: 49.29) Mean years of education (E: 15.17, C: 15.86)	
Interventions	Sessions use mSMT which is a validated CR protocol that trains people to use visualization and context to learn new information, 10 sessions over 5 weeks, duration not specified	
Outcomes	TimexGroup interaction not significant Significant effect for time accuracy for E Time and Group interactions significant for reaction time	
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CR: Cognitive rehabilitation	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "Participant randomization was achieved via a computerized random number generator"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Researchers and participants blinded to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Impellizzeri 2020

Study characteristics		

Impellizzeri 2020 *(Continued)*

Methods	Observer-blind, RCT
Participants	n = 30 Mean age (years) (E: 51.73, C: 51.33) Mean years of education (E: 11.47, C: 11.47)
Interventions	Experimental group: sessions included CCR and NMT (this included AMMT and MPC), 24 sessions over 8 weeks, 60 minutes long Control group: only received CCR for the same amount of sessions
Outcomes	Sig between group differences for SPART and SRT E more significant improvement in mental subset of MSQoL-54 E more significant improvement for both BDI and EAQ
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CCR: Conventional cognitive rehabilitation, NMT: Neurological Music Therapy, AMMT: Associative Mood and Memory Training, MPC: Music in Psychosocial Training and Counseling, SPART: 10/36 spatial recall test, SRT: Selective reminding test, BDI: Beck depression inventory, EAQ: Emotional awareness questionnaire

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "Patients were assigned in a 1:1 ratio using a computer-generated randomization list assessed by statisticians, which was blinded to the training allocation"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Single-blind Quote “Patients received instructions not to tell other patients anything about what they do during NMT training techniques” "Statisticians and clinical assessors were blinded to group allocation"
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Jønsson 1993

Study characteristics

Methods	Closed-envelope randomisation

Jønsson 1993 *(Continued)*

Participants	n = 40 (E: 20; C: 20) hospital inpatients; (16 + 16 completed) Mean age: 44.5 years (SD: 8.3) Education: 11.5 years (SD: 2.5) Gender: 19F, 21M Groups comparable on all variables, except visuospatial memory and visual perception (more impaired in E group) Mild-moderate cognitive impairments
Interventions	Individual treatment 1-1.5hours, 3 times/week; mean total hours: 17.2 (5.1) E: compensation (internal and external memory aids), substitution, direct training (puzzles, etc.) + neuropsychotherapy C: attention placebo: discussion and games
Outcomes	Intention-to-treat analysis not used Follow-up 1: E > C on visual perception (but could be due to regression towards the mean and ceiling effects) and BDI Follow-up 2: E > C on visuospatial memory and BDI (C group became more depressed)
Notes	E: Experimental; C: Control; SD: Standard deviation; BDI: Beck Depression Inventory

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomly recruited"
Allocation concealment (selection bias)	Low risk	Closed-envelope system
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Quote: "Patients were told there were 2 treatments, but not which was better"... and ... "Healthcare providers were not told of patients' allocation, but a few words would have given it away" ... and ... Quote:"At follow up we were in principle blinded to what kind of treatment patients had been given", but patient report/talk could have broken blind
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not stated
Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Low risk	None identified

Lincoln 2002

Study characteristics	
Methods	Single-blind RCT; independent phone randomisation Computer-generated numbers
Participants	n = 240 Randomised (A: 82; B: 79; C: 79) Completed (A: 77; B: 71; C: 73) Median age: 40 to 43 years Age left education: 16 years Groups comparable on baseline variables
Interventions	Individual treatment A: only baseline assessment with no feedback B: detailed cognitive assessment with feedback C: detailed cognitive assessment + feedback + internal and external memory aids
Outcomes	Intention-to-treat analysis used No significant differences between 3 groups on any measures at follow-up 1 or 2 for patient and relative data, except QoL (Questions 53 and 54 of the MSQOL-54) at follow-up 2
Notes	For this review A vs C compared; RCT: randomised controlled trial; QoL: Quality of life

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "telephoning an independent department who had a computer generated allocation list"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote: "An independent assessor, who was unaware of the group allocation, assessed the outcome at 4 and 8 months after randomisation"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Analysis was just on those who completed the outcome assessments, however it included those who did not get the intervention as planned
Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Low risk	None identified

Lincoln 2020

Study characteristics	
Methods	Multicentre, parallel-group, pragmatic RCT
Participants	n = 449 Randomised (E: 245, C: 204) Included in intermediate follow-up (E: 220, C: 182) Included in longer-term follow-up (E: 206, C: 170) Mean age (years) (E: 49.9, C: 48.9) Mean years of education (E: 14.2, C: 13.9)
Interventions	Sessions involved group CR with 4 to 6 participants per group with the intention of identifying the most appropriate strategies to help individuals overcome cognitive difficulties Control group: received usual care
Outcomes	Intent-to-treat analysis used for missing data At 6 and 12 months between group differences favoured E for EMQ-p, SRT (total recall at 6 months and delayed at 12 months), no other significant between group differences Adjusted significant between group differences for MSIS at 6 months but none at 12 months Sig differences favoured E for GHQ
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CR: Cognitive rehabilitation, EMQ-p: Everyday memory questionnaire-participant, MSIS: Multiple sclerosis impact scale, GHQ: General health questionnaire

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "Participants were individually randomised to the intervention or control group on a 6:5 ratio, to allow for clustering in the intervention group"
Allocation concealment (selection bias)	Low risk	Quote “The sequence of group allocations was concealed from the trial statistician until all participants had been allocated, and recruitment, data collection and all other trial-related assessments were complete”
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote “The RAs were not involved in recruitment to the trial of or delivery of the intervention to any of the participants that they followed up, and were blind to treatment allocation” “At the start of the appointment, the RA reminded the participants of the importance of them remaining blind to group allocation and asked participants not to discuss any aspects of their involvement in the trial”
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intent-to-treat analysis used

Lincoln 2020 *(Continued)*

Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Maggio 2020

Study characteristics		
Methods	Single-blind, RCT	
Participants	n = 60 Mean age (years) (E: 51.9, C: 48.2) Mean years of education (E: 14.1, C: 15.5)	
Interventions	Sessions involved VR training which consisted of providing CR in the semi-immersive VR system that stimulate real-life scenarios Control group: traditional CR through the face-to-face approach	
Outcomes	Significant improvements in both groups both intervention but greater differences in E Significant interaction for Group*Time	
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, VR: Virtual reality, CR: Cognitive rehabilitation,	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "All patients were randomized (using a block randomization with a block size of 22) into either traditional CR (representing the control group, CG; 30 patients, 13 men and 17 women), or the experimental group undergoing VR (EG: 30 patients, 18 men and 12 women)"
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Both assessors and therapists were blinded to the patient's allocation and treatment" Unclear whether participants were blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	None apparent
Other bias	Low risk	None identified

Mani 2018

Study characteristics	
Methods	Single-blind, RCT
Participants	n = 34 Included in follow-up (E: 17, C: 17) Mean age (years) (E: 35.29, C: 35.82) Mean years of education (E: 14.5, C: 14.6)
Interventions	Group-based CR consisting of 8 sessions over 4 weeks, 2 hours long, designed as a compensatory, problem- based, and integrated approach based on learning theory and an information processing model to enhance general cognitive function Control group: non-therapeutic session for the same frequency
Outcomes	Significant improvement for E in WMS-R and ACE, as well as MFQ
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CR: Cognitive rehabilitation, WMS-R: Weschlers memory scale-revised, ACE: Addenbrooke's cognitive examination, MFQ: Memory failures questionnaire

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "Patients were randomly allocated to study (A) and control (B) groups using a table of random numbers performed by the research fellow who was not involved in CR or the cognitive assessments"
Allocation concealment (selection bias)	Low risk	Quote "To eliminate potential selection bias, this randomized allocation was performed with allocation concealment"
Blinding (performance bias and detection bias) All outcomes	Low risk	"To minimize the chance of evaluation bias, all psychological assessments (pretest, post- test, and follow- up cognitive assessments) were made by a psychiatry resident who was blinded to the grouping of patients." Unclear whether patients were blinded
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Important to note that 1 experimental participant and two control group participants were missing at follow-up
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Mattioli 2016

Study characteristics	
Methods	RCT

Mattioli 2016 *(Continued)*

Participants	n = 41 Completed (E: 15, C: 17) Mean age (years) (E: 44.80, C: 44.88) Mean years of education (E: 12.12, C: 10.93)
Interventions	Sessions involved RehaCom software to improve specific areas of cognitive difficulty individualised for each participant, 30 sessions over 15 weeks, 60 minutes long Control group: received general CR
Outcomes	Sig improvement for E in SDMT, COWAL and SPART No significant differences for MSQoL-54 No significant differences for MADRS
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CR: Cognitive rehabilitation, SDMT: Symbol digit modalities test, COWAT: Controlled oral word association test, SPART: 10/36 spatial test

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "Randomization (according to a computer-generated list of random number) and statistical analysis of data were carried out by an independent center, from whom all the Centers received the patients' number."
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	High risk	Not blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	9 participants dropped out, no clear reasons given
Selective reporting (reporting bias)	High risk	Missing data for MSQoL, mFIS and MADRS at T0and T12. Authors claim significant improvement from T0 to T12 within group but no P values. Missing SDs.
Other bias	Low risk	None identified

Mendozzi 1998

Study characteristics

Methods	Single-blind, quasi-RCT
Participants	n = 30 randomly allocated to groups, n = 30 matched on age, gender, and education
Interventions	Computerised treatment

Mendozzi 1998 *(Continued)*

	A: Specific cognitive retraining B: Non-specific cognitive retraining C: Control group 15 bi-weekly sessions, 45 minutes
Outcomes	Intention-to-treat analysis not used Specific group improved on 7 outcome measures, non-specific on 1 measure
Notes	For this review A versus C compared, because B was not considered cognitive rehabilitation; RCT: randomised controlled trial

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Only half the sample randomised, with no stated generation method
Allocation concealment (selection bias)	Unclear risk	Quote: "Assignment by principal investigator, who was not involved in the CR or cognitive testing and scoring"
Blinding (performance bias and detection bias) All outcomes	Low risk	Single blinding: quote: "the tests were always administered and scored by the same investigator who was not involved in the clinical work and was unaware of the treatment assignments"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Participant data replaced mid-trial if dropped out
Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Unclear risk	1 participant in the specific cognitive retraining group discontinued retraining and was replaced by a new entry

Messinis 2017

Study characteristics

Methods	Mutlicente, RCT
Participants	n = 58 Mean age (years) (E: 46.03, C: 45.15) Mean years of education (E: 12.12, C: 12.73)
Interventions	Sessions involved individualised CR using RehaCom software based in a clinic Control group: usual care
Outcomes	Sig improvements for E in SRT and BVMT-R

Messinis 2017 *(Continued)*

Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CR: Cognitive rehabilitation, SRT: Selective reminding test, BVMT-R: Brief visuospatial memory test-revised

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "Eligible patients were randomized by a computer -generated, site-stratified, independent randomization schedule to either undergo cognitive rehabilitation (IG; intervention group) with the RehaCom software or were placed in the placebo arm (CG; control group) and spent the same portion of time (10 weeks) receiving usual clinical care"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	High risk	Quote "The participants and clinicians taking part in the assessments and intervention were not blind to the allocated treatments"
Incomplete outcome data (attrition bias) All outcomes	Low risk	All data available
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	High risk	6-month follow-up data were not collected for the control group

Messinis 2020

Study characteristics

Methods	Multi-site, sham-controlled, RCT
Participants	n = 36 Mean age (years) (E: 46.47, C: 45.29) Mean years of education (E: 13.89, C: 13.70)
Interventions	Sessions consisted of computerised individualised, domain and task-specific CR using RehaCom, 24 sessions over 8 weeks, 45 minutes long Control group: non-specific computerised activities e.g. solving puzzles for same frequency as E
Outcomes	Significant improvement for E in SRT and BVMT-R
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control Group, CR: Cognitive rehabilitation, SRT: Selective reminding test, BVMT-R: Brief visuospatial memory test-revised

Risk of bias

Bias	Authors' judgement	Support for judgement

Messinis 2020 *(Continued)*

Random sequence generation (selection bias)	Low risk	Quote "Eligible patients were randomized by a computer – generated, site stratified, independent randomization schedule to either undergo cognitive rehabilitation (IG; intervention group) with the RehaCom software (RehaCom Cognitive Therapy Software. https:// www.rehacom.co.uk), or were placed in the placebo arm (CG; control group) and spent the same portion of time (8-weeks) receiving usual clinical care plus sham cognitive intervention"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote: "Scoring of neuropsychological measures at baseline and post treatment was performed by two blind observers, in order to avoid inter-rater variability."
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Mousavi 2018a

Study characteristics	
Methods	Double-blind, RCT
Participants	n = 60 Mean age (years) (E: 40.55, C: 40.65, P: 41.25) Education means not given
Interventions	Sessions involved teaching participants to use memory aids, compensatory strategies, and then they were given homework to practice using the techniques in daily life, 8 sessions over 8 weeks, 60 minutes long Control group: participants given ordinary information on cognitive difficulties Placebo group: participants received relaxation techniques
Outcomes	Significant improvement in WM for E Results of GHQ not reported
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, P: Placebo group, WM: Working memory, GHQ: General health questionnaire

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "Testees were compared in three groups according to demographic characteristics and duration of the disease; and, finally, they were randomly

Mousavi 2018a *(Continued)*

		assigned to three groups—namely, experimental (n = 20), placebo (n = 20), and control (n = 20)—with the help of randomized software"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	High risk	Quote "It is worth mentioning that both participants and the statistics analyzer were blind to the purposes of the study" - this is not the same as being blind to treatment allocation despite the study being labelled "double-blind"
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	High risk	Only results from one outcome measure reported
Other bias	Low risk	None identified

Mousavi 2018b

Study characteristics	
Methods	Single-blind, RCT
Participants	n = 60 Mean age (years) (E: 40.55, C: 40.65, P: 41.25) Mean education not given
Interventions	Sessions involved teaching participants to use memory aids, compensatory strategies, and then they were given homework to practice using the technigues in daily life, 8 sessions over 8 weeks, 60 minutes long Control group: given usual care Placebo group: participants received relaxation techniques
Outcomes	Significant improvement at post intervention for E but not follow-up
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, P: Placebo group

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote "Participants were randomly assigned to three groups namely, experimental (n=20), placebo (n=20) and control (n=20) with the help of random allocation software"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Quote: "The follow-up studies were performed in the presence of the patients by the centre's psychologist. It was not possible to hide allocation from the therapist. The patients and the analyst were blind to the allocation."

Mousavi 2018b *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	No attrition according to CONSORT diagram
Selective reporting (reporting bias)	High risk	Only one outcome measure reported
Other bias	Low risk	None identified

Naeeni Davarani 2020

Study characteristics	
Methods	Double-blind, RCT
Participants	n = 60 Randomised (E: 30, C: 30) Completed (E: 28, C: 26) Mean age (years) (E: 39.31, C: 37.55) Mean years of education: (E: 14.9, C: 14.7)
Interventions	Sessions involved RehaCom software that is auto-adaptive so the level of ability will automatically increase or decrease, modules were chosen based on intended cognitive functions for rehabilitation Control: usual care
Outcomes	Significant effect of treatment for E for all memory measures including PASAT
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, PASAT: Paced auditory serial addition test

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Randomisation method not specified
Allocation concealment (selection bias)	Unclear risk	No information on this
Blinding (performance bias and detection bias) All outcomes	High risk	Quote: "Certainly, the patients who received no treatment knew they were in the control group, but the therapists who conducted cognitive tests did not know."
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Intention to treat not used. 2 experimental participants and 4 control participants dropped out, reasons not given
Selective reporting (reporting bias)	Low risk	All outcomes reported

Naeeni Davarani 2020 *(Continued)*

Other bias	Low risk	None identified

Pedulla 2016

Study characteristics	
Methods	RCT
Participants	n = 28 Mean age (years) (E: 49.0, C: 46.1) Mean years of education (E: 12.8, C: 10.7)
Interventions	Sessions involved computerised home-based CR using COGNI-TRAcK which delivers deliver intensive, automatically adaptive and monitored cognitive training, 40 sessions over 8 weeks, 30 minutes long Control group: non-adaptive computerised CR
Outcomes	Significant interaction of Group*Time found for 6 of 10 subtests used
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CR: Cognitive rehabilitation

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: “The enrolled patients were randomly assigned by a blinded psychologist to the study group or to the control group (see next section).” - Randomisation method not specified
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding (performance bias and detection bias) All outcomes	High risk	Not clear if participants were blinded or who conducted assessments
Incomplete outcome data (attrition bias) All outcomes	Low risk	All outcomes reported
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Perez-Martin 2017

Study characteristics	
Methods	Observer-blinded, RCT
Participants	n = 62

Perez-Martin 2017 *(Continued)*

	Mean age (years) (E: 44.93, C: 40.88) Mean years of education (E: 10.21, C: 11.59)
Interventions	Sessions focused on attention, processing speed, memory and executive functions through computerised and paper and pencil tasks, 12 sessions over 12 weeks, 60-75 minutes long Control group: received information about their cognitive status and a booklet containing general advice
Outcomes	Significant decrease in MSNQ score for E, sig improvement in verbal and working memory for E Both groups sig increase in MSQoL-54 Significant decrease in HADS on both depression and anxiety scales for E
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, MSNQ: Multiple sclerosis neuropsychological screening questionnaire

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomisation method not specified
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding (performance bias and detection bias) All outcomes	Low risk	Participants not blinded to group allocation Quote “The neuropsychologist who performed the posttreatment cognitive assessment was blinded to the group allocation of the patients”
Incomplete outcome data (attrition bias) All outcomes	Low risk	No attrition for post intervention, 8 lost to follow-up
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Pusswald 2014

Study characteristics

Methods	Alternating allocation
Participants	n = 40 (Intervention: 20, Control: 20) Both groups comparable on clinical and sociodemographic baseline characteristics
Interventions	Cognitive functional training, computer-based home training of divided attention, carried out 3/week for 30 minutes for 5 weeks alongside weekly 90-minute sessions in groups focusing on cognitive rehabilitation techniques and approaches, and included memory retraining.

Pusswald 2014 *(Continued)*

	Control group received no specific training.
Outcomes	Significant within-group effect on objective memory for intervention group when comparing before training to after training
Notes	None identified

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Alternating allocation
Allocation concealment (selection bias)	Unclear risk	No mention of allocation concealment
Blinding (performance bias and detection bias) All outcomes	High risk	Outcome assessor not blind
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No mention of incomplete data
Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Low risk	None identified

Rahmani 2020

Study characteristics

Methods	Quasi-experimental with a pre-test-post-test, 2-months follow-up, a placebo, and a control group design
Participants	n = 60 (12 participants randomised to each group) Mean age (years) (Comp: 30.17, Man: 29.41, Comb: 27.83, P: 31.16, C: 29.70) Mean years of education (Total: 14.7)
Interventions	Comp: only computerised CR Man: 0nly manual-based CR Comb: both computerised and manual-based CR Sessions included four main steps: remediation, substitution, accomodation, and assimilation P: physical rehabilitation C: usual care

Rahmani 2020 *(Continued)*

Outcomes	Significant improvements in PASAT for all three intervention groups
Notes	RCT: randomised controlled trial, Comp: Computer-based group, Man: Manual-based group, Comb: Combined group, P: Placebo group, C: Control group, CR: Cognitive rehabilitation, PASAT: Paced auditory serial addition test

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Randomisation method not specified
Allocation concealment (selection bias)	Unclear risk	No information given
Blinding (performance bias and detection bias) All outcomes	High risk	No mention of blinding
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Rilo 2018

Study characteristics

Methods	Non-blinded parallel group randomised trial with equal randomisation
Participants	n = 42 Randomised (E: 21, C: 21) Completed (E: 20, C: 20) Mean age (years) (E: 43.90, C: 43.67) Mean years of education (E: 13.00, C: 13.95)
Interventions	Sessions involved REHACOP which is an integrative cognitive rehabilitation programme based on the principles of restoration, compensation and optimisation, 39 sessions over 12 weeks, 60 minutes long
Outcomes	Significant improvements for intervention group, large effect size for working memory
Notes	E: Experimental group, C: Control group

Risk of bias

Bias	Authors' judgement	Support for judgement

Rilo 2018 *(Continued)*

Random sequence generation (selection bias)	Low risk	Quote "Patients were randomly assigned to each study condition using an online computer-generated random number at a ratio of 1:1"
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	High risk	Non-blinded study
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	One experimental participant and one control group participant dropped out
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Shahpouri 2019

Study characteristics	
Methods	Double-blinded, RCT
Participants	n = 56 Mean age (years) (E: 32.21, C: 30.46) Mean education not given
Interventions	Sessions involved the mnemonic approach which includes visual imagery, theological organisation, and relational strategies including mnemonics of fiction, the clues about the first word, chain connection, and the technique of PQRST, 10 sessions, 120 minutes long Control group: patients were requested to present their experiences of cognitive impairments, and cases with successful coping with new conditions were admired for the same frequency as E
Outcomes	Significant improvement of EMQ for E Significant. improvement for physical and mental scales for MSQoL-54 Significant improvement of BDI for E
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, PQRST: Preview, Question, Read, `Self-recitation`, and Test, EMQ: Everyday memory questionnaire, MSQoL-54: Multiple sclerosis quality of life-54, BDI: Beck depression inventory

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Therefore, each patient was provided with a particular number using the mentioned software that allocated him/her to either the control group or the intervention group. A random number was assigned to each patient, and individuals with even numbers were allocated to the intervention group"

Shahpouri 2019 *(Continued)*

Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	Low risk	Double-blind study
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Solari 2004

Study characteristics	
Methods	Independent randomisation; computer-generated, site-stratified schedule; double-blind
Participants	n = 82 Randomised: E: 42; C: 40 Analysed: E: 40; C: 37 Age: E: 46.2 years (SD: 9.2); C: 41.2 years (SD: 10.6) Education: E: 21 C: 20 high school+
Interventions	Individual treatment 45 minutes, 2 per week, 8 weeks Computerised programmes E: memory and attention retraining C: visuoconstruction and visuomotor co-ordination
Outcomes	Intention-to-treat analysis used No significant differences between groups on any measures at follow-up 1 or 2, when Bonferroni adjustments made
Notes	E: Experimental; C: Control; SD: standard deviation

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "were assigned to one of the two interventions by an independent randomisation unit, using a computer-generated, site-stratified, randomisation schedule."

Solari 2004 *(Continued)*

Allocation concealment (selection bias)	Low risk	see above
Blinding (performance bias and detection bias) All outcomes	Low risk	Participants, healthcare providers, and outcome assessors all blinded. Outcome assessor asked to guess participant group
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote:"Missing values were imputed according to the 'last observation carried forward' method"
Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Low risk	None identified

Stuifbergen 2012

Study characteristics	
Methods	Single-blind randomised controlled trial
Participants	n = 61 (Intervention: 34, Control: 27) Age range 24 to 60 years, mean: 47.95 Length of time since diagnosis range 1 to 29 years, mean: 12.2
Interventions	Group-based; MAPSS-MS (Memory, Attention, and Problem Solving Skills for People with Multiple Sclerosis). 8 weekly, 2-hour group sessions focused on building efficacy for use of compensatory strategies, and use of a computer-assisted training programme. Home-based practice using the computer program.
Outcomes	Significant difference between groups on CVLT-total (medium effect size) and Strategy subscale of the Multifactorial Memory Questionnaire (large effect size), E > C. Both groups improved over time on neuropsychological testing, ADLs, and use of compensatory strategies.
Notes	CVLT-total: California Verbal Learning Test; ADL: Activities of daily living; E: Experimental, C: Control

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Prior to the initiation of data collection, the data analysts for the project generated a random number sequence for randomization to intervention and control"
Allocation concealment (selection bias)	Low risk	Each allocation placed in sealed envelope prior to study start and opened by project director when participant randomised, to let them know their allocation

Stuifbergen 2012 *(Continued)*

Blinding (performance bias and detection bias) All outcomes	Low risk	Quote:"The staff members conducting neuropsychological assessments were blinded to participants' group assignment". States that those involved in intervention were not involved in collecting, entering, or analysing data
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intention to treat; participant analysed if completed baseline and attended at least 1 class. Missing values replaced using last observation carried forward. If participant missed time point 2, but completed 1 and 3, then 2 was an average of 1 and 3
Selective reporting (reporting bias)	Low risk	No selective reporting apparent, full analyses available
Other bias	Low risk	None identified

Stuifbergen 2018

Study characteristics	
Methods	Double-blind, RCT
Participants	n = 183 Randomised (E: 93, C: 90) Completed T2 (E: 78, C: 85) Complete T3 (E: 74, C: 82) Completed T4 (E: 67, C: 83) Mean age (years) (E: 49.8, C: 49.4) Mean years of education (E: 14.9, C: 14.9)
Interventions	Sessions involved the Luminosity programme that delivers interactive programs that run directly in standard web browsers and is designed to adapt to the individual user and offers novel, engaging, and challenging tasks within an integrated, hierarchical structure, 24 sessions over 8 weeks, 45-60 minutes long
Outcomes	Intention-to-treat analysis use Marginial improvements for subjective measures Immediate significant improvements for intervention group for CVLT and PASAT, at follow-up significant improvements for PASAT Significant Time*Group effect, IG scored sig. lower post intervention and 3 month follow up but not 6 month for CES-D
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CVLT-II: California verbal learning trials-second version, PASAT: paced auditory serial addition test, CES-D: Center for Epidemiological Studies-Depression
Risk of bias	

Bias	**Authors' judgement**	**Support for judgement**

Stuifbergen 2018 *(Continued)*

Random sequence generation (selection bias)	Low risk	Quote "Prior to the initiation of data collection, we used a 1:1 ratio to randomly assign participants to groups. Randomization assignment was recorded on a letter sealed in an opaque envelope. Following the completion of baseline testing, the project manager opened the next envelope in the sequence and assigned the participant to either the active intervention or the usual care comparison group."
Allocation concealment (selection bias)	Low risk	See above
Blinding (performance bias and detection bias) All outcomes	High risk	Quote "The testers conducting neuropsychological assessments were blinded to participants' group assignment and participants were not informed of their specific group assignment (intervention or comparison)"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Intent-to-treat analysis used
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Tesar 2005

Study characteristics	
Methods	Simple random sampling with independent allocation
Participants	n = 19 (E: 10; C: 9) Mild-moderate cognitive deficits Groups comparable on baseline variables
Interventions	Group treatment E: 12 1-hour sessions in 4 weeks; neuropsychological training programme; computer-based direct functional training internal and external memory C: rehabilitation only
Outcomes	Intention-to-treat analysis not used Significant differences between groups seen only on CKV and HAWIE-R (but practice effects as no parallel forms used?) No other significant differences on other measures Based on feedback interview, authors conclude treatment effectiveness
Notes	E: Experimental; C: Control; CKV: Computer-aided card-sorting procedure; HAWIE-R: Hamburg Wechsler Intelligence Test–Revised
Risk of bias	
Bias	**Authors' judgement** **Support for judgement**

Tesar 2005 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Quote: "simple random sampling"
Allocation concealment (selection bias)	Low risk	Quote: "Allocation to the two study groups (treated and control group) was done by a person who worked in an out-patient MS facility and who was not involved in the study"
Blinding (performance bias and detection bias) All outcomes	High risk	Quote: "The participants were aware of each intervention" but no indication of assessor blinding
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No mention of how dropout data handled
Selective reporting (reporting bias)	Low risk	No selective reporting apparent
Other bias	Low risk	None identified

Vilou 2020

Study characteristics	
Methods	Observer-blinded, RCT
Participants	n = 47 Mean age (years) (E: 33.5, C: 37.8) Mean education not given
Interventions	Computerised home-based CR, activities were set for the participants in advance and a trained psychologist called weekly to check on progress, 12 sessions over 6 weeks, unknown length of time Control group: not clear
Outcomes	Significant improvements in intervention group for BVMT, GVLT and TMT, these improvements could be seen at follow-up as well
Notes	RCT: randomised controlled trial, E: Experimental group, C: Control group, CR: Cognitive rehabilitation, BVMT-R: Brief visuospatial memory test-revised, GVLT: Greek verbal learning test, TMT: Trail making test

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote “Patients were randomized into the intervention group (n = 23) and the control group (n = 24) by an automated randomization software”
Allocation concealment (selection bias)	Low risk	see above
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote “Neuropsychological assessment was performed by an experienced neuropsychologist who was blinded to the patient's group allocation and was

Vilou 2020 *(Continued)*

		carried out in a quiet room with no distractions at the neuropsychology laboratory of the clinic"
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	None identified

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Aguirre 2019	Not MS, healthy control group
Aisen 1994	Non-RCT with a mixed aetiology group, non-memory
Akhtar 2006	Not MS
Aldrich 1995	Not just MS, non-RCT
Allen 1998	Non-RCT, no control
Amato 2014	Focus on attention, not memory
Barbarulo 2018	Non-RCT, no random allocation
Barker 2019	Non-RCT, no random allocation
Bhargav 2016	Not memory rehabilitation
Bonavita 2015	No memory focus, active control
Bove 2019	Non-RCT, no control group
Brenk 2008	Non-RCT, allocated participants by demographics
Brissart 2010	Non-RCT, no control group
Brissart 2013	Non-RCT
Cabrera-Gomez 2010	No memory focus, non-RCT
Campbell 2015	Conference poster presentation, no full text available
Canellopoulou 1998	Not memory rehabilitation, not MS control group
Chiaravalloti 2003	Non-RCT, healthy controls
Chiaravalloti 2012	Same sample as Chiaravalloti 2013

Study	Reason for exclusion
Chiaravalloti 2018	Not a memory rehabilitation
das Nair 2017	Conference poster presentation, no full text available
De Giglio 2014	Not a memory rehabilitation
Dobryakova 2014	Same sample as Chiaravalloti 2013
Ernst 2013	Non-RCT, healthy controls
Fiorotto 2015	Full text not in English, no translation available
Flavia 2010	Not memory rehabilitation
Goreover 2011	Not memory rehabilitation
Grasso 2017	Not a memory rehabilitation
Güçlü Altun 2015	Non-RCT, no control group
Guijarro-Castro 2017	Study protocol, no results available
Hanssen 2016	Not a memory rehabilitation
Harand 2017	Conference poster presentation, no full text available
Harand 2019	Study protocol, no results available
Iaffaldano 2015	Conference poster presentation, no full text available
Jimenez-Morales 2017	Full text not in English, no translation available
Kavaklioglu 2017	Conference poster presentation, no full text available
Lamargue 2020	Not MS, healthy control group
Leavitt 2014	Subgroup analysis of Chiaravalloti 2013
Lincoln 2015	Study protocol, results reported in later publication
Loewenstein 2004	Not MS: Alzheimer's disease
Mäntynen 2014	Not memory specific
Martin 2014	Subgroup analysis of das Nair 2012
Mattioli 2012	Not memory specific
Messinis 2015	Conference poster presentation, no full text available
Moore 2008	No rehabilitation, as intervention only involved 1 session of 1 hour
Nauta 2017	Study protocol, no results available
Nurova 2014	Conference poster presentation, no full text available

Study	Reason for exclusion
Penner 2018	Conference poster presentation, no full text available
Perez-Martin 2016	Conference poster presentation, no full text available
Pineau 2019	Non-RCT, no random allocation
Rilo 2015	Conference poster presentation, no full text available
Rilo 2016	Conference poster presentation, no full text available
Rilo 2017	Conference poster presentation, no full text available
Rodgers 1996	Non-RCT
Rosti-Otajärvi 2013a	Not memory specific
Rosti-Otajärvi 2013b	Not memory specific
Shatil 2010	Non-RCT
Thaut 2014	No cognitive rehabilitation
Thomas 2006	Non-RCT: systematic review
van der Putten 1999	Stroke and MS patients, non-RCT, non-memory
Veldkamp 2019	Not a memory rehabilitation
Vogt 2009	No MS control group, only healthy controls
Wilson 2001	The authors do not distinguish results for participants with MS from those for participants with acquired progressive brain injury; no MS control group
Zimmer 2018	Not a memory rehabilitation

MS: multiple sclerosis; RCT: randomised controlled trial

Characteristics of ongoing studies *[ordered by study ID]*

ISRCTN54901925

Study name	A randomised study of cognitive rehabilitation in multiple sclerosis
Methods	Randomised controlled trial
Participants	Planned sample size: 50 Adult Participant inclusion criteria 1. Diagnosis of MS by consultant neurologist to best current criteria. 2. Able and willing to give informed consent. 3. Cognitive impairment defined by scoring below 5th percentile on 1 or more of BICAMS scales (Langdon 2012) as identified at the clinic. 4. Willing to commit to 3x 45-minute computer training sessions for 6 weeks.

ISRCTN54901925 *(Continued)*

	5. Home PC fulfilling experimental spec. 6. Willing to attend total of 3 MRI scans at the University of Sussex MRI scanner. 7. Age between 18 and 70. Participant exclusion criteria 1. Significant change in medications in last 4 weeks. 2. Relapse recovery within last 4 weeks. 3. Sensorimotor dysfunction likely to interfere with PC interface. 4. Significant psychiatric history/condition. 5. Significant medical condition (other than MS), personal or social circumstances likely to influence cognition or study participation. 6. Women who are pregnant.
Interventions	Participants will be randomised to undergo either cognitive rehabilitation with RehaCom Software (3x 45-minute training sessions per week for 6 weeks) or be placed in the placebo arm to spend the same amount of time in the control condition (natural history DVDs). During this period, they will be expected to undertake 3 x 45-minute computer training sessions per week for the 6-week period. There will also be an MRI brain scan at baseline prior to undertaking the training. Following completion of the 6-week training period, both the full cognitive assessments and MRI scanning will be repeated immediately following the training period and again at approximately 3 to 6 months
Outcomes	**Primary outcome measures** 1. Objective cognitive performance: BICAMS (a 15-minute screening tool). 2. Quality of life: a. EQ-5D, a generic health-related quality of life scale (EuroQoL Group 1990) b. Functional Assessment of MS (FAMS), an MS-specific quality-of-life scale (Cella 1996) **Secondary outcome measures** MRI: The data will be acquired on the 1.5T Siemens machine. The following analyses will be completed: 1. Voxel-based morphometry 2. Tensor-based morphometry 3. Cortical thickness 4. Lesion load 5. Resting state analysis (default mode network) 6. Diffusion tensor imaging analysis
Starting date	November 2013
Contact information	Dr Waqar Rashid Department of Neurology Royal Sussex County Hospital Eastern Road Brighton BN25BE
Notes	

NCT03471338

Study name	Neuropsychological management of multiple sclerosis: benefits of a computerised semi-autonomous at-home cognitive rehabilitation programme
Methods	Randomised controlled trial
Participants	Inclusion Criteria • MS defined according to the McDonald criteria revised in 2010 • Men and women aged between 18 and 65 years • RR and SP forms • Duration of progression ≤ 25 years • EDSS ≤ 5.5 • Lack of disease activity as defined by the new Lublin criteria (2013) • Cognitive complaint and/or cognitive disorders according to the investigator's judgement • Impaired cognitive performance at least 1.65 SD below normative data at one test of the BCcogSEP battery • French native language • Owner of a laptop computer with Internet access • Signing of the informed consent Exclusion Criteria • - Other neurological, psychiatric or developmental diseases prior to the MS diagnosis • Cranial trauma sequelae • Chronic alcohol and/or drug consumption • EDSS > 6 • Relapse and/or treatment with corticosteroids within the past month • Persons deprived of liberty, minors, adults under wardship • Cognitive examination within the past 6 months (including in particular all or some of the tests proposed by this project) • Presence of dementia according to DSM V criteria, or of cognitive disorders preventing the patient from undergoing cognitive tests or performing cognitive rehabilitation exercises • Any visual or motor deficit preventing the patient from undergoing cognitive tests or performing cognitive rehabilitation exercises
Interventions	**Experimental: Experimental Group** Behavioural: Cognitive rehabilitation At-site inclusion visit: assessment of patient's eligibility by cognitive complaint questionnaire and BCcogSEP, VAPS and multiple errands test conducted by neuropsychologist. At-site baseline visit: assessment of quality of life (MUSIQOL), self-esteem (SEI), depression (MADRS), anxiety (HAMA), BICAMS: SDMT, CVLT-II, BVMTR, metacognition (MCQ-30), fatigue (EMIF-SEP), subjective sleep quality (PSQI) conducted by a neuropsychologist. At-home neuropsychological management (9 weeks): The patient performs the program (PRESCO software) on his computer autonomously at home at a rate of 3 sessions per week. A neuropsychologist performs at-home visits and weekly phone meetings to train the patient to the software, to encourage him to do exercises and to answer any software use-related questions. At-site follow-up visits: short and long-term retest of assessments performed in inclusion visit. **Sham Comparator: Standard Psychological care** Behavioral: Standard Psychological care At-site inclusion visit: assessment of patient's eligibility by cognitive complaint questionnaire and BCcogSEP, VAPS and multiple errands test conducted by neuropsychologist.

NCT03471338 *(Continued)*

	At-site baseline visit: assessment of quality of life (MUSIQOL), self-esteem (SEI), depression (MADRS), anxiety (HAMA), BICAMS: SDMT, CVLT-II, BVMTR, metacognition (MCQ-30), fatigue (EMIF-SEP), subjective sleep quality (PSQI) conducted by a neuropsychologist. At-home neuropsychological management (9 weeks): A neuropsychologist performs at-home visits and weekly phone meetings consisting in discussion of the patient's cognitive disorders. At-site follow-up visits: short and long-term retest of assessments performed in inclusion visit.
Outcomes	**Primary Outcome Measures:** • Efficacy of cognitive rehabilitation on quality of life at short term. Quality of life will be assessed by measuring the change of the scores of MUSIQOL (MUltiple Sclerosis International Quality Of Life) questionnaire between baseline and short-term visits. Efficacy will be assessed by comparing theses scores between groups A and B. • Efficacy of cognitive rehabilitation on quality of life at short term. Quality of life will be assessed by measuring the change of the scores of MUSIQOL (MUltiple Sclerosis International Quality Of Life) questionnaire between baseline and short-term visits. Efficacy will be assessed by comparing theses scores between groups A and B.
Starting date	March 2018
Contact information	Gilles Defer, Pr Telephone: 231064620 Ext. +33 Email: defer-gi@chu-caen.fr
Notes	

BICAMS: Brief International Cognitive Assessment for Multiple Sclerosis; **CVLT:** California Verbal Learning Test;**DSM-5:** Diagnostic and Statistical Manual of Mental Disorders; ; **EDSS:** Expanded Disability Status Scale; **MADRS:** Montgomery-Asberg Depression Rating Scale; **MRI:** Magnetic resonance imaging; **PSQI**: Pittsburgh Sleep Quality Index; **SD:** standard deviation; **SDMT**: Symbol Digit Modalities Test.

DATA AND ANALYSES

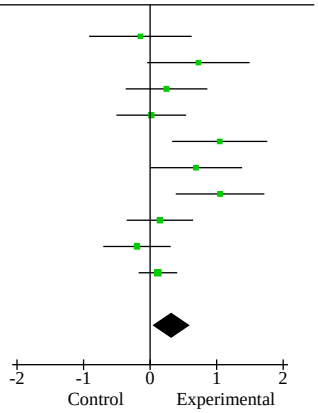
Comparison 1. Subjective memory measures

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Immediate	10	568	Std. Mean Difference (IV, Random, 95% CI)	0.32 [0.05, 0.58]
1.2 Intermediate	11	1045	Std. Mean Difference (IV, Random, 95% CI)	0.23 [0.11, 0.35]
1.3 Longer-term	5	775	Std. Mean Difference (IV, Random, 95% CI)	0.16 [0.02, 0.30]

Analysis 1.1. Comparison 1: Subjective memory measures, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chiaravalloti 2005	77.67	26.69	13	81.38	22.19	13	7.3%	-0.15 [-0.92 , 0.62]
Chiaravalloti 2019b	15.84	5.52	15	11.58	5.9	13	7.3%	0.73 [-0.04 , 1.50]
Chmelařová 2020	-61.27	20.52	26	-65.71	11.88	17	9.6%	0.25 [-0.37 , 0.86]
Goodwin 2020	-40.3	22.9	27	-40.7	21.3	29	11.2%	0.02 [-0.51 , 0.54]
Goverover 2018a	250.2	4.7	19	209.4	56.3	16	8.1%	1.05 [0.33 , 1.76]
Mani 2018	49.03	6.05	17	43.94	8.22	17	8.3%	0.69 [-0.01 , 1.38]
Mousavi 2018b	123.5	3.06	20	120.65	2.18	20	8.8%	1.05 [0.39 , 1.72]
Perez-Martin 2017	-23.87	11.83	30	-25.63	11.83	32	11.7%	0.15 [-0.35 , 0.65]
Stuifbergen 2012	-28.41	11.13	34	-26.15	11.56	27	11.5%	-0.20 [-0.70 , 0.31]
Stuifbergen 2018	41.7	11	93	40.4	11.1	90	16.2%	0.12 [-0.17 , 0.41]
Total (95% CI)			**294**			**274**	**100.0%**	**0.32 [0.05 , 0.58]**

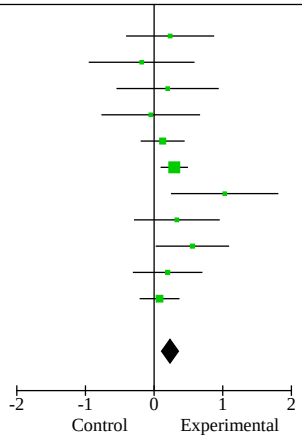
Heterogeneity: Tau² = 0.09; Chi² = 19.07, df = 9 (P = 0.02); I² = 53%
Test for overall effect: Z = 2.37 (P = 0.02)
Test for subgroup differences: Not applicable



Analysis 1.2. Comparison 1: Subjective memory measures, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carr 2014	-21.7	13.1	17	-25.8	19.9	21	3.7%	0.23 [-0.41 , 0.88]
Chiaravalloti 2005	79	29.31	13	83.92	22.53	13	2.5%	-0.18 [-0.95 , 0.59]
Chiaravalloti 2019b	16	18.11	15	12.5	16.22	13	2.7%	0.20 [-0.55 , 0.94]
das Nair 2012	-42.48	22.44	29	-41.4	18.95	10	2.9%	-0.05 [-0.77 , 0.67]
Lincoln 2002	-23.32	20.28	74	-26.11	24.02	76	14.7%	0.12 [-0.20 , 0.45]
Lincoln 2020	-37.6	23.4	214	-44.5	23.5	181	38.1%	0.29 [0.09 , 0.49]
Mani 2018	51.28	4.31	14	44.41	8.01	15	2.5%	1.03 [0.25 , 1.81]
Mousavi 2018b	121.2	1.57	20	120.7	1.38	20	3.9%	0.33 [-0.29 , 0.96]
Shahpouri 2019	-92.93	44.29	28	-112.57	21.14	28	5.3%	0.56 [0.02 , 1.09]
Stuifbergen 2012	28.41	11.13	34	26.15	11.56	27	5.9%	0.20 [-0.31 , 0.70]
Stuifbergen 2018	40.5	11.4	93	39.6	11.2	90	17.9%	0.08 [-0.21 , 0.37]
Total (95% CI)			**551**			**494**	**100.0%**	**0.23 [0.11 , 0.35]**

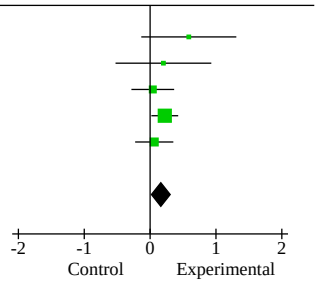
Heterogeneity: Tau² = 0.00; Chi² = 9.09, df = 10 (P = 0.52); I² = 0%
Test for overall effect: Z = 3.69 (P = 0.0002)
Test for subgroup differences: Not applicable



Analysis 1.3. Comparison 1: Subjective memory measures, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carr 2014	-17.3	11.2	15	-26.9	19.3	16	3.9%	0.59 [-0.13 , 1.31]
das Nair 2012	-40.44	22.55	27	-45	20.64	10	3.8%	0.20 [-0.53 , 0.93]
Lincoln 2002	-22.37	23.62	73	-23.3	21.86	73	19.2%	0.04 [-0.28 , 0.37]
Lincoln 2020	-37.9	22.7	210	-43.1	24	168	48.9%	0.22 [0.02 , 0.43]
Stuifbergen 2018	40.2	11	93	39.5	11.2	90	24.1%	0.06 [-0.23 , 0.35]
Total (95% CI)			**418**			**357**	**100.0%**	**0.16 [0.02 , 0.30]**

Heterogeneity: Tau² = 0.00; Chi² = 2.68, df = 4 (P = 0.61); I² = 0%
Test for overall effect: Z = 2.24 (P = 0.03)
Test for subgroup differences: Not applicable



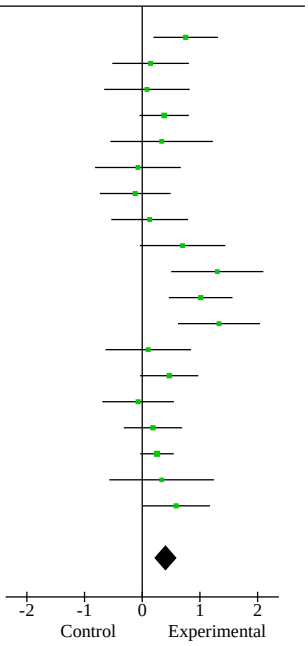
Comparison 2. Objective verbal memory

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
2.1 Immediate	19	922	Std. Mean Difference (IV, Random, 95% CI)	0.40 [0.22, 0.58]
2.2 Intermediate	6	753	Std. Mean Difference (IV, Random, 95% CI)	0.25 [0.11, 0.40]
2.3 Longer-term	4	619	Std. Mean Difference (IV, Random, 95% CI)	0.13 [-0.03, 0.29]

Analysis 2.1. Comparison 2: Objective verbal memory, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Arian Darestani 2020	56.54	14.29	27	47.12	9.84	26	5.9%	0.75 [0.20 , 1.31]
Campbell 2016	45.32	9.56	17	43.89	9.73	18	4.8%	0.14 [-0.52 , 0.81]
Chiaravalloti 2005	26.57	3.69	14	26.29	2.89	14	4.1%	0.08 [-0.66 , 0.82]
Chiaravalloti 2013	50.13	11.99	45	45.24	13.44	41	7.9%	0.38 [-0.05 , 0.81]
Chiaravalloti 2019a	51.67	11.11	9	47.36	13.19	11	3.1%	0.34 [-0.55 , 1.22]
Chiaravalloti 2019b	48.4	12.6	15	49.54	17.19	13	4.1%	-0.07 [-0.82 , 0.67]
Gich 2015	51.3	8.8	21	52.3	7.3	20	5.3%	-0.12 [-0.73 , 0.49]
Goverover 2018a	53.5	10.5	19	52.1	10.6	16	4.8%	0.13 [-0.54 , 0.80]
Hancock 2015	54.75	8.7	15	46.79	13.02	15	4.1%	0.70 [-0.04 , 1.44]
Impellizzeri 2020	40.25	7.64	15	31.41	5.39	15	3.7%	1.30 [0.50 , 2.10]
Messinis 2017	43.47	8.09	32	36.38	5.06	26	6.0%	1.01 [0.46 , 1.56]
Messinis 2020	58.1	8.3	19	47.35	7.5	19	4.4%	1.33 [0.62 , 2.04]
Pedulla 2016	39.79	11.75	14	38.33	15.13	14	4.1%	0.10 [-0.64 , 0.85]
Perez-Martin 2017	41.4	14.91	30	34	16.26	32	6.7%	0.47 [-0.04 , 0.97]
Rilo 2018	24.48	4.63	20	24.81	4.42	20	5.2%	-0.07 [-0.69 , 0.55]
Stuifbergen 2012	52.5	12.3	34	50.2	12.1	27	6.6%	0.19 [-0.32 , 0.69]
Stuifbergen 2018	53	12.6	93	49.9	11.5	90	10.4%	0.26 [-0.04 , 0.55]
Tesar 2005	52	8.2	10	48.2	13.1	9	3.0%	0.34 [-0.57 , 1.25]
Vilou 2020	63.7	17	23	54.4	14	24	5.6%	0.59 [0.00 , 1.17]
Total (95% CI)			**472**			**450**	**100.0%**	**0.40 [0.22 , 0.58]**

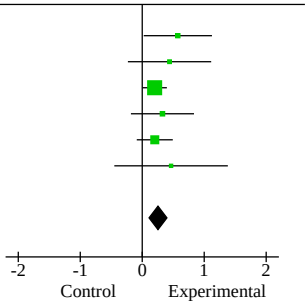
Heterogeneity: Tau² = 0.06; Chi² = 29.59, df = 18 (P = 0.04); I² = 39%
Test for overall effect: Z = 4.44 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 2.2. Comparison 2: Objective verbal memory, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Arian Darestani 2020	54	14.7	27	46.62	10.1	26	6.8%	0.57 [0.02 , 1.12]
Campbell 2016	51.99	7.56	17	47.95	10.1	18	4.6%	0.44 [-0.23 , 1.11]
Lincoln 2020	45.6	10.5	220	43.5	10.4	182	53.5%	0.20 [0.00 , 0.40]
Stuifbergen 2012	58.4	13.6	34	53.8	14.3	27	8.0%	0.33 [-0.18 , 0.84]
Stuifbergen 2018	57.2	12.3	93	54.7	12.3	90	24.6%	0.20 [-0.09 , 0.49]
Tesar 2005	56.9	13.1	10	50.4	13.6	9	2.5%	0.47 [-0.45 , 1.38]
Total (95% CI)			**401**			**352**	**100.0%**	**0.25 [0.11 , 0.40]**

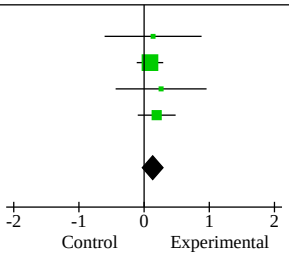
Heterogeneity: Tau² = 0.00; Chi² = 2.29, df = 5 (P = 0.81); I² = 0%
Test for overall effect: Z = 3.46 (P = 0.0005)
Test for subgroup differences: Not applicable



Analysis 2.3. Comparison 2: Objective verbal memory, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chiaravalloti 2019b	46.25	14.73	15	43.83	19.41	13	4.5%	0.14 [-0.61 , 0.88]
Lincoln 2020	47.5	10.9	206	46.5	11.3	170	60.7%	0.09 [-0.11 , 0.29]
Mattioli 2016	41.81	8.41	15	38.59	14.49	17	5.1%	0.26 [-0.44 , 0.96]
Stuifbergen 2018	56.1	12.9	93	53.6	12.9	90	29.7%	0.19 [-0.10 , 0.48]
Total (95% CI)			**329**			**290**	**100.0%**	**0.13 [-0.03 , 0.29]**

Heterogeneity: Tau² = 0.00; Chi² = 0.46, df = 3 (P = 0.93); I² = 0%
Test for overall effect: Z = 1.63 (P = 0.10)
Test for subgroup differences: Not applicable



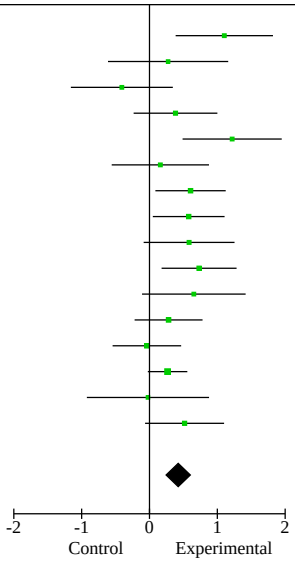
Comparison 3. Objective visual memory

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
3.1 Immediate	16	799	Std. Mean Difference (IV, Random, 95% CI)	0.42 [0.25, 0.60]
3.2 Intermediate	6	751	Std. Mean Difference (IV, Random, 95% CI)	0.20 [-0.11, 0.50]
3.3 Longer-term	4	619	Std. Mean Difference (IV, Random, 95% CI)	0.12 [-0.13, 0.37]

Analysis 3.1. Comparison 3: Objective visual memory, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	25.28	5.18	17	19.99	4.17	18	4.8%	1.10 [0.39 , 1.82]
Chiaravalloti 2019a	45.56	9.68	9	41.3	17.97	11	3.4%	0.27 [-0.61 , 1.16]
Chiaravalloti 2019b	13.8	4.46	15	15.77	4.95	13	4.5%	-0.41 [-1.16 , 0.34]
Chmelařová 2020	106.9	10.9	26	101.9	15.3	17	6.0%	0.38 [-0.23 , 1.00]
Goverover 2018a	15.6	2.6	19	12.3	2.7	16	4.7%	1.22 [0.49 , 1.95]
Impellizzeri 2020	17	4.76	15	16.25	4.35	15	4.8%	0.16 [-0.56 , 0.88]
Maggio 2020	16.6	9.68	30	11.7	5.83	30	7.7%	0.61 [0.09 , 1.12]
Messinis 2017	24.5	5.85	32	20.8	6.85	26	7.5%	0.58 [0.05 , 1.11]
Messinis 2020	18.89	4.7	19	15.88	5.4	17	5.4%	0.58 [-0.09 , 1.25]
Naeeni Davarani 2020	21.96	3.08	28	19.65	3.15	26	7.1%	0.73 [0.18 , 1.28]
Pedulla 2016	19.13	5.46	14	15.99	3.7	14	4.4%	0.65 [-0.11 , 1.42]
Perez-Martin 2017	22.77	5.56	30	21.38	4.14	32	8.0%	0.28 [-0.22 , 0.78]
Stuifbergen 2012	23.8	7.6	34	24.1	7.8	27	7.9%	-0.04 [-0.54 , 0.47]
Stuifbergen 2018	24.3	6.1	93	22.6	6.6	90	13.9%	0.27 [-0.02 , 0.56]
Tesar 2005	42.3	8	10	42.5	8.9	9	3.3%	-0.02 [-0.92 , 0.88]
Vilou 2020	27.5	10	23	22.5	9	24	6.6%	0.52 [-0.06 , 1.10]
Total (95% CI)			**414**			**385**	**100.0%**	**0.42 [0.25 , 0.60]**

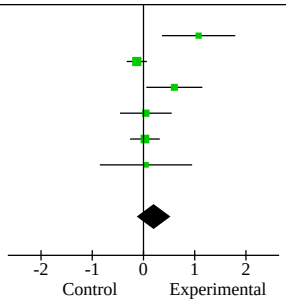
Heterogeneity: Tau² = 0.04; Chi² = 21.50, df = 15 (P = 0.12); I² = 30%
Test for overall effect: Z = 4.65 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 3.2. Comparison 3: Objective visual memory, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	27.92	5.07	17	22.19	5.32	18	11.3%	1.08 [0.36 , 1.79]
Lincoln 2020	19.1	5.3	217	19.8	5.4	182	25.9%	-0.13 [-0.33 , 0.07]
Naeeni Davarani 2020	20.69	2.85	28	18.77	3.43	26	15.1%	0.60 [0.06 , 1.15]
Stuifbergen 2012	24.9	6	34	24.6	6.9	27	16.3%	0.05 [-0.46 , 0.55]
Stuifbergen 2018	21.9	6.8	93	20.7	61	90	23.1%	0.03 [-0.26 , 0.32]
Tesar 2005	49	14.9	10	48.3	12.2	9	8.3%	0.05 [-0.85 , 0.95]
Total (95% CI)			**399**			**352**	**100.0%**	**0.20 [-0.11 , 0.50]**

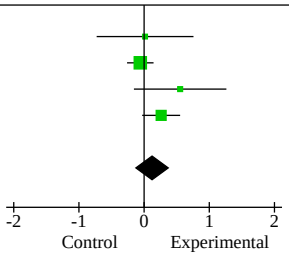
Heterogeneity: Tau² = 0.08; Chi² = 15.00, df = 5 (P = 0.01); I² = 67%
Test for overall effect: Z = 1.25 (P = 0.21)
Test for subgroup differences: Not applicable



Analysis 3.3. Comparison 3: Objective visual memory, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chiaravalloti 2019b	16.25	4.94	15	16.17	5.67	13	9.5%	0.01 [-0.73 , 0.76]
Lincoln 2020	20.1	4.9	206	20.4	5.4	170	45.7%	-0.06 [-0.26 , 0.14]
Mattioli 2016	20.19	4.49	15	17.06	6.28	17	10.3%	0.55 [-0.16 , 1.26]
Stuifbergen 2018	21.9	7	93	20.1	6.7	90	34.5%	0.26 [-0.03 , 0.55]
Total (95% CI)			**329**			**290**	**100.0%**	**0.12 [-0.13 , 0.37]**

Heterogeneity: Tau² = 0.02; Chi² = 5.00, df = 3 (P = 0.17); I² = 40%
Test for overall effect: Z = 0.96 (P = 0.33)
Test for subgroup differences: Not applicable



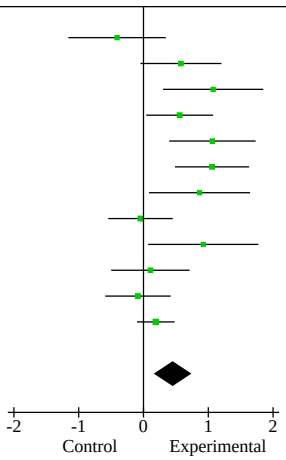
Comparison 4. Objective working memory

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
4.1 Immediate	12	655	Std. Mean Difference (IV, Random, 95% CI)	0.45 [0.18, 0.72]
4.2 Intermediate	8	821	Std. Mean Difference (IV, Random, 95% CI)	0.16 [-0.09, 0.40]
4.3 Longer-term	5	665	Std. Mean Difference (IV, Random, 95% CI)	0.04 [-0.11, 0.20]

Analysis 4.1. Comparison 4: Objective working memory, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chiaravalloti 2019b	13.8	4.46	15	15.77	4.95	13	6.8%	-0.41 [-1.16 , 0.34]
Chmelařová 2020	107.8	16.9	26	97.7	17.5	17	8.2%	0.58 [-0.05 , 1.20]
Impellizzeri 2020	25.33	4.6	15	18.97	6.71	15	6.6%	1.08 [0.30 , 1.85]
Maggio 2020	24.3	16.3	30	16	12.84	30	9.4%	0.56 [0.04 , 1.07]
Mousavi 2018a	24.55	2.98	20	20.85	3.79	20	7.7%	1.06 [0.40 , 1.73]
Naeeni Davarani 2020	45.58	8.1	28	36.27	9.25	26	8.8%	1.06 [0.49 , 1.63]
Pedulla 2016	33.07	11.12	14	24.12	8.85	14	6.6%	0.86 [0.08 , 1.64]
Perez-Martin 2017	29.7	15.48	30	30.44	16.08	32	9.7%	-0.05 [-0.54 , 0.45]
Rahmani 2020	23.17	3.51	12	19.78	3.58	12	6.0%	0.92 [0.07 , 1.77]
Rilo 2018	6.43	1.75	21	6.24	1.73	21	8.4%	0.11 [-0.50 , 0.71]
Stuifbergen 2012	34	9.3	34	34.9	11.5	27	9.6%	-0.09 [-0.59 , 0.42]
Stuifbergen 2018	33	12	93	30.7	12.1	90	12.3%	0.19 [-0.10 , 0.48]
Total (95% CI)			**338**			**317**	**100.0%**	**0.45 [0.18 , 0.72]**

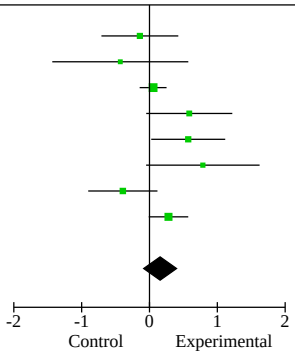
Heterogeneity: Tau² = 0.13; Chi² = 29.14, df = 11 (P = 0.002); I² = 62%
Test for overall effect: Z = 3.24 (P = 0.001)
Test for subgroup differences: Not applicable



Analysis 4.2. Comparison 4: Objective working memory, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
das Nair 2012	23.3	10	24	24.6	7.9	24	11.1%	-0.14 [-0.71 , 0.42]
Huiskamp 2016	0.87	0.14	9	0.93	0.12	7	4.8%	-0.43 [-1.43 , 0.57]
Lincoln 2020	36.6	16.2	217	35.7	17.6	178	23.8%	0.05 [-0.14 , 0.25]
Mousavi 2018a	22.85	3.08	20	20.9	3.43	20	9.6%	0.59 [-0.05 , 1.22]
Naeeni Davarani 2020	41.85	8.86	28	36.54	9.48	26	11.6%	0.57 [0.03 , 1.12]
Rahmani 2020	22.75	3.72	12	19.59	4.03	12	6.5%	0.79 [-0.05 , 1.62]
Stuifbergen 2012	34.2	9.8	34	38.1	9.8	27	12.5%	-0.39 [-0.90 , 0.12]
Stuifbergen 2018	35.1	12.4	93	31.6	12.5	90	20.1%	0.28 [-0.01 , 0.57]
Total (95% CI)			**437**			**384**	**100.0%**	**0.16 [-0.09 , 0.40]**

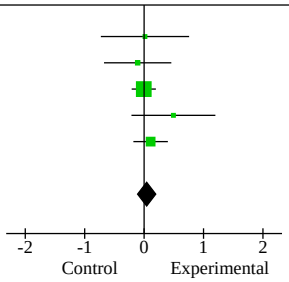
Heterogeneity: Tau² = 0.05; Chi² = 14.58, df = 7 (P = 0.04); I² = 52%
Test for overall effect: Z = 1.27 (P = 0.20)
Test for subgroup differences: Not applicable



Analysis 4.3. Comparison 4: Objective working memory, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chiaravalloti 2019b	16.25	4.94	15	16.17	5.67	13	4.2%	0.01 [-0.73 , 0.76]
das Nair 2012	25.2	8.8	24	26.1	7.5	24	7.3%	-0.11 [-0.67 , 0.46]
Lincoln 2020	36.4	17.8	205	36.5	17.7	169	56.2%	-0.01 [-0.21 , 0.20]
Mattioli 2016	43.25	8.9	15	37.18	14.16	17	4.7%	0.49 [-0.21 , 1.20]
Stuifbergen 2018	34.8	13.2	93	33.4	12.2	90	27.7%	0.11 [-0.18 , 0.40]
Total (95% CI)			**352**			**313**	**100.0%**	**0.04 [-0.11 , 0.20]**

Heterogeneity: Tau² = 0.00; Chi² = 2.26, df = 4 (P = 0.69); I² = 0%
Test for overall effect: Z = 0.55 (P = 0.58)
Test for subgroup differences: Not applicable



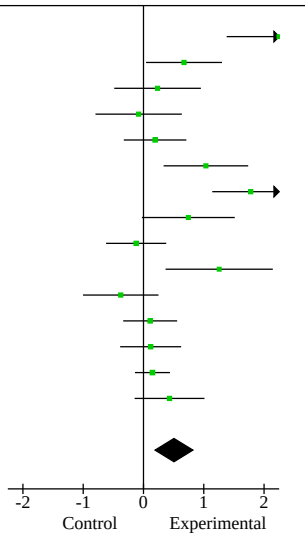
Comparison 5. Information processing

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
5.1 Immediate	15	808	Std. Mean Difference (IV, Random, 95% CI)	0.51 [0.19, 0.82]
5.2 Intermediate	8	933	Std. Mean Difference (IV, Random, 95% CI)	0.27 [0.00, 0.54]
5.3 Longer-term	5	723	Std. Mean Difference (IV, Random, 95% CI)	0.21 [-0.03, 0.45]

Analysis 5.1. Comparison 5: Information processing, Outcome 1: Immediate

	Experimental			Control				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Campbell 2016	47.33	5.05	17	37.58	3.3	18	5.4%	2.25 [1.38 , 3.11]	
Chmelařová 2020	-40	14.8	26	-58.2	38.3	17	6.7%	0.67 [0.04 , 1.30]	
Hancock 2015	53.13	10.79	15	49.4	19.16	15	6.2%	0.23 [-0.49 , 0.95]	
Impellizzeri 2020	26.24	10.83	15	26.93	4.69	15	6.2%	-0.08 [-0.80 , 0.64]	
Messinis 2017	40.03	7.08	32	37.43	18.38	26	7.3%	0.19 [-0.33 , 0.71]	
Messinis 2020	40.42	7.3	19	31.52	9.5	17	6.2%	1.04 [0.33 , 1.74]	
Naeeni Davarani 2020	40.19	3.52	28	32.42	5.02	26	6.6%	1.78 [1.14 , 2.42]	
Pedulla 2016	46.03	11.52	14	38.08	9.09	14	5.9%	0.74 [-0.03 , 1.51]	
Perez-Martin 2017	46.47	13.3	30	47.93	10.34	32	7.4%	-0.12 [-0.62 , 0.38]	
Rahmani 2020	-10.17	1.94	12	-13.08	2.5	12	5.2%	1.26 [0.37 , 2.15]	
Rilo 2018	42.62	12.46	20	47.52	13	20	6.7%	-0.38 [-1.00 , 0.25]	
Solari 2004	29.95	8.2	40	29.1	6.9	37	7.7%	0.11 [-0.34 , 0.56]	
Stuifbergen 2012	49.6	11.1	34	48.1	14	27	7.3%	0.12 [-0.39 , 0.62]	
Stuifbergen 2018	52.4	12.6	93	50.6	11.5	90	8.4%	0.15 [-0.14 , 0.44]	
Vilou 2020	50	12	23	44.5	13	24	6.9%	0.43 [-0.15 , 1.01]	
Total (95% CI)			**418**			**390**	**100.0%**	**0.51 [0.19 , 0.82]**	

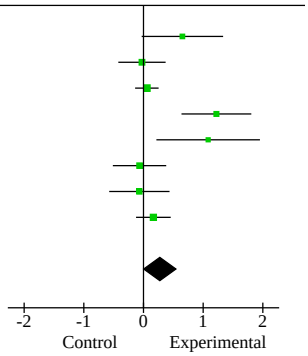
Heterogeneity: Tau² = 0.28; Chi² = 61.62, df = 14 (P < 0.00001); I² = 77%
Test for overall effect: Z = 3.16 (P = 0.002)
Test for subgroup differences: Not applicable



Analysis 5.2. Comparison 5: Information processing, Outcome 2: Intermediate

	Experimental			Control				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Campbell 2016	46.74	4.17	17	42.78	7.21	18	8.9%	0.65 [-0.03 , 1.33]	
Hanssen 2015	56.4	11.7	47	56.7	11.7	51	14.2%	-0.03 [-0.42 , 0.37]	
Lincoln 2020	41.4	12.1	220	40.7	12.7	181	18.5%	0.06 [-0.14 , 0.25]	
Naeeni Davarani 2020	38.5	3.79	28	32.96	5.1	26	10.4%	1.22 [0.64 , 1.81]	
Rahmani 2020	-10.83	1.69	12	-13.16	2.4	12	6.6%	1.08 [0.22 , 1.95]	
Solari 2004	30.59	8.2	40	31.1	6.9	37	13.1%	-0.07 [-0.51 , 0.38]	
Stuifbergen 2012	49.7	12.7	34	50.6	13.1	27	11.9%	-0.07 [-0.57 , 0.44]	
Stuifbergen 2018	52.8	13	93	50.7	12.2	90	16.5%	0.17 [-0.12 , 0.46]	
Total (95% CI)			**491**			**442**	**100.0%**	**0.27 [0.00 , 0.54]**	

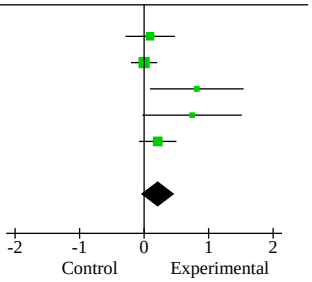
Heterogeneity: Tau² = 0.09; Chi² = 22.68, df = 7 (P = 0.002); I² = 69%
Test for overall effect: Z = 1.99 (P = 0.05)
Test for subgroup differences: Not applicable



Analysis 5.3. Comparison 5: Information processing, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Hanssen 2015	56.3	11.8	54	55.2	11.5	51	21.2%	0.09 [-0.29 , 0.48]
Lincoln 2020	39.9	11.9	205	39.9	12.8	170	34.5%	0.00 [-0.20 , 0.20]
Mattioli 2016	45	7.47	15	35.71	13.49	17	8.9%	0.82 [0.09 , 1.54]
Pedulla 2016	46.03	11.52	14	38.08	9.09	14	8.1%	0.74 [-0.03 , 1.51]
Stuifbergen 2018	54.6	12.2	93	52	12.4	90	27.4%	0.21 [-0.08 , 0.50]
Total (95% CI)			**381**			**342**	**100.0%**	**0.21 [-0.03 , 0.45]**

Heterogeneity: Tau² = 0.03; Chi² = 7.77, df = 4 (P = 0.10); I² = 49%
Test for overall effect: Z = 1.71 (P = 0.09)
Test for subgroup differences: Not applicable



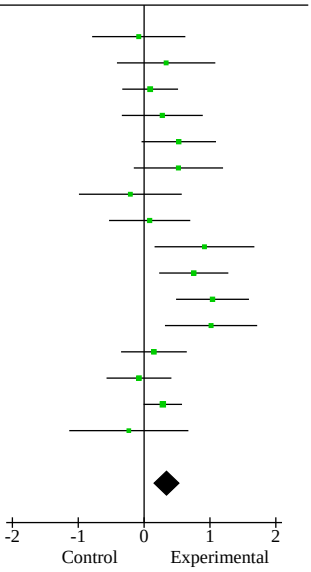
Comparison 6. Mood - Depression Scale

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
6.1 Immediate	16	853	Std. Mean Difference (IV, Random, 95% CI)	0.34 [0.15, 0.53]
6.2 Intermediate	10	1003	Std. Mean Difference (IV, Random, 95% CI)	0.20 [-0.06, 0.45]
6.3 Longer-term	7	891	Std. Mean Difference (IV, Random, 95% CI)	0.15 [-0.04, 0.34]

Analysis 6.1. Comparison 6: Mood - Depression Scale, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	-9.82	3.34	17	-9.5	4.35	14	4.9%	-0.08 [-0.79 , 0.63]
Chiaravalloti 2005	-6.21	6.2	14	-8.36	6.28	14	4.5%	0.33 [-0.41 , 1.08]
Chiaravalloti 2013	-55.05	15.7	45	-56.39	12.92	41	8.9%	0.09 [-0.33 , 0.52]
Chmelařová 2020	-12.42	10.03	26	-15.18	9.44	17	5.9%	0.28 [-0.34 , 0.89]
Goodwin 2020	-31.9	12.6	24	-39.7	16.2	26	6.6%	0.53 [-0.04 , 1.09]
Goverover 2018a	-53.1	15.8	19	-61.9	17.3	16	5.2%	0.52 [-0.16 , 1.20]
Hancock 2015	-3.63	2.58	15	-3.09	2.39	11	4.3%	-0.21 [-0.99 , 0.57]
Hildebrandt 2007	-10.3	8.5	17	-11	7.9	25	5.9%	0.08 [-0.53 , 0.70]
Impellizzeri 2020	-6.47	7.83	15	-13.87	7.88	15	4.4%	0.92 [0.16 , 1.67]
Maggio 2020	-7	7.83	30	-13	7.88	30	7.2%	0.75 [0.23 , 1.28]
Messinis 2017	-3.68	2.1	32	-6.7	3.6	26	6.7%	1.04 [0.49 , 1.59]
Messinis 2020	-3.68	2.1	19	-6.7	3.6	17	5.0%	1.02 [0.32 , 1.72]
Perez-Martin 2017	-5.57	3.93	30	-6.13	3.49	32	7.6%	0.15 [-0.35 , 0.65]
Solari 2004	-28.5	13.1	35	-27.6	8.9	29	7.7%	-0.08 [-0.57 , 0.41]
Stuifbergen 2018	-9.7	5.6	93	-11.4	6.3	90	11.8%	0.28 [-0.01 , 0.58]
Tesar 2005	-8.6	4.1	10	-7.7	3.2	9	3.4%	-0.23 [-1.14 , 0.67]
Total (95% CI)			**441**			**412**	**100.0%**	**0.34 [0.15 , 0.53]**

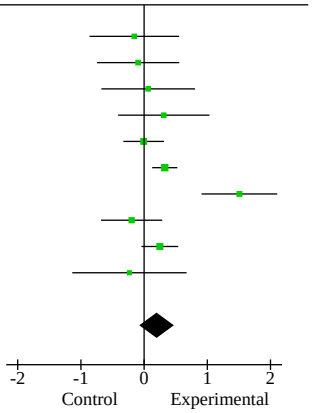
Heterogeneity: Tau² = 0.06; Chi² = 25.28, df = 15 (P = 0.05); I² = 41%
Test for overall effect: Z = 3.56 (P = 0.0004)
Test for subgroup differences: Not applicable



Analysis 6.2. Comparison 6: Mood - Depression Scale, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	-9.35	2.85	17	-8.79	4.21	14	7.5%	-0.15 [-0.86 , 0.55]
Carr 2014	-23.7	10.9	16	-22.7	9.9	21	8.3%	-0.09 [-0.75 , 0.56]
Chiaravalloti 2005	-6.79	8.15	14	-7.29	6.8	14	7.1%	0.06 [-0.68 , 0.81]
das Nair 2012	-14.28	2.79	29	-15.7	7.6	10	7.3%	0.31 [-0.41 , 1.03]
Lincoln 2002	-25.34	13.27	73	-25.24	14.6	76	13.9%	-0.01 [-0.33 , 0.31]
Lincoln 2020	-32.9	15.1	212	-37.8	14.8	183	16.0%	0.33 [0.13 , 0.53]
Shahpouri 2019	-11	6.86	28	-20.64	5.69	28	9.0%	1.51 [0.91 , 2.11]
Solari 2004	-28.03	12.87	34	-25.84	8.45	32	10.9%	-0.20 [-0.68 , 0.29]
Stuifbergen 2018	-9.9	6.2	93	-11.5	6.6	90	14.5%	0.25 [-0.04 , 0.54]
Tesar 2005	-8.6	4.1	10	-7.7	3.2	9	5.5%	-0.23 [-1.14 , 0.67]
Total (95% CI)			**526**			**477**	**100.0%**	**0.20 [-0.06 , 0.45]**

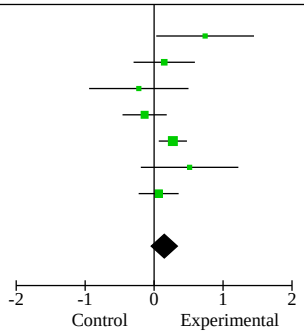
Heterogeneity: Tau² = 0.10; Chi² = 26.87, df = 9 (P = 0.001); I² = 66%
Test for overall effect: Z = 1.51 (P = 0.13)
Test for subgroup differences: Not applicable



Analysis 6.3. Comparison 6: Mood - Depression Scale, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carr 2014	-18.4	7	17	-25.3	10.9	16	6.1%	0.74 [0.03 , 1.45]
Chiaravalloti 2013	-54.44	15.62	40	-56.48	11.46	38	12.7%	0.15 [-0.30 , 0.59]
das Nair 2012	-15.93	8.61	29	-14.1	6.14	10	5.9%	-0.22 [-0.94 , 0.50]
Lincoln 2002	-27	15.7	73	-24.9	14.7	77	19.2%	-0.14 [-0.46 , 0.18]
Lincoln 2020	-33.9	16.1	209	-38.3	16.2	167	28.5%	0.27 [0.07 , 0.48]
Mattioli 2016	-6.75	5.97	15	-12.18	12.93	17	6.1%	0.51 [-0.19 , 1.22]
Stuifbergen 2018	-10.1	6.2	93	-10.5	5.9	90	21.3%	0.07 [-0.22 , 0.36]
Total (95% CI)			**476**			**415**	**100.0%**	**0.15 [-0.04 , 0.34]**

Heterogeneity: Tau² = 0.02; Chi² = 9.48, df = 6 (P = 0.15); I² = 37%
Test for overall effect: Z = 1.53 (P = 0.13)
Test for subgroup differences: Not applicable



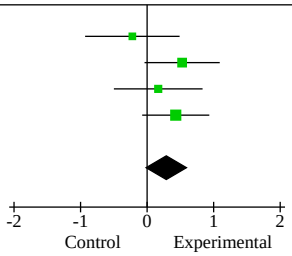
Comparison 7. Mood - Anxiety Scale

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
7.1 Immediate	4	178	Std. Mean Difference (IV, Random, 95% CI)	0.29 [-0.01, 0.59]
7.2 Intermediate	4	502	Std. Mean Difference (IV, Random, 95% CI)	0.16 [-0.15, 0.46]
7.3 Longer-term	3	448	Std. Mean Difference (IV, Random, 95% CI)	0.27 [-0.12, 0.65]

Analysis 7.1. Comparison 7: Mood - Anxiety Scale, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	-9.29	4.47	17	-8.21	5.06	14	17.8%	-0.22 [-0.93 , 0.49]
Goodwin 2020	-31.9	12.6	24	-39.7	16.2	26	27.7%	0.53 [-0.04 , 1.09]
Goverover 2018a	-39.4	12.6	19	-41.4	10.5	16	20.1%	0.17 [-0.50 , 0.83]
Perez-Martin 2017	-5.97	3.15	30	-7.41	3.44	32	34.4%	0.43 [-0.07 , 0.93]
Total (95% CI)			**90**			**88**	**100.0%**	**0.29 [-0.01 , 0.59]**

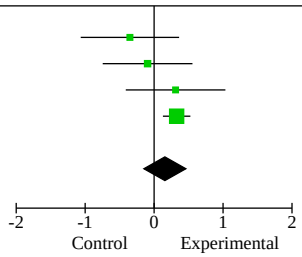
Heterogeneity: Tau² = 0.00; Chi² = 3.10, df = 3 (P = 0.38); I² = 3%
Test for overall effect: Z = 1.87 (P = 0.06)
Test for subgroup differences: Not applicable



Analysis 7.2. Comparison 7: Mood - Anxiety Scale, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	-8.53	4.38	17	-6.86	4.93	14	14.5%	-0.35 [-1.06 , 0.36]
Carr 2014	-23.7	10.9	16	-22.7	9.9	21	16.7%	-0.09 [-0.75 , 0.56]
das Nair 2012	-14.28	2.79	29	-15.7	7.6	10	14.2%	0.31 [-0.41 , 1.03]
Lincoln 2020	-32.9	15.1	212	-37.8	14.8	183	54.6%	0.33 [0.13 , 0.53]
Total (95% CI)			**274**			**228**	**100.0%**	**0.16 [-0.15 , 0.46]**

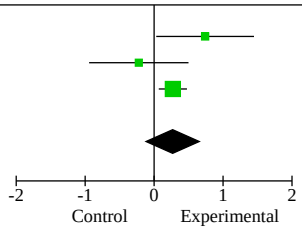
Heterogeneity: Tau² = 0.03; Chi² = 4.40, df = 3 (P = 0.22); I² = 32%
Test for overall effect: Z = 1.01 (P = 0.31)
Test for subgroup differences: Not applicable



Analysis 7.3. Comparison 7: Mood - Anxiety Scale, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carr 2014	-18.4	7	17	-25.3	10.9	16	20.8%	0.74 [0.03 , 1.45]
das Nair 2012	-15.93	8.61	29	-14.1	6.14	10	20.4%	-0.22 [-0.94 , 0.50]
Lincoln 2020	-33.9	16.1	209	-38.3	16.2	167	58.8%	0.27 [0.07 , 0.48]
Total (95% CI)			**255**			**193**	**100.0%**	**0.27 [-0.12 , 0.65]**

Heterogeneity: Tau² = 0.05; Chi² = 3.48, df = 2 (P = 0.18); I² = 43%
Test for overall effect: Z = 1.37 (P = 0.17)
Test for subgroup differences: Not applicable



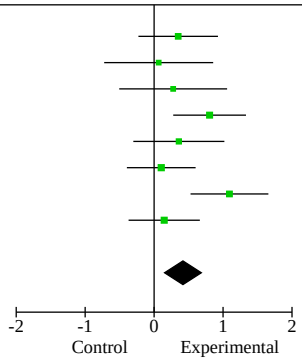
Comparison 8. Quality of life

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
8.1 Immediate	8	371	Std. Mean Difference (IV, Random, 95% CI)	0.42 [0.15, 0.68]
8.2 Intermediate	6	683	Std. Mean Difference (IV, Random, 95% CI)	0.30 [0.02, 0.58]
8.3 Longer-term	5	687	Std. Mean Difference (IV, Random, 95% CI)	0.17 [0.02, 0.32]

Analysis 8.1. Comparison 8: Quality of life, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Goodwin 2020	53.2	26.4	23	44.5	22.5	24	13.1%	0.35 [-0.23 , 0.93]
Goverover 2018a	19.19	18.3	10	18.3	8.3	16	8.5%	0.07 [-0.72 , 0.86]
Hancock 2015	-69.69	12.37	15	-73.65	15.73	11	8.7%	0.28 [-0.51 , 1.06]
Maggio 2020	74.6	28.55	30	58	3.7	30	14.6%	0.80 [0.28 , 1.33]
Messinis 2020	60.21	16.9	19	54.11	16.3	17	11.0%	0.36 [-0.30 , 1.02]
Perez-Martin 2017	26.87	11.83	30	25.63	11.83	32	15.6%	0.10 [-0.39 , 0.60]
Shahpouri 2019	67.77	15.12	28	50.9	15.32	28	13.5%	1.09 [0.53 , 1.66]
Solari 2004	-46.99	17.38	31	-49.26	12.44	27	15.0%	0.15 [-0.37 , 0.66]
Total (95% CI)			**186**			**185**	**100.0%**	**0.42 [0.15 , 0.68]**

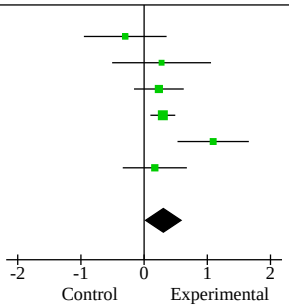
Heterogeneity: Tau² = 0.05; Chi² = 11.12, df = 7 (P = 0.13); I² = 37%
Test for overall effect: Z = 3.08 (P = 0.002)
Test for subgroup differences: Not applicable



Analysis 8.2. Comparison 8: Quality of life, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carr 2014	-77.2	30.7	16	-69	23.6	21	11.8%	-0.30 [-0.95 , 0.36]
Hanssen 2015	-69.69	12.37	15	-73.65	15.73	11	9.3%	0.28 [-0.51 , 1.06]
Lincoln 2002	-18.3	6.7	50	-20	7.7	50	20.1%	0.23 [-0.16 , 0.63]
Lincoln 2020	-22.3	6.2	217	-24.1	5.9	187	28.6%	0.30 [0.10 , 0.49]
Shahpouri 2019	67.77	15.12	28	50.9	15.32	28	14.2%	1.09 [0.53 , 1.66]
Solari 2004	-48.57	17.22	29	-51.18	13.06	31	15.9%	0.17 [-0.34 , 0.68]
Total (95% CI)			**355**			**328**	**100.0%**	**0.30 [0.02 , 0.58]**

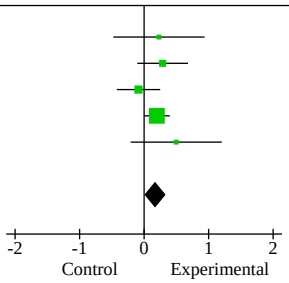
Heterogeneity: Tau² = 0.06; Chi² = 11.17, df = 5 (P = 0.05); I² = 55%
Test for overall effect: Z = 2.13 (P = 0.03)
Test for subgroup differences: Not applicable



Analysis 8.3. Comparison 8: Quality of life, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Carr 2014	-68.3	28	15	-74.6	25.4	16	4.5%	0.23 [-0.48 , 0.94]
Hanssen 2015	-18.3	7.2	53	-20.5	8	48	14.7%	0.29 [-0.10 , 0.68]
Lincoln 2002	45.42	11.94	66	46.52	13.19	70	20.0%	-0.09 [-0.42 , 0.25]
Lincoln 2020	-22.2	6.1	214	-23.4	6	173	56.2%	0.20 [-0.00 , 0.40]
Mattioli 2016	155.94	46.17	15	132.65	45.17	17	4.5%	0.50 [-0.21 , 1.20]
Total (95% CI)			**363**			**324**	**100.0%**	**0.17 [0.02 , 0.32]**

Heterogeneity: Tau² = 0.00; Chi² = 3.51, df = 4 (P = 0.48); I² = 0%
Test for overall effect: Z = 2.20 (P = 0.03)
Test for subgroup differences: Not applicable



Comparison 9. Activities of Daily Living

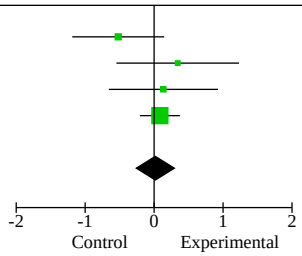
Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
9.1 Immediate	4	265	Std. Mean Difference (IV, Random, 95% CI)	0.02 [-0.26, 0.29]

Outcome or sub-group title	No. of studies	No. of partici-pants	Statistical method	Effect size
9.2 Intermediate	4	400	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.36, 0.24]
9.3 Longer-term	3	369	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.49, 0.27]

Analysis 9.1. Comparison 9: Activities of Daily Living, Outcome 1: Immediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	82.18	25.09	17	99.14	36.89	19	15.8%	-0.52 [-1.19 , 0.15]
Chiaravalloti 2019a	17.14	5.27	9	15.12	5.94	11	9.2%	0.34 [-0.55 , 1.23]
Goverover 2018a	102.5	27.5	10	98.3	31.5	16	11.5%	0.14 [-0.66 , 0.93]
Stuifbergen 2018	23.4	5	93	23	4.5	90	63.5%	0.08 [-0.21 , 0.37]
Total (95% CI)			**129**			**136**	**100.0%**	**0.02 [-0.26 , 0.29]**

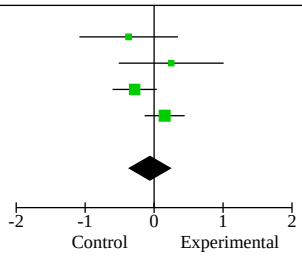
Heterogeneity: Tau² = 0.01; Chi² = 3.29, df = 3 (P = 0.35); I² = 9%
Test for overall effect: Z = 0.13 (P = 0.90)
Test for subgroup differences: Not applicable



Analysis 9.2. Comparison 9: Activities of Daily Living, Outcome 2: Intermediate

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Campbell 2016	89	30.99	17	101	32.4	14	13.5%	-0.37 [-1.08 , 0.34]
das Nair 2012	48.54	10.87	26	45.56	14.14	9	12.2%	0.25 [-0.51 , 1.01]
Lincoln 2002	40.87	18.39	74	45.82	16.49	77	35.7%	-0.28 [-0.60 , 0.04]
Stuifbergen 2018	23.8	4.8	93	23.1	4.3	90	38.6%	0.15 [-0.14 , 0.44]
Total (95% CI)			**210**			**190**	**100.0%**	**-0.06 [-0.36 , 0.24]**

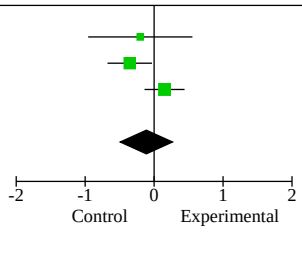
Heterogeneity: Tau² = 0.04; Chi² = 5.25, df = 3 (P = 0.15); I² = 43%
Test for overall effect: Z = 0.41 (P = 0.69)
Test for subgroup differences: Not applicable



Analysis 9.3. Comparison 9: Activities of Daily Living, Outcome 3: Longer-term

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
das Nair 2012	48.52	11.28	27	50.89	12.41	9	17.4%	-0.20 [-0.96 , 0.56]
Lincoln 2002	39.96	18.18	73	46.2	16.93	77	40.1%	-0.35 [-0.68 , -0.03]
Stuifbergen 2018	24.2	4.8	93	23.5	4.4	90	42.5%	0.15 [-0.14 , 0.44]
Total (95% CI)			**193**			**176**	**100.0%**	**-0.11 [-0.49 , 0.27]**

Heterogeneity: Tau² = 0.07; Chi² = 5.30, df = 2 (P = 0.07); I² = 62%
Test for overall effect: Z = 0.58 (P = 0.56)
Test for subgroup differences: Not applicable



ADDITIONAL TABLES

Table 1. Summary of findings continued

	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of evidence (GRADE)	Comments
	Assumed Risk	Corresponding Risk				
Outcomes	Control	Memory rehabiliation				
Subjective memory - immediate EMQ, MSNQ, CFQ, MFQ[a] Follow-up: within one month	-	The mean subjective memory measures - immediate in the intervention groups was **0.32 standard deviations higher** (0.05 to 0.58 higher)	-	568 (10 studies)	⊕⊕⊕⊖ **moderate** [b]	SMD 0.32 (0.05 to 0.58)
Subjective memory - longer-term EMQ, MSNQ, CFQ, MFQ[a] Follow-up: 6 months +	-	The mean subjective memory measures - long term in the intervention groups was **0.16 standard deviations higher** (0.02 to 0.30 higher)	-	775 (5 studies)	⊕⊕⊕⊕ **high**	SMD 0.16 (0.02 to 0.30)
Verbal memory - immediate CVLT, AVLT, HVLT, VLT, SRT, MUSIC[a] Follow-up: within one month	-	The mean objective verbal memory measures - immediate in the intervention groups was **0.4 standard deviations higher** (0.22 to 0.58 higher)	-	922 (19 studies)	⊕⊕⊖⊖ **low** [c,d]	SMD 0.40 (0.22 to 0.58)
Verbal memory - longer-term CVLT, AVLT, HVLT, VLT, SRT, MUSIC[a] Follow-up: 6 months +	-	The mean objective verbal memory measures - longer-term in the intervention groups was **0.13 standard deviations higher** (0.03 lower to 0.29 higher)	-	619 (4 studies)	⊕⊕⊕⊖ **moderate** [d]	SMD 0.13 (-0.03 to 0.29)
Visual memory - immediate BVMT-R, SPART, CMT, ROCF Follow-up: within one month	-	The mean objective visual memory measures - immediate in the intervention groups was **0.42 standard deviations higher** (0.25 to 0.6 higher)	-	799 (16 studies)	⊕⊕⊕⊖ **moderate** [f]	SMD 0.42 (0.25 to 0.60)
Visual memory - longer-term BVMT-R, SPART, CMT, ROCF Follow-up: 6 months +	-	The mean objective visual memory measures - longer-term in the intervention groups was **0.12 standard deviations higher** (0.13 lower to 0.37 higher)	-	619 (4 studies)	⊕⊕⊕⊕ **high**	SMD 0.12 (-0.13 to 0.37)
Working memory - immediate	-	The mean objective working memory measures -	-	655 (12 studies)	⊕⊕⊖⊖ **low** [h,p]	SMD 0.45

Table 1. Summary of findings continued *(Continued)*

PASAT, WAIS Follow-up: within one month		immediate in the intervention groups was **0.45 standard deviations higher** (0.18 to 0.72 higher)				(0.18 to 0.72)
Working memory - longer-term PASAT, WAIS Follow-up: 6 months +	-	The mean objective working memory measures - longer-term in the intervention groups was **0.04 standard deviations higher** (0.11 lower to 0.2 higher)	-	665 (5 studies)	⊕⊕⊕⊕ **high**	SMD 0.04 (-0.11 to 0.2)
Information processing - immediate SDMT Follow-up: within one month	-	The mean information processing measures - immediate in the intervention groups was **0.51 standard deviations higher** (0.19 to 0.82 higher)	-	808 (15 studies)	⊕⊕⊖⊖ **low** [j,p]	SMD 0.51 (0.19 to 0.82)
Information processing - longer-term SDMT Follow-up: 6 months +	-	The mean information processing measures - longer-term in the intervention groups was **0.21 standard deviations higher** (0.03 lower to 0.45 higher)	-	723 (5 studies)	⊕⊕⊕⊖ **moderate** [l]	SMD 0.21 (-0.03 to 0.45)
Depression (mood) - immediate GHQ, BDI, BDI-FS, Chicago Multiscale Depression Inventory, HADS, EAQ, CES-D, MADRS[a] Follow-up: within one month	-	The mean depression measures (mood) - immediate in the intervention groups was **0.34 standard deviations higher** (0.15 to 0.53 higher)	-	853 (16 studies)	⊕⊕⊕⊖ **moderate** [m]	SMD 0.34 (0.15 to 0.53)
Depression (mood) - intermediate GHQ, BDI, BDI-FS, Chicago Multiscale Depression Inventory, HADS, EAQ, CES-D, MADRS[a] Follow-up: 1 to 6 months	-	The mean depression measures (mood) - intermediate in the intervention groups was **0.20 standard deviations higher** (0.06 lower to 0.45 higher)	-	1003 (10 studies)	⊕⊕⊕⊖ **moderate** [m]	SMD 0.20 (-0.06 to 0.45)
Depression (mood) - longer-term GHQ, BDI, BDI-FS, Chicago Multiscale Depression Inventory, HADS, EAQ, CES-D, MADRS[a]	-	The mean depression measures (mood) - longer-term in the intervention groups was **0.15 standard deviations higher** (0.04 lower to 0.34 higher)	-	891 (7 studies)	⊕⊕⊕⊕ **high**	SMD 0.15 (-0.04 to 0.34)

Table 1. Summary of findings continued *(Continued)*

Follow-up: 1 to 6 months						
Anxiety (mood) - immediate GHQ, EAQ, STAI, HADS Follow-up: within one month	-	The mean anxiety measures (mood) - immediate in the intervention groups was **0.29 standard deviations higher** (0.01 lower to 0.59 higher)	-	178 (4 studies)	⊕⊕⊕⊕ **high**	SMD 0.29 (-0.01 to 0.59)
Anxiety (mood) - intermediate GHQ, EAQ, STAI, HADS Follow-up: 1 to 6 months	-	The mean anxiety measures (mood) - intermediate in the intervention groups was **0.16 standard deviations higher** (0.15 lower to 0.46 higher)	-	502 (4 studies)	⊕⊕⊕⊕ **high**	SMD 0.16 (-0.15 to 0.46)
Anxiety (mood) - longer-term GHQ, EAQ, STAI, HADS Follow-up: 6 months +	-	The mean anxiety measures (mood) - longer-term in the intervention groups was **0.27 standard deviations higher** (0.12 lower to 0.65 higher)	-	502 (4 studies)	⊕⊕⊕⊕ **high**	SMD 0.27 (-0.12 to 0.65)
Quality of life - immediate MSIS, MSQOL, SF-36, SF-12, SWLS, EQ-5D-5L[a] Follow-up: within one month	-	The mean quality of life measures - immediate in the intervention groups was **0.42 standard deviations higher** (0.15 to 0.68 higher)	-	371 (8 studies)	⊕⊕⊕⊕ **high**	SMD 0.42 (0.15 to 0.68)
Quality of life - longer-term MSIS, MSQOL, SF-36, SF-12, SWLS, EQ-5D-5L[a] Follow-up: 6 months +	-	The mean quality of life measures - longer-term in the intervention groups was **0.17 standard deviations higher** (0.02 to 0.32 higher)	-	687 (5 studies)	⊕⊕⊕⊖ **moderate[o]**	SMD 0.17 (0.02 to 0.32)
Activities of daily living - immediate EADL[a] Follow-up: within one month	-	The mean activities of daily living measures - immediate in the intervention groups was **0.02 standard deviations higher** (0.26 lower to 0.29 higher)	-	265 (4 studies)	⊕⊕⊕⊕ **high**	SMD 0.02 (-0.26 to 0.29)
Activities of daily living - longer-term EADL[a] Follow-up: 6 months +	-	The mean activities of daily living measures - longer-term in the intervention groups was **0.11 standard deviations lower** (0.49 lower to 0.27 higher)	-	369 (3 studies)	⊕⊕⊕⊕ **high**	SMD -0.11 (-0.49 to 0.27)

Table 1. Summary of findings continued *(Continued)*

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; SMD: standardised mean difference

GRADE Working Group grades of evidence
High quality ⊕⊕⊕⊕**:** further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality ⊕⊕⊕⊖**:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality ⊕⊕⊖⊖**:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality ⊕⊖⊖⊖**:** we are very uncertain about the estimate.

[a] **CMT:** Contextual Memory Text; **EAQ:** Emotional awareness questionnaire;**EMQ:** Everyday Memory Questionnaire;**HADS:** Hospital Anxiety and Depression Scale; **STAI:** State Trait Anxiety Inventory; **MSNQ:** Multiple Sclerosis Neuropsychological Screening Questionnaire;**MFQ:** Memory Functioning Questionnaire;**RBMT:** Rivermead Behavioural Memory Test;**CVLT:** California Verbal Learning Test; **AVLT:** Auditory Verbal Learning Test; **HVLT:** Hopkins Verbal Learning Test; **VLT:** Verbal Learning Test; LNNB: Luria-Nebraska Neuropsychological Battery; **BRBNT:** Brief Repeatable Battery of Neuropsychological Tests; **GHQ:** General Health Questionnaire; **BDI:** Beck Depression Inventory; **BDI-FS:** Beck Depression Inventory-Fast Screen; **EADL:** Extended Activities of Daily Living; **MSIS:** Multiple Sclerosis Impact Scale; **FAMS:** Functional Assessment of Multiple Sclerosis;**MSQOL:** Multiple Sclerosis Quality of Life; **PASAT:** Paced auditory serial addition test; **SF-36:** 36-Item Short Form Health Survey; **SF-12:** 12-Item Short Form Health Survey.

[b] 1 of 10 studies had possible risk of bias related to random sequence generation, and in 2 of the 10 it was unclear. Allocation concealment was possible in 1 study, and unclear in 3 of the 10 studies. Blinding was a potential source of bias in 2 studies, and unclear in 1 of the 10 studies. Incomplete outcome data may have been biased in 1 study, an unclear in 3 of the 10 studies. Selective reporting may have been biased in 1 study.
Downgraded by 1 due to 95% confidence intervals including no effect, and the upper or lower confidence intervals limit crosses an effect size of 0.5 in either direction.

[c] 1 study had possible risk of bias related to random sequence generation, and in 5 of 19 studies this was unclear. Allocation concealment was potentially biased in 1 study, and unclear in 6 of 19 studies. Blinding was a potential source of bias in 7 studies. Incomplete outcome data may have biased 2 of 19 studies and was unclear in 6 of 19 studies. Selective reporting may have been bias in 1 study. May have been other sources in of bias in 1 study and unclear in 1 study.

[d] All or nearly all of the studies used a list-learning task as an objective measure of verbal memory, which has poor ecological validity.

[f] 5 of 16 studies showed unclear potential risk of bias related to random sequence generation. 6 of 16 studies showed unclear potential risk of bias related to allocation concealment. 7 of 16 studies showed potential risk of bias related to blinding. 1 study showed potential risk of bias related to incomplete outcome data, 4 of the 16 studies showed unclear risk of bias. May have other source of bias in 1 study.

[h] 5 of 12 studies showed unclear potential risk of bias related to random sequence generation. 6 of 12 studies showed unclear risk of bias related to allocation concealment. 7 of 12 studies showed possible risk of bias related to blinding procedures. 1 study showed potential risk of bias related to incomplete data, and 3 of 12 studies were unclear risk of bias. 1 study had potential risk of bias related to selective reporting.

[j] 5 of 15 studies showed unclear risk of bias related to random sequence generation. 6 of 15 studies showed unclear risk of bias related to allocation concealment. 8 of 15 studies showed potential risk of bias related to blinding procedures. 2 of 15 studIes showed potential risk of bias related to blinding procedures, and 3 of 15 were unclear risk of bias. 1 study showed potential risk of bias related to incomplete data. 1 study showed potential risk of bias related to other bias.

[l] 2 of 5 studies showed unclear risk of bias related to random sequence generation. 1 study showed potential risk of bias related to allocation concealment, 1 study showed unclear risk of bias. 3 of 5 studies showed potential risk of bias related to blinding procedures, 1 study showed unclear risk of bias. 1 study showed potential risk of bias related to incomplete data, 1 study showed unclear risk of bias. 1 study showed potential risk of bias related to selective reporting.

[m] 2 of 16 studies showed potential risk of bias related to random sequence generation, 3 of 16 studies showed unclear risk. 1 study showed potential risk of bias relating to allocation concealment, 6 of 16 studies showed unclear risk of bias. 5 of 16 studies showed potential risk of bias relating to blinding procedures. 3 of 16 studies showed potential risk of bias relating to incomplete data, 3 of 13 studies showed unclear risk of bias. 1 study showed potential risk of bias relating to selective reporting. 1 study showed potential risk of bias relating to other bias.

[o] 1 study showed unclear risk of bias related to random sequence generation, blinding procedures and incomplete outcome data, as well as high risk of bias relating to allocation concealment. 1 study showed high risk of bias relating to blinding procedures, incomplete data and selective reporting.

[p] Inconsistency with results, statistical heterogeneity>50%

Table 2. Sensitivity analysis

Outcome	No. of studies	No. of participants	Effect size SMD (95% CI)	Heterogeneity (I^2)	Test for overall effect
Subjective memory - immediate	2	E = 127 C = 117	0.03 [-0.24, 0.31]	10%	Z = 0.22 (P = 0.82)
Subjective memory - intermediate	6	E = 396 C = 343	0.25 [0.11, 0.40]	0%	Z = 3.39 (P = 0.0007)
Subjective memory - longer-term	4	E = 325 C = 294	0.19 [0.03, 0.36]	0%	Z = 2.33 (P = 0.03)
Verbal memory - immediate	5	E = 100 C = 96	0.72 [0.24, 1.19]	59%	Z = 2.96 (P = 0.003)
Verbal memory - intermediate	2	E = 254 C = 209	0.22 [0.03, 0.40]	0%	Z = 2.32 (P = 0.02)
Verbal memory - longer-term	N/A				
Visual memory - immediate	5	E = 100 C = 94	0.27 [-0.01, 0.56]	0%	Z = 1.86 (P = 0.06)
Visual memory - intermediate	2	E = 251 C = 209	-0.11 [-0.29, 0.08]	0%	Z = 1.14 (P = 0.25)
Visual memory - longer-term	N/A				
Working memory - immediate	2	E = 49 C = 42	0.46 [-0.68, 1.59]	84%	Z = 0.79 (P = 0.43)
Working memory - intermediate	4	E = 284 C = 236	-0.06 [-0.28, 0.15]	11%	Z = 0.59 (P = 0.56)
Working memory - longer-term	2	E = 229 C = 193	-0.02 [-0.21, 0.17]	0%	Z = 0.18 (P = 0.86)
Information processing - immediate	4	E = 131 C = 120	0.29 [-0.04, 0.62]	40%	Z = 1.72 (P = 0.05)
Information processing - intermediate	4	E = 294 C = 245	0.02 [-0.14, 0.19]	0%	Z = 0.28 (P = 0.78)
Information processing - longer-term	N/A				
Depression - immediate	4	E = 93	0.55 [0.03, 1.07]	65%	Z = 2.07 (P = 0.04)

Table 2. Sensitivity analysis *(Continued)*

		C = 87			
Depression - intermediate	6	E = 392 C = 350	0.29 [-0.10, 0.67]	79%	Z = 1.45 (P = 0.15)
Depression - longer-term	4	E = 328 C = 270	0.14 [-0.20, 0.48]	63%	Z = 0.80 (P = 0.42)
Anxiety - immediate	N/A				
Anxiety - intermediate	3	E = 257 C = 214	0.29 [0.11, 0.48]	0%	Z = 3.11 (P = 0.002)
Anxiety - longer-term	3	E = 255 C = 193	0.27 [-0.12, 0.65]	43%	Z = 1.37 (P = 0.17)
Quality of life - immediate	4	E = 101 C = 96	0.49 [0.06, 0.91]	54%	Z = 2.25 (P = 0.02)
Quality of life - intermediate	5	E = 340 C = 317	0.31 [-0.01, 0.62]	64%	Z = 1.90 (P = 0.06)
Quality of life - longer-term	3	E = 295 C = 259	0.12 [-0.05, 0.30]	5%	Z = 1.37 (P = 0.17)
Activities of daily living - immediate	N/A				
Activities of daily living - intermediate	2	E = 100 C = 86	-0.13 [-0.60, 0.33]	37%	Z = 0.56 (P = 0.57)
Activities of daily living - longer-term	2	E = 100 C = 86	-0.33 [-0.63, -0.03]	0%	Z = 2.18 (P = 0.03)

E: Experimental; **C:** Control; **SMD:** Standardised mean difference.

APPENDICES

Appendix 1. Keywords

{attention*} OR {cognition} OR {cognition disorder*} OR {cognitive} OR {concentration} OR {distract*} OR {alert*} AND {training} OR {retraining} OR {therap*} OR {rehabilitation} OR {treatment*} OR {therapeutic*} OR {computer assisted therap*} OR {computer*} OR {neuropsychological test*} OR {neurorehabilitation} OR {neuropsychological rehabilitation} OR {rehabilitation} OR {cognition} OR {neurological system and disorders} OR {memory} OR {cognitive retraining}

WHAT'S NEW

Date	Event	Description
4 November 2020	New citation required and conclusions have changed	Updated search completed 6 September 2020. Twenty-nine studies have been added to the review, bringing the total to 44 studies and conclusions have changed.
4 November 2020	New search has been performed	New review authors have been added to the review team. In this version of the review, the quality of the evidence from the included studies was assessed using GRADE approach and a summary of findings table was added. There are nine different outcomes instead of the previous five and three measured timepoints instead of the previous two.

HISTORY

Protocol first published: Issue 10, 2010
Review first published: Issue 3, 2012

CONTRIBUTIONS OF AUTHORS

RdN and NBL conceptualised the protocol for the review. LT ran the searches and collected the studies. LT, NE, DW, JMM, LS reviewed the studies, which were verified by RdN. LT wrote the review with input from NE, DW, JMM, LS, KJP, NBL and RdN.

DECLARATIONS OF INTEREST

RdN, JMM, KJP and NBL have conducted memory rehabilitation studies in MS that have been included in this review.

RdN and NE have been funded by NIHR for a programme grant on cognitive screening and rehabilitation.

NE has received lecture fees from Biogen and participated in paid advisory board for Biogen, Roche and Merck where cognition was discussed.

LT, DW, and LS have nothing to declare.

RdN is the Chair of the NIHR Research for Patient Benefit East Midlands Research Advisory Committee. He has received funding to prepare and deliver lectures (speakers bureau) on cognitive rehabilitation in multiple sclerosis from Novartis, Merck, and Biogen.

SOURCES OF SUPPORT

Internal sources

- No sources of support provided

External sources

- MS Society , UK

  The MS Society has funded a PhD studentship for Lauren Taylor. This update forms one part of her PhD.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We did not perform subgroup analyses because data were not available.

INDEX TERMS

Medical Subject Headings (MeSH)

Audiovisual Aids; Memory Disorders [etiology] [*rehabilitation]; Multiple Sclerosis [*complications]; Randomized Controlled Trials as Topic; Therapy, Computer-Assisted [methods]

MeSH check words

Humans