CADTH

CADTH Common Drug Review

Clinical Review Report

SATRALIZUMAB (ENSPRYNG)

(Hoffmann-La Roche Limited)

Indication: Neuromyelitis optica spectrum disorder

Service Line:	CADTH Common Drug Review
Version:	Final (with redactions)
Publication Date:	June 2021
Report Length:	92 Pages

CADTH

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

CADTH

Table of Contents

Tables

Figures

Abbreviations

AE	adverse event
AQP4	aquaporin 4
ARR	annualized relapse rate
BOCF	baseline observation carried forward
CI	confidence interval
CrI	credible interval
EDSS	Expanded Disability Status Scale
EQ-5D-3L	EuroQol 5-Dimensions 3-Levels questionnaire
FACIT-F	Functional Assessment of Chronic Illness Therapy–Fatigue
FSS	Functional System Score
HR	hazard ratio
IgG	immunoglobulin G
ITC	indirect treatment comparison
ITT	intention to treat
MID	minimal important difference
MMRM	mixed model for repeated measures
MS	multiple sclerosis
NMA	network meta-analysis
NMO	neuromyelitis optica
NMOSD	neuromyelitis optica spectrum disorder
PY	patient-year
SD	standard deviation
SF-36v2	Short Form (36) Health Survey version 2
VAS	Visual Analogue Scale
WDAE	withdrawal due to adverse event

Executive Summary

An overview of the submission details for the drug under review is provided in Table 1.

Table 1: Submitted for Review

Item	Description
Drug product	Satralizumab (Enspryng), 120 mg/mL pre-filled syringe for subcutaneous injection
Indication	As monotherapy or in combination with immunosuppressive therapy for the treatment of NMOSD in adult and adolescent patients who are anti–AQP4 seropositive. Satralizumab is not intended for acute treatment of an NMOSD relapse.
Reimbursement request	As per indication
Health Canada approval status	NOC
Health Canada review pathway	Priority review
NOC date	June 1, 2020
Sponsor	Hoffmann-La Roche Limited

AQP4 = aquaporin 4; NMOSD = neuromyelitis optica spectrum disorder; NOC = Notice of Compliance.

Introduction

Neuromyelitis optica spectrum disorder (NMOSD) is a rare, debilitating, immune-mediated, demyelinating disorder of the central nervous system. It is typically characterized by acute attacks or relapses that primarily cause damage to the optic nerves and spinal cord, which can result in blindness, weakness in the arms and legs, numbness or tingling, pain and discomfort, bladder and bowel dysfunction, paraplegia, and increased overall mortality.[1-4] Most of the disability in NMOSD is incurred through relapses; thus, prevention of relapses is a key goal of therapy. The discovery of aquaporin 4 immunoglobulin G (AQP4 IgG) was key to understanding the pathogenesis of NMOSD. The AQP4 antibody is found in 70% to 90% of patients with NMOSD and is a defining criteria.[1,3,5,6] Estimates of the incidence and prevalence of NMOSD range from 0.053 to 0.40 per 100,000 people and 0.51 to 4.4 per 100,000 people, respectively.[7,8] No Canadian-specific prevalence estimates are currently available.

Off-label immunosuppressants, such as azathioprine, mycophenolate mofetil, and rituximab, are used in Canada to prevent relapses, although the treatment approach differs by province and territory due in part to differential access to these drugs. These immunosuppressants have significant failure rates, and patients will still experience severe relapses that cause them to accrue disability. Non-specific immunosuppressants also have safety concerns, particularly with longer-term use and in younger patients. Eculizumab (Soliris) was the first drug approved in Canada for NMOSD, but it is limited clinically by its IV route of administration and potential adverse effects, including increased risk of meningococcal infection.

Satralizumab, a monoclonal antibody, is indicated as monotherapy, or in combination with immunosuppressive therapy, for the treatment of NMOSD in adult and adolescent patients who are AQP4 antibody positive.[9] Satralizumab is not intended for acute treatment of an NMOSD relapse.[9] Satralizumab is available as a 120 mg/mL single-use, pre-filled syringe and the recommended dose is 120 mg by subcutaneous (SC) injection at weeks 0, 2, and 4, and then every 4 weeks thereafter.[9]

The objective of this report is to perform a systematic review of the beneficial and harmful effects of satralizumab as monotherapy or in combination with immunosuppressive therapy for the treatment of NMOSD in adult and adolescent patients who are AQP4 antibody positive.

Stakeholder Engagement

The information in this section is a summary of input provided by the patient groups that responded to CADTH's call for patient input and from clinical experts consulted by CADTH for the purpose of this review.

Patient Input

The Multiple Sclerosis (MS) Society of Canada provided input for CADTH's review of satralizumab. The MS Society of Canada gathered information for its submission through an online survey that received responses from 37 people, including 25 respondents (68%) who had been diagnosed with NMOSD.

Patients with NMOSD report experiencing pain, muscle weakness, paralysis, loss of vision, and bladder or bowel control problems caused by relapses. The accrued disability leads to employment instability or loss, increased need for assistance or caregiving, loss of independence, isolation, cognitive decline, and increased mobility challenges.

Patients hope that satralizumab will reduce attacks and reduce disability, which were highlighted as important therapeutic gaps with current NMOSD drug treatments. The input noted that treatment with satralizumab has the potential to allow people living with NMOSD to remain in the workforce, sustain family and social roles and responsibilities longer, improve their quality of life, and decrease the need for family or paid caregivers.

Clinician Input

The panel of clinical experts emphasized the need for safe and effective relapse-prevention treatments for patients with NMOSD, as early intervention to eliminate relapses is key to averting disability and improving longer-term outcomes for patients.

The experts consulted indicated that satralizumab could be used as a first-line treatment for NMOSD but could also be used after inadequate response or intolerance to immunosuppressants. The panel acknowledged that comparative clinical data are not available at this time to optimally guide the position of satralizumab in the treatment algorithm. Satralizumab is suitable for patients with a confirmed diagnosis of NMOSD who are AQP4 antibody positive. It may be used as monotherapy or in combination with immunosuppressants.

Treatment response is determined by the elimination of relapses or a decrease in the frequency or severity of relapses. According to the experts, relapses are identified clinically, based on a neurological examination and patient-reported symptoms. The clinical experts agreed that diagnosis of NMOSD and prescribing of satralizumab should be limited to neurologists with specific expertise in NMOSD or demyelinating disorders.

Clinical Evidence

Pivotal Studies and Protocol Selected Studies

Description of Studies

The systematic review included 2 pivotal, double-blind randomized controlled trials (RCTs) designed to evaluate the safety and efficacy of satralizumab versus placebo in patients with neuromyelitis optica (NMO) or NMOSD and included patients who were AQP4 IgG positive or negative (Study 898 and Study 900). Patients were randomized to placebo or satralizumab 120 mg by SC injection at weeks 0, 2, and 4 and every 4 weeks thereafter. The primary outcome in both trials was the time to first adjudicated, protocol-defined relapse for the overall, intention-to-treat (ITT) population. Key secondary outcomes included the change from baseline to week 24 in pain Visual Analogue Scale (VAS) scores and the change in Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) scores.

Study 898 enrolled 83 adults and adolescents (12 years to 74 years of age), of whom 55 (66%) were included in the AQP4 antibody–positive subgroup (i.e., the indicated population). The patients enrolled had at least 2 relapses in the past year (1 of which occurred in the last 12 months) and all received background immunosuppressant treatment of azathioprine, mycophenolate mofetil, or corticosteroids during the trial.

Study 900 enrolled 95 adults aged 18 to 74 years who had at least 1 relapse in the past year, including a first attack. The AQP4 IgG–positive subgroup included 64 patients (67%).

Across both trials, the mean age of patients in the AQP4 antibody–positive subgroup ranged from [redacted] years per treatment group, and [redacted] of patients were female. Most of the patients were White ([redacted]) or Asian ([redacted]). The mean annualized relapse rate (ARR) at baseline was [redacted] in Study 898 and [redacted] in Study 900, with a median baseline Expanded Disability Status Scale (EDSS) score of [redacted] in the AQP4 antibody–positive subgroup.

The CADTH review focused on the results in the AQP4 IgG–positive subgroup, as this is the indicated population in Canada.

Efficacy Results

The primary outcome for both trials was the time to first adjudicated, protocol-defined relapse for the ITT population. A protocol-defined relapse was any new or worsening neurological symptoms attributable to NMO or NMOSD that persisted for a minimum of 24 hours, were not attributable to confounding clinical factors, and met 1 of 5 predefined criteria for a 1.0- or 2.0-point increase in EDSS or Functional System Score (FSS).

For the ITT population, the time to first adjudicated, protocol-defined relapse was statistically significantly different favouring satralizumab versus placebo when administered as add-on therapy to immunosuppressants (hazard ratio [HR] = 0.38; 95% confidence interval [CI], 0.16 to 0.88; P = 0.018), or as monotherapy (HR = 0.45; 95% CI, 0.23 to 0.89; P = 0.018) in Study 898 and Study 900, respectively.

For the AQP4 IgG–positive subgroup in Study 898, 43% (12 of 28) patients in the placebo plus immunosuppressant group, and 11% (3 of 27) in the satralizumab plus immunosuppressant group experienced an adjudicated, protocol-defined relapse, with an

HR of 0.21 (95% CI, 0.06 to 0.75; P = 0.0086; not controlled for type I error rate) (Table 2). In the AQP4 IgG–positive subgroup of Study 900, 57% (13 of 23) of patients in the placebo group and 22% (9 of 41) of patients in the satralizumab group experienced a protocol-defined relapse (HR = 0.26; 95% CI, 0.11 to 0.63; P = 0.0014; not controlled for type I error rate).

No statistically significant differences were detected between groups for the change from baseline to week 24 in pain VAS score or FACIT-F scores for the ITT populations (key secondary outcomes). Pain and fatigue data for the AQP4 IgG–positive subgroup are shown in Table 2 and were generally consistent with the results in the overall study populations.

The trials did not report health-related quality of life or disability outcomes for the AQP4 IgG–positive subgroup. No consistent differences were found between groups based on the ITT population in either study for the change from baseline in EDSS score, modified Rankin Scale score, visual acuity, EuroQol 5-Dimensions questionnaire (EQ-5D), or Short Form (36) Health Survey (SF-36). The ARR was reported for the ITT population only. These data suggest a reduction in ARR for satralizumab versus placebo in Study 898; however, these data were not controlled for the type I error rate and should be considered as supportive evidence only. Moreover, the ARR is likely under-reported due to the design of the trials, where patients were withdrawn after a relapse. No information on productivity or health care resource utilization was reported in either study.

Harms Results

The proportion of patients who experienced an adverse event ranged from 75% to 95% in the placebo groups and from 90% to 92% in the satralizumab groups. After adjusting for follow-up time, the rate of adverse events was 495 to 514 events per 100 patient-years (PYs) among those assigned to placebo, and from 474 to 485 events per 100 PYs to those who received satralizumab (Table 3). The most common adverse events were urinary tract infections (17% to 25% of patients), upper respiratory tract infection (14% to 24%), headache (10% to 24%), nasopharyngitis (3% to 24%) and injection-related reactions (5% to 16%). The rate of infections ranged from 150 to 163 events per 100 PYs among those randomized to placebo, and from 100 to 133 events per 100 PYs to those who received satralizumab.

Serious adverse events were reported in 16% to 21% of patients assigned to placebo, and 17% to 19% of patients who received satralizumab, with a serious adverse event rate of 15 to 20 events per 100 PYs, and 12 to 17 events per 100 PYs in the placebo and satralizumab groups, respectively. More patients stopped treatment due to adverse events in the add-on therapy trial (Study 898: placebo = 12%; satralizumab = 7%) than in the monotherapy trial (Study 900: placebo 3%, satralizumab 2%) (Table 3).

No deaths, hepatotoxicity, or anaphylaxis events were reported in either study. Injection-related reactions were reported by 5% to 16% of patients; however, no patient stopped treatment due to these adverse events.

Table 2: Summary of Key Efficacy Results From Pivotal and Protocol Selected Studies, AQP4 IgG–Positive Subgroup

	Study 898 AQP4 IgG–positive subgroup		Study 900 AQP4 IgG–positive subgroup	
	Placebo plus IST N = 28	Satralizumab plus IST N = 27	Placebo N = 23	Satralizumab N = 41
Time to first protocol-defined relapse (adjudicated)				
Number of patients contributing to the analysis	28	27	23	41
Number of patients with relapse, n (%)	12 (42.9)	3 (11.1)	13 (56.5)	9 (22.0)
HR versus placebo (95% CI)[a]		0.21 (0.06 to 0.75)		0.26 (0.11 to 0.63)
P value (log-rank)		0.0086[b]		0.0014[b]
Change from baseline to week 24 in pain VAS[c]				
Number of patients contributing to the analysis	28	27	23	40
Adjusted mean change from baseline, mean (95 % CI)	−5.3 (−11.2 to 0.6)	5.5 (−0.5 to 11.6)	−8.5 (−20.1 to 3.1)	−1.5 (−12.5 to 9.5)
Difference in adjusted means versus placebo (95% CI)[d]		10.8 (2.4 to 19.2)		7.1 (−4.0 to 18.1)
P value		0.0132[b]		0.207[b]
Change from baseline to week 24 in FACIT-F score[e]				
Number of patients contributing to the analysis	28	27	23	40
Adjusted mean change from baseline, mean (95% CI)	2.7 (0.5 to 4.8)	−0.3 (−2.5 to 1.9)	4.4 (0.4 to 8.4)	6.5 (2.7 to 10.3)
Difference in adjusted means versus placebo (95% CI)[d]		−2.9 (−6.0 to 0.1)		2.1 (−1.6 to 5.9)
P value		0.059[b]		0.264[b]

ANCOVA = analysis of covariance; AQP4 = aquaporin 4; ARR = annualized relapse rate; BOCF = baseline observation carried forward; CI = confidence interval; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; HR = hazard ratio; IgG = immunoglobulin G; IST = immunosuppressive therapy; VAS = Visual Analogue Scale.

[a] HR and 95% CI based on Cox proportional hazards model and P value based on a log-rank test for the AQP4 IgG–positive subgroup. In Study 898, the analyses were stratified by baseline ARR (1 versus > 1) and geographic region (Asia versus Europe or other), and in Study 900, analyses were stratified by prior therapy (B-cell–depleting therapy versus ISTs or other) and most recent attack (first attack versus relapse).

[b] P value has not been adjusted for multiple testing (i.e., the type I error rate has not been controlled).

[c] Pain VAS is scored from 0 to 100, with higher scores representing worse pain. For the between-group comparison, the negative difference in means favours satralizumab over placebo.

[d] Analysis based on ANCOVA model including treatment, baseline value, and stratification factors (Study 898: baseline ARR and geographic region; Study 900: prior therapy and most recent attack type) with BOCF for missing data.

[e] FACIT-F scale is scored from 0 to 52, with lower scores representing more fatigue. For the between-group comparison, positive difference in means favours satralizumab over placebo.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Table 3: Summary of Key Safety Results From Pivotal and Protocol Selected Studies

	Study 898 safety population				Study 900 safety population			
	Placebo plus IST N = 42		Satralizumab plus IST N = 41		Placebo N = 32		Satralizumab N = 63	
	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]
Patients with ≥ 1 AEs								
AEs	40 (95)	514.3	37 (90)	485.2	24 (75)	495.2	58 (92)	473.9
SAEs	9 (21)	20.2	7 (17)	11.5	5 (16)	14.8	12 (19)	17.4
Stopped treatment due to AE	5 (12)	NR	3 (7)	NR	1 (3)	NR	1 (2)	NR
Death	0	0	0	0	0	0	0	0
Notable harms								
Infection and infestations (SOC)	26 (62)	149.6	28 (68)	132.5	14 (44)	162.6	34 (54)	99.8
Serious infection	3 (7)	5.0	2 (5)	2.6	3 (9)	9.9	6 (10)	5.2
Potential opportunistic infection	5 (12)	35.3	4 (10)	10.2	5 (16)	17.3	3 (5)	2.6
Injection-related reaction	2 (5)	3.4	5 (12)	21.7	5 (16)	17.3	8 (13)	13.9

AE = adverse event; IST = immunosuppressive therapy; NR = not reported; PY = patient-year; SAE = serious adverse event; SOC = system organ class.

[a] Number of patients with an AE.

[b] Number of AEs per 100 PYs of follow-up (multiple occurrences of the same event in 1 patient counted multiple times). Study 898 total PYs: placebo = 59.5, satralizumab = 78.5. Study 900 total PYs: placebo = 40.6, satralizumab = 115.2.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Other Considerations

The clinical experts consulted by CADTH indicated there is potential off-label use of satralizumab in patients younger than 12 years of age and those who are AQP4 IgG negative.

Although satralizumab is not approved for use in patients who are negative for AQP4 antibodies, the clinical experts stated that the mechanism of action of satralizumab provides multiple mechanisms of immunomodulation and may impact the mechanisms affecting AQP4 IgG–negative NMOSD patients. Moreover, patients who are AQP4 IgG negative have no approved or soon-to-be approved drugs available to them, as eculizumab, and the FDA-approved inebilizumab, are only approved for use in AQP4 IgG–positive patients.

Pediatric-onset NMOSD is very rare, but the disabilities associated with this disease that occur at a young age (i.e., loss of vision) can have an important lifelong impact. The experts stated that despite the limited clinical data, there is potential for off-label use in children younger than 12 years of age.

Critical Appraisal

The available evidence consisted of 2 double-blind, placebo-controlled event-driven trials that evaluated the safety and efficacy of satralizumab as monotherapy or in combination with immunosuppressants, in adult and adolescent patients with NMOSD (95 patients in Study 900 and 83 patients in Study 898). Both trials used accepted methods to randomize

patients and although allocation was likely adequately concealed, all prognostic or effect modifiers may not be balanced between groups in the overall population due to the small sample size and, in the subgroup of AQP4 IgG–positive patients, due to the lack of stratification at randomization. Imbalances between treatment groups in age, sex, and racial distribution, body mass index, and prior therapies were observed in Study 900 but were not thought to impact key outcomes after consultation with clinical experts.

The primary outcome in both trials was the time to first protocol-defined relapse that was confirmed by the blinded clinical event committee. Use of adjudication is expected to increase the validity and objectivity of the outcome, as it reduces inter-site variability in assessments and over-reporting bias that may have influenced attending physician–determined relapses, as the need for immediate treatment of relapses could impact the classification of an event as a relapse. The study design inherently emphasizes the efficacy of satralizumab on the first relapse, but it is not designed to assess its efficacy pertaining to subsequent relapses or the impact of relapses on symptoms, disability, or health-related quality of life. As such, the trials may not fully capture the change in pain VAS, FACIT-F, EDSS, EQ-5D, or SF-36 scores over time. Many patients did not have pain VAS or FACIT-F score data at 24 weeks (17% to 38%) due to patients being withdrawn after experiencing a relapse or other early discontinuations. Those with missing data had baseline values carried forward, which likely biases the results in favour of satralizumab, as patients who are missing due to relapse likely have worse outcomes. Disability or health-related quality of life measure data that were missing for 14% to 41% of patients at 24 weeks, with no imputation for missing data. Given the magnitude of missing data, there is potential for the validity of these results to be affected.

While ARR results were reported and are a clinically relevant end point, the trials do not capture data on subsequent relapses (occurring after 30 days of the first relapse) because patients were censored and, therefore, subsequent relapses would not have been captured, thereby likely underestimating the ARR. ARR was not part of the statistical testing hierarchy and the P value was not controlled for type I error rate; thus, these data should be considered as supportive evidence for the effect of satralizumab in the overall population. Data for ARR were not reported for the AQP4 IgG–positive subgroup.

With respect to external validity, clinical expert input to CADTH considered the baseline demographics and disease characteristics of the patients enrolled in the trials to be consistent with patients seen in the Canadian clinical setting although, in Study 898, the frequency of use of corticosteroids as longer-term relapse-prevention therapy was not consistent with Canadian clinical practice. Although adjudicated protocol-defined relapses are thought to be the more robust and reproducible measure, there may be issues with the generalizability of the results to clinical practice, where strict criteria are not used to identify relapses. Given the lack of head-to-head studies comparing satralizumab with eculizumab or other immunosuppressants, determining the comparative efficacy and the optimal place in therapy for satralizumab may be challenging.

Indirect Comparisons

The sponsor submitted an indirect treatment comparison (ITC) that estimated the relative treatment effects and safety of satralizumab versus eculizumab or inebilizumab. Bayesian network meta-analysis (NMA) methods were used to combine data from 4 RCTs, including an analysis in patients who were AQP4 IgG positive based on pooled subgroup data from the 2 pivotal satralizumab trials. The NMA results for the time to first protocol-defined relapse did not differentiate between satralizumab and eculizumab, or between

satralizumab and placebo, and showed wide 95% credible intervals (CrIs) and high uncertainty. A similar pattern of results was observed for the analyses of ARR, proportion of relapse-free patients at 48 weeks, change in EDSS score at 48 weeks, withdrawals due to adverse events, and rate of serious infections.

Although the NMA was conducted using accepted statistical methods, there were many differences between populations, study designs, effect modifiers, and end point definitions in these 4 trials, which present severe limitations to the analyses. Due to the sparse network, which was based on subgroup data, and the clinical heterogeneity between trials, the results of the NMA are highly uncertain and, thus, no conclusions can be drawn on the comparative efficacy and safety of satralizumab versus eculizumab in patients who are AQP4 IgG positive.

Other Relevant Evidence

Longer-term safety and efficacy data were reported for the open-label extension phase of Study 898 and Study 900. Patients were eligible for the extension period if they experienced a protocol-defined relapse (both studies) or received rescue therapy (Study 898) during the double-blind period, or if they completed the double-blind period without a relapse. During the extension period, all patients received open-label satralizumab 120 mg SC at weeks 0, 2, and 4 and every 4 weeks thereafter.

The interim data reported included a summary of harms for 145 patients who received satralizumab in either the double-blind or extension period of each study (up to October 2018) and supplemental pooled safety and efficacy data for the double-blind and extension period up to June 2019 (N = 166).

Evidence from the extension period suggests acceptable tolerability of satralizumab administered every 4 weeks. The longer-term harms data are consistent with the double-blind period, with infections being the most commonly reported adverse event (92 to 146 infections reported per 100 PYs of follow-up). The rate of serious adverse events ranged from 13 to 15 events per 100 PYs, and discontinuation of treatment due to adverse events ranged from 0.6 to 3.8 events per 100 PYs. The data are limited by selection bias, lack of blinding, and lack of comparator group, which may affect the internal or external validity of the results.

Conclusions

In patients with AQP4 antibody–positive NMOSD, fewer patients treated with satralizumab experienced an adjudicated relapse relative to placebo when administered as monotherapy or in combination with immunosuppressants. The between-group differences were considered clinically meaningful based on clinical expert input.

The 2 pivotal trials did not demonstrate an effect for satralizumab on pain or fatigue symptoms measured using a VAS for pain or FACIT-F score at 24 weeks. No conclusions can be drawn on the impact of satralizumab on disability or health-related quality of life due to limitations in the design of the trials and the extent of missing data. No data were available to assess the effects of productivity or health care resource utilization.

Infections were the most commonly reported adverse event in the double-blind and open-label extension periods. Safety data were limited by the small sample size of the trials and the lack of blinding, comparator group, and potential selection bias for the extension period.

Head-to-head trials comparing satralizumab with other immunosuppressants are lacking. The sponsor-submitted indirect comparison that estimated the relative treatment effects and safety of satralizumab versus eculizumab was limited by the sparse network and clinical heterogeneity between trials. The results of the indirect comparison were highly uncertain and, thus, no conclusions can be drawn on the comparative efficacy and safety of satralizumab versus eculizumab in patients who are AQP4 antibody positive.

Introduction

Disease Background

NMOSD is a rare, immune-mediated demyelinating disorder of the central nervous system known to be an astrocytopathy that typically causes marked injury to optic nerves, spinal cord, and the brainstem, with potential damage to the cortex, hypothalamus, and other central nervous system structures. It is distinct from MS, despite overlapping clinical features.[1,2,5] NMOSD is a debilitating disease and is typically characterized by acute attacks or relapses of optic neuritis and longitudinally extensive transverse myelitis, although additional clinical characteristics are now recognized, such as area postrema syndrome.[1-3] Optic neuritis involves inflammation of the optic nerve; it causes eye pain and vision loss and can occur unilaterally or bilaterally. Transverse myelitis is inflammation of the spinal cord that may cause sensorimotor impairment, which may result in weakness in the arms and legs, numbness or tingling, pain and discomfort, and bladder and bowel dysfunction. The natural history of the disease is most often relapsing, where patients experience an episode and then may demonstrate some recovery, followed by further episodes and partial recovery while progressively accruing disability.[1,2,5] In some patients, the first episode is severe enough to cause permanent disability. Most of the disability in NMOSD is incurred through relapses rather than progression (as is the case with MS). Relapses in NMOSD can result in blindness, paraplegia, and increased overall mortality.[1,4] A relapse is defined as the development of new signs and/or symptoms that prompt a change or addition of treatments such as immunosuppressants, plasma exchange, or IV immunoglobulin. The diagnosis of NMOSD now typically occurs during the first episode. In the past, when NMOSD was less recognized, the diagnosis may have been delayed or initially misclassified as MS. Input from patients and their caregivers highlighted the debilitating nature of the damage caused by NMOSD relapses and the resulting impact on vision and mobility that leads to a loss of independence, which alters every aspect of daily life.

NMOSD disproportionally affects females and those with coexisting autoimmune diseases.[4,5] Systematic reviews based on data from several countries estimated the incidence and prevalence of NMO to range from 0.053 to 0.40 per 100,000 people and 0.51 to 4.4 per 100,000 people, respectively.[7,8] No Canadian-specific estimates were identified in either study. It is unclear if these data are representative of NMOSD in Canadians, as the criteria for NMOSD are broader than those for NMO. People of Asian and African ancestry are at increased risk of NMOSD, and those with African ancestry have higher rates of mortality.[4,12] A recent study on overall mortality based on data from 2 large US clinics estimated the mortality rate to be 7%, which differs substantially from the mortality rate described in older studies (22% to 32%).[12]

NMOSD was previously referred to as Devic disease and, until 2004, was suspected to be a severe form of MS.[4,5] The discovery of AQP4 immunoglobulin G (AQP4 IgG) was key to the understanding of the pathogenesis of NMOSD and was an important factor in distinguishing it from MS[3,13] This antibody binds to AQP4, an abundant water channel in the central nervous system expressed on astrocytes.[3] AQP4 IgG is found in 70% to 90% of patients with NMOSD.[1,3,5,6] Other antibodies, such as myelin oligodendrocyte glycoprotein antibodies, may also be involved in the pathology of NMOSD; however, the evidence is limited compared with AQP4 IgG.[3,5] For a description and critical appraisal of the evidence evaluating detection tests for AQP4 IgG in patients with NMOSD, please refer to Appendix 5 in the CADTH Clinical Review Report for Soliris.[14]

In Canada, patients are typically diagnosed by a neurologist or physician with expertise in demyelinating disorders. The criteria currently used in Canada are based on the 2015 diagnostic criteria established by the International Panel for NMO Diagnosis.[15] There are separate criteria for patients who test positive for AQP4 IgG and for those who test negative for AQP4 IgG or whose status is unknown. A diagnosis of NMOSD for patients who test positive for AQP4 IgG involves 1 core clinical characteristic (i.e., optic neuritis, acute myelitis, area postrema syndrome, acute brainstem syndrome, symptomatic narcolepsy, or acute diencephalic clinical syndrome with NMOSD-typical MRI lesions, or symptomatic cerebral syndrome with NMOSD-typical brain lesions) and the exclusion of alternative diagnoses.[5] As the testing for AQP4 IgG has evolved and become more available, it is now possible to identify a broader range of patients and to identify patients much earlier in the disease course, which allows for earlier treatment and possibly less disability. Few patients who are AQP4 IgG positive have monophasic disease; this is more often seen in those who are AQP4 IgG negative. A diagnosis of NMOSD for patients who test negative for AQP4 IgG (or have unknown AQP4 IgG status) requires more stringent clinical and MRI criteria.[5]

Standards of Therapy

The goals of treatment relate to 3 broad areas: the prevention of relapses, the treatment of relapses, and the treatment of residual symptoms following an episode. The focus of this review is on the prevention of relapses based on the indication and place in therapy for satralizumab.

The therapeutic management of NMOSD in Canada is not based on a specific clinical guideline and, until recently, there were no Health Canada–approved drugs for the treatment of patients with NMOSD. Treatment differs by province and territory based in part on differential access to drugs (e.g., mycophenolate mofetil and rituximab). According to the clinical experts consulted, first-line therapies for the prevention of relapses in NMOSD include rituximab, azathioprine, and mycophenolate mofetil (Table 4). Other therapies that may be used to prevent relapses are tocilizumab, methotrexate, cyclophosphamide, mitoxantrone, cyclosporine, oral corticosteroids (prednisone), or bortezomib. The evidence for the use of these drugs comes primarily from observational studies, except for 1 trial conducted with rituximab versus azathioprine.[16] Rituximab may exert its therapeutic effect on patients with NMOSD through B-cell–mediated humoral immunity[17,18] and has been shown to be superior to azathioprine for NMOSD in 1 open-label RCTs.[19] Azathioprine is a purine analogue that interferes with DNA synthesis of rapidly proliferating cells. It has been widely used as a first-line immunosuppressant medication for autoimmune diseases.[17] Azathioprine was first studied in patients with NMOSD in 1998, where it was found to have a benefit in reducing disability.[17] Gastrointestinal and hematological side effects are associated with its use. Mycophenolate mofetil was developed to be a specific immunosuppressive drug with limited side effects by targeting guanosine more than adenosine.[17] Mycophenolate mofetil is widely used as an immunosuppressant for the treatment of autoimmune diseases and NMOSD, with fewer adverse effects than other therapies, such as azathioprine.[17]

Eculizumab was the first drug approved in Canada for the prevention of relapses based on data from a phase III RCT in adults with NMOSD who were AQP4 IgG positive.[20] In August 2020, the CADTH Canadian Drug Expert Committee (CDEC) recommended that eculizumab be reimbursed for the treatment of NMOSD if specific clinical and pricing criteria

were met.[21] Table 4 provides a summary of key drugs used to prevent relapses in patients with NMOSD.

There is no cure for NMOSD. In their input to CADTH, patients expressed that the currently available therapies for NMOSD only offered a temporary solution. Patients voiced the need for a new drug that reduces relapses and disability and is easy to administer, as currently available therapeutics fail to do so.

Drug

Satralizumab is a humanized immunoglobulin G2 monoclonal antibody that acts as an immunosuppressant by binding to the interleukin-6 receptor, thereby preventing downstream signaling through this receptor.[9] It is indicated as monotherapy or in combination with immunosuppressive therapy for the treatment of NMOSD in adult and adolescent patients who are AQP4 IgG positive.[9] Satralizumab is not intended for acute treatment of an NMOSD relapse.[9]

Satralizumab is available as a 120 mg/mL single-use, pre-filled syringe, and the recommended dose is 120 mg by SC injection at weeks 0, 2, and 4 and then every 4 weeks thereafter.[9]

Satralizumab underwent priority review by Health Canada and received a Notice of Compliance on June 1, 2020. The sponsor has requested reimbursement as per the indication.

Table 4: Key Characteristics of Satralizumab, Eculizumab, Rituximab, Azathioprine, and Mycophenolate

Characteristic	Satralizumab	Eculizumab	Rituximab	Azathioprine	Mycophenolate
Mechanism of action	Monoclonal antibody that blocks interleukin-6 receptor	Monoclonal antibody that specifically binds to the complement protein C5	Monoclonal antibody that specifically binds to the transmembrane antigen CD20	Immuno-suppressant	Immuno-suppressant
Indication[a]	Treatment of NMOSD (monotherapy or with IST) in patients ≥ 12 years of age who are AQP4 IgG seropositive; satralizumab is not intended for acute treatment of an NMOSD relapse	Treatment of NMOSD in adults who are AQP4 IgG seropositive; eculizumab is not intended for acute treatment of an NMOSD relapse	No Health Canada indication for the treatment of NMOSD	No Health Canada indication for the treatment of NMOSD	No Health Canada indication for the treatment of NMOSD
Route of administration	SC	IV	IV	Oral	Oral, IV

Characteristic	Satralizumab	Eculizumab	Rituximab	Azathioprine	Mycophenolate
Recommended dosage	120 mg by SC injection at weeks 0, 2, and 4, followed by 120 mg every 4 weeks	900 mg IV weekly for the first 4 weeks followed by 1,200 mg for the fifth dose 1 week later, then 1,200 mg every 2 weeks	Rheumatoid arthritis protocol: 1,000 mg IV infusion, followed 2 weeks later by the second 1,000 mg IV infusion Lymphoma protocol: 375 mg/m^2 IV infusion weekly for 4 weeks	2 mg/kg per day to 3 mg/kg per day	Myfortic: 720 mg (four 180 mg or two 360 mg tablets) administered twice daily (1.440 g total daily dose) Cellcept: 1 g to 3 g daily, administered orally or intravenously twice a day
Serious adverse effects or safety issues	Infections Monitor liver enzymes and neutrophils	Serious or fatal meningococcal infections	Infusion reactions, progressive multifocal leukoencephalopathy, tumour lysis syndrome, hepatitis B virus, mucocutaneous reactions, infections, cardiovascular events	Leukopenia, thrombocytopenia, macrophage activation syndrome, infection, carcinogenic, hepatotoxicity, fetal harm	Infection, lymphoma, fetal harm

AQP4 = aquaporin 4; IgG = immunoglobulin G; IST = immunosuppressive therapy; NMOSD = neuromyelitis optica spectrum disorder; SC = subcutaneous.

[a] Health Canada–approved indication.

Source: Product monograph for Enspryng,[9] Soliris,[22] Rituxan,[23] Imuran,[24] Myfortic,[25] and Cellcept.[26]

Stakeholder Engagement

Patient Group Input

This section was prepared by CADTH staff based on the input provided by patient groups.

About the Patient Groups and Information Gathered

The MS Society of Canada was the sole patient group to provide input for CADTH's review of satralizumab. It aims to provide support and services for patients living with MS and allied diseases, such as NMOSD. This patient group provides programs and services for patients living with NMOSD and, more broadly, advocates and funds research for those living with MS.

A disclosure of any conflicts of interest for the MS Society of Canada is available on the CADTH website. The MS Society of Canada gathered information for its submission through an online survey that was available in English and French on the MS Society of Canada website and its Facebook page in October 2020. Additionally, this survey was shared with the Guthy-Jackson Charitable Foundation in the US, which supports patients living with NMOSD and funds research for this disease. In total, there were 37 survey respondents, 86% of whom were female; 40% were aged 55 to 64 years, 27% were aged 45 to 54 years, and 27% were aged 35 to 44 years. Approximately two-thirds (25 respondents; 68%) were diagnosed with NMOSD; 4 respondents were patients with MS and 1 respondent was a caregiver for a patient with MS. Of the respondents who disclosed their disease duration, 2 respondents had been living with NMOSD for less than 2 years, 6 had been living with NMOSD for between 2 and 4 years, and 14 and 3 respondents had been living with this disease between 5 and 10 years, or between 11 and 20 years, respectively.

Disease Experience

If a patient is experiencing a relapse, inflammatory attacks on the optic nerves cause swelling, pain, and loss of vision, while damage to the spinal cord causes weakness or paralysis in the legs and arms associated with loss of sensation and bladder and bowel control problems. With every relapse, damage accrued to the optic nerve and/or spinal cord causes accumulating disability. Every day, patients struggle with "pain and fatigue," "pins and needles sensations," "poor bladder control," inability to stay in high temperature settings for a prolonged period of time, blindness, and "limited daily activities and work."

Due to the nature of the disease, the MS Society of Canada indicates that patients experience employment instability or loss, increased need for assistance or caregiving, loss of independence, isolation, cognitive decline, and increased mobility challenges. Nine of the 11 patient quotes presented by the MS Society of Canada mentioned they either had to quit their job or required assisted services in their work. Regarding the impact of NMOSD, 1 respondent stated, "I remain positive, but it's made a massive impact. I have poor bladder control, can't walk unassisted and am currently on sick leave from work." Another patient said: "I'm unable to work, I'm constantly in pain. The past 5 years has been very difficult."

Experience With Treatment

Prior to 2019, there were no treatments with a Health Canada indication for treatment of NMOSD; eculizumab was the first therapy indicated for this condition. Of the respondents who identified their drug treatments, 13 said they had experience with rituximab, 6 had previously taken steroids, 3 had taken azathioprine, and 1 respondent had experience with eculizumab. The MS Society of Canada stated that treatment with IV immunoglobulin or plasma exchange has been used to remove antibodies. Moreover, medications and therapies are used by patients to treat symptoms such as neuropathy, other pain, stiffness, muscle spasms, and bladder and bowel control problems.

Regarding the effectiveness of current therapies, 14 respondents felt their therapy was effective, 2 reported no perceived effectiveness, and 7 respondents reported they did not know whether their therapy was effective. Respondents identified concerns regarding limited access to treatments, as off-label treatments are not always covered through private or public health plans.

None of the respondents had experience with satralizumab.

Improved Outcomes

Patients reported value in the availability of another approved treatment for NMOSD. They stated that the administration of satralizumab (self-administered as a once-monthly SC injection) fills an unmet need compared with eculizumab (administered via infusion in a specialized clinic every 2 weeks). Patients hope the new treatment will reduce attacks and reduce disability, which was highlighted as an important therapeutic gap for current drug treatments for NMOSD.

The MS Society of Canada indicated that treatment with satralizumab has the potential to allow people living with NMOSD to remain in the workforce, sustain family and social roles and responsibilities longer, improve their quality of life, and decrease the need for caregiving (family caregiver or paid caregiver).

Clinician Input

All CADTH review teams include at least 1 clinical specialist with expertise in the diagnosis and management of the condition for which the drug is indicated. Clinical experts are a critical part of the review team and are involved in all phases of the review process (e.g., providing guidance on the development of the review protocol, assisting in the critical appraisal of clinical evidence, interpreting the clinical relevance of the results, and providing guidance on the potential place in therapy). In addition, as part of the satralizumab review, a panel of 4 clinical experts from across Canada was convened to characterize unmet therapeutic needs, assist in identifying and communicating situations where there are gaps in the evidence that could be addressed through the collection of additional data, promote the early identification of potential implementation challenges, gain further insight into the clinical management of patients living with a condition, and explore the potential place in therapy of the drug (e.g., potential reimbursement conditions). A summary of this panel discussion is presented below.

Unmet Needs

NMOSD is a severe, debilitating disease characterized by relapses that may result in permanent and life-changing neurological deficits. Off-label immunosuppressants such as azathioprine, mycophenolate mofetil, and rituximab are used to prevent relapses, although the treatment approach differs by province and territory, in part due to differential access to these drugs. These immunosuppressants have significant failure rates, and patients will still experience severe relapses that cause them to accrue disability. Non-specific immunosuppressants also have safety concerns, particularly with longer-term use and in younger patients.

Eculizumab (Soliris) was the first drug approved in Canada for patients with NMOSD who are AQP4 antibody positive, but it is limited by the inconvenience of its IV route of administration and concerns regarding potential adverse effects, including an increased risk of meningococcal infection.

The experts emphasized the need for safe and effective relapse-prevention treatments for patients with NMOSD, as early intervention to eliminate relapses is key to averting disability and improving longer-term outcomes for patients.

Place in Therapy

The panel indicated that satralizumab could be used as a first-line treatment for patients who are anti–AQP4 antibody positive but could also be used after inadequate response or intolerance of immunosuppressants. The panel acknowledged that comparative clinical data are not available at this time to optimally guide the position of satralizumab in the treatment algorithm.

Satralizumab could be used as a monotherapy or as an add-on to corticosteroids and other immunosuppressants. It is not intended for the management of acute relapses.

Patient Population

The clinical experts stated that satralizumab is suitable for patients with a confirmed diagnosis of NMOSD who are anti–AQP4 antibody positive. Additional patient characteristics that might favour the use of satralizumab over existing therapies may include patients who have had a severe initial attack, patients who have experienced multiple or severe relapses, or patients who had a relapse while receiving immunosuppressant therapy.

A diagnosis of NMOSD requires input from a specialist in MS and demyelinating diseases and is based on international diagnostic criteria. Due to greater understanding of and awareness of NMOSD among clinicians, misdiagnosis or delays in diagnosis are less common in Canada. However, the experts expressed some concerns regarding the test accuracy for anti-AQP4 antibodies in Canada, which may prevent patients with a false-negative test from accessing satralizumab.

Assessing Response to Treatment

Treatment response is determined by the elimination of relapses or a decrease in the frequency or severity of relapses. The experts stated that relapses are identified clinically, based on a neurological examination and patient-reported symptoms. An MRI may also be used to assess relapses.

There is no formal guidance on how often treatment response should be assessed. Frequency may vary from every 3 to 4 months after initiating therapy to every 6 to 12 months with maintenance therapy.

Discontinuing Treatment

According to the clinical experts, failure to observe an improvement in relapse frequency or severity may prompt a treatment change. In particular, a recurrence of relapses that is severe and associated with a meaningful change in EDSS score or functionality may warrant a change in therapy. In addition, treatment may be stopped if adverse effects are intolerable. Although drug discontinuation may be considered for patients with substantial deficits, the experts stated that continuation of treatment may be warranted for those who still have function that may contribute to their independence.

The use of the EDSS scoring system to determine treatment response has well-recognized limitations, as this instrument does not fully capture the disability associated with NMOSD; it is relatively insensitive to changes in non-ambulatory disability, particularly in the upper end of the range. For example, changes in visual acuity or loss of the use of upper limbs may not be captured by the EDSS scoring system for individuals who already require a walker, but it can have a profound impact on patients' functioning and quality of life.

Prescribing Conditions

The clinical experts agreed that diagnosis of NMOSD and prescribing of satralizumab should be limited to neurologists with specific expertise in NMOSD or demyelinating disorders. In some circumstances, subsequent care may be managed by a general neurologist.

Clinical Evidence

The clinical evidence included in the review of satralizumab is presented in 3 sections. The first section, the systematic review, includes pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those studies that were selected according to an a priori protocol. The second section includes indirect evidence from the sponsor and indirect evidence selected from the literature that met the selection criteria specified in the review. The third section includes sponsor-submitted long-term extension studies and additional relevant studies that were considered to address important gaps in the evidence included in the systematic review.

Systematic Review (Pivotal and Protocol Selected Studies)

Objectives

To perform a systematic review of the beneficial and harmful effects of satralizumab as monotherapy or in combination with immunosuppressive therapy for the treatment of NMOSD in adult and adolescent patients who are anti–AQP4 antibody positive.

Methods

Studies selected for inclusion in the systematic review included pivotal studies provided in the sponsor's submission to CADTH and Health Canada, as well as those meeting the selection criteria presented in Table 5.

Table 5: Inclusion Criteria for the Systematic Review

Patient population	Adult and adolescent patients (≥ 12 years) with NMOSD who are anti–AQP4 antibody positive. Subgroups: • treatment experience • mobility-related impairment versus vision-related impairment
Intervention	Satralizumab 120 mg pre-filled syringe for subcutaneous injection at weeks 0, 2, and 4 then every 4 weeks, alone or in combination with immunosuppressive therapy
Comparators	One or more of the following treatments: • eculizumab • rituximab[a] • azathioprine[a] • mycophenolate mofetil[a] • tocilizumab[a] • methotrexate[a] • cyclophosphamide[a] • mitoxantrone[a] • cyclosporine[a] • prednisone[a] • bortezomib[a] • placebo
Outcomes	**Efficacy outcomes** • Relapse (e.g., time to first relapse, relapse rate) • Disability (e.g., worsening neurologic disability, visual acuity)[b] • HRQoL[b] • Productivity (e.g., attend school or work)[b]

	• Symptoms (e.g., pain, fatigue, bladder or bowel function, sexual dysfunction, respiratory)[b] • Health care resource utilization **Harms outcomes** AEs, SAEs, WDAEs, mortality, infections, hepatotoxicity, injection-site AEs, hypersensitivity
Study design	Published and unpublished phase III and IV RCTs

AE = adverse event; AQP4 = aquaporin 4; HRQoL = health-related quality of life; NMOSD = neuromyelitis optica spectrum disorder; RCT = randomized controlled trial; SAE = serious adverse event; WDAE = withdrawal due to adverse event.

[a] This drug does not have a Health Canada indication for the treatment of patients with NMOSD.

[b] These outcomes were identified as being of particular importance to patients in the input received by CADTH from patient groups.

The literature search for clinical studies was performed by an information specialist using a peer-reviewed search strategy according to the *PRESS Peer Review of Electronic Search Strategies* checklist (https://www.cadth.ca/resources/finding-evidence/press).[27]

Published literature was identified by searching the following bibliographic databases: MEDLINE All (1946–) through Ovid and Embase (1974–) through Ovid. The search strategy was comprised of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were Enspryng (satralizumab). Clinical trials registries were searched: the US National Institutes of Health's clinicaltrials.gov, World Health Organization's International Clinical Trials Registry Platform (ICTRP) search portal, Health Canada's Clinical Trials Database, and the European Union Clinical Trials Register.

No filters were applied to limit the retrieval by study type. Retrieval was not limited by publication date or by language. Conference abstracts were excluded from the search results. See Appendix 1 for the detailed search strategies.

The initial search was completed on November 19, 2020. Regular alerts updated the search until the meeting of CDEC on March 17, 2021.

Grey literature (literature that is not commercially published) was identified by searching relevant websites from the *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* checklist (https://www.cadth.ca/grey-matters).[28] Included in this search were the websites of regulatory agencies (US Food and Drug Administration and European Medicines Agency). Google was used to search for additional internet-based materials. See Appendix 1 for more information on the grey literature search strategy.

Two CADTH clinical reviewers independently selected studies for inclusion in the review based on titles and abstracts, according to the predetermined protocol. Full-text articles of all citations considered potentially relevant by at least 1 reviewer were acquired. Reviewers independently made the final selection of studies to be included in the review, and differences were resolved through discussion.

Findings From the Literature

A total of 2 studies were identified from the literature for inclusion in the systematic review (Figure 1). The included studies are summarized in Table 6. A list of excluded studies is presented in Appendix 2.

Figure 1: Flow Diagram for Inclusion and Exclusion of Studies

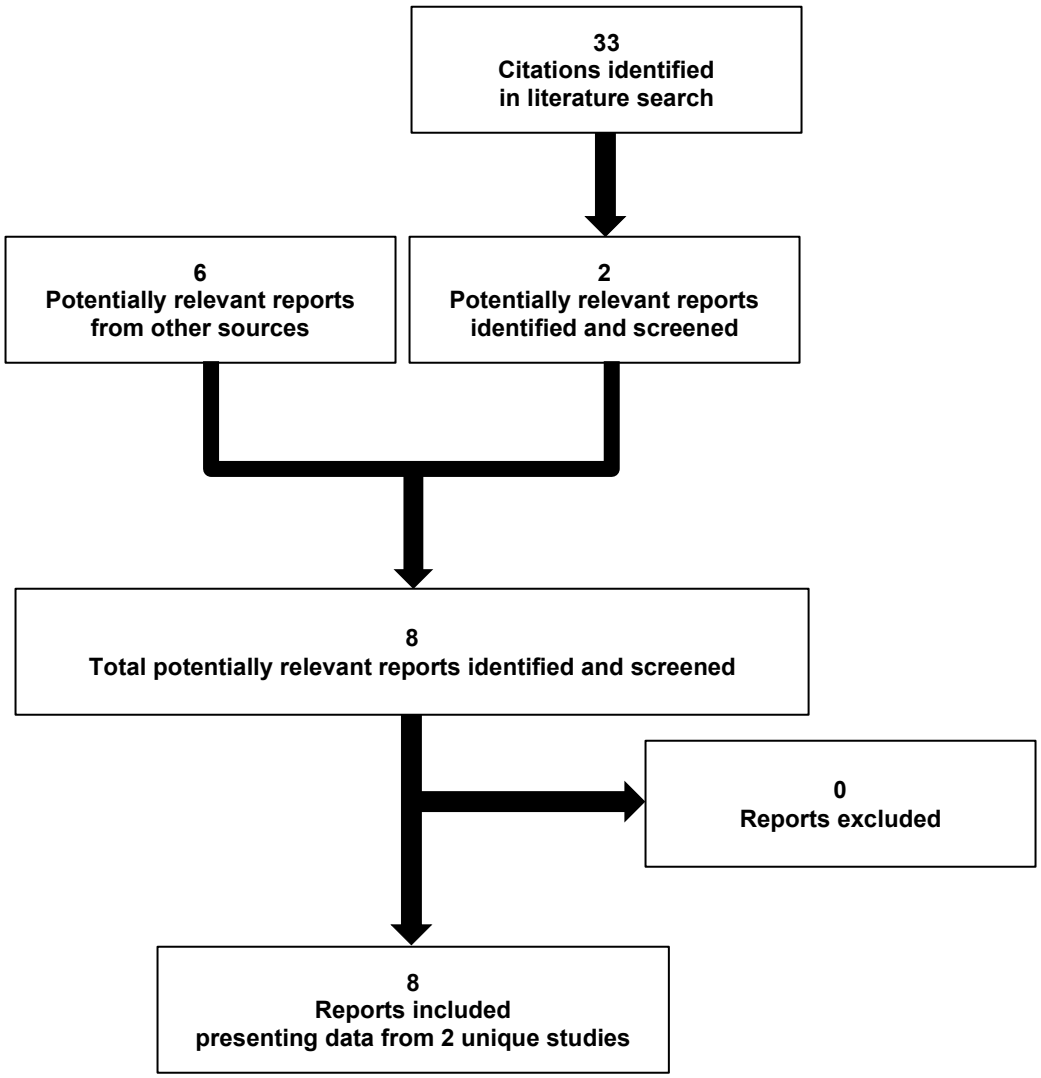


Table 6: Details of Included Studies

	Detail	Study 898 (BN40898 or SAkuraSky)	Study 900 (BN40900 or SAkuraStar)
Designs and Populations	**Study design**	DB RCT	DB RCT
	Locations	Europe, Asia, US	US, Canada, Asia, Europe
	Patient enrolment dates	February 20, 2014 to June 6, 2018	August 5, 2014 to April 2, 2017
	Data cut-off	June 6, 2018	October 12, 2018
	Randomized (N)	83	95
	Inclusion criteria	Patients (12 to 74 years) with NMOSD (Wingerchuck [2007] criteria) with anti-AQP4 antibodies at screening or NMO (defined by Wingerchuck [2006] criteria): • at least 2 relapses in the last 2 years, 1 of which occurred in the past 12 months • EDSS score of 0 to 6.5 • receiving azathioprine, mycophenolate mofetil, or oral corticosteroids (for adults), or corticosteroids plus azathioprine or mycophenolate mofetil (in adolescents), at stable doses for 8 weeks prior to baseline • maximum 30% were anti–AQP4 antibody negative	Adults (18 to 74 years) with NMSOD (Wingerchuck [2007] criteria) with anti-AQP4 antibodies at screening or NMO (defined by Wingerchuck [2006] criteria): • at least 1 relapse in the last 12 months (including the first attack) • EDSS score of 0 to 6.5 • maximum 30% were anti–AQP4 antibody negative
	Exclusion criteria	• Prior treatment with IL-6 inhibitor (e.g., tocilizumab, alemtuzumab, total body irradiation, or bone marrow transplant) • Prior treatment with anti-CD20, eculizumab, belimumab, interferon, natalizumab, glatiramer acetate, fingolimod, teriflunomide, or dimethyl fumarate within 6 months prior to baseline • Prior treatment with anti-CD4, cladribine, or mitoxantrone within the past 2 years • Surgery within the past 4 weeks • Active infection within the past 4 weeks, active tuberculosis, chronic active hepatitis B or C, active interstitial lung disease, diverticulitis • Other demyelinating disease or progressive multifocal leukoencephalopathy • Low WBC, neutrophil, lymphocyte, or platelet count; elevated liver enzymes (> 1.5 ULN) • Received live or live-attenuated vaccine within 6 weeks • Drug or alcohol abuse within 1 year prior • Suicidal ideation within 6 months or suicide attempt in past 3 years • Malignancy in past 5 years • Severe allergic reaction to biologic drug • Other serious uncontrolled disease that may preclude participation	• Relapse onset within 30 days prior to baseline • Prior treatment with IL-6 inhibitor (e.g., tocilizumab, alemtuzumab, total body irradiation, or bone marrow transplant) • Prior treatment with anti-CD20, eculizumab, belimumab, interferon, natalizumab, glatiramer acetate, fingolimod, teriflunomide, or dimethyl fumarate within 6 months prior to baseline • Prior treatment with anti-CD4, cladribine, cyclophosphamide or mitoxantrone within the past 2 years • Surgery within the past 4 weeks • Active infection within the past 4 weeks, active tuberculosis, chronic active hepatitis B or C, active interstitial lung disease, diverticulitis • Other demyelinating disease or progressive multifocal leukoencephalopathy • Low WBC, neutrophil, lymphocyte, or platelet count; elevated liver enzymes (> 1.5 ULN) • Received live or live-attenuated vaccine within 6 weeks • Drug or alcohol abuse within 1 year prior • Suicidal ideation within 6 months or suicide attempt in past 3 years • Malignancy in past 5 years • Severe allergic reaction to biologic drug • History of Stevens-Johnson syndrome • Other serious uncontrolled disease that may preclude participation

	Detail	Study 898 (BN40898 or SAkuraSky)	Study 900 (BN40900 or SAkuraStar)
Drugs	**Intervention**	Satralizumab 120 mg SC at weeks 0, 2, and 4, then every 4 weeks plus baseline IST	Satralizumab 120 mg SC at weeks 0, 2, and 4, then every 4 weeks
	Comparator(s)	Placebo SC injection plus baseline IST	Placebo SC injection
Duration	**Phase**		
	Screening	28 days	28 days
	DB	Until 26 protocol-defined relapse events	Until 44 protocol-defined relapse events or 1.5 years after the last patient was enrolled
	Follow-up	Safety follow-up in adults: up to 24 weeks (1 year for adolescents)	Safety follow-up: up to 24 weeks
Outcomes	**Primary end point**	Time to first protocol-defined relapse event	Time to first protocol-defined relapse event
	Secondary and exploratory end points	Key secondary: • change from baseline to week 24 in pain VAS • change from baseline to week 24 in FACIT-F Other secondary: • change from baseline in SF-36, EQ-5D-3L • ARR • proportion of relapse-free patients • change in modified Rankin Scale • change in Zarit Burden Interview • change in EDSS • change in visual acuity (Snellen chart) Harms	Key secondary: • change from baseline to week 24 in pain VAS • change from baseline to week 24 in FACIT-F Other secondary: • change from baseline in SF-36, EQ-5D-3L • ARR • proportion of relapse-free patients • change in modified Rankin Scale • change in Zarit Burden Interview • change in EDSS • change in visual acuity (Snellen chart, low-contrast visual acuity) • Timed 25-foot walk test Harms
Notes	**Publications**	Yamamura et al. (2019)[29]	Traboulsee et al. (2020)[30]

AQP4 = aquaporin 4; ARR = annualized relapse rate; DB = double blind; EDSS = Expanded Disability Status Scale; EQ-5D-3L = EuroQol 5-Dimensions 3-Levels questionnaire; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; IL-6 = interleukin-6; IST = immunosuppressive therapy; NMO = neuromyelitis optica; NMOSD = neuromyelitis optica spectrum disorder; NR = not reported; RCT = randomized controlled trial; SC = subcutaneous; SF-36v2 = Short Form (36) Health Survey version 2; ULD = upper limit of normal; VAS = Visual Analogue Scale; WBC = white blood count.

Note: Four additional reports were included: CADTH Common Drug Review submission for Enspryng,[31] FDA Clinical Report,[32] FDA Statistical Report,[33] Health Canada Reviewer's Report.[34]

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Description of Studies

Two pivotal, double-blind randomized placebo-controlled trials met the inclusion criteria for the systematic review: Study BN40898 or SAkuraSky (referred to in this report as Study 898), and BN40900 or SAkuraStar (referred to in this report as Study 900) (Table 6).

The objective of Study 898 was to evaluate the efficacy and safety of satralizumab versus placebo, in addition to baseline immunosuppressant treatment, in adults and adolescent patients with NMO and NMOSD. This double-blind, parallel design trial randomized 83 patients (1:1) to satralizumab 120 mg or placebo via an interactive web or voice response system. Randomization was stratified by baseline ARR (ARR of 1; ARR > 1) and geographic region (Asia; Europe or other). The event-driven trial was to stop once 26

primary outcome relapse events were reported. Study 898 enrolled patients from 34 sites in Asia, Europe, and the US (no Canadian sites).

Study 900 was designed to evaluate the efficacy and safety of satralizumab monotherapy compared with placebo in adults with NMO or NMOSD. The double-blind parallel design trial randomized a total of 95 patients (2:1) to satralizumab 120 mg or placebo using an interactive web or voice response system, with randomization stratified by prior therapy received (B-cell–depleting therapy versus immunosuppressant or other treatments) and by the most recent attack (first attack versus relapse). The trial was to stop when 44 primary outcome relapse events had occurred or 1.5 years since the last patient was randomized. Study 900 was conducted in 44 sites in the US, Canada (3 sites; 11 patients),[32] Asia, and Europe.

Both RCTs included an open-label extension phase where all patients received satralizumab.

Populations

Inclusion and Exclusion Criteria

Study 898 and Study 900 enrolled patients who met the diagnostic criteria for NMOSD or NMO as listed in Appendix 3 (Table 24), and who had an EDSS score of 6.5 or less. Both studies enrolled patients who were AQP4 IgG positive or negative, with a maximum of 30% of patients who were seronegative (Table 6).

Study 898 included patients who were 12 to 74 years of age who had at least 2 relapses in the past year (1 of which occurred in the last 12 months) and who were receiving azathioprine, mycophenolate, and/or corticosteroids at stable doses for the past 8 weeks. In contrast, Study 900 enrolled adults 18 to 74 years old who had at least 1 relapse in the past year, including a first attack.

In both studies, the key exclusion criteria were active infection, interstitial lung disease, diverticulitis, other demyelinating disease or progressive multifocal leukoencephalopathy, drug or alcohol abuse within the past year, abnormal laboratory values, recent suicidal ideation, or recent surgery. Exclusion criteria related to drugs and vaccines are outlined in the interventions section.

This report will focus on the subgroup of patients who were AQP4 IgG positive, as this is the population that received Health Canada approval.

Baseline Characteristics

The demographics for the ITT and AQP4 IgG–positive subgroup were generally well balanced between groups in Study 898 (Table 7). Of the 83 patients enrolled in Study 898, 55 patients (66%) were included in the AQP4 subgroup. The mean age of patients in the AQP4 subgroup was [redacted] for the placebo and satralizumab groups, including [redacted] adolescent patients in the placebo group [redacted] and [redacted] adolescent patients in the satralizumab group [redacted]. [redacted] patients were female, [redacted] were Asian, and [redacted] were White in the placebo and satralizumab groups, respectively. There were fewer patients diagnosed according to NMO criteria (versus NMOSD criteria) in the placebo group [redacted] than in the satralizumab group [redacted]. The baseline mean ARR was [redacted] in both groups, and the median baseline EDSS score was [redacted] in the placebo group and [redacted] in the satralizumab group. More patients were receiving corticosteroids ([redacted] for placebo

and satralizumab groups, respectively) or azathioprine ([redacted]), than mycophenolate mofetil ([redacted]) or mycophenolate mofetil plus corticosteroids ([redacted]).

In Study 900, 23 of the 32 patients randomized to placebo (72%) and 41 of the 63 randomized to satralizumab (65%) were included in the AQP4 IgG–positive subgroup (Table 8). Some imbalances in the patient demographic and disease characteristics were present. In the placebo and satralizumab groups of the AQP4 IgG–positive subgroup, the mean age was 40.1 years (standard deviation [SD] = 11.5) and 46.0 years (SD = 12.0), respectively; 96% and 76% were female. Patients were White (57% in the placebo group and 46% in the satralizumab group), Asian (26% and 17%), Black (13% and 27%), or another race (4% and 10%). Most patients were diagnosed according to NMO criteria (64%) rather than NMOSD criteria (36%). Most patients had experienced more than 1 attack, with only 17% of patients receiving placebo and 12% of patients receiving satralizumab having been enrolled after their first attack. At baseline, the mean ARR was [redacted], and the median EDSS score was 3.5 (range = 1.0 to 6.5) and 4.0 (range =1.5 to 6.5) in the placebo and satralizumab groups, respectively. In the AQP4 IgG–positive subgroup, 17% and 12% had received prior B-cell–depleting therapy, and 83% and 88% had received either no prior therapy or other immunosuppressants. In the overall population of Study 900, [redacted] had not received prior immunosuppressant treatments for NMOSD relapse prevention and were considered treatment-naive, including [redacted] in the placebo group and [redacted] in the satralizumab group.[34] It is unclear what percentage of patients in the AQP4 IgG–positive subgroup were treatment-naive.

Table 7: Summary of Baseline Characteristics for Study 898

Characteristic	ITT population		AQP4 IgG–positive subgroup	
	Placebo plus IST N = 42	Satralizumab plus IST N = 41	Placebo plus IST N = 28	Satralizumab plus IST N = 27
Mean age, years (SD)	43.4 (12.0)	40.8 (16.1)	[redacted]	[redacted]
Age < 18 years, n (%)	3 (7)	4 (10)	[redacted]	[redacted]
Female, n (%)	40 (95)	37 (90)	[redacted]	[redacted]
Race, n (%)				
White	21 (50)	24 (59)	[redacted]	[redacted]
Asian	18 (43)	17 (41)	[redacted]	[redacted]
Black	2 (5)	0	[redacted]	[redacted]
Other	1 (2)	0	[redacted]	[redacted]
BMI (kg/m^2)	23.9 (5.9)	23.5 (4.9)	[redacted]	[redacted]
Diagnosis, n (%)				
NMO	28 (67)	33 (80)	[redacted]	[redacted]
NMOSD	14 (33)	8 (20)	[redacted]	[redacted]
AQP4 IgG positive, n (%)	28 (67)	27 (66)	[redacted]	[redacted]
Baseline ARR, mean (SD)	1.50 (0.60)	1.48 (0.63)	[redacted]	[redacted]
Baseline ARR, median (range)	1.5 (1.0 to 3.0)	1.5 (1.0 to 3.5)	[redacted]	[redacted]
Baseline EDSS, mean (SD)	3.6 (1.32)	3.8 (1.57)	[redacted]	[redacted]

Characteristic	ITT population		AQP4 IgG–positive subgroup	
	Placebo plus IST N = 42	Satralizumab plus IST N = 41	Placebo plus IST N = 28	Satralizumab plus IST N = 27
Baseline EDSS, median (range)	3.5 (1.5 to 6.5)	3.5 (1.0 to 6.5)	[redacted]	[redacted]
Baseline treatment, n (%)				
Azathioprine	13 (31)	16 (39)	[redacted]	[redacted]
Mycophenolate mofetil	8 (19)	4 (10)	[redacted]	[redacted]
Corticosteroids, oral	20 (48)	17 (42)	[redacted]	[redacted]
Azathioprine plus corticosteroids	0	3 (7)	[redacted]	[redacted]
Mycophenolate mofetil plus corticosteroids	1 (2)	1 (2)	[redacted]	[redacted]

AQP4 = aquaporin 4; ARR = annualized relapse rate; BMI = body mass index; EDSS = Expanded Disability Status Scale; IgG = immunoglobulin G; IST = immunosuppressive therapy; ITT = intention to treat; NMO = neuromyelitis optica; NMOSD = neuromyelitis optica spectrum disorder; SD = standard deviation.

Source: Clinical Study Report for Study 898.[10]

Table 8: Summary of Baseline Characteristics for Study 900

Characteristic	ITT population		AQP4 IgG–positive subgroup	
	Placebo N = 32	Satralizumab N = 63	Placebo N = 23	Satralizumab N = 41
Mean age, years (SD)	40.5 (10.5)	45.3 (12.0)	40.1 (11.5)	46.0 (12.0)
Female, n (%)	31 (97)	46 (73)	22 (96)	31 (76)
Race, n (%)				
White	22 (69)	37 (59)	13 (57)	19 (46)
Asian	6 (19)	8 (13)	6 (26)	7 (17)
Black	3 (9)	13 (21)	3 (13)	11 (27)
Other	1 (3)	5 (8)	1 (4)	4 (10)
BMI (kg/m^2)	26.2 (7.0)	28.5 (8.6)	24.9 (6.0)	28.5 (8.9)
Diagnosis, n (%)				
NMO	24 (75)	47 (75)	15 (65)	26 (63)
NMOSD	8 (25)	16 (25)	8 (35)	15 (37)
AQP4 IgG positive, n (%)	23 (72)	41 (65)	23 (100)	41 (100)
Baseline ARR, mean (SD)[a]	1.05 (0.50)	0.94 (0.48)	[redacted]	[redacted]
Baseline ARR, median (range)[a]	1.0 (0.5 to 2.5)	1.0 (0.5 to 2.5)	[redacted]	[redacted]
Baseline EDSS, mean (SD)	3.7 (1.6)	3.9 (1.5)	3.4 (1.6)	4.0 (1.5)
Baseline EDSS, median (range)	3.5 (1.0 to 6.5)	4.0 (1.5 to 6.5)	3.5 (1.0 to 6.5)	4.0 (1.5 to 6.5)
Most recent attack, n (%)				
First attack	4 (13)	7 (11)	4 (17)	5 (12)
Relapse	28 (88)	56 (89)	19 (83)	36 (88)
Prior therapy, n (%)				

Characteristic	ITT population		AQP4 IgG–positive subgroup	
	Placebo N = 32	Satralizumab N = 63	Placebo N = 23	Satralizumab N = 41
B-cell–depleting therapy	4 (13)	8 (13)	4 (17)	5 (12)
Immunosuppressant/ other[b]	28 (88)	55 (87)	19 (83)	36 (88)

AQP4 = aquaporin 4; ARR = annualized relapse rate; BMI = body mass index; EDSS = Expanded Disability Status Scale; IgG = immunoglobulin G; ITT = intention to treat; NMO = neuromyelitis optica; NMOSD = neuromyelitis optica spectrum disorder; SD = standard deviation.

[a] Baseline ARR derived using relapses having imputed onset dates 2 years prior to screening.[34]

[b] The immunosuppressant/other category for the ITT population includes 10 patients (31%) in the placebo group and 23 (37%) in the satralizumab group who had not received prior immunosuppressant treatments for NMOSD relapse prevention and were considered treatment-naive.

Source: Clinical Study Report for Study 900[11] and Health Canada Reviewer's Report.[34]

Interventions

Study Drug

In both studies, patients received placebo or satralizumab 120 mg by SC injection at weeks 0, 2, and 4 and then every 4 weeks, administered by the investigator at the study site. Patients were monitored at the study centre for at least 1 hour post dose.

The patients, investigators, study staff, and sponsor were blinded to treatment allocation. Vials of placebo solution for injection were identical in composition, colour, appearance, and packaging to the satralizumab solution for injection. To maintain blinding to treatment allocation, some laboratory results remained blinded to site staff, study monitor, or the sponsor (e.g., satralizumab serum concentration, C-reactive protein, interleukin-6). Fibrinogen levels were not blinded to the investigator involved in patient care or the safety monitor.

In both studies, treatment was stopped or interrupted if the patient developed the following: serious infection, decreased or persistent neutropenia, low platelet count or thrombocytopenia, elevated transaminases, anaphylactic or other serious hypersensitivity reaction, or malignancy. Patients were withdrawn if they missed 3 consecutive doses of the study drug, became pregnant, or experienced unacceptable toxicity. Patients who stopped study drug treatment were followed until the withdrawal visit and were then discontinued from the study.

Concurrent Medications

In Study 898, all patients continued on background immunosuppressant therapy during the trial. Stable doses of 1 of the following treatments were required: azathioprine (maximum 3 mg/kg per day); mycophenolate mofetil (maximum 3,000 mg per day); oral corticosteroids (maximum 15 mg per day prednisone equivalent). Combination therapy of oral corticosteroids with azathioprine or mycophenolate mofetil was allowed for adolescents. Dose reductions of background therapies were allowed for safety reasons but an increase in dose or change in treatment was not permitted. In this trial, patients were prohibited from receiving immunosuppressants other than azathioprine and mycophenolate mofetil (from 8 weeks prior to baseline).

During Study 900, patients were prohibited from receiving immunosuppressants (e.g., azathioprine, cyclosporine, methotrexate, mycophenolate mofetil, tacrolimus), and corticosteroids or IV immunoglobulin (except for rescue therapy for clinical relapse).

In both studies, patients were prohibited from receiving live or live-attenuated vaccines within 6 weeks of baseline and during the trials. Patients were required to stop any previous treatment with anti-CD20, eculizumab, anti–B lymphocyte stimulator monoclonal antibody (e.g., belimumab), and any other treatment for the prevention of MS relapse at least 6 months prior to baseline. Other prohibited medications included the following: other investigational drugs within the last 3 months; anti-CD4, cladribine, cyclophosphamide, or mitoxantrone within the past 2 years; interleukin-6 inhibitory therapy (e.g., tocilizumab); alemtuzumab, total body irradiation, or bone marrow transplant in the patient's lifetime.

In both studies, pain medications, such as pregabalin, gabapentin, carbamazepine, clonazepam, duloxetine, and tramadol plus acetaminophen were allowed, with dose or drug adjustments permitted if pain control was insufficient or for safety reasons.

Rescue Therapy

Patients in both studies who experienced a clinical or protocol-defined relapse were permitted to receive rescue therapy. Rescue therapies included IV corticosteroids, IV immunoglobulin, and/or apheresis (plasma exchange and plasmapheresis). Study 900 also allowed oral corticosteroids as rescue therapy for tapering.

In Study 898, patients who received rescue therapy for either a clinical or protocol-defined relapse were to stop the blinded study drug and enter the extension period during which they received open-label satralizumab (31 to 60 days after the relapse onset once the disease had stabilized).

In Study 900, patients who received rescue therapy for a protocol-defined relapse were entered into the extension period and received open-label satralizumab. However, patients who experienced a relapse who did not meet the criteria for a protocol-defined relapse were allowed to continue in the double-blind period. The sponsor stated that this change was made based on comments from the FDA in order to minimize dropouts from the double-blind period.

In Study 898, a change in background therapy was allowed for patients who experienced a relapse. Permitted maintenance treatments included: azathioprine up to 3 mg/kg per day with or without oral corticosteroids (up to 1 mg/kg per day prednisone equivalent), mycophenolate mofetil up to 3,000 mg per day with or without oral corticosteroids (up to 1 mg/kg per day prednisone equivalent), or oral corticosteroids up to 1 mg/kg per day (prednisone equivalent).

Outcomes

A list of efficacy end points identified in the CADTH review protocol that were assessed in the clinical trials included in this review is provided in Table 9. These end points are further summarized below. A detailed discussion and critical appraisal of the outcome measures is provided in Appendix 4.

Table 9: Summary of Outcomes of Interest Identified in the CADTH Review Protocol

Outcome measure	Study 898	Study 090
Time to protocol-defined relapse	Primary	Primary
Change from baseline to week 24 in pain VAS	Key secondary	Key secondary
Change from baseline to week 24 in FACIT-F score	Key secondary	Key secondary
Change from baseline in SF-36v2 PCS, MCS, and domains	Other secondary	Other secondary
Change in EQ-5D-3L index score	Other secondary	Other secondary
ARR (based on first relapse event)	Other secondary	Other secondary
Change in modified Rankin Scale score	Other secondary	Other secondary
Change in EDSS score	Other secondary	Other secondary
Change in visual acuity (Snellen chart and low-contrast Sloan letter chart)	Other secondary	Other secondary

ARR = annualized relapse rate; EDSS = Expanded Disability Status Scale; EQ-5D-3L = EuroQol 5-Dimensions 3-Levels questionnaire; FACIT-F = Functional Assessment of Chronic Illness–Fatigue; MCS = mental component score; PCS = physical component score; SF-36v2 = Short Form (36) Health Survey version 2; VAS = Visual Analogue Scale.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

The primary outcome in both trials was protocol-defined relapse, confirmed by the clinical event committee. A protocol-defined relapse was any new or worsening neurological symptoms attributable to NMO or NMOSD that persisted for a minimum of 24 hours and that were not attributable to confounding clinical factors. Symptoms that recurred within 31 days were considered part of the same relapse. The relapse event had to meet 1 of the following:

- An increase of at least:
  - 1.0 point on the EDSS score, or a 2.0 point increase if the baseline EDSS was zero
  - 2.0 points on 1 of the appropriate FSS
  - 1.0 point on 2 or more of the appropriate FSS if the baseline score was 1 or more
  - 1.0 point in single eye FSS when the baseline score in that eye was 1 or more

The most recent EDSS and FSS score was the basis of comparison for assessing the change score. A qualifying FSS change was one that affected at least 1 of the following functional systems: pyramidal, cerebellar, brainstem, sensory, bowel or bladder, or visual (single eye). Sexual dysfunction and cerebral function did not suffice to establish a protocol-defined relapse. To be included in the primary analysis, relapses had to be assessed within 7 days of the patient reporting symptoms to the study site. In addition to study visits, patients were contacted weekly by phone and were asked about any change in symptoms or other signs of a potential relapse.

In both studies, the blinded clinical event committee consisted of 3 neurologists and/or ophthalmologists who had expertise in the diagnosis and treatment of NMOSD, expertise in EDSS assessment and scoring, and experience with MS clinical trials. Separate blinded outcome assessors, who were not involved in patient care, were used to evaluate the EDSS and FSS for the primary efficacy measure. Patients were asked to discuss with the outcome assessors only those symptoms related to the EDSS and FSS. Treating investigators managed patient care and monitored safety.

Sensitivity analyses were conducted using alternate definitions of relapse. A clinical relapse was defined as any relapse reported by the investigator. Treated clinical relapse included

any clinical relapse that resulted in the patient receiving rescue treatment. A treated clinical relapse of optic neuritis was defined as a relapse treated with rescue therapy and judged by the sponsor to be optic neuritis.

The key secondary outcomes in both trials was the change from baseline to week 24 in the pain VAS and FACIT-F scale.

The VAS for pain captures the self-rating of the current intensity of pain using a visual "thermometer," 100 mm in length, that ranges from no pain (best imaginable health state) to pain as bad as it could be (worst imaginable health state). Limited data on the reliability of the VAS as a subjective measure of pain was identified in patients with MS, with no information on reliability, validity, responsiveness, or minimal important difference (MID) found in patients with NMOSD (Appendix 4).

The FACIT-F is a 13-item questionnaire that measures a patient's level of daily fatigue over the past week. Scores for each item are summed for an overall score that ranges from 0 to 52, where 0 is the worst possible score and 52 the best, which indicates less fatigue. There were no studies identified that evaluated the validity, reliability, responsiveness, or MID for the FACIT-F scale in patients with NMOSD or MS (Appendix 4).

Health-related quality of life was assessed using the Short Form (36) Health Survey version 2 (SF-36v2) and EuroQol 5-Dimensions 3-Levels questionnaire (EQ-5D-3L).

The SF-36 is a 36-item, general health status instrument that consists of 8 health domains: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health.[35] The SF-36 also provides 2 component summaries, the physical component summary and the mental component summary, derived from aggregating the 8 domains according to a scoring algorithm. All of the domain scores are based on a scale of 0 to 100, with higher scores indicating higher quality of life. No estimates of the MID were found for patients with NMOSD. In general use, a change of 2 points in the SF-36 physical component score and 3 points in the SF-36 mental component score indicates a clinically meaningful improvement, as determined by the patient.[36]

The EQ-5D-3L questionnaire is a generic, preference-based, health-related quality of life measure.[37] It includes 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension is divided into 3 levels (1, 2, 3) representing "no problems," "some problems," and "extreme problems," respectively. The 5 questions are scored and together contribute to an EQ-5D index (utility) score of between 0 and 1, where 0 represents death and 1 represents perfect health. In both studies, the EQ-5D index score was based on a US preference–weighted algorithm (range = −0.2 to 1). No MID for patients with NMOSD was identified; the MID range for the general population is 0.033 to 0.074.[38]

The change in EDSS score and modified Rankin Scale score was reported as a secondary outcome in both trials. The EDSS is a quantitative measure of disability that is based on a standard neurological examination. Disability in several functional systems is assessed (i.e., pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, cerebral, and other), with each system assigned an ordinal rating ranging from 0 to 5 or 6. These FSS scores are combined with information concerning gait and the use of assisted devices to assign an overall score. The EDSS is an ordinal scale that ranges from 0 points (normal neurological examination) to 10 points (death) that increases in half-point increments once an EDSS of 1.0 has been reached. EDSS steps 1.0 to 4.5 refer to people who are fully ambulatory, with

steps 5.0 to 9.5 defined by impairment to ambulation. The full scale is described in Table 27 (Appendix 4).

The modified Rankin Scale is a generic, clinician-reported scale for measuring the degree of disability or dependence in the daily activities of people who have suffered a neurological disability. The scale ranges from 0 (no disability) to 6 (death). No studies evaluating validity, reliability, or the MID in patients with NMOSD or MS were identified, but the instrument is reliable and has been validated in patients who have suffered a disability due to a stroke.[39]

Statistical Analysis

Both trials used similar methods to analyze outcomes. The statistical methods are summarized in Table 10 and Table 11.

The primary outcome of time to first protocol-defined relapse was analyzed using a 2-sided log-rank test stratified by randomization stratification factors. Randomization was stratified by baseline ARR (1 versus > 1) and geographical region (Asia versus Europe or other) in Study 898, and in Study 900 by prior therapy (B-cell–depleting therapy versus immunosuppressants or other) and most recent attack prior to baseline (patient's first attack versus relapse). The Kaplan-Meier method was used to plot the time to relapse. A Cox proportional hazard model (stratified by randomization factors) was used to calculate the HR and 95% CI. A number of sensitivity analyses were conducted for the time to first relapse using different definitions of relapse as described in Table 10.

Different censoring rules were applied in Study 898 and Study 900 in the time to relapse analyses. In Study 898, patients were censored at the earliest of the following: the end of the double-blind study, switching or increasing baseline treatment, receiving rescue therapy for clinical relapse, or the withdrawal visit (for those who left the study early). For each patient, the double-blind period ended the day before they received their first dose of satralizumab in the open-label extension period or until their withdrawal visit (which was up to 12 weeks after the last dose of the study drug). The study was to end once 26 protocol-defined relapses had occurred.

In Study 900, patients were censored at the following: upon discontinuation from the double-blind period, at the cut-off date (for patients continuing to the end of the study), or at the date they entered the open-label extension period after having a clinical relapse (prior to the implementation of protocol 5 [on November 5, 2015], which specified that patients could enter the extension only after an adjudicated protocol-defined relapse). This was an event-driven trial that initially planned to enroll 70 patients and to stop once 19 protocol-defined relapses had occurred. The sample size and stopping criteria were amended twice after enrolment began in August 2014. First, to increase the sample size to 90 patients and increase the number of primary relapse events to 44 (March 2016) and, second, to add another stopping criterion that allowed the study to terminate once the last patient randomized had been treated for 1.5 years (June 2018). According to the FDA Statistical Report, the sponsor justified the first change due to a higher-than-expected early relapse rate, and justified the second change based on a lower-than-expected relapse rate and a prolonged double-blind period.[33] The timing of these changes to the protocol is relevant to the potential unblinding events that are discussed in the Critical Appraisal section of this report.

The key secondary outcomes of change from baseline to week 24 in pain VAS and FACIT-F scores were analyzed using an analysis of covariance model (ANCOVA) with baseline measurement and randomization stratification factors as covariates and using baseline observation carried forward (BOCF) for missing data. Sensitivity analyses were conducted using alternate methods to handle missing data (Table 10).

The change from baseline in SF-36, EQ-5D, modified Rankin Scale, EDSS, and visual acuity (Study 900 only) results were analyzed using mixed model for repeated measures (MMRM) methods that included the baseline score, randomization stratification factors, week, treatment, and treatment-by-week interaction. These analyses included those patients who reported a baseline value and at least 1 post-baseline outcome measurement.

The ARR was estimated using the first protocol-defined relapse event for each patient as of the event or censor date. The ARR was calculated as the total number of relapses per group divided by the number of PYs of follow-up, with a 95% CI based on a Poisson distribution. For the comparison between groups, the ARR was analyzed using a negative binomial regression model with relapse number as the response variable, treatment group and randomization stratification factors as covariates, and log-transformed exposure time as an offset.

Relapse-free rates at 6 months were calculated based on Kaplan-Meier estimates.

The type I error rate was controlled in both studies using a serial gatekeeping method that included the primary outcome and 2 key secondary outcomes (pain VAS and FACIT-F) analyzed in hierarchical order for the ITT population. The P values were not to be presented if statistical significance was not met for an end point that was higher in the hierarchy. The authors state that no statistical inferences can be drawn from other secondary outcomes that were not part of the statistical hierarchy. Pre-planned subgroup analyses of interest to this review included: AQP4 IgG status at screening, baseline treatment (Study 898 only), and prior treatments (Study 900 only). Subgroup analyses were not part of the statistical testing hierarchy.

Study 898 was predicted to have 80% power (alpha of 0.05) based on a sample size of 70, randomized 1:1 with a total of 26 relapse events, and a 2-year withdrawal rate of 10%. These calculations assumed an HR of 0.335 for the time to first relapse for satralizumab versus placebo, and an annual hazard rate of 0.4184 for the distribution of time to first relapse in the placebo group.

Study 900 had 80% power to detect a difference between groups in the time to protocol-defined relapse based on a planned sample size of 90 (randomized 2:1) with 44 relapse events and a 2-year dropout rate of 10% (2-sided log-rank test, alpha of 0.05). The HR for satralizumab versus placebo was assumed to be 1.0 for the initial 2 months and then 0.25 afterward, and the distribution of the time to a protocol-defined relapse for the placebo group was assumed to follow an annual hazard rate of 1.1295.

Table 10: Statistical Analysis of Efficacy End Points in Study 898

End point	Statistical model	Adjustment factors	Sensitivity analyses
Study 898			
Time to first protocol-defined relapse	• Kaplan-Meier method • 2-sided log-rank test • Cox proportional hazards model	Stratified by baseline ARR (1 versus > 1) and geographical region (Asia versus Europe or other) No imputation for missing study visits	Time to first event sensitivity analyses were based on the same statistical model as the primary analysis: • clinical relapse • treated clinical relapse • protocol-defined relapse (based on EDSS or FSS change from baseline) • protocol-defined relapse regardless of 7-day EDSS or FSS assessment limit • clinical relapse of optic neuritis
Change in VAS pain score from baseline to week 24	ANCOVA (BOCF)	Baseline value; Stratified by baseline ARR (1 versus > 1) and geographical region (Asia versus Europe or other)	• ANCOVA with multiple imputation for missing data (random hot deck) • MMRM, including visit, treatment-by-visit interaction, baseline measurement, and stratification factors (baseline ARR, geographic region)
Change in FACIT-F score from baseline to week 24			
Proportion of relapse-free patients	Kaplan-Meier estimates	NR	NR
ARR	Negative binomial regression model with log exposure time as offset	Stratification factors (baseline ARR, geographical region)	NR
Change in SF-36v2 PCS, MCS, and domains	MMRM (patients with baseline value and at least 1 post-baseline value)	Baseline score, stratification factors (baseline ARR and geographical region), week, treatment, treatment-by-week interaction	NR
Change in EQ-5D-3L index score			
Change in modified Rankin Scale score			
Change in EDSS score			
Change in visual acuity (Snellen chart)	Descriptive data only	NR	NR

ANCOVA = analysis of covariance; ARR = annualized relapse rate; BOCF = baseline observation carried forward; EDSS = Expanded Disability Status Scale; EQ-5D-3L = EuroQol 5 Dimension 3 level; FACIT-F = Functional Assessment of Chronic Illness–Fatigue; FSS = Functional System Score; IST = immunosuppressive therapy; MCS = mental component score; MMRM = mixed model for repeated measures; mRS = modified Rankin Scale; NR = not reported; PCS = physical component score; SF-36v2 = Short Form (36) Health Survey version 2; VAS = Visual Analogue Scale.

Source: Clinical Study Report for Study 898.[10]

Table 11: Statistical Analysis of Efficacy End Points in Study 900

End point	Statistical model	Adjustment factors	Sensitivity analyses
Study 900			
Time to first protocol-defined relapse	• Kaplan-Meier method • 2-sided log-rank test • Cox proportional hazards model	Stratified by prior therapy (B-cell–depleting therapy or IST/other) and recent attack prior to baseline (first attack or relapse) No imputation for missing study visits	Time to first event sensitivity analyses were based on the same statistical model as the primary analysis: • protocol-defined relapse by affecting medications[a] • protocol-defined relapse regardless of 7-day EDSS or FSS assessment limit • first clinical relapse • treated clinical relapse • clinical relapse of optic neuritis • protocol-defined relapse (weighted log-rank test)
Change in VAS pain score from baseline to week 24	ANCOVA (BOCF)	Baseline value, stratification factors (prior therapy, recent attack prior to baseline)	• Multiple imputations for missing data • Imputation using open-extension baseline or withdrawal visit values for missing data • MMRM
Change in FACIT-F score from baseline to week 24			
Proportion of relapse-free patients	Kaplan-Meier estimate	NR	NR
ARR	Negative binomial regression model with log exposure time as offset	Stratification factors (prior therapy, recent attack prior to baseline)	NR
Change in SF-36v2 PCS, MCS, and domains	MMRM (patients with baseline value and at least 1 post-baseline value)	Baseline value, stratification factors (prior therapy, recent attack prior to baseline), week, treatment, treatment-by-week interaction	NR
Change in EQ-5D-3L index score			
Change in modified Rankin Scale Score			
Change in EDSS score			
Change in visual acuity (Snellen chart and low-contrast Sloan letter chart)			

ANCOVA = analysis of covariance; ARR = annualized relapse rate; BOCF = baseline observation carried forward; EDSS = Expanded Disability Status Scale; EQ-5D-3L = EuroQol 5-Dimensions 3-Levels questionnaire; FACIT-F = Functional Assessment of Chronic Illness–Fatigue; FSS = Functional System Score; IST = immunosuppressive therapy; MCS = mental component score; MMRM = mixed model for repeated measures; mRS = modified Rankin Scale; NR = not reported; PCS = physical component score; SF-36v2 = Short Form (36) Health Survey version 2; VAS = Visual Analogue Scale.

[a] Censored at the first start date of the following affecting medications: relapse-prevention therapy, rescue therapy, or systemic corticosteroid for other indication taken for more than 5 days.

Source: Clinical Study Report for Study 900.[11]

Analysis Populations

Efficacy analyses in both studies were based on the ITT population, which included all randomized patients, analyzed according to the group they were randomized to. The safety population included all randomized patients who received at least 1 dose of the study drug, analyzed according to the drug they received.

Results

Patient Disposition

Disposition data for the overall population in studies 898 and 900 is presented in Table 12. No disposition data for the AQP4 IgG–positive subgroup was reported in the Clinical Study Reports.

A total of 96 patients were screened for inclusion in Study 898 and 83 patients were randomized (86%). Violation of eligibility criteria was stated as the main reason for screen failure, with no further details provided. In the placebo group, 10 patients (24%) were discontinued from the study due to adverse events (12%), refusal of treatment (5%), withdrawal of consent (5%), or eligibility deviation (2%). Three patients (7%) were discontinued from the satralizumab group, all due to adverse events.

Of the 168 patients screened for Study 900, 95 patients (57%) met the inclusion criteria and were randomized. The main reason for screen failure was violation of eligibility criteria (no details provided). Eleven patients were discontinued from the study, including 4 (13%) in the placebo group and 7 (11%) in the satralizumab group. The most common reasons for discontinuation were withdrawal of consent (6% and 3% for placebo and satralizumab, respectively), adverse events (3% and 2%), and other reasons (3% and 3%).

Table 12: Patient Disposition

	Study 898 (ITT)		Study 900 (ITT)	
	Placebo plus IST	Satralizumab plus IST	Placebo	Satralizumab
Screened, N	96		168	
Randomized, N (%)	83 (86)		95 (57)	
	42	41	32	63
Discontinued from study, N (%)	10 (24)	3 (7)	4 (13)	7 (11)
Reason for discontinuation, N (%)				
Adverse events	5 (12)	3 (7)	1 (3)	1 (2)
Eligibility deviation	1 (2)	0	0	0
Refused treatment	2 (5)	0	0	1 (2)
Withdrawal of consent	2 (5)	0	2 (6)	2 (3)
Protocol deviation	0	0	0	1 (2)
Other	0	0	1 (3)	2 (3)
ITT, N	42	41	32	63
PP, N	39	35	30	56
Safety, N	42	41	32	63

IST = immunosuppressive therapy; ITT = intention to treat; PP = per protocol.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Exposure to Study Treatments

The treatment duration was longer in the satralizumab group than in the placebo group in both studies.

In Study 898, the median treatment duration was 33 weeks (59.5 PYs) in the placebo group and 107 weeks (78.5 PYs) in the satralizumab group. Forty-three percent of patients in the placebo group and 59% of patients in the satralizumab group received at least 48 weeks of the study drug. Median adherence to treatment was 100% in both groups, ranging from 96% to 125% in the placebo group, and from 89% to 133% in the satralizumab group.

In Study 900, the median treatment duration was 54.6 weeks (40.6 PYs) in the placebo group and 92.3 weeks (115.2 PYs) in the satralizumab group. Fifty-three percent of patients in the placebo group and 73% in the satralizumab group received at least 48 weeks of the study drug. Three patients had less than 80% adherence to the study drug, including 1 patient in the placebo group and 2 in the satralizumab group (withdrawn due to nonadherence; dose interruption due to adverse event of vertigo). Median adherence to treatment was 100% in both groups and ranged from 77% to 200% in the placebo group and from 67% to 150% in the satralizumab group. Five patients with greater than 120% adherence to treatment were entered into the open-extension period shortly after receiving the loading doses of the study drug.

The FDA reported that 17 (53%) patients in the placebo group and 21 (33%) in the satralizumab group received rescue therapy in Study 900 (ITT).[33] In Study 898, 25 patients who received placebo (60%) and 18 who received satralizumab (44%) were administered rescue therapy during the trial.

Efficacy

Only those efficacy outcomes and analyses of subgroups identified in the review protocol are reported subsequently. See Appendix 3 for detailed efficacy data.

Relapse

More patients in the placebo group than in the satralizumab group experienced an adjudicated, protocol-defined relapse in the ITT population of Study 898 and Study 900 (Table 13). The HR for the time to first protocol-defined relapse was 0.38 (95% CI, 0.16 to 0.88; P = 0.018) in favour of satralizumab plus immunosuppressant versus placebo plus immunosuppressant (Study 898). In Study 900, the HR for the time to first protocol-defined relapse was 0.45 (95% CI, 0.23 to 0.89; P = 0.018) for satralizumab versus placebo. Both studies met their primary end point and showed statistically significant differences favouring satralizumab versus placebo for the ITT populations (Table 13).

The treatment effects for the overall population were driven mainly by the AQP4 IgG–positive subgroup, which in Study 900 reported a statistically significant treatment by AQP4 status interaction term (P = 0.02). The analysis of subgroups according to AQP4 status at baseline was pre-planned.

For the AQP4 IgG–positive subgroup in Study 898, 43% (12 of 28) of patients in the placebo plus immunosuppressant group, and 11% (3 of 27) of those in the satralizumab plus immunosuppressant group, experienced an adjudicated, protocol-defined relapse, with an HR of 0.21 (95% CI, 0.06 to 0.75; P = 0.0086; not controlled for type I error rate) (Table 13). The Kaplan-Meier plot of the time to first protocol-defined relapse in the

AQP4 IgG–positive subgroup shows that the groups start to separate after approximately 24 weeks of therapy (Figure 2).

In the AQP4 IgG–positive subgroup of Study 900, 57% (13 of 23) of patients in the placebo group, and 22% (9 of 41) of patients in the satralizumab group, experienced a protocol-defined relapse, with an HR of 0.26 (95% CI, 0.11 to 0.63; P = 0.0014; not controlled for type I error rate). The Kaplan-Meier plot for the AQP4 IgG–positive subgroup shows the treatment groups starting to separate within the first 12 weeks of the study (Figure 3).

The time to first protocol-defined relapse in the AQP4 IgG–negative subgroup was as follows. For Study 898, the HR was 0.66 (95% CI, 0.20 to 2.23; N = 28). For Study 900, the HR was 1.19 (95% CI, 0.30 to 4.78; N = 31). Additional pre-planned subgroup analyses for the ITT population are shown in Figure 6 and Figure 7 (Appendix 3).

The time to first clinical relapse (non-adjudicated relapse determined by the individual investigator) was reported for the ITT and AQP4 IgG–positive populations (Table 13). In the AQP4 IgG–positive subgroup of Study 898, [redacted] of patients in the placebo and satralizumab groups reported a clinical relapse ([redacted]). In Study 900, [redacted] of patients in the placebo and satralizumab groups, respectively, experienced a clinical relapse, with an HR of [redacted].

A number of other sensitivity analyses that used alternate definitions of relapse were reported for the ITT population (see Table 10 and Table 11 for a description of the analyses). Although point estimates consistently favoured satralizumab versus placebo in Study 898, the 95% CI included the null for time to protocol-defined relapse based on the change in EDSS relative to baseline values, treated clinical relapse, and treated clinical relapse of optic neuritis. The 95% CI excluded the null for time to protocol-defined relapse if the EDSS assessment took place outside the 7-day limit. In Study 900, the sensitivity analyses for time to protocol-defined relapse, censored for concomitant medications, the relapses not assessed within the 7-day limit, and the treated clinical relapse, all showed point estimates that favoured satralizumab, with a 95% CI that excluded the null. Time to clinical relapse of optic neuritis reported a 95% CI that included the null.

Table 13: Time to First Protocol-Defined or Clinical Relapse for Study 898 and Study 900

	ITT population		AQP4 IgG–positive subgroup	
	Placebo	Satralizumab	Placebo	Satralizumab
Time to first protocol-defined relapse[a]				
Study 898[b]				
Number of patients contributing to the analysis	42	41	28	27
Number of patients with relapse, n (%)	18 (42.9)	8 (19.5)	12 (42.9)	3 (11.1)
HR versus placebo (95% CI)		0.38 (0.16 to 0.88)		0.21 (0.06 to 0.75)
P value (log-rank test)		0.0184		0.0086[c]
Study 900[d]				
Number of patients contributing to the analysis	32	63	23	41
Number of patients with relapse, n (%)	16 (50.0)	19 (30.2)	13 (56.5)	9 (22.0)
HR versus placebo (95% CI)		0.45 (0.23 to 0.89)		0.26 (0.11 to 0.63)
P value (log-rank)		0.0184		0.0014[c]

	ITT population		AQP4 IgG–positive subgroup	
	Placebo	Satralizumab	Placebo	Satralizumab
Time to first clinical relapse[e]				
Study 898[b]				
Number of patients contributing to the analysis	42	41	[redacted]	[redacted]
Number of patients with relapse, n (%)	27 (64.3)	18 (43.9)	[redacted]	[redacted]
HR versus placebo (95% CI)		0.59 (0.33 to 1.08)		[redacted]
P value (log-rank test)		[redacted]		[redacted]
Study 900[d]				
Number of patients contributing to the analysis	32	63	[redacted]	[redacted]
Number of patients with relapse, n (%)	17 (53.1)	31 (49.2)	[redacted]	[redacted]
HR versus placebo (95% CI)		0.74 (0.41 to 1.35)		[redacted]
P value (log-rank test)		[redacted]		[redacted]

AQP4 = aquaporin 4; ARR = annualized relapse rate; CI = confidence interval; EDSS = Expanded Disability Status Scale; HR = hazard ratio; IgG = immunoglobulin G; ITT = intention to treat; NE = not estimable.

[a] Protocol-defined relapse: Adjudicated by the Clinical Endpoint Committee; EDSS assessment performed within 7 days of relapse reporting.

[b] Study 898 HR and 95% CI based on Cox proportional hazards model and P value based on a log-rank test. Analyses were stratified by baseline ARR (1 versus > 1) and geographic region (Asia versus Europe or other) for the ITT population (column 1) or the AQP4 IgG–positive subgroup (column 2). Both treatment groups also received background immunosuppressive therapy.

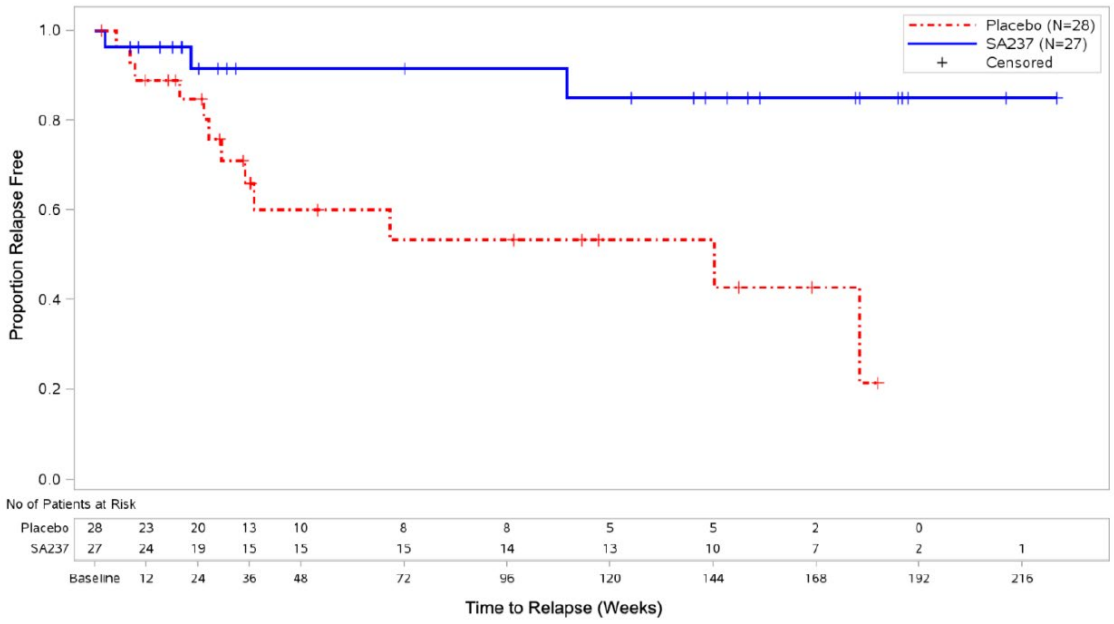
[c] P value has not been adjusted for multiple testing (i.e., the type I error rate has not been controlled).

[d] Study 900 HR and 95% CI based on Cox proportional hazards model and P value based on a log-rank test. Analyses were stratified by prior therapy (B-cell–depleting therapy versus immunosuppressants or other) and most recent attack (first attack versus relapse) for the ITT population (column 1) or the AQP4 IgG–positive subgroup (column 2).

[e] Clinical relapse was defined as any relapse reported by the investigator (not adjudicated).

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

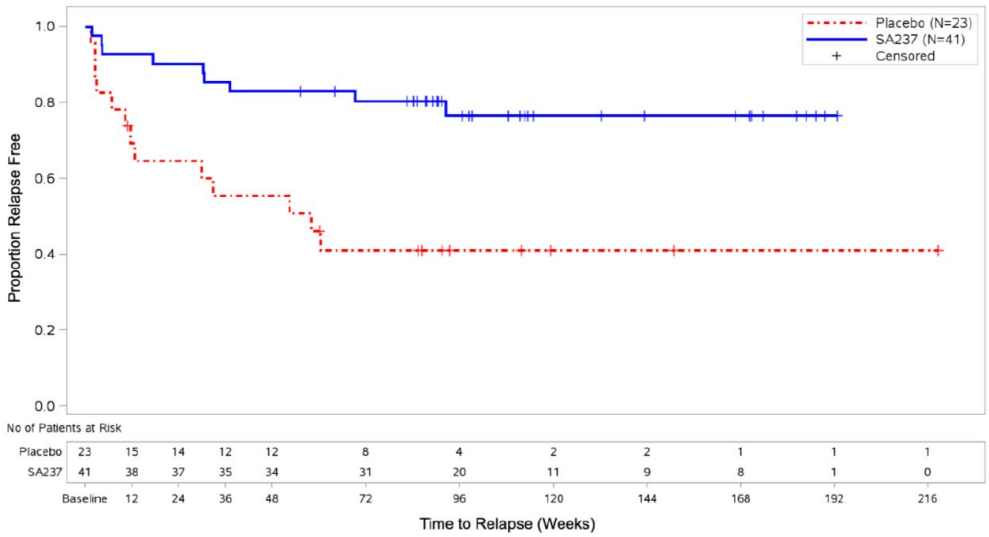
Figure 2: Time to Protocol-Defined Relapse in the AQP4 IgG–Positive Subgroup of Study 898



AQP4 = aquaporin 4; IgG = immunoglobulin G; SA237 = satralizumab.

Source: Reproduced from the Health Canada Reviewer's Report (p. 20).[34]

Figure 3: Time to Protocol-Defined Relapse in the AQP4 IgG–Positive Subgroup of Study 900



AQP4 = aquaporin 4; IgG = immunoglobulin G; SA237 = satralizumab.

Source: Reproduced from the Health Canada Reviewer's Report (p. 27).[34]

The proportion of patients who were relapse-free at week 48 and 96 and the ARR for the ITT population are presented in Table 25 (Appendix 3). In Study 900 (no background immunosuppressive therapy), the adjusted ARR was 2.0 (95% CI, 0.7 to 6.1) for the placebo group, compared with 0.6 (95% CI, 0.2 to 1.4) for the satralizumab group, with an adjusted ARR ratio of 0.3 (95% CI, 0.1 to 1.1; P = 0.067; not adjusted for type I error rate). In Study 898, where all patients were receiving background immunosuppressant therapy, the adjusted ARR was 0.5 (95% CI, 0.2 to 1.3) in the placebo group and 0.1 (95% CI, 0.1 to 0.4) in the satralizumab group. The adjusted ARR ratio was 0.3 (95% CI, 0.1 to 0.8; P = 0.018; not adjusted for type I error and should be considered as supportive evidence only).

Disability

Disability was measured using the modified Rankin Scale and EDSS scores and reported for the ITT population in both studies (MMRM model). In Study 900, visual acuity data were measured based on a 20-foot Snellen chart with logarithm of the minimum angle of resolution (logMAR) visual acuity scoring. The data were reported descriptively in Study 898 and analyzed using an MMRM model in Study 900. These outcomes were not part of the statistical testing hierarchy. No data were found for the AQP4 IgG–positive subgroups.

In Study 898, the difference in adjusted means for the change from baseline to week 24 in the modified Rankin Scale score was 0.02 (95% CI, −0.23 to 0.27; P = 0.88) for satralizumab versus placebo. For the change from baseline to week 24 in the EDSS score, the adjusted difference in means was 0.11 (95% CI, −0.28 to 0.49; P = 0.57). For these outcomes, 24-week data were missing for 31% and 29% of patients in the placebo and satralizumab groups, respectively.

In Study 900, there was no consistent difference between the 2 groups in the modified Rankin Scale or EDSS scores. The MMRM analysis of the change from baseline to week 24 in the modified Rankin Scale scores reported a difference in adjusted means of 0.17 (95% CI, −0.14 to 0.47; P = 0.29) for satralizumab versus placebo. The difference in the adjusted means for the change from baseline to week 24 in the EDSS scores was −0.17 (95% CI, −0.50 to 0.16; P = 0.31) for satralizumab versus placebo. At 24 weeks, data were missing for 14% to 16% of patients in the satralizumab group and 38% to 41% of patients in the placebo group.

In both studies, there was no apparent difference observed in the logMAR visual acuity score between the 2 groups. At 24 weeks in Study 898, the median within-group change from baseline in logMAR visual acuity score was zero for both eyes in both treatment groups. The mean change from baseline ranged from −0.064 (SD = 0.197) for the right eye in the placebo group to 0.059 (SD = 0.319) for the left eye in the satralizumab group. In Study 900, the difference in adjusted means for the logMAR score for the right eye was 0.12 (95% CI, −0.12 to 0.35; P = 0.33) and left eye was −0.04 (95% CI, −0.27 to 0.19; P = 0.75) for satralizumab versus placebo. The sponsor reported no apparent difference between groups in visual acuity measured using the low-contrast Sloan letter chart. Visual acuity data at 24 weeks were missing from 29% of patients per group in Study 898, and from 16% and 38% of patients treated with satralizumab and placebo, respectively, in Study 900.

Health-Related Quality of Life

Health-related quality of life was measured using the SF-36 and the EQ-5D-3L instruments and reported for the ITT population (MMRM model). These outcomes were not part of the statistical testing hierarchy. No data were found for the AQP4 IgG–positive subgroups.

In Study 898, there was no consistent difference between treatment groups in the change from baseline in the SF-36 physical component score, mental component score, or individual domains, nor for the EQ-5D index score. The difference in the adjusted mean change from baseline to week 24 for the SF-36 mental component score was −2.3 (95% CI, −6.4 to 1.8; P = 0.26) and −1.4 for the physical component score (95% CI, −4.3 to 1.5; P = 0.35) for satralizumab versus placebo. At week 24, the difference in adjusted means for the change in EQ-5D index score was −0.045 (95% CI, −0.124 to 0.034; P = 0.26) for satralizumab versus placebo. Data were missing for 29% to 33% of patients per treatment group at 24 weeks.

In Study 900, the difference in adjusted means for the change from baseline to week 24 in the SF-36 mental component score was 3.4 (95% CI, −0.8 to 7.7; P = 0.11), and for the physical component score was −1.0 (95% CI, −4.0 to 1.9; P = 0.48). The difference in adjusted means for the change from baseline to week 24 in the EQ-5D index score was −0.003 (95% CI, −0.087 to 0.080; P = 0.94). Data were missing for 38% of patients in the placebo group and 14% to 16% of patients in the satralizumab group at week 24.

Productivity

Neither study reported data on productivity.

Symptoms

The change from baseline to week 24 in pain VAS scores and FACIT-F scores in the ITT population were key secondary outcomes in both trials and part of the statistical testing hierarchy. At 24 weeks, data for pain or fatigue was observed for 29 patients per group in Study 898; thus, baseline values were imputed for 31% and 29% of patients in the placebo and satralizumab groups, respectively.[33] In Study 900, 24-week data were missing for 12 patients (38%) in the placebo group and 11 patients (17%) in the satralizumab group and were imputed using BOCF.[33]

No statistically significant differences in pain or fatigue scores favouring satralizumab were detected in either the ITT population or the AQP4 IgG–positive subgroups in studies 898 or 900 (Table 14, Table 15). In the AQP4 IgG–positive subgroup, the difference in adjusted means for the change in VAS pain scores was [redacted] in Study 898, and [redacted] in Study 900 for satralizumab versus placebo. The difference in adjusted means for the change in FACIT-F scores was [redacted] in Study 898 and [redacted] in Study 900 for the AQP4 IgG–positive subgroup.

Table 14: Change in Pain VAS From Baseline to Week 24 in Study 898 and Study 900

Change in baseline to week 24 in VAS for pain[a]	ITT population		AQP4 IgG–positive subgroup	
	Placebo	Satralizumab	Placebo	Satralizumab
Study 898[b]	**N = 42**	**N = 41**	**N = 28**	**N = 27**
Number of patients contributing to the analysis	42	41	[redacted]	[redacted]
Baseline, mean (SE)	34.6 (4.0)	27.6 (4.4)	[redacted]	[redacted]
Adjusted mean change from baseline, mean (SE or 95% CI)	−3.5 (SE 2.4)	2.9 (SE 2.4)	[redacted]	[redacted]
Difference in adjusted means versus placebo (95% CI)		6.4 (−0.3 to 13.0)		[redacted]
P value		0.060		[redacted]
Study 900[d]	**N = 32**	**N = 63**	**N = 23**	**N = 41**
Number of patients contributing to the analysis	32	62	[redacted]	[redacted]
Baseline, mean (SE)	27.6 (5.4)	31.7 (3.7)	[redacted]	[redacted]
Adjusted mean change from baseline, mean (SE or 95% CI)	−5.9 (SE 4.8)	−2.7 (SE 4.3)	[redacted]	[redacted]
Difference in adjusted means versus placebo (95% CI)		3.2 (−5.1 to 11.5)		[redacted]
P value		0.444		[redacted]

ANCOVA = analysis of covariance; AQP4 = aquaporin 4; ARR = annualized relapse rate; BOCF = baseline observation carried forward; CI = confidence interval; IgG = immunoglobulin G; ITT = intention to treat; NR = not reported; SE = standard error; VAS = Visual Analogue Scale.

[a] Pain VAS is scored from 0 to 100, with higher scores representing worse pain. For the between-group comparison, the negative difference in means favours satralizumab over placebo.

[b] Study 898 analysis based on ANCOVA model, including treatment, baseline value, and stratification factors (baseline ARR and geographic region) with BOCF for missing data. Both groups also received background immunosuppressant therapy.

[c] P value has not been adjusted for multiple testing (i.e., the type I error rate has not been controlled).

[d] Study 900 analysis based on ANCOVA model including treatment, baseline value, and stratification factors (prior therapy and most recent attack) with BOCF for missing data.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Table 15: Change in FACIT-F Score From Baseline to Week 24 in Study 898 and Study 900

Change from baseline to week 24 in FACIT-F score[a]	ITT population		AQP4 IgG–positive subgroup	
	Placebo	Satralizumab	Placebo	Satralizumab
Study 898[b]	**N = 42**	**N = 41**	**N = 28**	**N = 27**
Number of patients contributing to analysis	42	41	[redacted]	[redacted]
Baseline, mean (SE)	33.9 (1.8)	34.7 (1.7)	[redacted]	[redacted]
Adjusted mean change from baseline to week 24, mean (SE or 95% CI)	2.2 (SE 0.9)	0.1 (SE 1.0)	[redacted]	[redacted]
Difference in adjusted means versus placebo (95% CI)		−2.1 (−4.8 to 0.6)		[redacted]
P value		0.122		[redacted]
Study 900[d]	**N = 32**	**N = 63**	**N = 23**	**N = 41**
Number of patients contributing to analysis	32	62	[redacted]	[redacted]
Baseline, mean (SE)	29.7 (2.3)	30.6 (1.5)	[redacted]	[redacted]

Change from baseline to week 24 in FACIT-F score[a]	ITT population		AQP4 IgG–positive subgroup	
	Placebo	Satralizumab	Placebo	Satralizumab
Adjusted mean change from baseline to week 24, mean (SE or 95% CI)	3.6 (SE 1.8)	5.7 (SE 1.6)	[redacted]	[redacted]
Difference in adjusted means versus placebo (95% CI)		2.1 (−1.0 to 5.2)		[redacted]
P value		0.182		[redacted]

ANCOVA = analysis of covariance; AQP4 = aquaporin 4; ARR = annualized relapse rate; BOCF = baseline observation carried forward; CI = confidence interval; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; IgG = immunoglobulin G; ITT = intention to treat; NR = not reported; SE = standard error.

[a] FACIT-F scale is scored from 0 to 52, with lower scores representing more fatigue. For the between-group comparison, the positive difference in means favours satralizumab over placebo.

[b] Study 898 analysis based on ANCOVA model, including treatment, baseline value, and stratification factors (baseline ARR and geographic region) with BOCF for missing data. Both groups also received background immunosuppressant therapy.

[c] P value has not been adjusted for multiple testing (i.e., the type I error rate has not been controlled).

[d] Study 900 analysis based on ANCOVA model, including treatment, baseline value, and stratification factors (prior therapy and most recent attack) with BOCF for missing.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Health Care Resource Utilization

Neither study reported data on health care resource utilization.

Harms

Only those harms identified in the review protocol are reported subsequently. See Table 16 and Table 17 for detailed harms data. All harms data were reported based on the safety population. No safety data were reported for the AQP4 IgG–positive subgroup.

Adverse Events

The proportion of patients who experienced an adverse event ranged from 75% to 95% in the placebo groups and from 90% to 92% in the satralizumab groups. After adjusting for follow-up time, the rate of adverse events was 495 to 514 events per 100 PYs among those assigned to placebo, and from 474 to 485 events per 100 PYs to those who received satralizumab (Table 16). The most common adverse events were urinary tract infections (17% to 25% of patients), upper respiratory tract infection (14% to 24%), headache (10% to 24%), nasopharyngitis (3% to 24%), and injection-related reactions (5% to 16%).

Serious Adverse Events

Serious adverse events were reported in 16% to 21% of patients assigned to placebo and 17% to 19% of patients who received satralizumab, with a serious adverse event rate of 15 to 20 events per 100 PYs, and 12 to 17 events per 100 PYs in the placebo and satralizumab groups, respectively (Table 17). Infections were the most common serious adverse event reported in both studies (5% to 10% of patients).

Stopped Treatment Due to Adverse Events

In Study 898, 5 patients (12%) in the placebo group stopped treatment due to adverse events (serious adverse event: breast cancer, hepatic cancer, autoimmune thrombocytopenia, lymphopenia; non-serious adverse event: leukopenia). Three patients (7%) in the satralizumab group stopped treatment (non-serious adverse event: increased hepatic enzymes, decreased neutrophil count, moderate urticarial). More patients in the satralizumab group had treatment interrupted due to adverse events than in the placebo

group (17% versus 10%). Infections were the most comment event leading to dose interruption (placebo = 5%; satralizumab = 12%).

In Study 900, 1 patient (2%) who received satralizumab, discontinued treatment due to a serious adverse event of pneumonia. One patient (3%) in the placebo group stopped treatment due to systemic lupus erythematosus. In the double-blind period, 28% of patients in the placebo group and 25% in the satralizumab group had their treatment interrupted due to adverse events. Infections were the most comment event leading to dose interruption (placebo = 19%; satralizumab = 13%).

Mortality

No deaths were reported during the double-blind period of Study 898 and Study 900.

Notable Harms

In Study 898, the frequency (62% versus 68%) and rate of infections or infestations (150 versus 133 events per 100 PYs) was similar in the placebo and satralizumab groups, respectively, where treatment was administered as add-on therapy to background immunosuppressants (Table 17). The most common events in both groups were nasopharyngitis (17% and 24%), upper respiratory tract infection (14% and 24%), and urinary tract infection (17% in both groups) (Table 16). The proportion of patients who reported a serious infection was 7% and 5%, with a rate of 5.0 and 2.6 events per 100 PYs in the placebo and satralizumab groups, respectively. Of these patients, 1 patient in the placebo group and 2 in the satralizumab group had their treatment interrupted due to serious infections. Four patients in the placebo group (10%) and 1 in the satralizumab group (2%) received IV antibiotics for an infection. Potential opportunistic infections were reported by 12% and 10% in the placebo and satralizumab groups, respectively, with a rate of 35.3 and 10.2 events per 100 PYs.

In Study 900, 44% of patients in the placebo group and 54% in the satralizumab group reported an infection or infestation (163 versus 100 events per 100 PYs) (Table 17). The most common infections in the placebo and satralizumab groups were urinary tract infection (25% versus 18%), upper respiratory tract infection (19% versus 16%), nasopharyngitis (3% versus 14%) and influenza (6% versus 8%, respectively) (Table 16). Serious infections were reported in 9% and 10% of patients in the placebo and satralizumab groups, respectively (9.9 events per 100 PYs versus 5.2 events per 100 PYs). One patient stopped satralizumab due to a serious adverse event of pneumonia, and 2 patients per group had treatment interrupted due to infection. One patient in the placebo group (3%) and 6 patients (10%) in the satralizumab group had an infection that required IV antibiotics; all infections resolved. Potential opportunistic infections were reported in 16% and 5% of patients in the placebo and satralizumab groups, respectively, with a rate of 17.3 and 2.6 events per 100 PYs.

In Study 898, 2 patients (5%) in the placebo group reported injection-related reactions, whereas 5 patients (12%) in the satralizumab reported injection-related reactions (Table 17). In Study 900, 5 patients in the placebo group (16%) and 8 patients (13%) in the satralizumab group had injection-related reactions. None of the injection-related reactions resulted in treatment discontinuation in either study.

No patients in either study met the criteria for drug-induced liver injury, defined as a level of aspartate aminotransferase or alanine aminotransferase that is 3 or more times greater than the upper limit of normal, with a total bilirubin level that is 2 or more times greater than the upper limit of normal or clinical signs of jaundice. No anaphylaxis events were reported in either study.

Table 16: Summary of Common Adverse Events for Study 898 and Study 900

	Study 898 (safety population)				Study 900 (safety population)			
	Placebo plus IST N = 42		Satralizumab plus IST N = 41		Placebo N = 32		Satralizumab N = 63	
	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]
Any adverse event								
Patients with ≥ 1 AE	40 (95)	514.3	37 (90)	485.2	24 (75)	495.2	58 (92)	473.9
Most common events,[c] n (%)								
Urinary tract infection	7 (17)		7 (17)		8 (25)		11 (18)	
Upper respiratory tract infection	6 (14)		10 (24)		6 (19)		10 (16)	
Headache	4 (10)		10 (24)		4 (13)		10 (16)	
Nasopharyngitis	7 (17)		10 (24)		1 (3)		9 (14)	
Injection-related reactions	2 (5)		5 (12)		5 (16)		8 (13)	
Constipation	7 (17)		2 (5)		2 (6)		3 (5)	
Pharyngitis	3 (7)		4 (10)		0		1 (2)	
Influenza	4 (10)		0		2 (6)		5 (8)	
Gastritis	0		4 (10)		NR		NR	
Leukopenia	4 (10)		6 (15)		0		2 (3)	
Anemia	5 (12)		3 (7)		0		2 (3)	
Lymphopenia	4 (10)		3 (7)		0		2 (3)	
Back pain	5 (12)		4 (10)		3 (9)		4 (6)	
Arthralgia	0		4 (10)		1 (3)		10 (16)	
Hypercholesterolemia	5 (12)		4 (10)		0		2 (3)	
Pyrexia	5 (12)		0		0		1 (2)	
Nausea	3 (7)		3 (7)		2 (6)		11 (18)	
Fatigue	1 (2)		2 (5)		2 (6)		7 (11)	

	Study 898 (safety population)				Study 900 (safety population)			
	Placebo plus IST N = 42		Satralizumab plus IST N = 41		Placebo N = 32		Satralizumab N = 63	
	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]
Rash	2 (5)		0		1 (3)		9 (14)	
Pain in extremity	3 (7)		1 (2)		3 (9)		9 (14)	

AE = adverse event; NR = not reported; IST = immunosuppressive therapy; PY = patient-year.

[a] Number of patients with an AE.

[b] Number of AEs per 100 PYs of follow-up (multiple occurrences of the same event in 1 patient counted multiple times). Total PYs for Study 898: placebo = 59.5; satralizumab = 78.5. Total PYs for Study 900: placebo = 40.6; satralizumab = 115.2.

[c] Frequency ≥ 10%.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Table 17: Summary of SAEs, WDAEs, Deaths, and Notable Harms for Study 898 and Study 900

	Study 898 (safety population)				Study 900 (safety population)			
	Placebo plus IST N = 42		Satralizumab plus IST N = 41		Placebo N = 32		Satralizumab N = 63	
	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]
SAE								
Patients with ≥ 1 SAE	9 (21)	20.2	7 (17)	11.5	5 (16)	14.8	12 (19)	17.4
Most common events by SOC[c]								
Infections and infestations	3 (7)		2 (5)		3 (9)		6 (10)	
Blood and lymphatic system disorders	3 (7)		1 (2)		NR		NR	
Injury, poisoning, or procedural complication	0		2 (5)[d]		0		2 (3)	
Neoplasms benign, malignant, and unspecified	2 (5)		0		NR		NR	
Nervous system disorders	0		1 (2)		2 (6)		0	
Stopped treatment due to AEs								
Patients who stopped treatment due to AEs	5 (12)		3 (7)		1 (3)		1 (2)	

	Study 898 (safety population)				Study 900 (safety population)			
	Placebo plus IST N = 42		Satralizumab plus IST N = 41		Placebo N = 32		Satralizumab N = 63	
	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]	n (%)[a]	AEs per 100 PYs[b]
Deaths								
Deaths	0	0	0	0	0	0	0	0
Notable harms								
Infection and infestations (SOC)	26 (62)	149.6	28 (68)	132.5	14 (44)	162.6	34 (54)	99.8
Serious infection	3 (7)	5.0	2 (5)	2.6	3 (9)	9.9	6 (10)	5.2
Potential opportunistic infection	5 (12)	35.3	4 (10)	10.2	5 (16)	17.3	3 (5)	2.6
Injection-related reaction	2 (5)	3.4	5 (12)	21.7	5 (16)	17.3	8 (13)	13.9
Hepatotoxicity[e]	0	0	0	0	0	0	0	0
Hypersensitivity	NR	NR	NR	NR	NR	NR	NR	NR
Anaphylaxis	0	0	0	0	0	0	0	0

AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; IST = immunosuppressive therapy; NR = not reported; PY = patient-year; SAE = serious adverse event; SOC = system organ class; ULN = upper limit of normal; WDAE = withdrawal due to adverse event.

[a] Number of patients with an AE.

[b] Number of AEs per 100 PYs of follow-up (multiple occurrences of the same event in 1 patient counted multiple times). Study 898 total PYs: placebo = 59.5; satralizumab = 78.5; Study 900 total PYs: placebo = 40.6; satralizumab = 115.2.

[c] Reported in 5% or more of patients per group.

[d] Both events were fractures.

[e] Drug-induced liver injury defined as AST or ALT ≥ 3 times ULN with total bilirubin ≥ 2 times ULN or clinical signs of jaundice.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Critical Appraisal

Internal Validity

Both trials used accepted methods to randomize patients to treatment groups that included an interactive web or voice response system, with randomization stratified by ARR and region for Study 898, and by prior NMOSD therapy and most recent attack (i.e., patient's first or recurrent event) in Study 900. Although subgroup analysis by AQP4 status at baseline was a pre-planned analysis, randomization was not stratified by this factor; thus, it is possible there are imbalances between treatment groups in this subgroup. Based on the data presented in the Clinical Study Reports, the baseline patient characteristics appear to be similar between groups for both the ITT and AQP4 IgG–positive subgroup in Study 898; however, in Study 900, the treatment groups appear to differ in age, sex, racial distribution, body mass index, and prior therapies. The clinical expert consulted did not anticipate that the differences observed would bias the results in favour of satralizumab. The FDA also noted differences between groups in Study 900, but concluded they were likely due to chance and the small sample size and did not expect the differences to affect the key results.[32] The percentage of patients who withdrew was similar between groups in Study 900 (13% and 11%), but was higher in the placebo plus immunosuppressant group than in the satralizumab plus immunosuppressant group (24% versus 7%) in Study 898.

Both trials were double blinded, with identical-looking placebo and active study drug treatments. Both used blinded outcome assessors who were not involved in patient care to evaluate patients who reported potential relapse events. There were no notable differences in the adverse event profile; thus, the risk of unblinding due to adverse effects was not a major concern. Some of the patients' laboratory data were concealed from study personnel to prevent inadvertent unblinding; however, the FDA noted that fibrinogen levels were not concealed from treating investigators or the sponsor's medical monitoring team.[33] Other interleukin-6 inhibitors are known to reduce fibrinogen levels; thus, the treating investigators may have been able to infer treatment allocation from patients' fibrinogen levels. The FDA noted that 2 clinical science members from Chugai Pharmaceutical (study partner and sponsor) used fibrinogen levels to infer patients' treatment allocation, which was used to calculate the HR for time to relapse for both studies.[33] The accuracy of the inferred treatment allocation was considered high and ranged from 73% to 100%, and the estimated results were shared with a number of colleagues. The FDA Statistical Report states that the shared efficacy estimates may have influenced the protocol changes made to Study 900, which changed the sample size and stopping criteria of the study. "Based on the Bias Assessment Report included in this BLA [biologics license application] submission, the attempts of using fibrinogen data for efficacy purpose allowed for a highly accurate prediction of study treatment groups, which likely changed the course of the study to give the study drug an opportunity to show its effect under an optimized sample size and study duration" (FDA Statistical Review, page 6).[33] Based on the original planned sample size and stopping criteria (70 patients and 19 relapse events), the FDA calculated that no statistically significant differences would have been detected between groups. Of note: the investigators responsible for assessing EDSS and the adjudication committee members were blinded to fibrinogen levels; thus, no bias in determining protocol-defined relapse events was identified by the FDA. Although the investigators responsible for patient care had access to fibrinogen levels and may have inferred treatment allocation, the sponsor analyzed imbalances in the collection and reporting of data sensitive to bias and did not identify any bias in data reporting.[33] Health Canada also noted that the results of Study 898

became available while Study 900 was ongoing and may have influenced protocol changes related to the study's stopping criteria.[34]

The primary outcome in both trials was the time to first protocol-defined relapse that was confirmed by the blinded clinical event adjudication committee, which consisted of 3 neurologists and/or ophthalmologists with expertise in the diagnosis and treatment of NMOSD. Use of adjudication is expected to increase the validity and objectivity of the outcome assessment, as it reduces the inter-site variability in assessments and over-reporting bias that may have influenced the relapses determined by the attending physician, as the need for immediate treatment of relapses could impact the classification of an event as a relapse. In both studies, a protocol-defined relapse included any new or worsening neurological symptoms attributable to NMO or NMOSD that persisted for a minimum of 24 hours and that were not attributable to confounding clinical factors and that met pre-specified criteria for an increase in EDSS or FSS scores. The clinical expert consulted considered the definition to be acceptable for the clinical trial, although they stated that in clinical practice, strict criteria based on EDSS or FSS scores are not used, and even minor neurological changes attributed to a relapse would be treated early and aggressively in order to prevent disability accrual. In the trials, a number of sensitivity analyses were run using different definitions of relapse. In general, the point estimates favoured satralizumab versus placebo but, for some sensitivity analyses, the CIs included the null. This included an analysis of clinical relapses, as determined by the treating physician. While this definition of relapse may more closely reflect clinical practice, the weekly monitoring of patients for relapses, and the option to move to the open-label satralizumab extension period for patients with a treated clinical relapse in Study 898 may have influenced the reporting of these events.

The time to first protocol-defined relapse for the AQP4 IgG–positive subgroup was not part of the statistical testing hierarchy, and therefore the P values are not controlled for the type I error rate. The Clinical Study Report does not report if any testing was conducted to determine if the proportional hazards assumption was met for the time to relapse outcomes. However, the Kaplan-Meier graphs did not suggest any major violations of the proportional hazards assumption.

Study 898 and Study 900 were designed as time-to-event trials, where patients completed the trial after having a protocol-defined relapse (or treated clinical relapse for Study 898). This design inherently emphasizes the efficacy of satralizumab on the first relapse, but it is not designed to assess its efficacy on subsequent relapses. While ARR results were reported and are a clinically relevant end point, the trials do not capture data on subsequent relapses (occurring after 30 days of the first relapse) because patients were censored and therefore subsequent relapses would not have been captured, thereby likely underestimating the ARR. ARR was not part of the statistical testing hierarchy and the P value was not controlled for type I error rate; thus, any differences detected should be considered as supportive evidence for the effect of satralizumab in the overall population.

The key secondary outcomes in both studies were the change from baseline to week 24 in pain and fatigue, measured using a VAS and the FACIT-F score. The FDA statistical reviewer stated the trials were not designed to fully capture the change in either of these outcomes.[33] A large number of patients did not have pain VAS or FACIT-F score data at 24 weeks due to patients being withdrawn after experiencing a relapse or other early discontinuations. Patients with missing data had baseline values carried forward (Study 898: 29% to 31%; Study 900: [redacted]) which the FDA stated likely biases the results,

particularly for those missing due to relapse.[33] Since patients who relapse would be expected to have worse outcomes, and since relapses were more frequent in the placebo group, using BOCF likely biases in favour of satralizumab. The FDA conducted additional analyses with no imputation for missing data (i.e., observed case data) and these analyses showed results similar to the sponsor's data, with no statistically significant difference detected between groups in the ITT population.[33]

Although both trials reported data for the change in disability or health-related quality of life measures, these outcomes were limited by the extent of missing data, which was due in part to the withdrawal of patients who experienced a relapse. The frequency of missing data was high for the EDSS, modified Rankin Scale, visual acuity, EQ-5D, and SF-36 data, which ranged from [redacted] at 24 weeks. Moreover, there was no imputation for missing data in the MMRM analyses and these outcomes were not part of the statistical testing hierarchy. Although EDSS is considered a standard measure of disability for patients with NMOSD, it has known limitations and may not fully capture the disability experienced by patients with this rare disease.

Study 898 enrolled a total of 7 adolescents ([redacted] of whom were AQP4 IgG positive) which was insufficient to accurately calculate the HR for the time to protocol-defined relapse. Health Canada approved the drug in adolescent patients aged 12 or older based on clinical pharmacology data and extrapolation of efficacy and safety from adult patients with NMOSD.[34]

The dose of satralizumab used in the clinical trials was consistent with the Canadian product monograph, and the drugs and doses used as background therapy in Study 898 were acceptable, according to the clinical expert consulted. Although the proportion of patients receiving corticosteroids as relapse-prevention therapy was higher than would be expected in Canada, there were no between-group imbalances in the use of background therapies. The comparator in both studies was placebo; thus, evidence is lacking for the direct comparative efficacy and safety of satralizumab versus eculizumab or other immunosuppressants.

No information was available on the impact of satralizumab on productivity or health care resource utilization. Safety data were limited by the small sample size and follow-up time; therefore, longer-term safety is uncertain. Median treatment duration was shorter for patients who received placebo compared with satralizumab (33 to 55 weeks versus 92 to 107 weeks); thus, exposure time should be considered when comparing the proportion of patients who experienced an adverse event in each treatment group.

External Validity

The clinical expert consulted by CADTH considered the baseline demographics and disease characteristics of the patients enrolled in the trials to be generally consistent with patients seen in the Canadian clinical setting. Approximately 50% of patients in Study 900 were from the US and 12% were from Canada, whereas most patients in Study 898 were from Asia or Europe (no Canadian sites). Overall, 86% of patients screened were enrolled in Study 898, but only 57% were enrolled in Study 900, with few details provided on the characteristics of those who were excluded. The preponderance of female study participants was consistent with the natural history of the disease. In addition, the inclusion criteria of the trials created study populations that were likely to experience a relapse based on their historical relapses. This is a practical point when designing a clinical trial in order to ensure sufficient relapses occur within the planned duration of the study. However, it is

unclear how well the frequency of relapses in the 2 pivotal trials reflect the broader NMOSD patient population in clinical practice.

Dosing of satralizumab and background therapies was consistent with clinical practice, although the expert consulted stated that in Canada, corticosteroids would not typically be used as longer-term preventive therapy. In Study 898, almost half of the patients were receiving monotherapy with corticosteroids, which was more frequent than would be expected in Canada. Of note, the pre-filled syringe device was not available during the double-blind study period, but it was used to administer satralizumab doses during the extension period.

In both trials, the treatment effect of satralizumab was assessed based on the time to first relapse, which was protocol-defined and committee-adjudicated. The generalizability of the observed effect to real-world clinical practice would largely rely on the agreement between the more stringent protocol-defined relapses and clinical relapse as determined by individual attending physicians. Some differences were noted in the treatment effects based on protocol-defined versus clinical relapse. Although adjudicated relapses are thought to be the more robust and reproducible measure, there may be issues with the generalizability of the results to clinical practice, where strict criteria are not used to identify relapses.

Due to the recognized limitations of the design of the studies, the benefit on patient-important outcomes such as disability, visual acuity, and health-related quality of life was uncertain. This impacts the extrapolation of the benefit that a reduction in relapses has on longer-term disability or health-related quality of life in the target population. Determining the comparative efficacy and the optimal place in therapy for satralizumab may be challenging, given the lack of head-to-head studies comparing satralizumab with eculizumab or other immunosuppressants.

Indirect Evidence

Objectives and Methods for the Summary of Indirect Evidence

No direct evidence was found that compared satralizumab with other relapse-prevention treatments for NMOSD; therefore, a search was conducted for indirect evidence that met the patient, intervention, comparator, and outcome criteria listed in Table 5. A focused literature search for ITCs in patients with NMOSD was run in MEDLINE All (1946–) and Embase (1974–) on November 19, 2020. No limits were applied to the search. A single researcher screened the search results for relevant studies. Three articles were found in the literature, but none met the inclusion criteria for this review.

The sponsor supplied an indirect comparison that was used to inform the pharmacoeconomic model.[31] This section provides a summary and critical appraisal of the submitted indirect comparison.

Description of Indirect Comparison

Methods of ITC

Objectives

The objective of the indirect comparison was to estimate the relative treatment effects of satralizumab and other therapies that have or could potentially have a regulatory licence for the treatment of NMOSD at the launch of satralizumab.

Study Selection Methods

The authors of the ITC conducted a systematic literature review for English language RCTs, prospective single-arm studies or case series, or prospective comparative observational studies in patients (≥ 12 years of age) with NMOSD. Interventions included satralizumab, eculizumab, inebilizumab, rituximab, azathioprine, mycophenolate mofetil, methotrexate, tocilizumab, mitoxantrone, cyclophosphamide, IV immunoglobin, and corticosteroids, compared with each other or placebo. Efficacy outcomes of interest included time to relapse, ARR, proportion of relapse-free patients, change in pain, fatigue, health-related quality of life, and disability, as measured in the satralizumab pivotal trials. Adverse events, withdrawals due to adverse events, infections, and other harms were also examined.

Multiple databases were searched (MEDLINE, PubMed, Embase, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, Database of Abstracts of Reviews of Effects), as well as clinical trial registries, neurological conference proceedings, and reference lists of relevant systematic reviews. Searches were conducted in January and February 2019 with an update in July 2019.

No information was provided on the methods used to conduct screening or data extraction. The study quality of the RCTs was assessed using the Cochrane risk-of-bias tool.

ITC Analysis Methods

The ITC authors conducted a feasibility study to evaluate if it was appropriate to combine data from the identified studies. The study designs, populations enrolled, interventions and co-interventions, outcome definitions, and follow-up time of the trials were compared to assess for heterogeneity. Although the systematic review included several treatments for NMOSD, the ITC was limited to RCTs of drugs that did have, or could potentially have, regulatory approval for the treatment of NMOSD at the time of satralizumab's launch. Several sources of heterogeneity were identified by the ITCs authors; however, they opted to conduct the ITC.

Bayesian NMA methods were used to combine data from RCTs (Table 18). A random-effects model was selected for the base-case analysis due to the clinical heterogeneity across trials. Two networks were constructed: 1 that included patients who were AQP4 IgG positive, and a second based on the ITT population of the included studies. In the AQP4 IgG–positive network, the 2 satralizumab studies (studies 898 and 900) were pooled into 1 node, whereas in the ITT network, each satralizumab study was kept as a separate node. The ITC authors stated that the AQP4 IgG–positive network was the preferred model for the comparison between satralizumab and eculizumab, and justified pooling the subgroup data from the 2 satralizumab trials in order to create a population that was more similar to the PREVENT trial.

Vague priors were used for all parameters except for the between-study variance, which used informative priors that were selected from the meta-analyses by Turner et al.[40,41] The authors stated that informative priors were used due to the sparse network. Because the EDSS ranges from 0 to 10, the prior for the change in EDSS score was bound from −10 to 10. The analysis was conducted on the Roche BEE environment using R version 3.4.2 or above and JAGS version 4.6.0. A total of 50,000 iterations were run with a burn-in of 12,500 iterations. No information was provided on how convergence was assessed.

Sensitivity analyses were run, including a fixed-effects model, and a random-effects model that used alternate informative priors (twice as large between-trial heterogeneity). The authors identified the following potential effect modifiers: baseline ARR, AQP4 status, concomitant immunosuppressant therapy, age, sex, race, weight, and active disease. Other than AQP4 status, the authors stated that due to the small number of trials and sparsity of the data, it was not feasible to conduct additional sensitivity or meta-regression analyses to explore the impact of these potential effect modifiers. Model fit was assessed based on deviance information criterion values (a difference of fewer than 5 points was deemed inconsequential) and residual deviance.

The outcomes analyzed included the time to first protocol-defined relapse (as defined in each study), ARR, proportion of relapse-free patients at 48 weeks, change in EDSS at 48 weeks, withdrawal due to adverse events, all-cause discontinuation, and rate of serious infection. Additional analyses were run using the data for time to first clinical relapse and treated clinical relapse from the satralizumab studies, and the closest comparable relapse definition from the other trials. The authors of the ITC stated that it was not possible to assess other outcomes due to a lack of data from the comparator trials.

Table 18: ITC Analysis Methods

	Sponsor-submitted ITC
ITC methods	Bayesian random-effects model as follows:[a] • time to relapse: normal likelihood, proportional hazards model • proportion relapse-free, adverse events: binomial likelihood, logit link • ARR, adverse event rate: Poisson likelihood, log link • change in EDSS: normal likelihood
Priors	Vague priors were used except for between-study variance, which used informative priors (Turner [2012][41] and Turner [2015][40])
Assessment of model fit	DIC, residual deviance
Assessment of consistency	NA (no closed loops)
Assessment of convergence	NR
Outcomes	Time to first protocol-defined relapse,[b] clinical relapse, and treated clinical relapse; ARR; proportion of relapse-free patients at 48 weeks; change in EDSS at 48 weeks; WDAE; all-cause discontinuation; rate of serious infection
Follow-up time points	End of study (relapse, ARR, safety), 48 weeks or end of study (EDSS and proportion relapse-free)
Construction of nodes	• Network 1 for AQP4 IgG–positive population: pooled satralizumab studies into 1 node • Network 2 for ITT population: separate nodes for satralizumab monotherapy (Study 900) and add-on to IST (Study 898) trials

	Sponsor-submitted ITC
Sensitivity analyses	Fixed-effects model; random-effects model with alternate informative priors
Subgroup analysis	AQP4 IgG positive, ITT population
Methods for pairwise meta-analysis	NA

AQP4 = aquaporin 4; ARR = annualized relapse rate; DIC = deviance information criterion; EDSS = Expanded Disability Status Scale; IgG = immunoglobulin G; IST = immunosuppressive therapy; ITC = indirect treatment comparison; ITT = intention to treat; NA = not applicable; NR = not reported; WDAE = withdrawal due to adverse event.

[a] For safety analyses that did not converge, a continuity correction was applied by adding 0.5 to cells with 0 events and models were run on the log odds ratio scale with a normal likelihood.

[b] Used to inform the pharmacoeconomic model.

Source: CADTH Common Drug Review submission for Enspryng.[31]

Results of ITC

Summary of Included Studies

A total of 7 RCTs and 28 non-RCTs met the inclusion criteria for the systematic review; however, the ITC was limited to those drugs that are expected to receive regulatory approval for use in NMOSD. As a result, 4 RCTs were included in the ITC: 2 satralizumab studies (Study 898 [SAkuraSky][29] and Study 900 [SAkuraStar][30]); 1 eculizumab study (PREVENT[20]); and 1 inebilizumab study (N-MOmentum[42]).

There were differences across trials in the populations enrolled based on age, diagnosis of NMO or NMOSD and AQP4 IgG positive status, EDSS score, frequency of recent attacks, and prior and concurrent immunosuppressants (Table 19). In 2 trials, patients were receiving no background immunosuppressants during the study period;[30,42] in 1 study, all patients were on background immunosuppressants,[29] and 1 study included a mixed population (76% received immunosuppressants).[20] All studies were event-driven trials that assessed the time to the first relapse as the primary outcome; however, the definitions of relapse varied across trials. The trial duration was similar for the eculizumab and satralizumab studies (approximately 90 weeks), but not for the inebilizumab trials, which were approximately 28 weeks in duration. The sample size ranged from 55 to 230 patients.

The authors rated all 4 trials as having a low risk of bias for most items on the Cochrane risk-of-bias tool. Study 898, Study 900, and the N-MOmentum trials were rated as having an unclear risk of bias related to selective outcome reporting. In addition, the N-MOmentum study was rated as having an unclear risk of bias related to incomplete outcome data. All 4 trials were rated as having other potential sources of bias. For N-MOmentum, this was due to the early termination of the study and baseline imbalances by race or ethnicity. In the PREVENT study, baseline imbalances in the proportion diagnosed with NMO (versus NMOSD) and with prior rituximab treatment were noted. In addition, the proportion of patients who discontinued was higher in the eculizumab group (17%) than in the placebo group (6%). Studies 898 and 900 were not powered or designed to test for differences for subgroups based on AQP4 status, and randomization was not stratified by this factor.

Table 19: Assessment of Homogeneity for Sponsor-Submitted ITC

	PREVENT (eculizumab)	Study 898 (satralizumab)	Study 900 (satralizumab)	N-Momentum (inebilizumab)
Number of patients	143	83 ITT (55 in AQP4 subgroup)	95 ITT (64 in AQP4 subgroup)	230
Clinical trial eligibility criteria	• Adults • NMO or NMOSD • AQP4 IgG positive only • ≥ 2 relapses in past year or ≥ 3 in past 2 years with 1 in past year • EDSS ≤ 7.0	• Patients ≥ 12 years • NMO or NMOSD • AQP4 IgG positive or negative • 2 relapses in past 2 years with 1 in past year • EDSS ≤ 6.5	• Adults • NMO or NMOSD • AQP4 IgG positive or negative • 1 relapse in past year • EDSS ≤ 6.5	• Adults • NMO • AQP4 IgG positive or negative • 1 attack in past year or 2 attacks in past 2 years • EDSS ≤ 8.0
Disease severity	Baseline ARR 2.0	Baseline ARR 1.5	Baseline ARR 1.4	NR
Comparators	Placebo (± background IST)	Placebo plus background IST	Placebo	Placebo
Treatment history	• Background IST allowed) at stable doses (24% had no background IST) • No prior rituximab or mitoxantrone for 3 months or IVIG for 3 weeks or prednisone doses greater than 20 mg per day or the equivalent	• Receiving background IST therapy (azathioprine, mycophenolate mofetil, and/or corticosteroids) • Excluded patients who had received other treatments for NMOSD or MS in past 6 months or longer	• No background therapy • Excluded patients who had received other treatments for NMOSD or MS in past 6 months or longer	• No background therapy • Excluded patients currently on IST
Definitions of end points	• Protocol-defined relapse: new onset or worsening of neurological symptoms with an objective change on examination that persisted for more than 24 hours and symptoms that were attributable to NMOSD, with onset preceded by at least 30 days of stability; adjudicated by central committee • On-trial investigator determined relapse (not adjudicated)	• Protocol-defined relapse that met EDSS criteria and was assessed within 7 days of onset; adjudicated by central committee • Clinical relapse defined as any relapse reported by investigator (not adjudicated) • Treated clinical relapse (not adjudicated)	• Protocol-defined relapse that met EDSS criteria and was assessed within 7 days of onset; adjudicated by central committee • Clinical relapse defined as any relapse reported by investigator (not adjudicated) • Treated clinical relapse (not adjudicated)	• Protocol-defined relapse based on 18 study-specific clinical or MRI criteria for NMO, adjudicated by central committee
Timing of end point evaluation or trial duration	Median treatment duration 91 weeks (placebo: 53 PYs; eculizumab: 173 PYs)	Median treatment duration placebo: 33 weeks (59.5 PY); satralizumab: 107 weeks (78.5 PYs)	Median treatment duration placebo: 55 weeks (40.6 PYs); satralizumab: 92 weeks (115.2 PYs)	Study duration 28 weeks

	PREVENT (eculizumab)	Study 898 (satralizumab)	Study 900 (satralizumab)	N-Momentum (inebilizumab)
Withdrawal frequency	Differential losses in placebo (6%) and eculizumab group (17%)	Differential losses in placebo (24%) and satralizumab group (7%)	• Placebo: 13% • Satralizumab: 11%	• Placebo: 4% • Inebilizumab: 3%
Study design	Event-driven: Stopped early (after 23 of 24 planned relapses had occurred)	Event-driven	Event-driven (or 1.5 years of follow-up)	Event-driven: Stopped early for efficacy
Percentage with relapse in placebo (ARR)	43% (0.35)	43% (0.32)	50% (0.41)	39% (NR) 28-week follow-up

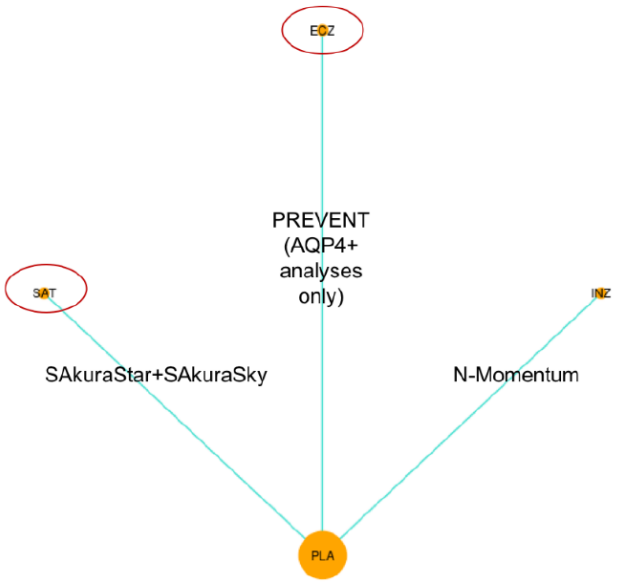
AQP4 = aquaporin 4; ARR = annualized relapse rate; EDSS = Expanded Disability Status Scale; IgG = immunoglobulin G; IST = immunosuppressive therapy; ITC = indirect treatment comparison; ITT = intention to treat; IVIG = intravenous immunoglobulin; NMO = neuromyelitis optica; NMOSD = neuromyelitis optica spectrum disorder; MS = multiple sclerosis; NR = not reported; PY = patient-year.

Source: CADTH Common Drug Review submission for Enspryng.[31] Pittock et al. (2019),[20] Cree et al. (2019),[42] and Clinical Study Reports for Study 898[10] and Study 900.[11]

Results

The network for the AQP4 IgG–positive population is shown in Figure 4. The NMA for time to first protocol-defined relapse and withdrawals due to adverse effects included 4 trials (3 nodes), and all other analyses included 3 trials (2 nodes; no data for inebilizumab).

Figure 4: Network Diagram for NMA in AQP4 IgG–Positive Population



AQP4+ = aquaporin 4 positive; ECZ = eculizumab; IgG = immunoglobulin G; INZ = inebilizumab; NMA = network meta-analysis; PLA = placebo; SAT = satralizumab.

Source: CADTH Common Drug Review submission for Enspryng.[31]

The NMA results for the time to first protocol-defined relapse did not differentiate between satralizumab, eculizumab, and inebilizumab, and showed wide 95% CrIs, indicating the lack of precision of the results (Table 20). Contrary to the head-to-head studies, no statistically significant differences were found between satralizumab and placebo based on the

random-effects NMA. A similar pattern of results was observed for the analyses of ARR, proportion of relapse-free patients at 48 weeks, change in EDSS score at 48 weeks, withdrawals due to adverse events, and rate of serious infections (Table 20).

For all the outcomes reported, model fit was similar (i.e., difference in DIC values < 5) for the base-case random-effects model, the fixed-effects model, and the sensitivity analysis model that used alternate priors for the between-study heterogeneity parameter. Across outcomes, the results of the sensitivity analyses were similar to the base-case analysis, with the fixed-effects model showing a narrower CrI. However, the authors of the ITC stated that the requirements of the fixed-effects model were too stringent for this rare disease; thus, a random-effects model was more appropriate. The analyses of time to first clinical relapse or treated clinical relapse showed findings similar to the time to protocol-defined relapse.

Table 20: Results for NMA of AQP4 IgG–Positive Patients

Satralizumab versus comparator	Time to protocol-defined relapse[a] HR (95% CrI)	ARR[b] rate ratio (95% CrI)	Proportion of relapse-free patients at 48 weeks[b] OR (95% CrI)	Change in EDSS at 48 weeks difference (95% CrI)	WDAE[a] OR (95% CrI)	Serious infections[b] rate ratio (95% CrI)
Placebo	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Eculizumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

AQP4 aquaporin 4; ARR = annualized relapse rate; CrI = credible interval; EDSS = Expended Disability Status Score; HR = hazard ratio; IgG = immunoglobulin G; NMA = network meta-analysis; OR = odds ratio; RCT = randomized controlled trial; WDAE = withdrawal due to adverse events.

[a] Random-effects model included data from 4 RCTs.

[b] Random-effects model included data from 3 RCTs.

Source: CADTH Common Drug Review submission for Enspryng.[31]

Critical Appraisal of ITC

The ITC was based on a systematic literature search; however, only 4 RCTs were included in the NMA and these trials enrolled a small number of patients with this rare disease. As a result, the network was sparse and did not include off-label immunosuppressants that are currently used to treat NMOSD. Moreover, the NMA most relevant to this review was based on AQP4 subgroup data from Study 898 and Study 900, which further limited the sample size (55 and 64 patients). Studies 898 and 900 were not designed or powered to test for differences for this subpopulation. Randomization was not stratified by AQP4 status in either Study 898 or Study 900; consequently, the analyses based on these subgroups break randomization. Furthermore, it is unclear if pooling the data from the satralizumab trials was appropriate, given the study design and population differences between these studies. Data from the inebilizumab trial was included in the analyses despite its short duration (28 weeks) relative to the eculizumab and satralizumab trials. The model of withdrawals due to adverse events did not converge due to zero events in some cells; thus, a continuity correction was used (i.e., 0.5 events added to zero cells) as per the planned statistical analysis. However, this correction is known to introduce bias.

The feasibility of conducting this analysis appears to have been questioned by the authors of the ITC, who stated “there were many differences between populations, study designs, effect modifiers and end point definitions in these 4 trials which present severe

limitations…as they cannot be adjusted for (p. 19)."[31] Although the authors used accepted methods to conduct the NMA, the presence of substantial clinical heterogeneity between trials calls into question whether it was appropriate to run the NMA. There is a high degree of uncertainty in the results of the NMA, which is reflected in the wide CrI observed. As the authors of the ITC state in their conclusions, "Credible intervals were very wide in this rare disease setting, creating so much uncertainty that the models are not that informative" (p. 70).[31]

No conclusions could be drawn from the indirect comparison due to the limitations of the NMA; thus, the comparative efficacy and safety of satralizumab versus eculizumab remains unknown.

Other Relevant Evidence

This section includes the open-label extension periods of Study 898 and Study 900, which were included in the sponsor's submission to CADTH and were considered to address important gaps in the evidence included in the systematic review.

Long-Term Extension Period for Study 898 and Study 900

Two open-label extension periods for Study 898 and Study 900 were summarized to provide evidence regarding safety and efficacy data for a minimum of at least 48 weeks and 96 weeks for studies 898 and 900, respectively.[10,11] In addition, the sponsor submitted a pooled efficacy and safety analysis of Study 898 and Study 900.[43]

Methods

Patients were eligible to enter the extension period at least 31 days after they experienced a protocol-defined relapse or a clinical relapse treated with rescue therapy (Study 898), or a committee-adjudicated protocol-defined relapse (Study 900) after the disease stabilized. Patients who completed the double-blind period without experiencing a relapse were eligible to enter the extension period 4 weeks after the administration of their last dose of the double-blind study drug.

During the extension periods, patients received open-label satralizumab injections at the same dosing schedule as the double-blind period (120 mg SC pre-filled syringe at weeks 0, 2, and 4, and every 4 weeks thereafter).

In the extension period of Study 898, patients were permitted to continue the same baseline and maintenance therapies as in the double-blind period. For Study 900, there was no description of maintenance therapies that were permitted during the extension period. Lastly, for both studies, patients were permitted to receive rescue therapy and pain medications as per the double-blind period.

For both studies, the all-treated population is presented and includes all patients who received at least 1 dose of satralizumab in either the double-blind or extension periods, starting on the date of the first satralizumab administration. There was no baseline demographic or disease characteristic data available for the all-treated population for the extension period.

The following data are included in this section: summary of harms for the all-treated population (N = 145) based on interim analyses with a data cut-off of June 2018 for Study 898 and October 2018 for Study 900, and supplemental pooled safety and efficacy

data for the overall satralizumab treatment period, which includes data from the double-blind and extension period up to June 7, 2019 (N = 166).

For the summary of harms, the adverse events per 100 PYs was calculated in the same way as described in the main body of this report. The pooled efficacy and safety data were submitted to CADTH as a conference presentation slide deck and, as such, there was limited information regarding the statistical methods utilized. In the pooled analysis, time to first investigator-assessed protocol-defined relapse in the combined double-blind and extension periods was presented. This includes relapses considered by the investigator to meet protocol-defined relapse criteria, but these events were not confirmed by an independent clinical event committee. The HR and stratified P value were reported for the pooled efficacy data.

Patient Disposition

The patient disposition for the extension periods as reported in the Clinical Study Report (up to October 2018) is summarized in Table 21. There were [redacted] patients [redacted] and [redacted] patients who entered the extension period of Study 898 and Study 900, respectively. Overall, [redacted] of patients in the placebo groups, versus [redacted] of patients in the satralizumab groups in Study 898 and Study 900, respectively, entered the extension period.

Withdrawals from the extension periods were similar between groups for each study, with [redacted] of patients withdrawing from the extension periods. Of the patients initially allocated to the placebo group during the double-blind period, [redacted] in Study 898 and Study 900, respectively, withdrew from the extension period due to lack of efficacy.

No disposition data were reported for the pooled efficacy or safety data.

Table 21: Patient Disposition for Extension Period of Study 898 and Study 900

	Study 898		Study 900	
	Placebo	Satralizumab	Placebo	Satralizumab
Entered the OLE period from the DB period, N	[redacted]	[redacted]	[redacted]	[redacted]
Withdrew from the OLE, n (%)[a]	[redacted]	[redacted]	[redacted]	[redacted]
Reason for withdrawal, n (%)[a]				
Adverse event	[redacted]	[redacted]	[redacted]	[redacted]
Lack of efficacy	[redacted]	[redacted]	[redacted]	[redacted]
Refused treatment or did not cooperate	[redacted]	[redacted]	[redacted]	[redacted]
Withdraw of consent	[redacted]	[redacted]	[redacted]	[redacted]
All-treated population, N	[redacted]	[redacted]	[redacted]	[redacted]

DB = double blind; OLE = open-label extension.

[a] Percentages are derived from the number of patients entering the extension period in each group. Includes data up to June 2018 for Study 898 and October 2018 for Study 900.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Summary of Harms

The summary of harms data for the all-treated population (N = 145, including data up to October 2018) is presented in Table 22. The median treatment duration was 140 weeks (range = 4 to 224), and 96 weeks (range = 5 to 206) for the all-treated population of Study 898 and Study 900, respectively. The median number of satralizumab doses received was 37 (range = 3 to 58) and 26 (range = 2 to 51) in Study 898 and Study 900, respectively.

The occurrence of adverse events was similar between studies 898 and 900 (94% versus 95%). Patients in Study 898 reported more serious adverse events than in Study 900 (22% versus 6%). Mental status change occurred in 2 patients in Study 900; otherwise, no serious adverse events were reported in more than 1 patient in either trial. A greater proportion of patients stopped treatment due to adverse events in Study 898 than in Study 900 (9% versus 1%). There were no deaths reported in either study. Regarding notable harms of interest to the CADTH review, more patients in Study 898 presented with infections and infestations (79% versus 55%), with serious infections (11% versus 8%), and with potential opportunistic infections (15% versus 5%). However, the frequency of injection-related reactions was similar in both studies (14% and 15%). There were no cases of anaphylaxis or hepatotoxicity in either study.

Table 22: Summary of Harms — All-Treated Population up to October 2018

	Study 898		Study 900	
	At least 1 dose of satralizumab N = 65		At least 1 dose of satralizumab N = 80	
	n (%)	AEs per 100 PYs	n (%)	AEs per 100 PYs
Adverse event				
Patients with ≥ 1 AE	61 (94)	457.4	76 (95)	452.7
Patients with ≥ 1 SAE	14 (22)	14.6	13 (6)	12.9
Mental status change	0	0	2 (3)	NR
Patients who stopped treatment due to adverse events	6 (9)	3.8	1 (1)	0.6
Infections and infestations (SOC)	2 (3)	NR	1 (1) [pneumonia]	NR
Investigations (SOC)	2 (3)	NR	NR	NR
Skin and subcutaneous disorders (SOC)	1 (2)	NR	NR	NR
Vascular disorders (SOC)	1 (2)	NR	NR	NR
Deaths	0	0	0	0
Notable harms				
Infections and infestations (SOC)	51 (79)	145.7	44 (55)	92.0
Serious infection	7 (11)	5.1	6 (8)	3.5
Potential opportunistic infection	10 (15)	10.2	4 (5)	2.3
Injection-related reactions	9 (14)	16.5	12 (15)	15.2
Anaphylaxis	0	0	0	0

	Study 898		Study 900	
	At least 1 dose of satralizumab N = 65		At least 1 dose of satralizumab N = 80	
	n (%)	AEs per 100 PYs	n (%)	AEs per 100 PYs
Hepatotoxicity[a]	0	0	0	0
Hypersensitivity	NR	NR	NR	NR

AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; NR = not reported; PY = patient-year; SAE = serious adverse event; SOC = system organ class; ULN = upper limit of normal.

[a] Drug-induced liver injury defined as AST or ALT ≥ 3 times ULN with total bilirubin ≥ 2 times ULN or clinical signs of jaundice.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Pooled Harms Data for Study 898 and Study 900

Supplemental pooled harms data for Study 898 and Study 900, including data up to June 2019, is presented in Table 23. This data includes 166 patients who received at least 1 dose of satralizumab (equivalent to 437.7 PYs), and where the longest drug exposure was 5.5 years.

Overall, 92% of patients experienced an adverse event, with upper respiratory tract infections (23%), nasopharyngitis (22%), urinary tract infections (18%), and headache (16%) occurring most often. Serious adverse events were reported in 21% of patients, and 4% of patients withdrew due to an adverse event. Of the harms of special interest identified in the CADTH review protocol, infections and serious infections occurred in 66% and 10% of patients, respectively, while injection-related reactions were evident in 13% of patients. There were no deaths or anaphylaxis events reported.

Table 23: Summary of Pooled Harms up to June 2019

	Overall satralizumab treatment period[a] N = 166	
	n (%)	AEs per 100 PYs
AEs		
Patients with ≥ 1 AE	153 (92.2)	418.8
Upper respiratory tract infection	38 (22.9)	25.1
Headache	27 (16.3)	11.0
Nasopharyngitis	37 (22.3)	20.1
Urinary tract infection	29 (17.5)	18.5
Patients with ≥ 1 SAE	35 (21.2)	12.6
Patients who stopped treatment due to adverse events	7 (4.2)	1.8
Deaths	0	0
Notable harms, n (%)		
Infections	109 (65.7)	112.4
Serious infections	16 (9.6)	3.0
Potential opportunistic infection	NR	NR
Injection-related reactions	21 (12.7)	12.1
Hypersensitivity	NR	NR

	Overall satralizumab treatment period[a] N = 166	
	n (%)	AEs per 100 PYs
Anaphylaxis	0	0
Hepatotoxicity	NR	NR

AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; NR = not reported; PY = patient-year; SAE = serious adverse event.

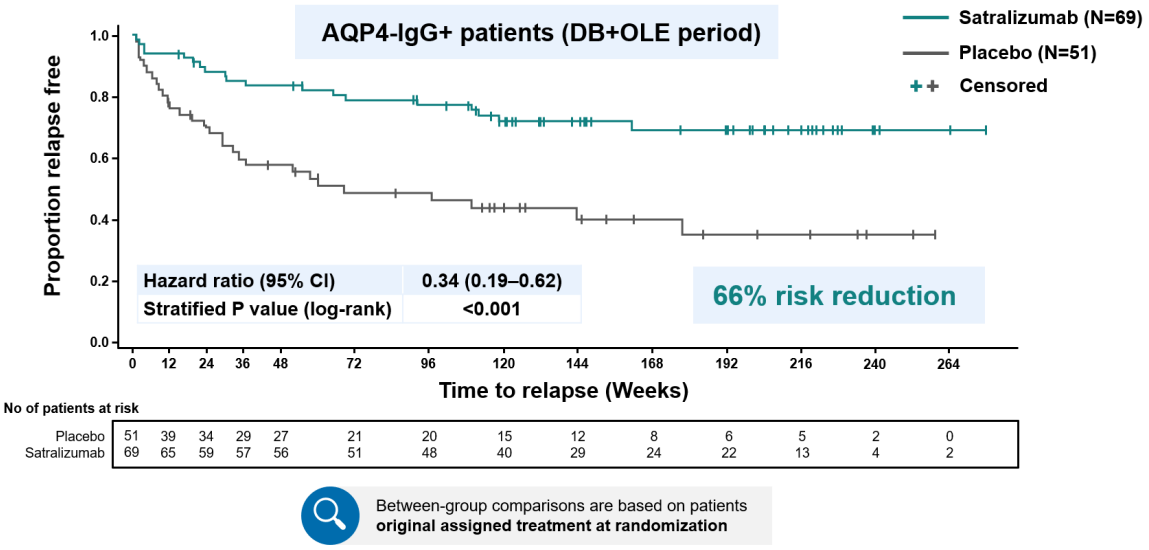
[a] The overall satralizumab treatment period includes cumulative data from the double-blind and extension periods up to a data cut-off of June 7, 2019.

Source: Additional sponsor-submitted data.[43]

Pooled Efficacy Data for Study 898 and Study 900

Figure 5 outlines the time to first investigator-assessed protocol-defined relapse for the AQP4 IgG–positive subgroup in the overall satralizumab period (cumulative data up to June 2019). The HR for patients treated with satralizumab compared with placebo was 0.34 (95% CI, 0.19 to 0.62; P < 0.001), in favour of satralizumab.

Figure 5: Pooled Time to Relapse in the AQP4 IgG–Positive Subgroup of Study 898 and Study 900



AQP4-IgG+, aquaporin-4 autoantibody seropositive; CI, confidence interval; DB, double-blind; ITT, intent to treat; OLE, open-label extension; PDR, protocol-defined relapse. The relapse data shown in the figure relate to investigator-assessed PDRs.

AQP4 = aquaporin 4; CI = confidence interval; DB = double blind; IgG+ = immunoglobulin G positive; OLE = open-label extension.

Note: Time to first investigator-assessed protocol-defined relapse in the combined double-blind and extension periods includes relapses considered by the investigator to meet the protocol-defined relapse criteria, but these events were not confirmed by an independent clinical event committee.

Source: Additional sponsor-submitted data.[43]

Critical Appraisal

Evidence from the extension data suggests acceptable tolerability of satralizumab 120 mg every 4 weeks. The longer-term harms data are consistent with the double-blind period, with infections being the most commonly reported adverse event.

Limitations of the extension data include potential selection bias, lack of blinding, and lack of comparator group, which may have affected the internal validity of the safety and efficacy results. Since completion of the double-blind period (or relapse) was an eligibility criterion for the extension period, patients who discontinued those trials due to adverse events or death were excluded. Thus, the population enrolled may be more tolerant of satralizumab and, with a survival bias, resulted in fewer adverse events being reported. In addition, the lack of blinding could have introduced bias in the reporting of subjective adverse events in favour of satralizumab if patients believed the drug was beneficial.

The definition of relapse used in the pooled efficacy analysis was different from that used in the double-blind period. In the pooled analysis, all relapse events were evaluated by the investigator and were not adjudicated by an independent committee. Assessment of relapses may also be affected by the lack of blinding, as expectations of therapy may affect reporting and evaluation of relapse symptoms. It is uncertain if pooling the trials was appropriate, given the differences between studies in background therapies, patient characteristics, and extension period entry criteria.

The patients enrolled in the extension period represent a select subset of patients with NMOSD, which may impact the data's external validity.

Discussion

Summary of Available Evidence

The systematic review included 2 pivotal, double-blind RCTs designed to evaluate the safety and efficacy of satralizumab 120 mg versus placebo in patients with NMO or NMOSD. Study 898 enrolled 83 adults and adolescents, of whom 55 (66%) were included in the AQP4 IgG–positive subgroup (i.e., the indicated population). All patients received background immunosuppressant treatment of azathioprine, mycophenolate mofetil, or corticosteroids during the trial. Study 900 enrolled 95 adults, including 64 (67%) who were AQP4 IgG positive. The primary outcome in both trials was the time to first adjudicated, protocol-defined relapse.

Among patients who were AQP4 IgG positive, the mean age of patients enrolled ranged from [redacted] years to [redacted] years per treatment group, including [redacted] of patients who were female, across the 2 trials. Most of the patients enrolled were White ([redacted]) or Asian ([redacted]). The mean ARR at baseline was [redacted] in Study 898 and [redacted] in Study 900, with a median baseline EDSS score of [redacted]

This review also includes a summary and appraisal of the sponsor-submitted ITC that compared the safety and efficacy of satralizumab versus eculizumab, as well as longer-term data from the open-label extension phase of Study 898 and Study 900.

Interpretation of Results

Efficacy

The pivotal studies met the primary end point and showed a statistically significant difference in the time to first protocol-defined relapse favouring satralizumab versus placebo in the overall study population, which included patients who were AQP4 IgG positive and negative. The between-group difference was driven mainly by the AQP4 IgG–positive subgroup, which showed a clinically relevant difference in the time to adjudicated relapse. According to the clinical experts consulted, the observed treatment effects were clinically important, as avoidance of relapses is the key goal of therapy. The treatment effects were similar in Study 898, where all patients were receiving background immunosuppressant therapy, and in the trial with no background relapse-prevention treatments (Study 900). Sensitivity analyses that used alternate definitions of relapse were generally consistent and showed point estimates favouring satralizumab; however, the CIs did not consistently exclude the null. This was the case for the analysis of time to clinical relapse, defined as any relapse identified by the investigator involved in patient care. For these analyses, more patients in the placebo group experienced a clinical relapse than in the satralizumab group, but the time-to-event analyses did not exclude the null. The primary outcome of time to protocol-defined relapse was considered more robust, as adjudication by a blinded event committee is expected to increase the validity and objectivity of the outcome by reducing inter-site variability and over-reporting bias.

Study 898 and Study 900 were designed as time-to-event trials, where patients completed the trial after having a protocol-defined relapse (or treated clinical relapse for Study 898). This design inherently emphasizes the efficacy of satralizumab on the first relapse, but it is not designed to assess its efficacy pertaining to subsequent relapses. While ARR results were reported and are a clinically relevant end point, the trials do not capture data on subsequent relapses (occurring after 30 days of the first relapse) because patients were

censored and therefore subsequent relapses would not have been captured, thereby likely underestimating the ARR. ARR was not part of the statistical testing hierarchy and the P value was not controlled for the type I error rate; thus, the data should be considered as supportive evidence for the effect of satralizumab in the overall population.

The trials failed to demonstrate a difference between groups in pain and fatigue symptoms measured at 24 weeks using the VAS for pain and FACIT-F scores for the ITT populations (key secondary outcomes). Given that satralizumab is not expected to directly impact pain or fatigue, the clinical expert indicated that the lack of effect observed was not surprising. Other secondary outcomes of the change in EDSS score, EQ-5D, or SF-36 scores also did not detect a difference between groups for the ITT population. The study, however, was not designed to fully capture the change in symptoms, disability, or health-related quality of life over time, as many patients did not have 24-week outcome data due to patients being withdrawn after experiencing a relapse or other early discontinuations. For the analyses of pain and fatigue, baseline values were carried forward for [redacted] of patients who had missing data at 24 weeks. Use of BOCF likely biases the results, particularly for those missing due to relapse who may be expected to have worse outcomes. This bias likely favours satralizumab, as there were more patients in the placebo group who experienced a relapse. A substantial percentage of patients had missing disability, visual acuity, and health-related quality of life scores at 24 weeks ([redacted]), with no imputation for missing data in the MMRM analyses. Given the magnitude of missing data, there is potential for the validity of these results to be affected. Moreover, there were no EDSS, visual acuity, or health-related quality of life data reported for the indicated population. As a result of these limitations, the impact of satralizumab on short-term disability or health-related quality of life is unclear, and it is uncertain if the reduction in relapses will translate to a benefit on longer-term irreversible disability. Neither study collected data on productivity or health care resource utilization; thus, the impact of satralizumab on these outcomes is unknown.

With respect to external validity, the clinical expert consulted by CADTH considered the baseline demographics and disease characteristics of the patients enrolled in the trials to be generally consistent with patients seen in the Canadian clinical setting although, in Study 898, the frequency of use of corticosteroids as longer-term relapse-prevention therapy was not consistent with Canadian clinical practice. Although adjudicated protocol-defined relapses are thought to be the more robust and reproducible measure, there may be issues with the generalizability of the results to clinical practice, where strict criteria are not used to identify relapses.

No head-to-head studies comparing satralizumab with eculizumab or other immunosuppressants were identified. The sponsor submitted an indirect comparison that estimated the relative treatment effects and safety of satralizumab versus eculizumab or inebilizumab. Bayesian NMA methods were used to combine data from 4 RCTs, including a network limited to patients who were AQP4 IgG positive that used pooled subgroup data from the 2 pivotal satralizumab trials. The NMA results for the time to first protocol-defined relapse did not differentiate between satralizumab, eculizumab, and inebilizumab, or between satralizumab and placebo, and showed wide 95% CrIs and high uncertainty. A similar pattern of results was observed for the analyses of ARR, proportion of relapse-free patients at 48 weeks, change in EDSS score at 48 weeks, withdrawals due to adverse events, and the rate of serious infections.

Although the NMA was conducted using accepted statistical methods, there were many differences between populations, study designs, effect modifiers, and end point definitions

in these 4 trials, which present severe limitations to the analyses. Due to the sparse network, which was based on pooled subgroup data, and the clinical heterogeneity between trials, the results of the NMA are highly uncertain and, thus, no conclusions can be drawn on the comparative efficacy and safety of satralizumab versus eculizumab in patients who are AQP4 IgG positive. Give the lack of direct evidence, and the limitations of the sponsor-submitted indirect comparison, determining the comparative efficacy and the optimal place in therapy for satralizumab may be challenging.

Harms

The occurrence of adverse events was generally similar between groups within trials. The proportion of patients who experienced an adverse event ranged from 75% to 95% in the placebo groups and from 90% to 92% in the satralizumab groups. After adjusting for follow-up time, the rate of adverse events was 495 to 514 events per 100 PYs among those assigned to placebo, and from 474 to 485 events per 100 PYs to those who received satralizumab. The most common adverse events were infections, which were reported by 62% to 68% of patients in the combination therapy trial (133 to 150 events per 100 PYs) and from 44% to 54% of patients in the monotherapy trials (100 to 163 events/100 PYs).

Serious adverse events were reported in 16% to 21% of patients assigned to placebo, and 17% to 19% of patients who received satralizumab, with a serious adverse event rate of 15 to 20 events per 100 PYs, and 12 to 17 events per 100 PYs in the placebo and satralizumab groups, respectively. More patients stopped treatment due to adverse events in the add-on therapy trial (Study 898: placebo 12%; satralizumab 7%) than in the monotherapy trial (Study 900: placebo 3%; satralizumab 2%). No deaths, hepatotoxicity, or anaphylaxis events were reported in either study.

Median treatment duration was shorter for patients who received placebo compared with satralizumab (33 to 55 weeks versus 92 to 107 weeks); thus, exposure time should be considered when comparing the proportion of patients who experienced an adverse event in each treatment group. Any numerical differences in the frequency of adverse events should be interpreted cautiously, given the differences in follow-up time and the small sample size of the trials.

Evidence from the extension period suggests acceptable tolerability of satralizumab administered every 4 weeks. The longer-term harms data are consistent with the double-blind period, with infections being the most commonly reported adverse event (92 to 146 infections reported per 100 PYs of follow-up). The rate of serious adverse events ranged from 13 to 15 events per 100 PYs, and discontinuation of treatment due to adverse events ranged from 0.6 to 3.8 events per 100 PYs. The data are limited by selection bias, lack of blinding, and lack of a comparator group, which may affect the internal or external validity of the results.

The sponsor-submitted indirect comparison had a number of limitations that affect the validity of the results; thus, no conclusions can be drawn on the comparative safety of satralizumab versus eculizumab.

Other Considerations

The clinical experts consulted by CADTH indicated there is the potential off-label use of satralizumab in patients less than 12 years of age and those who are AQP4 IgG negative. Although satralizumab is not approved for use in patients who are negative for AQP4 antibodies, the clinical experts stated that the mechanism of action of satralizumab (inhibition of interleukin-6) provides multiple mechanisms of immunomodulation and may impact the mechanisms affecting AQP4 negative NMOSD patients. Moreover, patients who are AQP4 negative have no approved or soon-to-be approved drugs available to them as eculizumab, and the FDA-approved inebilizumab, are only approved for use in AQP4 IgG–positive patients.

Pediatric-onset NMOSD is very rare, but the disabilities associated with this disease that occur at a young age (i.e., loss of vision) can have an important lifelong impact. The experts stated that despite the limited clinical data, there is potential for off-label use in children less than 12 years of age.

Conclusions

In patients with AQP4 antibody–positive NMOSD, fewer patients treated with satralizumab experienced an adjudicated relapse, relative to placebo, when administered as monotherapy or in combination with immunosuppressants. The between-group differences were considered clinically meaningful based on clinical expert input.

The 2 pivotal trials did not demonstrate an effect for satralizumab on pain or fatigue symptoms measured using a VAS for pain or FACIT-F scale at 24 weeks. No conclusions can be drawn on the impact of satralizumab on disability or health-related quality of life, due to limitations in the design of the trials and the extent of missing data. No data were available to assess the effects on productivity or health care resource utilization.

Infections were the most commonly reported adverse event in the double-blind and open-label extension periods. Safety data were limited by the small sample size of the trials, the lack of blinding and comparator group, and the potential selection bias for the extension period.

Head-to-head trials comparing satralizumab with other immunosuppressants are lacking. The sponsor-submitted indirect comparison that estimated the relative treatment effects and safety of satralizumab versus eculizumab was limited by the sparse network and clinical heterogeneity between trials. The results of the indirect comparison were highly uncertain and, thus, no conclusions can be drawn on the comparative efficacy and safety of satralizumab versus eculizumab in patients who are AQP4 antibody positive.

Appendix 1: Literature Search Strategy

OVERVIEW	
Interface:	Ovid
Databases:	MEDLINE All (1946-present) Embase (1974-present) **Note:** Subject headings and search fields have been customized for each database. Duplicates between databases were removed in Ovid.
Date of Search:	November 19, 2020
Alerts:	Bi-weekly search updates until project completion
Study Types:	No search filters were applied
Limits:	No date or language limits were used Conference abstracts: excluded

SYNTAX GUIDE	
/	At the end of a phrase, searches the phrase as a subject heading
*	Before a word, indicates that the marked subject heading is a primary topic; or, after a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
MeSH	Medical Subject Heading
exp	Explode a subject heading
.ti	Title
.ab	Abstract
.dq	Candidate term word (Embase)
.ot	Original title
adj#	Requires terms to be adjacent to each other within # number of words (in any order)
.hw	Heading word; usually includes subject headings and controlled vocabulary
.kf	Author keyword heading word (MEDLINE)
.kw	Author keyword (Embase)
.pt	Publication type
.mp	Mapped term
.rn	Registry number
.yr	Publication year
medall	Ovid database code: MEDLINE All, 1946 to present, updated daily
oemezd	Ovid database code; Embase, 1974 to present, updated daily

MULTI-DATABASE STRATEGY	
Line #	**Search Strategy**
1	(Enspryng* or satralizumab* or sapelizumab* or sa237 or sa 237 or YB18NF020M).ti,ab,kf,ot,hw,rn,nm.
2	1 use medall
3	*satralizumab/
4	(Enspryng* or satralizumab* or sapelizumab* or sa237 or sa 237).ti,ab,kw,dq.
5	3 or 4
6	5 use oemezd
7	6 not (conference review or conference abstract).pt.
8	*2 or 7
9	remove duplicates from 8

CLINICAL TRIALS REGISTRIES		
ClinicalTrials.gov	Produced by the US National Library of Medicine. Targeted search used to capture registered clinical trials. [Search -- Studies with results Enspryng OR satralizumab],	
WHO ICTRP	International Clinical Trials Registry Platform, produced by the World Health Organization. Targeted search used to capture registered clinical trials. [Search terms – Enspryng OR satralizumab]	
Health Canada's Clinical Trials Database	Produced by Health Canada. Targeted search used to capture registered clinical trials. [Search terms – Enspryng OR satralizumab]	
EU Clinical Trials Register	European Union Clinical Trials Register, produced by the European Union. Targeted search used to capture registered clinical trials. [Search terms – Enspryng OR satralizumab]	

Grey Literature

Search dates:	November 10 to 16, 2020
Keywords:	Enspryng, satralizumab, neuromyelitis optica spectrum disorder
Limits:	None
Updated:	Search updated prior to the completion of stakeholder feedback period

Relevant websites from the following sections of the CADTH grey literature checklist *Grey Matters: A Practical Tool For Searching Health-Related Grey Literature* (https://www.cadth.ca/grey-matters) were searched:

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Drug and Device Regulatory Approvals
- Advisories and Warnings
- Drug Class Reviews
- Clinical Trials Registries
- Databases (free)
- Health Statistics
- Internet Search

Appendix 2: Excluded Studies

No studies were excluded at the second stage of screening.

Appendix 3: Detailed Outcome Data

Table 24: NMO and NMOSD Diagnostic Criteria

NMO	NMOSD

NMO as defined by Wingerchuk et al. (2006), which required the following:
• Optic neuritis
• Acute myelitis
• At least 2 of 3 supportive criteria:
  ○ contiguous spinal cord lesion identified on an MRI scan extending over 3 vertebral segments
  ○ brain MRI not meeting diagnostic criteria for MS
  ○ NMO immunoglobulin G (anti–AQP4 IgG) seropositive status
	NMOSD as defined by either of the following criteria with anti–AQP4 IgG seropositive status at screening (Wingerchuk, 2007):
	• Idiopathic single or recurrent events of longitudinally extensive myelitis (≥ 3 vertebral segment spinal cord MRI lesion)
	• Optic neuritis: recurrent or simultaneous bilateral[a]

AQP4 = aquaporin 4; IgG = immunoglobulin G; MS = multiple sclerosis; NMO = neuromyelitis optica; NMOSD = neuromyelitis optica spectrum disorder.

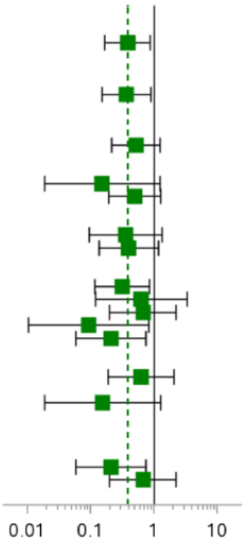
[a] Criteria for Study 900 also included single optic neuritis.

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Figure 6: Subgroup Analyses of Time to Protocol-Defined Relapse for Study 898 — Intention-to-Treat Population

Time to First Relapse (Protocol Defined Relapse) during Double-Blind Period by Subgroup, Intent-to-Treat Population
Protocol: SA-307JG

Baseline Risk Factors	Total	Placebo (N=42)		SA237 (N=41)		Hazard Ratio	95% CI	p-value (Log-rank)	Interaction p-value	Plot Hazard Ratio + 95% CI
	n	n	Events	n	Events					
All Patients	83	42	18	41	8	0.378	(0.164, 0.875)	0.0184		
Age Category									0.9374	
<18	7	3	1	4	1	0.000	(0.000, NE)	0.1573		
>=18	76	39	17	37	7	0.362	(0.149, 0.878)	0.0192		
Race Category									0.0965	
Japanese	21	10	3	11		0.000	(0.000, NE)	0.0499		
Non-Japanese	62	32	15	30	8	0.514	(0.213, 1.244)	0.1337		
Region*									0.2688	
ASIA	34	18	7	16	1	0.150	(0.018, 1.231)	0.0419		
EUROPE/OTHER	49	24	11	25	7	0.495	(0.191, 1.283)	0.1400		
Baseline ARR**									0.8994	
1	40	20	8	20	3	0.354	(0.094, 1.341)	0.1105		
>1	43	22	10	21	5	0.396	(0.134, 1.167)	0.0823		
NMO/NMOSD and AQP4 Status									0.4069***	
NMO	61	28	12	33	6	0.314	(0.115, 0.852)	0.0169		
NMOSD	22	14	6	8	2	0.628	(0.118, 3.329)	0.5813		
NMO and AQP4 Negative	28	14	6	14	5	0.663	(0.197, 2.235)	0.5047		
NMO and AQP4 Positive	33	14	6	19	1	0.092	(0.010, 0.827)	0.0110		
NMO/NMOSD and AQP4 Positive	55	28	12	27	3	0.208	(0.058, 0.750)	0.0086		
Baseline Treatment									0.4773	
AZATHIOPRINE	29	13	7	16	5	0.621	(0.188, 2.051)	0.4307		
MYCOPHENOLATE MOFETIL	12	8	2	4	1	0.000	(0.000, NE)	0.1025		
ORAL CSs	37	20	8	17	1	0.152	(0.018, 1.253)	0.0462		
AZATHIOPRINE+ORAL CSs	3			3	1	NE	(NE , NE)	NE		
MYCOPHENOLATE MOFETIL+ORAL CSs	2	1	1	1		0.000	(0.000, NE)	0.3173		
AQP4 Status - ELISA									0.1469	
Positive	55	28	12	27	3	0.208	(0.058, 0.750)	0.0086		
Negative	28	14	6	14	5	0.663	(0.197, 2.235)	0.5047		



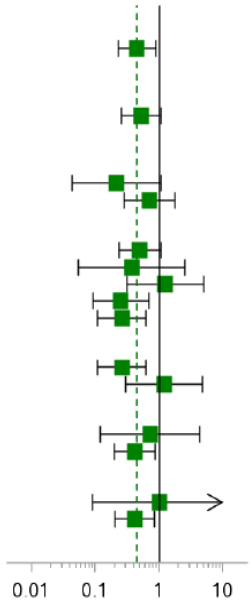
AQP4=Aquaporin-4, ARR=Annualized Relapse Rate, ELISA=Enzyme-Linked Immunosorbent Assay, NMO=Neuromyelitis Optica, NMOSD=Neuromyelitis Optica Spectrum Disorder, ORAL CSs=Oral Corticosteroids.
Protocol Defined Relapse: Adjudicated by the Clinical Endpoint Committee, EDSS assessment performed within 7 days of relapse reporting.
Model is stratified by baseline ARR (1, >1) and geographical region (Asia, Europe/other).
* Stratified by baseline ARR only. ** Stratified by geographic region only.
*** Interaction p-value only for NMO vs NMOSD subgroups.

Source: Clinical Study Report for Study 898.[10]

Figure 7: Subgroup Analyses of Time to Protocol-Defined Relapse for Study 900 — Intention-to-Treat Population

Time to First Relapse (Protocol Defined Relapse) during Double-Blind Period by Subgroup, Intent-to-Treat Population
Protocol: SA-309JG

Baseline Risk Factors	Total	Placebo (N=32)		SA237 (N=63)		Hazard Ratio	95% CI	p-value (Log-rank)	Interaction p-value	Plot Hazard Ratio + 95% CI
	n	n	Events	n	Events					
All Patients	95	32	16	63	19	0.450	(0.228, 0.889)	0.0184		
Geographic Region1									0.0709	
ASIA	10	5	3	5		0.000	(0.000, NE)	0.0101		
EUROPE/US/OTHER	85	27	13	58	19	0.519	(0.252, 1.068)	0.0701		
Geographic Region2									0.1068	
ASIA	10	5	3	5		0.000	(0.000, NE)	0.0101		
EUROPE/OTHER	27	11	6	16	3	0.212	(0.043, 1.056)	0.0371		
NORTH AMERICA	58	16	7	42	16	0.706	(0.285, 1.749)	0.4502		
NMO/NMOSD and AQP4 Status									0.6013***	
NMO	71	24	13	47	17	0.496	(0.234, 1.048)	0.0611		
NMOSD	24	8	3	16	2	0.369	(0.054, 2.539)	0.2986		
NMO and AQP4 Negative	30	9	3	21	10	1.250	(0.312, 5.005)	0.7525		
NMO and AQP4 Positive	41	15	10	26	7	0.251	(0.091, 0.689)	0.0043		
NMO/NMOSD and AQP4 Positive	64	23	13	41	9	0.261	(0.108, 0.627)	0.0014		
Anti-AQP4 Status									0.0223	
POSITIVE	64	23	13	41	9	0.261	(0.108, 0.627)	0.0014		
NEGATIVE	31	9	3	22	10	1.192	(0.298, 4.775)	0.8036		
Prior therapy*									0.5790	
B-CELL DEPLETING THERAPY	12	4	2	8	3	0.715	(0.119, 4.296)	0.7130		
IMMUNOSUPPRESANTS/OTHERS	83	28	14	55	16	0.415	(0.199, 0.868)	0.0159		
Most recent attack**									0.4842	
FIRST ATTACK	11	4	1	7	3	0.995	(0.090, 11.018)	0.9970		
RELAPSE	84	28	15	56	16	0.417	(0.205, 0.851)	0.0132		



AQP4=Aquaporin-4,NMO=Neuromyelitis Optica,
NMOSD=Neuromyelitis Optica Spectrum Disorder.
Protocol Defined Relapse: Adjudicated by the Clinical Endpoint Committee, EDSS assessment performed within 7 days of relapse reporting.
Model is stratified by Prior therapy (B-cell depleting therapy or Immunosuppressants/Others) and Most recent attack (First Attack or Relapse).
* Stratified by Most recent attack only. ** Stratified by Prior therapy only.*** interaction p-value only for NMO vs NMOSD subgroups.

AQP4 = aquaporin 4; CI = confidence interval; EDSS = Expanded Disability Status Scale; NMO = neuromyelitis optica; NMOSD = neuromyelitis optica spectrum disorder; SA237 = satralizumab; vs = versus.

Source: Clinical Study Report for Study 900.[11]

Table 25: Relapse-Free Patients and ARR in Study 898 and Study 900

	Study 898 (ITT)		Study 900 (ITT)	
	Placebo plus IST N = 42	Satralizumab plus IST N = 41	Placebo N = 32	Satralizumab N = 63
Protocol-defined relapse				
Relapse-free patients (%)				
Week 48	66.0	88.9	61.9	76.1
Week 96	58.7	77.6	51.2	72.1
ARR[a]				
Patient-years at risk	[redacted]	[redacted]	[redacted]	[redacted]
Number of patients with relapse	[redacted]	[redacted]	[redacted]	[redacted]
Unadjusted ARR (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Adjusted ARR (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Adjusted ARR ratio (95% CI)		[redacted]		[redacted]
P value		[redacted]		[redacted]

ARR = annualized relapse rate; CI = confidence interval; IST = immunosuppressive therapy; ITT = intention to treat.

[a] Total number of patient-years were calculated using patient event date or censor date. Unadjusted ARR calculated by taking the total number of relapses for all patients and dividing by the total number of patient-years followed. The adjusted ARR was adjusted for randomization stratification factors, with log-transformed time to censor or event included as an offset variable, negative binomial regression model.

[b] P value has not been adjusted for multiple testing (i.e., the type I error rate has not been controlled).

Source: Clinical Study Reports for Study 898[10] and Study 900.[11]

Appendix 4: Description and Appraisal of Outcome Measures

Aim

To describe the following outcome measures and review their measurement properties, including validity, reliability, responsiveness to change, and MID:

- relapse
- VAS for pain
- FACIT-F scale
- SF-36
- EQ-5D-3L
- modified Rankin Scale
- EDSS

Findings

Table 26: Summary of Outcome Measures and Their Measurement Properties

Outcome measure	Type	Conclusions about measurement properties	MID
Relapse	There were 2 methods to evaluate relapses: either by the attending physician or adjudicated via a committee of experts. A generic clinically assessed relapse was defined as a new onset of neurologic symptoms or worsening of existing neurologic symptoms with an objective change (clinical sign) on neurologic examination that persists for more than 24 hours, as confirmed by the treating physician. The signs and symptoms must be attributed to NMO (i.e., not caused by an identifiable cause, such as infection, excessive exercise, or excessively high ambient temperature). The relapse must be preceded by at least 30 days of clinical stability.	Not applicable in the context of measurement scales, although internal validity is reinforced by the blinding of the outcome assessors.	Not identified.
VAS for pain	Generic, 100 mm visual scale that captures pain intensity from “no pain” to “pain as bad as it could be.”	No studies on patients with NMOSD were found. In a cross-sectional study that included 52 patients with MS and 52 healthy patients as controls, participants were asked to score 15 statements related to pain, anxiety,	No MID studies were found for patients with NMOSD or MS.

Outcome measure	Type	Conclusions about measurement properties	MID
		fatigue, and QoL on an electronic VAS using either a smartphone or a tablet. Only in patients with MS did the electronic VAS showed acceptable test-retest reliability (ICC > 0.7), displayed floor and ceiling effects in more than 15% of patients, and had high random error but no systematic error. This study is limited in that test-retest reliability was evaluated within the same session.[44]	
FACIT– Fatigue scale	Five-point, 13-item, ordinal scale measuring a patient's level of daily fatigue over a 7-day recall period.	No studies on patients with NMOSD or MS were found.	No MID studies were found for patients with NMOSD or MS.
SF-36	A generic self-reported questionnaire consisting of 8 domains: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health. The SF-36 also yields 2 summary measures of physical health (the PCS measure) and mental health (MCS measure) derived from scale aggregates. Higher global scores are associated with better quality of life.	No studies on patients with NMOSD were found. The instrument has been validated in patients with MS and neurological disabilities. One HTA systematic review[45] with 7 studies and 3,142 patients showed acceptable reliability (Cronbach alpha of 0.70 for all subscales) and validity (with correlations ranging from 0.5 to 0.81) for all domains. Two studies showed good-to-excellent internal consistency for the total instrument and for all subscales (within PCS and MCS) with the exception of social function. Correlations between SF-36 subscales and impairment measures were weak (correlation coefficient ranging from 0.1 to 0.3). Inter-rater reliability between patients with disabilities and their caretakers was moderate (level of agreement ranging from 0.41 to 0.6). **Responsiveness:** No conclusions could be drawn regarding responsiveness from a study of 100 patients with MS.[46]	No MID studies were found for patients with NMOSD. Indirect evidence from patients with MS was obtained. MID ranges for the SF-36 domains were as follows: 4 to 9 points for physical functioning, 6 to 8 for role physical, and 6 to 7 for social functioning; for the PCS score, the MID was consistently 6.[47] There were no data available for the other domains.

Outcome measure	Type	Conclusions about measurement properties	MID
EQ-5D-3L	Generic preference-based HRQoL instrument consisting of a composite index score of 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.	One systematic review[48] assessing the EQ-5D (9 studies) in patients with MS: • Content validity. The EQ-5D includes certain domains, such as walking (mobility) and mood (anxiety/depression), that patients considered important to their quality of life, yet other critical domains such as fatigue and cognition are not included in the EQ-5D. • Convergent validity of impairment (gait, speed, severity) was moderate (pooled correlation estimate = 0.35; 95% CI, 0.25 to 0.45). For activity limitations, the pooled correlation was 0.51 (95% CI, 0.45 to 0.57). When EQ-5D was compared against measures evaluating HRQoL, the correlation value was 0.56 (95% CI, 0.54 to 0.59). • Discriminative validity was evaluated in 3 studies. The mobility item lacked discriminative ability. The EQ-5D was able to differentiate between all EDSS levels except between levels 3 and 4. • Test–retest reliability: The intra-class correlation coefficient for test–retest reliability of the EQ-5D was 0.81 (acceptable).	None identified for patients with NMOSD. An MID range of 0.033 to 0.074 is acceptable for the general population.[38]
mRS	The mRS is a generic, commonly used clinician-reported scale for measuring the degree of disability or dependence in the daily activities of people suffering with a neurological disability. The scale ranges from 0 (no disability) to 6 (death).	No studies on NMOSD or MS patients evaluating validity or reliability were identified. The instrument is reliable and has been well validated in patients suffering disability from stroke,[39] implying, however, an issue with its construct validity when applied to patients with NMOSD.	None identified for patients with NMOSD or MS.
EDSS	Ordinal clinical rating scale that ranges from 0 (normal neurologic examination) to 10 (death) in half-point increments. The KFS (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, cerebral, other) and ambulation are rated in the context of a standard neurological examination, and then these ratings (KFS scores) are used	No studies on NMOSD patients were identified. Validity has been established in patients with MS and it is usually used as the gold standard for evaluating new scales.[49] Reliability has been established in patients with MS and has low-to-moderate values, with inter-rater kappa values	No MID specific for NMOSD was found. Indirect estimates can be obtained from patients with MS; 1 such study found that a change of 1.5 points as a single score was considered enough deterioration from the patient perspective.[50] This

Outcome measure	Type	Conclusions about measurement properties	MID
	in conjunction with observations and information concerning the patient's mobility, gait, and use of assistive devices to assign an EDSS score.	between 0.32 and 0.76 for EDSS and between 0.23 and 0.58 for the individual functional systems. For scores below 3.5, reliability is regarded as good.[49]	was in agreement with a second study characterizing a 1.5-point increase from baseline 0 as important; from a baseline of 1 to 5.5, a 1-point increase was considered important, and from a baseline score of ≥ 6, a 0.5-point increase was considered important.[51]

CI = confidence interval; EDSS = Expanded Disability Status Scale; EQ-5D = EuroQol 5-Dimensions questionnaire; EQ-5D-3L = EuroQol 5-Dimensions 3-Levels questionnaire; HRQoL = health-related quality of life; ICC = intra-class correlation coefficient; FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue; HTA = health technology assessment; KFS = Kurtzke functional systems; MCS = mental component summary; MID = minimal important difference; mRS = modified Rankin Scale; MS = multiple sclerosis; NMO = neuromyelitis optica; NMOSD = neuromyelitis optica spectrum disorder; PCS= physical component summary; QoL = quality of life; SF-36 = Short Form (36) Health Survey; SRD = standardized response difference; SRM = standardized response mean; VAS = Visual Analogue Scale.

Source: Kos (2017),[44] Riemsma RP (2001),[45] Pfennings LE (1999),[46] Robinson (2009),[47], Kuspinar (2014),[48] Rooney (2019),[52] Banks (2007),[39] Meyer-Moock (2014),[49] de Groot (2006),[50] Goldman (2019).[51]

Visual Analogue Scale for Pain

The VAS for pain captures the self-rating for the current intensity of pain using a visual "thermometer" 100 mm in length. It ranges from no pain (best imaginable health state) to pain as bad as it could be (worst imaginable health state). An example of this scale is found in Figure 8. There was no information given in the Clinical Study Reports regarding the recall of the VAS for pain scale administered in Study 898 or Study 900.[10,11]

Figure 8: Visual Analogue Scale for Pain

Source: Clinical Study Report for Study 898.[10]

Measurement Properties

In a cross-sectional observational study that included 52 patients with MS and 52 healthy patients (controls), participants were asked to score 15 statements related to pain, anxiety, fatigue, and quality of life on an electronic VAS using either a smartphone or a tablet. In only patients with MS, the electronic VAS showed acceptable test–retest reliability (intra-class correlation coefficient > 0.7), displayed floor and ceiling effects in more than 15% of patients, and had high random error but no systematic error. This study is limited in that test–retest reliability was evaluated within the same session.[44]

Minimal Important Difference

None identified for patients with NMOSD or MS.

FACIT-F Scale

The FACIT-F scale is a 13-item tool that measures a patient's level of daily fatigue over the past week. The scale used within Study 898 and Study 900 is highlighted in Figure 9. The level of fatigue is measured on a 5-point scale ranging from 0 indicating "not at all" to 4, which indicates "very much."

Measurement Properties

There were no studies identified that evaluated the validity, reliability, responsiveness, or MID for the FACIT-F scale in patients with NMOSD or MS.

Figure 9: FACIT-F Scale

Below is a list of statements that other people with your illness have said are important. **Please circle or mark one number per line to indicate your response as it applies to the past 7 days.**

		Not at all	A little bit	Some-what	Quite a bit	Very much
HI7	I feel fatigued	0	1	2	3	4
HI12	I feel weak all over	0	1	2	3	4
An1	I feel listless ("washed out")	0	1	2	3	4
An2	I feel tired	0	1	2	3	4
An3	I have trouble starting things because I am tired	0	1	2	3	4
An4	I have trouble finishing things because I am tired	0	1	2	3	4
An5	I have energy	0	1	2	3	4
An7	I am able to do my usual activities	0	1	2	3	4
An8	I need to sleep during the day	0	1	2	3	4
An12	I am too tired to eat	0	1	2	3	4
An14	I need help doing my usual activities	0	1	2	3	4
An15	I am frustrated by being too tired to do the things I want to do	0	1	2	3	4
An16	I have to limit my social activity because I am tired	0	1	2	3	4

FACIT-F = Functional Assessment of Chronic Illness Therapy–Fatigue.

Source: Clinical Study Report for Study 898.[10]

The Short Form (36) Health Survey

The SF-36 is a generic health assessment questionnaire that is used to study the impact of chronic disease on health-related quality of life. The multi-item questionnaire contains 8 dimensions: physical functioning, role physical, bodily pain, general health, vitality, social role functioning, role emotional, and mental health.[35] SF-36 also provides 2 component summaries, the physical component summary (PCS) and the mental component summary (MCS), which are created by aggregating the 8 domains according to a scoring algorithm. The PCS and MCS and 8 dimensions are each measured on a scale of 0 to 100, which are T scores (mean of 50 and SD of 10) that have been standardized to the US general population. Thus, a score of 50 on any scale would be at the average or norm of the general US population, and a score 10 points lower (i.e., 40) would be 1 SD below the norm. An increase in score indicates improvement in health status on any scale. Most questions within the SF-36 ask for a 4-week recall, with 1 question prompting patients to recall over a year, while the rest of the questions are not based on a previous point in time.

Measurement Properties

No studies on patients with NMOSD were found. The instrument has been validated in patients with MS and neurological disabilities. One HTA (systematic) review[45] that included 7 studies and 3,142 patients showed moderate reliability (Cronbach alpha of 0.70 for all subscales) and validity (with correlations ranging from 0.5 to 0.81). Two studies showed good-to-excellent internal consistency for the total instrument and for all subscales, with the exception of social function. Correlations between SF-36 subscales and impairment measures were weak. Inter-rater reliability between patients with disabilities and their caretakers was moderate.

Responsiveness of this instrument was evaluated in 100 MS patients by completing the SF-36 scale 5 times over 2 years. In addition, this study assessed the reliability (intra-class correlation coefficient > 0.7) of the SF-36 tool. However, the authors concluded that in slowly progressive diseases (such as MS), evaluation of responsiveness is not feasible, as all the responsiveness measures used (smallest real difference, standardized response mean, effect size) were not greater than what can be attributed to random error or “noise” within the population.[46]

Minimal Important Difference

No estimates of the MID were found for patients with NMOSD. In general use, a change of 2 points in the SF-36 PCS and 3 points in the SF-36 MCS indicates a clinically meaningful improvement, as determined by the patient.[36]

Limitations

Summary scores of SF-36 in the MS patient population should be reported and interpreted with caution. This is the result of the inability to explain variability in the social functioning and SF-36 component scores. In addition, the SF-36 has been reported to overestimate the mental health of MS patients on the mental health summary scale.

Given that there is no direct assessment in patients with NMOSD, convergence and discriminative validity might be an issue, given the absence of studies in patients with this clinical condition.

The EuroQol 5-Dimensions 3-Levels Questionnaire

The EQ-5D-3L is a generic, standardized patient self-administered instrument that provides a simple, descriptive profile and a single index value for health status.[37] The EQ-5D comprises 5 dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension consists of 3 levels (some, moderate, extreme problems), generating a total of 243 theoretically possible health states. The response period is the day of assessment only.

Measurement Properties

No studies in patients with NMOSD were found. However, 1 systematic review[48] assessing the EQ-5D (9 studies) in patients with MS was available. In terms of the content validity of the EQ-5D, the instrument includes domains such as walking (mobility) and mood (anxiety/depression) that patients considered important to their quality of life, yet other critical domains such as fatigue and cognition are not included in EQ-5D.

Convergent validity of impairment (gait, speed, severity) was moderate (pooled correlation estimate = 0.35; 95% CI, 0.25 to 0.45). For activity limitations, the pooled correlation was 0.51 (95% CI, 0.45 to 0.57). When EQ-5D was compared against measures evaluating health-related quality of life, the correlation value was 0.56 (95% CI, 0.54 to 0.59). Discriminative validity was evaluated in 3 studies. The mobility item lacked discriminative ability. The EQ-5D was able to differentiate between all EDSS levels, except between levels 3 and 4.

In terms of reliability, the test–retest intra-class correlation coefficient of the EQ-5D was found to be acceptable, with a value of 0.81.

Minimal Important Difference

None identified for patients with NMOSD. It has been found that an MID range of 0.033 to 0.074 is acceptable for the general population.[38]

Limitations

Some issues with content validity for patients with MS and in consequence with NMOSD.

The Modified Rankin Scale

The modified Rankin Scale is a generic, commonly used clinician-reported scale for measuring the degree of disability or dependence in the daily activities of people who have suffered a neurological disability. The scale ranges from 0 (no disability) to 6 (death).

The scale runs from 0 to 6, running from perfect health without symptoms to death.

0 = No symptoms.

1 = No significant disability. Able to carry out all usual activities, despite some symptoms.

2 = Slight disability. Able to look after own affairs without assistance, but unable to carry out all previous activities.

3 = Moderate disability. Requires some help, but able to walk unassisted.

4 = Moderately severe disability. Unable to attend to own bodily needs without assistance, and unable to walk unassisted.

5 = Severe disability. Requires constant nursing care and attention, bedridden, incontinent.

6 = Dead.

Measurement Properties

No studies evaluating validity or reliability in patients with NMOSD or MS were identified. The instrument is reliable and has been well validated in patients suffering disability due to a stroke.[39]

Minimally Important Difference

None identified for patients with NMOSD or MS.

Limitations

The inter-judge reproducibility seems better if the assessment is tied with a semi-structured conversation. The convergence validity in patients with stroke has been assessed by comparing it to the scales for disability in the Barthel index. Given there is no direct assessment of this tool in patients with NMOSD, the construct and convergence validity remains unclear.

The Expanded Disability Status Scale

The EDSS is an ordinal clinical rating scale that ranges from 0 (normal neurologic examination) to 10 (death) in half-point increments. The Kurtzke functional systems (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, cerebral, other) and ambulation are rated in the context of a standard neurological examination, and FSSs are created for each system. Each FSS score is an ordinal clinical rating ranging from 0 to 5 or 6. These ratings are then used in conjunction with observations and information concerning the patient's mobility, gait, and use of assistive devices to assign an EDSS score. In NMOSD, it is used to assess the severity of patient relapse.

EDSS steps 1.0 to 4.5 refer to people who are fully ambulatory. A patient's disability can involve different functional systems reflected, for example, in an EDSS score of 4.0 (e.g., bilateral vision loss, severe ataxia, paresis in at least 2 limbs, and marked reduction in sensation in at least 1 limb) or involve different functional systems that may or may not be reflected in the EDSS score. For example, following a relapse of NMOSD, a range of changes in EDSS scores is possible, from 0 for an area postrema relapse (symptoms not captured by EDSS) to a higher score that reflects impairment of ambulation. EDSS steps 5.0 to 9.5 are defined by impairment to ambulation.

The EDSS is a method of quantifying disability in MS that replaced the Disability Status Scale used previously. The functional systems of the EDSS are: pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual; cerebral, and other.

Table 27: The Kurtzke Expanded Disability Status Scale

0.0	Normal neurological examination
1.0	No disability, minimal signs in 1 FS
1.5	No disability, minimal signs in more than 1 FS
2.0	Minimal disability in 1 FS
2.5	Mild disability in 1 FS or minimal disability in 2 FS
3.0	Moderate disability in 1 FS, or mild disability in 3 or 4 FS; fully ambulatory
3.5	Fully ambulatory but with moderate disability in one FS and more than minimal disability in several others
4.0	Fully ambulatory without aid, self-sufficient, up and about some 12 hours a day despite relatively severe disability; able to walk without aid or rest some 500 metres
4.5	Fully ambulatory without aid, up and about much of the day, able to work a full day, may otherwise have some limitation of full activity or require minimal assistance; characterized by relatively severe disability; able to walk without aid or rest some 300 metres
5.0	Ambulatory without aid or rest for about 200 metres; disability severe enough to impair full daily activities (work a full day without special provisions)
5.5	Ambulatory without aid or rest for about 100 metres; disability severe enough to preclude full daily activities
6.0	Intermittent or unilateral constant assistance (cane, crutch, brace) required to walk about 100 metres with or without resting
6.5	Constant bilateral assistance (canes, crutches, braces) required to walk about 20 metres without resting
7.0	Unable to walk beyond approximately 5 metres even with aid, essentially restricted to wheelchair; wheels self in standard wheelchair and transfers alone; up and about in wheelchair some 12 hours a day
7.5	Unable to take more than a few steps; restricted to wheelchair; may need aid in transfer; wheels self but cannot carry on in standard wheelchair a full day; may require motorized wheelchair
8.0	Essentially restricted to bed or chair or perambulated in wheelchair, but may be out of bed itself much of the day; retains many self-care functions; generally, has effective use of arms
8.5	Essentially restricted to bed much of day; has some effective use of arms, retains some self-care functions
9.0	Confined to bed; can still communicate and eat
9.5	Totally helpless bed patient; unable to communicate effectively or eat/swallow
10.0	Death due to MS

FS = functional system; MS = multiple sclerosis.

Source: Clinical Study Report for Study 898.[10]

Measurement Properties

No studies on patients with NMOSD were identified. One systematic review with 54 studies addressing the validity and reliability of the EDSS was identified in the literature.[49] Validity has been established in patients with MS and it is usually used as the gold standard for evaluating new scales in the MS population. However, there have been some criticisms related to its reliability.

In this same study, reliability has been assessed as being low to moderate, with inter-rater kappa values of between 0.32 and 0.76 for EDSS and between 0.23 and 0.58 for the

individual functional systems. For scores below 3.5, reliability is regarded as good (inter-rater reliability evaluated with a kappa statistic less than 0.4).

The review found that EDSS is sensitive to change in disease progression in patients with MS.

Minimal Important Difference

No MID specific for NMOSD was found. Indirect estimates can be obtained from patients with MS, where 1 study found that a change of 1.5 points as a single score was considered enough deterioration from the patient perspective.[50] This was in agreement with a second study that characterized a 1.5-point increase from baseline 0 as important; from a baseline of 1 to 5.5, a 1-point increase was considered important, and from a baseline score of 6 or greater, a 0.5-point increase was considered important.[51]

Limitations

No information was found on the measurement properties of the EDSS in patients with NMOSD. The scale has shown low to moderate reliability in patients with MS.

References

1. Wingerchuk DM, Lennon VA, Lucchinetti CF, Pittock SJ, Weinshenker BG. The spectrum of neuromyelitis optica. *Lancet Neurol.* 2007;6(9):805-815.
2. Sellner J, Boggild M, Clanet M, et al. EFNS guidelines on diagnosis and management of neuromyelitis optica. *Eur J Neurol.* 2010;17(8):1019-1032.
3. Collongues N, Ayme-Dietrich E, Monassier L, de Seze J. Pharmacotherapy for neuromyelitis optica spectrum disorders: current management and future options. *Drugs.* 2019;79(2):125-142.
4. Huda S, Whittam D, Bhojak M, Chamberlain J, Noonan C, Jacob A. Neuromyelitis optica spectrum disorders. *Clin Med (Lond).* 2019;19(2):169-176.
5. Wingerchuk DM, Banwell B, Bennett JL, et al. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology.* 2015;85(2):177-189.
6. Hamid SHM, Whittam D, Mutch K, et al. What proportion of AQP4-IgG-negative NMO spectrum disorder patients are MOG-IgG positive? A cross sectional study of 132 patients. *J Neurol.* 2017;264(10):2088-2094.
7. Marrie RA, Gryba C. The incidence and prevalence of neuromyelitis optica: a systematic review. *Int J MS Care.* 2013;15(3):113-118.
8. Etemadifar M, Nasr Z, Khalili B, Taherioun M, Vosoughi R. Epidemiology of neuromyelitis optica in the world: a systematic review and meta-analysis. *Mult Scler Int.* 2015;2015:174720.
9. Enspryng (satralizumab injection), solution, 120 mg/mL, subcutaneous [product monograph]. Mississauga (ON): Hoffmann-La Roche Limited; 2020 Jun 1.
10. Clinical Study Report: BN40898. A multicenter, randomized, addition to baseline treatment, double-blind, placebo-controlled, phase III study to evaluate the efficacy and safety of satralizumab in patients with Neuromyelitis Optica (NMO) and NMO Spectrum Disorder (NMOSD)- report no. 1089823 [internal sponsor's report]. Basel (CH): F. Hoffman La-Roche Ltd; 2019 May 27.
11. Clinical Study Report: study BN40900. A multicenter, randomized, double-blind, placebo-controlled, phase III study to evaluate the efficacy and safety of satralizumab as monotherapy in patients with Neuromyelitis Optica (NMO) and Neuromyelitis Optica Spectrum Disorder (NMOSD) study - report no. 1089825 [internal sponsor's report]. Basel (CH): F. Hoffman La-Roche Ltd; 2019 June 19.
12. Mealy MA, Kessler RA, Rimler Z, et al. Mortality in neuromyelitis optica is strongly associated with African ancestry. *Neurol Neuroimmunol Neuroinflamm,.* 2018;5(4):e468.
13. Lennon VA, Wingerchuk DM, Kryzer TJ, et al. A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. *Lancet.* 2004;364(9451):2106-2112.
14. CADTH Common Drug Review. Clinical review report eculizumab (Soliris) for NMOSD. Ottawa (ON): CADTH; 2020: https://cadth.ca/sites/default/files/cdr/clinical/sr0640-soliris-nmosd-clinical-review-report.pdf. Accessed 2021 Jan 14.
15. Wingerchuk DM, Banwell B, Bennett JL, et al. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology.* 2015;85(2):177-189.
16. Huang W, Wang L, Zhang B, Zhou L, Zhang T, Quan C. Effectiveness and tolerability of immunosuppressants and monoclonal antibodies in preventive treatment of neuromyelitis optica spectrum disorders: a systematic review and network meta-analysis. *Mult Scler Relat Disord.* 2019;35:246-252.
17. Kessler RA, Mealy MA, Levy M. Treatment of neuromyelitis optica spectrum disorder: acute, preventive, and symptomatic. *Curr Treat Options Neurol.* 2016;18(1):2.
18. Etemadifar M, Salari M, Mirmosayyeb O, et al. Efficacy and safety of rituximab in neuromyelitis optica: review of evidence. *J Res Med Sci.* 2017;22:18.
19. Nikoo Z, Badihian S, Shaygannejad V, Asgari N, Ashtari F. Comparison of the efficacy of azathioprine and rituximab in neuromyelitis optica spectrum disorder: a randomized clinical trial. *J Neurol.* 2017;264(9):2003-2009.

20. Pittock SJ, Berthele A, Fujihara K, et al. Eculizumab in aquaporin-4-positive neuromyelitis optica spectrum disorder. *N Engl J Med.* 2019;381(7):614-625.

21. CADTH Canadian Drug Expert Committee (CDEC) final recommendation: eculizumab (Soliris — Alexion Pharma Canada Corp.). Ottawa (ON): CADTH; 2020 Aug 24: https://cadth.ca/sites/default/files/cdr/complete/SR0640%20Soliris%20-%20CDEC%20Final%20Recommendation%20August%2024%2C%202020%20for%20posting.pdf. Accessed 2020 Dec 22.

22. Soliris (eculizumab for injection): 30 mL parenteral solution (10 mg/mL) [product monograph]. Zürich (CH): Alexion Pharma GmbH; 2019 Oct 29: https://pdf.hres.ca/dpd_pm/00053762.PDF. Accessed 2020 Dec 11.

23. Rituxan (rituximab): 10 mg/mL intravenous infusion [product monograph]. Mississauga (ON): Hoffmann-La Roche Ltd.; 2019 Oct 10: https://pdf.hres.ca/dpd_pm/00053554.PDF. Accessed 2020 Dec 11.

24. Imuran (azathioprine): tablets USP (50 mg), azathioprine sodium for injection, manufacturer's standard (50 mg azathioprine per vial) [product monograph]. Oakville (ON): Aspen Pharmacare Canada Inc.; 2020 Jun 5: https://pdf.hres.ca/dpd_pm/00056739.PDF. Accessed 2020 Dec 11.

25. Myfortic (mycophenolic acid enteric-coated tablet): 180 mg, 360 mg (as mycophenolate sodium) [product monograph]. Dorval (QC): Novartis Pharmaceuticals Canada Inc; 2018 May 10: https://pdf.hres.ca/dpd_pm/00045374.PDF. Accessed 2020 Dec 11.

26. CellCept (mycophenolate mofetil): capsules – 250 mg film-coated tablets – 500 mg powder for oral suspension – 200 mg/mL (when reconstituted). CellCept i.v. (mycophenolate mofetil): for injection (as hydrochloride) – 500 mg/vial [product monograph]. Mississauga (ON): Hoffmann-La Roche Ltd.; 2018 Nov 14: https://pdf.hres.ca/dpd_pm/00048798.PDF. Accessed 2020 Dec 11.

27. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol.* 2016;75:40-46.

28. Grey matters: a practical tool for searching health-related grey literature. Ottawa (ON): CADTH; 2019: https://www.cadth.ca/grey-matters. Accessed 2020 Nov 9.

29. Yamamura T, Kleiter I, Fujihara K, et al. Trial of satralizumab in neuromyelitis optica spectrum disorder. *N Engl J Med.* 2019;381(22):2114-2124.

30. Traboulsee A, Greenberg BM, Bennett JL, et al. Safety and efficacy of satralizumab monotherapy in neuromyelitis optica spectrum disorder: a randomised, double-blind, multicentre, placebo-controlled phase 3 trial. *Lancet Neurol.* 2020;19(5):402-412.

31. Drug Reimbursement Review sponsor submission: Enspryng (satralizumab), solution, 120 mg/mL, subcutaneous [internal sponsor's package]. Mississauga (ON): Hoffmann-La Roche Limited; 2020 Oct 22.

32. Center for Drug Evaluation Research. Clinical review(s). *Enspryng (satralizumab) subcutaneous. Company: Genentech, Inc. Application No. 761149. Approval date: 8/14/2020 (FDA approval package)*. Rockville (MD): U. S. Food and Drug Administration (FDA); 2020: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/761149Orig1s000MedR.pdf. Accessed 2020 Dec 8.

33. Center for Drug Evaluation Research. Statistical review(s). *Enspryng (satralizumab) subcutaneous. Company: Genentech, Inc. Application No. 761149. Approval date: 8/14/2020 (FDA approval package)*. Rockville (MD): U. S. Food and Drug Administration (FDA); 2020 Aug 15: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/761149Orig1s000StatR.pdf. Accessed 2020 Nov 9.

34. Health Canada reviewer's report: Enspryng (satralizumab) [internal sponsor's report]. Ottawa (ON): Therapeutics Products Directorate, Health Canada; 2020.

35. Frendl DM, Ware JE, Jr. Patient-reported functional health and well-being outcomes with drug therapy: a systematic review of randomized trials using the SF-36 health survey. *Med Care.* 2014;52(5):439-445.

36. User's manual for the SF-36v2 health survey. In: Maruish ME, ed. Lincoln (RI): QualityMetric Incorporated; 2011.

37. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(3):199-208.

38. Sinnott PL, Joyce VR, Barnett PG. Preference measurement in economic analysis. Guidebook. Menlo Park (CA): Health Economics Research Center; 2007.

39. Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke.* 2007;38(3):1091-1096.

40. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JP. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Stat Med.* 2015;34(6):984-998.

41. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *Int J Epidemiol.* 2012;41(3):818-827.

42. Cree BAC, Bennett JL, Kim HJ, et al. Inebilizumab for the treatment of neuromyelitis optica spectrum disorder (N-MOmentum): a double-blind, randomised placebo-controlled phase 2/3 trial. *Lancet.* 2019;394(10206):1352-1363.

43. Hoffmann-La Roche Limited response to 2020 Nov 25 CDR request for additional information regarding the Enspryng (satralizumab) DRR review [internal additional sponsor's information]. Mississauga (ON): Hoffmann-La Roche Limited; 2020 Nov 27.

44. Kos D, Raeymaekers J, Van Remoortel A, et al. Electronic visual analogue scales for pain, fatigue, anxiety and quality of life in people with multiple sclerosis using smartphone and tablet: a reliability and feasibility study. *Clin Rehabil.* 2017;31(9):1215-1225.

45. Riemsma RP, Forbes CA, Glanville JM, Eastwood AJ, Kleijnen J. General health status measures for people with cognitive impairment: learning disability and acquired brain injury. *Health Technol Assess.* 2001;5(6):1-100.

46. Pfennings LE, van der Ploeg HM, Cohen L, Polman CH. A comparison of responsiveness indices in multiple sclerosis patients. *Qual Life Res.* 1999;8(6):481-489.

47. Robinson D, Jr., Zhao N, Gathany T, Kim LL, Cella D, Revicki D. Health perceptions and clinical characteristics of relapsing-remitting multiple sclerosis patients: baseline data from an international clinical trial. *Curr Med Res Opin.* 2009;25(5):1121-1130.

48. Kuspinar A, Mayo NE. A review of the psychometric properties of generic utility measures in multiple sclerosis. *Pharmacoeconomics.* 2014;32(8):759-773.

49. Meyer-Moock S, Feng YS, Maeurer M, Dippel FW, Kohlmann T. Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. *BMC Neurol.* 2014;14:58.

50. de Groot V, Beckerman H, Uitdehaag BM, et al. The usefulness of evaluative outcome measures in patients with multiple sclerosis. *Brain.* 2006;129(Pt 10):2648-2659.

51. Goldman MD, LaRocca NG, Rudick RA, et al. Evaluation of multiple sclerosis disability outcome measures using pooled clinical trial data. *Neurology.* 2019;93(21):e1921-e1931.

52. Rooney S, McFadyen DA, Wood DL, Moffat DF, Paul PL. Minimally important difference of the fatigue severity scale and modified fatigue impact scale in people with multiple sclerosis. *Mult Scler Relat Disord.* 2019;35:158-163.