Title

The role of serum neurofilament light (sNfL) as a biomarker in multiple sclerosis: insights from a systematic review

Journal

Journal of Neurology

Authors

Mark S. Freedman*[1], Ahmed Abdelhak[2], Mohit K. Bhutani[3], Jason Freeman[4], Sharmilee Gnanapavan[5], Salman Hussain[3], Sheshank Madiraju[3], Friedemann Paul[6,7]

1. University of Ottawa, Department of Medicine and The Ottawa Hospital Research Institute, Ottawa, ON, Canada
2. Weill Institute for Neurosciences, Department of Neurology, University of California San Francisco, San Francisco, CA, USA
3. Novartis Healthcare Pvt. Ltd., Hyderabad, India
4. Novartis Pharmaceuticals, East Hanover, New Jersey, USA
5. Department of Neurology, Barts Health NHS Trust, London, England, UK
6. NeuroCure Clinical Research Center and Clinical and Experimental Multiple Sclerosis Research Center, Department of Neurology, Charité University Medicine, Berlin, Germany
7. Experimental and Clinical Research Center, Max Delbrück Center for Molecular Medicine and Charité University Medicine Berlin, Berlin, Germany

Corresponding author*

Mark S. Freedman

Professor of Medicine (Neurology)

University of Ottawa

The Ottawa Hospital Research Institute

Ottawa, ON, Canada

Email: mfreedman@toh.ca

Table S1: Search algorithm (conducted via the ovid.com interface on 14th September 2023)

Facet	#	Search terms	Hits
Disease terms	1	exp multiple sclerosis/	203534
	2	multiple sclerosis.mp.	252745
	3	1 or 2	252747
Outcomes	4	("neurofilament" or "neurofilament light chain" or "nfl" or "neurofilament-light chain" or "sNfL" or "pNfL").mp.	32496
Disease terms + outcomes	5	3 and 4	2980
Deduplication	6	remove duplicates from 5	2116
Disease terms + outcomes; removed duplicates; filters applied; English language and humans	7	limit 6 to human [Limit not valid in CDSR,CCTR; records were retained]	1791
	8	limit 7 to humans [Limit not valid in CDSR,CCTR; records were retained]	1791
	9	limit 8 to English language [Limit not valid in CDSR; records were retained]	1771

Table S2: Study characteristics and quality assessment score

Author year	Centers	Diagnostic criteria	Population	Assay method	Risk of bias tool and score	
RCT and extension, if applicable						
Bar-Or 2023 [21] (APLIOS)	MC	2010 revised McDonald criteria	Majority RMS	ADVIA® Centaur NfL assay (Siemens)	Cochrane RoB 2.0	Low risk in all domains
Bar-Or 2023 [44] (OPERA 1 & II)	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix) using HD-X analyzer		Low risk in all domains
Bar-Or 2023 [44] (ORATORIO)	MC	2005 revised McDonald criteria	PMS	Simoa® NF-light™ Kit (Quanterix) using HD-X analyzer		Low risk in all domains
Calabresi 2021 [80] (ADVANCE)	MC	2005 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)		Low risk in all domains
Calabresi 2021 [80] (AFFIRM)	MC	2001 McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix)		Low risk in all domains
Chow 2023 [35]	SC	2010 revised McDonald and Lublin (2014) criteria	PMS	Simoa assay (Quanterix)		Low risk in all domains
Comabella 2022 [37]	SC	Schumacher criteria	PMS	Simoa assay (Quanterix) using HD-1 analyzer		Low risk in all domains
Cutter 2023 [50]	MC	2001 McDonald criteria	Majority RMS	Simoa Human Neurology 4-Plex A assay (Quanterix) using HD-1 analyzer		Low risk in all domains
Fox 2022 [102]	MC	2017 revised McDonald criteria	Majority RMS	NR		Low risk in all domains
Harris 2021 [55] (RADIANCE)	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer		Low risk in all domains
Harris 2021 [55] (SUNBEAM)	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer		Low risk in all domains
Harris 2022 [74]	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix)		Low risk in all domains
Hauser 2020 [88] (ASCLEPIOS I)	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix)		Low risk in all domains

Author year	Centers	Diagnostic criteria	Population	Assay method	Risk of bias tool and score
Hauser 2020 [88] (ASCLEPIOS II)	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix)	Low risk in all domains
Hauser 2023 [43]	MC	2010 revised McDonald criteria	Majority RMS	Atellica® Immunoassay Analyzer part of Antelleca solution (Siemens)	NA[a]
Kuhle 2019 [41] (FREEDOMS)	MC	2005 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)[b]	Low risk in all domains
Kuhle 2019 [41] (TRANSFORMS)	MC	2005 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)[b]	Low risk in all domains
Kuhle 2020 [39]	MC	Poser criteria	Majority RMS	Simoa assay (Quanterix)	Low risk in all domains
Kuhle 2022 [75]	MC	2005 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)	Low risk in all domains
Kuhle 2022 [76] (OPTIMUM)	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix)	Low risk in all domains
Kuhle 2023 [36]	MC	2010 revised McDonald criteria and 2013 revised Lublin criteria	Majority RMS	Simoa® NF-light™ Kit	Low risk in all domains
Leppert 2022 [96] (EXPAND)	MC	2010 revised McDonald criteria	PMS	Simoa assay	Low risk in all domains
Leppert 2022 [96] (INFORMS)	MC	2005 revised McDonald criteria	PMS	Simoa assay	Low risk in all domains
Ziemssen 2022 [54] (ASCLEPIOS I & II)	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix)	Low risk in all domains
Non-RCT[c]					
Abdelhak 2023 [15] (EPIC)	SC	2001 McDonald criteria	All MS types	Simoa assay (Quanterix) using HD-1 analyzer	Fair
Abdelhak 2023 [15] (SMSC)	MC	2005, 2010 revised McDonald criteria, Poser criteria	All MS types	Simoa assay (Quanterix) using HD-1 analyzer	Fair
Akgün 2021 [78]	SC	NR	Majority RMS	Simoa assay	Poor

Author year	Centers	Diagnostic criteria	Population	Assay method	Risk of bias tool and score
Anderson 2020 [53]	NR	2017 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer	Fair
Barro 2018 [16]	MC	1996 Lublin criteria, 2001, 2005, 2017 revised McDonald criteria	All MS types	Simoa assay (Quanterix)	Fair
Benkert 2022 [51]	MC	2001 McDonald criteria, 2005 and 2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix)	Fair
Bove 2023 [70]	NR	2017 revised McDonald criteria	Majority RMS	NR	Fair
Bridel 2021 [79]	SC	NR	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix)	Fair
Brune 2022 [45]	MC	2010 revised McDonald criteria	All MS types	Simoa assay (Quanterix)	Fair
Bsteh 2020 [85]	SC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using SR-X analyzer	Fair
Chitnis 2018 [94]	SC	2010, 2017 revised McDonald criteria	All MS types	Simoa assay (Quanterix)	Fair
Cohen 2019 [109]	MC	2005 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)	Fair
Dal-Bianco 2021 [46]	SC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix) using SR-X analyzer	Fair
de Flon 2019 [91]	MC	2010 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)	Fair
Delcoigne 2020 [86]	MC	NR	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix)	Fair
Disanto 2021 [68]	SC	2010 revised McDonald criteria	All MS types	Simoa® NF-light™ Kit (Quanterix) using HD-X analyzer	Fair
Fedičová 2023 [58]	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer	Fair
Fernandez 2023 [98]	NR	NR	Majority RMS	NR	Fair

Author year	Centers	Diagnostic criteria	Population	Assay method	Risk of bias tool and score	
Fernández-Velasco 2022 [110]	MC	2017 revised McDonald criteria	PMS	Simoa® NF-light™ Kit (Quanterix)	Modified Downs and Black checklist[d]	Fair
Ferraro 2020 [87]	MC	2010 revised McDonald criteria	All MS types	Simoa assay (Quanterix) using HD-1 analyzer		Poor
Gafson 2019 [92]	SC	2010 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)		Fair
Giarraputo 2021 [48]	NR	2010 revised McDonald criteria	PMS	Simoa Neurology 4-Plex B assay (NF-light) (Quanterix) using SR-X analyzer		Poor
Häring 2020 [42]	MC	2005 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)[b]		Fair
Jakimovski 2020 [49]	SC	2010 revised McDonald criteria	All MS types	Simoa assay		Fair
Lin 2021 [56]	MC	2017 revised McDonald criteria	All MS types	Simoa assay (Quanterix)		Fair
Longbrake 2021 [111]	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix)		Fair
Maltby 2023 [112]	MC	NR	All MS types	NR		Fair
Manouchehrinia 2020 [89]	MC	2001 McDonald criteria, 2005 and 2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix)		Fair
Mao-Draayer 2022 [113]	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Kit (Quanterix)		Fair
Masanneck 2022 [63]	MC	2010 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer		Fair
Mattioli 2020 [90]	NR	2010 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix) using SR-X analyzer		Fair
Novakova 2017 [22]	MC	Revised McDonald criteria	All MS types	Simoa® NF-light™ Kit (Quanterix)		Poor
Gimenez 2023 [71]	NR	NR	Majority RMS	Simoa assay		Poor
Olsson 2021 [114]	SC	2017 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix)		Fair
Paolicelli 2022 [38]	SC	2013 revised Lublin criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using SR-X analyzer		Fair

Author year	Centers	Diagnostic criteria	Population	Assay method	Risk of bias tool and score
Pauwels 2022 [67]	MC	2010 revised McDonald criteria	All MS types	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer	Fair
Schaefer 2023 [107]	MC	NR	All MS types	Single molecule array with HD1 Neurology 4-Plex A Advantage Kit (Quanterix)	Fair
Sehr 2019 [69]	SC	NR	All MS types	Simoa® NF-light™ Kit (Quanterix) using HD-1 analyzer	Fair
Seiberl 2023 [72]	SC	NR	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using SR-X analyzer	Fair
Sejbaek 2019 [93]	MC	2010 revised McDonald criteria	Majority RMS	Simoa assay (Quanterix)	Fair
Sotirchos 2022 [17]	SC	2017 revised McDonald criteria	All MS types	Simoa assay (Quanterix) using HD-1 analyzer	Fair
Sotirchos 2023 [40]	MC	Physician confirmed	All MS types	Atellica® solution platform using acridinium-ester immunoassay (Siemens)	Fair
Stenberg 2022 [77]	SC	NR	Majority RMS	Simoa assay	Poor
Tiu 2022 [62]	SC	2017 revised McDonald criteria	Majority RMS	Simoa assay	Fair
Uher 2021 [81]	MC	2005, 2017 revised McDonald criteria	Majority RMS	Simoa assay	Fair
Uphaus 2021 [104],	SC	2010, 2017 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer	Fair
Vollmer 2021 [83]	MC	2010 revised McDonald criteria	Majority RMS	NR	Good
Walo-Delgado 2021 [84]	SC	NR	Majority RMS	Simoa assay (Quanterix) using SR-X analyzer	Fair
Wessels 2023 [73]	SC	NR	Majority RMS	NR	Poor
Wiendl 2023 [115]	MC	NR	Majority RMS	NR	Fair
Zhou 2022 [47]	SC	2017 revised McDonald criteria	Majority RMS	Simoa® NF-light™ Advantage Kit (Quanterix) using HD-1 analyzer	Poor

[a]This was pooled analysis of similarly designed RCTs. The quality assessment was performed for individual studies to avoid duplication; [b]Assay using the capture mAB 47:3 and the biotinylated detector mAB 2:1 from UmanDiagnostics and transferred onto the Simoa HD-1 instrument (Quanterix); [c]Non-RCTs

included prospective and retrospective studies, case-control studies, open-label extensions of multiple RCTs, and noncomparative clinical studies; [d]Study quality was assessed based on overall score as excellent (26–28); good (20–25); fair (15–19); and poor (≤14).
Abbreviations: mAB, monoclonal antibody; MC, multicenter; MS, multiple sclerosis; NA, not applicable; NR, not reported; PMS, progressive multiple sclerosis; RCT, randomized controlled trial; RMS, relapsing multiple sclerosis; SC, single center.
Note: All references are provided within the manuscript

Table S3: Impact of DMTs on sNfL levels (in pg/mL) across the included studies

Author year	Treatment effect
Relapsing multiple sclerosis (studies with ≥80% RRMS)	
Bar-Or 2023 [44] (OPERA 1 & II)	**Ocrelizumab vs. IFNβ-1a** • Greater reduction in sNfL with ocrelizumab (% reduction in geometric mean) o 12 weeks: −20.7% vs. −13.7% o 24 weeks: −31.7% vs. −20.5%, $p < 0.0001$ o 48 weeks: −39.4% vs. −28.3%, $p < 0.01$ o 72 weeks: −43.0% vs. −27.6% o 96 weeks: −43.7% vs. −30.2%
Bar-Or 2023 [21] (APLIOS)	**Ofatumumab** • 12 weeks: Consistent decline in sNfL levels from baseline
Bove 2023 [70]	**Ofatumumab** • Significant reduction following 24 weeks of treatment o Baseline (mean): 9.39 o 4 weeks (mean): 10.63 o 12 weeks (mean): 9.93 o 24 weeks (mean): 8.72, $p = 0.0305$ vs. baseline o 36 weeks (mean): 8.62, $p = 0.0398$ vs. baseline o 48 weeks (mean): 7.98, $p < 0.0001$ vs. baseline
Cutter 2023 [50]	**IFNβ-1a, glatiramer acetate, and IM IFNβ-1a + glatiramer acetate** • 6 months: Significant decrease in the proportion of patients with sNfL ≥16 (all $p < 0.05$) • 12 and 36 months: Results consistent to 6 months ($p < 0.05$ to $p < 0.001$)
Fedičová 2023 [58]	**DMTs** • Change from baseline at 12 months (median = −10.3%, IQR: −37.4% to 25.0%)
Fernandez 2023 [98]	**DMTs** • Significant decrease in sNfL levels following treatment with DMTs, more sharply with high efficacy drugs **Injectables vs. orals vs. monoclonals** • sNfL Z-score at baseline (mean [SD]): 1.57 vs. 1.28 vs. 1.74 • sNfL Z-score at 1 year (mean [SD]): 1.01 vs. 0.88 vs. 0.87 •
Gimenez 2023 [71]	**Dimethyl fumarate vs. natalizumab** • Baseline (mean [SD]): 9.2 [6.3] vs. 9.0 [4.3] • 6 months (mean [SD]): 8.4 [5.9] vs. 7.9 [4.2] • 12 months (mean [SD]): 11.6 [15.3] vs. 8.4 [2.8]
Hauser 2023 [43], Alvarez 2023 [66]	**Ofatumumab continuation vs. teriflunomide to ofatumumab switching at 24 months** • Reduced sNfL levels were maintained following continuous ofatumumab treatment; further sNfL levels were reduced after switching of treatment from teriflunomide to ofatumumab

Author year	Treatment effect
	○ Baseline (median): 8.26 vs. 10.42 ○ 6 months (geometric mean): 8.31 vs. 9.07, p < 0.001 ○ 24 months (geometric mean): 8.50 vs. 8.23 ○ 48 months (geometric mean): 8.60 vs. 8.38
Seiberl 2023 [72]	**Cladribine** • Baseline (mean [SD]): 24.7 [23.8] • 12 months (mean [SD]): 8.8 [6.2], p = 0.0008
Wiendl 2023 [115]	**Cladribine** • Change from baseline at 12 months (median): −25.22% • Change from baseline at 24 months (median): −23.23%
Wessels 2023 [73]	**Natalizumab vs. ocrelizumab** • Baseline (median): 14.61 vs. 9.45 • 12 months (median): 7.64 vs. 7.80 • 18 months (median): 7.25 vs. 8.50 • 24 months (median): 7.86 vs. 7.50 • 36 months (median): 6.75 vs. 7.80
Benkert 2022 [51]	**Treated vs. untreated** • 12 months: sNfL levels decreased rapidly in treated patients while levels fell marginally in untreated patients **Monoclonal antibodies vs. oral therapies/platform therapies** • Alemtuzumab, natalizumab, ocrelizumab, and rituximab led to a higher decrease in sNfL levels compared with oral therapies (i.e., dimethyl fumarate, fingolimod, siponimod, and teriflunomide) • Longitudinal sNfL Z-scores remained elevated with platform compounds (interferons and glatiramer acetate, p < 0.0001 for the interaction term between treatment category and treatment duration)
Fox 2022 [102]	**Vidofludimus calcium 45 mg vs. vidofludimus calcium 30 mg vs. placebo** • Change from baseline at 24 weeks (median): −20.5% vs. −17.0% vs. 6.5%
Harris 2022 [74]	**Ozanimod vs. placebo** • Baseline (median [IQR]): 11.0 [7.7 to 15.0] vs. 11.7 [8.2 to 16.3] • Change from baseline at 4 weeks (median [IQR]): −7.9% [−18.9 to 14.5] vs. NR • Change from baseline at 24 weeks (median [IQR]): −15.9% [−32.0 to 1.8], p < 0.0001 vs. NR
Kuhle 2022 [52], Kuhle 2023 [36]	**Evobrutinib 75 mg qd/75 mg bid vs. evobrutinib 25 mg qd/placebo** • 6 months: Dose-dependent reduction in sNfL levels during the 24-week double-blind period • 36 months: Reduced levels were maintained • Gd+ T1 activity: Significant reduction in both high (≥11.36 pg/mL) and low (<11.36 pg/mL) sNfL groups (relative reduction = 69.2%, p = 0.0018 for high sNfL and 69.4%, p = 0.0018 for low sNfL) • New or enlarging Gd+ T2 activity: Significant reduction in both high (≥11.36 pg/mL) and low (<11.36 pg/mL) sNfL groups (relative reduction = 54.0%, p = 0.0458 for high sNfL and 73.4%, p = 0.0012 for low sNfL)

Author year	Treatment effect

Kuhle 2022 [75]	**Alemtuzumab vs. IFNβ-1a**
	• sNfL levels were significantly lower following 6 months of treatment with alemtuzumab vs. IFNβ-1a
	  ○ Baseline (median [IQR]): 31.7 [17.1 to 60.4] vs. 31.4 [17.5 to 61.1], p = 0.57
	  ○ 6 months (median [IQR]): 17.2 [9.7 to 24.7] vs. 21.4 [14.4 to 33.9], p < 0.0001
	  ○ 12 months (median [IQR]): 14.2 [8.9 to 22.9] vs. 17.7 [11.9 to 29.2], p = 0.0014
	  ○ 18 months (median [IQR]): 13.2 [8.4 to 18.8] vs. 15.6 [9.5 to 24.7], p = 0.0123
	  ○ 24 months (median [IQR]): 13.2 [8.6 to 19.5] vs. 18.7 [12.6 to 27.7], p < 0.0001
	  ○ 84 months (median [IQR]): 12.7 vs. NA

Kuhle 2022 [76]	**Ponesimod vs. teriflunomide** • Baseline (mean [SD]): 14.9 [15.66] vs. 15.8 [21.17] • 108 weeks (mean [SD]): 8.3 [4.28] vs. 11.4 [7.96]
Masanneck 2022 [63]	**DMTs** • Baseline (mean [SD]): 14.5 (17.5) • 6 months (mean [SD]): 10.3 (7.3) • Second follow-up (12 months after): 8.0 (4.7), p = 0.008 vs. baseline
Mao-Draayer 2022 [113]	**Newly started fingolimod vs. continuous fingolimod** • Change from baseline at 12 months (mean [SD]): −3.73 [11.2] vs. 0.67 [8.39]
Paolicelli 2022 [38]	**Cladribine** • Baseline (mean [SD]): 21.78 [14.75] • 24 weeks (mean [SD]): 13.01 [6.31], p = 0.01
Tiu 2022 [62]	**DMTs** • Baseline (median [range]): 20.5 [3.2 to 208][a] • 3 months (median [range]): 12.7 [2.9 to 49.8][a] • 6 months (median [range]): 10.5 [2.77 to 31.7][a]
Zhou 2022 [47]	**Teriflunomide** • Baseline (median [IQR]): 35.82 [47.50] • 6 months (median [IQR]): 31.78 [22.42] • 12 months (median [IQR]): 24.79 [11.72]

Ziemssen 2022 [54], Alvarez 2023 [66]	**Ofatumumab vs. teriflunomide**
	• Risk difference for 3mCDP among high sNfL and low sNfL: −17.6%, p = 0.468 with ofatumumab and −10.7%, p = 0.589 with teriflunomide
	• Risk difference for 6mCDP among high sNfL and low sNfL: −15.2%, p = 0.571 with ofatumumab and −15.7%, p = 0.491 with teriflunomide
	• Ofatumumab decreased T2 lesion formation vs. teriflunomide, showing relative reductions of 82% and 87% in groups with high and low baseline sNfL levels, respectively
	• Ofatumumab reduced sNfL levels compared with teriflunomide
	  ○ Baseline (median): 9.93 vs. 9.63
	  ○ 3 months (geometric mean): 9.62 vs. 10.38, p < 0.001

Author year	Treatment effect
	○ 12 months (geometric mean): 8.03 vs. 1.0.25, p < 0.001 ○ 24 months (geometric mean): 7.96 vs. 9.97, p < 0.001
Akgün 2021 [78]	**Fingolimod** • 12 months: 35% decrease in baseline sNfL (modeled mean at baseline = 9.8, 95% CI: 7.7 to 12.5 vs. modeled mean at 12 months = 6.4, 95% CI: 5.7 to 7.1) • 24 months: No relevant changes
Bridel 2021 [79]	**Natalizumab** • Baseline (median [IQR]): 14.8 [10.0 to 27.1] • 3 months (median [IQR]): 11.1 [8.4 to 16.0] • 12 months (median [IQR]): 7.9 [5.9 to 11.0] • 24 months (median [IQR]): 7.9 [5.7 to 10.5] • 5.2 years (median [IQR]): 8.9 [5.6 to 11.3]
Calabresi 2021 [100]	**Peginterferon beta-1a vs. placebo** • Change from baseline at 48 weeks (mean): −9.5% vs. 6.8%, p < 0.01
Harris 2021 [55]	**IFNβ-1a vs. ozanimod** **SUNBEAM** • Median percentage change at 12 months: −13.4% vs. −22.8% with ozanimod 0.46 mg (p = 0.0003 vs. IFNβ-1a), and −26.9% with ozanimod 0.92 mg (p < 0.0001 vs. IFNβ-1a) **RADIANCE** • Median percentage change at 24 months: −15.5% vs. −19.7% with ozanimod 0.46 mg (p = 0.0024 vs. IFNβ-1a), and −23.5% with ozanimod 0.92 mg (p = 0.0001 vs. IFNβ-1a)
Longbrake 2021 [111]	**Dimethyl fumarate** • Change from baseline at 96 weeks (mean [SD]): −19% [34]
Olsson 2021 [114]	**DMTs** • Reduction from baseline at 12 months (mean [95% CI]): −31% [−41% to −19%], p < 0.001 ○ **First line DMTs (teriflunomide, dimethyl fumarate, glatiramer acetate, peginterferon beta-1a)** ▪ Mean (95% CI): −18% (−30% to −5%), p = 0.011 ○ **Second line DMTs (fingolimod, natalizumab, ocrelizumab/rituximab, ofatumumab, daclizumab, cladribine)** ▪ Mean (95% CI): −51% (−65% to −31%), p < 0.001
Srpova 2021 [57] Uher 2021 [81]	**IFNβ-1a** • Baseline (median [IQR]): 22.68 [12.62 to 39.89] • 1 month (median [IQR]): 17.70 [10.99 to 31.05] • 12 months (median [IQR]): 13.86 [9.51 to 21.29] • 24 months (median [IQR]): 12.48 [8.61 to 18.00] • 36 months (median [IQR]): 12.24 [8.96 to 16.49]
Vollmer 2021 [83]	**Ocrelizumab** • Baseline (geometric mean): 14.5

Author year	Treatment effect
	• 48 weeks (geometric mean): 6.41
Walo-Delgado 2021 [84]	**Dimethyl fumarate** • Median (95% CI) decrease from baseline at 3 months: 4.0 (2.4 to 5.6), $p < 0.0001$ • Median (95% CI) decrease from baseline from 3 to 12 months: 6.7 (5.5 to 8.3), $p < 0.0001$
Bsteh 2020 [85]	**DMT initiation/escalation** • 3 years (median [IQR]): 3.6 [2.2 to 5.4] decline from prior to post-DMT initiation/escalation
Delcoigne 2020 [86]	• DMT choices were associated with degree of reduction in sNfL levels, supporting the role of sNfL as a drug-response marker • Alemtuzumab resulted in maximum decline in sNfL levels **Alemtuzumab[b]** • Baseline (median [IQR]): 10.5 [6.3 to 24.8] • 4 to 24 weeks (median [IQR]): 6.9 [5.4 to 8.8] **Dimethyl fumarate[b]** • Baseline (median [IQR]): 11.1 [8.2 to 15.6] • 4 to 24 weeks (median [IQR]): 8.3 [6.8 to 10.7] **Fingolimod[b]** • Baseline (median [IQR]): 12.3 [8.7 to 16.9] • 4 to 24 weeks (median [IQR]): 9.6 [7.6 to 11.8] **Natalizumab[b]** • Baseline (median [IQR]): 15.5 [9.9 to 26.9] • 4 to 24 weeks (median [IQR]): 8.7 [7.3 to 11.8] **Rituximab[b]** • Baseline (median [IQR]): 12.3 [9.7 to 18.2] • 4 to 24 weeks (median [IQR]): 9.6 [7.9 to 11.5] **Teriflunomide[b]** • Baseline (median [IQR]): 9.0 [7.0 to 12.2] • 4 to 24 weeks (median [IQR]): 10.0 [7.2 to 13.0]
Häring 2020 [42]	**Fingolimod** • Baseline (geometric mean): 29.7 • 12 months (geometric mean): 17.72 • 24 months (geometric mean): 17.96

Hauser 2020 [88] (ASCLEPIOS I)	**Ofatumumab vs. teriflunomide**
	• Reduction in sNfL levels was higher with ofatumumab compared with teriflunomide
	  ○ Baseline (mean [SD]): 13.3 [13.2] vs. 11.7 [9.3]
	  ○ 3 months (geometric mean [95% CI]): 8.8 [8.5 to 9.1] vs. 9.4 [9.1 to 9.8], $p = 0.01$
	  ○ 12 months (geometric mean [95% CI]): 7.0 [6.7 to 7.3] vs. 9.6 [9.2 to 10.1], $p < 0.001$
	  ○ 24 months (geometric mean [95% CI]): 6.9 [6.6 to 7.2] vs. 9.0 [8.6 to 9.5], $p < 0.001$

Author year	Treatment effect

Hauser 2020 [88] (ASCLEPIOS II)	**Ofatumumab vs. teriflunomide**
	• Reduction in sNfL levels was higher with ofatumumab compared with teriflunomide
	  ○ Baseline (mean [SD]): 14.7 [18.2] vs. 13.4 [14.0]
	  ○ 3 months (geometric mean [95% CI]): 8.9 [8.6 to 9.2] vs. 10.0 [9.7 to 10.4], p < 0.001
	  ○ 12 months (geometric mean [95% CI]): 7.1 [6.8 to 7.4] vs. 9.5 [9.1 to 10.0], p < 0.001
	  ○ 24 months (geometric mean [95% CI]): 6.8 [6.5 to 7.1] vs. 9.0 [8.6 to 9.4], p < 0.001

Mattioli 2020 [90]	**IFNβ-1a** • Baseline (mean): 7.52 • 1 year (mean): 7.18, p = 0.44 vs. baseline
Cohen 2019 [109]	**Fingolimod any dose vs. fingolimod 0.5 mg** • Baseline (geometric mean): 28.97 vs. 32.63 • End of core study (geometric mean): 25.1 vs. 19.55 • 14 years (geometric mean): 17.19 vs. 19.84
de Flon 2019 [91]	**Rituximab** • Baseline (mean [SD]): 9.73 (7.04) • 12 months (mean [SD]): 7.94 (3.36), p = 0.055 • 24 months (mean [SD]): 7.99 (3.36), p = 0.046 • 36 months (mean [SD]): 8.04 (3.12), p = 0.088 • 48 months (mean [SD]): 7.87 (3.67), p = 0.052 • 60 months (mean [SD]): 9.69 (5.01), p = 0.296
Gafson 2019 [92]	**Dimethyl fumarate** • Change from baseline at 15 months: 40% decline (mean [SD] = 7.83 [3.94] vs. 13.2 [18.56] at baseline)
Kuhle 2019 [41]	**Fingolimod vs. placebo** **FREEDOMS** • 6 months: 35.4% (30.6 to 19.6 pg/mL) vs. 9% (29.1 to 26.7 pg/mL) decline • 24 months: 43.0% (31.4 to 18.0 pg/mL) vs. 4% (28.2 to 26.9 pg/mL) decline **Fingolimod vs. IFN-β-1a** **TRANSFORMS** • 6 months: 36% (28.5 to 18.4 pg/mL) vs. 14% (24.8 to 21.5 pg/mL) decline • 12 months: 39% (28.2 to 17.1 pg/mL) vs. 17% (24.9 to 20.7 pg/mL) decline

Sejbaek 2019 [93]	**Dimethyl fumarate vs. placebo**
	• Significant reduction following dimethyl fumarate compared with placebo
	  ○ Baseline (mean [SD]): 16.4 [14.4] vs. 17.5 [14]
	  ○ 12 months (mean [SD]): 7.4 [3.1], p < 0.0001 vs. 16.6 [14.0], p > 0.99

Novakova 2017 [22]	**DMTs** • Baseline (median [range]): 16.9 [1.9 to 420.0][c] • 12 months (median [range]): 12.1 [2.2 to 40.4][c]

Author year	Treatment effect
All multiple sclerosis subtypes (RRMS, SPMS, PPMS, CIS)	
Abdelhak 2023 [15], Canto 2019 [60]	**Treated vs. untreated** • 36 months: High potency therapies[d] resulted in more significant decrease of sNfL levels compared with that in untreated patients (β = 0.922, 95% CI: 0.868 to 0.980, p < 0.01) **High potency drugs[d] vs. untreated and platform therapies** • 60 months: High potency therapies[d] resulted in greater decreases in sNfL levels compared with that in patients who were untreated or received platform therapies[e] (vs. untreated: β = 0.946, 95% CI, 0.915 to 0.976, p < 0.001; vs. platform: β = 0.972, 95% CI, 0.948 to 0.998, p = 0.04)
Sotirchos 2023 [40]	**DMTs** • Active treatment was associated with lower odds of elevated sNfL
Maltby 2023 [112]	**Cladribine** • Mean sNfL Z-score at baseline = 0.58 • Mean sNfL Z-score at 30 months = −0.2, p = 0.003
Moreira Ferreira 2022 [116], Chitnis 2018 [94]	**High-efficacy early DMT vs. lower-efficacy early DMT[a,f]** • Change from baseline at 3 years (mean [SD]): −0.35 [0.83] vs. −0.29 [0.75], p = 0.49
Pauwels 2022 [67]	**DMTs** • sNfL had no association with DMTs
Sehr 2019 [69]	**Fingolimod** • Baseline (mean): 8.42[b] • 4 months (mean): 7.36[b], p = 0.009 • 12 months (mean): 7.37[b] • 24 months (mean): 5.66[b]
Progressive multiple sclerosis (SPMS, PPMS)	

Bar-Or 2023 [44] (ORATORIO)	**Ocrelizumab vs. placebo**
	• Greater reduction in sNfL with ocrelizumab (% reduction in geometric mean)
	  ○ 12 weeks: −12.4% vs. −5.4%
	  ○ 24 weeks: −14.9% vs. −2.5%
	  ○ 48 weeks: −17.6% vs. −1.9%, significant reduction vs. placebo
	  ○ 72 weeks: −16.5% vs. −2.1%, significant reduction vs. placebo
	  ○ 96 weeks: −19.0% vs. −1.9%, significant reduction vs. placebo
	  ○ 120 weeks: −20.2% vs. −6.7%, significant reduction vs. placebo

Chow 2023[35]	**Dimethyl fumarate vs. placebo** • Change from baseline at 48 weeks (mean [95% CI]): −0.15 [−4.4 to −4.1] vs. 0.30 [−1.9 to 2.4][c] **Continued dimethyl fumarate vs. placebo to dimethyl fumarate** • Change from 48–96 weeks (mean [95% CI]): −1.6 [−6.1 to 2.8] vs. −0.17 [−3.1 to 2.8][c]
Comabella 2022 [37]	**IFNβ-1b** • 24 months: Slight decrease in sNfL levels (β = −0.13, 95% CI −0.19 to 0.07, p = 0.02)

Author year	Treatment effect
Fernández-Velasco 2022 [110]	**Ocrelizumab** • Median [IQR] sNfL Z-score at baseline = 0.569 [−0.094 to 1.801] • Median [IQR] sNfL Z-score at 6 months = 0.228 [−0.358 to 1.282]
Leppert 2022 [96]	**Siponimod vs. placebo** Lower sNfL levels in siponimod-treated patients **EXPAND** • 12 months: −10.5%, p = 0.0118 • 24 months: −12.4%, p = 0.0012 • 36 months: −22.4%, p = 0.0071 **INFORMS** • 12 months: −9.1%, p = 0.0494 • 24 months: −18.2%, p ≤ 0.0001 • 36 months: −10.8%, p = 0.0377
Novakova 2017 [22]	**DMTs** • Baseline (median [range]): 23.6 [10.8 to 313][c] • 12 months (median [range]): 22.7 [10.0 to 180.0][c]

[a]Unit not reported; [b]Reported as pg/L and levels were normalized to age 40 years; [c]Reported as ng/L; [d]Natalizumab, rituximab, mitoxantrone, cyclophosphamide, fingolimod, and dimethyl fumarate; [e]IFNβ-1b, IFNβ-1a, and glatiramer acetate, monthly pulsed dose glucocorticoids, azathioprine, mycophenolate mofetil, and teriflunomide; [f]High-efficacy treatments (fingolimod, natalizumab, ocrelizumab, rituximab), lower-efficacy treatments (dimethyl fumarate, glatiramer acetate, interferons, teriflunomide).

Note: sNfL levels reported in pg/mL; pg/mL and ng/L were considered equivalent units as 1 pg/mL = 1 ng/L. Some studies reported use of log normal sNfL levels for analysis; however, no information was reported for other studies. All references are provided within the manuscript.

Abbreviations: 3mCDP, 3-month confirmed disability progression; 6mCDP, 6-month confirmed disability progression; bid, twice daily; CI, confidence interval; CIS, clinically isolated syndrome; DMT, disease-modifying therapy; Gd+, gadolinium-enhancing; IFNβ, interferon beta; IM, intramuscular; IQR, interquartile range; NA, not available; NF, neurofilament; NR, not reported; PPMS, primary progressive multiple sclerosis; qd. once daily; RRMS, relapsing-remitting multiple sclerosis; SD, standard deviation; sNfL, serum neurofilament light chain; SPMS, secondary progressive multiple sclerosis.

Table S4: List of studies reporting association of sNfL and disease-worsening parameters

Author year	Key outcomes relevant to the current SLR
Relapsing multiple sclerosis (studies with ≥80% RRMS)	
Abdelhak 2023 [15], Abdelhak 2022 [64]	**Confirmed disability progression with clinical relapse** **EPIC cohort** • An NfL Z-score >1.0 was associated with a 91% higher risk for diagnosing CDW-R in ~12.6 months (HR = 1.91, 95% CI: 0.94 to 3.87, p = 0.07) **SMSC cohort** • An NfL Z-score >1.0 was associated with a 70% higher risk for diagnosing CDW-R in ~11.0 months (HR = 1.70, 95% CI: 1.10 to 2.61, p = 0.02) **Confirmed disability progression with no clinical relapse** **EPIC cohort** • An NfL Z-score >1.0 was associated with a 40% higher risk for diagnosing CDW-NR in ~12 months (HR = 1.40, 95% CI: 1.06 to 1.85, p = 0.02) **SMSC cohort** • An NfL Z-score >1.0 was associated with a 49% higher risk for diagnosing CDW-NR in ~21 months (HR = 1.49, 95% CI: 1.20 to 1.84, p < 0.001) **PARA/PIRA** • sNfL levels (mean [SD]) at baseline were higher in patients with PARA (26.9 [1.8]) vs. PIRA (22.7 [1.7], $P_{Bonferroni}$ = 0.013) • High baseline sNfL values predicted PARA (HR = 2.3, 95% CI: 1.1 to 5.1, p = 0.037) but not PIRA (HR = 0.97, 95% CI: 0.7 to 1.3)

Bar-Or 2023 [44] (OPERA 1 & II)	**EDSS**
	• Higher baseline sNfL levels were independently associated with higher EDSS scores (effect on $\log_{10}$ sNfL in multiple linear regression model = 0.02, 95% CI, 0.0 to 0.03, p = 0.0414)
	Disease progression
	• High NfL (>10.6 pg/mL) at Week 48 compared with low sNfL was significantly associated with risk for future 24-week CDP in patients receiving ocrelizumab (2 years, p = 0.018, 5 years and 10 years, p < 0.001)
	PIRA
	• Baseline NfL level was associated with PIRA, but only in patients without disease activity. Effect of a 2-fold higher baseline sNfL was seen in:
	  ○ Patients without disease activity: HR = 2.44, 95% CI: 1.58 to 3.76, p < 0.0001
	  ○ All patients: HR = 1.24, 95% CI: 0.91 to 1.69, p = 0.1792
	SDMT
	• Baseline NfL level was not associated with SDMT scores. Effect of 2-fold higher baseline sNfL was seen in:
	  ○ Patients without disease activity: HR = 0.82, 95% CI: 0.54 to 1.24, p = 0.3533
	  ○ All patients: HR = 0.89, 95% CI: 0.70 to 1.13, p = 0.3231
	9-HPT
	• Baseline sNfL level was associated with 9-HPT but only in patients without disease activity (HR = 2.10, 95% CI: 1.24 to 3.53, p = 0.0054)

Author year	Key outcomes relevant to the current SLR
Bar-Or 2023 [21] (APLIOS)	**NEDA-3** • The proportion of patients with NEDA-3 was higher among those with below baseline median sNfL (65.4%) compared with those with above (21.7%) or crossing (50.0%) the baseline sNfL median levels
Fedičová 2023 [58]	**EDSS** • sNfL dynamics were significantly correlated with EDSS score at the follow-up visit ($r = 0.34$, $p < 0.001$) • Patients with annual sNfL increase of >10% had a significantly higher number of patients with EDSS worsening compared with patients who had either any annual decrease or an annual increase of up to 10% in sNfL levels (42.2% vs. 6.3%, $p < 0.001$) **NEDA-3 and NEDA-4** • sNfL dynamic variables correlated with NEDA-3 status (AUC = 0.813, 95% CI: 0.726 to 0.9, $p < 0.001$; sensitivity and specificity were 77% and 74%, respectively) • Lower sNfL dynamics were associated with a higher probability of achieving NEDA-3 status, with a cutoff level for sNfL dynamics of 11% • Receiver operating characteristics analysis showed that a sNfL annual change ≥10% correlated with absence of NEDA-3 status ($p < 0.001$, AUC = 0.92) and absence of NEDA-4 status ($p < 0.001$, AUC = 0.839)
Fernandez 2023 [98]	**NEDA-3** • sNfL changes at Year 1 were predictors of loss of NEDA-3 at Year 2 (univariate analysis, OR = 1.36, $p = 0.012$) • sNfL increase from baseline at Year 1 was a predictor of loss of NEDA-3 at Year 2 (univariate analysis, OR = 2.19, $p = 0.010$, multivariate analysis, OR = 2.84, $p = 0.010$) o The association was more pronounced for the treatment-naive patients; patients with an increase in sNfL at Year 1 had 10 times higher risk of losing NEDA-3 (OR = 11.47, 95% CI: 2.05 to 64.11, $p = 0.005$) o Patients with sNfL Z-score ≥1.5 had 4 times higher risk of losing NEDA-3 at Year 2 (OR = 4.735, 95% CI: 1.16 to 19.35, $p = 0.03$), adjusting for clinical and radiological activity **Disability progression** • Among patients with NEDA-3 at Year 1, any increase in sNfL at Year 1 from baseline increased 5 times the risk of having inflammatory activity and/or disability progression (OR = 5.44, 95% CI: 1.34 to 22.07, $p = 0.02$)
Benkert 2022 [51], Abdelhak 2023 [15] (SMSC cohort)	**EDSS** • Higher sNfL Z-scores were associated with a greater probability of EDSS worsening (OR = 1.11, 95% CI: 1.03 to 1.21, $p = 0.0093$) **Disease activity** • sNfL Z-scores >1.5 were associated with an increased risk of future clinical or MRI disease activity in all patients with MS (OR = 3.15, 95% CI: 2.35 to 4.23, $p < 0.0001$) and in those considered stable with NEDA (OR = 2.66, 95% CI: 1.08 to 6.55, $p = 0.034$) • Increased Z-scores outperformed absolute raw sNfL cutoff levels for diagnostic accuracy **EDA-3** • Higher sNfL Z-scores were associated with a greater probability of EDA-3 (OR = 1.43, 95% CI: 1.31 to 1.57, $p < 0.0001$)
Benkert 2022 [51] (EIMS, IMSE, COMBAT-MS cohorts)	**EDSS** • Higher sNfL Z-scores were associated with a higher probability of EDSS worsening in the following year (OR = 1.12, $p < 0.01$) **EDA** • Higher sNfL Z-scores were associated with a higher probability of EDA-3 in the following year (OR = 1.33, $p < 0.001$)

Author year	Key outcomes relevant to the current SLR
	• An incremental increase in the risk of EDA-3 in the following year was observed with increasing sNfL Z-score cutoffs with an up to 2.1-fold risk in patients with sNfL above vs. below the 97.7th percentile (Z-score >2.0) • Patients with NEDA-3 with sNfL levels above the 93.3rd percentile (Z-score >1.50) had a 2.64-fold (95% CI: 1.30 to 5.37, p = 0.0074) higher risk of experiencing EDA-3 in the following year
Brune 2022 [45]	**Disease worsening** • High sNfL concentrations (≥8 pg/mL) at baseline were associated with an increased risk of disease worsening at the median 2-year follow-up (OR = 2.8, 95% CI: 1.5 to 5.3, p = 0.001) **9-HPT and T25FWT** • Higher sNfL concentrations were significantly associated with slower performance on both 9-HPT (r_p = 0.24, p = 0.003) and T25FWT (r_p = 0.31, p < 0.001) at follow-up
Masanneck 2022 [63]	**Disease activity** • Neither NfL levels at baseline or at one of the first two follow-ups nor a change in NfL levels over time showed any significant correlation with the occurrence of loss of NEDA-3, PIRA, and EDSS progression **RAW** • sNfL levels had weak significant association with RAW (coefficient = 0.03, 95% CI: 0.01 to 0.05, p = 0.01)
Pauwels 2022 [67]	**Disease worsening** • Median levels of sNfL were higher in patients with vs. without EDSS-Plus worsening; however, it did not reach significance (p = 0.11)

Tiu 2022 [62]	**MoCA**
	• Moderate negative correlation was found between raw baseline sNfL levels and 1-year follow-up MoCA scores (r = −0.33, p = 0.019), 3- months (r = −0.32, p = 0.021) and 6-month follow-up (r = −0.42, p > 0.001), as well as sNfL Z-scores at 3 months follow-up (r = −0.32, p = 0.022)
	SDMT
	• Moderate negative correlation was found between 6-month follow-up raw sNfL levels and 1-year follow-up oral SDMT scores (r = −0.36, p = 0.01)
	BVMT-R
	• Weak-to-moderate negative correlation was found between 6-month follow-up raw sNfL levels and BVMT-R test scores
	  ○ BVMT-R total score T1–T3 1-year follow-up: r = −0.289, p = 0.042
	  ○ BVMT-R DR 1-year follow-up: r = −0.286, p = 0.049

Ziemssen 2022 [54]	**CDP** • Patients with high and low sNfL did not differ in the risk of 3mCDP or 6mCDP
Akgün 2021 [78]	**EDSS** • Depending on the EDSS score, sNfL levels were higher in patients with EDSS score >5 **SDMT** • Baseline sNfL levels were negatively correlated with SDMT (r = −0.218, p < 0.05)
Bridel 2021 [79]	**EDSS** sNfL levels, both at baseline or Year 1, did not predict EDSS or EDSS-Plus progression at the final follow-up visit

Author year	Key outcomes relevant to the current SLR

Calabresi 2021 [100]	**EDSS**
	• A decline in sNfL was associated with a 4-year change in EDSS
	  ○ No sNfL decrease (levels remained ≥16 pg/mL) vs. sNfL decrease (levels decreased to ≤16 pg/mL); EDSS association was evident at 9 months (least square means [95% CI]):
	  ▪ 3 months: 0.321 (−0.149 to 0.791), p = 0.179
	  ▪ 6 months: 0.237 (−0.271 to 0.746), p = 0.357
	  ▪ 9 months: 0.530 (0.019 to 1.041), p = 0.042
	  ▪ 12 months: 0.513 (0.072 to 0.954), p = 0.023

Dal-Bianco 2021 [46]	**SDMT** • High sNfL levels were associated with lower SDMT Z-score (Spearman's rank correlation coefficient = 0.531, p < 0.004)
Harris 2021 [55], Harris 2022 [117]	**NEDA** **SUNBEAM/RADIANCE** • Greater sNfL reduction was associated with NEDA **SUNBEAM** **SDMT** • Baseline sNfL levels and SDMT score had a slightly negative association (Kendall's correlation = −0.10, 95% CI: −0.14 to −0.06) • Higher median percentage reduction in sNfL concentration was associated with higher 12-month mean change from baseline in SDMT score
Uher 2021 [81], Srpova 2021 [57], Friedova 2020 [118]	**EDSS** • High sNfL levels were associated with higher odds of having EDSS worsening in the following year (8.0% vs. 2.8%; β_{OR} = 3.70, 95% CI: 1.09 to 12.60, p = 0.036) • sNfL showed a weak association with baseline EDSS (rho = 0.21, p = 0.01) • In a repeated-measures analysis, EDSS score was not associated with percentage changes in sNfL **Clinical disease activity** • High sNfL levels were associated with higher odds of clinical disease activity (absence of relapse and/or disease worsening) compared with low sNfL levels in the following year (45.3% vs. 26.2%; β_{OR} = 2.51, 95% CI: 1.29 to 4.90, p = 0.007) **Patients with EDA-3** • Higher sNfL levels were associated with higher odds of EDA-3 in the following year compared with low sNfL levels (86.5% vs. 57.9%; β_{OR} = 4.25, 95% CI: 2.02 to 8.95, p = 0.0001) **Patients with NEDA-3** • High sNfL levels were associated with a higher frequency of clinical disease activity (absence of relapse and/or disease worsening) compared with low sNfL levels (21.4% vs. 13.3%; β_{OR} = 1.72, 95% CI: 0.42 to 7.09, p = 0.45) • Patients with higher sNfL showed numerically higher disease activity (EDA-3) in the following year compared with those with low sNfL (57.1% vs. 31.1%; β_{OR} = 2.55 95% CI: −0.78 to 8.39, p = 0.12) • Patients with loss of NEDA-3 status within 36 months showed higher sNfL levels over follow-up among those with active MS **CVLT-II** • Higher sNfL levels were not associated with an increased risk of cognitive decline; however there was a trend for a greater risk of CVLT-II decline in patients with higher 1-year sNfL levels (OR = 15.8, 95% CI: 1.7 to 147.0, unadjusted p = 0.015)

Author year	Key outcomes relevant to the current SLR
	PASAT • An association was observed between elevated sNfL levels at 2 years and a decline in PASAT-3 scores by Year 9 (OR = 3.9, 95% CI: 0.8 to 19.0, p = 0.091, q = 0.198)
Uphaus 2021 [104], Steffen 2023 [105]	**Disease progression** • In patients with RFP compared with no RFP, sNfL levels were elevated at both baseline (median = 10.8, IQR: 7.7 to 15.0 pg/mL vs. median = 7.2, IQR: 4.5 to 12.5 pg/mL, p = 0.017) and 6-year follow-up (10.0 [6.4 to 13.2] pg/mL vs. 6.9 [5.1 to 9.1] pg/mL, p = 0.008) • In a multivariable logistic regression model, increased sNfL levels at baseline (OR = 1.02, 95% CI: 1.01 to 1.04, p = 0.012) remained an independent risk factor for RFP and predicted individual RFP risk with an accuracy of 82% as revealed by support vector machine • Patients with sNfL ≥7.3 pg/mL showed an increased risk of RFP at follow-up in the time to event analysis (log-rank p = 0.0135) • Similarly, Cox regression analysis revealed a 190% increased risk of experiencing RFP in these patients (HR = 2.90, 95% CI: 1.19 to 7.03, p = 0.009) • The sNfL follow-up/baseline ratio was increased in SPMS-converters (1.16 [0.89 to 1.70] vs. 0.96 [0.75 to 1.23], p = 0.011). This was confirmed by a multivariable logistic regression model, as the sNfL follow-up/baseline ratio remained in the model (OR = 1.476, 95% CI: 1.078 to 2.019, p = 0.015) and individual sNfL follow-up/baseline ratios showed a predictive accuracy of 72% as revealed by machine learning **EDSS** • A significant correlation between sNfL and Year 0 EDSS (r = 0.104, p = 0.199) and Year 6 EDSS (r = 0.053, p = 0.5131) and EDSS change over time (r = −0.024, p = 0.769) was lacking **NEDA-3** (including development of new persistent T1 hypointense lesions) • After multivariable correction, decreased sNfL levels (OR = 0.883, 95% CI: 0.819 to 0.952, p = 0.001) were associated with NEDA-3 T1 status at Year 6 **EDA** • Patients with sNfL levels ≤8.6 pg/mL showed a 76% risk reduction for EDA and development of new T1 hypointense lesions at Year 6 (HR = 0.244, 95% CI: 0.142 to 0.419, p < 0.001) • Median time until EDA T1 was reduced from 93 months in patients with Year 0 sNfL ≤8.6 pg/mL (95% CI: 81 to 103) to 78 months (95% CI: 68 to 86) in patients with sNfL >8.6 pg/mL
Walo-Delgado 2021 [84]	**NEDA** • Multivariate analysis results reported that low baseline sNfL levels (≤12 pg/mL) increased the probability of NEDA at 12 months (OR = 5.8, CI: 1.82 to 15.6, p = 0.002) • Multivariate analysis adjusted for the presence of baseline Gd+ lesions, the number of Gd+ lesions, and NEDA status in the previous year confirmed that sNfL ≤7 pg/mL at 3 months remained a significant predictor of NEDA status at 12 months (OR = 4.8, 1.6 to 14.2, p = 0.005)
Anderson 2020 [53]	**EDSS** • Log-transformed sNfL concentration at diagnosis was modestly associated with baseline EDSS score (β = 0.272, 95% CI: 0.051 to 0.494, p = 0.016)

Author year	Key outcomes relevant to the current SLR
	• However, no significant association was found between baseline sNfL and 5-year EDSS change (β = −0.180, 95% CI: −0.436 to 0.076, p = 0.167) nor when patients were categorized according to whether baseline sNfL was </>13.7 pg/mL (β = −0.26, 95% CI: −0.87 to 0.34, p = 0.389)
Delcoigne 2020 [86], Piehl 2018 [23]	**EDSS** • In the univariate analysis, baseline NfL was associated with EDSS at baseline (β = 0.032, SE = 0.015, p = 0.032); however, the association was not significant following multivariate analysis (β = 0.014, SE = 0.017, p = 0.40) **SDMT** • Baseline sNfL levels were negatively associated with SDMT score (log sNfL and SDMT $\beta_{univariate} = 0.989$, $p \leq 0.001$; $\beta_{multivariate} = 0.991$, $p \leq 0.001$)
Häring 2020 [42]	**EDSS worsening** • A single high baseline sNfL (≥30 pg/mL) compared with low baseline sNfL (<30 pg/mL) had a 2-fold increase in the hazard of reaching EDSS ≥4.0 (HR = 2.19, 95% CI: 1.21 to 3.97, p = 0.0098) o Predictive value increased over 12 months and 24 months when geometric mean of NfL_{long} was measured (12 months, HR = 2.78, 95% CI: 1.51 to 5.10, p = 0.0010; 24 months, HR = 7.91, 95% CI: 2.99 to 20.92, p < 0.0001) **6mCDP** • High baseline sNfL (≥30 pg/mL) compared with low baseline sNfL (<30 pg/mL) was not predictive of 6mCDP (HR = 1.54, 95% CI: 0.91 to 2.61, p = 0.1059 o sNfL levels were associated with accelerated 6mCDP only at 24 months when geometric mean of NfL_{long} was measured over 12 months and 24 months (12 months, HR = 1.53, 95% CI: 0.89 to 2.62, p = 0.1217; 24 months, HR = 3.14, 95% CI: 1.38 to 7.11, p = 0.0061) **T25FWT** • High baseline sNfL (≥30 pg/mL) compared with low baseline sNfL (<30 pg/mL) was not predictive of 20% worsening in the T25FWT (HR = 1.06, 95% CI: 0.67 to 1.68, p = 0.7988) o sNfL levels were associated with 20% worsening in the T25FWT only at 24 months when geometric mean of NfL_{long} was measured over 12 months and 24 months (12 months, HR = 1.10, 95% CI: 0.65 to 1.84, p = 0.7269; 24 months, HR = 3.05, 95% CI: 1.38 to 6.70, p = 0.0056) **PASAT** • A single high sNfL at baseline compared with low sNfL were not predictive of 20% worsening in the PASAT (HR = 1.48, 95% CI: 1.64 to 3.39, p = 0.3539) o sNfL levels were associated with 20% worsening in the PASAT only at 12 months when geometric mean of NfL_{long} was measured over 12 months and 24 months (12 months, HR = 2.59, 95% CI: 1.04 to 6.47, p = 0.0410; 24 months, HR = 3.03, 95% CI: 0.72 to 12.69, p = 0.1300) **9-HPT** • A single high baseline sNfL, compared with low sNfL, did not predict a 20% decline in the PASAT at any time point
Kuhle 2020 [39]	**EDSS** • 3-year and 4-year sNfL levels were associated with changes in EDSS score at Year 8 (r = 0.27, p < 0.05 and r = 0.26, p < 0.05, respectively); 4-year sNfL levels was also associated with EDSS score changes at Year 15 (r = 0.3, p < 0.05)

Author year	Key outcomes relevant to the current SLR
	• Risk of reaching an EDSS score of 6.0 after 8 years of follow-up was significantly increased in patients in the upper sNfL tertile compared with the lowest tertile (3-year, OR = 11.0, 95% CI: 2.0 to 114.6, $p < 0.01$; 4-year, OR = 7.3, 95% CI: 2.0 to 33.3, $p < 0.01$) • Risk of reaching an EDSS score of 6.0 after 15 years of follow-up was significantly only for 4-year sNfL levels (OR = 4.9, 95% CI: 1.4 to 20.4, $p < 0.05$)
Manouchehrinia 2020 [89]	**EDSS** • High sNfL was associated with increased adjusted rates of EDSS worsening ranging between 1.4 (95% CI: 1.1 to 1.8) and 1.7 (95% CI: 1.4 to 2.3) • High sNfL was associated with the risk of reaching a sustained EDSS score of 3.0, with adjusted rates ranging between 1.5 (95% CI: 1.2 to 1.8) and 1.55 (95% CI: 1.3 to 1.8) over all percentile cutoffs (all $p < 0.001$). Similar increases were observed for the risk of sustained EDSS score 4.0 • Risk of reaching sustained EDSS score 6.0 and conversion to SPMS was not consistently significant
Kuhle 2019 [41], Sormani 2019 [106]	**Disease worsening** **FREEDOMS** • NfL at 6 months correlated with the cumulative risk of 6mCDP (HR = 1.83, $p = 0.012$) • High (>60 pg/mL) vs. low (<30 pg/mL) baseline NfL levels were associated with 1.9 times higher risk of 3mCDP (HR = 1.94, 95% CI: 0.97 to 3.87, $p = 0.0605$)
Chitnis 2018 [94], Bose 2023 [59], Galetta 2021 [95]	**EDSS** • The correlation between sNfL and the EDSS during the 2 years was mild but statistically significant ($r_s = 0.15$, $p = 0.009$) • Neither baseline sNfL nor follow-up biomarker levels were significantly associated with the 10-year EDSS **RRMS conversion to SPMS** • Individually, baseline sNfL levels were not significantly associated with the odds of developing SPMS (OR = 0.64, 95% CI: 0.31 to 1.36) • However, when sNfL and sGFAP were modeled together, higher baseline sGFAP was associated with developing SPMS (OR = 3.3, 95% CI: 1.1 to 10.6, $p = 0.04$)
All multiple sclerosis subtypes (RRMS, SPMS, PPMS, CIS)	
Maltby 2023 [112]	**NEDA-3** • Patients with high starting sNfL Z-score were less likely to achieve NEDA-3 compared with those with a normal score (OR = 2.35, range: 1.17 to 4.77)
Meier 2023 [61], Disanto 2017 [14], Abdelhak 2023 [15]	**EDSS** • sNfL was independently associated with EDSS assessments ($\beta = 1.105$, $p < 0.001$) • The proportion of patients experiencing EDSS worsening within 12 months after sampling gradually increased with increasing sNfL percentile category: ○ 6.7% for samples <80th percentile to ~15% for samples >97.5th percentile (OR = 2.41, 95% CI: 1.07 to 5.42, $p = 0.034$) **CDP** • Patients with high sNfL levels (i.e., Z-score >1.3) showed a 2-fold increased risk of future CDP (HR = 2.26, 95% CI: 1.24 to 4.14, $p = 0.008$) vs. patients with low sNfL levels

Author year	Key outcomes relevant to the current SLR
	PIRA • Baseline sNfL levels had prognostic ability for future PIRA (HR = 1.77, 95% CI: 1.11 to 2.83, p = 0.02). However, the results were nonsignificant after adjustment for age, sex, BMI, and disease duration (HR = 1.90, 95% CI: 0.86 to 4.19, p = 0.11)
Sotirchos 2023 [40]	**Disease worsening** • Clinical disability was worse in those with elevated sNfL compared with those with normal sNfL, as evidenced by higher self-reported disability (adjusted OR—moderate vs. mild disability = 1.39, 95% CI: 1.15 to 1.67, p < 0.001; severe vs. mild disability = 2.26, 95% CI: 1.85 to 2.75, p < 0.001) and worse neuroperformance (adjusted difference in Z-scores—walking speed = −0.54, 95% CI: −0.80 to −0.28; manual dexterity = −0.45, 95% CI: −0.58 to −0.33; processing speed = −0.30; 95% CI: −0.38 to −0.22, p < 0.001 for all)
Pauwels 2022 [67]	**Disease worsening** • Median levels of sNfL were higher in patients with vs. without EDSS-Plus worsening (r = 0.21, p = 0.03) • High sNfL levels were associated with a higher risk for EDSS-Plus worsening (univariate, HR = 1.045, 95% CI: 1.019 to 1.071, p < 0.001, multivariate, HR = 1.046, 95% CI: 1.018 to 1.075, p < 0.001) • Patients with high sNfL (≥12.19 ng/L) had a significantly shorter time to EDSS-Plus worsening compared with patients with low sNfL (p = 0.045); however, the significance was lost when five patients with a relapse at baseline were excluded from the analysis (p = 0.058) **T25FWT** • sNfL levels correlated with baseline T25FW score (r_s = 0.29, p < 0.001) **9-HPT** • sNfL levels correlated with baseline 9-HPT dominant (r_s = 0.20, p = 0.047) and 9-HPT nondominant (r_s = 0.26, p = 0.009) scores

Lin 2021 [56]	**EDSS**
	• Abnormal sNfL showed no association with future confirmed EDSS worsening
	• No association was found for sNfL combined with any of the three optical coherence tomography parameters (thin GCIPL/pRNFL/thick INL)
	NEDA-3
	• Abnormal baseline sNfL[a] alone was associated with a higher risk of violating NEDA-3 (HR = 2.28, 95% CI: 1.27 to 4.09, p = 0.006)
	• Compared with abnormal sNfL alone, an even higher risk of violating NEDA-3 was associated with
	  ○ Abnormal sNfL + thin GCIPL (HR = 3.61, 95% CI: 1.77 to 7.36, p < 0.001)
	  ○ Abnormal sNfL + thin pRNFL (HR = 2.63, 95% CI: 1.21 to 5.70, p = 0.015)
	  ○ Abnormal sNfL + thick INL (HR = 3.05, 95% CI: 1.32 to 7.05, p = 0.009)

Jakimovski 2020 [49]	**EDSS** • Baseline sNfL levels predicted 5-year EDSS scores (r = 0.25, q = 0.012) • In the cross-sectional analysis using follow-up data, sNfL levels were significantly associated with the EDSS score (r = 0.356, q = 0.002) **Disability** • sNfL levels were cross-sectionally associated with walking speed (r = 0.235, q = 0.036), manual dexterity (r = 0.337, q = 0.002), and CPS (r = −0.265, q = 0.012)

Author year	Key outcomes relevant to the current SLR
	Cognition • Patients with cognitive impairment had higher follow-up sNfL levels (median = 27.2 vs. 20.6 pg/mL, p = 0.016) • Absolute change in sNfL over the 5 years leading to the follow-up ○ Cognitive examination was significantly greater in patients with vs. without cognitive impairment (median 4.8 vs. 0.7 pg/mL, p = 0.04)
Canto 2019 [60], Abdelhak 2023 [15]	**EDSS** • Baseline sNfL levels showed significant associations with EDSS score (β = 1.080, 95% CI: 1.047 to 1.114, p < 0.001; i.e., 8.0% higher sNfL levels per EDSS step) • A significant interaction was noted between EDSS worsening and change in sNfL levels over time (β = 1.015, 95% CI: 1.007 to 1.023, p < 0.001) between progressors and nonprogressors, indicating a steeper trajectory of sNfL levels in progressors. This result remained significant after correction for age, sex, and disease duration • At the last visit, sNfL levels showed a univariable association with EDSS score (β = 1.095, 95% CI: 1.071 to 1.120, p < 0.001). The association remained significant in multivariate analysis • sNfL levels categorized according to extreme percentiles were not associated with subsequent EDSS worsening, nor were they associated with previous EDSS worsening
Barro 2018 [16]	**EDSS worsening** • sNfL levels >90th percentile were associated with increased odds of EDSS worsening at the next visit compared with levels below the 90th percentile (estimated β_{OR} = 2.577, 95% CI: 1.553 to 4.278, p < 0.001, n = 677 observations) • In the multivariable model, sNfL above the 90th percentile (β_{OR} = 2.786, 95% CI: 1.609 to 4.826, p < 0.001, n = 677 observations) was also a significant predictor of EDSS worsening in the subsequent year • Probability of EDSS worsening gradually increased with higher sNfL percentile category • Univariable analyses showed significant positive associations of sNfL with EDSS (β_{mult} = 1.094, 95% CI: 1.070 to 1.120, p < 0.001) • Multivariable model analysis confirmed the association of higher sNfL levels with higher EDSS, whereas higher values of progressive vs. relapsing MS were no longer statistically significant
Chitnis 2018 [94], Barro 2022 [119], Barro 2023 [120]	**6mCDP** • sNfL was associated with the risk of 6mCDP [HR = 1.78, 95% CI: 1.13 to 1.69, p = 0.002]. However, sNfL levels did not predict future 6mCDW in any age group when evaluated separately • sNfL levels at Year 2 only was correlated with Year 10 EDSS (r_s = 0.21, p = 0.04) **SDMT and T25FW** • sNfL was significantly associated with concurrent SDMT (adjusted mean change in SDMT score = −4.5, 95% CI: −8.7 to −0.2, p = 0.039) • sNfL predicted decline in SDMT score, particularly in active patients (adjusted change in slope = −1.14, 95% CI: −1.83 to −0.44, p = 0.001) • No significant associations of either annual or averaged yearly sNfL with 10-year SDMT score and T25FW were observed
Progressive multiple sclerosis (SPMS, PPMS)	
Bar-Or 2023 [44] (ORATORIO)	**EDSS** • Higher baseline NfL levels were independently associated with higher EDSS (effect on $\log_{10}$ sNfL in multiple linear regression model = 0.02, 95% CI: 0.0 to 0.04, p = 0.0682)

Author year	Key outcomes relevant to the current SLR
	Disease progression • High NfL at Week 48 (>7.5 pg/mL) was significantly associated with the risk for future 24-week CDP in patients receiving ocrelizumab
Brune 2022 [45]	**SDMT score** • Higher sNfL concentrations at baseline were significantly associated with lower baseline scores on the SDMT (r_p = –0.32, p = 0.03)
Comabella 2022 [37]	**EDSS** • Baseline sNfL levels were not associated with EDSS progression • In univariable analysis, a baseline sNfL cutoff value of 10.2 pg/mL discriminated between long-term progressors and nonprogressors with a 75% sensitivity and 67% specificity (adjusted OR = 7.8, 95% CI: 1.8 to 46.4, p = 0.01) • In univariable analysis, a cutoff increase of 5.1 pg/mL in sNfL levels between baseline and 6 years also discriminated between long-term progressors and nonprogressors with a 71% sensitivity and 86% specificity. ○ A cutoff increase of 5.1 pg/mL sNfL in medium term remained significant in the adjusted logistic regression (OR = 49.4, 95% CI: 4.4 to 2×10^3, p = 0.008), although with high variability
Leppert 2022 [96]	**Disease progression** **EXPAND** • Higher baseline sNfL levels were associated with higher EDSS score (geometric mean ratio = 1.065, 95% CI: 1.038 to 1.093, p < 0.0001), more Gd+ lesions (geometric mean ratio = 1.441, 95% CI: 1.347 to 1.541, p < 0.0001), and higher T2 lesion load (geometric mean ratio = 1.007, 95% CI: 1.005 to 1.009, p < 0.0001) • High vs. low baseline sNfL levels were associated with significantly higher risks of confirmed 3-month (32%; HR = 1.32, 95% CI: 1.09 to 1.61) and 6-month (26%; HR = 1.26, 95% CI: 1.01 to 1.57) disability progression, earlier wheelchair dependence (50%; HR = 1.50, 95% CI: 0.96 to 2.34), cognitive decline (41%; HR = 1.41, 95% CI: 1.09 to 1.84), and higher rates of brain atrophy (mean change at Month 24 = −0.92) • Baseline sNfL levels were associated with future disability progression and the degree of brain atrophy regardless of presence or absence of acute disease activity **INFORMS** • Higher baseline sNfL levels were associated with higher EDSS scores (geometric mean ratio = 1.087, 95% CI: 1.029 to 1.148, p < 0.0030), more Gd+ lesions (geometric mean ratio = 1.571, 95% CI: 1.306 to 1.890, p < 0.0001), and higher T2 lesion load (geometric mean ratio = 1.014, 95% CI: 1.008 to 1.020, p < 0.0001) • High vs. low baseline sNfL levels were associated with significantly higher risks of confirmed 3-month (49%; HR = 1.49, 95% CI: 1.05 to 2.12) and 6-month (48%; HR = 1.48, 95% CI: 1.01 to 2.17) disability progression, earlier wheelchair dependence (197%; HR = 2.97, 95% CI: 1.44 to 6.10), and higher rates of brain atrophy (mean change at Month 24 = −1.39) • Baseline sNfL levels were associated with future disability progression and the degree of brain atrophy regardless of presence or absence of acute disease activity
Pauwels 2022 [67]	**Disease worsening** • Median levels of sNfL were higher in patients with vs. without EDSS-Plus worsening (r = 40, p = 0.04)

Author year	Key outcomes relevant to the current SLR
Giarraputo 2021[48]	**Clinical worsening** • No association between baseline sNfL levels and subsequent clinical worsening in walking speed (OR = 0.249, 95% CI: 0.021 to 1.710, p = 0.194), walking distance (OR = 0.510, 95% CI: 0.056 to 3.433, p = 0.502), or balance (OR = 0.780, 95% CI: 0.109 to 5.047, p = 0.790) was observed
Chitnis 2018 [94], Barro 2023 [97]	**Disease activity** • sNfL was higher in patients with disease activity in the 2 years before baseline (adjusted β = 1.21, 95% CI: 1.04 to 1.42, p = 0.016) and during the first 2 years of follow-up (adjusted β = 1.17, 95% CI: 1.01 to 1.36, p = 0.042) **EDSS** • sNfL was associated with baseline EDSS score (adjusted β = 1.08, 95% CI: 1.00 to 1.15, p = 0.041) **6mCDP on EDSS** • sNfL was not significantly associated with time to 6mCDP in all patients • sNfL was not associated with the risk of future 6mCDP and PIRA in either active or nonactive patients

[a]In this study, abnormal sNfL was defined as sNfL levels >80th percentile of age-corrected reference values.

Note: sNfL levels reported in pg/mL; pg/mL and ng/L were considered equivalent units as 1 pg/mL = 1 ng/L. Some studies reported use of log normal sNfL levels for analysis; however, no information was reported for other studies. **Note**: All references are provided within the manuscript.

Abbreviations: 3mCDP, 3-month confirmed disability progression; 6mCDP, 6-month confirmed disability progression; 6mCDW, 6-month confirmed disability worsening; 9-HPT, 9-Hole Peg Test; AUC, area under the curve; BMI, body mass index; BVMT-R, Brief Visuospatial Memory Test-Revised; BVMTR-DR, Brief Visuospatial Memory Test-Revised—Delayed Recall; CDP, confirmed disability progression; CDW-NR, confirmed disability worsening with no clinical relapse; CDW-R, confirmed disability worsening with clinical relapse; CI, confidence interval; CIS, clinically isolated syndrome; CPS, cognitive processing speed; CVLT-II, California Verbal Learning Test-II; EDA, evidence of disease activity; EDSS, Expanded Disability Status Scale; GCIPL, ganglion cell and inner plexiform layer; Gd+, gadolinium-enhancing; HR, hazard ratio; INL, inner nuclear layer; IQR, interquartile range; MoCA, Montreal Cognitive Assessment; MRI, magnetic resonance imaging; MS, multiple sclerosis; NEDA, no evidence of disease activity; NfL, neurofilament light chain; OR, odds ratio; PARA, progression associated with relapse activity; PASAT, Paced Auditory Serial Addition Test; PIRA, progression independent of relapse activity; PPMS, primary progressive multiple sclerosis; pRNFL, peripapillary retinal nerve fiber layer; RAW, relapse-associated worsening; RFP, relapse-free progression; RRMS, relapsing-remitting multiple sclerosis; SD, standard deviation; SDMT, Symbol Digit Modalities Test; SE, standard error; sGFAP, serum glial fibrillary acidic protein; SLR, systematic literature review; sNfL, serum neurofilament light chain; SPMS, secondary progressive multiple sclerosis; T25FWT, Timed 25-Foot Walk Test