scientific reports

Check for updates

OPEN

A systematic review and meta-analysis of socio-cognitive impairments in multiple sclerosis

Mandy Roheger[1,2], Lydia Grothe[1], Laura Hasselberg[1], Matthias Grothe[1,3]✉ & Marcus Meinzer[1,3]

Socio-cognitive impairment is frequent in multiple sclerosis (MS). However, little is known about the relationship between other potentially relevant clinical symptoms (i.e., cognition, depression, fatigue) and the degree of socio-cognitive impairment, and neural mechanisms underlying socio-cognitive deficits in MS. Therefore, we meta-analytically quantified socio-cognitive impairment in MS. A systematic literature search in MEDLINE Ovid, Web of Science Core Collection, CENTRAL, and PsycInfo was conducted until December 2022. Studies investigating affective or cognitive theory of mind (a/cToM), visual perspective taking (VPT) and social decision making (SDM) in MS patients relative to healthy controls were included. Risk-of-bias (RoB) was assessed using the CLARITY group "Tool for Assessing RoB in Cohort Studies". Mediation analysis investigated the contribution of clinical symptoms to socio-cognitive impairment. In total, n = 8534 studies were screened, 58 were included in the systematic review, 27 in the meta-analyses. Most studies were rated with a moderate RoB. Meta-analyses confirmed impairment of both aToM and cToM in MS patients, with larger effect sizes for aToM. Mediation analysis demonstrated that higher levels of fatigue selectively predicted the degree of cToM impairment. There was insufficient data available to quantify impairment in other socio-cognitive domains. Fourteen structural and functional imaging studies were identified and characterized by substantial heterogeneity. Summarized, this study confirmed substantial socio-cognitive impairment in MS and highlights the potential exacerbating role of comorbid clinical symptoms. We identify several evidence gaps that need to be addressed in future large-scale studies using comprehensive and coordinated assessments of socio-cognitive parameters, potential mediators, and neural correlates.

Trial registration: The pre-registered review protocol can be assessed at www.crd.york.ac.uk/PROSPERO/ (ID: CRD42020206225).

Multiple sclerosis (MS) is a chronic and progressive neurodegenerative disease leading to heterogenous neurological deficits[1], including up to 70% of patients suffering from cognitive impairment[2], fatigue[3] or depression[4]. Besides these well-documented symptoms, impairment of social cognition (SC), an umbrella term describing how people process, store, and apply information relevant to social interactions[5], have also been reported in MS[6,7]. For example, MS-patients may have problems understanding the emotions of others (affective Theory of Mind , aToM), or their cognitive states, beliefs, thoughts, or intentions (cognitive Theory of Mind, cToM)[8]. Several, recent meta-analyses have demonstrated moderate effect sizes for impairment of both aToM and cToM in patients with MS compared to healthy controls[9–11]. In the clinical presentation of MS-patients, there is a strong interplay between cognition, fatigue, and depression, which often complicates diagnostic evaluation and initiating adequate treatment[12]. However, the potential impact of these symptoms on socio-cognitive impairment has

[1]Department of Neurology, University Medicine Greifswald, Walther Rathenau Str. 49, 17489 Greifswald, Germany. [2]Ambulatory Assessment in Psychology, Department of Psychology, Carl Von Ossietzky University Oldenburg, Oldenburg, Germany. [3]These authors contributed equally: Matthias Grothe and Marcus Meinzer. ✉email: matthias.grothe@med.uni-greifswald.de

not yet been addressed. Moreover, while there is an extensive literature on the functional and structural brain correlates underlying impaired cognition[8], fatigue[13], and depression[14] in MS, only a few studies have used brain imaging methods, such as functional and structural magnetic resonance imaging (MRI), to investigate the neural underpinnings of socio-cognitive deficits in MS.

To address these open issues, we initially performed a systematic review of all studies assessing SC in patients with MS across three broad socio-cognitive domains (i.e., social perception, social understanding and social decision making[15]), investigated methodological biases and conducted a meta-analysis to quantify the degree of socio-cognitive impairment in MS relative to healthy controls. In a second step, we conducted a meta-regression analyses to identify potential mediators of socio-cognitive impairment in MS. Finally, we also provide an overview of studies that also investigated the functional and structural correlates of SC impairment in MS using MRI based measures.

Results

Search results
The search strategy yielded 10,615 articles (including the results of the update search). After deduplication, 8,534 unique articles had their titles and abstracts assessed for eligibility. From these articles, 8,424 articles were ineligible. 117 full-texts were assessed further and 58 studies included in the systematic review (Table 1), 14 of which also provided information on imaging data and the neural correlates of SC (Table 2). 27 studies were included in pairwise meta-analyses, 18 studies providing information on both social cognition and clinical outcomes in the mediation analysis. The PRISMA flow-diagram[16] in Fig. 1 illustrates the study selection process.

Included studies
In the following paragraphs, the results of the systematic review will be presented (1.1) as well as the systematic review results of studies including data on cognition, depression, and fatigue (1.2). In the following, the risk of bias analysis will be presented (1.3) and the results of the meta-analyses (1.4), which are divided in results for the analysis of cToM (1.4.1) and aToM (1.4.2). These results are followed by on overall sensitivity analysis (1.5). In a second step, we investigated the impact of depression, fatigue and cognitive status on socio-cognitive impairment in a meta-regression (2) and present the results on the neural correlates of socio-cognitive decline (3), again separately for aToM (3.1) and cToM (3.2).

Systematic review
Table 1 provides an overview of all included studies detailing demographic and clinical sample characteristics, socio-cognitive tasks used and the main results. Overall, 58 studies that investigated SC in patients with MS were included. Of those, all except for two studies[17,18] provided information on the type of MS. The majority of patients were diagnosed with relapsing–remitting MS (RR-MS). MS is generally more prevalent in women[19], which is also reflected in skewed sex distributions of the included study samples (with one exception[20]).

All except one study[21] used ToM as one of their outcomes. Cognitive ToM was assessed in 31 studies and assessed using either variations of the faux-pas task (a task comprising a situation/context where one character (the speaker) makes a statement that is unintentionally offensive to the listener because the speaker has a false belief), a false-belief task (two different types: first-order false-belief tasks involve attribution about other's false belief with regard to real events; second-order false-belief tasks assess what people think about other people's thoughts), a strange stories task (a selection of stories that test pretense, jokes, lies, white lies, misunderstandings, persuasion, appearance/reality, figures of speech, irony, double bluffs, contrary emotions, and forgetting) or a video test (a series of videotaped vignettes of social scenarios followed by questions about thoughts, feelings, and/or intentions of the characters). Cognitive TOM was found to be impaired in patients with MS in 26 of 31 studies. Affective ToM was measured in 52 studies using different types of emotion recognition or processing tasks, in which either the eyes or the whole face of a person were presented. Participants were then asked to label the emotion represented by the eyes/face. 42 out of 52 studies showed an impairment in emotion recognition and/or emotion processing in MS patients compared to healthy controls. Two studies investigated VPT[22,23], both showing significantly reduced imitation and poorer accuracy in perspective taking than healthy controls. SDM was investigated in two studies[21,24] using either a moral/conventional distinction task (in which social situations and moral and conventional transgressions were presented, as well as an authority jurisdiction)[24], or vignettes consisting of moral dilemmas (e.g. choice of whether or not to harm a person to save five other people)[21]. Results showed a lack of distinction between conventional and moral judgement in patients with MS[24], as well as a reduced moral permissibility in patients with MS[21].

Systematic review of studies including data on cognition, depression, and fatigue
Notably, even though information on the level of depressive symptoms, fatigue, and degree of cognitive impairment are essential for treatment decisions, only 18 out of 58 studies reported data on all three domains. All of them showed higher fatigue and depressive scores in patients with MS compared to healthy controls. In all studies, depression was described as mild because patients with severe symptoms were generally excluded. Several cognitive domains were reported as impaired in patients in MS compared to healthy controls, including executive functions[25,26], semantic fluency[27], and IQ scores[27]. Three studies[17,28,29] included depressive and fatigue scores as control variables in their analyses; two of them found that even after controlling for these variables, patients with MS performed significantly worse on aToM[28] and cToM tasks[28,29], one study found no significant influence of these variables on impaired ToM performance[17]. Two studies showed a significant negative correlation between aToM performance and depressive symptoms (higher depressive scores were associated with lower aToM scores)[1,30]. Only one study did not demonstrate correlations between socio-cognitive measures

Study and year	Sample number	Type of MS	Age (in years, Mean (SD))	Disease duration (in years, Mean (SD))	Gender (F:M)	Education (in years, Mean ± SD)	EDSS (Mean ± SD or median (range))	Depression score (and assessment tool) (Mean ± SD or median (range))	Fatigue score (and assessment tool)	Cognitive Tasks	Social Cognitive Task	Result
Adamaszek 2022*	MS:11 HC:11	RRMS: 11	37.4 ± 11.0 36.5 ± 8.5	8.5 ± 4.3	7:4 7:4	n.p.	2.7 (1.6)	n.p.	n.p.	PASAT, TAP, Wisconsin Card Sorting, WMS-R	Tübinger Affect Battery (Emotion recognition)	MS patients showed significant impairments in emotion recognition (subtests facial expression discrimination and facial expression matching).
Banati 2010	MS: 40 HC: 35	RRMS: 37 RSPMS: 3	36.2 ± 9.4 33.4 ± 7.8	Short-term MS (1-7y): 3.8 ± 2.9 Long-term MS (8-18y):13.9 ± 5.3	29:11 18:17	n.p.	low EDSS 1-2: 1.1 ± 0.35 high EDSS 2.5-4.5: 2.9 ± 1.0	BDI: 9.5 ± 7.4 4.1 ± 3.4	n.p.	Wechsler Adult Intelligence Scale-Revised	Faux-Pas Test; Baron-Cohen's Adult Eyes and Faces Test; Empathy	ToM impairment in patients with MS.
Batista 2017a*	MS. 60 HC: 60	RRMS: 50 SPMS:10	37.2 ± 7.5 36.1 ± 9.4	10.6 ± 6.6	40:20 40:20	13.2 ± 4.0 14.0 ± 3.9	2.0 (IQR: 1.5)	BDI: n.r.	MFIS; n.r.	SDMT-RA; PASAT-RA; BVMT-R; CVLT; JOLO; COWA; WCST	RMIE, ToM Videos test	Patients with MS have impairment on social cognition (RMET and video), independent from educational level.
Batista 2017b*	MS: 60 HC: 60	RRMS: 50 SPMS: 10	37.2 ± 7.5 36.1 ± 9.4	10.6 ± 6.6	40:20 40:20	13.2 ± 4.0 14.0 ± 3.9	2.0 (IQR 1.5)	BDI; n.r.	MFIS; n.r.	SDMT-RA; PASAT-RA BVMT-R; CVLT; JOLO COWAT; WCST	RMIE, ToM Videos test	Patients with MS have impairment on social cognition (RMET and video). Performance on ToM tests in patients with MS was not associated with age, sex, years of education, disease duration, EDSS, depression or fatigue.
Batista 2018	MS: 60 HC: 60	RRMS: 50 SPMS: 10	37.2 ± 7.5 36.1 ± 9.4	10.6 ± 6.6	40:20 40:20	13.2 ± 4.0 14.0 ± 3.9	2.0 (IQR 1.5)	BDI: 9.5 ± 7.0 3.8 ± 3.9	MFIS; 33.8 ± 19.6 16.5 ± 15.1	COWAT;WCST;Stroop Test; TMT-A/B; Interpretation of proverbs; RAPM	RMET, ToM Videos test	This study suggests a dissociation of executive functions and ToM in MS.
Beatty 1989	MS: 21 HC: 19	CPMS	52.0 (n.r.) 51.1 (n.r.)	18.40 (n.r.)	n.r.	14.0 (n.r.) 14.4 (n.r.)	6.60 (n.r.)	BDI: 7.30 ± 6.60 2.40 ± 4.30	n.p.	MMSE	BFRT POFA	Patients with PPMS were less accurate in judging emotional expressions and discriminating neutral faces.
Berneiser 2014	MS: 61 HC: 53	RRMS: 47 SPMS: 11 PPMS: 3	42.2 (range 19-68) 38.5 (range 20-67)	6.10 (range 1-21)	44:17 33:20	14.0 (range 9-24) n.r.	3.65 (range 1-8)	BDI: 9.21 (n.r.) 4.93 (n.r.)	MS-FS: 3.37 (n.r.)	PASAT	FAB	Patients with MS showed a poor performance in all subtests that required emotion recognition.
Bisecco 2020*	MS: 41 HC: 25	RRMS	34.2 ± 10.3 37.8 ± 12.0	8.8 ± 8.2	27:14 18:7	n.p.	1.5 (0-6.5)	BDI: median 7.0 (range 0-18) median 4.0 (range 0-14)	n.p.	SDMT	ToM Picture Sequencing Task (TMPS), RMIE	Patients were impaired in TMPS but not RMIE compared to controls. TMPS was correlated to cognition. TMPS scales were correlated to functional connectivity changes.
Bruno 2022	MS. 36 HC: 42	RRMS	39.2 ± 10.2 37.1 ± 10.7	9.3 ± 7.23	n.p.	16.6 ± 3.3 16.8 ± 2.7	1.5 ± 1.5	BDI-II: 13.8 ± 11.1 6.6 ± 7.2	MFIS: 39.0 ± 20.0	n.p.	RMET, Faux-pas Test, Perspective Taking	Poorer performance in ToM and perspective taking tests by patients with MS.
Cecchetto 2014	MS: 30 HC: 30	RRMS	34.2 ± 6.2 32.5 ± 6.4	9.1 ± 6.7	21:9 21:9	14.7 ± 2.0 15.2 ± 3.1	2.0 ± 1.0	BDI: 4.5 ± 4.7 4.2 ± 3.4	FSS: 3.9 ± 1.9	TMT; Corsi; Digit span; Phonemic Fluency; Benton RAO's Brief Repeatable Battery	Facial and Bodily Emotion Recognition Task	Facial emotion recognition can be impaired in patients with MS.
Chanial 2020	MS: 21 HC: 21	RRMS: 19 SPMS: 1 PPMS: 1	38.8 ± 5.5 33.9 ± 7.0	10.0 ± 7.0	15:6 16:5	13.1 ± 2.4 15.1 ± 2.3	4.2 ± 2.0 n.r.	n.p.	n.p.	MoCa; BDAE; MT 86;LAOC; SAOC; LAWC SAWC; TC	Theory of Mind Task	Performance in social cognition tasks was significantly lower in patients with MS.
Czekoova 2019*	MS: 43 HC: 43	RRMS	35.8 ± 8.0 34.7 ± 11.0	7.5 ± 4.4	31:12 25:18	n.p.	2.5 (0-6)	BDI: 10.5 ± 8.2 7.4 ± 6.1	MFIS: 29.7 ± 13.0 18.7 ± 13.6	SDMT	Perspective taking, imitation, RMIE	Patients showed significantly less imitation and poorer accuracy in perspective taking, but emotion recognition was similar in the MS patients and controls. Gray matter volume was higher in the HC relative to the MS patients in several regions.
Dulau 2017	MS: 60 HC: 65	RRMS: 30 SPMS: 15 PPMS: 15	46.5 ± 10.6 43.2 ± 9.3	14.4 ± 9.4	35:25 46:14	12.9 ± 3.3 12.5 ± 2.8	4.0 (n.r.)	BDI: 10.8 ± 7.9 7.8 ± 6.6	n.p.	FaFCSR; PASAT; IPS; CSCT; D2 Stroop 45; VF WCST	Theory of Mind tasks, Emotion Recognition, Emotional Awareness, Emotional Fluency	SC impairment was found in all phenotypes and was more prominent in cognitively impaired patients.
Ehrlé 2020	MS: 46 HC: 46	RRMS	38.7 ± 10.0 39.2 ± 10.8	3.4 ± 1.7	37:9 37:9	10.2 ± 4.0 10.8 ± 2.30	3.4 ± 1.7 n.r.	n.p.	n.p.	Digit Span, WAIS-III, fluency, Boston Naming, PASAT, Wisconsin Card Sorting.	Facial Emotion Recognition, Discrimination, Theory of Mind	The study suggests a lack of distinction between conventional and moral judgment in MS.

(continued)

3

Fereydouni 2019	MS: 51 HC: 51	RRMS	34.9 ± 7.9 31.8 ± 9.4	6.0 (n.r.)	42:9 44:7	n.p.	n.r.	GHQ: 9.8 ± 4.1 9.9 ± 3.9	n.p.	n.p.	Facial Emotion Recognition Task (FAB)	Facial emotion recognition is impaired at early stages of MS.
Genova 2016	MS: 15 HC: 15	RRMS: 10 SPMS: 2 PPMS: 2 SRMS: 1	49.5 ± 8.0 38.9 ± 13.1	18.0 ± 10.3	11:4 5:10	15.0 ± 1.8 14.7 ± 2.3	n.p.	n.p.	n.p.	WASI-II;WAIS-IV; SDMT; PASAT; CVLT	Theory of Mind (TASIT)	The results of the study indicate that the MS group was significantly impaired on the interpretation of lies and sarcasm compared to HCs.
Genova 2020	MS: 17 HC: 15	SPMS: 4 PPMS: 9 PRMS: 2	51.9 ± 9.3 45.6 ± 11.7	13.8 ± 9.5	9:8 11:4	15.6 ± 2.2 15.9 ± 2.3	n.p.	n.p.	n.p.	BVMT-R; CWIs; JOLO; PASAT SDMT	Theory of Mind (TASIT)	The findings suggest individuals with progressive MS were impaired across multiple social cognition domain.
Gleichgerrcht 2015	MS: 38 HC: 38	RRMS	42.3 ± 11.3 39.3 ± 8.1	1.6 ± 8.7	33:5 33:5	15.4 ± 2.8 15.7 ± 1.8	1.7 ± 1.6	n.p.	n.p.	n.p.	Moral dilemmas, Empathy (Interpersonal Reactivity Index)	Relative to control, patients exhibited decreased levels of other-oriented empathy.
Goitia 2020	MS:36 HC: 42	RRMS	39.2 ±10.2 37.1 ±10.7	9.3 ± 7.3	30:6 29:13	16.6 ± 3.3 16.8 ± 2.7	n.r.	BDI: n.r.	MFIS: 39.0 ±20.0 19.1 ±15.3	WCST,VF, TMT-B, WAIS-III	Theory of Mind, Faux-pas Test	Social cognition is significantly reduced in patients with MS.
Golde 2020*	MS: 30 HC: 30	RRMS	40.2 ± 9.9 39.6 ± 8.4	8.2 ± 5.0	18:12 19:11	12.2 ± 1.4 12.0 ± 1.4	1.8 (0-4)	HADS: 2.4 ± 2.9 1.6 ± 2.3	FSMC: 50.0 ±21.8 27.6 ± 6.7	SDMT VLMT TAP WAIS-II BVMT	Movie for the Assessment of Social Cognition, Emotion Recognition (implicit and explicit FacePuzzle task), Empathy (Multifaceted Empathy Test)	Patients with MS showed significant deficits in implicit emotion recognition.
Hälbig 2020	MS:29 HC:29	RRMS: 21 CIS: 8	Only subgroups reported	Only subgroups reported	Only subgroups reported	Only subgroups reported	Only subgroups reported	BDI / HADS: n.r.	MFIS: n.r.	MMSE BRB-NT	Emotional processing	Emotional processing changes may be present in early MS.
Henry 2009	MS: 27 HC: 30	n.r.	47.0 ± 11.0 44.3 ± 9.6	7.0 ± 6.1	18:9 19:11	15.0 ± 3.4 14.8 ± 2.6	1.9 ± 2.0	GDS: 9.1 ± 6.7 7.6 ± 6.4	n.p.	SEFCI (Delayed recall, vocabulary, abstraction, SDMT)	RMET-R	MS participants were significantly impaired on the ToM task, and presented with specific deficits decoding facial emotions of anger and fear.
Henry 2011	MS: 64 HC: 30	RRMS	42.4 ± 9.8 38.6 ±13.9	9.1 ± 5.4	50:14 21:9	11.1 ± 3.1 12.4 ± 3.3	2.3 ± 1.7	BDI: 6.2 ± 5.1 2.2 ± 1.9	MFIS: 79.4 ± 37.8	IQ, WAIS-R/7SF, BSAT WAIS-R/SS	Theory of Mind Task, Facial Emotion Recognition Task	MS patients performed significantly worse than controls in emotion recognition and all ToM tasks.
Henry 2015	MS: 64 HC: 30	RRMS	42.5 ± 10.0 38.6 ± 13.8	9.1 ± 5.0	50:14 21:9	10.8 ± 3.0 12.4 ± 3.0	2.4 ± 1.8	BSI-FS: 2.30 ± 2.6 0.8 ± 1.1	EMIF-SEP: 79.4 ± 37.8	WAIS-R/7SF; BSAT; FT	Theory of Mind Task (Faux-pas), Emotion Recognition Task	Emotional impairment is observed at early stages of the disease in the absence of cognitive dysfunction.
Henry 2017	MS: 62 HC: 33	RRMS: 31 SPMS: 15 PPMS: 16	46.8 ± 10.9 43.7 ± 10.5	11.4 ± 9.4	36:26 24:9	10.3 ± 2.3 11.0 ± 2.6	3.8 ± 1.8	HADS: 6.4 ± 4.0 4.5 ± 3.3	EMIF-SEP: 48.2 ±22.0 36.6 ±20.1	WAIS-III, WCST, VF, BFRT	Theory of Mind Task (Faux-pas), Emotion Recognition Task	The results suggested that there may be qualitative differences in social cognition difficulties among the phenotypes of MS.
Henry 2021	MS: 106 HC: 53	n.r.	42.0 ±11.7 38.3 ±13.0	8.5 ± 8.6	68:38 33:20	11.0 ± 2.6 11.5 ± 2.6	3.0 ± n.r.	HADS 5.9 ± 4.0 4.8 ± 5.1	EMIF-SEP: 50.0 ± 22.6 38.4 ± 19.3	DSB ; WCST; PF	Emotion Recognition Task, Theory of Mind Task (False-belief, Faux-pas)	PwMS performed significantly more poorly in FER and ToM.
Ignatova 2020	MS: 36 HC: 36	RRMS	high EDSS 43.7 ± 8.5 n=18 low EDSS 41.9 ± 11.6 n=18 HC 42.4 ± 12.3	high EDSS 11.2 ± 7.5 low EDSS 7.1 ± 4.3	high EDSS 11:7 low EDSS 13:5 HC 24:12	high EDSS 12.7 ± 1.7 low EDSS 13.2 ± 2.3 HC 12.2 ± 3.5	high EDSS 4.6 ± 1.0 low EDSS 1.9 ± 0.8 HC n.p.	BDI-II: high EDSS 14.4 ± 7.2 low EDSS 14.2 ± 10.8 HC 9.6 ± 7.6	n.p.	VF; FITMR; VSO; TMT-A; TMT-B; SDMT; PASAT 2	Theory of Mind Task, Emotion Recognition	It was found that, during the course of MS, deterioration of both social cognitive skills and basic cognitive abilities occurs, which is parallel to physical disability.
Isernia 2019	MS: 42 HC: 26	RRMS: 26 SPMS: 8 PPMS: 8	52.4 ± 10.3 51.4 ± 12.4	21.2 ± 11.0	24:18 19:7	12.8 ± 4.2 14.9 ± 4.9	6.0 (n.r.)	BDI-II: 8.9 ± 8.7 5.0 ± 4.9	n.p.	MoCa; BRB-NT	Theory of Mind Task (Strange Stories, Faux-pas), Emotion Recognition (RMET)	Statistically significant groups differences in cognitive but not affective ToM were found, with a lower performance in PrMS than those with a RRMS disease course.
Isernia 2020*	MS:35 HC: 21	RRMS: 16 PMS: 19 (9 PP, 10 SP)	50.6 ± 11.5 54.6 ± 11.2 53.1 ± 10.4	14.0 ± 10.3 22.21 ± 10.8	10:6 13:6 10:11	12.6 ± 2.7 11.4 ± 4.6 13.4 ± 3.2	4.3 (1-6.5) 6.4 (4-8)	BDI: 8.9 ± 8.6 10.0 ± 7.4 3.5 ± 4.2	n.p.	MoCa BRB	Strange stories, Faux-pas, RMIE, Movies for the Assessment of Social Cognition	Progressive MS patients had lower performance in cognitive ToM composite scores. Cognitive ToM correlated with cognition in progressive MS patients. No differences in affective ToM composite scores. Progressive MS patients revealed significant correlations between cognitive ToM scores and DTI metrics.

(continued)

Jehna 2010	MS: 20 HC: 23	RRMS: 7 SPMS: 1 CIS: 12	36.4 ± 9.3 28.2 ± 6.9	RRMS 8.0 ± 7.86 SPMS 6.0 ± n.r. CIS 0.6 ± 0.6	13:7 18:5	13.6 ± 1.9 14.3 ± 1.8	1.8 ± 0.9	n.p.	n.p.	FST	Facial Emotion Recognition	The patients demonstrated significant decreased reaction-times regarding emotion recognition tests compared to HC.
Jehna 2011*	MS: 15 HC: 15	RRMS	29.5 ± 9.6 30.3 ±10.6	7.3 ± 6.5	10:5 10:5	13.8 ± 2.9 15.4 ± 3.0	1.70 ± 1.1 (median 2.0, range 0-3.5)	BDI 5.4 ± 5.8 5.6 ± 8.4	n.p.	BRB, Wisconsin Card Sorting Test	Facial expressions labeling, facial expression matching task	Patients with MS and controls had comparable behavioral data, but MS patients demonstrated increased fMRI activations compared to HC.
Koubiyr 2021*	MS.20 HC:15	RRMS	37.8 ± 9.0 33.5 ± 7.8	9.8 ± 6.0	16:4 12:3	n.p.	1.5 (0 - 4.5)	HADS 2.7 ± 2.2 1.2 ± 1.6	MFIS 33.0 ± 21.4	TMT, Go/No-go task, Wisconsin Card Sorting, Word List Generation	Faux pas, false belief task, RMIE	MS patients and controls showed similar performance each social cognition task. Patients had alterations in the left amygdala functional and structural connectivity.
Kraemer 2013	MS:25 HC:25	RRMS	30.9 ± n.r. 33.4 ± n.r.	14.9 ± 3.0 months	15:10 11:14	11.5 ± n.r. 11.9 ± n.r.	n.r.	BDI-I: 9.2 ± 1.1 4.4 ± 0.9	n.p.	Stroop test, TMT-A; TMT-B; NLT	Theory of Mind (MASC), Empathy	The findings suggest that theory of mind and empathy are deficient even at early stages of RRMS.
Kumcu 2022	MS: 44 HC: 51	RRMS	35.6 ± 8.4 35.1 ± 7.5	8.2 ± 5.6	33:11 35:16	12.7 ± 4.0 13.0 ± 4.0	1.2 ± 1.0	BDI-II: 11.8 ± 7.4 10.6 ± 10.4	n.p.	MoCa	Benton Face Recognition Test	MS patients showed a worse performance in face recognition.
Labbe 2021*	MS.68 HC:50	RRMS	37.4 ± 11.2 38.0 ± 10.8	5.1 ± 3.7	48:20 26:24	n.p.	1.0 (1 - 4.5)	BDI: 21.0 ± 15.4 15.5 ± 7.4	FSS: 2.9 ± 1.7 2.0 ± 0.8	SDMT; PASAT; CVLT; BVMT-R; FAS; Stroop	Mini-Social Cognition and Emotional Assessment (shortened faux-pas and facial expression recognition task)	Patients with MS showed significant deficits in facial expression recognition, but not in the faux-pas task compared to controls. Regional gray matter volume reduction in the left insula and both medial frontal regions correlated with social cognition abilities in patients and controls.
Labbe 2020*	MS:45 HC:47	RRMS	36.3 ±10.2 37.6 ±12.0	4.3 ± 3.7	22:23 21:26	n.p.	1.0 (0-5)	n.p.	n.p.	BICAMS	Mini-Social Cognition and Emotional Assessment (shortened faux-pas and facial expression recognition task)	Patients with MS showed significant deficits in facial expression recognition, but not in the faux-pas task compared to controls.
Lancaster 2019	MS: 15 HC: 15	SPMS: 4 PPMS: 11	48.9 ± 8.6 45.6 ±11.7	14.4 ± 9.1	7:8 11:4	15.9 ± 2.1 15.9 ± 2.3	n.p.	n.p.	n.p.	BICAMS SDMT CVLT-II BVMT-R	Cognitive and affective ToM	The findings demonstrate that ToM deficits in progressive MS may be limited to cognitive ToM, while affective ToM is conserved.
Lenne 2014	MS: 55 HC: 21	RRMS	39.8 ± 7.7 36.4 ±10.1	7.5 ± 5.9	44:11 13:8	13.5 ± 2.8 14.1 ± 2.5	2.0(range: 0-6)	BDI-II: 12.8 ± 10.2 2.6 ± 4.0	n.p.	SRT PASAT DDS RDS	Emotion Recognition	Deficit in emotion recognition was independent from disability.
Massano 2021	Pediatric-onset MS: 30 Adult-aged onset MS: 30 Adult-onset disease duration matched MS: 30	Pediatric onset, adult-onset MS, RRMS, SPMS	POMS: 30.4 ±8.8 AOAMS: 33.9 ± 4.9 AODMS: 39.0 ±7.8 HC-Age-matched: 30.8 ± 7.1 HC-disease matched: 36.4 ± 8.0	POMS: 15.5 ± 9.1 AOAMS: 7.9 ± 4.6 AODMS: 12.1 ± 6.1	POMS: 21:9 AOAMS: 20:10 AODMS: 20:10 HC-A: 21:9 HC-B: 21:9	POMS: 13.3 ± 2.8 AOAMS: 13.9 ± 3.5 AODMS: 14.1 ± 3.9 HC-A: 14.0 ± 4.2 HC-B: 14.6 ± 4.1	POMS: 1.5(IR:1.1) AOAMS: 2.0(1.0) AODMS: 2.0(1.1)	n.p.	n.p.	BICAMS (CVLT, BVMT, SDMT)	Emotion Recognition, Theory of Mind	Patients with POMS were more prone to develop impairment on classic cognitive domains than on ToM ability, when compared with AOMS patients.
Mike 2013*	MS: 49 HC: 24	RRMS: 44 SPMS: 5	39.8 ± 9.3 36.7 ± 7.3	9.5 ± 6.2	31:18 13:11	n.p.	2.4 ± 1.7	BDI; n.r.	n.p.	n.p.	Baron-Cohen's Adult Eyes and Faces Test; Faux-Pas Test	Patients with MS performed significantly poorer in the Faces test and in the Eyes test but not in Faux pas test after correction for anxiety, depression, and gender, compared to controls.
Montembeault 2022	MS: 29 HC:27	RRMS: 20 PPMS: 2	MS young: 32.6 ± 5.2 MS old: 64.1 ± 5.1 HC young: 37.3 ± 7.4 HC old: 64.9 ± 5.7	MS young: 9.2 ± 3.1 MS old: 21.7 ± 17.3	24:7 18:9	MS young: 15.7 ± 2.4 MS old: 13.9 ± 1.9 HC young: 15.8 ± 2.2 HC old: 15.8 ± 2.1	MS young: 0.8 ± 1.2 MS old: 3.8 ± 2.4	BDI-FS MS young: 1.1 ± 1.2 MS old: 3.1 ± 3.0 HC young: 0.2 ± 0.4 HC old: 0.8 ± 1.1	MFIS young: 15.1 ± 13.1 MFIS old: 13.1 ± 14.7	MoCa, SDMT; BTA	Experimental emotion recognition task	HC performed better in recognition of facial emotions compared to patients with MS.
Neuhaus 2018	MS: 35 HC: 34	RRMS: 25 SPMS: 2 PPMS: 8	43.8 ± 12.1 43.9 ± 12.5	12.9 ± 9.6	22:13 22:12	n.p.	n.p.	HADS: n.p.	FSMC: n.p.	SDMT VLMT 4 BVMT	Geneva Social Cognition Scale, including ToM, Emotion Recognition	The impairment in the group was restricted to high order and affective social cognition tasks and independent of general cognitive performance, EDSS, disease duration and depression.
Ouellet 2010	MS: 41 HC: 20	RRMS: 22 SPMS: 13 CPMS: 3 uMS: 1	Only subgroups reported	Only subgroups reported	27:14 10:10	Only subgroups reported	Only subgroups reported	BDI: Only subgroups reported	n.p.	Bells Test, PASAT, WAIS-III, RAVLT, Clock drawing, TMT A/B, OWAT, Zoo map test, Card sorting test	ToM tasks (Strange stories task, faux-pas task, video task)	MS patients with cognitive impairments were found to have more difficulties attributing mental states to others than did cognitively intact MS patients and HCs.

(continued)

Study	Sample	MS type	Age	Disease duration	Sex (f:m)	EDSS	Depression/Anxiety	Fatigue	Cognitive tests	Social cognition tests	Results
Pfaff 2021	MS: 25 HC: 27	RRMS	42.8 ± 9.9 41.5 ± 10.0	10.4 ± 8.9	25:0 27:0	13.7 ± 2.1 14.3 ± 3.1	1.6 ± 1.6	EHD: 22.0 ± 6.0 15.4 ± 4.2	MFIS: 45.2 ± 24.2 20.2 ± 7.3	SPART, WAIS-IV, TAP, SDMT, fNART	Emotion recognition (FAB)	The MS group had more difficulties in recognizing emotions compared to the HC group.
Phillips 2011	MS: 32 HC: 33	RRMS: 27 SPMS: 3 PPMS: 2	44.0 ± 9.2 44.4 ± 9.8	7.9 ± 5.5	28:4 24:9	15.3 ± 3.4 16.7 ± 3.5	n.p.	HADS: 5.2 ± 4.6 2.2 ± 2.0	n.p.	FAS Fluency Memory task from SEFCI SART	Experimental emotion Perception task	The results indicate a specific deficit in decoding static and dynamic information about emotion in MS, as compared to non-emotional information.
Pinto 2012	MS: 56 HC: 56	RRMS:48 SPMS: 3 PPMS:5	38.9 ±10.3 37.4 ±10.6	n.p.	32:24 31:25	13.2 ± 4.5 13.8 ± 4.1	2.5 ± 2.0	HADS: 4.5 ± 4.0 2.9 ± 3.1	n.p.	MMSE, Digit Span, Corsi-Block Test, Fluency, Wisconsin Card Sorting	Emotion Recognition	No significant differences were found between MS and HC on ERT's behavioral and oculomotor measures.
Pitteri 2019*	MS: 31 HC: 38	RRMS	36.3 ± 7.6 37.1 ± 8.9	7.0 ± 4.5	24:7 28:10	13.4 ± 3.4 14.6 ± 3.4	1.0 ± 3.5	DASS: n.r.	n.p.	BRB, Stroop test	RMIE, Facial affect recognition (naming), empathy quotient	Patients had fewer correct response in the RMIE; the FAR (fear, anger), and the EQ. None of the cognitive measures predicted significantly the SC abilities.
Pöttgen 2013	MS: 45 HC: 45	RRMS: 31 SPMS: 8 PPMS: 6	42.4 ± 10.7 42.5 ± 10.5	8.5 ± 6.2	31:14 31:14	14.2 ± 2.8 13.7 ± 2.1	3.47 (n.r.)	HADS: 7.4 ± 14.4 n.r.	n.p.	MCVITB, SDMT, VeLMT LPSst4	MASC	The results suggest impaired social cognition in MS.
Prochnow 2011	MS: 35 HC: 61	RRMS: 5 SPMS: 29 PPMS: 1	48.2 ±10.2 33.5 ± 11.5	9.2 ± 8.4	12:23 24:37	10.8 ± 2.7 12.6 ± 2.4	6.0 (n.r.)	11.4 ± 7.7 6.1 ± 5.5	n.p.	MMSE, BFRT FST	PCFAE ESFT	Relative to HC, PwMS were impaired in facial affect recognition on four of the six Ekman basic emotions, except happiness and disgust.
Radlak 2021	MS: 53 HC: 31	RRMS:29 SPMS: 21 PPMS: 3	50.0 ±10.4 50.2 ±11.5	11.9 ± 8.0	38:15 19:12	14.6 ± 2.9 15.8 ± 3.3	n.p.	n.p.	n.p.	Fluency, Digit Symbol Task	The Awareness of Social Inference Test, Florida Affect Battery, Emotion Recognition	PwMS performed worse than demographically matched controls on all measures of emotion perception.
Raimo 2017	MS: 40 HC: 40	RRMS: 36 SPMS: 2 PPMS: 2	40.6 ± 11.5 40.2 ± 11.4	8.2 ± 7.5	29:11 31:9	13.1 ± 3.8 13.0 ± 3.6	2.4 ± 1.5 n.p.	HDRS: 12.7 ± 10.4 4.8 ± 4.3	FSS: 39.7 ± 17.1 n.p.	MMSE BRB-NT	ToM, Emotion Recognition	MS patients performed significantly worse than controls on tasks assessing cognitive and affective ToM, in verbal and nonverbal modality.
Realmuto 2019	MS: 45 HC: 45	RRMS	34.2 ± 7.7 33.0 ± 7.7	9.7 ± 6.2	31:14 32:13	13.5 ± 2.5 13.3 ± 3.1	2.1 ± 1.5 n.p.	HADS: 13.4 ± 6.9 n.p.	FSS: 61.0 ± 35.2 n.p.	BICAMS, CET Stroop Task	Emotion Recognition, Empathy	The majority of patients showed deficits at non-social tasks, particularly in the executive domains.
Roca 2014	MS: 18 HC: 16	RRMS	40.7 ± 9.5 40.9 ± 10.0	5.1 ± 3.8	n.p.	14.8 ± 3.6 16.5 ± 1.6	0.6 ± 1.0 n.r.	BDI: n.r.	MFIS: n.r.	PASAT, FrAB DSF, DSB, VF WCST, TMTB	Cognitive and affective ToM	The patients showed deficits in cognitive ToM, but their affective ToM seemed to be spared.
Sofologi 2019	MS: 25 HC: 30	PPMS	43.6 ± 5.4	n.p.	13:12	n.p.	n.p.	BDI: n.r.	n.p.	MoCa, NST, VF	ToM	The findings indicate that patients with PPMS show decline in emotion recognition and social inference abilities.
Turner 2021	MS: 20 HC: 20	RRMS: 17 SPMS: 1 PPMS: 2	47.3 ± 11.0 44.3 ± 11.4	n.p.	16:4 14:6	12.5 ± 1.8 13.3 ± 2.0	n.p.	HADS: 5.5 ± 3.0 3.2 ± 3.4	n.p.	n.p.	Emotion Recognition, Emotional Empathy, TASIT_S	PwMS performed worse in one subtest of social cognition.
Weinstein 1996	MS: 26 HC: 26	n.p.	n.p.	n.r.	n.r.	n.r.	n.p.	n.p.	n.p.	n.p.	n.r.	MS patients exhibited deficits on both visual and auditory measures of emotion identification with greater impairment for negative emotional stimuli.
Yap 2022	MS: 50 HC: 25	RRMS: 21 PPMS: 24	49.4 ± 9.4 44.1 ±10.9	12.3 ± 10.3	26:24 14:11	16.0 ± 2.7 16.4 ± 2.2	3.5 (range: 1.0-6.5)	HADS: 5.0 ± 0.5 2.8 ±0.7	MFIS: 33.6 ± 17.0	WASI-II, SDMT, WAIS-IV, CVLT-II, D-KEFS, CWIT	Cambridge Mindreading Battery, EsCOT, RMET	No differences between HC and patients with MS.
Yokote 2021*	MS: 20 HC: 27	RRMS:16 SPMS: 4	42.0 ±12.0 39.0 ± 8.6	10.0 ± 6.6	16:4 21:6	n.p.	2.0 (0-6)	HADS: n.r.	FSS: n.r.	SDMT	RMIE, Faux pas	RMIE and FPT performances were impaired in patients compared to controls. FPT but bot RMIE scores were significantly associated with volume of subcortical structures.

Table 1. Study overview: describes demographic and clinical characteristics, clinical symptoms, socio-cognitive tasks used and the main results. Studies shaded in grey provide data on depressive level, fatigue, and cognitive scores of patients with MS. MS, Multiple sclerosis; HC, Healthy controls; BDI: Beck's Depression Inventory; RRMS, relapsing–remitting multiple sclerosis; RSPMS, relapsing secondary progressive multiple sclerosis; EDSS, Expanded Disability Status Scale; n.r. , not reported; IQ, intelligence quotient; IQR, Interquartile range; MFIS, Modified Fatigue Impact Scale; ADS-L, Allgemeine Depressionsskala-Langform; AIT, Attribution of intentions test; APACS, Assessment of Pragmatic Abilities and Cognitive Substrates; AToM, Affective Theory of Mind; ATT, Advanced Test of ToM; BDI (-FS) , Beck Depression Inventory (-Fast Screen); BERT, behavioral emotion recognition test; BFRT, Benton Facial Recognition Test; BICAMS, Brief International Cognitive Assessment for MS; BRB-A/NT, Brief Repeatable Battery version A/of Neuropsychological Test;BSAT, Brixton Spacial Anticiation Test; BVMT (-R) , Brief Visuospatial Memory Test (-Revised); C&I, Conversation and Insinuation; CI, cognitive impairment; Cimp, cognitively impaired; Cint, cognitively intact; CIS, clinically isolated syndrome; COWAT, Controlled Oral Word Association Test; CPMS, chronic progressive MS; CrT, crossed taping; CSCT, Computerized Speed Cognitive Test; CToM, Cognitive Theory of Mind; CVLT, California Verbal Learning Test; CWIs, Color-Word Intereference subtest; DAGPVT, Dynamic Age and Gender Perception Video Task; DASS, Depression Anxiety Stress Scale; DDS, direct digit span; DEPVT, Dynamic Emotion Perception Video Task; DMT, decision making test; DSB, Digit span backward; DSCT, Digit Symbol Coding Task; DSF, Digit span forward; EAT, Emotion Attribution Task; EDSS, extended disability status scale; EEI, Expressive emotional intensity (adapted from Edman); EET, Emotion Evaluation Test; EMIF-SEP, Échelle modifiée d'Impact de la fatigue-sclérose en plaques; ESFT, Ekman-60-Faces test; FAB, Florida Affect Battery; FAR, Facial Affect Recognition FaFCSR, French adaptation of the Grober and Buschke Free and Cued Selecive Reminding Test; FED, Facial emotion discrimination; FEEST, Facial Expressions of Emotion: Stimuli and Test; FEM, Facial emotion matching; FEN, Facial emotion naming; FER, Facial emotion recognition; FER-FC, FER-forced choice; FER-FR, FER-free recall; FES, Facial emotion selection; FID, Facial identity discrimination; FITMR, Five items 10-min recall; FOFBT, First-order false belief task; FPi/e, FacePuzzle implicit/explicit; FPT, Faux Pas test; FPT-C/A, Faux Pas Test-cognitive/affective; FPT-I/E, Faux Pas Test-Intention/Emotion; FrAB, Frontal Assessment Battery; FSMC, Fatigue Scale for Motor and Cognitive Functions; FSS, Fatigue Severity Scale; FST, Faces Symbol Test; FT, Fluency test; GDS, Geriatric Depression Scale; GeSoCS, Geneva Social Cognition Scale; GHQ, general health questionnaire; G/NG, Go/No Go; HADS, Hospital Anxiety and Depression Scale; HDRS, Hamilton Depression Rating Scale; HC, healthy control; HCMRI, HCs tested in the MRI; HCPC, HCs tested on the PC; HI, Humor identification; IAPS, International Affective Picture System; IPS, Information processing speed; IQR, interquartile range; IRI (EC/PD/PT/F) , Interpersonal Reactivity Index (empathic concern/personal distress/perspective taking/fantasy); JOLO, Judgment of Line Orientation Test; LAOC, Lexical awareness in OC; LAWC, Lexical awareness in WC; LEAS, Levels of Emotional Awareness Scale; LF, letter fluency; LPSst4, Subtest 4 of the Leistungsprüfsystem; LT, labelling test; MASC, Movie for the Assessment of Social Cognition; MASC-T/I/A/C, MASC-Thoughts/Intention/Affective/Cognitive; MCVIT, Multiple Choice Vocabulary Intelligence Test; MFIS, modified fatigue impact scale; Mini-SEA, Mini-Social cognition & Emotional Assessment; MMSE, Mini Mental Status Examination Test; MoCa, Montreal Cognitive Assessment; MS-FS, MS-specific fatigue scale; MS, multiple sclerosis; n.d. , no data; NLT, number-letter task; n.p. , not performed; NST, Number Series test; OC, oral comprehension; PA, Picture Arrangement (subtest of WAIS-III); PASAT (-RA) , (Rao Adaptation of the) Paced Auditory Serial Addition Test; PatMRI, Patients tested in the MRI; PatPC, Patients tested on a PC; PCFAE, Test of Perceptual Competence of Facial Affect Recognition; PF, phonetic fluency; POFA, Pictures of Facial Affect; PPMS, primary progressive MS; PMS, progressive MS; PRMS, progressive-relapsing MS; PwMS, Patients with MS; Q-IDS, Quick Inventory of Depressive Symptoms; RAPM, The Raven's Advanced Progressive Matrices; RDS, reverse digit span; RMET(-R) , Reading the Mind in the Eyes Test (revised); RRMS, relapsing–remitting MS; RSPMS, relapsing secondary progressive MS; RTET, reaction times for emotion recognition test; RTGender, times for the gender test; RTPPET, reaction times for the Posner paradigm emotional test; SAOC, Syntaxic awareness in OC; SART, Sustained Attention to Response Task; SAWC, Syntaxic awareness in WC; SC, social cognition; SDMT (-RA) , (Rao Adaptation of the) Symbol Digit Modalities Test; SEFCI, Screening Examination for Cognitive Impairment; SET(-Tot) , Empathy Test (Total score); SF, semantic fluency; SI-E, Social Inference-Enriched; SI-M, Social Inference-Minimal; SIFET, Static Images of Facial Emotion Task; SIPPT, Static Identity Perception Photograph Task; SOFBT, Second-order false belief task; SPART (-D) , (delayed recall of the) Spacial Recall Test; SPMS, secondary progressive MS; SRT (LTS/CLTR/D) , Selective Reminding Test (Long Term Storage/Consistent Long Term Retrieval/Delayed recall); SST (MT/PT) , Strange Stories Task (Mental task/Physical task); SST-D/W/M/E, Strange Stories Task-Doublebluff/Whitelie/Misunderstanding/Emotions; TAP, Test Battery of Attentional Performance; TASIT, The Awareness of Social Interference Test; TC, Text comprehension; TMPS, Theory of Mind Pictures Sequencing Test; TMT(A/B) , The Trail Making Test (A/B); ToM, Theory of Mind; (c/e/min/mis) ToM, correct/excessive/minor/missing ToM; uMS, undetermined MS; VAMA (t/c/a) , Virtual Assessment of Mentalising Ability (total/cognitive/affective); VeLMT, Verbal Learning and Memory Test; VELS, Voice Emotion Labeling Subtest of the FAB; VF, Verbal fluency; VLMT, Visual Learning and Memory Test; VPT, visual perspective taking; VRI, Verbal Reasoning Index; VSO, visual spacial orientation; WAIS-IV, Wechsler Adult Intelligence Scale; WAIS-R/DSCS, Digit-symbol coding subtest of the WAIS-R; WAIS-R/7 SF, The Ward seven-subtest of the revised WAIS; WAIS-R/SS, similarities subtest of the WAIS-R; WASI-II, Wechsler Abbreviated Scale Intelligence; WC, written comprehension; WCST, Wisconsin Card Sorting Test; WLG, Word List Generation, Notes: (1) Batista et al.[28,41,43] used the same data from patients with MS and healthy controls. Data from Batista et al.[41,43] are expressed in percent, 2018 in points. (2) Carotenuto (2018): The SET-Test was acquired from 33/42 patients. (3) Cecchetto et al.[32]: The FSS was acquired from 26/30 patients. Studies with asterisk (*) provide imaging information and are therefore also included in Table 2.

Study	Design			Results
	Structural MRI	Task-fMRI	rs-fMRI	
Affective Theory of Mind				
Adamaszek et al. 2022	x			Facial affect matching negatively correlated with total white matter lesion volume
Batista et al. 2017a	x			Eyes test performance negatively correlated with total white matter lesion volume Eyes test performance did not correlate with total grey matter Eyes test performance positively correlated with regional grey matter volume in bilateral: Amygdala Entorhinal cortex Superior parietal gyrus Anterior cingulate cortex Posterior cingulate gyrus Superior temporal gyrus Fusiform gyrus Medial orbitofrontal cortex Putamen Supramarginal gyrus Videos test performance not correlated with regional grey mater volume
Batista et al. 2017b	x			Eyes test performance negatively correlated with total white matter lesion volume Eyes test performance correlated with DTI metrics (positive with FA, negative with MD) in: Corpus callosum Bilateral fornix stria Bilateral tapetum Bilateral uncinate fasciculus Left inferior cerebellar peduncle
Bisecco et al. 2020	x		x	Eyes test performance scores did not correlate with functional connectivity of the default mode, bilateral fronto-parietal executive, salience, cerebellar and limbic networks Eyes test performance scores did not correlate with regional gray matter volumes
Golde et al. 2020	x		x	Implicit emotion recognition negatively correlated with functional connectivity of the fusiform gyrus with lateral occipital gyrus, but not fatigue or motor/sensory impairments Implicit emotion recognition was correlated with lower FA values across widespread white matter networks of both hemispheres
Isernia et al. 2020	x			Affective ToM composite scores did not correlate with white matter DTI metrics (FA/MD) in patients Affective ToM composite scores did not correlate with cortical thickness in RRMS/PMS patients
Jehna et al. 2011		x		Facial emotion recognition task revealed increased activation in patients compared to healthy controls for anger and disgust in: Precuneus Posterior cingulate cortex
Koubiyr et al. 2022	x	x	x	Eyes test performance positively correlated with resting-state functional connectivity between: Left amygdala and left frontal pole Left amygdala and paracingulate gyrus Eyes test performance positively correlated with task-based functional connectivity between: Left amygdala and left infratentorial Left amygdala and temporal regions Structural connectivity (MD) was increased in patients compared to controls between: Left amygdala and left parahippocampal gyrus Left amygdala and temporal fusiform cortex
Labbe et al. 2020			x	Facial affect recognition task revealed decreased functional connectivity in patients with MS compared to controls between: Left orbitofrontal cortex and right superior temporal gyrus Left fusiform cortex and several left and right temporal, occipital and frontal areas Medial frontal cortex and right operculum
Labbe et al. 2021	x			Reduced facial affect recognition performance was positively correlated with reduced regional gray matter volume in patients with MS compared to controls in:

(continued)

				Left insular cortex Left caudate Right thalamus Bilateral frontal cortices Bilateral parietal lobes

Mike et al. 2013	x			Faces test and eyes test correlated positively with total T1-lesion volume but not T2-lesion volume
				Faces test positively correlated with regional T1-lesion load in:
				    Corpus callosum
				    Left Corona radiata
				    Right inferior longitudinal fasciculus
				    Right inferior fronto-occipital fasciculus
				    Uncinate fasciculus
				Faces test positively correlated with regional T2-lesion load in:
				    Corpus callosum
				    Left Fornix/ Stria terminalis
				Eyes test positively correlated with regional T1-lesion load in:
				    Corpus callosum
				Eyes test did not correlate with regional T2-lesion load
				Faces test positively correlated with cortical thickness (corrected for age, sex, disability, depression, anxiety) in:
				    Fusiform face area
				    Right entorhinal cortex
				Eyes test positively correlated with cortical thickness (corrected for age, sex, disability, depression, anxiety) in:
				    Left anterior inferior temporal gyrus
				    Left fusiform face area
				    Right frontal eye field

Pitteri et al. 2019	x			Faces test performance negatively correlated with cortical lesion volume in the amygdale Performance for fear and anger negatively correlated with cortical lesion volume in the amygdale Eyes test performance negatively correlated with cortical lesion volume in the amygdale Emotion quotient did not correlate with cortical lesion volume in the amygdale Emotion quotient negatively correlated with total cortical lesion volume
Yokote et al. 2021	x			Eyes test did not correlate with regional cortical thickness or subcortical volumes
Cognitive Theory of Mind Tasks				
Batista et al. 2017a	x			Video Test performance negatively correlated with total white matter lesion volume Video Test performance did not correlate with total grey matter Video Test performance did not correlate with regional grey mater volume.

Batista et al. 2017b	x			Video Test performance negatively correlated with total white matter lesion volume
				Video Test performance correlated with white matter tract DTI metrics (positive with FA, negative with MD) in:
				    Corpus callosum
				    Fornix stria
				    Uncinate fasciculus
				    Left tapetum
				    Right superior fronto-occipital fasciculus

Bisecco et al. 2020	x		x	ToM pictures sequencing task reciprocity score positively correlated with functional connectivity in:
				    Right middle temporal gyrus
				ToM pictures sequencing task first order score negatively correlated with functional connectivity in:
				    Posterior cingulate cortex
				ToM pictures sequencing task total score negatively correlated with functional connectivity in:
				    Cingulate cortex
				ToM pictures sequencing task reciprocity score negatively correlated with functional connectivity in:
				    Cingulate cortex
				    Right superior temporal gyrus
				ToM pictures sequencing task scores did not correlate with regional grey matter volumes

Czekóova et al. 2019	x			Imitation performance was positively correlated with grey matter volume in:
				    Left thalamus
				    Left anterior insula

(continued)

				Perspective taking accuracy was positively correlated with grey matter volume in: Left putamen
Isernia et al. 2020	x			Cognitive ToM composite scores correlated with white matter tract DTI metrics (positive with FA, negative with MD) in progressive MS-patients in: Superior and inferior fasciculus Forceps major Thalamic radiation Cognitive ToM composite scores negatively correlated with white matter tract DTI MD metrics in progressive MS-patients in: Right fronto-occipital fasciculus Bilateral superior longitudinal fasciculus Bilateral cortico-spinal tract Left uncinate Corpus callosum Cognitive ToM composite scores did not correlate with white matter DTI metrics in relapsing remitting MS-patients Cognitive ToM composite scores did not correlate with cortical thickness in MS patients
Labbe et al. 2020			x	Faux-pas task revealed increased functional connectivity in patients with MS compared to controls in: Left cerebellum and right occipital cortex Right amygdala and left occipital cortex
Labbe et al. 2021	x			Faux-pas task revealed reduced regional grey matter volume in patients with MS compared to controls in: Left insular cortex Right caudate Right thalamus Bilateral cingulate Bilateral frontal cortices Bilateral parietal lobes
Mike et al. 2013	x			Faux-pas task did not correlate with total T1-lesion or T2-lesion volume Faux-pas task did not correlate with cortical thickness
Yokote et al. 2021	x			Faux-pas task scores were positively correlated with volumes of the: Right thalamus Left pallidum Faux-pas task scores were positively correlated with cortical thickness in: Left fusiform gyrus Left orbitofrontal cortex Left temporo-parietal junction Left superior temporal gyrus

MRI = magnetic resonance imaging, fMRI = functional magnetic resonance imaging, rs = resting-state, ToM = theory of mind, FA = fractional anisotropy, MD = mean diffusivity, DTI = diffusion tensor imaging

Table 2. Results of studies that investigated neural correlates of socio-cognitive deficits using structural and functional MRI measures. MRI, Magnetic resonance imaging, fMRI, Functional magnetic resonance imaging, rs, Resting-state, ToM, Theory of mind, FA, Fractional anisotropy, MD, Mean diffusivity, DTI, Diffusion tensor imaging.
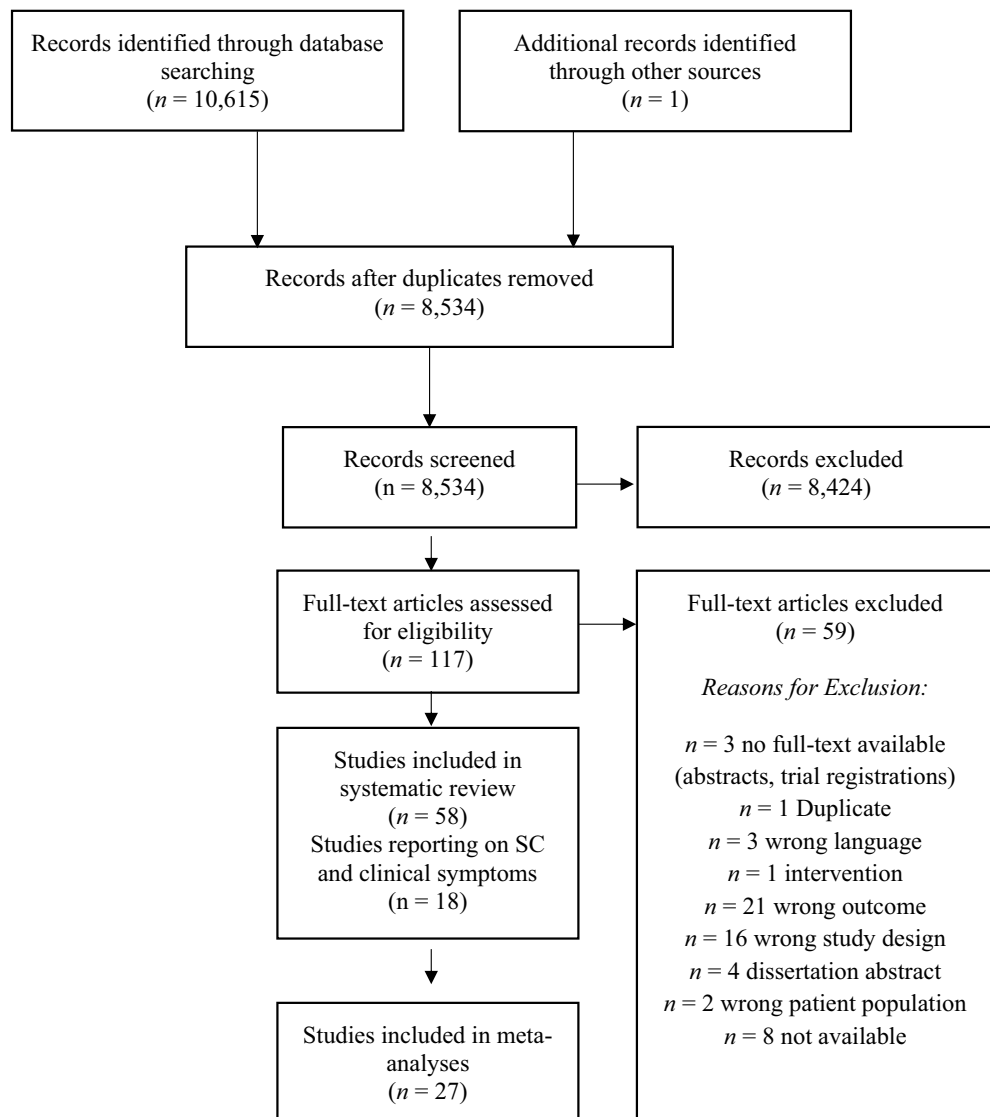
and depression[31]. Three studies showed a significant positive correlation between aToM and cognition (higher aToM scores were associated with better performance on long-term storage and retrieval tasks, the Symbol-digit-modalities test[32]; and semantic fluency and IQ[27]). Reduced performance in ToM correlated with poorer executive function, intellectual ability and episodic memory[31], and cToM and IQ and semantic fluency[27]. Pfaff et al. showed that inhibition and divided attention measures were predictive of difficulties in identifying facial emotions (aToM) in MS patients[33], Montembeault et al. reported a positive correlation between general cognition (Montreal Cognitive Assessment, MoCA), but not with attention, measured with the Brief Test of Attention, the Symbol-Digit-Modalities Test, or the Stroop Inhibition task[34].

Risk of bias assessment
Results of the RoB Assessment are shown in Table 3. Overall, only one domain was rated with a low RoB in all studies concerning the assessment of the study outcome, namely "Can we be confident in the assessment of the outcome?". Most studies did not match their groups for possible confounders (e.g., comorbid depression) or controlled for possible confounding variables, leading to either an unsure RoB when no data was provided or to high RoB in cases were confounders were not assessed.

Meta-analyses
Meta-analyses were only possible for the primary outcome (ToM), because there was not enough data available for the remaining socio-cognitive domains (SDM, VPT). Due to the high heterogeneity of outcomes and ToM constructs, we decided to calculate separate analyses for cToM (Results of Faux-Pas Tests and results of Video tests) and aToM (often labelled as emotion recognition). Furthermore, if applicable, we subdivided the meta-analysis in the respective tests used for assessment to ensure better comparability.

```
┌──────────────────────────────┐   ┌──────────────────────────────┐
│ Records identified through   │   │ Additional records identified│
│ database searching           │   │ through other sources        │
│ (n = 10,615)                 │   │ (n = 1)                      │
└──────────────────────────────┘   └──────────────────────────────┘

          ┌──────────────────────────────────────────┐
          │ Records after duplicates removed          │
          │ (n = 8,534)                               │
          └──────────────────────────────────────────┘

          ┌──────────────────────┐    ┌──────────────────────┐
          │ Records screened      │──▶ │ Records excluded      │
          │ (n = 8,534)           │    │ (n = 8,424)           │
          └──────────────────────┘    └──────────────────────┘

          ┌──────────────────────┐    ┌──────────────────────────────┐
          │ Full-text articles    │──▶ │ Full-text articles excluded   │
          │ assessed for          │    │ (n = 59)                      │
          │ eligibility           │    │                               │
          │ (n = 117)             │    │ Reasons for Exclusion:        │
          └──────────────────────┘    │                               │
                                       │ n = 3 no full-text available  │
          ┌──────────────────────┐    │ (abstracts, trial registrations)│
          │ Studies included in   │    │ n = 1 Duplicate               │
          │ systematic review     │    │ n = 3 wrong language          │
          │ (n = 58)              │    │ n = 1 intervention            │
          │ Studies reporting on  │    │ n = 21 wrong outcome          │
          │ SC and clinical       │    │ n = 16 wrong study design     │
          │ symptoms (n = 18)     │    │ n = 4 dissertation abstract   │
          └──────────────────────┘    │ n = 2 wrong patient population│
                                       │ n = 8 not available           │
          ┌──────────────────────┐    └──────────────────────────────┘
          │ Studies included in   │
          │ meta-analyses         │
          │ (n = 27)              │
          └──────────────────────┘
```

Figure 1. PRISMA Diagram of the study selection process.

Cognitive theory of mind

Faux-Pas test

Overall, ten studies were included in the meta-analysis that included variations of the faux-pas test and compared patients with MS and healthy controls. Four used the faux-pas task by Baron-Cohen et al.[35], five the faux-pas task by Stone et al.[36], and one study using a faux-pas task included in a Social Cognition battery. Results showed that healthy controls performed significantly better than MS patients across the different tasks: SMD = (−0.50), 95% CI (−0.85) to (−0.16), I^2 = 79%) and also in each of the two task versions where sufficient data was available for separate analyses (Baron-Cohen task: SMD = (−0.35), 95% CI −0.86 to 0.16, I^2 = 78%; Stone Task: SMD = (−0.70), 95% CI −1.26 to 0.14, I^2 = 84%). A forest plot for the outcome is displayed in Fig. 2a.

Video test

Seven studies were included that investigated differences in performance using four different types of ToM Video Tests between healthy controls and patients with MS. Overall results are reported, because there was not enough data available to consider individual outcomes. The overall result indicates that healthy controls performed significantly better on Video Tests than patients with MS (SMD = (−0.70), 95% CI (−1.21) to (−0.30), I^2 = 75%). The forest plot for the outcome is displayed in Fig. 2b.

Affective theory of mind/emotion recognition

Twenty studies were included in the meta-analyses that investigated aToM tasks. The overall effect size of the random effects analysis was −0.75(CI:(−0.93) to (−0.57), favoring healthy controls. Nine studies used the Baron-Cohen Adult Eyes Test, showing better performance in healthy controls compared to MS patients (SMD = −0.83,

Study	Was selection of exposed and non-exposed cohorts drawn from the same population?	Can we be confident in the assessment of exposure?	Did the study match exposed and unexposed for all variables that are associated with the outcome of interest or did the statistical analysis adjust for these prognostic variables?	Can we be confident in the assessment of the presence or absence of prognostic factors?	Can we be confident in the assessment of outcome?
Adamszek (2022)					
Banati et al.[6]					
Batista et al.[40]					
Batista et al.[27]					
Batista et al.[42]					
Beatty (1989)					
Berneiser et al.[1]					
Bisecco et al.[45]					
Bruno et al.[22]					
Cecchetto et al.[31]					
Chanial (2019)					
Czekoova et al.[21]					
Dulau (2017)					
Ehrlé et al.[23]					
Fereyduoni (2019)					
Genova (2016)					
Genova (2020)					
Gleichgerrcht et al.[20]					
Goitia et al.[28]					
Golde et al.[39]					
Hälbig (2020)					
Henry (2009)					
Henry et al.[36]					
Henry et al.[26]					
Henry (2021)					
Ignatova (2020)					
Isernia (2019)					
Isernia et al.[47]					
Jehna (2010)					
Jehna (2011)					
Kraemer et al.[38]					
Kumcu (2022)					
Koubiyr et al.[25]					
Labbe et al.[29]					
Labbe et al.[29]					
Lancaster (2019)					
Lenne (2014)					
Mike et al.[43]					
Montembeault et al.[33]					
Neuhaus (2018)					
Ouellet et al.[37]					
Patil (2017)					
Pfaff et al.[32]					
Phillips (2011)					
Pinto (2012)					
Pitteri et al.[58]					
Pöttgen (2013)					
Prochnow et al.[19]					
Radlak (2021)					
Raimo (2017)					
Realmuto (2019)					
Roca (2014)					
Sofolgi (2019)					
Turner (2021)					
Weinstein et al.[17]					
Yap et al.[30]					
Yokote et al.[46]					

Table 3. Risk of bias assessment. Red color indicates a high risk of bias, yellow color indicates a medium risk of bias, green color indicates a low risk of bias (dark green: no concerns, light green: small concerns, but still a low risk of bias), assessed with the "Tool for Assessing Risk of Bias in Cohort Studies" by the CLARITY Group[60].

95% CI (− 1.15) to (− 0.51), I^2 = 77%). Two studies used the Emotion Recognition Florida Affective Battery (SMD = − 0.87, 95% CI (− 1.15) to (− 0.59), I^2 = 0%); four studies used the Facial Expression of Emotion (FEEST) Tests (SMD = − 0.55, 95% CI (− 0.93) to (− 0.17), I^2 = 71%) and four studies used similar experimental emotional recognition tasks that were grouped together as they all used different tests (SMD = − 0.71, 95% CI (− 1.15) to (− 0.59), I^2 = 0%), all of them showing a better performance in healthy controls than patients with MS. The forest plot is displayed in Fig. 3.

Sensitivity analyses

The results of the sensitivity analyses using fixed effects models are displayed in Figs. 2a,b and 3 and did not show any significant differences compared to the random effects models.

Meta-regression analyses: Impact of depression, fatigue and cognitive status on socio-cognitive impairment

Meta-regressions analyses could be conducted for cToM(faux-pas), cTom(videos), and aToM, see Table 4. Due to the substantial heterogeneity of assessments that were used to quantify cognitive impairment, fatigue and depression, as well as insufficient data, we were not able to integrate the variable cognition, as well as depression for cToM(faux-pas) and fatigue for cTom(videos). In the analysis on cToM(faux-pas), that included three studies[27,29,37], R^2 was 100% (meaning 100% of the difference in true effect sizes can be explained by the predictor fatigue), and the intercept as well as the predictor fatigue were significant (intercept: − 1.73 ($p < 0.001$), fatigue: 0.02 ($p < 0.05$)). This means that performance on cTom (faux-pas) decreased with higher fatigue levels. In the meta-regression on cToM(video), four studies were included[28,38–40]. Yet, the overall model did not reach significance (R^2 = 0%), so the predictor depression did not explain any additional variance in the results. Five studies that investigated aToM were included[1,17,22,27,37]. The intercept (the difference in performance between healthy controls and patients with MS when all predictors have a value of 0) was significant (− 1.82, $p < 0.05$) and favored healthy controls, similar to the other results. Both predictors, depression (0.13), and fatigue (0.01) did not reach significance. The overall R^2 of the model was 28.73% and the I^2 was at 32.76%, indicating only low heterogeneity.

Neural Correlates of socio-cognitive decline

Fourteen studies provided information on the neural correlates of socio-cognitive impairment in MS. Among those, thirteen studies investigated aToM, nine studies investigated cToM (for a detailed overview please see Table 2).

aToM
Eleven studies used structural MRI to investigate neural correlates of aToM using a number of different outcomes (e.g., eyes or faces tests, aToM composite scores). Of those, four studies demonstrated positive correlations between total white matter lesion volume and aToM performance[41–44], one with total grey matter volume[43]. Using more regional approaches, two studies revealed association between aToM and white matter lesions[44] or other DTI derived metrics (e.g., fractional anisotropy, mean diffusivity[41]), highlighting the potential role of disconnection of regions associated with socio-cognitive processing. Six studies that used voxel-based morphometry or assessed regional cortical thickness demonstrated positive correlations between aToM and integrity of the amygdala, fronto-temporal and other regions[26,28,44,45], whereas two studies could not confirm these findings[46,47].

Four studies used task-free resting-state functional imaging[26,30,40,46] and three of these studies highlighted the contribution of frontal and temporal networks to aToM impairment in MS. Specifically, they demonstrated a negative correlation between implicit emotion recognition performance and functional connectivity of the fusiform gyrus with lateral occipital gyrus[40], a positive association between RMET performance and functional connectivity between the left amygdala and frontal pole/paracingulate cortex[26], or decreased functional connectivity between fronto-temporal regions in patients compared to controls during a facial affect recognition task[30]. On the other hand, Bisecco et al.[46] did not find any correlations between the RMET performance and functional connectivity of the default mode, bilateral fronto-parietal executive, salience, cerebellar and limbic networks.

cToM
Eight out of nine studies investigated neural correlates of cToM using structural MRI. Three studies investigated the association to the total white matter lesion volume, with two studies demonstrating a negative association between the lesion volume and the video test performance[28,43], and one study without an association between the lesion volume and the faux-pas task ability[44]. One study failed to show a significant correlation between cToM performance and total grey matter volume[28].

Regional grey matter volume reductions were investigated in seven studies. Of those, three studies demonstrated a significant correlation between reduced volume of the thalamus and reduced cToM performance[22,45,47] and additional positive correlations between grey matter volumes in different cortical regions (e.g., insula, frontal cortex, temporal and parietal cortex) and cToM. Four studies did not find any association between performance and regional grey matter integrity[43,44,46,48]. Two studies demonstrated an association between altered white matter DTI metrices (i.e., reduced FA and higher MD), especially in the corpus callosum and the superior fasciculus[41,48]. Two studies used resting-state fMRI and found that increased functional connectivity between the occipital cortex and the cerebellum/amygdala in MS-patients compared to controls was correlated with better cToM performance[30], as well as positive and negative correlations between different subscales of a cToM picture sequencing task and functional connectivity of the right middle temporal and (posterior) cingulate cortex[46].

(a) Forest Plots of cognitive Theory of Mind Tasks : Faux-Pas Tasks
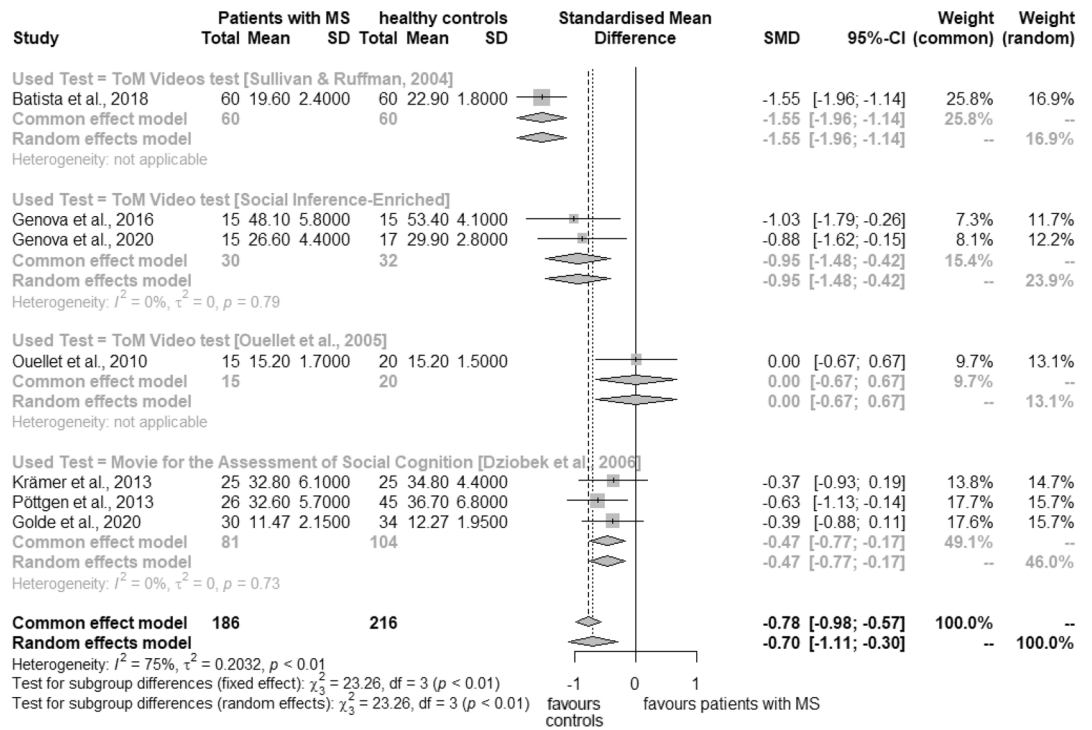
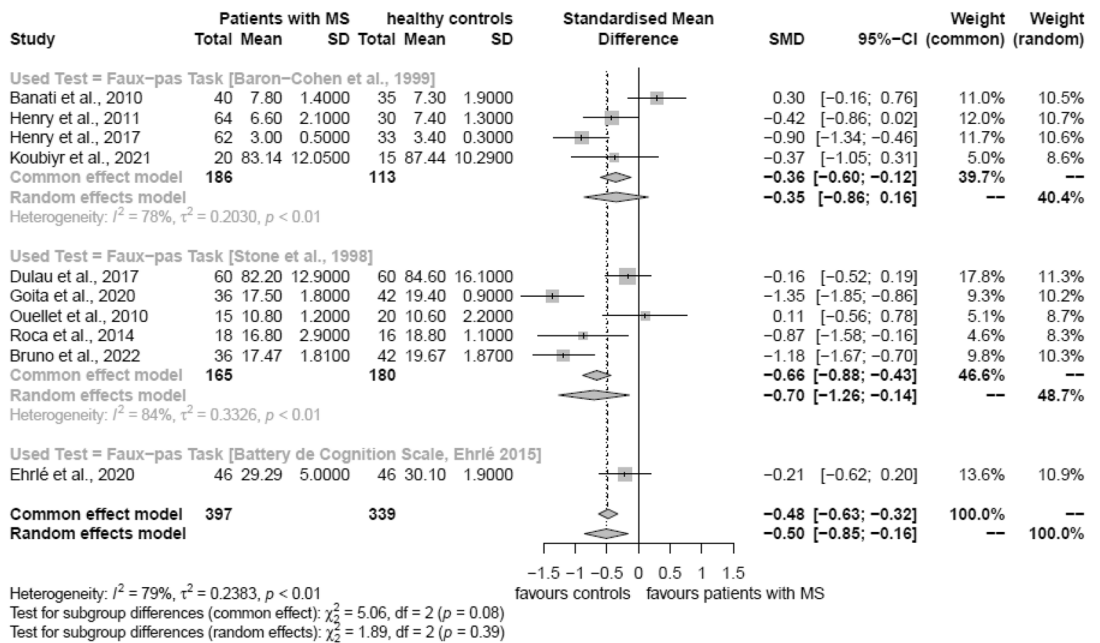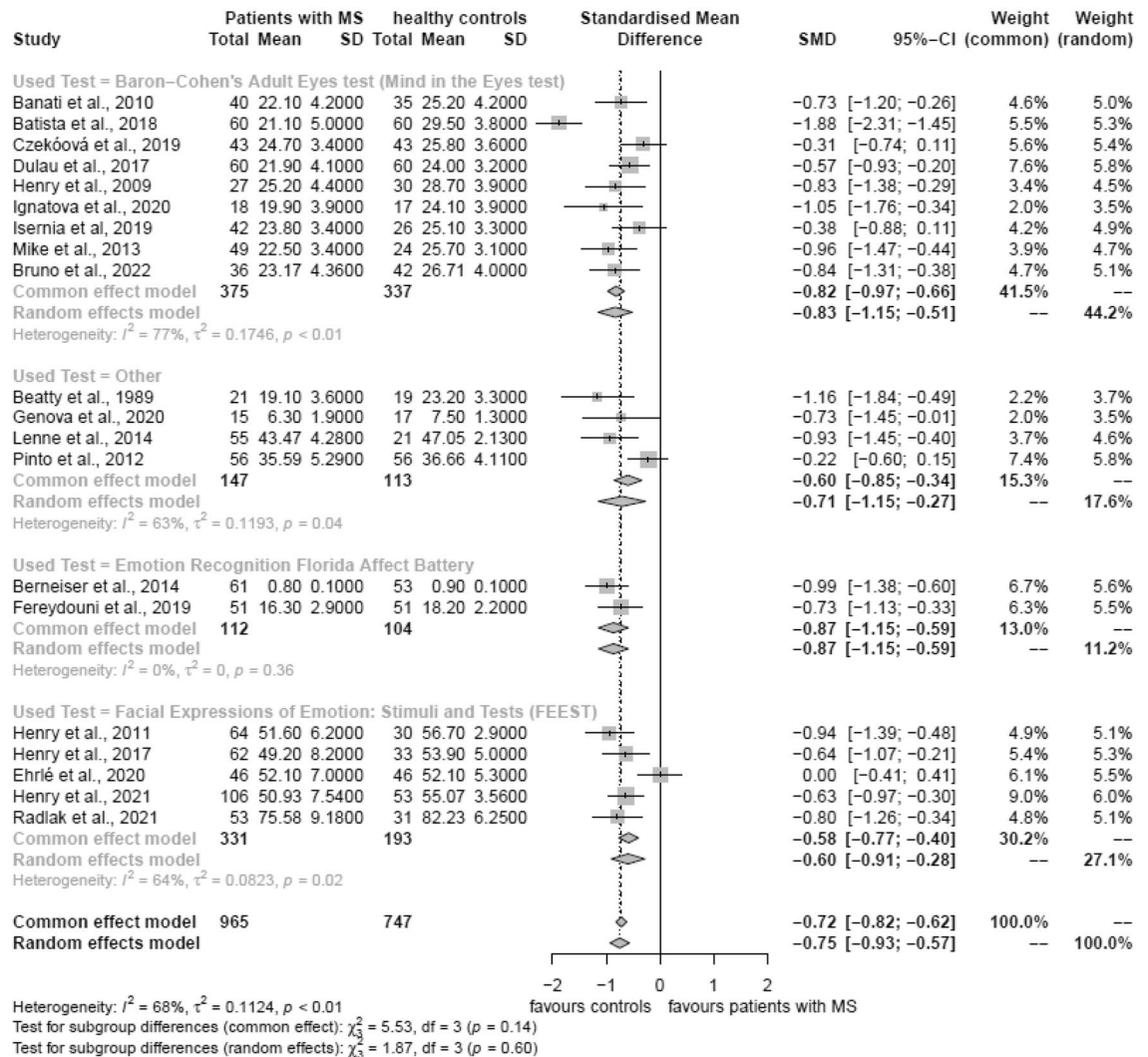## (b) Forest Plots of cognitive Theory of Mind Tasks :Video Test

Figure 2. Forest plots of cognitive theory of mind tasks.

Figure 3. Forest plots of affective theory of mind tasks.

Models	cToM	cToM	aToM
	Faux-Pas test $k=3$	Video test $k=4$	Emotion recognition $k=5$
	Estimate (SE)	Estimate (SE)	Estimate (SE)
Intercept	−1.73 ($p<.001$)	−1.03 ($p=.188$)	−1.82 ($p<.05$)
Depression	–	0.05 ($p=.527$)	0.13 ($p=.111$)
Fatigue	0.02 ($p<.05$)	–	0.01 ($p=.256$)
Cognitive status	–	–	–
I^2/R^2	0%/100%	88.62%/0.00%	32.76%/28.73%

Table 4. Results of meta-regressions. For the analysis of cToM: Faux-Pas test, only three studies could be included providing data on fatigue due to missing data in the other eligible studies and possible moderators. Goita et al. (2020) used the MFIS to assess fatigue, Henry et al.[27,37] used the EMIF-Sep to assess fatigue. For the analysis cToM: Video Test, we could include four studies providing data on depression. Batista et al.[28], Ouellet et al.[38] and Krämer et al. (2013) used Beck's Depressive Inventory (BDI) to assess depression, Golde et al.[40] used the Hospital Anxiety and Depression Scale: Subscale Depression (HADS-D). In the analysis aToM: Emotion Recognition, we could include five studies providing data on depression and fatigue. Three studies used the BDI and the MFIS[1,22, 37], and two studies used the HADS-D and the EMIF-Sep[27], Henry et al. (2021) to assess depressive symptoms and fatigue. Interpretation: I^2 = after inclusion of the predictor, XX% of the variability in our data can be attributed to the remaining between-study heterogeneity. R^2 = XX% of the difference in true effect sizes can be explained by the predictor. Intercept = The difference in the outcome variable between our two groups when all integrated predictors have a value of 0. Significant values are in [bold].

Four imaging studies provided information on depression, fatigue and cognition, with three studies investigating aToM[26,40,45], and two cToM[22,45], but used different measures to quantify depression, fatigue and cognition, that complicate meaningful comparison of the results.

Discussion

The present systematic review and meta-analysis confirms previous reports demonstrating socio-cognitive impairment in patients with MS[9–11] and suggests that comorbid cognitive and affective symptoms or fatigue can further exacerbate these impairments. The vast majority of eligible studies investigated different aspects of ToM and approximately 80% of the included studies reported impairment of either affective or cognitive ToM in MS patients relative to healthy control groups. Our meta-analyses demonstrated more pronounced impairment for aToM (ES = 0.8) compared to cToM (ES = 0.05-0.07). This pattern is in line with results of previous meta-analyses, that also demonstrated more pronounced impairment of aToM, especially for the RMET and facial emotion recognition tasks, compared to cToM (i.e., faux-pas tasks)[9,11]. While only 4/58 eligible studies investigated different aspects of social cognition in MS (VPT/SDM), all of them reported significant impairment compared to healthy control groups. This highlights the need to further investigate other socio-cognitive processes than ToM in MS and to determine the potential interplay with other clinical symptoms (i.e., depression, fatigue or cognitive status). This was not possible in the present study, due to the small number of available studies.

Eighteen of the included studies provided additional information on clinical symptoms that may impact on socio-cognitive impairment, but only three studies controlled for these variables in their analyses[17,27,28]. Nonetheless, the overall pattern of results from individual studies suggests that depression, fatigue and cognitive impairment can contribute to socio-cognitive impairment in MS. This was further supported by the results of our meta-regression analyses that demonstrated a specific contribution of fatigue to the degree of impairment in cToM, but not aToM. However, future research is needed, to systematically investigate whether specific clinical symptoms exacerbate the degree of impairment in different aspects of socio-cognitive functioning and to determine causal relationships between them.

It needs to be acknowledged that the results of this study are based on a relatively small number of studies and are therefore to be interpreted with caution. Nonetheless, our study included about 30% more studies compared to the most recent previous meta-analysis by Lin et al.[11]. Two earlier meta-analyses published in 2016 included only about half the number of studies[9,10], which highlights an emerging interest in this topic. This not unsurprising, because intact social functioning has been linked to relationship and vocational success, and better life satisfaction in healthy individuals[49]. Moreover, socio-cognitive impairment can have a profound negative impact on social participation, resulting in loneliness and poor mental health[50], which may be particularly detrimental in individuals attempting to cope with progressive conditions like MS. Nonetheless, the direct contribution of socio-cognitive impairment to reduced quality-of-life (QoL) in MS is currently unclear. For example, while Philips et al.[51] demonstrated that emotion regulation capacity was positively correlated with higher QoL in MS patients, others failed to demonstrate independent contributions of socio-cognitive impairment to QoL (e.g.[52]). Such discrepancies are likely explained by mutual interdependencies of both social cognition and QoL with clinical symptoms that are frequent in MS[53]. However, only about ~ 30% of our included studies reported information on specific socio-cognitive outcomes and (substantially varying degrees of) cognition, fatigue and depression. Only three studies controlled for these symptoms in their analyses, none reported associations with QoL. Thus, future systematic research is needed to disentangle the complex interactions between socio-cognitive impairment and cognition, fatigue and depression, and how they affect real-life outcomes, including QoL or the ability to cope with disease progression.

Finally, the systematic review of the anatomical and functional brain correlates underlying socio-cognitive impairment in MS revealed substantial heterogeneity between studies with regard to characteristics of the included patients, imaging methods, and outcome measures. As for the behavioral studies described above, the majority of imaging studies focused on different aspects of ToM (cTOM: 9 studies, aTOM: 13 studies). With regard to imaging methods, twelve studies used structural imaging and investigated global or regional grey and white matter changes. Only five studies employed functional MRI. Despite partially conflicting findings, these studies demonstrated that lesions affecting major cortical or subcortical hubs (e.g., orbito-frontal or insular cortex, the amygdala) within task-relevant regions of the "social brain"[54] or domain-general networks (e.g., ventral/dorsal attention, salience or default networks) can be related to the degree of specific socio-cognitive impairment. Similarly, several studies demonstrated the contribution of white matter pathways (e.g., corpus callosum, uncinate fasciculus, superior longitudinal fasciculus) or functional connectivity changes in specific networks to socio-cognitive impairment. However, aside from the general heterogeneity (including paradigms, methods for data acquisition and analyses, patient characteristics, etc.), the interpretation of neural findings in the included studies is often further complicated because the observed local or network level findings partially overlap with those reported in the much more extensive literature on neural underpinnings of cognition, depression and fatigue in MS[3,55]. Moreover, only four imaging studies provided additional information about these potentially conflicting variables, which were also not directly related to the imaging results. Therefore, results of individual studies need to be interpreted with caution.

In sum, the present study demonstrates substantial impairment of socio-cognitive processes in MS and highlights the potential mediating role of comorbid clinical symptoms. We identify several current evidence gaps and larger scale studies using comprehensive and coordinated assessments of socio-cognitive parameters (e.g., similar to current efforts for establishing core outcome parameters for clinical trials, https://www.comet-initiative.org/), potential mediators and neural correlates are urgently needed.

Methods

The present systematic review and meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline[16]. The pre-registered review protocol can be assessed at www.crd.york.ac.uk/PROSPERO/ (ID: CRD42020206225).

Systematic search, study selection and eligibility criteria

A systematic electronic search was conducted in MEDLINE Ovid, Web of Science Core Collection, CENTRAL, and PsycInfo up to 31st August 2020 with the following keywords: multiple sclerosis, theory of mind, mind reading, social cognition, social cognitive deficits, emotional expression, facial emotion, empathy, social decision making. An update search was conducted on the 15th December 2022. Our search string for MEDLINE Ovid is provided as an example in Supplementary Table 1.

Three review authors (MR, LG, LH) screened all obtained titles and abstracts according to pre-defined criteria using the Covidence Software (https://www.covidence.org/). Full-texts were again screened for studies meeting the inclusion criteria. Disagreements between the reviewers were solved by discussion.

We included studies that investigated social cognition in male and female patients ≥ 18 years old with multiple sclerosis diagnosis (all diagnostic types) compared to a healthy control group. We defined ToM as our primary outcome, because ToM is a key aspect of social cognition, adequate ToM performance is critical for establishing proper social interaction and also relevant for coping with chronic conditions such as MS[56]. ToM is defined as the ability to attribute mental states to others or the ability to understand and predict others' behaviour based on their mental states and is the most frequently studied socio-cognitive process across development and in healthy and pathological aging. Please note, separate meta-analyses were calculated for cToM and aToM to reduce heterogeneity and because both are supported by partially different neural networks[57]. Secondary outcomes were chosen to represent two additional major socio-cognitive domains: social perception (recognizing others as "living persons" via the analysis of perceptual information including e.g. visual perspective taking), and social decision-making (using the obtained social information for social decision making)[15]. All socio-cognitive outcomes needed to be tested with standardized tests to be included in our review. If more than one assessment was conducted, only the first timepoint was considered. Studies specifically assessing empathy were not considered because of the highly heterogeneous nature of this concept (e.g., different aspects of empathy are associated with different neural networks) and overlap with emotion processing and ToM[58].

Data extraction

Three review authors [MR, LG, LH] extracted the data using a study specific, standardized data extraction sheet. Disagreements were discussed with all authors until consensus was reached. We contacted $n = 13$ authors for missing data. Only four replies were received, two authors provided data[1,59].

Quality assessment

We assessed risk of bias (RoB) for each included study using the first six signaling questions of the "Tool for Assessing Risk of Bias in Cohort Studies" by the CLARITY Group[60]. Signaling questions can either be answered with "definitely yes" (low RoB), "probably yes", "probably no", "definitely no" (high RoB). Note, that three signaling questions were not applicable to our research question and studies. Two review authors (LG, LH) individually assessed RoB for each study. If no consensus could be reached, a third author (MR) was involved.

Meta-analyses

We conducted random-effects pairwise meta-analyses to investigate the degree of SC impairment in patients with MS relative to healthy controls. Data was clustered according to our four outcomes aToM, cToM, SCD and VPT. For each outcome, we also clustered studies according to the tests that were used for assessments (e.g., aToM: Baron-Cohen's Adult Eyes Test, Emotion Recognition Florida Affective Battery, and the Facial Expression of Emotion Test, FEEST). Meta-analyses were only calculated if $n \geq 3$ studies were available.

Data analysis was conducted using R. For all analyses, the alpha level was set at 0.05. Standardized mean differences (SMD) were used as effect sizes, because constructs (e.g., ToM) were assessed with different tests. The mean score of the dependent variable, the mean standard deviation, and the number of included participants in each group were used to calculate SMD.

To address heterogeneity, we used the I^2 statistic. As recommended in the Cochrane Handbook for systematic reviews of interventions[61], heterogeneity was interpreted as: 0–40%: not important/low heterogeneity; 30–60%: moderate heterogeneity; 50–90%: substantial heterogeneity; 75–100%: considerable heterogeneity. A funnel plot for identifying possible publication bias was calculated in analyses including ≥ 10 studies. Sensitivity analysis were calculated using fixed effect models to control for small-study effects. If the effect estimates of both, the fixed and random effects model are similar, then any small-study effects have little effect on the effect estimate.

To further assess the impact of cognitive status, depressive symptoms, and fatigue on socio-cognitive abilities in patients with MS, meta-regression analyses were conducted using aToM and cToM as outcome variables and cognitive scores (measured via the neuropsychological test that was most frequently reported in the included studies) and depressive symptoms and fatigue (both measured with standardized questionnaires) as possible predictors. Meta-regressions on SDM and VPT could not be conducted as there was not enough data reported in the studies (Note: this analysis requires correlations between the investigated outcome variable and all possible predictors that are included in the model).

Data availability

All data generated or analyzed in this study are included in the published article [and its supplementary information files]. Aggregated data can be shared by the corresponding author on reasonable request.

Code availability

Code can be shared by the corresponding author on reasonable request.

References

1. Thompson, A. J., Baranzini, S. E., Geurts, J., Hemmer, B. & Ciccarelli, O. Multiple sclerosis. *Lancet* 391(10130), 1622–1636. https://doi.org/10.1016/s0140-6736(18)30481-1 (2018).
2. Benedict, R. H. B., Amato, M. P., DeLuca, J. & Geurts, J. J. G. Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues. *Lancet Neurol.* 19(10), 860–871. https://doi.org/10.1016/S1474-4422(20)30277-5 (2020).
3. Penner, I.-K. & Paul, F. Fatigue as a symptom or comorbidity of neurological diseases. *Nat. Rev. Neurol.* 13, 662–675 (2017).
4. Patten, S. B. *et al.* Major depression in multiple sclerosis: A population-based perspective. *Neurology* 61, 1524–1527 (2003).
5. Brüne, M. & Brüne-Cohrs, U. Theory of mind–evolution, ontogeny, brain mechanisms and psychopathology. *Neurosci. Biobehav. Rev.* 30, 437–455 (2006).
6. Berneiser, J. *et al.* Impaired recognition of emotional facial expressions in patients with multiple sclerosis. *Mult. Scler. Relat. Disord.* 3, 482–488 (2014).
7. Banati, M. *et al.* Social cognition and Theory of Mind in patients with relapsing-remitting multiple sclerosis. *Eur. J. Neurol.* 17, 426–433 (2010).
8. Benedict, R. H. B. *et al.* Cognitive impairment in multiple sclerosis: Clinical management, MRI, and therapeutic avenues. *Lancet Neurol.* 19, 860–871 (2020).
9. Cotter, J. *et al.* Social cognition in multiple sclerosis: A systematic review and meta-analysis. *Neurology* 87, 1727–1736 (2016).
10. Bora, E. *et al.* Social cognition in multiple sclerosis: A meta-analysis. *Neuropsychol. Rev.* 26, 160–172 (2016).
11. Lin, X. *et al.* Social cognition in multiple sclerosis and its subtypes: A meta-analysis. *Mult. Scler. Relat. Disord.* 52, 102973 (2021).
12. Penner, I.-K. Evaluation of cognition and fatigue in multiple sclerosis: daily practice and future directions. *Acta Neurol. Scand.* 134(Suppl 200), 19–23 (2016).
13. Palotai, M. & Guttmann, C. R. Brain anatomical correlates of fatigue in multiple sclerosis. *Mult. Scler.* 26, 751–764 (2020).
14. Masuccio, F. G. *et al.* Imaging and depression in multiple sclerosis: A historical perspective. *Neurol. Sci.* 42, 835–845 (2021).
15. Arioli, M., Crespi, C. & Canessa, N. Social cognition through the lens of cognitive and clinical neuroscience. *Biomed. Res. Int.* 2018, 4283427 (2018).
16. Moher, D. *et al.* Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ* 339, b2535 (2009).
17. Henry, A. *et al.* Social cognition and executive functioning in multiple sclerosis: A cluster-analytic approach. *J. Neuropsychol.* 16, 97–115 (2022).
18. Weinstein, K. M., Patterson & Stephen, M. R. *Hemispheric Asymmetries and Processing of Affective Stimuli: Contribution of Callosal Communication* (1996).
19. Harbo, H. F., Gold, R. & Tintoré, M. Sex and gender issues in multiple sclerosis. *Ther. Adv. Neurol. Disord.* 6, 237–248 (2013).
20. Prochnow, D. *et al.* Alexithymia and impaired facial affect recognition in multiple sclerosis. *J. Neurol.* 258, 1683–1688 (2011).
21. Gleichgerrcht, E., Tomashitis, B. & Sinay, V. The relationship between alexithymia, empathy and moral judgment in patients with multiple sclerosis. *Eur. J. Neurol.* 22, 1295–1303 (2015).
22. Czekóová, K. *et al.* Impaired self-other distinction and subcortical gray-matter alterations characterize socio-cognitive disturbances in multiple sclerosis. *Front. Neurol.* 10, 525 (2019).
23. Bruno, D., Galiani, A., Golfeder, M., *et al.* Perspective taking deficits and their relationship with theory of mind abilities in patients with relapsing-remitting multiple sclerosis (RRMS). *Appl. Neuropsychol. Adult* 1–11 (2022).
24. Ehrlé, N. *et al.* Social norms in patients with relapsing-remitting multiple sclerosis: Impairment of the moral/conventional distinction?. *Soc. Neurosci.* 15, 630–640 (2020).
25. Henry, A. *et al.* Facial emotion recognition and Faux Pas interpretation in multiple sclerosis. *Brain Impairment* 16, 158–172 (2015).
26. Koubiyr, I. *et al.* Amygdala network reorganization mediates the theory of mind performances in multiple sclerosis. *J. Neurosc.i Res.* 100, 537–550 (2022).
27. Henry, A. *et al.* Social cognition abilities in patients with different multiple sclerosis subtypes. *J. Int. Neuropsychol. Soc.* 23, 653–664 (2017).
28. Batista, S. *et al.* Theory of mind and executive functions are dissociated in multiple sclerosis. *Arch. Clin. Neuropsychol.* 33, 541–551 (2018).
29. Goitia, B. *et al.* The relationship between executive functions and fluid intelligence in multiple sclerosis. *PLoS One* 15, e0231868 (2020).
30. Labbe, T. P. *et al.* Social cognition in Multiple Sclerosis is associated to changes in brain connectivity: A resting-state fMRI study. *Mult. Scler. Relat. Disord.* 45, 102333 (2020).
31. Yap, S. M., Davenport, L., Cogley, C., *et al.* Word finding, prosody and social cognition in multiple sclerosis. *J. Neuropsychol.* (2022).
32. Cecchetto, C. *et al.* Facial and bodily emotion recognition in multiple sclerosis: The role of alexithymia and other characteristics of the disease. *J. Int. Neuropsychol. Soc.* 20, 1004–1014 (2014).
33. Pfaff, L. *et al.* Emotional experience is increased and emotion recognition decreased in multiple sclerosis. *Sci. Rep.* 11, 21885 (2021).
34. Montembeault, M. *et al.* Multimodal emotion perception in young and elderly patients with multiple sclerosis. *Mult. Scler. Relat. Disord.* 58, 103478 (2022).
35. Baron-Cohen, S., O'Riordan, M., Jones, R., *et al.* *A New Test of Social Sensitivity: Detection of Faux Pas in Normal Children and Children with Asperger Syndrome* (1999).
36. Stone, V. E., Baron-Cohen, S. & Knight, R. T. Frontal lobe contributions to theory of mind. *J. Cogn. Neurosci.* 10, 640–656 (1998).
37. Henry, A. *et al.* Social cognition impairments in relapsing-remitting multiple sclerosis. *J. Int. Neuropsychol. Soc.* 17, 1122–1131 (2011).
38. Ouellet, J. *et al.* Assessment of social cognition in patients with multiple sclerosis. *J. Int. Neuropsychol. Soc.* 16, 287–296 (2010).
39. Kraemer, M. *et al.* Theory of mind and empathy in patients at an early stage of relapsing remitting multiple sclerosis. *Clin. Neurol. Neurosurg.* 115, 1016–1022 (2013).
40. Golde, S. *et al.* Distinct functional connectivity signatures of impaired social cognition in multiple sclerosis. *Front. Neurol.* 11, 507 (2020).
41. Batista, S. *et al.* Disconnection as a mechanism for social cognition impairment in multiple sclerosis. *Neurology* 89, 38–45 (2017).

42. Adamaszek, M. *et al.* Clinical and neurophysiological patterns of impairments to emotion attention and empathy in multiple sclerosis. *J. Integr. Neurosci.* **21**, 7 (2022).
43. Batista, S. *et al.* Impairment of social cognition in multiple sclerosis: Amygdala atrophy is the main predictor. *Mult. Scler.* **23**, 1358–1366 (2017).
44. Mike, A. *et al.* Disconnection mechanism and regional cortical atrophy contribute to impaired processing of facial expressions and theory of mind in multiple sclerosis: A structural MRI study. *PLoS One* **8**, e82422 (2013).
45. Labbe, T. P. *et al.* Regional brain atrophy is related to social cognition impairment in multiple sclerosis. *Arq. Neuropsiquiatr.* **79**, 666–675 (2021).
46. Bisecco, A. *et al.* Resting-state functional correlates of social cognition in multiple sclerosis: An explorative study. *Front. Behav. Neurosci.* **13**, 276 (2019).
47. Yokote, H., Okano, K. & Toru, S. Theory of mind and its neuroanatomical correlates in people with multiple sclerosis. *Mult. Scler. Relat. Disord.* **55**, 103156 (2021).
48. Isernia, S. *et al.* Theory of mind network in multiple sclerosis: A double disconnection mechanism. *Soc. Neurosci.* **15**, 544–557 (2020).
49. Amdurer, E. *et al.* Long term impact of emotional, social and cognitive intelligence competencies and GMAT on career and life satisfaction and career success. *Front. Psychol.* **5**, 1447 (2014).
50. Eramudugolla, R. *et al.* Social cognition and social functioning in MCI and dementia in an epidemiological sample. *J. Int. Neuropsychol. Soc.* **28**, 661–672 (2022).
51. Phillips, L. H. *et al.* Difficulties with emotion regulation in multiple sclerosis: Links to executive function, mood, and quality of life. *J. Clin. Exp. Neuropsychol.* **36**, 831–842 (2014).
52. Grothe, M. *et al.* Testing social cognition in multiple sclerosis: Difference between emotion recognition and theory of mind and its influence on quality of life. *Brain Behav.* **11**, e01925 (2021).
53. Fernández, O. *et al.* Patient characteristics and determinants of quality of life in an international population with multiple sclerosis: Assessment using the MusiQoL and SF-36 questionnaires. *Mult. Scler.* **17**, 1238–1249 (2011).
54. Frith, C. D. The social brain?. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 671–678 (2007).
55. Sumowski, J. F. *et al.* Cognition in multiple sclerosis: State of the field and priorities for the future. *Neurology* **90**, 278–288 (2018).
56. Chalah, M. A. *et al.* Theory of mind in multiple sclerosis: A neuropsychological and MRI study. *Neurosci. Lett.* **658**, 108–113 (2017).
57. Molenberghs, P. *et al.* Understanding the minds of others: A neuroimaging meta-analysis. *Neurosci. Biobehav. Rev.* **65**, 276–291 (2016).
58. Walter, H. Social cognitive neuroscience of empathy: Concepts, circuits, and genes. *Emotion Rev.* **4**, 9–17 (2012).
59. Pitteri, M. *et al.* Social cognition deficits and the role of amygdala in relapsing remitting multiple sclerosis patients without cognitive impairment. *Mult. Scler. Relat. Disord.* **29**, 118–123 (2019).
60. Higgins, J. P. T. & Green, S. (eds) *Cochrane Handbook of Systematic Reviews of Interventions* 1st edn. (Wiley, 2008).
61. Higgins, J., Thomas, J., Chandler, J. *et al.* (eds) *Cochrane Handbook for Systematic Reviews of Interventions* 2nd edn. (Wiley, 2019).

Author contributions

M.R., M.G., and M.M. were responsible for study conception and design. L.G., L.H. and M.R. did the systematic search and screening, M.R. is responsible for data analysis, M.R., M.R. and M.M. are responsible for interpretation of the results and drafting of the manuscript. All authors reviewed the results and read and approved the final version of the manuscript.

Funding

Competing interests

MG is author of one of the included studies, however, he was not involved in the data extraction and quality rating of this study. Other than that, the authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-53750-5.

Correspondence and requests for materials should be addressed to M.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.