THE LANCET
Neurology

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Zhang C, Zhang M, Qiu W, et al. Safety and efficacy of tocilizumab versus azathioprine in highly relapsing neuromyelitis optica spectrum disorder (TANGO): an open-label, multicentre, randomised, phase 2 trial. *Lancet Neurol* 2020; **19:** 391–401.

Tocilizumab versus azathioprine in highly relapsing neuromyelitis optica spectrum disorder (TANGO): a multicenter, randomised, head-to-head comparison trial

Supplementary Appendix: Table of contents

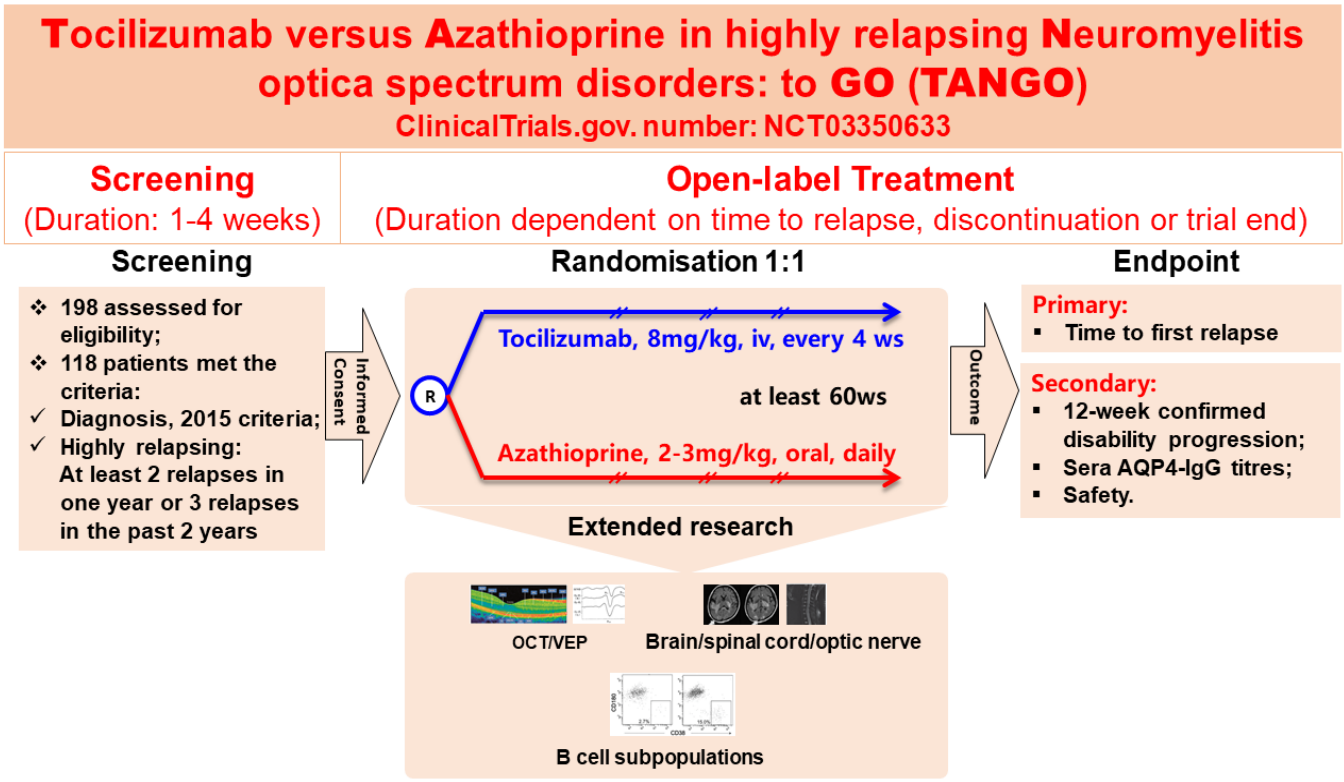
Trial Overview

Trial Registration

The first patient was enrolled on November 1, 2017. We first submitted our registration on November 10, 2017. After several rounds of additional information, the registration was approved on November 22, 2017. Prior to this date, we had 20 patients enrolled. At that time, a large number of patients decided to participate this trial for the grave concern about progression of their disease.

TANGO study design

Additional Figure 1: TANGO Study design flow diagram

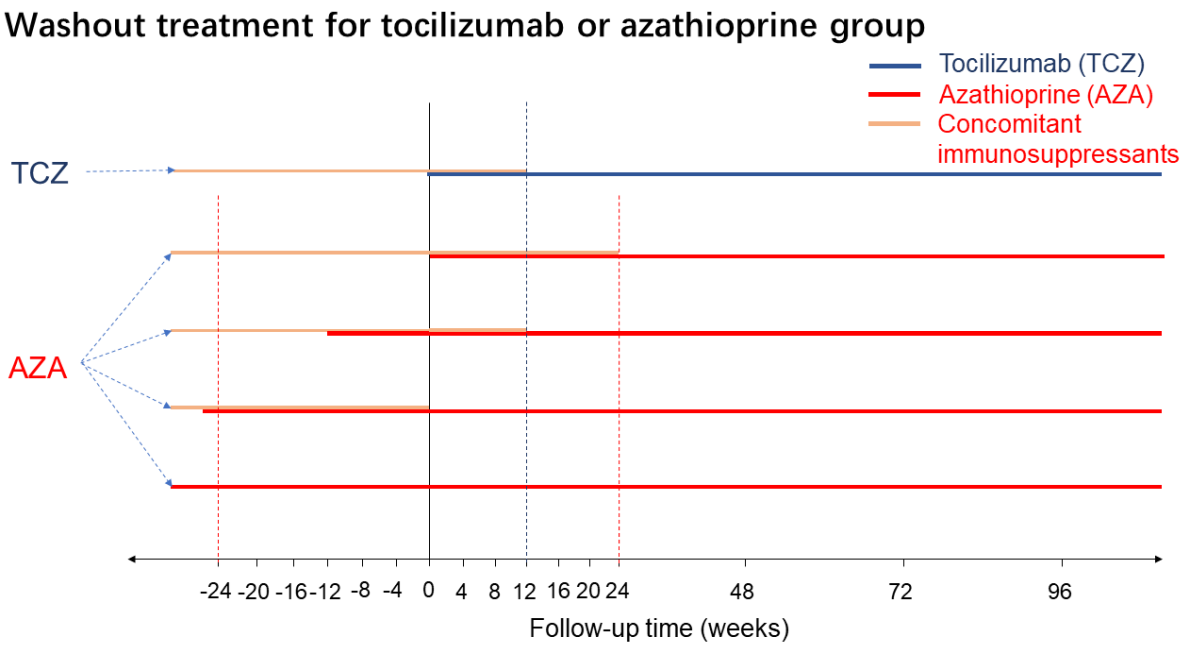


The TANGO trial was an open-label, randomised controlled trial, with a mean follow-up period of at least 60 weeks. The trial enrolled patients with highly active NMOSD. Patients meeting eligibility criteria were randomly assigned in a 1:1 ratio to receive either tocilizumab or azathioprine. The primary outcome was time to first relapse. The secondary outcomes included 12-week confirmed disability progression and changes of serum AQP4-IgG titres measured by live cell-based assay. Safety was assessed in the intention-to-treat population. Several exploratory efficacy measures included: 1) time to onset of confirmed disability progression (CDP) at 24 weeks; 2) change of high-contrast visual acuity (VA) from baseline; 3) change of low-contrast letter acuity (LCLA) from baseline; 4) change of average retinal nerve fiber layer (RNFL) thickness from baseline measured by spectral-domain optical coherence tomography (SD-OCT); 5) change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT; 6) change of P100 latency and amplitude from baseline measured in visual evoked potentials (VEP); 7) number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI); 8) change of peripheral blood B cell subsets counts measured by flow cytometry.

Patients experienced a washout period of baseline immunosuppressants before tocilizumab or azathioprine was used as monotherapy (Figure Additional 2). In the azathioprine group, patients without prior azathioprine treatment took concomitant immunosuppressants (oral corticosteroids, mycophenolate mofetil, cyclophosphamide, or methotrexate) for the initial 24 weeks. Thereafter, azathioprine was used as monotherapy. Patients receiving azathioprine for < 24 weeks prior to randomisation received concomitant immunosuppressants until 24 weeks, then azathioprine was used as monotherapy. Patients receiving azathioprine for ≥ 24 weeks before randomisation used azathioprine as monotherapy directly, without concomitant immunosuppressants. Patients in the tocilizumab group received concomitant immunosuppressants for the first 12 weeks, and thereafter tocilizumab was used as

monotherapy. Patients could complete the trial owing to a relapse. No new immunosuppressants were permitted during the monotherapy period unless a relapse had occurred or there was a compelling medical need for treatment adjustment.

Additional Figure 2



TANGO study Expert Panel (designated as the central review committee) for Relapse Assessment

Fu-Dong Shi (Chair), Department of Neurology, Tianjin Neurological Institute, Tianjin Medical University General Hospital, Tianjin Medical University, Tianjin, China; National Clinical Research Center for Neurological Diseases of China, Tiantan Hospital, Capital Medical University, Beijing, China;

Jeffrey L. Bennett, Departments of Neurology and Ophthalmology, Programs in Neuroscience and Immunology, University of Colorado Denver School of Medicine, Aurora, CO, USA.

Xinghu Zhang, National Clinical Research Center for Neurological Diseases of China, Tiantan Hospital, Capital Medical University, Beijing, China;

Meini Zhang, Department of Neurology, First Hospital of Shanxi Medical University, Taiyuan, China;

Wei Qiu, Department of Neurology, Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China;

Chuishui Yu, School of Radiology, Tianjin Medical University, Tianjin, China;

Yangtai Guan, Department of Neurology, RenJi Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China;

Huiqing Dong, Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China;

Lihua Wang, Department of Neurology, The Second Affiliated Hospital of Harbin Medical University, Harbin, China.

TANGO study Statistician Panel

Wenli Lu (Chair); Yongjie Chen; Lihui Zhou,From School of Public Health, Tianjin Medical University, Tianjin, China.

List of collaborators of TANGO study:

Beijing, Municipality

Beijing Tiantan Hospital, Capital Medical University - principal investigator: Xinghu Zhang; investigators: De-Cai Tian, Tian Song, Xindi Li, Xiran Liu.

Datong, ShanXi Province

The Third People's Hospital of Datong - principal investigator: Hongshan Ma; investigators: Xuejiao Wang, Zhigang Cui, Huiming Zhang, Xiaomin Liu, Yanjun Liu.

Guangzhou, GuangDong Province

The Third Affiliated Hospital of Sun Yat-sen University - principal investigator: Wei Qiu; investigators: Zhengqi Lu, Jing Li, Ting Shen, Yuge Wang, Rui Li.

Taiyuan, ShanXi Province

First Hospital of Shanxi Medical University - principal investigator: Meini Zhang; investigators: Wen Yang, Juan Dong, Ruxin Xue.

Tianjin, Municipality

Tianjin Medical University General Hospital - principal investigator: Fu-Dong Shi; investigators: Chao Zhang, Dongmei Jia, Chun-Sheng Yang, Tian-Xiang Zhang, Meng Yuan, Yan Feng, Li Yang.

Tianjin, Municipality

Tianjin Huanhu Hospital - principal investigator: Zilong Zhu; investigators: Lei Chen, Jie Qin, Yuwang Li, Lu Han.

We appreciate all the faculties blinded to the randomisation and clinical treatments, including EDSS raters, OCT/VEP technicians, laboratory technicians, MRI technicians and raters.

Additional methodology

Inclusion criteria

1. Male or female patients ≥ 18 years old.
2. Diagnosis of NMOSD as defined by 2015 Criteria.
3. Historical Relapse (as defined by this protocol) of at least 2 relapses in the last 12 months or 3 relapses in the last 24 months, with at least 1 relapse in the 12 months prior to the screening.
4. Able and willing to give written informed consent and comply with the requirements of the study protocol.
5. EDSS ≤ 7·5.
6. Men and women of reproductive potential must agree to use a highly effective method of birth control from screening to 6 months after final dose of the experimental drugs.

Exclusion Criteria

1. Current evidence or known history of clinically significant infection (Herpes simplex virus, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human immunodeficiency virus, Hepatitis viruses, Syphilis, etc).
2. Pregnant, breastfeeding, or potential/planned child-bearing during the course of the study.
3. Patients will not participate in any other clinical therapeutic study or will not have participated in any other experimental treatment study within 30 days of screening.
4. Participation in another interventional trial within the last 3 months.
5. Heart or kidney insufficiency.
6. Tumor disease currently or within last 5 years.
7. Clinically relevant liver, kidney or bone marrow function disorder.
8. Receipt of rituximab or any experimental B-cell depleting agent within 6 months prior screening and B-cells below the lower limit of normal measured by flow cytometry.
9. Prior relapses or toxic intolerance on azathioprine treatment (adequate dosage and follow-up period observation).
10. The patient has a heterozygous or homozygous mutation in the thiopurine methyltransferase (TPMT) gene.

Definition of a relapse of neuromyelitis optica spectrum disorder

The protocol defined requisites to diagnose a new relapse. In practice, they are related in parallel. The following criteria indicate detailed conditions for relapses in the trial.

1. Acute optic neuritis (ON):

(1) Acute vision loss with or without eye pain and optic papillae edema

(2) Abrupt abnormal visual field associated with optic nerve damage

(3) At least 5 characters drop in 100% high-contrast visual acuity (monocular)

(4) New relative afferent pupillary defect (RADP) or loss of a previously documented RAPD in fellow eye

(5) Confirmatory MRI finding in the corresponding optic nerve

The relapse of ON will meet one of the following combinations:

(1)+(3)+(4); (2)+(3)+(4); (1)+(3)+(5); (2)+(3)+(5); (3)+(5).

2. Acute myelitis:

The relapse of myelitis will meet one of the following criteria:

(1) At least 1-point worsening in EDSS if baseline EDSS >5·5

(2) At least 1·5-point worsening in EDSS if baseline EDSS ≤ 5·5

(3) At least 2-point worsening in at least a relevant functional system score (pyramidal, sensory, bladder/bowel)

(4) 0·5-point worsening in EDSS if baseline EDSS >5·5 or 1-point worsening in EDSS if baseline EDSS ≤ 5·5 or at least 1-point worsening in at least 2 relevant functional system scores and confirmatory MRI finding in spinal cord

3. Area postrema syndrome: Unexplained hiccups or nausea and vomiting and confirmatory MRI finding in the area postrema

4. Acute brainstem syndrome: Cinical presentations caused by brainstem lesions, i.e., syndrome of medial longitudinal fasciculus, limbic weakness, bulbar palsy and confirmatory MRI finding in brainstem

5. Acute diencephalic clinical syndrome or symptomatic narcolepsy: Symptomatic narcolepsy, excessive sleepiness or other diencephalic symptoms and confirmatory MRI finding in diencephalon

6. Symptomatic cerebral syndrome: Clinical presentations caused by cortical or subcortical lesions, i.e., hemiparalysis, aphasia and confirmatory MRI finding in the brain.

MRI criteria was necessary if a relapse was suspected but with clinical change not meeting clinical criteria. No clinical change in EDSS score, with or without new or enlarging T2 MRI lesions or T1 Gadolium-enhancing lesions would not count as a relapse. All the confirmatory MRI findings must be consistent with related clinical changes (Table Additional 1).

Table Additional 1: MRI criteria for confirmation of a relapse

Confirmatory MRI findings
Optic nerve MRI
• Unilateral or bilateral new or enlarged lesions on horizontal T2-weighted within optic nerve or optic chiasm; relatively long lesions (e.g., those extending more than half the distance from orbit to chiasm) and those involving the posterior aspects of the optic nerves or the chiasm are associated with NMOSD • T1 gadolinium enhancement within optic nerve or optic chiasm
Spinal cord MRI
• New lesions on sagittal and horizontal T2-weighted (standard T2-weighted, proton density, or STIR sequences) • Enlarged lesions on sagittal and horizontal T2-weighted (standard T2-weighted, proton density, or STIR sequences) • Gadolinium enhancement of the lesion on T1-weighted sequences (no specific distribution or pattern of enhancement is required)
Cerebral MRI
• New or enlarged lesions in brain on T2-weighted sequences • T1 gadolinium enhancement in brain **Typical brain lesion patterns met at least one of the following areas:** • Lesions involving the dorsal medulla (especially the area postrema), either small and localized, often bilateral, or contiguous with an upper cervical spinal cord lesion • Periependymal surfaces of the fourth ventricle in the brainstem/cerebellum • Lesions involving the hypothalamus, thalamus, or periependymal surfaces of the third ventricle • Large, confluent, unilateral, or bilateral subcortical or deep white matter lesions • Long (1/2 of the length of the corpus callosum or greater), diffuse, heterogeneous, or edematous corpus callosum lesions • Long corticospinal tract lesions, unilateral or bilateral, contiguously involving internal capsule and cerebral peduncle • Extensive periependymal brain lesions, often with gadolinium enhancement

Sera AQP4-IgG titres detection by live cell-based assay: detailed protocol

Sera AQP4-IgG titres were determined via cell-based assay (CBA) using human embryonic kidney (HEK293T) cells transiently transfected with plasmid (pcDNA3.1) containing full-length human AQP4 cDNA. AQP4-transfected and non-transfected HEK293T cells were grown on cover slips to 70% confluency and then fixed. The cover slips were cut into 5-millimeter square fragments and stored at -80℃. Prior to use, coverslips were permeabilized with phosphate-buffered saline (PBS)-0.2%Tween 20 and blocked with 3% bovine serum albumin. Sera were diluted at 1:10 in PBS-0.2%Tween 20 and incubated on coverslips for 1 hour at room temperature. Coverslips were then washed in PBS, incubated for 30 min with goat anti-human IgG (1:500, Invitrogen), washed again in PBS, and evaluated by immunofluorescence microscopy. Two independent masked assessors classified each sample as positive or negative based on the intensity of surface immunofluorescence in direct comparison with non-transfected cells and control samples. Once confirmed, the $AQP4^+$ sera were then serially diluted by twofold from 1:60 to 1:3840 to determine the titers. The final titer was defined as the sample dilution value for which specific fluorescence was barely but clearly identifiable and expressed as the corresponding dilution value. **Figure Additional 3** shows the representative fluorescence images stained by serum from an $AQP4^+$ subject at different dilution folds.

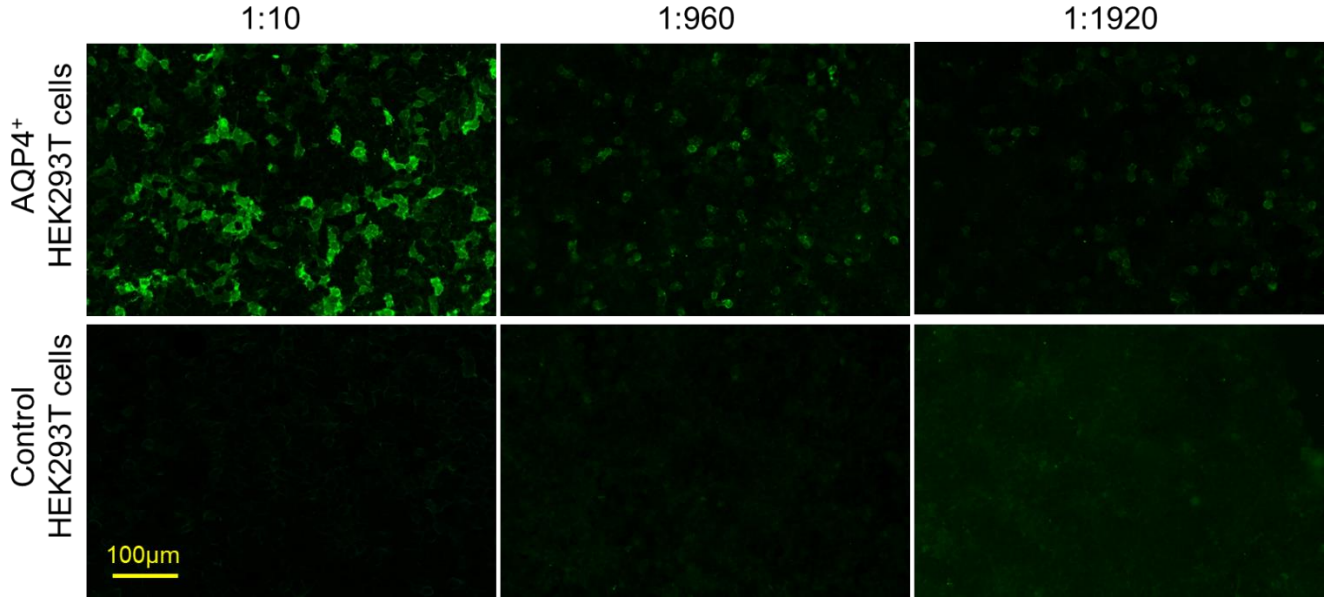


Figure Additional 3: Representative AQP4-IgG detection by live CBA.

HEK293T cells transfected with human AQP4 cDNA (upper row). Non-transfected HEK293T cells present on the same slide were used as controls (lower row). Positive samples were diluted in different dilutions to ascertain the maximum dilution titer. Scale bar: 100 μm.

Study data

Table S1. Primary statistical analyses and adjusted analyses by center effect for the primary and secondary endpoints.

End point	Tocilizumab (N=59)	Azathioprine (N=59)	Hazard Ratio or Difference (95% CI) *	p Value	Adjusted Hazard Ratio or Difference (95% CI) †	Adjusted p Value
Primary outcome						
First relapse, n (%)	8 (14%)	28 (47%)	0·236 (0·107 to 0·518)	<0·0001	0.236 (0·107 to 0·520)	0·0003
Secondary outcomes						
12-week confirmed disease progression	5 (8%)	15 (25%)	0·288 (0·105 to 0·795)	0·0087	0·288 (0·104 to 0·793)	0.0162
Serum AQP4-IgG titres‡						
Change from baseline	-240 (-720 to -240)	0 (-240 to 0)	-240 (-480 to -240)	0·0004	-240 (-480 to -240)	0·0004
Percentage change	-50% (-75 to -25)	0 (-33 to 0)	-33% (-50 to -17)	<0·0001	-33% (-50 to -17)	<0·0001

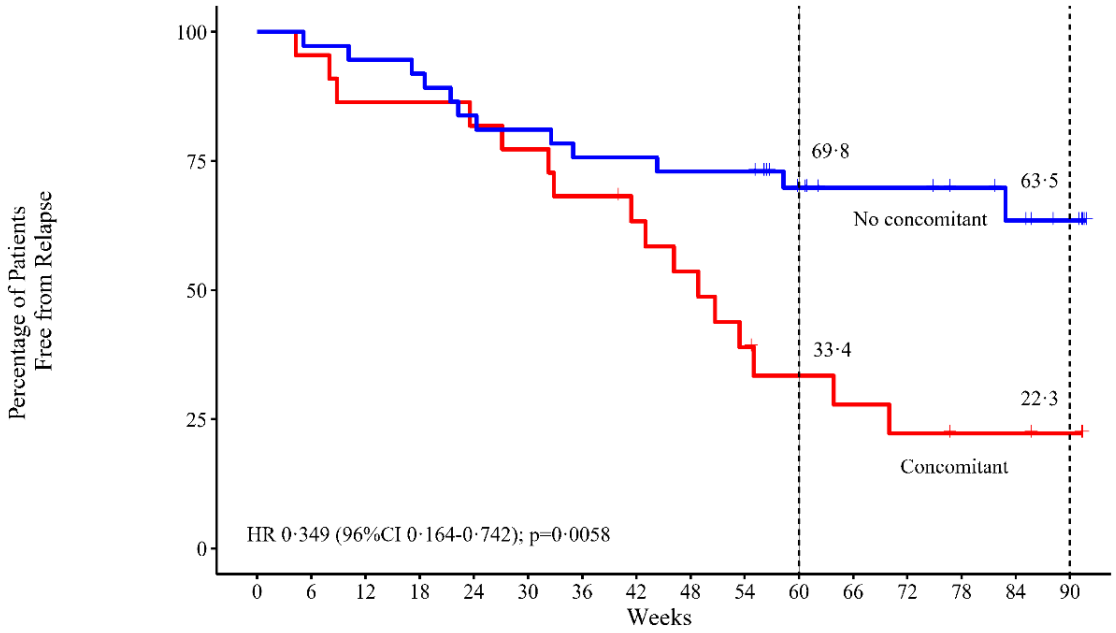
* The difference is for the tocilizumab group as compared with the azathioprine group. CI denotes confidence interval.

† The adjusted analyses took center as a random effect factor into account in Cox model.

‡ The cutoff of serum AQP4-IgG titres by live cell-based assay was 1:10. Values are indicated as median and interquartile range (IQR, 25th and 75th percentile).

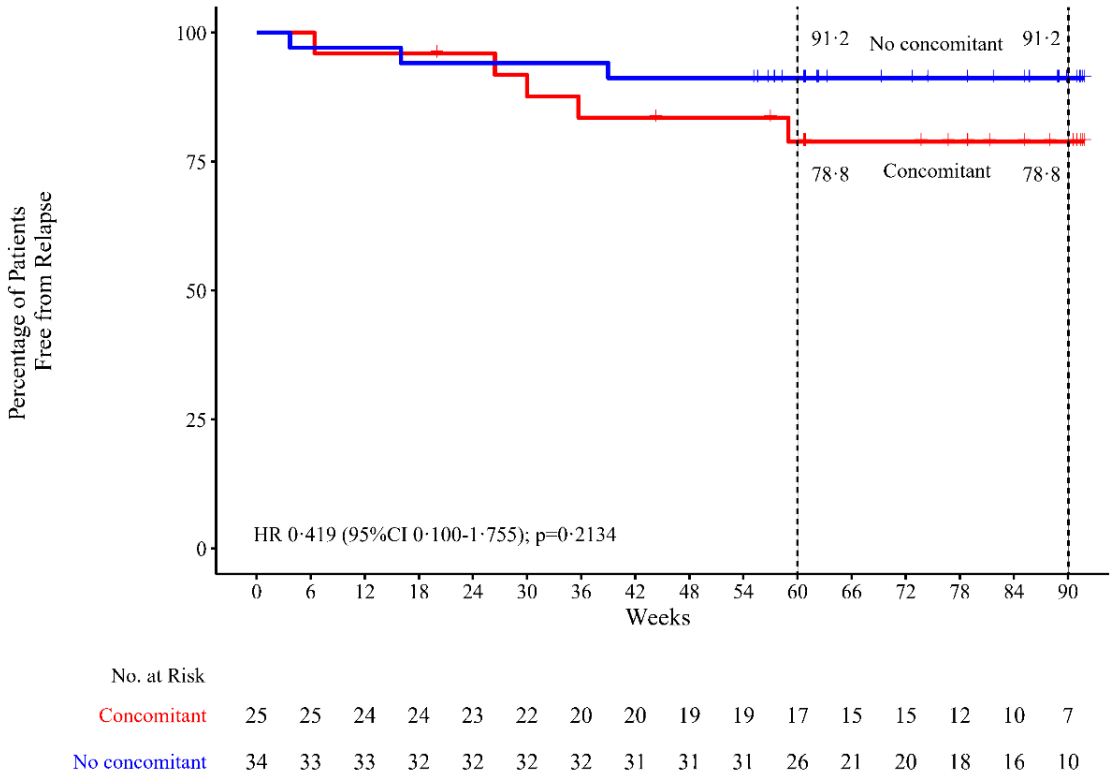
Figure S1: Time to first relapse (primary outcome) in the ITT analysis.

A



No. at Risk																
Concomitant	22	21	19	19	18	17	15	13	11	8	6	5	4	3	3	2
No concomitant	37	36	35	34	31	30	28	28	27	27	20	16	16	13	10	7

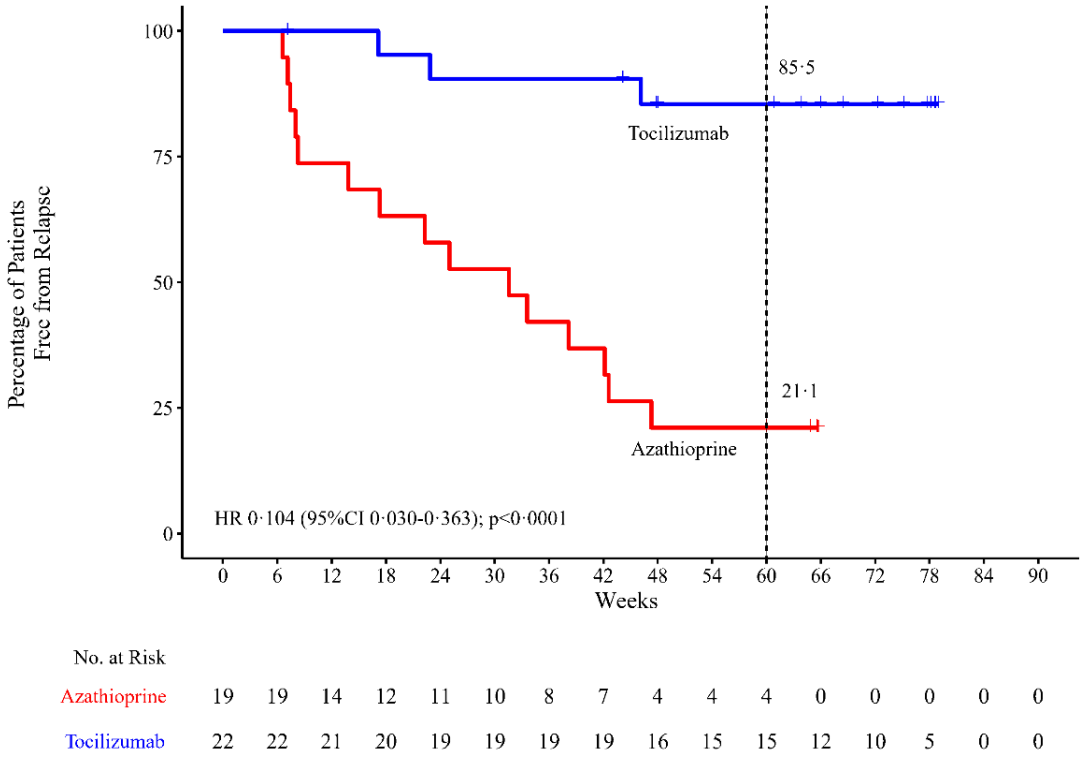
B



Kaplan-Meier plots of relapse-free survival among patients with NMOSD receiving azathioprine (A) or tocilizumab (B). The pre-specified patients were divided into two subgroups: NMOSD with concomitant autoimmune diseases (red line) and NMOSD without concomitant autoimmune diseases (blue line). The tick marks indicate censoring data. Data for patients who did not have a relapse were censored at the end of the trial period. ITT, intent-to-treat.

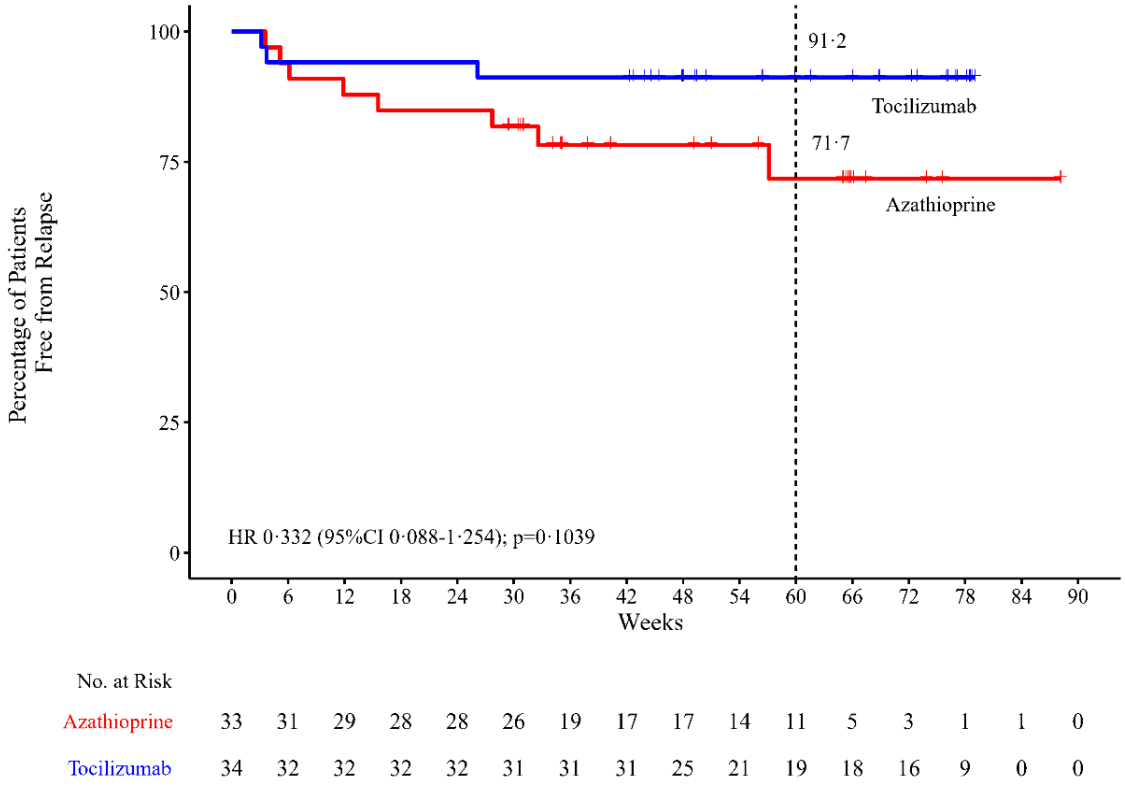
Figure S2: Per-protocol analysis of time to first relapse of NMOSD with/without concomitant autoimmune diseases.

A



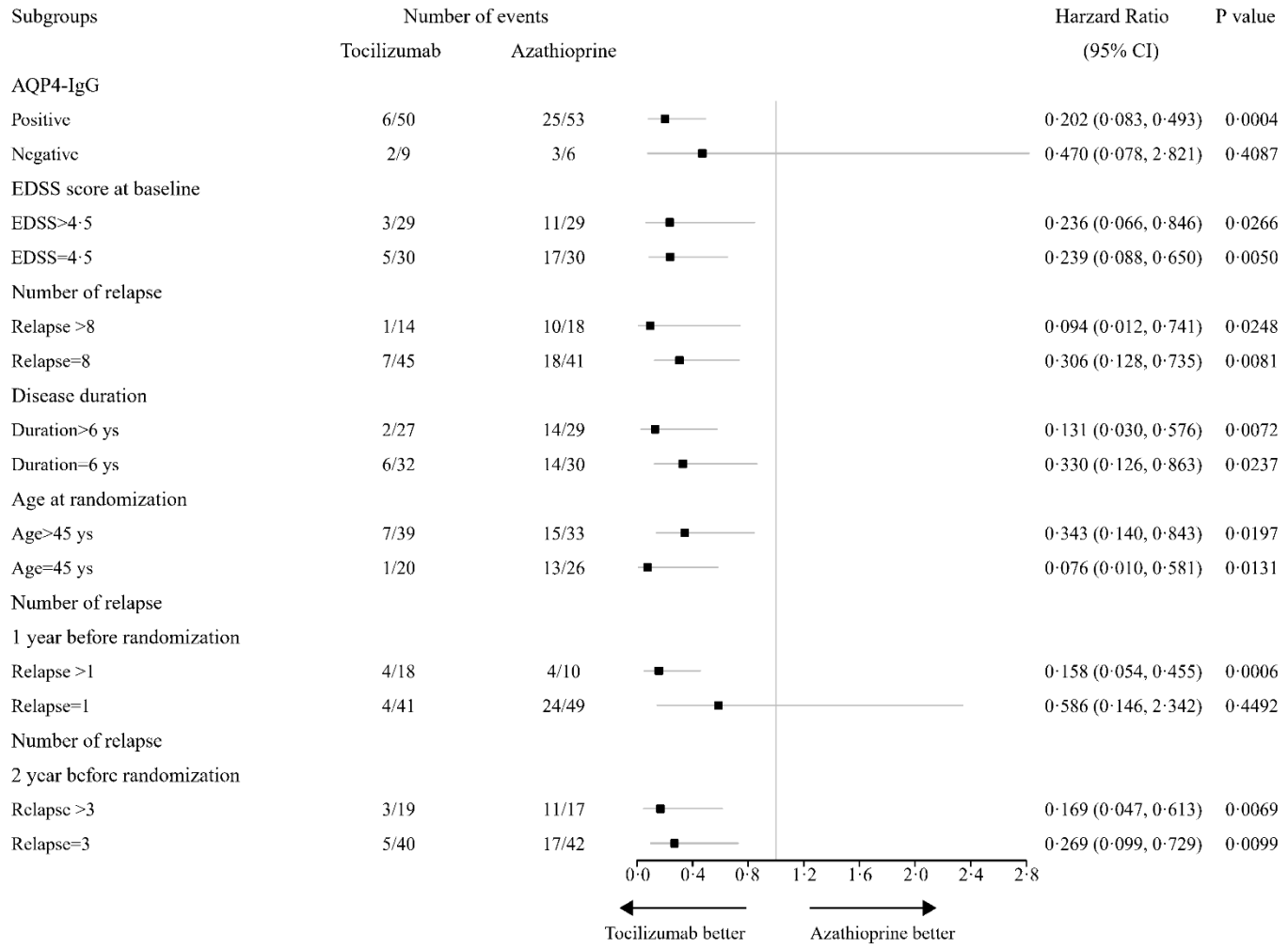
No. at Risk																
Azathioprine	19	19	14	12	11	10	8	7	4	4	4	0	0	0	0	0
Tocilizumab	22	22	21	20	19	19	19	19	16	15	15	12	10	5	0	0

B



Kaplan-Meier plots of relapse-free survival among NMOSD patients with (A) and without (B) concomitant autoimmune diseases in the per-protocol analysis for the primary outcome. After a protocol-defined washout period, the patients received tocilizumab or azathioprine treatment as monotherapy. The tick marks indicate censored data. Data for patients who did not have a relapse were censored at the end of the trial period.

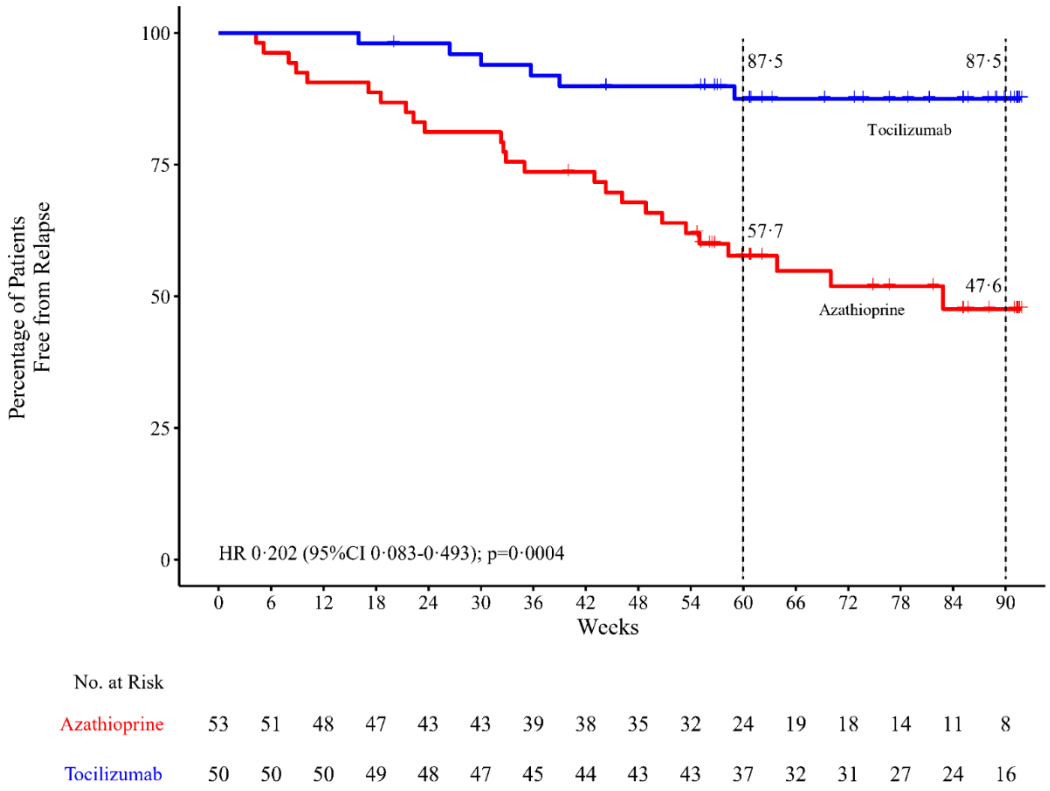
Figure S3: Subgroup analyses of the hazard ratio for relapse risk.



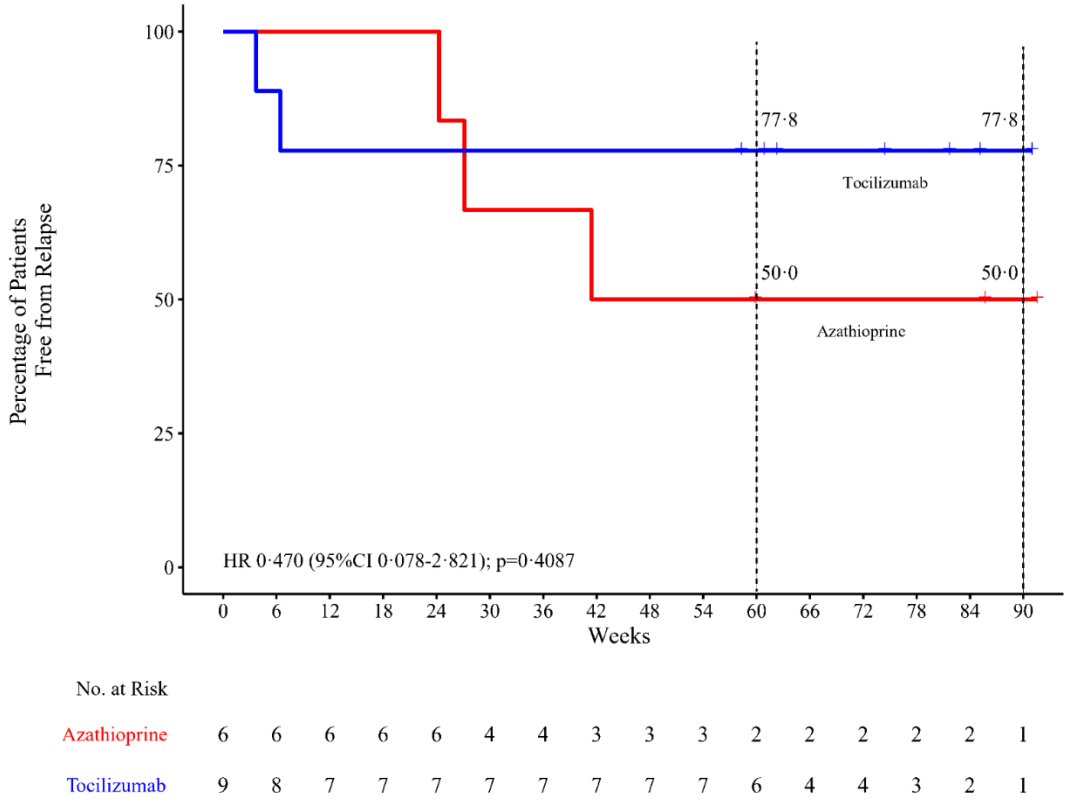
Forest plots of comparative treatment efficacy among patient subgroups. The subgroups were stratified by 6 factors: baseline EDSS (≤ 4·5 and > 4·5), total number of historical relapses (≤ 8 and > 8), disease duration (≤ 6 ys and > 6 ys), age at randomisation (≤ 45 and > 45), and number of relapses at 1 or 2 years before randomisation. CI, confidence interval. Hazard Ratio (HR) and 95% CI for reduced disease activity by subset between the tocilizumab group and the azathioprine group.

Figure S4: Time to first relapse by AQP4-IgG status in the ITT analysis.

A



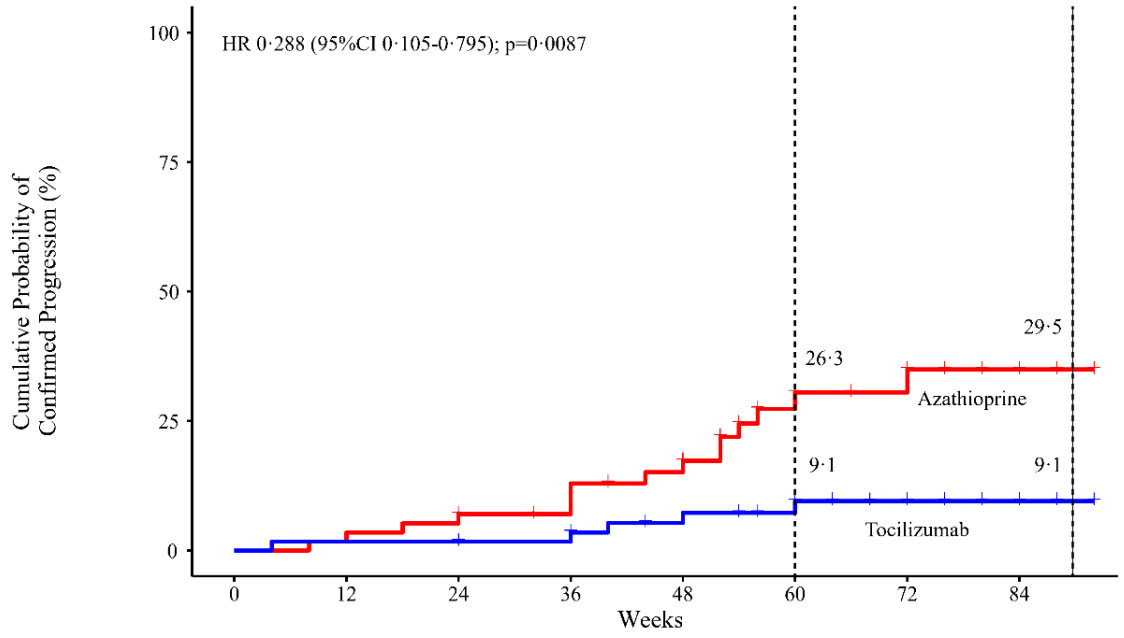
B



No. at Risk															
Azathioprine	6	6	6	6	6	4	4	3	3	3	2	2	2	2	2	1
Tocilizumab	9	8	7	7	7	7	7	7	7	7	6	4	4	3	2	1

Kaplan-Meier plots of relapse-free survival among NMOSD patients with AQP4-IgG positive (A) and AQP4-IgG negative (B) in the ITT analysis. The tick marks indicate censoring data. Data for patients who did not have a relapse were censored at the end of the trial period. ITT, intent-to-treat.

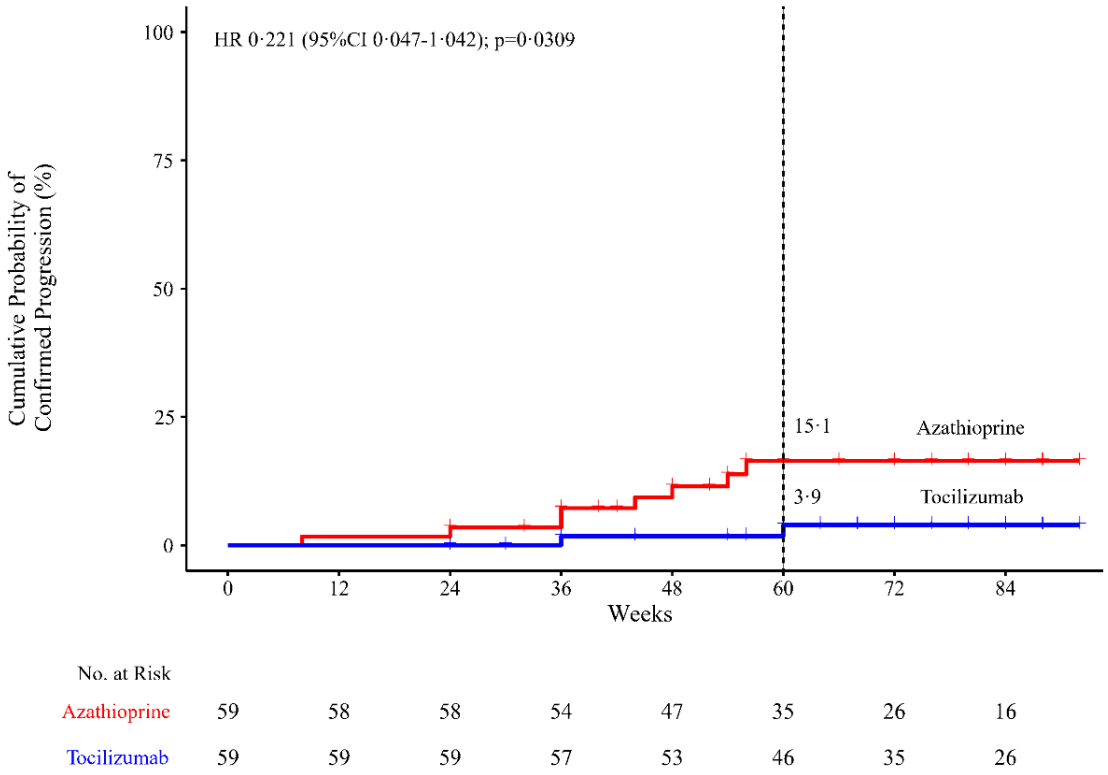
Figure S5:12-week confirmed disability progression between the tocilizumab group and the azathioprine group.



No. at Risk								
Azathioprine	59	58	56	52	46	32	23	15
Tocilizumab	59	58	58	57	52	44	35	26

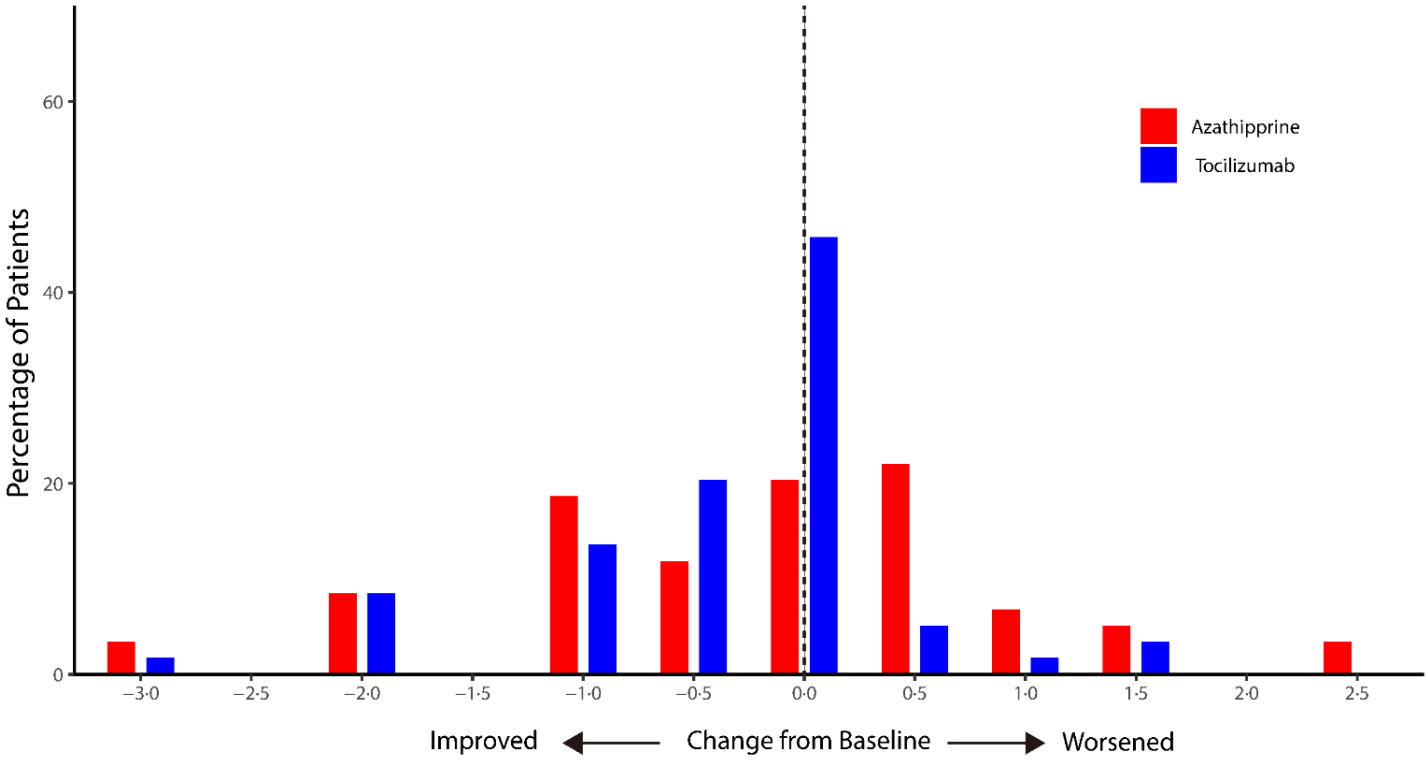
12-week confirmed disability progression outcome in the ITT Population. Shown is the cumulative probability of clinical disability progression (as defined by an increase in the score on the Expanded Disability Status Scale) confirmed after 12 weeks in a time-to-event analysis. P value was calculated with the use of the log-rank test.

Figure S6:24-week confirmed disability progression between the tocilizumab group and the azathioprine group.



24-week confirmed disability progression outcome in the ITT Population. Shown is the cumulative probability of clinical disability progression (as defined by an increase in the score on the Expanded Disability Status Scale) that was confirmed after at least 24 weeks in time-to-event analysis. P value was calculated with the use of the log-rank test.

Figure S7: Changes of EDSS scores at the end of the TANGO study from the baseline.



Distributions of changes in the scores on the Expanded Disability Status Scale (EDSS) from baseline to the end of the trial (ITT Population). The mean change of EDSS score was -0·32 ± 0·72 in the tocilizumab group versus -0·13 ± 1·05 in the azathioprine group. A difference of -0·20 (95%CI, -0·52 to -0·13) favored the tocilizumab group; however, the magnitude was not statistically significant (p=0·242). Fewer participants had EDSS score worsening from baseline with tocilizumab than with azathioprine (relative risk 3·667 [95% CI 1·603-8·387]; p=0·0005).

Table S1: Summary of the relapses.

Sex	Age (yrs)*	Historical Diseases*	AQP4-IgG	Immunosuppressants*	Baseline EDSS Score*	Day of Onset*	Relapse Type	Trial Period	Therapeutics after the trial
Tocilizumab group									
Female	60	NMOSD, SS, GD	+	PSL+MMF	5·5	413	Acute Myelitis	Monotherapy	RTX§
Female	52	NMOSD, AD, DVT	-‡	PSL+AZA‡	5·5	45	Acute Right Optic Neuritis+ Acute Myelitis	Therapy in combination	PSL+RTX
Male	54	NMOSD, ONFH†	-	None	4·5	26	Acute Myelitis	Monotherapy	MMF
Female	49	NMOSD	+	PSL+CTX	4	112	Acute Myelitis (MRI confirmed)	Monotherapy	PSL+MMF
Female	47	NMOSD, CI, HBP, AD	+	MMF	4	210	Acute Myelitis (MRI confirmed)	Monotherapy	PSL+RTX
Female	59	NMOSD, gMG	+	PSL+CTX	4	250	Acute Myelitis	Monotherapy	PSL+MMF
Female	49	NMOSD	+	PSL+AZA	4	273	Acute Myelitis	Monotherapy	PSL+RTX
Female	19	NMOSD, SS	+	PSL+AZA	6	179	Area Postrema Syndrome+ Acute Myelitis	Monotherapy	Death
Azathioprine group									
Female	48	NMOSD, SS, HBP	-	PSL+AZA	3·5	190	Acute Myelitis (MRI confirmed)	Monotherapy	PSL+AZA
Female	54	NMOSD, HBP, DVT	+	MMF	4	71	Acute Myelitis	Therapy in combination	PSL+RTX
Female	65	NMOSD	+	PSL+AZA	5·5	156	Acute Myelitis	Monotherapy	Death
Female	30	NMOSD	+	PSL+AZA	3·5	245	Acute Left Optic Neuritis	Monotherapy	RTX
Female	52	NMOSD, SS	+	MPSL+TAC	5	355	Acute Myelitis (MRI confirmed)	Monotherapy	RTX
Female	27	NMMOSD	-	AZA	4	170	Area Postrema Syndrome+ Acute Myelitis (MRI confirmed)	Monotherapy	TCZ
Female	51	NMOSD, SS	+	PSL+AZA	6	342	Acute Right Optic Neuritis+ Acute Myelitis	Monotherapy	RTX
Female	30	NMOSD, GD	+	PSL+AZA	4·5	490	Acute Myelitis	Monotherapy	PSL+AZA
Female	61	NMOSD, SS	+	PSL+AZA	6	374	Acute Myelitis	Monotherapy	PSL+MMF
Female	53	NMOSD, SS, HBP, Severe Osteoporosis	+	AZA	6	385	Acute Myelitis	Monotherapy	RTX
Female	54	NMOSD, MCTD	+	MMF	3	447	Area Postrema Syndrome (MRI confirmed)	Monotherapy	PSL+MMF
Female	19	NMOSD	+	PSL+AZA	4·5	310	Acute Left Optic Neuritis	Monotherapy	TCZ
Female	43	NMOSD, SS, ONFH	+	PSL+AZA	7·5	323	Acute Myelitis+ Symptomatic cerebral syndrome (MRI confirmed)	Monotherapy	RTX
Female	36	NMOSD, AD	-	PSL+AZA	4·5	290	Acute Myelitis (MRI confirmed)	Monotherapy	TCZ
Male	21	NMOSD	+	MMF	3·5	580	Acute Myelitis (MRI confirmed)	Monotherapy	PSL+AZA

Female	61	NMOSD, DM, CI	+	PSL+AZA	5·5	228	Acute Myelitis	Monotherapy	TCZ
Female	61	NMOSD, HBP, Gastric Ulcer	+	PSL+MMF	4·5	408	Acute Myelitis	Monotherapy	PSL
Female	64	NMOSD, MCTD	+	PSL+AZA	5	62	Acute Myelitis+ Bilateral Optic Neuritis+ Acute Diencephalic Clinical Syndrome	Monotherapy	TCZ
Female	40	NMOSD, gMG	+	PSL+CTX	4	301	Acute Left Optic Neuritis+ Acute Myelitis	Monotherapy	PSL+RTX
Male	65	NMOSD, HBP, CI	+	PSL+AZA	5·5	36	Acute Brainstem Syndrome (MRI confirmed)	Monotherapy	MPSL+MMF
Female	54	NMOSD	+	PSL+MMF	5·5	150	Acute Myelitis	Therapy in combination	BTZ + Melphalan + PSL[¶]
Female	63	NMOSD, SLE, HBP, Severe Osteoporosis	+	MMF	4·5	226	Acute Myelitis	Monotherapy	RTX
Female	21	NMOSD, HT	+	PSL+AZA	4·5	165	Acute Left Optic Neuritis (MRI confirmed)	Monotherapy	MPSL+MMF
Female	21	NMOSD, GD	+	AZA	3	56	Acute Right Optic Neuritis	Monotherapy	TCZ
Female	51	NMOSD, RA, HBP, DM	+	MTX	4·5	230	Acute Myelitis	Monotherapy	PSL+AZA
Female	18	NMOSD	+	PSL+AZA	3·5	130	Acute Left Optic Neuritis+ Acute Brainstem Syndrome	Monotherapy	TCZ
Female	38	NMOSD	+	PSL+AZA	5	120	Acute Myelitis	Therapy in combination	TCZ
Female	23	NMOSD, SLE	+	PSL+AZA	3	30	Acute Myelitis	Therapy in combination	MPSL+RTX

[*] Baseline characteristics.

[‡]Amongst the patients with AQP4-IgG negative, 2 of 9 patients in the tocilizumab group and 3 of 6 patients in the azathioprine group had a new relapse during the trial (Chi-square test, P=0·329). No difference in relapse activity was found between the two groups in the small number of patients who were AQP4-IgG seronegative.

[†] The abbreviations of the diseases are as follows: AD, Atopic dermatitis; AF, Atrial fibrillation; AH, Autoimmune hemolytic anemia; CHD, Coronary heart disease; CI, Cerebral infarction; DM, Diabetes mellitus; DVT, Deep venous thrombosis; HBP, Hypertension; HT, Hashimoto's thyroiditis; GD, Grave's disease; MCTD, Mixed connective tissue disease; gMG, generalized Myasthenia gravis; ONFH, Osteonecrosis of the Femoral Head; NMOSD, RA, Rheumatic arthritis; SLE, systemic lupus erythematosus; SS, Sjögren's syndrome.

[‡] The abbreviations of the immunosuppressants are as follows: AZA, Azathioprine; BTZ, Bortezomib; CTX, Cyclophosphamide; MMF, Mycophenolate mofetil; MPSL, Methylprednisolone; MTX, Methotrexate; PSL, Prednisolone; RTX, Rituximab; TAC, tacrolimus; TCZ, Tocilizumab.

[§] RTX was used with lower-dose regimen. Rituximab was intravenously administered at the dose of 100 mg/day for 3 or 5 consecutive days, until the frequency of CD19 positive B cells in peripheral blood mononuclear cells, as measured with flow cytometry, decreased to 1% or lower. The regimen was followed by additional infusion of the same dosage depending on circulating B-cell repopulation recovered over 1%.

[¶] The patient was diagnosed with multiple myeloma after an on-trial relapse of NMOSD. She received combined chemotherapy of VMP regimen (Bortezomib+Melphalan +Prednisolone) and partial remission was achieved at the last follow-up.

Table S2: Patients who withdrew from the trial.

Sex	Age (yrs)	Diagnostic Diseases	Baseline Immunosu-ppresssants	Baseline EDSS Score	Days in Trial	Adverse Events	Trial Period	Reasons for withdrawal
Tocilizumab group								
Female	63	NMOSD, SLE, Hypertension	PSL+MMF	6·5	310	Edema limbs, nasopharyngitis, urinary tract infection, hyperhidrosis, headache, hypertension, hemorrhagic stroke	Therapy in combination	Severe hemiplegia due to acute hemorrhagic stroke
Female	19	NMOSD, SS	PSL+AZA	6	182	anemia, abdominal distension, nausea, vomiting, abdominal pain, dyspepsia, fever, nasopharyngitis, bronchial infection, back pain	Monotherapy	Death due to severe relapse involving medulla
Azathioprine group								
Female	41	NMOSD, GD	4·4	3·5	280	Nasopharyngitis, neutrophil count decreased, ALT increased, AST increased, GGT increased, headache, depression, rash maculo-papular	Therapy in combination	Severe hepatic dysfunction
Female	40	NMOSD, RA	1·1	3·5	383	Anemia, nausea, platelet count decreased, neutrophil count decreased, ALT increased, AST increased	Therapy in combination	Severe myelosuppression
Female	65	NMOSD	1·8	5·5	170	Dyspepsia, gastric discomfort, nausea, encephalitis infection and meningitis, arthralgia, spasticity, hyperhidrosis	Monotherapy	Death due to post-relapse Listeria monocytogenes meningoencephalitis

Table S3: Comparison of effects of tocilizumab and azathioprine on visual function assessments [*]

	Affected eyes [†]		Difference 95% CI	p value	Unaffected eyes [‡]		Difference 95% CI	p value
	Tocilizumab	**Azathioprine**			**Tocilizumab**	**Azathioprine**		
LogMAR visual acuity [§]	0.0022 (0.0084)	0.0117 (0.0418)	-0.0095 (-0.0191, 0.0002)	0.0558	-0.0002 (0.0019)	-0.0014 (0.0135)	0.0012 (-0.0032, 0.0056)	0.5796
High-contrast letter score (100%)	-0.0874 (0.3886)	-0.4426 (1.8910)	0.3553 (-0.0833, 0.7938)	0.1110	0.0043 (0.0126)	0.0009 (0.0103)	0.0034 (-0.0300, 0.0367)	0.8398
Low-contrast letter score (2.5%)	-0.0361 (0.2473)	-0.1473 (0.4574)	0.1113 (-0.0078, 0.2304)	0.0667	-0.0082 (0.0153)	-0.0246 (0.0120)	0.0164 (0.0292, 0.1415)	0.4190
RNFL thickness, µm [⁑]	-0.1731 (0.3218)	-0.3407 (0.6925)	0.1676 (-0.0097, 0.3449)	0.0637	-0.1130 (0.4017)	0.0069 (0.3916)	-0.1199 (-0.2936, 0.0542)	0.1745
VEP P100 latency, ms	0.1045 (0.3887)	0.5577 (1.4163)	-0.4532 (-0.7909, -0.1156)	**0.0091**	0.0836 (0.2583)	0.0851 (0.2305)	-0.0015 (-0.1093, 0.1063)	0.9784
VEP P100 amplitude, µV	-0.0673 (0.2448)	-0.0371 (0.0824)	-0.0302 (-0.0954, 0.0351)	0.3599	-0.0161 (0.1428)	-0.0297 (0.2028)	0.0136 (-0.0647, 0.0919)	0.7298

RNFL=retinal nerve fibre layer. VEP=visual evoked potential.

[*] Monocular visual function was assessed. We separated the eyes into affected and unaffected eyes. All the values were given as mean rate of per month (SD) from baseline to end of the study.

[†] One attack of optic neuritis (ON) occurred in the tocilizumab group. 3 attacks of ON occurred in the azathioprine group.

[‡] Six attacks of ON occurred in the azathioprine group. No attacks of ON occurred in the tocilizumab group.

[§] An increase in LogMAR visual acuity from baseline represented impairment of vision. An increase in high-contrast letter score and low-contrast letter score represented recovery.

[⁑] A decrease (negative value) represented loss of RNFL thickness. A decrease of macular volume represented degeneration of macula.

Study code: NCT03350633

The following supplement contains the protocols of TANGO trial:

1. Original TANGO study protocol (v1.0)
2. Final TANGO study protocol (v2.0)
3. Summary of changes to TANGO study protocol
4. Original TANGO study statistical analysis plan (v1.0)
5. Final TANGO study statistical analysis plan (v2.0)
6. Summary of changes to TANGO study statistical analysis plan

INVESTIGATOR'S AGREEMENT

PROTOCOL TITLE: A RANDOMIZED, CONTROLLED, OPEN-LABEL, MULTI-CENTER TRIAL TO COMPARE THE SAFETY AND EFFICACY OF TOCILIZUMAB VERSUS AZATHIOPRINE IN PATIENTS WITH HIGHLY RELAPSING NEUROMYELITIS OPTICA SPECTRUM DISORDER (NMOSD): A HEAD-TO-HEAD COMPARATIVE STUDY

PROTOCOL NUMBER: 2017kylc005

I have received and read the Investigator's Brochure for the trial. I have read the 2017kylc005 study protocol and agree to conduct the study as outlined. I agree to maintain the confidentiality of all information received or developed in connection with this protocol. I agree to conduct the trial in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with International Conference on Harmonisation (ICH)/Good Clinical Practice (GCP) and applicable regulatory requirements.

Printed Name of Investigator

_____

Signature of Investigator

_____

Date

_____

SYNOPSIS

Supported by:
Grants from Tianjin Medical University Clinical Research Project (2017kylc005); the National Key Research and Development Program of China (2018YFC1312200); the Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing; National Science Foundation of China (91642205, 81830038 and 81601019)

The Investigational Medicinal Products (IMPs):
Tocilizumab and Azathioprine

Title of Trial:
A Randomized, Controlled, Open-label, Multi-Center Trial to Compare the Safety and Efficacy of Tocilizumab versus Azathioprine in Patients with Highly Relapsing Neuromyelitis Optica Spectrum Disorder (NMOSD): a Head-to-Head Comparative Study

Abbreviated Trial Name:
TANGO

Trial Rationale:
NMOSD is a severe, relapsing immune-mediated inflammatory disorder of the central nervous system (CNS), characterized by attacks of optic neuritis and longitudinally extensive transverse myelitis. The typical marker of NMOSD is the presence of serum autoantibody, termed NMO-IgG, directly against the extracellular domain of the water channel protein aquaporin-4 (AQP4) expressed on astrocytic endfeet. The AQP4-IgG is pathological and may cause astrocytic toxicity, thus rapid progression of disability. Relapse prevention is of paramount importance to reduce the risk of disability accumulation. Currently there are no approved therapies for the treatment of NMOSD. Azathioprine, as a first-line immunosuppressant in many other autoimmune diseases, is also widely used for the prevention of relapse in NMOSD patients. However, the efficacy and safety of azathioprine vary in different individuals. Though it is administered with oral corticosteroids, many patients with NMOSD still relapsed during long-term follow-ups. Recently, interleukin-6 (IL-6) is presumed to be critical in the pathogenesis of NMOSD because it is significantly elevated in the serum and cerebrospinal fluid (CSF) of NMOSD patients and promotes AQP4-IgG production by plasmablasts. Tocilizumab, a humanized IL-6 receptor monoclonal antibody, may provide therapeutic benefits through inhibition of IL-6 signaling pathway. In previous case series studies, tocilizumab reduced annualized relapse rate (ARR) and disability in highly active NMOSD patients. In this way, TANGO trial is intended to compare the safety and efficacy of tocilizumab with azathioprine in patients with highly relapsing NMOSD.

Trial Centers:
5 centers in North China and 1 center in South China

Investigators:
A list containing all Investigators will be provided when all the patients have been screened.

Study period:
Estimated date of the first patient enrolled: Nov, 2017
Estimated date of the last patient completed: Sep, 2019
The mean follow-up period was set 60 weeks.

Objectives:

Primary outcome measures:

- Time to first relapse from the baseline in a time-to-event analysis

Secondary outcomes measure:

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

Exploratory efficacy measures including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks

- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

Methodology:
This is an Investigator-initiated, randomized, parallel-group, multi-center, open-label, time-to-event trial to compare the safety and efficacy of tocilizumab with azathioprine in patients with highly relapsing NMOSD. Eligible patients will be randomized 1:1 to one of two parallel treatment arms: 1) tocilizumab or 2) azathioprine. Patients will receive tocilizumab at the dose of 8 mg/kg every 4 weeks or azathioprine 2-3 mg/kg every day. Tocilizumab will be administered intravenously over an approximate 60 minutes duration. Adjustment of the infusion rate and symptomatic treatments (prednisone and diphenhydramine) are permitted to manage infusion-related reactions. Azathioprine will be initiated at 25 mg/d orally and increased stepwise by 25 mg/d increments until the target dose is reached. Patients experiencing medication-related side effects during the loading period are allowed symptomatic treatments, with the exception of any new immunosuppressants. Patients receive their final, stable dosage of azathioprine, as daily maintenance until relapse, discontinuation, or the end of the trial.

Patients randomized to azathioprine receive concomitant immunosuppressants (oral corticosteroids, mycophenolate mofetil, cyclophosphamide, or methotrexate) for their initial 24 weeks of treatment based on the following schema: 1) patients without prior azathioprine treatment receive 24 weeks of concomitant immunosuppressants, 2) patients receiving azathioprine for < 24 weeks before randomization receive supplementary immunosuppressants until they reach 24 weeks of azathioprine therapy, 3) patients receiving azathioprine for ≥ 24 weeks before randomization receive no concomitant immunosuppressants. All azathioprine patients continue medication as monotherapy after 24 weeks of combined treatment. Patients in the tocilizumab group receive concomitant immunosuppressants for the first 12 weeks, and thereafter tocilizumab is used as monotherapy. The total duration of therapy is 60 weeks since randomization.

Based on the estimated baseline demographics and high relapse rate before the trial, the study is designed to continue until a mean follow-up of 60-week in the overall trial arrives. The overall trial duration is estimated to take approximately 2 years including enrollment. In this time-to-event trial, the trial duration for an individual patient will vary depending on when the patient enters the trial and on the patient's outcome. The course of the trial for an individual patient will consist of **Screening Period**, **Study Period**, and **Safety Follow-up Period**. The end of the study visit for an individual patient will take place when one of the following conditions is met, whichever comes first: (a) the patient experiences a definite relapse and early termination; or (b) when the last participating patient completes the last scheduled visit or when the Investigators decide to discontinue the study or development program. Patients were encouraged to receive sufficient adjusted rituximab or other therapeutic regimen (tocilizumab if possible) after completion of the end of study visit.

Screening Period (1-4 weeks)
At the Screening Visit, after informed consent form (ICF) is obtained from the patient, the patient will be screened for trial eligibility through medical history review, demographic data, and laboratory assessments. The medical history review will include confirmation of the diagnosis of

NMOSD. Relapses within the last 2 years prior to screening must be assessed by the Investigator to determine if they meet the inclusive criteria as specified by this protocol. Detailed information related to relapses within the 2 years prior to the Screening Visit will also be collected. This includes date of onset, clinical presentation for each relapse (e.g., optic neuritis [ON], transverse myelitis [TM]/longitudinally extensive transverse myelitis [LETM], acute brainstem syndrome, acute diencephalic clinical syndrome, symptomatic cerebral syndrome), and Expanded Disability Status Scale (EDSS) score at the baseline. For each historical relapse, the treatment for acute attacks and the regimen to prevent relapses will also be collected and recorded, including the dosage and the course of corticosteroids, the immunosuppressants, intravenous immunoglobulin (IVIG), plasma exchange (PE) or high-dose methylprednisolone (HDMP) treatment. The interval days between the occurrence of the last relapse and randomization will be calculated and recorded.

The baseline immunosuppressants for relapse prevention are allowed during the first 12 weeks for tocilizumab or 24 weeks for azathioprine of the trial under certain restrictions. No new immunosuppressants are permitted during the trial unless the patient experiences a relapse.

Immunosuppressants for the purpose of relapse prevention or treatment of a relapse prior to the screening visit and all other medications for symptomatic treatment within 30 days of screening will be reviewed and recorded on the case report form (CRF). Patients who experience a relapse during the Screening Period will be considered a screening failure. Such patients have the chance to be re-screened after receiving treatment for the relapse and when the patient is medically stable.

Deficient thiopurine methyltransferase (TPMT) activity may suggest an increase in the risk of azathioprine-related side effects.To reduce the risk of dropout for the patients in the azathioprine group, TPMT genetic polymorphisms including four variant TPMT alleles TPMT*2 (G238C), TPMT*3A (A719G/G460A), TPMT*3B (G460A), and TPMT*3C (A719G) for all the patients will be detected. Presence of TPMT mutation (homozygous or heterozygous) will exclude the patients from the trial.

Study Period (5 weeks - 60 weeks and later)
Randomization:
All patients who are cleared for randomization by the Investigator will be randomized on Day 1 on a 1:1 basis to the Tocilizumab Arm or the Azathioprine Arm. This is a completely randomized grouping experiment. The random numbers produced by a binomial distribution determine the group distribution of patients. The randomization will be across centers. Patients will receive either tocilizumab or azathioprine according to the randomization assignment.

The treatment duration for an individual patient varies in this time-to-event trial. All patients must remain on randomized treatment assignment until the end of study or early termination due to special reasons.

Relapse Evaluation:
Identification of potential relapse is critical for patient disability and for the integrity of the trial. Patients will receive a Patient Education Brochure that describes all the signs and symptoms of a definite attacks of NMOSD and the Package Inserts of the IMPs. The ICF contains instructions to contact the study site and the Investigator or the Investigator at the first sign or symptom of a potential relapse. The Investigator should review, in detail, this information and any additional warning signs of a relapse specific to that patient's clinical picture at each visit. Patients should be evaluated within 24 hours for a possible relapse by the Investigator. All reports of possible relapses and actions taken for dealing with a possible relapse must be recorded.

As part of the Relapse Evaluation, a blinded EDSS rater will perform the Kurtzke neurological assessment to determine the Functional System Scores (FSS) and EDSS score, and also finish the

assessments of low-contrast letter scores measured with retro-illuminated 2.5% Sloan letter chart and best corrected high-contrast logMAR visual acuity measured using retro-illuminated Early Treatment Diabetic Retinopathy Study (ETDRS) chart at 2.52 m, respectively.

High resolution SD-OCT images are acquired using identical protocols across centers. The technicians who conduct SD-OCT are unaware of the trial-group assignment. VEP images are also recorded by neuroelectrophysiologists who are also blinded to the trial-group assignment.

MRI is obtained on 3T scanner including enhancement with identical scanning protocols to identify new lesions associated with relapses. MRI scans are analyzed independently at a central MRI reading center (School of Radiology, Tianjin Medical University) by staff members who are also unaware of the trial-group assignment.

Blood will be collected for measurements of serum AQP4-IgG and B cell subsets by the blinded technician. If applicable, CSF samples will also be collected.

An Investigator is composed by the Principle Investigators (PI) across the centers. The Investigator will submit the medical record of the potential relapse in the patients and the Investigator will make the decision as to whether the clinical signs, symptoms and the change meet the definition of a relapse. The Investigator will decide whether each possible relapse meets the pre-defined objective criteria for a relapse. If the event is confirmed as a relapse, the patient must receive immediate treatments, according to the recommended standardized acute relapse treatment regimen.

Follow-up Relapse Evaluation Visits will be performed every 4 weeks after the onset of relapse. Additional unscheduled Follow-Up Relapse Evaluation Visits outside the specified time points are permitted at the discretion of the Investigator. As this trial is a time-to-event trial, patients who experience a relapse will be discontinued from this trial after completion of the week-24 Relapse Evaluation Visit, which also serves as the end of study visit. That is, patients who has a relapse will have a prolonged 24-week follow up.

Patients who complete this trial, either because of a relapse or because the trial is ended when a mean 60-week follow up is achieved, may continue to receive tocilizumab or azathioprine, if they had no relapses.

Safety Follow-up Period (12 weeks)

If a patient withdraws from the trial at any time after receiving any amount of tocilizumab or azathioprine, or does not wish to receive tocilizumab or azathioprine after completion of this trial, the patient will be required to complete an early termination or end of study visit at the time of withdrawal and a further follow-up visit 12 weeks after the last dose of tocilizumab or azathioprine for safety measures.

If a patient is discontinued due to an adverse event (AE), the event will be followed until it is resolved or in the opinion of the Investigator is determined medically stable.

Number of Patients (planned):

This is a randomized open-label parallel-group study, to evaluate the safety and efficacy of tocilizumab and azathioprine in patients with highly relapsing NMOSD. As such, the study is based on observing relapse events. With 1:1 randomization for the trial groups, we calculated that 118 patients would provide a power of 80% to determine the pre-specified between-group difference on the basis of a two-sided log-rank test at a 5% level of significance, assuming a 10% dropout rate.

Primary Efficacy Endpoint

The primary efficacy endpoint is time to first relapse from the baseline in a time-to-event analysis.

The trial will be considered to have met its primary efficacy objective if a statistically significant difference is observed between the tocilizumab treatment group and the azathioprine group. The comparison of the treatment groups for the primary endpoint will use a log-rank test. The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model.

Secondary Efficacy Endpoints

During the Study Period, Baseline is defined as the last available assessment prior to tocilizumab or azathioprine treatment for all patients regardless of their treatment group.

The comparison of the treatment groups for secondary efficacy endpoint - 12 weeks confirmed disability progression and serum AQP4-IgG titres will use a log-rank test.

Diagnosis and main criteria for inclusion:

Inclusion Criteria:

1. Male or female patients ≥ 18 years old
2. Diagnosis of NMOSD as defined by 2015 Criteria by Wingerchuk et al.
3. Historical Relapse (as defined by this protocol) of at least 2 relapses in the last 12 months or 3 relapses in the last 24 months, with at least 1 relapse in the 12 months prior to the screening
4. Able and willing to give written informed consent and comply with the requirements of the study protocol.
5. EDSS ≤ 7.5
6. Men and women of reproductive potential must agree to use a highly effective method of birth control from screening to 6 months after final dose of the experimental product.

Exclusion Criteria:

1. Current evidence or known history of clinically significant infection (Herpes simplex virus, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human immunodeficiency virus, Hepatitis viruses, Syphilis, et al)
2. Pregnant, breastfeeding, or child-bearing potential during the course of the study
3. Patients will not participate in any other clinical therapeutic study or will not have participated in any other experimental treatment study within 30 days of screening
4. Participation in another interventional trial within the last 3 months
5. Heart or kidney insufficiency
6. Tumor disease currently or within last 5 years
7. Clinically relevant liver, kidney or bone marrow function disorder
8. Receipt of rituximab or any experimental B-cell depleting agent within 6 months prior screening and the frequency of CD 19^+ B cells in peripheral blood mononuclear cells, as measured with flow cytometry, exceeded 1%.

- **Criteria for evaluation:**
- **Efficacy:**
- Duration of treatment commences with the randomization. The study period defines the time
- period for assessment of the trial endpoints. A 25% difference in relapse-free probability between the two groups is to be observed. Patients will continue to receive tocilizumab or azathioprine concomitant with corticosteroids or other immunosuppressants from this trial after completion of study. When the trial is stopped, all data from all patients will be collected, and the database cleaned, locked, and analyzed. Data from the study period will be used for efficacy analysis.
- Relapses will be monitored closely throughout the trial and evaluated.
- Disability will be assessed by the EDSS scores comparing the change from the baseline.
- The blinded EDSS Rater will perform the Kurtzke neurological assessment to determine the Functional System Scores (FSS) and EDSS score, and also finish the assessments of low-contrast letter scores measured with retro-illuminated 2.5% Sloan letter chart and best corrected high-contrast logMAR visual acuity measured using retro-illuminated Early Treatment Diabetic Retinopathy Study (ETDRS) chart at 2.52 m, respectively.

- Serum AQP4-IgG titers will be performed by the blinded technician at the protocol specified time points.
- RNFL thickness and GCC volume by SD-OCT and P100 latency and amplitude in VEP at specified time points will also be collected by the blinded technician.
- MRI including enhancement at relapses will be conducted. MRI scans will be analyzed independently at a central MRI reading center by staff members who are also unaware of the trial-group assignment.
- Peripheral B cell subsets will be analyzed by the blinded flow cytometry technician at the protocol specified time points.

Safety:

The safety of tocilizumab or azathioprine will be assessed based on the all treatment-emergent adverse events (TEAEs), Common Terminology Criteria for Adverse Events (CTCAEs 5.0), serious adverse events (SAEs), and the changes from baseline through trial completion in vital signs, electrocardiogram (ECG), routine clinical laboratory tests, and pregnancy tests for female patients of childbearing potential.

Statistical methods:

Analyses will be produced for the study period in order to compare the tocilizumab group with the azathioprine group. The analyses will include efficacy and safety.

Efficacy:

Efficacy analyses will be performed on the Full Analysis Set (FAS) population as well as on the Per-Protocol (PP) population.

FAS Population:

The population on which primary, secondary and additional exploratory efficacy analyses will be performed. This set consists immunosuppressants of all patients who are randomized to treatment and who have received at least 1 dose of drug. Patients will be compared for efficacy according to the treatment they were randomized to receive, irrespective of the treatment they actually received.

PP Population:

The per-protocol population is a subset of the full analysis set population, excluding patients with major protocol deviations. The PP population will include all patients who:

- Have no major protocol deviations or key inclusion/exclusion criteria deviations that might potentially affect efficacy
- Patients who took at least 80% of the required treatment doses while they were in the study period
- Patients who received tocilizumab or azathioprine as monotherapy during the study period

Primary Efficacy Analysis for the Study Period:

Note: During the study period, baseline is defined as the last available assessment prior to tocilizumab or azathioprine treatment for all patients regardless of treatment group.

The primary efficacy endpoint is time to first relapse as defined in the protocol. The trial will be considered to have met its primary efficacy objective if a statistically significant difference is observed between the tocilizumab treatment group and the azathioprine group. The comparison of the treatment groups for the primary endpoint will use a log-rank test. Confidence intervals and p-values will be presented. Kaplan-Meier curves for both treatment groups will be produced. Hazard ratio and risk reduction will be summarized. A sensitivity analysis will be performed on time to first relapse (as identified by the Investigator) using a log-rank test including strata for the randomization stratification variables. In addition, a sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator, and randomization stratification variables as the covariates in the model.

Secondary Efficacy Analysis for the Study Period:
The secondary efficacy analyses will use the available data from the study period.

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

Additional Exploratory Efficacy Analysis including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord MRI
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

The time to onset of confirmed disability progression (12-week confirmation [weeks]) is defined as the time from baseline to the onset of the first disability progression that is confirmed at the next regularly scheduled visit ≥ 12 weeks after the initial disability progression. Baseline for the time to onset of confirmed disability is the date of randomization. Disability progression is defined as an increase of ≥ 1.0 point from baseline EDSS score if the baseline EDSS value is ≤ 5.5 points (inclusive) or an increase of ≥ 0.5 points if the baseline EDSS value is > 5.5 points. Assessments within 30 days after a protocol-defined relapse will not be used for confirmation of confirmed disability progression. The non-confirmatory EDSS assessments (between the initial disability progression and the confirmation of disability progression should also fulfill the requirements of the progression.The comparison of the treatment groups for secondary efficacy endpoint--12 weeks confirmed disability progression will use a log-rank test including strata for the randomized intervention variable. The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model.

Cell based assay and Fluorescence immunoprecipitation assay will be used to dynamically evaluate the serum AQP4-IgG titers of patients in this study at the time of baseline, 12 weeks, 24 weeks, 36 weeks, 48 weeks and 60 weeks from the randomization. Wilcoxon rank sum test is used to compare the serum AQP4-IgG titers of two groups.

Other exploratory efficacy endpoints listed in the protocol will also be analyzed but will be reported in separate study reports or publications.

Safety:
Safety analyses include all patients who receive at least 1 dose of trial drug. Patients will be compared for safety according to the treatment they actually received. All adverse events and other safety information including untreated patients collected after the signing of informed consent will be reported in listings, as applicable.

Safety for the Study Period:
For the Study Period, adverse events will be summarized by incidence, system organ class (SOC), preferred term (PT), seriousness, severity, relationship to treatment, and by treatment group.

Concomitant medications will be summarized by treatment group.

Changes from Baseline in vital signs and laboratory assessments (chemistry and hematology) will be summarized by treatment group. Likewise, shift tables (L [low], N [normal], H [high]) by treatment group will be produced for clinical laboratory tests and pregnancy tests will be summarized in patient listings.

TABLE OF CONTENTS

List of Tables

GLOSSARY OF ABBREVIATIONS

ADE	Adverse drug events
AE	Adverse event
APRIL	A proliferation-inducing ligand
AQP4	Water channel protein aquaporin aquaporin-4
ARR	Annualized relapse rate
AZA	Azathioprine
BAFF	B-cell activating factor
BP	Blood pressure
CBC	Complete blood count
CDP	Confirmed disability progression
CFDA	China Food and Drug Administration
CNS	Central nervous system
CRF	Case report form
CSF	Cerebrospinal fluid
CTCAEs	Common Terminology Criteria for Adverse Events
Cz	Midline central
DMARDs	Disease-modifying anti-rheumatic drugs
DTI	Diffusion tensor imaging
ECG	Electrocardiogram
EDSS	Expanded Disability Status Scale
EFNS	European Federation of neurological Societies
EMA	European Medicines Agency
ETDRS	Early Treatment Diabetic Retinopathy Study
EULAR	European League Against Rheumatism
FAS	Full Analysis Set
fMRI	Resting-state functional MRI
FSS	Functional System Scores
FUs	Fluorescent units
Fz	Midline frontal
GCC	Ganglion cell complex
GCP	Good Clinical Practice
H	High
HEK293	Human embryonic kidney 293
HR	Heart rate
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
IEC	Independent Ethics Committee
IL-6	Interleukin-6
IPND	International Panel for NMOSD Diagnosis
IRB	Institutional Review Board
iTregs	Regulatory T cells
IV	Intravenous
IVIG	Intravenous immunoglobulin
IVMP	Intravenous methylprednisolone
L	Low
LCLA	Low-contrast letter acuity
LETM	Longitudinally extensive transverse myelitis
MedDRA	Medical Dictionary for Regulatory Activities
MMF	Mycophenolate mofetil
MRI	Magnetic resonance imaging

MS	Multiple sclerosis
MTX	Methotrexate
N	Normal
NMOSD	Neuromyelitis Optica Spectrum Disorder
ON	Optic neuritis
Oz	Occipital midline
PBs	Plasmablasts
PE	Plasma exchange
PH	Proportional hazards
PI	Principal Investigator
PP	Per-Protocol
PT	Preferred Term
RA	Rheumatic arthritis
RNFL	Retinal nerve fiber layer
RR	Respiration rate
RTX	Rituximab
SAEs	Serious adverse events
SAP	Statistical Analysis Plan
SD-OCT	Spectral-domain optical coherence tomography SD
sJIA	Systemic juvenile idiopathic arthritis
SOC	System organ class
SWI	Susceptibility weighted imaging
TCZ	Tocilizumab
TEAEs	Treatment-emergent adverse events
TGF	Transforming growth factor
TM	Transverse myelitis
TPMT	Thiopurine methyltransferase
VA	Visual acuity
VAS	Visual analogue scale
VEP	Visual evoked potentials
WBC	White blood cell

5. BACKGROUND

5.1 Neuromyelitis Optica Spectrum Disorder (NMOSD)

Neuromyelitis optica spectrum disorder (NMOSD), also known as neuromyelitis optica (NMO), is an autoimmune inflammatory disease of the central nervous system (CNS), which preferentially affects the optic nerves and spinal cord[1]. Though it was previously thought to be a subtype of multiple sclerosis, clinical features, neuroimaging, immunological, and histopathological characteristics distinguish NMOSD from classical MS since the presence of specific serum autoantibody aquaporin-4 (AQP4-IgG) was found in 2004[2]. More than 60%-80% of the patients with NMOSD have detectable AQP4-IgG (also known as NMO-IgG) in the serum. For the patients with negative AQP4-IgG in the serum, the diagnostic criteria have been refined in 2015 to increase the diagnostic accuracy to a large extent[3].

Unlike MS, most patients with NMOSD experience a severe relapsing course[4], thus risk of increased progression in disability makes NMOSD a devastating disease in terms of social and economic burden. NMOSD is a quite important cause of irreversible neurological disability in Chinese young adults. To date, there are few definite epidemic data worldwide for NMOSD patients[5]. But NMOSD is more common in certain regions of Asia, Africa and Latin America, where MS is relatively rare[6,7]. This is different from European countries. Similarly, till now there is no sufficient epidemic data of NMOSD in China. In some clinics specialized in NMOSD, the number of NMOSD patients is about 10 times as many as that of MS[8].

5.2 Clinical Features of NMOSD

The clinical hallmarks of NMOSD are acute optic neuritis and transverse myelitis that involves more than 3 vertebral levels, described as longitudinally extensive transverse myelitis (LETM). These clinical events can occur either simultaneously or in isolation. Signs and symptoms attributable to lesions beyond the optic nerves and spinal cord can also occur in patients with NMOSD and are reported in about 15% of patients. So, the revised diagnostic criteria were released by the International Panel for NMOSD Diagnosis (IPND) in 2015. NMOSD was stratified by AQP4-IgG serostatus, that is, AQP4-IgG-serosepositive NMOSD and AQP4-IgG-seronegative NMOSD (or unknown serostatus)[3]. In AQP4-IgG-seropositive NMOSD, if AQP4-IgG is reliably positive (cell-based assay is preferred) and alternative diagnoses are excluded, only core clinical characteristic (optic neuritis, acute myelitis or brain syndrome) is required for the diagnosis. Brain syndromes, such as area postrema syndrome manifested by intractable hiccup, nausea and vomiting, acute brainstem syndrome, symptomatic narcolepsy or acute diencephalic clinical syndrome with NMOSD-typical diencephalic MRI lesions and symptomatic cerebral syndrome with NMOSD-typical brain lesions. About 60%-80% of NMOSD patients are AQP4-IgG-seropositive. On the other hand, for the diagnosis of AQP4-IgG-seronegative NMOSD, more stringent criteria were set to exclude a variety of diseases mimicking NMOSD, although exclusion of alternative diagnoses was imperative (see *Appendix 1 and Appendix 2*).

Most frequently, NMOSD develops as a recurrent disorder, which predominantly affects women.The majority of patients suffer from a recurrent course (80 – 90%) whilst monophasic (10 – 20%) and primary or secondary progressive courses are rare. Clinical features are often severe involvements of optic nerve and/or spinal cord and also many other symptoms and signs. Acute to subacute onset of blindness in one or both eyes, preceded or followed by a severe transverse or ascending myelitis. The spinal cord lesions are often necrotizing, more likely to lead to permanent than those of typical demyelination in MS. If myelitis extends up into the brainstem causing respiratory failure and death. Uncommon features such as vomitting and hiccups reflect damage in the area postrema. If brainstem is involved, the patient may have vertigo, hearing loss, diplopia, ptosis or nystagmus. Hypothalamic involvement, could cause hypersomnolence, hypothermia or hyperthermia[9,10].

Once a relapsing course has been established in NMOSD, recurrent optic neuritis and myelitis attacks result in stepwise accumulation of neurological disability. Within 5 years, more than 50% of such

patients are functionally blind (visual acuity worse than 20/200) or have lost the ability to ambulate without assistance. In NMOSD patients, the disability accumulation is associated with relapses. Therefore, relapse prevention is paramount for successful treatment of relapsing NMOSD. As NMOSD has the potential to cause significant disability, the prognosis of relapsing NMOSD is poor. The 5-year mortality of NMOSD was reported to be 30%. 50% individuals sustain permanent severe disability, visual (blind in one or both eyes) or ambulatory (requiring a wheelchair). Most deaths result from neurogenic respiratory failure secondary to a high cervical cord or brainstem lesion. Frequent early relapses predict a poor prognosis[11].

5.3 Different Treatment Options in NMOSD

Acute NMOSD relapses are generally treated with high-dose intravenous (IV) steroids. Plasma exchange (PE) or intravenous immunoglobulin (IVIG) often used as a rescue therapy for those who do not respond. After acute phase, supportive treatments against relapse are necessary. But there is no therapy approved for the prevention of NMOSD. Of the off-label prescription of immunosuppressive agents, corticosteroids, azathioprine (AZA), mycophenolate mofetil (MMF) and rituximab (RTX) are commonly used for long-term prophylaxis[12-14]. However, a significant number of patients will continue to have attacks resulting in additional and permanent neurological deficits and disability when using these empirical drugs[15]. Empirically, AZA is the most frequently used to prevent relapses in NMOSD. However, a well-defined protocol for choosing the most optimal treatment is lacking[16]. Given the seriousness of the disease, limitations of currently available therapies, and the limited options for treatment, there remains a significant unmet medical need for an effective and safe treatment for NMOSD. There are no streamline disease modified drugs as MS in European or North American countries. Many patients with NMOSD in China received irregular treatments for prevention, or even received no preventive treatments. As a result, higher rate of relapses was common in these patients. To reduce relapse rate and improve the disability of NMOSD patients, clinical trials, especially that can provide high-level evidence, are urgent on comparing the efficacy and safety of the drugs available in China.

5.4 Efficacy and Tolerability of AZA for Preventing Relapses of NMOSD

A review published in 2016 and the guideline by the European Federation of neurological Societies (EFNS) issued in 2010 recommended AZA, a purine analog that blocks deoxyribonucleic acid synthesis particularly of rapidly proliferating B and T lymphocytes, as one of the first-line therapies for NMOSD[16,17]. However, these recommendations were only based on studies involving AZA with few comparator groups; and therefore, the reported measures of therapeutic effect may be remarkably biased. For instance, several NMOSD case series involving patients who took AZA had shown that post-treatment ARR was significantly reduced compared to pre-treatment ARR and there was a trend towards improved or stabilized EDSS scores at post-treatment compared to EDSS scores at pret-reatment[18,19]. However, these beneficial results of AZA should be brought to proper perspective since pre-treatment and post-treatment measurements are not independent of each other and these are affected by natural processes and patient characteristics. To control for these variables, between-group comparison must be performed so the effects of the intervention can be discerned. More studies suggested that AZA may be inferior to RTX in terms of reduction of ARR and reduction of EDSS while fewer other studies revealed no difference in terms of ARR and EDSS[20]. In the comparison between AZA and MMF groups, few studies did not find significant differences in ARR or and EDSS[21].

Adverse drug events (ADE) associated with AZA were seemingly frequent and may contribute to patient non-adherence to prescribed medication. In the pooled analyses of ADE, there was significant number of patients in the AZA group who had experienced any adverse event, including leukopenia, hair loss, elevated liver enzymes, hepatotoxicity or severe myelosuppression compared to MMF[22]. The increased occurrence of ADE in patients who took AZA may be due to possible genetic mutations at the TPMT*3C in those individuals which causes decreased levels of thiopurine methyltransferase (TPMT) leading to toxicity. Although the AZA treated individuals with TPMT*3C heterozygous or

homozygous genetic mutation may contribute to these adverse events, only 10% of the side effects may be elucidated by TPMT insufficiency[23].

Since AZA is often used in conjunction with corticosteroids, the efficacy of AZA as monotherapy, remains unclear. Due to the overall limited therapeutic evidence for AZA treatment, it is highly recommended that the implementation of well-conducted, randomized, controlled clinical trials is imperative to determine with certainty the role of AZA for patients with NMOSD.

5.5 Role of IL-6 Signaling Pathway in NMOSD

NMOSD is thought to be immunological distinct from MS. Abnormal levels of multiple cytokines in the serum and CSF have been reported in NMOSD patients[24]. Of the diversity of cytokines, only the IL-6 level showed significant elevation in NMOSD in serum analysis. The CSF level of IL-6 was significantly increased in NMOSD compared with MS and other neurological diseases. Importantly, the CSF IL-6 level had a significant correlation with the CSF glial fibrillary acidic protein (GFAP) level and CSF cells, and a weak correlation with AQP4-IgG titers[25]. The GFAP level is remarkably high in the CSF of AQP4-IgG-seropositive NMOSD patients during acute exacerbations whereas the CSF GFAP is not elevated at all in typical multiple sclerosis. CSF IL-6 levels are directly associated with CSF GFAP, reflecting IL-6-induced astrocytic damage. The amount of AQP4-IgG present in the CNS is associated with astrocyte injury and inflammatory responses during NMOSD attacks[26]. AQP4-IgG selectively induced IL-6 production by AQP4-positive astrocytes and that IL-6 signaling to endothelial cells decreases barrier function, increases chemokine production, and enhances leukocyte transmigration under flow. These effects were reversed after application of IL-6 neutralizing antibody[27]. In NMOSD, increased CSF matrix metalloproteinase-2, likely induced by interleukin-6 signaling, may disrupt the blood-brain barrier and enable serum AQP4-IgG migration into the CNS[28]. The crosstalk between IL-6 and AQP4-IgG plays important roles in pathogenesis of NMOSD.

Previous studies have proved that $CD19^{int}CD27^{high}CD38^{high}CD180^{-}$ plasmablasts (PBs), which are a subset of terminal stage of B cells, are capable of producing autoantibodies, including AQP4-IgG. The survival of PBs and the production of AQP4-IgG are both promoted by IL-6, but not B-cell activating factor (BAFF) and a proliferation-inducing ligand (APRIL)[29]. IL-6 was originally cloned as B-cell stimulatory factor-2. IL-6 activates a receptor complex consisting of the IL-6 receptor (IL-6R) and the signal-transducing receptor subunit gp130. IL-6 can bind to both a transmembrane form and a soluble form of IL-6R, which can interact with gp130 to trigger downstream signal transduction and gene expression. Apart from B cells, IL-6 exerts pleiotropic effects on other immune cells. Th1 cells can be differentiated from naïve T cells by IL-6 and transforming growth factor (TGF)-β, whereas IL-6 inhibits the differentiation into TGF-β induced regulatory T cells (iTregs). IL-6 dysregulation is involved in the development of many autoimmune diseases such as systemic lupus erythematosus (SLE), rheumatic arthritis (RA). In SLE, in vivo blockade of the IL-6 receptor by tocilizumab decreases lymphocyte activation and restores B and T cell homoeostasis by either blocking differentiation and/or trafficking in patients with SLE and leads to normalization of the abnormal B and T cell subsets seen at baseline[30]. Inhibition of IL-6 receptor IL-6 signal-blockade therapy may serve as a promising option to decrease disease activity in NMOSD.

5.6 Therapeutic Strategy of Tocilizumab (TCZ) as the First IL-6R Monoantibody Inhibitor

In rheumatic arthritis (RA), characterized by systemic and joint inflammation, TCZ showed protective efficacy against joint destruction and suppressed disease activity. TCZ is the sole biologic that, as monotherapy, shows greater efficacy than standard disease-modifying anti-rheumatic drugs (DMARDs) and methotrexate (MTX). TCZ add-on or switch to TCZ monotherapy was studied for patients with active RA despite MTX. According to the 2016 European League Against Rheumatism (EULAR) updated recommendations for the management of RA, IL-6 pathway inhibitors may have some advantages (i.e., safety) compared with other biological agents in patients who cannot use

conventional synthetic DMARDs as co-medications[31]. This provides favorable information of TCZ in the treatment of NMOSD. Based on the available safety data in the RA trials, the adverse effects of TCZ have
been shown to be manageable, reversible, and usually not treatment limiting. In addition, based on the RA safety data, it is justified to conduct a clinical study in adult NMOSD patients, using TCZ doses of 8 mg/kg every 4 weeks in hospitals.

Indeed, treatment with TCZ ameliorated AQP4-IgG production in the serum of NMOSD patients and suppressed relapse rate and neuropathic pain in patients with refractory NMOSD, suggesting its clinical efficacy[32-34]. In a retrospective observational study, TCZ significantly reduced EDSS, active MRI lesions and also serum AQP4-IgG titers. But attacks occurred when TCZ dosage was reduced[35]. Though TCZ may be a promising drug for preventing acute attacks in patients with NMOSD, the adjusted dosage and long-term efficacy and safety after TCZ is initiated warrant an investigation in large cohorts of patients with NMOSD.

6. STUDY RATIONAL

To date, there is no standard treatment for prevention of NMOSD relapses. Evidence for choice of drugs available still lacks. We aim to look for efficient regimen with high level of evidence. This study is a pivotal head-to-head clinical trial comparing the safety and efficacy of TCZ and AZA. All patients will be permitted to use immunosuppressive therapy for a washout period after randomization. Efficacy of monotherapy of TCZ or AZA afterwards will also be evaluated. Opinions vary widely among Investigators regarding ethics of placebo-controlled studies for maintenance treatment of NMOSD[36].Though US FDA strongly prefers a pivotal monotherapy, placebo-controlled arm, it is in direct contrast to rulings by the European Medicines Agency (EMA), China Food and Drug Administration (CFDA) and the viewpoints of academic and patient-advocacy group for such a devastating disease. AZA is a widely used therapy, even if it was not sufficiently supported by trial evidence and considered experimental rather than standard of care. To protect the patients from more relapses, AZA shows better efficacy than placebo. By combining the available data on first-line treatments AZA and RTX, risks of relapses could be calculated, with AZA 53% and RTX 18%[11]. As previous case series studies showed that TCZ reduced relapse rates of patients who are refractory to RTX treatment, we assume TCZ is no inferior or at least as effective as RTX.

6.1 Statement on the Use of AZA as a Comparator

NMOSD is a severe disabling disease. More than 85% of the patients will have a new relapse within one year if sufficient immunosuppressants is not maintained. To date, no FDA-approved disease-modifying treatments have shown effectiveness with long-term tolerance in NMOSD. AZA has been recommended as the first-line treatment in many countries. But its efficacy varies, especially in most cases it is used with concomitant corticosteroids. The maintenance corticosteroid therapy is of great importance at the beginning of NMOSD treatment, because of the gradual onset of the effects of AZA, which can take several months to exert its full activity. It was challenging to determine the solitary effects of AZA since this regimen was typically administered with oral corticosteroids. It should be noted that the sole administration of steroids may reduce relapses in NMOSD. Therefore, cautious interpretation of efficacy data should be performed due to variability of the interventions[37]. In addition, AZA is metabolized by the enzyme TPMT. Those who have low TPMT activity are more likely to experience side effects, including risk of irreversible myelosuppression. The mutation rate of TPMT in our population limited the bias against using AZA as a treatment modality. So we will detect TPMT gene single nucleotide polymorphisms at Screening Period to reduce the dropout rate for AZA during the trial. After AZA treatment, 54.5% patients reached a relapse-free state in at least two-year follow ups. But steroid-sparing effects of AZA therapy are limited. Overall, the safety issues of AZA may refine it to extensively usage. Its efficacy as monotherapy remained unclear. Head-to-head comparison studies are needed to define the patient groups that will profit most from AZA.

6.2 TCZ or AZA Dose for NMOSD

Empirical data from RA clinical trials support the now approved dose regimen of 8 mg/kg/4w. Baseline concomitant immunosuppressants would be discontinued at 12 weeks, when TCZ will be used as monotherapy completely. Because it may take at least 12 weeks for TCZ to reduce serum IgG1 level significantly compared to baseline in SLE patients receiving TCZ[30], we assume TCZ exert its efficacy at 12 weeks.

In the AZA group, patients will receive oral AZA 2-3 mg/kg every day as a maintenance. AZA will be given to patients while their concomitant corticosteroids or other immunosuppressants are being tapered down until withdrawn. Patients may start at 25 mg/day with the incremental dosages: 50 mg/day, 75 mg/day, 100 mg/day, 125 mg/day, 150 mg/day, to 2-3 mg/kg/day with good tolerability. They must remain on that dose for the duration of the study or until the patient experiences a relapse. It may take 3-6 months for AZA to become biologically active[17]. So patients in the AZA group took concomitant immunosuppressants for 24 weeks, during which period AZA achieved stable efficacy. Thereafter, AZA would be used as monotherapy. If the patients had used AZA for at least 24 weeks before randomization, concomitant immunosuppressants would be discontinued at baseline.

6.3 Rationale for Early Concomitant Usage of Corticosteroids and Immunosuppressants

Usually, initial or recurrent episodes are usually treated with high-dose intravenous methylprednisolone (1 g daily for three to five consecutive days). In many countries of the EU, the intravenous therapy with methylprednisolone is followed by an oral taper and needs to be performed slowly[16]. Some patients experience clinical worsening when prednisone is reduced below 5 -15 mg/day. So immunosuppressant would be added. Taking AZA as an example, as the treatment may only take full effect after 3 - 6 months, it should initially be combined with oral steroid therapy, as oral steroids have been beneficial to suppress disease activity in NMOSD[17].

7. TRIAL OUTCOMES

7.1 Primary Objective

The primary objective of this trial is to compare the safety and efficacy of TCZ treatment as compared to AZA in highly relapsing NMOSD patients based on time to first relapse and relapse risk reduction.

7.2 Secondary Objectives

The secondary objectives of this trial are as follows:

- Reduced risk reduction of confirmed disability progression (CDP) for at least 12 weeks by TCZ compared to AZA
- To determine whether TCZ or AZA could reduce serum AQP4-IgG titers from baseline to 60 weeks in AQP4-IgG positive patients
- To compare the overall safety and tolerability of TCZ to AZA

7.3 Efficacy Endpoints

7.3.1 Primary Efficacy Endpoint

Time to first relapse for TCZ compared with AZA.

7.3.2 Secondary Efficacy Endpoints (TCZ outcomes as compared with AZA outcomes):

The secondary efficacy analyses will use the available data from the study period.

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG from baseline to 60 weeks
- Overall safety and tolerability of TCZ or AZA

7.3.3 Additional Exploratory Efficacy Endpoints:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord MRI
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

8. OVERALL TRIAL DESIGN AND PLAN

This is an Investigator-initiated, randomized, controlled, parallel-group, multi-center, time-to-event trial to evaluate the safety and efficacy of TCZ as compared with AZA in patients with highly relapsing NMOSD. The study will consist of the following periods: Screening Period, Study Period and Safety Follow-up Period.

8.1 Screening Period

At the Screening Visit, after informed consent is obtained from the patient, the patient will be screened for trial eligibility through medical history review, demographic data, electrocardiogram (ECG) and laboratory assessments. The medical history review will include confirmation of the diagnosis of NMOSD. Detailed information on relapses within the last 2 years prior to screening must be assessed by the Investigator to determine if they meet the definition for historical relapses as specified by this protocol. Detailed information related to relapses within the last 2 years will be collected and recorded in the case report form (CRF), if available. This includes date of onset and its clinical presentation for each relapse (corresponding to 6 core symptoms defined by Wingerchuk et al;, 2015; see *Appendix 1* and *Appendix 2*); and EDSS score at the following time points: prior to relapse, at nadir and during recovery, for severity of relapse and recovery. Start/stop dates and dose regimen of all immunosuppressants and non-drug therapies taken for relapse prevention or treatment of a relapse will also be collected and recorded. If PE or IVIG was administered for treatment of a relapse, the number of cycles of PE or IVIG will also be collected. Information on all other previous historical relapses including relapse onset date and its clinical presentations and treatment received for the acute relapse and/or to prevent relapse will also be collected, if available. Only validated diagnostic tests performed by a qualified laboratory are acceptable.

TPMT genotyping will be performed in all the patients. If the patient has homozygous or heterozygous mutation in TPMT gene, then the patient will not be enrolled. If the patient had AZA-related toxicity which caused discontinuation or relapses when AZA was used, the patient will be considered a screening failure.

Supportive immunosuppressants for relapse prevention are allowed during the first 24 weeks in the AZA group or 12 weeks in the TCZ group. The following immunosuppressants are allowed either as mono-therapy or in combination such as corticosteroids, AZA, MMF, MTX, tacrolimus, cyclosporine and cyclophosphamide. No new immunosuppressants are permitted during the trial unless a patient experiences a relapse. Immunosuppressants and/or therapies for NMOSD relapses (either acute treatment or prevention) prior to screening and all other medications taken within 30 days of the Screening Visit will be reviewed and recorded on the CRF.

Patients who experience a relapse during the Screening Period will be considered a screening failure. Such patients may be re-screened for enrollment into the trial after receiving treatment for the relapse and when, in the opinion of the Investigator, the patient is medically stable. The patient must meet the enrollment criteria in order to enter the trial.

8.2 Study Period

All patients who are cleared for randomization by the Investigator will be randomized on Day 1 in a ratio of 1:1 to the TCZ Arm or the AZA Arm. A randomization worksheet will be provided by the School of Public Health of Tianjin Medical University who are to ensure proper randomization. The randomization will be across centers.

Patients will receive either TCZ or AZA according to the randomization assignment and the treatment described in Section 10. The treatment duration for an individual patient varies in this time-to-event trial. All patients must remain on randomized treatment assignment until the end of the Study Period visit.
All the patients we screen have never used TCZ before.

Patients randomized to azathioprine received concomitant immunosuppressants (oral corticosteroids, mycophenolate mofetil, cyclophosphamide, or methotrexate) for their initial 24 weeks of treatment based on the following schema: 1) patients without prior azathioprine treatment received 24 weeks of concomitant immunosuppressants, 2) patients receiving azathioprine for < 24 weeks before randomization received supplementary immunosuppressants until they had reached 24 weeks of azathioprine therapy, 3) patients receiving azathioprine for ≥ 24 weeks before randomization received no concomitant immunosuppressants. All azathioprine patients continued medication as monotherapy after 24 weeks of combined treatment. Patients in the tocilizumab group received concomitant immunosuppressants for the first 12 weeks, and thereafter tocilizumab was used as monotherapy.

The end of the study visit for an individual patient will take place when one of the following conditions is met, whichever comes first: (a) the patient experiences a definite relapse and early termination; or (b) when the last participating patient completes the last scheduled visit or when the Investigator decide to discontinue the study or development program. Identification of potential relapse is critical for patient safety and for the trial. Any potential relapse will be evaluated according to the information below.

Follow-up visits to monitor the course of the relapse and disability progression by EDSS will be performed at 4, 8, 12, 16, 20, 24 weeks after the onset of relapse. Unscheduled Follow-up Relapse Evaluation Visits are permitted and will be made at the discretion of the Investigator. All reports of possible relapses and actions taken for the possible relapse must be documented in the patient's medical chart or source documents and recorded in the CRF.

As this is a time-to-event trial, patients who experience a relapse will be discontinued from this trial after completion of the Week 24 Relapse Evaluation Visit. Thus, the Week 24 Relapse Evaluation Visit also serves as the end of the study visit for these patients. For patients who do not have relapses, the end of the study visit will be defined as when the last participating patient completes the last scheduled visit or when the Investigator decide to discontinue the study or development program. Patients who complete the trial either because of a relapse or because of the trial is ended may be encouraged receive RTX (or TCZ, if possible) treatment.

8.2.1 Relapse Evaluation

Patients will be educated and directed to contact the study site at the first sign or symptom of a potential relapse. Patients should be evaluated within 24 hours of notification of the Investigator of a possible relapse, and no later than 24 hours. All potential relapses must be evaluated by both the Investigator and the blinded EDSS Rater. All reports of possible relapses and actions taken for the possible relapse must be documented in the patient's CRF. At each Relapse Evaluation Visit, the blinded EDSS Rater will perform the Kurtzke neurological assessment and document the Functional System Scores (FSS) and the EDSS score, including 100% high contrast visual acuity and 2.5% low contrast letter acuity. The Investigator will perform a complete neurological examination. The blinded SD-OCT rater will perform SD-OCT was also performed for patients who had optic neuritis. The blinded VEP rater will perform VEP was also performed for patients who had optic neuritis. MRI images will be obtained according to this protocol. Blood will also be collected.

The Investigator determines if the clinical signs, symptoms and neurological change (objective findings on examination) meet the definition for a relapse as outlined in this protocol. After all specified relapse evaluation procedures are complete and the relapse is confirmed, the Investigator may initiate the recommended relapse treatment outlined in this protocol and change the supportive immunosuppressants if needed.

All relapses should be reported in the CRF at the Relapse Evaluation Visit. Relapses that meet the criteria of a severe adverse event (SAE) should also be reported as a SAE.

To monitor the course of the relapse, Follow-Up Relapse Evaluation Visits will be performed at 4, 8, 12, 16, 20, 24 weeks after the onset of relapse. Unscheduled Follow-Up Relapse Evaluation Visits are permitted at the discretion of the Investigator and must be documented in the patient's CRF. During the Relapse Evaluation Period, high-dose intravenous methylprednisolone (IVMP) administration visits may overlap with the Relapse Evaluation Visit and/or Follow-Up Relapse Evaluation Visits.

8.3 Safety Follow-up Period (Post-Treatment)

If patients withdraw from this trial after receiving any amount of TCZ or AZA, a follow-up visit for safety assessments is required at 12 weeks after the last dose of TCZ or AZA. If a patient is discontinued due to an adverse event (AE), the event will be followed until it is resolved or, in the opinion of the Investigator, is determined medically stable.

8.4 Standard Protocol Definitions

8.4.1 Historical Relapse

Historical relapses are the relapses that occurred prior to the Screening Visit, including the onset of the first attack. They are defined as a new onset of neurological symptoms or worsening of existing neurological symptoms with an objective change on neurological examination (confirmed by EDSS, CSF examination, neuroelectrophysiology, MRI findings or all of them) that persisted for more than 24 hours that required treatment. Treatment for relapses includes IVMP, PE or IVIG. Events that occur within a 30-day interval are considered as one relapse.

8.4.2 Definitions for Relapses during the Trial

Relapses are acute attacks that occur during the trial, defined as a new onset of neurological symptoms or worsening of existing neurological symptoms with an objective change on neurological examination that persists for more than 24 hours as confirmed by the Investigator. The signs and symptoms must be attributed to NMOSD, 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. Investigator is not required to wait 24 hours prior to initiating treatment for the relapse.

There are 6 core clinical characteristics in NMOSD. The relapse of each will meet the following:

Optic neuritis:

- The patient has acute vision loss with or without eye pain and optic papillae edema
- Abrupt abnormal visual field associated with optic nerve damage
- At least 5 characters drop in 100% high contrast visual acuity
- New relative afferent pupillary defect
- Abnormal Visual Evoked Potential
- Confirmatory MRI finding in optic nerve
- Exclude anterior segment lesions, retinal lesions, macular lesions, ametropia, glaucoma
- Exclude optic neuropathy caused by ischemia, compression and infiltration, trauma, toxicity and nutritional metabolic agents

Acute Myelitis:

- At least 0.5-point worsening in EDSS if baseline EDSS >5.5 or at least 1 point worsening in EDSS if baseline EDSS ≤ 5.5
- Confirmatory MRI finding in spinal cord

Area postrema syndrome

- Unexplained hiccups or nausea and vomiting
- Confirmatory MRI finding in the area postrema

Acute brainstem syndrome

- Clinical presentations caused by brainstem lesions, i.e., syndrome of medial longitudinal fasciculus, limbic weakness, bulbar palsy
- Confirmatory MRI finding in brainstem

Acute diencephalic clinical syndrome or symptomatic narcolepsy

- Symptomatic narcolepsy, excessive sleepiness or other diencephalic symptoms
- Confirmatory MRI finding in diencephalon

Symptomatic cerebral syndrome

- Clinical presentations caused by cortical or subcortical lesions, i.e., hemiparalysis, aphasia
- Confirmatory MRI finding in the brain

All the confirmatory MRI findings need to be consistent with neuroimaging characteristic of NMOSD (See *Appendix 2*).

8.4.3 The Washout Period of Baseline Immunosuppressants after Randomization

After randomization, patients with AZA group will continue to use AZA if AZA was used for relapse prevention since the last relapse. If AZA has been used for more than 24 weeks when randomization, then AZA will be used as monotherapy, withdrawing any other concomitant immunosuppressants or steroids. If AZA has not been used before or less than 24 weeks, AZA will be used concomitant with the regimen at the time of randomization until concomitant period achieve 24 weeks. If the patient is randomized into TCZ group, then immunosuppressants or steroids at randomization will be gradually withdrawn in 12 weeks.

8.4.4 The Responsibility of the Investigator

The Investigator will be responsible for the overall patient management including patient eligibility evaluation, supervision of the experimental drug administration, recording and treating of AEs and monitoring of safety assessment. At the time of a relapse, the Investigator will perform a complete neurological examination and determine if a patient experiences a relapse and may treat the patient's relapse according to the recommended treatment in Section 10. Treatment for relapse and any changes in the immunosuppressants following a relapse is at the discretion of the Investigator.

8.4.5 The EDSS Rater (Blinded)

The EDSS Rater, an independent neurologist, will be blinded to the randomization and treatments throughout the trial including at the time of a relapse. The rater will perform a complete Kurtzke neurological exam and document the FSS and the EDSS score.

8.4.6 Independent Laboratory Assessors

The laboratory assessors roles are responsible for the specified examination of the corresponding index. They will be blinded to the randomization and treatments throughout the trial including at the time of a relapse.

8.4.7 MRI imaging

MRI including enhancement at relapses will be conducted. MRI scans will be analyzed independently at a central MRI reading center by staff members who are also unaware of the trial-group assignment.

8.5 Schedule of Assessments

8.5.1 Screening Period

Table 1. Screening Period (1-4w)

Trial visit	**V1**
Screening Period Duration	0-4w
Informed consent	×
Demography	×
Physical or surgical history	×
NMOSD history	×
Thiopurine methyltransferase genotyping[a]	×
Pregnancy test (serum) [b]	×
Inclusion/Exclusion criteria[c]	×
Vital signs	×
Physical examination[d]	×
Expanded Disability Status Scale (EDSS)	×
Concomitant diseases and medications[e]	×
Electrocardiogram	×
Clinical laboratory tests[f]	×
Serum AQP4-IgG	×
High-contrast visual acuity (100%)	
Low-contrast letter acuity (2.5%)	
Spectrum Domain-Optical Coherence Tomography (SD-OCT)[g]	
VEP[h]	
MRI[i]	

8.5.2 Study Period

Table 2.Treatment Period

Trail visit	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26
Trial Week	0w	4w	8w	12w	16w	20w	24w	28w	32w	36w	40w	44w	48w	52w	56w	60w	64w	68w	72w	76w	80w	84w	88w	92w	96w
Trial day	D 1	28±7	56±7	84 ±7	112 ±7	140 ±7	168 ±7	196 ±7	224 ±7	252 ±7	280 ±7	308 ±7	336 ±7	364 ±7	392 ±7	420 ±7	448 ±7	476 ±7	504±7	532 ±7	560 ±7	588 ±7	616 ±7	644 ±7	672 ±7
Demography																									
Physical/Surgical history																									
Thiopurine Methyltransferase genotyping[a]																									
Randomization	×																								
Inclusion/Exclusion criteria[c]																									
Vital signs	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Physical examination[d]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Expanded Disability Status Scale(EDSS)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Concomitant diseases and medications[e]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Electrocardiogram	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Clinical laboratory tests[f]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Serum AQP4-IgG	×			×			×			×			×			×			×			×			×

High-contrast visual acuity (100%)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Low-contrast visual acuity (2.5%)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Spectrum Domain-Optical Coherence Tomography (SD-OCT)[g]	×			×			×			×			×			×			×			×			×
VEP[h]（If necessary）	×			×			×			×			×			×			×			×			×
MRI[i]	×	If necessary, according to clinical needs														×	If necessary, according to clinical needs							×
B cell subsets	×			×			×			×			×			×			×			×			×
Toxicity assessment/Adverse Events[j]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Patient Education Brochure [k]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
IMPs administration[l]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

8.5.3 Relapse Evaluation Period

Table 3. Relapse Evaluation Period

Assessment	**Relapse Visit**	**Follow-up Relapse Visit**						
Trial Week	Within 24 Hours	+W4	+W8	+W12	+W16	+W20	+W24	Unscheduled visit
Vital signs	×	×	×	×	×	×	×	×
Physical examination[d]	×	×	×	×	×	×	×	×
Expanded Disability Status Scale (EDSS)	×	×	×	×	×	×	×	×
Concomitant diseases and medications[e]	×	×	×	×	×	×	×	×
Electrocardiogram	×	×	×	×	×	×	×	×
Clinical laboratory tests[f]	×						×	×
Serum AQP4-IgG	×			×			×	×
High-contrast visual acuity (100%)	×	×	×	×	×	×	×	×
Low-contrast letter acuity (2.5%)	×	×	×	×	×	×	×	×
Spectrum Domain- Optical Coherence Tomography (SD-OCT)[g]	×			×			×	
VEP[h]（If necessary）	×			×			×	
MRI[i]	×						×	
B cell subsets	×			×			×	×
Toxicity assessment/Adverse Events[j]	×	×	×	×	×	×	×	×
Survival[m]	×	×	×	×	×	×	×	×

8.5.4 Safety Follow-up Period (Post-Treatment)

Table 4. Safety Follow-up Period (Post-Treatment)

Assessment	Follow-up Visit
Trial Week	+12 w
Vital signs	×
Physical examination[d]	×
Expanded Disability Status Scale (EDSS)	×
Concomitant diseases and medications[e]	×
Electrocardiogram	×
Clinical laboratory tests[f]	×
Pregnancy test (serum) [b]	×
Spectrum Domain-Optical Coherence Tomography (SD-OCT)[g]	
VEP[h]（If necessary）	
MRI[i]	
B cell subsets	
Toxicity assessment/Adverse Events[j]	×
Survival[m]	×

Footnotes：

a. The serum thiopurine methyltransferase activity mainly reflects the ability of the patient to metabolize azathioprine in the liver. If the activity of the enzyme is low, it indicates that the possibility of liver function damage after the application of azathioprine is significantly increased, which may result in the patient not being able to complete follow-up is completed. Therefore, this test is one of the basic conditions for screening patients.

b. Women who are fertile before menopause must have a negative urine or serum pregnancy test within 7 days prior to the first dose. If the result is positive, the subject will not be eligible to participate in the study. If a pregnancy condition is suspected during the test, the test should be repeated.

c. If the assessment is performed within 7 days prior to the first dose and the listed exclusion criteria are met (if applicable), it does not need to be repeated on the first dose, unless the Investigator believes that significant changes may occur.

d. Physical examination includes heart rate, blood pressure, respiratory rate, body temperature, height, weight, and neurological examination. If weight is measured at screening, there is no need to repeat unless there is a clinical indication.

e. The conbined diseases can be obtained from the history of the patient, including all the diseases, irrespective of autuoimmune diseases or other systemic diseases. The collected combined medication data, including drug dose, route of administration, dosing schedule, start date, indications, end date, and end cause, must be recorded from the time the study drug is used until one month after discontinuation. The Investigator gets this information from the patient's hospital follow-up.

f. Hematology examinations such as blood routine and blood biochemistry during the screening period within 7 days before the first administration, followed by blood tests such as blood routine and blood biochemistry should be studied within 7 days before and after the time point specified in the flow chart.

g. Retinal OCT examination must be carried out regularly according to the visiting cycle, recording the thickness of the retinal nerve fiber layer,ganglion cell complex, etc. If clinical recurrence is suspected during the non-visit period, or new lesions are suspected to occur in any part, OCT should be reviewed appropriately and in a timely manner, with the same regular visit.

h. FF-VEP checks include visually induced potential p100 wave and incubation period. If clinical recurrence is suspected during the non-visit period, or new lesions are suspected to occur in any part, FF-VEP should be reviewed appropriately and promptly, with the same regular visit.

i. Neuroimaging examination must be carried out regularly according to protocol, and the number and location of lesions must be recorded. If clinical recurrence is suspected during the non-visit period, or new

lesions are suspected to occur in any area, appropriate and timely imaging examination should be carried out. These imaging examinations include head, neck, thoracic MRI flat-sweep and gadolinium enhancement examination.
j. If any conscious abnormal symptoms, new or worsening neurological symptoms (e.g. limb numbness, weakness, pain, etc.) occurs, the patient should be given appropriate medical assistance immediately. Any symptoms are treated as clinically routine and, if defined, reported as adverse events or severe adverse events (SAEs). The collection of adverse events begins when the patient signs the informed consent form. After the study is terminated, all unhealed adverse events or serious adverse events need to be tracked until resolved, unless the researchers believe that the patient's own illness is unlikely to ease. All new adverse events and serious adverse events within 30 days after the last dose of the study were reported and tracked as described above. Record the most severe levels of toxic reactions in each course of treatment in the CRF.
k.Patient Education Brochure describes all the signs and symptoms of a definite attacks of NMOSD and the Package Inserts of the IMPs. Patients can make a record of all the discomfort during the trial. At each visit throughout the trial, the Investigator will ensure that the patient has the Patient Education Brochure.
l. Patients received oral azathioprine (2-3 mg/kg/d) or intravenous tocilizumab (8 mg/kg/4 weeks). Azathioprine was initiated at 25 mg/d and increased stepwise by 25 mg/d increments until the target dose was reached. Patients experiencing medication-related side effects during the loading period were allowed symptomatic treatments, with the exception of any new immunosuppressants. Patients received their final, stable dosage of azathioprine, as daily maintenance until relapse, discontinuation, or the end of the trial. Tocilizumab was administered intravenously over an approximate 60 minutes duration.
m. The Investigator should obtain information about the patient’s overall survival until the end of the study.

8.6 Schedules of Assessments and Procedures

8.6.1 Screening Examination for Eligibility (Screening Visit [Visit 1]) Occurs 1-4 weeks Prior to Baseline

Once a patient has written the informed consent, the following evaluations will be performed to make sure she or he will fulfill the entry criteria.The following examinations will be performed within 1-4 weeks prior to randomization to determine patient eligibility for participation in the trial:
The Investigator will perform the following:

- Signed informed consent form is indispensible. Note: the rial duration for an individual patient in this time-to-event trial will vary, i.e., the trial duration for in individual patient depends on the time of relapse or the time of trial closure
- Demography
- Physical or surgical history
- NMOSD history: Review the signs and symptoms of potential NMOSD relapse with the patient and instruct the patient to contact the study site at the first signs or symptoms of potential relapse
- AQP4-IgG test history confirmed by cell-based assay (CBA) method
- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS (See *Appendix 3*). Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity (See *Appendix 4*).
- Thiopurine methyltransferase genotyping
- Vital signs include assessments of systolic and diastolic blood pressure (BP), temperature, respiration rate (RR) and heart rate (HR)
- Electrocardiogram (ECG)
- Concomitant diseases and medications
- Clinical laboratory tests (hematology and chemistry) (See *Appendix 5*)
- Pregnancy test (serum human chorionic gonadotropin) must be performed on all women of child-bearing age. Note: if the patient is taking/using contraceptive medication/device, please be sure to record the medication or device in the CRF.
- Provide the brochure for the patients that describes the potential signs and symptoms of NMOSD relapse, package inserts of IMPs and the contact information of the study center.

8.6.2 Study Period

Visit intervals during the Study Period are every four weeks (every 28 days ± 7 days).

8.6.2.1 Baseline ([Visit 2/Day 1]

Once all of the Baseline visit procedures have been performed and the eligibility criteria have been confirmed, the patient will be randomized to one of the two treatment group: TCZ or AZA. The subject randomization numbers will be generated by the randomization center, School of Public Health, Tianjin Medical University and the numbers are to be allocated sequentially in the order in which the subjects are enrolled according to the specification document agreed with the randomization center. As confirmation, the site will be provided with a verification of each patient's randomization. the following tests and procedures will be completed at the Baseline visit (Visit 2/Day 1):

The Investigator will perform the following assessments:

- Review and assess the patient for potential signs or symptoms indicative of relapse
- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Bi-ocular SD-OCT
- Bi-ocular VEP
- Protocol-defined MRI
- Record the dosage of the IMPs
- Provide the brochure for the patients that describes the potential signs and symptoms of NMOSD relapse, package inserts of IMPs, and the contact information of the study center

8.6.2.2 Routine follow-up evaluation (From Visit 3 [Week 4] Through the end of the study or early termination)

The Investigator will review and assess the patient for any potential signs or symptoms indicative of relapse. The following tests and procedures will be completed at every visit until the end of the study or early termination:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Serum AQP4-IgG at protocol-defined visit
- Bilateral monocular SD-OCT at protocol-defined visit
- Bilateral monocular VEP at protocol-defined visit
- Protocol-defined MRI at protocol-defined visit
- Record the dosage of the IMPs
- Provide the brochure for the patients that describes the potential signs and symptoms of NMOSD relapse, package inserts of IMPs, and the contact information of the study center
- Body weight will be measured at 60 weeks or end of study visit or early termination visit

8.6.2.3 Washout Period

The patients will experience a washout period during which baseline immunosuppressants will taper until discontinuation.

- The patients in the TCZ group must taper immunosuppressants till withdrawal at 12 weeks.
- The patients in the AZA group will also discontinue concurrent immunosuppressants at 24 weeks. After withdrawal of concurrent immunosuppressants, patients in both groups will receive mono-therapy without other immunosuppressants. If the patients had used AZA for at least 24 weeks before randomization, concomitant immunosuppressants would be discontinued at baseline. If the patients had received AZA treatment with concomitant immunosuppressants for less than 24 weeks before randomization, the patients will continue to use concomitant immunosuppressants until concomitant period achieved 24 weeks. Then concomitant immunosuppressants will be discontinued.

After the washout period, TCZ or AZA will be administered as monotherapy. No other or new immunosuppressants will be given.

8.6.2.4 Missed Visits

Patients may fail to return for a scheduled visit due to specific factors. In this event, the Investigator must contact the patients by telephone. The reason why the patients cannot get to the study site must recorded in the CRF. If a potential relapse or an AE is considered, patients are strongly encouraged to ask the Investigator for help to assure the missing visit was or not due to a potential relapse or an AE. If the patients cannot come to the study site for examination, then the patients will be instructed to see the local neurologist. The Investigator must contact the local neurologist to obtain as much information as possible about the patient's medical and neurological condition, and provide clinical guidance. The Investigator must record all the relevant medical information.

If the patients decide to withdraw from the trial at any time, a further 12-week follow-up is needed.

8.6.3 Relapse Evaluation Visit (Within 24 Hours)

Patients will be educated and directed to contact the study site if they have new symptoms or worsening on the existing symptoms. The Investigator and the blinded EDSS Rater.will evaluate within 24 hours to determine a possible relapse.

The Investigator will perform the following assessments:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Serum AQP4-IgG at protocol-defined visit
- Bilateral monocular SD-OCT at protocol-defined visit
- Bilateral monocular VEP at protocol-defined visit
- Protocol-defined MRI at protocol-defined visit
- Record the dosage of the IMPs

The Investigator determines if the clinical signs, symptoms and neurological change (objective findings on the examination) meet the definition for relapses as outlined in this protocol. The Investigator will evaluate all the information done as above to assure the relapses. If the relapse is confirmed, the Investigator may initiate the recommended treatment regimen for confirmed relapse

outlined in this protocol (see *Section 10*).After acute attack, patients will be encouraged to receive potent treatments to prevent relapses.

8.6.4 Follow-up Relapse Evaluation Visits (Weeks 4, 8, 12, 16, 20, 24)

If the patients have confirmed relapses, then they must be discontinued from this trial and may experience a prolonged 24-week visit. Week 24 Follow-up Relapse Evaluation Visit will serve as the end of visit.

The following tests and procedures will be completed at these Relapse Evaluation Period visits:

The Investigator will perform the following:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Serum AQP4-IgG at protocol-defined visit
- Bilateral monocular SD-OCT at protocol-defined visit
- Bilateral monocular VEP at protocol-defined visit
- Protocol-defined MRI at protocol-defined visit

8.6.5. Unscheduled Follow-up Relapse Evaluation Visits

Unscheduled Follow-Up Relapse Evaluation Visits outside the specified visits are permitted at the discretion of the Investigator. Procedures, tests and assessments listed under the Relapse Evaluation Visit will be performed at the discretion of the Investigator. Any tests, procedures or assessments performed at the unscheduled visit must be recorded in the CRFs.

8.6.6. Safety Follow-up Period (Post-Treatment)

If a patient withdraws from the trial at any time during the Study Period after receiving any amount of experimental drugs, a follow-up visit for safety assessments is required at 12 weeks after the last dose of IMPs. The following tests and procedures will be completed at the Safety Follow-up Visit:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Pregnancy test (serum human chorionic gonadotropin) will be re-performed on all women of child-bearing age. Note: if the patient is taking/using contraceptive medication/device, please be sure to record the medication or device in the CRF.
- Any new AEs or changes in AEs since the previous visit will be evaluated and recorded

If a patient discontinued from the study due to an AE, the event will be followed until it is resolved or, in the opinion of the Principle Investigator, is determined medically stable.

8.7 Number of Patients

A maximum of 118 NMO patients will be randomized in a ratio of 1:1 (TCZ: AZA) across 6 centers. Patients will be randomized to 1 of 2 treatment groups, 59 for each group.

8.8 Treatment Assignment

If all 118 patients are enrolled, approximately 59 patients will be randomized to TCZ and 59 patients will be randomized to AZA. Randomized patients who discontinue after initiation of study treatment will not be replaced. All patients will remain on their assigned treatment until the end of study visit.

8.9 Criteria for Trial Termination

8.9.1 End of Trial for an Individual Patient

The trial will be ended for a patient when one of the following conditions is met, whichever comes first:

- The patient experiences a definite relapse and early termination
- When the last participating patient completes the last scheduled visit or when the Investigators decide to discontinue the study or development program.

8.9.2 End of Trial for All Patients

The end of the trial has been defined as the date at which the last data point from the last patient is collected.or when the Investigators decide to discontinue the study or development program. The end of trial is defined as completion of the end of study by all patients. The end of study visit is to be completed as soon as possible, preferably within 1 month of the end of the trial notification, with the exception of patients who experience a relapse who will continue until completion of the Relapse Evaluation Period.

9. SELECTION AND WITHDRAWAL OF PATIENTS

9.1 Patient Inclusion Criteria

1. Male or female patients ≥ 18 years old
2. Diagnosis of NMOSD as defined by 2015 Criteria.
3. Historical Relapse (as defined by this protocol) of at least 2 relapses in the last 12 months or 3 relapses in the last 24 months, with at least 1 relapse in the 12 months prior to the screening
4. Able and willing to give written informed consent and comply with the requirements of the study protocol.
5. EDSS ≤ 7.5
6. Men and women of reproductive potential must agree to use a highly effective method of birth control from screening to 6 months after final dose of the experimental drugs.

9.2 Patient Exclusion Criteria

1. Current evidence or known history of clinically significant infection (Herpes simplex virus, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human immunodeficiency virus, Hepatitis viruses, Syphilis, etc.)
2. Pregnant, breastfeeding, or child-bearing potential during the course of the study
3. Patients will not participate in any other clinical therapeutic study or will not have participated in any other experimental treatment study within 30 days of screening
4. Participation in another interventional trial within the last 3 months
5. Heart or kidney insufficiency
6. Tumor disease currently or within last 5 years
7. Clinically relevant liver, kidney or bone marrow function disorder
8. Receipt of rituximab or any experimental B-cell depleting agent within 6 months prior screening and B-cells below the lower limit of normal measured by flow cytometry

9.3 Patient Discontinuation Criteria

The Investigator has the right to discontinue a patient from study drug or withdraw a patient from the study at any time. In addition, patients have the right to voluntarily discontinue study drug or withdraw from the study at any time for any reason. Reasons for discontinuation of study drug or withdrawal from the study may include, but are not limited to, the following:

- Patient withdrawal of consent at any time
- Any medical condition that the Investigator determines may jeopardize the patient's safety if he or she continues in the study
- Investigator determines it is in the best interest of the patient
- Patient non-compliance, defined as significant non-compliance with instructions for home administration of study drug

Patients who discontinue study drug prematurely will be asked to return to the clinic and may undergo follow-up assessments. A follow-up visit for safety assessment is required at 12 weeks after the last dose of experimental drug administration. The primary reason for premature study drug discontinuation should be documented on the CRF. Patients who discontinue study drug prematurely will not be replaced.

If a patient is discontinued due to an AE, the event will be followed until it is resolved or in the opinion of the PI the patient is determined to be medically stable. Every effort will be made to undertake protocol-specified safety follow-up procedures.

9.4 Withdrawal from Study

Every effort should be made to obtain information on patients who withdraw from the study. The

primary reason for withdrawal from the study should be documented on the CRF. Patients will not be followed for any reason after consent has been withdrawn. Patients who withdraw from the study will not be replaced.

10.MEDICATIONS FOR THE PATIENTS

10.1 Investigational Midicinal Product (IMP)

10.1.1 Investigational Midicinal Product (IMP) Dosage and Administration

Each vial of TCZ contains TCZ 80 mg for IV administration. TCZ (8 mg/kg) will be administered to the patient intravenously over approximately 60 mins via an infusion pump. If an AE occurs during the administration of the TCZ, the infusion may be slowed or stopped at the discretion of the Investigator, depending upon the nature and severity of the event. The AE must be recorded in the patient's CRF.

Each tablet of AZA contains 50 mg for oral administration. AZA (2-3 mg/kg) will be administered orally to the patient. AZA was initiated at 25 mg/d and increased in stepwise by 25 mg/d increments until the target dose was reached.

10.1.2 IMP Dose Modifications, Interruptions and Delays

No dose modifications are foreseen.

10.1.2.1 AZA Administration Adjustment

If the patients had medication-related side effects during the loading period, symptomatic treatment was allowed with the exception of any new immunosuppressants. Patients received their maximal tolerated dose of AZA (1-2 mg/kg/d would be also permitted due to side effects), as daily maintenance until relapse, trial discontinuation, or the end of the trial.

10.1.2.2 TCZ administration adjustment

Slowing of the infusion rate or interruption of the infusion, may be necessary in the event of an infusion reaction. In rare cases, TCZ treatment may need to be Discontinued. Handling of infusion related reaction will depend on the intensity of symptoms. In the event that a patient experiences a mild to moderate (Common Terminology Criteria for Adverse Events [CTCAE] Grade 1 or 2; (see *Appendix 6*) non-allergic infusion-related event, the infusion rate should be reduced to half the rate being given at the time of onset of the event. Once the event has resolved, the Investigator should wait for 30 minutes while delivering the infusion at the reduced rate. If tolerated, the infusion rate may then be increased to the next closest rate on the patient's infusion schedule and the rate increments resumed.

Patients who experience a severe infusion-related event (CTCAE Grade 3) or a complex of flushing, fever and throat pain symptoms, should have their infusion interrupted immediately and should receive aggressive symptomatic treatment (such as i.v. diphenhydramine 20-40 mg or 5-10 mg of dexamethasone or both). The infusion should be re-started only after all the symptoms have disappeared. The initial infusion rate at restart should be half of the infusion rate that was in progress at the time of onset of the reaction.

Patients who experience a life threatening infusion-related event (CTCAE Grade 4) during an infusion should have their infusion immediately stopped and should receive appropriate treatment (including i.v. diphenhydramine 20-40 mg or 5-10 mg of dexamethasone or both, use of resuscitation medications and equipment that must be available and used as clinically indicated). These patients will be withdrawn from treatment and should enter the Safety Follow-up Period.

Patients in the TCZ group may prolong infusion interval if they are not tolerable to the dosage of 8 mg/kg, especially when the patients have active infecious diseases.

10.1.3 IMP Packaging and Labeling

TCZ is manufactured by Chugai Pharma Manufacturing Co., Ltd., affiliated by Roche Pharma (Schweiz) Ltd. and purchased from the third Pharmacy. AZA is manufactured by Excella GmbH & Co.KG and purchased from the third Pharmacy. They are stored in the Pharmacy of Tianjin Medical

University General Hospital. Both vials and tablets will be labeled according to the protocol and local regulatory requirements. The drugs will be transported to each participating trial center upon receipt of all required essential documents.

Table 5. Investigational Medicinal Product (IMP)

	IMPs	
Product Name	**Tocilizumab**	**Azathioprine**
Dosage Form	Concentrate solution for infusion	Solid tablet
Unit Dose	80 mg	50 mg
Route of Administration	Intravenous Infusion	Oral
Physical Description	4 mL vial	-
Manufacturer	Chugai Pharma Manufacturing Co., Ltd., affiliated by Roche Pharma (Schweiz) Ltd.	Excella GmbH & Co.KG

10.1.4 IMPs Storage

IMPs kit will have a booklet label describing the contents and a place for the pharmacist to record the patient number, patient initials and Investigator name.

Upon arrival at the center, TCZ should be promptly removed from the cooler and stored in refrigerated conditions at 2°C to 8°C with minimal light exposure. The pharmacist should immediately record the receipt of TCZ or AZA. AZA should be stored at room temperature.

10.2 Concomitant Therapy

Any concomitant medications (including prescription drugs, over-the-counter medications, herbal/homeopathic medications, preventive vaccines, vitamins, and nutritional supplements) must be recorded in the CRF. A description of the type of drug, amount, duration, and reason for administration of drug must be documented. Adverse events that are judged to be related to the administration of a concomitant medication must also be documented on the CRF.

10.2.1 Concomitant Immunosuppressants

Immunosuppressants are permitted to be used in the washout period before experimental drug is given as monotherapy. Immunosuppressants including AZA, MMF, MTX, methotrexate, tacrolimus, cyclosporine or cyclophosphamide are permitted.

If a patient enters the trial receiving steroids either as mono-therapy or in combination with another immunosuppressant, the daily steroid dose will not be more than prednisone 20 mg daily (or equivalent).

No new immunosuppressants or switch to another immunosuppressant is permitted during the trial unless the patient experiences a relapse. After a relapse there are no restrictions on adjustments or changes of immunosuppressants.

10.2.2 Glucocorticoid-Induced Osteopenia/Osteoporosis Prevention and Treatment

Patients should receive oral calcium and 25-hydroxy vitamin D supplementation unless contraindicated (calcium 1200−1500 mg and vitamin D 800−1000 IU daily in divided doses). Unless contraindicated, bisphosphonate therapy (e.g., alendronate 70 mg weekly) will also be administered at the discretion of the physician-investigator for the prevention of glucocorticoid-induced osteoporosis.

Participants with documented osteoporosis will be treated with approved drugs for osteoporosis according to local practice or clinical guidelines.

10.2.3 Other Permitted Medications

Symptomatic treatments are permitted during the course of the trial for underlying conditions. For example, oxcarbazepine could be used for neuropathic pain.

10.2.4 Prohibited Medications

Treatment with any cell-depleting therapies, biologic agents (e.g., tumor necrosis factor antagonists), Janus kinase inhibitors, alkylating agents such as chlorambucil, bone marrow transplantation with total lymphoid irradiation, or extracorporeal photopheresis is prohibited during the study.

Immunization with a live or attenuated vaccine is prohibited within 4 weeks of baseline for the duration of the patient's study participation and 12 weeks after administration of the last dose of study drug.

10.3 Recommended Standardized Relapse Treatment

For this protocol, the treatment for relapse is at the discretion of the Investigator. The following standardized treatment regimen for a confirmed relapse is recommended.

1. One-gram IV methylprednisolone (IVMP) administered daily for 3-5 days followed by an oral prednisone tapering. If the patient improves, then continue the trial assessments as per the schedule of this protocol.
2. If the patient's condition does not sufficiently improve or the neurological symptoms worsen, therapeutic plasma exchange (PE, five to seven cycles; 1 - 1.5 plasma volume per exchange) can be performed.
3. If there is no or minimal response to methylprednisolone, Intravenous Immunoglobulin (IVIG) will be allowed at the discretion of the Investigator. IVIG 0.4 g/kg/d for 5 days is recommended for treatment of attacks that do not respond to IVMP.

11. ASSESSMENT OF EFFICACY

11.1 Efficacy Parameters

Duration of treatment commences with the first experimental drugs administration. The Study Period defines the time period for assessment of the trial endpoints. At the point when the trial is stopped, all data from all patients will be collected, database cleaned, locked, and analyzed. Data from the Study Period will be used for efficacy analysis.

11.1.1 Relapses

Accurate identification and evaluation of confirmed relapses is critical for the integrity of the study. The primary efficacy endpoint is time-to-first relapse. The secondary efficacy endpoints include Time to onset of confirmed disability progression (CDP) for at least 12 weeks, determination of serum AQP4-IgG titers from baseline to 60 weeks and overall safety and tolerability of experimental drugs.

Pre-treatment historical relapses will be reviewed by the Investigator to determine if they meet criteria for Historical Relapse as defined by this protocol. Relapses will be monitored throughout the trial. The Investigator will review, in detail, the signs and symptoms of a potential relapse with the patient at each visit. Patients will be instructed to contact the study site at the first sign or symptom of a relapse. Patients should be evaluated within 24 hours of notification of the Investigator.

All potential relapses must be evaluated by both the Investigator and EDSS Rater. The Investigator will make the decision as to whether the clinical signs, symptoms and neurological changes (objective findings on exam) meet the protocol definition of relapse and may treat the patient's relapse according to the Recommended Standardized Relapse Treatment. The relapse treatment is at the Investigator's discretion. All investigations/tests related to the relapse evaluation should be recorded in the CRF.

Follow-up Relapse Evaluation Visits to monitor the course of the relapse until stabilization will be made according at the Investigator's discretion. All reports of possible relapses and actions taken must be documented in the patient's source documents and recorded in the CRF. Relapses that do not meet the criteria for SAE should be reported as part of the Relapse Evaluation visits, and not as AEs.

The Investigator at each center are responsible for evaluating patient eligibility, supervising the administration of study medication, recording and managing adverse events, assessing relapses. The Investigator identify relapses according to the following criteria: a new onset of neurological symptoms or worsening of existing neurological symptoms with an objective change on neurological examination that persisted for more than 24 hours, signs and symptoms attributable to NMOSD rather than other causes, and onset preceded by at least 30 days of clinical stability. Patients will be evaluated within 24 hours after a possible relapse, which is confirmed by the Investigator and again following intervals of 4 weeks until 24 weeks by the Investigator and by EDSS raters. The EDSS raters, who are also unaware of trial-group assignments, are not involved in patient care. Confirmed relapses could be treated with IVMP or IVIG at the Investigator's discretion.

11.1.2 Disability

Disability will be assessed based on the EDSS scores comparing the change from baseline at the end of trial in the two treatment groups. An EDSS Rater who is blinded to all other trial and patient clinical data will be responsible for performing the EDSS assessments throughout the trial at the protocol specified time points as well as at visits during the Relapse Evaluation Period.

To investigate the efficacy of TCA compared with AZA in patients with relapsing NMOSD, as measured by the time to onset of sustained confirmed disability progression (CDP) over the treatment period, defined as an increase in EDSS that is sustained for at least 12 or 24 weeks, based on regularly scheduled visits. Disability progression is defined as an increase of ≥ 1.0 point from the baseline EDSS when the baseline score is 5.5 or less, and ≥ 0.5 when the baseline score is more than 5.5, that is not attributable to another etiology (e.g., fever, concurrent illness, or concomitant medication).

Confirmation of disability progression must occur at a regularly scheduled visit that is at least 12 or 24 weeks after the initial disease progression. The non-confirmatory EDSS assessments (if any) between the initial and confirmation of disability progression should be at least as high as the minimum change required for progression.

11.1.3 Neurological Functions

Neurological function will be assessed based on EDSS. To assess vision precisely, low-contrast letter scores are measured using retro-illuminated 2.5% Sloan letter chart (Precision Vision, La Salle, IL, USA) using best refractive correction for each eye at 2.52 m[38]. Best corrected high-contrast logMAR visual acuity is measured using retro-illuminated Early Treatment Diabetic Retinopathy Study chart (Precision Vision, La Salle, IL, USA) at 2.52 m. The patient can miss one letter per row to score that row. When no letters could be correctly identified, a score of 1.7 is assigned by the masked researcher[39].

11.1.4 Serum AQP4-IgG Titers

For AQP4-IgG positive NMOSD patients, serum AQP4-IgG titers will be determined by cell-based assay (CBA)[40,41] at scheduled visits.

11.1.4.1 Plasmid construction

We construct AQP4 M23 Plasmid DNA and purify from E. coli XL-2 grown in LB-medium using the Endo Free Plasmid Maxi Kit (QIAGEN). The detailed procedures follow the manufacturer's instructions.

11.1.4.2 Transfection

- Human embryonic kidney (HEK293) cells are seeded on cover glasses with a density of 1×10^5 cells in 12 well.
- When the adherent cells' confluence reaches 70%-80%, the cultured HEK293 cells will be transfected with AQP4 M23 plasmid using Lipofectamine 2000 (Invitrogen) the detailed transfection procedures are followed the manufacturer's protocol.
- After 48 h of incubation, the transfected cells are fixed with PBS and 4% paraformaldehyde. Coated cover glasses are then cut into millimeter-sized fragments. And store in -80℃.

11.1.4.3 Immunofluorescence assay

Each slide coated with AQP4 transfected cells and non-transfected cells, are incubated with the patients' blood samples.

- Incubate each slide with PBS-0.3%Triton for 5min at room temperature.
- Dilute sample with PBS-0.3% Triton into different concentration ratio:1:10,1:60,1:240,1:480,1:720,1:960,1:1440,1:1920, 1:3840, 1:7680.
- The slides are coated with 3%BSA for 30min.
- Incubate with PBS-0.3%Triton for 1h at room temperature.
- Wash with PBS 3 times for 30mins.
- Incubate with goat anti-human IgG (Invitrogen) for 30mins, and then washed as before.

Two independent assessors, who are unaware of patients' timeline classified every blood sample as positive or negative based on the intensity of surface immunofluorescence in direct comparison with non-transfected cells and control samples as positive or negative.

11.1.4.4 Fluorescence Immunoprecipitation Assay (FIPA) for serum AQP4-IgG titers

The changes of serum AQP4-IgG titers will also be confirmed by fluorescence immunoprecipitation assay (FIPA)[42,43]. The protocol for serum AQP4-IgG titers by FIPA is as follows:

- $2\text{-}3\times10^6$ Human embryonic kidney (HEK293) cells were planted in 60 cm^2 dish.
- When the adherent cells' confluence reached 70%, the cultured HEK293 cells were transfected with EGFP-tagged AQP4 M1/M23 plasmid using Lipofectamine 2000 (Invitrogen).

- The detailed transfection procedures are as followed:
  1）Dilute Lipofectamine 2000 60ul in 1.5ml Opti-MEM Medium.
  2）Dilute 15ug DNA in 1.5ml Opti-MEM Medium.
  3）Add diluted DNA to diluted Lipofectamine® 2000 Reagent (1:2 ratio).
  4）Incubate 10 min.
  5）Add DNA-lipid complex to cells.
- After 48 h of incubation, the transfected cells are lysed with extraction buffer (10 mM Tris‐HCl pH 7.5,100 mM NaCl, 1 mM EDTA, 1% Triton X-100), adding 1ml buffer per flask, shaking the flask.
- The supernatant is acquired after centrifugation (4°C 15min, 13000 rpm). And the extracted EGFP-AQP-4 protein is assayed, the fluorescent units (FUs) are assayed and adjusted to 500‐600 FU/200μL extract.
- Preparation of Protein A Sepharose: using 5% ultra-low fetal calf serum (Gibco) to block the sepharose (75μL/case).
- 20μL serum is incubated with 200μL extracted EGFP-AQP-4 protein at 4°C overnight.
- The antibody IgG is precipitated by the addition of 75μL Protein A Sepharose (IgG:sepharose=1:1).
- Washing beads with extraction buffer 3 times (1000 rpm, 3min).
- The beads with 150μL extraction buffer are removed to 96-well black microtiter plates and the FU values (Excitation 488 nm, Emission 507 nm) are obtained on a Microplate reader.

11.1.5 SD-OCT

High resolution spectral domain OCT images were acquired at baseline and the last follow-up using identical protocols across sites (RTVUE100-2, Optovue lnc, Fremont, CA, USA). Appropriate quality assurance was undertaken to ensure comparability, with acceptable inter-rater coefficients of variation for measurements of the RNFL (0.51%) and macular volume (0.45%). RNFL measurements used a 3.45 mm diameter circle scan. A fast macular volume scan (20×20° field, 25 horizontal B scans, ART9) was also done. Scans were excluded if they had a signal strength of less than 25 dB or violated international consensus quality control criteria. If there was a relapse of optic neuritis, OCT of the affected eye will be performed 3 months later.

11.1.6 VEP

Visual evoked potentials (VEP) to reversal achromatic checks were recorded at baseline, relapses and the last follow-up according to International Federation of Neurophysiology guidelines on a Synergy system (EMG & Evoked Potential Response Unit, Nicolet, NE, USA) in standard background office lighting. Responses were recorded from the occipital midline (Oz) using midline frontal (Fz) as reference and midline central (Cz) as ground. Latency and amplitude of the P100 component were measured to one decimal place in the replicates. Participants with absent VEP latencies or amplitudes were assigned a value of 200 and 0, respectively.

11.1.7 MRI

Brain and spinal cord MRI scans will be obtained in all patients at baseline and at week 60, as well as the relapse period. T1 weighted gadolinium-enhancing lesions or new/enlarging hyperintense T2 lesions related to relapses will also performed during relapse period. To evaluate the atrophy of CNS, additional sequences will be done. The scan range of brain includes the whole brain, and the scan range of spine includes C1~C7. The following sequences were performed in head. 3D-T1, 3D-T2 FLAIR, 2D-T2WI, Diffusion tensor imaging (DTI), Resting-state functional MRI (fMRI) , Arterial spin labeling (ASL), and Quantitative susceptibility mapping (QSM) or susceptibility weighted imaging (SWI) was optional sequence. Sagittal PD-T2WI, axial 3D T2*, 3D-T1 and DTI performed in cervical spinal cord. All the scans will be performed by trained and certified MRI technicians.
MRI scans will be analyzed by a centralized center for relapses and CNS atrophy. The centralized reading center is blinded to the treatment assignment and clinical information of the patients. At the

Investigator's site, only the local radiologist/technician assigned to the study may have access to the MRI scans, except at baseline and relapses, when the Investigator may view the MRI scan. If safety concern is considered, the Investigator is permitted to have access to the MRI results.

11.1.8 Flow Cytometry (FACS)

FACS will include (but is not limited to) the following cells:

- Total B cells ($CD19^{pos}$)
- B cell subsets, e.g. memory B cells, naïve B cells, double negative B cells, double positive B cells, and antibody secreting cells.

12. ASSESSMENT OF SAFETY

12.1 Safety Parameters and Definitions

Safety assessments will consist of monitoring and recording adverse events, including serious adverse events and non-serious adverse events of special interest; measurement of protocol-specified safety laboratory assessments; measurement of protocol-specified vital signs; and other protocol-specified tests that are deemed critical to the safety evaluation of the study.

12.1.1 Demographic/Medical History

At Visit 1 (Screening), patients' initials, date of birth, race or ethnic origin and sex will be collected, medical history will be reviewed, and data will be recorded. Medical history including relevant medical/surgical history and NMOSD history will be reviewed and recorded.

12.1.2 Vital Signs

Vital signs will be measured at every visit and will include assessments of systolic and diastolic BP, temperature, RR and HR. Vital signs will be obtained after the patient has been supine or seated for at least 5 minutes. Ideally, each patient's BP should be measured using the same arm. Systolic and diastolic BPs will be documented in mmHg. Temperature will be obtained in degrees Celsius. HR will be documented in beats per minute.

12.1.3 Weight and Height

Body weight will be measured in pounds or kilograms. Height will be measured in inches or centimeters.

12.1.4 Physical Examination

The complete physical examination will include assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination. For consistency, all efforts should be made to have the physical examination performed by the same qualified study staff at these visits.

12.1.5 Electrocardiogram (ECG)

A 12-lead ECG will be conducted. Additional ECG assessments are permitted, at the Investigator's discretion. The Investigator will be responsible for reviewing the ECG to assess whether the ECG is within normal limits and to determine the clinical significance of the results. These assessments will be indicated on the CRF.

12.1.6 Laboratory Assessments

Patients will have biologic samples collected for analysis of various parameters. Chemistry panel, complete blood count (CBC) and hepatic function measures, renal function measures, and serum pregnancy test will be prepared. In the NMOSD patient population that may also be treated with immunosuppressants that are known to affect WBC counts, close monitoring of cell counts is imperative.

Routine hematology laboratory assessment including CBC will be performed at various time points as specified by the protocol and should be reviewed as soon as the lab result is available. Additional assessments to monitor WBC counts can be performed at the discretion of the Investigator as medically indicated. Treatment of leukopenia is at the discretion of the Investigator and consultation with hematologist is encouraged as medically indicated.

AEs and events related to the patients' underlying disease that have occurred during the trial will be collected at every visit.

Any clinically significant, abnormal laboratory result is to be reported as an AE.

12.2 Adverse and Serious Adverse Events

12.2.1 Definition of Adverse Events

12.2.1.1 Adverse Event

According to the International Conference of Harmonisation (ICH) guideline for Good Clinical Practice, an adverse event is any untoward medical occurrence in a patient or clinical investigation patient administered a pharmaceutical product,which does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign, including an abnormal laboratory finding, symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the medicinal product. Pre-existing conditions which worsen during a study are to be reported as AEs.

All the events were recorded and classified according to Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0 (see *Appendix 6*). For AEs with a causal relationship to the experimental drugs, the Investigator must follow-up on the outcome of the event until the event or sequelae either resolve or stabilize.

12.2.1.2 Assessment of Severity of Adverse Events

Table 6 provides guidance for assessing adverse event severity.The adverse event severity grading scale for the NCI CTCAE (v5.0) will be used for assessing adverse event severity. Table 6 will be used for assessing severity for adverse events that are not specifically listed in the NCI CTCAE.

Table 6. Adverse Event Severity Grading Scale

Grade Severity	Description
Grade 1	Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; or intervention not indicated
Grade 2	Moderate; minimal, local, or non-invasive intervention indicated; or limiting age-appropriate instrumental activities of daily living*
Grade 3	Severe or medically significant, but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; or limiting self-care activities of daily living**
Grade 4	Life-threatening consequences or urgent intervention indicated
Grade 5	Death related to adverse event

Activities of Daily Living (ADL)
*Instrumental ADL refer to preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.
**Self care ADL refer to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden.

Adverse events not listed by the CTCAE will be graded using the following criteria:
Grade 1: Discomfort noticed but no disruption of normal daily activity
Grade 2: Discomfort sufficient to reduce or affect normal daily activity
Grade 3: Inability to work or perform normal daily activity
Grade 4: Represents an immediate threat to life (falling into the category of SAE).
In the CRF, adverse events will be reported at each visit.

12.2.1.3 Serious Adverse Events (Immediately Reportable to the Sponsor)

A Serious Adverse Event is any experience that suggests a significant hazard, contraindication, side effect or precaution. It is any AE that, at any dose, fulfils at least one of the following criteria:

- Fatal (i.e., the adverse event actually causes or leads to death)
- Life threatening (refers to an event in which the patient was at immediate risk of death at the time of the event; it does not refer to an event which could hypothetically have caused a death had it been more severe)
- Requires or prolongs inpatient hospitalization
- Results in persistent or significant disability/incapacity (i.e., the adverse event results in substantial disruption of the patient's ability to conduct normal life functions)

- Congenital anomaly/birth defect in a neonate/infant born to a mother exposed to study drug
- medically significant or requires intervention to prevent one or other of the outcomes listed above

When the NMOSD relapse results in hospitalization for any reason other than for routine treatment of the relapse (such as for a treatment course beyond the standard treatment described in or when hospitalization is prolonged, the NMOSD relapse should be considered a SAE.

The terms “severe” and “serious” are not synonymous. Severity refers to the intensity of an adverse event (rated as mild, moderate, or severe, or according to CTCAE criteria); the event itself may be of relatively minor medical significance (such as severe headache without any further findings). Severity and seriousness need to be independently assessed for each adverse event recorded on the CRF.

Serious adverse events are required to be reported by the Investigator within 24 hours after learning of the event.

12.2.2 Hospitalization

AEs that are associated with hospitalization or prolongation of hospitalization are considered SAEs. All admissions to a health care facility meet the criteria, even if for less than 24 h.
Criteria for seriousness are also met if transfer within the hospital is done to receive more intense medical/surgical care (e.g., from the medical floor to the Intensive Care Unit).
Hospitalization does not include the following:

- Rehabilitation facility
- Nursing facility
- Emergency Room
- Same day surgery

Hospitalization or prolongation of hospitalization not associated with an AE is not an SAE, examples include:

- Admission for a pre-existing condition not associated with either a new AE or with worsening of a pre-existing AE
- Protocol-specified admission
- Pre-planned admission

If a relapse of NMOSD result in hospitalization for the patient, then AE can be recorded. If prolonged hospitalization occurs, then a SAE can be recorded.

12.2.3 Causality Assessment

An Investigator causality assessment (Unrelated, Unlikely, Possible, Probable, or Definite) must be provided for all AEs (both serious and non-serious). This assessment must be recorded in the CRF and any additional SAE forms as appropriate. The definitions for the causality assessments appear below.

- Not related (unrelated): This relationship suggests that there is no association between the IMP and the reported event
- Unlikely related: This relationship suggests that the clinical picture is highly consistent with a cause other than the IMP but attribution cannot be made with absolute certainty and a relationship between the experimental drug and the AE cannot be excluded with complete confidence
- Possibly related: This relationship suggests that treatment with the IMP may have caused or contributed to the AE, i.e., the event follows a reasonable temporal sequence from the time of drug administration and/or follows a known response pattern to the experimental drug, but could also have been produced by other factors
- Probably related: This relationship suggests that a reasonable temporal sequence of the event with the IMP administration and the likely association of the event with the IMP. This will be based upon the known pharmacological action of the IMP, known or previously reported adverse reactions to the IMP or class of drugs, or judgment based on the Investigator’s clinical experience
- Definitely related: Temporal relationship to the IMP, other conditions (concurrent illness,

concurrent medication reaction, or progression/expression of disease state) do not appear to explain event, corresponds with the known pharmaceutical profile, improvement on discontinuation, re-appearance on re-challenge

12.2.4 Lack of Efficacy

Since TCZ or AZA treatment in relapsing NMOSD patients is not approved by Food and Drug Administration and CFDA, lack of efficacy is not definite if reported as an AE.The determination of relapses of NMOSD will be based on Investigator's assessment at every visit. If there is any uncertainty as to whether an event is due to disease progression, it should be reported as an adverse event.

12.2.5 Pregnancy

Female patients should take all appropriate precautions to avoid becoming pregnant during this study. In the event that a female patient becomes pregnant during the study or within 90 days after taking the last dose of study drug, she will be instructed to immediately inform the Investigator. The CRF should be completed by the Investigator within 24 hours after learning of the pregnancy. But pregnancy should not be recorded on the Adverse Event CRF. The Investigator should discontinue study drug and counsel the patient, discussing the risks of the pregnancy and the possible effects on the fetus. Monitoring of the patient should continue until conclusion of the pregnancy. Information on the health and well being of the baby will also be collected. Whether the drug is excreted in the semen is unknown. Therefore, pregnancy occurring in the partner of a male patient participating in the study should also be reported to the Investigator.

12.3 Reporting Requirements for Adverse Events

The Investigator must collect all AEs observed, obtained by direct questioning or volunteered from the trial patient.

To further evaluate the adverse events, detailed information about these events should be documented and reported that can be gathered within 24 hours. For non-serious AE, the reporting period starts following the first dose of the IMPs and continues through the last study visit including the Safety Follow-up Visit. AEs, particularly causally related, are to be followed until the event or sequelae resolve or are determined to be medically stable.

The Investigator should follow each adverse event until the event has resolved to baseline grade or better, the event is assessed as stable by the Investigator, the patient is lost to follow up, or the patient withdraws consent. Every effort should be made to follow all serious adverse events considered to be related to study drug or trial-related procedures until a final outcome can be reported.

During the study period, resolution of adverse events (with dates) should be documented on the Adverse Event CRF and in the patient's medical record to facilitate source data verification.

12.3.1 Reporting Requirements for Serious Adverse Events

For SAEs, the Investigator must be notified immediately or within 24 hours of the Investigator site becoming aware of the event, regardless of presumed relationship to the drugs. If the event meets criteria for a fatal or life-threatening SAE, the Investigator should notify Tianjin Medical University General Hospital immediately.

Emergency Medical Contacts
Center Number: [redacted]
Investigator's phone: [redacted]

The Investigator must complete, sign, and date the SAE pages, verify the accuracy of the information recorded on the SAE pages with the corresponding source documents, and send a copy via Email or fax to contact information provided below:
Email: [redacted]
Fax: [redacted]

Additional follow-up information, if required or available, should be emailed or faxed to Tianjin Medical University General Hospital within 24 hours of the Investigator becoming aware of this additional information. Follow-up information should be recorded on the SAE CRF and placed with the original SAE information and kept with the appropriate section of the original subject records and/or trial file.
For all SAEs the Investigator must provide the following:

- Appropriate and requested follow-up information in the time frame detailed above
- Causality of the serious event(s)
- Outcome of the serious event(s)
- Medical records and laboratory/diagnostic information

All the SAEs must be reported to the local Independent Ethics Committee by the Investigators.

12.4 Warnings and Precautions Management of Specific Adverse Events

12.4.1 Opportunistic Infections and Serious Infections

Physicians should exercise caution when considering the use of IMPs in patients with a history of recurring infections or with underlying conditions, which may predispose patients to infections.IMPs should not be administered to patients with active infection. The signs and symptoms of infection should be considered when evaluating a patient for a potential infection, especially for the possible reactivation of viral and other serious infections (e.g., Epstein−Barr virus or varicella-zoster virus). Patients must be instructed to contact the Investigator immediately when any symptoms suggesting infection appear, to ensure rapid evaluation and appropriate treatment. If a patient develops a serious infection, the IMPs will be interrupted until the infection is controlled. The Investigator should consider the benefits and risks to the patient before resuming treatment.

12.4.2 Malignancies

Although no direct association of the IMPs with the malignancies was established, malignancies have been identified as a concern for IMPs in this trial. Once malignancies is suspected, diagnostic procedures will be started to determine if the patients have malignancies. IMPs must be discontinued for patients with malignancies.

12.4.3 Elevated Liver Enzymes

Elevated liver enzymes have been reported in the patients treated with tocilizumab or azathioprine in previous clinical trials and clinical practice. So monitoring liver enzymes is necessary for safety consideration.

Patients withdrawn from the study because of elevated liver function test results must have repeat tests performed as clinically indicated until levels return to baseline values. If the patient’s liver function test results have not returned to normal or to the patient’s baseline level within 6 months (or sooner if deemed necessary by the Investigator), a specialist referral is recommended and an ultrasound should be considered.

12.4.4 Neutropenia

Decrease in neutrophil has been observed following treatment with TCZ or AZA in patients with other autoimmune diseases. Administration of Recombinant Human Granulocyte will be appropriate, if necessary.

12.5 Follow-up Period for Adverse Events

The Investigator must collect all AEs observed, obtained by direct questioning or volunteered from the trial patient. Withdrawal due to an AE or SAE must be clearly differentiated from withdrawal due to other reasons.

For non-serious AE, the reporting period starts following the first dose of the IMPs (Day 1, Visit 2) and continues through the last study visit including the safety Follow-Up Visit. AEs, particularly causally related, are to be followed until the event or sequelae resolve or are determined to be medically stable.

For SAEs the reporting begins following the patient's signing of the ICF (providing consent to participate in the trial) and continues through 12 weeks after the last IMP dose. No time limit on reporting SAEs that are thought to be possibly or probably or definitely related to the IMP.

13.STATISTICAL METHOD AND PLANNED ANALYSES

13.1 General considerations

TANGO is a randomized, multi-center, open-label, two-arm, time-to-event trial comparing tocilizumab and azathioprine in patients with NMOSD. At completion of the trial, the School of Public Health of Tianjin Medical University will analyze the trial data. Further elaboration of statistical issues is provided in the Statistical Analysis Plan (SAP).

The School of Public Health of Tianjin Medical University will be responsible for data collection and editing, reviewing and validating all the information in the CRFs, statistical analysis, and generation of the clinical report.

Prior to locking the database, all data editing will be complete and decisions regarding the evaluability of all patient data for inclusion in the statistical analysis for the Per-Protocol (PP) Set will be made. The rationale for excluding any data from the statistical analyses will be prospectively defined, and classification of all or part of a patient's data as non-evaluable will be completed and documented before the database is locked and before the statistical analysis is begun. The statistical analysis will not begin until the entire database is locked and signed off.

The School of Public Health of Tianjin Medical University will perform the statistical analysis of the data derived from this trial. The analysis will be performed using the SAS® statistical software system Version 9.4.

All summary statistics will be computed and displayed by treatment group and scheduled assessment time. Summary statistics for continuous variables will minimally include n, mean, standard deviation, minimum, median, and maximum. For categorical variables, frequencies and percentages will be presented. Graphical displays will be provided as appropriate.

13.2 Determination of Sample Size

This is a randomized, open-label, multi-center, parallel-controlled trial to evaluate tocilizumab in NMO patients with a primary endpoint of time to first relapse. As such, the trial is based on observing relapse events.

The sample size and power calculation assumptions for this time to first event trial are as follows:

- Log-rank test for comparison of tocilizumab to azathioprine
- 1:1 randomization (tocilizumab:azathioprine)
- Power 80%
- Two-sided 5% level of significance
- Drop-out rate 10%
- Accrual period of approximately 60 weeks
- Relapse-free rate of 85% for tocilizumab and 60% for azathioprine at 12 months (hazard ratio = 0.318)

13.3 Analyses Sets

13.3.1 Full Analysis Set (FAS)

The population on which primary, secondary and other efficacy analyses will include all patients who are randomized to treatment and who have received at least 1 dose of the experimental drug. Patients will be compared for efficacy according to the treatment they were randomized to receive, irrespective of the treatment they actually received.

13.3.2 Per-Protocol (PP) Set

The PP Set is a subset of the FAS, excluding patients with major protocol deviations. The PP population will include all patients who:

- Have no major protocol deviations or key inclusion/exclusion criteria deviations that might

potentially affect efficacy
- Patients who took at least 80% of the required treatment doses while they were in the Study Period (for patients who have relapses any dosing after the relapse will not be included in this calculation)
- Patients who received the treatment of IMP as monotherapy

The PP population will be fully described in the statistical analysis plan, and patients identified prior to database lock.

13.3.3 Safety Set

Safety analyses will be performed on the Safety Set Population. The Safety Population includes all patients who receive at least 1 dose of the experimental drug. Patients will be compared for safety according to the treatment they actually received.

Patients who have signed informed consent but are not treated in the trial are not in the Safety Population. However, if these patients report AEs or SAEs, these events will be summarized separately in tables and listings as appropriate.

13.4 Demographics and Baseline Characteristics

All demographic and baseline characteristics information will be summarized using the following sets: Full Analysis, PP, and Safety Sets. No formal hypothesis testing will be performed. Summary statistics will be presented by treatment group and overall.
Medical history and medical history related to NMOSD will be summarized by treatment group. Listings related to medical history will also be produced.

13.5 Subject Disposition and Treatment Compliance

The number of patients screened, randomized, treated, completing the trial, and included in the safety and efficacy analysis sets will be tabulated by counts and percentage of patients by treatment group and overall. Reasons for any patient withdrawals will be provided.
Treatment compliance with the experimental drug will be summarized using descriptive statistics. The extra usage of the experimental drug for patients who are treated with PE during the trial will be summarized and listings will be produced.

13.6 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by treatment group. Listings of prior and concomitant medications will be produced. Supportive immunosuppressants are allowed during the trial under certain restrictions. The following immunosuppressants are allowed either as mono-therapy or in combination: corticosteroids, AZA, MMF, methotrexate, tacrolimus, cyclosporine and cyclophosphamide. Thus, supportive IST will be summarized by treatment group. Listings of supportive immunosuppressants will be produced. Changes in immunosuppressants during the trial will be summarized.

For patients with relapses, the use of IVMP and IVIG will be summarized by treatment group. Listings of IVMP and IVIG usage by patients with relapses will be produced. Medications will be coded using the World Health Organization Drug Dictionary (WHO Drug).

13.7 Efficacy Analyses

Analyses will be produced for the Study Period in order to compare the tocilizumab group with azathioprine group. The analyses will include efficacy and safety analyses. Efficacy analyses will be performed on the FAS population as well as the PP population. The analysis are prespecified by two subgroups of the patients: with and without concomitant autoimmune diseases.

13.8 Primary Efficacy Endpoint

The primary efficacy endpoint is time to first relapse. The trial will be considered to have met its primary efficacy objective if a statistically significant difference is observed between the tocilizumab treatment group and the azathioprine group. Confidence intervals and p-values will be presented. Hazard ratio and risk reduction will be summarized. A sensitivity analysis will be performed on time to first relapse (as identified by the Investigator) using a log-rank test including strata for the randomization stratification variables. In addition, a sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator, and randomization stratification variables as the only covariates in the model. A second sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator, randomization stratification variables, and region as the only covariates in the model. Region will be defined based on the sites for the trial and will include North America, South America, Europe, Asia-Pacific, and Other as applicable. In the event that the number of patients in some regions is too small to permit modeling, the smaller regions will be pooled together.

An additional sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator and randomization stratification variables as covariates and will also include withdrawals due to AEs as outcomes events (i.e., relapses).
An additional sensitivity of the treatment groups for the primary endpoint will use a log-rank test including strata for the randomization stratification variables for the FAS patients with a follow-up assessment (i.e., FAS patients without a follow-up assessment will be excluded from the analysis).
An additional sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator and randomization stratification variables as covariates only including the relapses asse
ssed within 24 hours. Confirmed relapses assessed after 24 hours will be censoring events based on the patient's last date in the study.

In order to account for relapses that are not evaluated according to the protocol (i.e., within 24 hours of the onset of relapse sign and symptoms), a Cochran-Mantel-Haenszel (CMH) test will be used assessing tocilizumab and azathioprine treatment and patient evaluated within 24 hours (yes, no) for patients with confirmed relapses. Summaries of the severity of the confirmed relapses versus patient evaluated within 24 hours (yes, no) by treatment group will also be produced.
Kaplan-Meier curves for both treatment groups will be produced. Likewise, Kaplan-Meier curves for the strata within each treatment group will be produced.

We pre-specify that the patients in each group will be categorized into two subgroups, the patients with and without concomitant autoimmune diseases.

13.9 Secondary Efficacy Analysis

During the Study Period, Baseline is defined as the last available assessment prior to treatment for all patients regardless of their treatment group.
Unless otherwise specified, the secondary efficacy analyses will use the available data from the Study Period. Hypothesis testing comparing tocilizumab treatment with azathioprine treatment for the secondary efficacy analyses will be performed:

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

The time to onset of confirmed disability progression (12-week confirmation [days]) is defined as the time from baseline to the onset of the first disability progression that is confirmed at the next regularly scheduled visit ≥ 12 weeks after the initial disability progression. If the patient has an infection, dosing may not occur on the Day 1 visit.

Baseline for the time to onset of confirmed disability is the date of randomization, independent of the date of first dosing. For example, a subject with delayed dosing may receive the first dose 2 weeks after the baseline visit; if an EDSS score is recorded between the randomization date and the date of the first dose, this value will be considered for the date of initial disability progression. Disability progression is defined as an increase of ≥ 1.0 point from baseline EDSS score if the baseline EDSS value is ≤ 5.5 points (inclusive) or an increase of ≥ 0.5 points if the baseline EDSS value is > 5.5 points. Assessments within 30 days after a protocol-defined relapse will not be used for confirmation of confirmed disability progression. The non-confirmatory EDSS assessments (between the initial disability progression and the confirmation of disability progression should also fulfill the requirements of the progression. Otherwise, the initial disability progression is not confirmed.

The baseline EDSS value is the average score of the EDSS assessment at screening and baseline (Day 1 visit) up to and including the date of randomization. If one of the values are missing, the non-missing values will be used as baseline.

A blinded rater at each study site will assess EDSS for all patients at the site at screening, baseline, every 12 weeks (regularly scheduled visit) during the treatment period of the study, during the safety follow-up period, at any unscheduled visits, and at withdrawal-from-treatment and end-of-study visits. Additional EDSS assessments for individual patients may be requested between visits (i.e., during an NMOSD relapse).

13.10 Other Efficacy Endpoints

Additional exploratory efficacy measures including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

13.11 Safety Analyses

Safety analyses will be performed on the Safety Population. The Safety Population includes all patients who receive at least 1 dose of the experimental drug. Patients will be compared for safety according to the treatment they actually received.

During the Study Period, Baseline is defined as the last available assessment prior to treatment for all patients regardless of their treatment group.

All AEs and other safety information including untreated patients collected after the signing of informed consent will be reported in listings, as applicable.

AEs will be summarized by incidence, terms of Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, seriousness, severity, relationship to treatment, and by treatment group. Concomitant medications will be summarized by treatment group.
Changes from Baseline in vital signs, laboratory assessments (chemistry and hematology) will be

summarized by treatment group. Likewise, shift tables (L [low], N [normal], H [high]) by treatment group will be produced for clinical laboratory tests and pregnancy tests will be summarized in patient listings.

13.11.1 Physical Examinations and Vital Signs

Physical examinations will be summarized by visit and treatment group. Vital signs (systolic and diastolic BP, temperature, and sitting or supine HR), height, and weight and changes from baseline in vital signs (including height and weight) will be summarized by visit and by treatment group. Listings of physical exams and vital signs will be produced.

13.11.2 Laboratory Assessments

Changes from Baseline in laboratory assessments (chemistry and hematology,) will be summarized by visit and treatment group. Likewise, shift tables (L [low], N [normal], H [high]) by visit and treatment group will be produced for clinical laboratory tests. Listings of laboratory data will be produced.

13.11.3 Adverse Events

SAEs occurring from the signing of informed consent and prior to the initiation of tocilizumab or azathioprine treatment (pre-treatment SAEs) will be summarized by treatment group.

Treatment-emergent AEs (TEAEs) are AEs that onset after the start of treatment in the trial. TEAEs will be summarized by incidence, terms of Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, seriousness, severity, relationship to treatment, and by treatment group. SAEs will be summarized by treatment group. TEAEs and SAEs will be summarized by gender and treatment group, by race and treatment group, and by region (of the world) and treatment group.

AEs and general medical/surgical histories will be coded by Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, corresponding to primary SOC of the Medical Dictionary for Regulatory Activities (MedDRA) (version 22.0).

13.11.4 Other Safety Endpoints

ECG results will be summarized in patient listings. Pregnancy tests will be summarized in patient listings.

13.12 Significance Levels

For all analyses, the tocilizumab treated group will be compared to the azathioprine group and all hypothesis testing will be two-sided and performed at the 0.05 level of significance, unless otherwise specified. Estimates of treatment effect on efficacy parameters will be accompanied by two-sided 95% confidence intervals for the effect size.

13.13 Missing or Invalid Data

For secondary and tertiary efficacy analyses, missing post-Baseline efficacy and safety data will not be imputed unless indicated in the described analysis in the SAP.

14. ETHICAL CONSIDERATIONS

14.1 Compliance with Laws and Regulations

This study will be conducted in full conformance with the ICH E6 guideline for Good Clinical Practice and the principles of the Declaration of Helsinki, or the laws and regulations of the country in which the research is conducted, whichever affords the greater protection to the individual. The study will comply with the requirements of the ICH E2A guideline (Clinical Safety Data Management: Definitions and Standards for Expedited Reporting). Studies conducted in China will comply with National Medical Products Administration regulations and applicable local laws.

14.2 Institutional Review Board (IRB) and Institutional Ethics Committee (IEC)

This protocol, the Informed Consent Forms (ICF), any information to be given to the patient, and relevant supporting information must be approved or given a favorable opinion in writing by an IRB or IEC as appropriate. The Investigator must submit a copy of the written approval to the Tianjin Medical University General Hospital before he or she can enroll any subject into the trial.

The Principal Investigator (PI) is responsible for providing written summaries of the status of the study to the IRB/IEC annually or more frequently in accordance with the requirements, policies, and procedures established by the IRB/EC. Investigators are also responsible for promptly informing the IRB/IEC of any protocol amendments.

Progress reports and notifications of serious adverse drug reactions will be provided to the IRB or IEC according to local regulations and guidelines.

14.3 Written Informed Consent

The PI(s) at each trial site will ensure that the patient is given full and adequate oral and written information about the nature, purpose, possible risk and benefit of the trial. The patient should be given the opportunity to ask detailed questions and allowed time to consider the benefits and risks. Patients must be informed that they are free to discontinue from the trial at any time.

The patient’s signed and dated informed consent must be obtained before conducting the trial. All signed and dated Consent Forms must remain in each patient’s study file or in the site file and must be available for verification by study monitors at any time.

The Consent Forms should be revised whenever there are changes to study procedures or when new information becomes available that may affect the willingness of the patient to participate.The final revised ICF must be approved by IRB/IEC. Patients must be re-consented to the most current version of the ICF (or to a significant new information/findings addendum in accordance with applicable laws and IRB/IEC policy) during their participation in the study.

15. RETENTION AND INSPECTION OF RECORDS

The IRB/IEC will be allowed to inspect facilities and records relevant to this study. The Investigator agrees to allow IRB/IEC to monitor the IMPs storage area, experimental drug stocks, experimental drug accountability records, patient charts and trial source documents, and other records relative to trial conduct.The Investigator must maintain adequate and accurate records to enable the conduct of the study to be fully documented and the study data to be subsequently verified.all documentation relating to the trial will be kept for at least 5 years following the discontinuance of the test article for investigation.

16. CONFIDENTIALITY

Patient medical information obtained by this study is confidential. The Investigators maintain confidentiality standards by coding each patient enrolled in the study through assignment of a unique patient identification number. This means that patient names are not included in data sets. Medical information may be given to a patient's personal physician or other appropriate medical personnel responsible for the patient's welfare, for treatment purposes.

Data generated by this study must be available for inspection upon request by representatives of the collaborators, and the IRB/EC for each study site, as appropriate.

17. REFERENCES

1. Wingerchuk DM, Lennon VA, Lucchinetti CF, Pittock SJ, Weinshenker BG. The spectrum of neuromyelitis optica. The Lancet Neurology 2007;6:805-15.
2. Lennon VA, Wingerchuk DM, Kryzer TJ, et al. A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. Lancet 2004;364:2106-12.
3. Wingerchuk DM, Banwell B, Bennett JL, et al. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. Neurology 2015;85:177-89.
4. Pandit L, Asgari N, Apiwattanakul M, et al. Demographic and clinical features of neuromyelitis optica: A review. Multiple sclerosis 2015;21:845-53.
5. Mealy MA, Wingerchuk DM, Greenberg BM, Levy M. Epidemiology of neuromyelitis optica in the United States: a multicenter analysis. Archives of neurology 2012;69:1176-80.
6. Bukhari W, Prain KM, Waters P, et al. Incidence and prevalence of NMOSD in Australia and New Zealand. Journal of neurology, neurosurgery, and psychiatry 2017;88:632-8.
7. Flanagan EP, Cabre P, Weinshenker BG, et al. Epidemiology of aquaporin-4 autoimmunity and neuromyelitis optica spectrum. Annals of neurology 2016;79:775-83.
8. ZhangBao J, Zhou L, Li X, et al. The clinical characteristics of AQP4 antibody positive NMO/SD in a large cohort of Chinese Han patients. Journal of neuroimmunology 2017;302:49-55.
9. Nagaishi A, Takagi M, Umemura A, et al. Clinical features of neuromyelitis optica in a large Japanese cohort: comparison between phenotypes. Journal of neurology, neurosurgery, and psychiatry 2011;82:1360-4.
10. Sato DK, Lana-Peixoto MA, Fujihara K, de Seze J. Clinical spectrum and treatment of neuromyelitis optica spectrum disorders: evolution and current status. Brain pathology 2013;23:647-60.
11. Levy M. The ethics of placebo controlled clinical trials in NMO - A balance of risks. Multiple sclerosis and related disorders 2015;4:512-4.
12. Papadopoulos MC, Bennett JL, Verkman AS. Treatment of neuromyelitis optica: state-of-the-art and emerging therapies. Nature reviews Neurology 2014;10:493-506.
13. Kimbrough DJ, Fujihara K, Jacob A, et al. Treatment of Neuromyelitis Optica: Review and Recommendations. Multiple sclerosis and related disorders 2012;1:180-7.
14. Stellmann JP, Krumbholz M, Friede T, et al. Immunotherapies in neuromyelitis optica spectrum disorder: efficacy and predictors of response. Journal of neurology, neurosurgery, and psychiatry 2017;88:639-47.
15. Mealy MA, Wingerchuk DM, Palace J, Greenberg BM, Levy M. Comparison of relapse and treatment failure rates among patients with neuromyelitis optica: multicenter study of treatment efficacy. JAMA neurology 2014;71:324-30.
16. Sellner J, Boggild M, Clanet M, et al. EFNS guidelines on diagnosis and management of neuromyelitis optica. European journal of neurology 2010;17:1019-32.
17. Trebst C, Jarius S, Berthele A, et al. Update on the diagnosis and treatment of neuromyelitis optica: recommendations of the Neuromyelitis Optica Study Group (NEMOS). Journal of neurology 2014;261:1-16.
18. Elsone L, Kitley J, Luppe S, et al. Long-term efficacy, tolerability and retention rate of azathioprine in 103 aquaporin-4 antibody-positive neuromyelitis optica spectrum disorder patients: a multicentre retrospective observational study from the UK. Multiple sclerosis 2014;20:1533-40.
19. Costanzi C, Matiello M, Lucchinetti CF, et al. Azathioprine: tolerability, efficacy, and predictors of benefit in neuromyelitis optica. Neurology 2011;77:659-66.
20. 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. Journal of neurology 2017;264:2003-9.
21. Xu Y, Wang Q, Ren HT, et al. Comparison of efficacy and tolerability of azathioprine, mycophenolate mofetil, and cyclophosphamide among patients with neuromyelitis optica spectrum disorder: A prospective cohort study. Journal of the neurological sciences 2016;370:224-8.
22. Chen H, Qiu W, Zhang Q, et al. Comparisons of the efficacy and tolerability of mycophenolate mofetil and azathioprine as treatments for neuromyelitis optica and neuromyelitis optica spectrum

disorder. European journal of neurology 2017;24:219-26.
23. Chen D, Lian F, Yuan S, et al. Association of thiopurine methyltransferase status with azathioprine side effects in Chinese patients with systemic lupus erythematosus. Clinical rheumatology 2014;33:499-503.
24. Uzawa A, Mori M, Kuwabara S. Cytokines and chemokines in neuromyelitis optica: pathogenetic and therapeutic implications. Brain pathology 2014;24:67-73.
25. Uzawa A, Mori M, Arai K, et al. Cytokine and chemokine profiles in neuromyelitis optica: significance of interleukin-6. Multiple sclerosis 2010;16:1443-52.
26. Sato DK, Callegaro D, de Haidar Jorge FM, et al. Cerebrospinal fluid aquaporin-4 antibody levels in neuromyelitis optica attacks. Annals of neurology 2014;76:305-9.
27. Takeshita Y, Obermeier B, Cotleur AC, et al. Effects of neuromyelitis optica-IgG at the blood-brain barrier in vitro. Neurology(R) neuroimmunology & neuroinflammation 2017;4:e311.
28. Uchida T, Mori M, Uzawa A, et al. Increased cerebrospinal fluid metalloproteinase-2 and interleukin-6 are associated with albumin quotient in neuromyelitis optica: Their possible role on blood-brain barrier disruption. Multiple sclerosis 2017;23:1072-84.
29. Chihara N, Aranami T, Sato W, et al. Interleukin 6 signaling promotes anti-aquaporin 4 autoantibody production from plasmablasts in neuromyelitis optica. Proceedings of the National Academy of Sciences of the United States of America 2011;108:3701-6.
30. Shirota Y, Yarboro C, Fischer R, Pham TH, Lipsky P, Illei GG. Impact of anti-interleukin-6 receptor blockade on circulating T and B cell subsets in patients with systemic lupus erythematosus. Annals of the rheumatic diseases 2013;72:118-28.
31. Smolen JS, Landewe R, Bijlsma J, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Annals of the rheumatic diseases 2017;76:960-77.
32. Kieseier BC, Stuve O, Dehmel T, et al. Disease amelioration with tocilizumab in a treatment-resistant patient with neuromyelitis optica: implication for cellular immune responses. JAMA neurology 2013;70:390-3.
33. Ayzenberg I, Kleiter I, Schroder A, et al. Interleukin 6 receptor blockade in patients with neuromyelitis optica nonresponsive to anti-CD20 therapy. JAMA neurology 2013;70:394-7.
34. Araki M, Matsuoka T, Miyamoto K, et al. Efficacy of the anti-IL-6 receptor antibody tocilizumab in neuromyelitis optica: a pilot study. Neurology 2014;82:1302-6.
35. Ringelstein M, Ayzenberg I, Harmel J, et al. Long-term Therapy With Interleukin 6 Receptor Blockade in Highly Active Neuromyelitis Optica Spectrum Disorder. JAMA neurology 2015;72:756-63.
36. Weinshenker BG, Barron G, Behne JM, et al. Challenges and opportunities in designing clinical trials for neuromyelitis optica. Neurology 2015;84:1805-15.
37. Watanabe S, Misu T, Miyazawa I, et al. Low-dose corticosteroids reduce relapses in neuromyelitis optica: a retrospective analysis. Multiple sclerosis 2007;13:968-74.
38. Balcer LJ, Raynowska J, Nolan R, et al. Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. Multiple sclerosis 2017;23:734-47.
39. Raftopoulos R, Hickman SJ, Toosy A, et al. Phenytoin for neuroprotection in patients with acute optic neuritis: a randomised, placebo-controlled, phase 2 trial. The Lancet Neurology 2016;15:259-69.
40. Pittock SJ, Lennon VA, McKeon A, et al. Eculizumab in AQP4-IgG-positive relapsing neuromyelitis optica spectrum disorders: an open-label pilot study. The Lancet Neurology 2013;12:554-62.
41. Zhang C, Tian DC, Yang CS, et al. Safety and Efficacy of Bortezomib in Patients With Highly Relapsing Neuromyelitis Optica Spectrum Disorder. JAMA neurology 2017;74:1010-2.
42. Waters PJ, McKeon A, Leite MI, et al. Serologic diagnosis of NMO: a multicenter comparison of aquaporin-4-IgG assays. Neurology 2012;78:665-71; discussion 9.
43. Wang J, Li X, Zhang DQ, et al. Quantitative analysis of aquaporin-4 antibody in longitudinally extensive transverse myelitis. Journal of neuroimmunology 2015;278:26-9.

18. APPENDICES

Appendix 1: 2015 DIAGNOSTIC CRITERIA FOR NMOSD
Appendix 2: NEUROIMAGING CHARACTERISTICS OF NMOSD
Appendix 3: KURTZKE EXPANDED DISABILITY STATUS SCALE (EDSS)
Appendix 4: VISUAL ACUITY TEST
Appendix 5: CLINICAL LARORATORY TESTS
Appendix 6. COMMON TERMINOLOGY CRITERIA FOR ADVERSE EVENTS (CTCAE 5.0)

APPENDIX 1. 2015 DIAGNOSTIC CRITERIA FOR NMOSD

NMOSD diagnostic criteria for adult patients
Diagnostic criteria for NMOSD with AQP4-IgG 1. At least 1 core clinical characteristic 2. Positive test for AQP4-IgG using best available detection method (cell-based assay strongly recommended) 3. Exclusion of alternative diagnoses
Diagnostic criteria for NMOSD without AQP4-IgG or NMOSD with unknown AQP4-IgG status 1. At least 2 core clinical characteristics occurring as a result of one or more clinical attacks and meeting all of the following requirements: a. At least 1 core clinical characteristic must be optic neuritis, acute myelitis with LETM, or area postrema syndrome b. Dissemination in space (2 or more different core clinical characteristics) c. Fulfillment of additional MRI requirements, as applicable 2. Negative tests for AQP4-IgG using best available detection method, or testing unavailable 3. Exclusion of alternative diagnoses
Core clinical characteristics 1. Optic neuritis 2. Acute myelitis 3. Area postrema syndrome: episode of otherwise unexplained hiccups or nausea and vomiting 4. Acute brainstem syndrome 5. Symptomatic narcolepsy or acute diencephalic clinical syndrome with NMOSD-typical diencephalic MRI lesions 6. Symptomatic cerebral syndrome with NMOSD-typical brain lesions
Additional MRI requirements for NMOSD without AQP4-IgG and NMOSD with unknown AQP4-IgG status 1. Acute optic neuritis: requires brain MRI showing normal findings or only nonspecific white matter lesions, OR optic nerve MRI with T2-hyperintense lesion or T1-weighted gadolinium-enhancing lesion extending over >1/2 optic nerve length or involving optic chiasm 2. Acute myelitis: requires associated intramedullary MRI lesion extending over ≥ 3 contiguous segments (LETM) OR ≥ 3 contiguous segments of focal spinal cord atrophy in patients with history compatible with acute myelitis 3. Area postrema syndrome: requires associated dorsal medulla/area postrema lesions 4. Acute brainstem syndrome: requires associated periependymal brainstem lesions

Abbreviations: AQP4 = aquaporin-4; IgG = immunoglobulin G; LETM = longitudinally extensive transverse myelitis lesions; NMOSD = neuromyelitis optica spectrum disorders.

APPENDIX 2. NEUROIMAGING CHARACTERISTICS OF NMOSD

Neuroimaging characteristics of NMOSD
Spinal cord MRI, acute
LETM lesion associated with acute TM
• Increased signal on sagittal T2-weighted (standard T2-weighted, proton density, or STIR sequences) extending over 3 or more complete vertebral segments • Central cord predominance (more than 70% of the lesion residing within the central gray matter) • Gadolinium enhancement of the lesion on T1-weighted sequences (no specific distribution or pattern of enhancement is required)
Other characteristic features that may be detected
• Rostral extension of the lesion into the brainstem • Cord expansion/swelling • Decreased signal on T1-weighted sequences corresponding to region of increased T2-weighted signal
Optic nerve MRI
Unilateral or bilateral increased T2 signal or T1 gadolinium enhancement within optic nerve or optic chiasm; relatively long lesions (e.g., those extending more than half the distance from orbit to chiasm) and those involving the posterior aspects of the optic nerves or the chiasm are associated with NMO
Cerebral MRI: NMOSD-typical brain lesion patterns (increased signal on T2-weighted MRI sequences unless otherwise noted)
• Lesions involving the dorsal medulla (especially the area postrema), either small and localized, often bilateral, or contiguous with an upper cervical spinal cord lesion • Periependymal surfaces of the fourth ventricle in the brainstem/cerebellum • Lesions involving the hypothalamus, thalamus, or periependymal surfaces of the third ventricle • Large, confluent, unilateral, or bilateral subcortical or deep white matter lesions • Long (1/2 of the length of the corpus callosum or greater), diffuse, heterogeneous, or edematous corpus callosum lesions • Long corticospinal tract lesions, unilateral or bilateral, contiguously involving internal capsule and cerebral peduncle • Extensive periependymal brain lesions, often with gadolinium enhancement

Abbreviations: LETM = longitudinally extensive transverse myelitis lesions; NMOSD = neuromyelitis optica spectrum disorders; STIR = short tau inversion recovery.

APPENDIX 3. KURTZKE EXPANDED DISABILITY STATUS SCALE (EDSS)

(1) Visual function examination

Visual function	L	R	note
Vision (after correction)			The visual acuity score is obtained by reading the Snellen eye chart at a distance of 5 meters from the patient standing. The recognition error cannot exceed one.
Field of view (coarse measurement)			0 = normal 1 = only physical signs, visual field defects only found during formal visual field examination 2 = moderate, the patient can feel the visual field defect, and the visual field examination reveals that the patient is not completely hemian 3=severe, completely ametropic or blind
Visual field defect (perimeter results)			0=none 1 = small, only found in formal inspections 2=large, the patient has a complaint
Optic papilla pale			0=none 1=Yes

0. Normal
1. The papillary pale and / or small blind spots and / or poor eye is less than 1.0 but better than 0.67
2. poor eye corrected visual acuity 0.67-0.34
3. The poor eye has a large blind spot and / or moderate visual field defect and / or 0.33-0.21
4. The poor eye has obvious visual field defect and / or visual acuity 0.2-0.1; 3 plus good eye is less than or equal to 0.33
5. The poor eye is less than 0.1; 4 plus the good eye is less than or equal to 0.33
6.5 plus good eye is less than or equal to 0.33

FS=

FS score correction:：

FS	6	5	4	3	2	1
After conversion FS	4	3	3	2	2	1

（2）brain stem function

Cranial nerve	score	note
Eye movement disorder		0=None 1 = only signs, no complaints (obscure, double vision) 2 = mild, the patient may have a slight EOM disorder that is aware of it, or a clinical examination reveals an incomplete dyskinesia of the eye, but the patient is unaware 3 = moderate, the patient can be aware of the obvious eye movement incomplete movement disorder, or any movement can not fully move when looking at a certain direction 4=Severe, any movement can not fully move when looking in more than one direction
Nystagmus		0=none 1 = only signs or mild: nystagmus when gazing, equivalent to brain stem FS =

		1 2 = moderate, continuous nystagmus at 30° horizontal or vertical gaze, but no nystagmus in situ, patients may or may not complain 4=Severity: In-situ continuous nystagmus, or gross continuous nystagmus in any direction with limited vision, complete internuclear ophthalmoplegia with continuous nystagmus during abduction, vibrating hallucinations
Trigeminal nerve injur		0=none 1 = only signs 2 = mild: clinical numbness can be detected and the patient has a complaint 3 = moderate: trigeminal nerve 1, 2 or 3 branch area tip / blunt discernment disorder, or trigeminal neuralgia (at least 1 episode in the past 24h) 4=Severe: One side or both sides of the face are completely pointed/bluntly dysfunctional
Facial paralysis		0 = no 1 = only signs 2 = mild: clinically detectable facial paralysis, and the patient has a complaint 3=Moderate: Incomplete facial paralysis, such as incomplete closure of the eyelids 4=Severe: unilateral or bilateral complete facial paralysis, resulting in increased ocular fissure, redness of the conjunctiva, difficulty in drinking water
Hearing disorder		0=none 1 = only physical signs: unilateral or bilateral ear hearing finger friction is weakened, Weber experiment is not centered, but the patient has no complaint 2 = mild: same as 1, but the patient has a complaint 3=Moderate: One-sided or two-sided ears cannot hear finger friction, and hear a few whispers 4=Severe: almost all whispering numbers are heard
Dysphonic disorder		0=none 1 = only signs 2 = mild: clinically detectable enough sound disorder, the patient has a complaint 3=Moderate: obvious enough obstacles in daily conversation, affecting understanding 4=Severity: a language that cannot be understood 5=Word can't
Swallowing disorder		0=none 1 = only signs 2 = mild: swallowing thin fluid barrier 3 = moderate: swallowing liquid and solid food barriers 4 = Severe: continuous swallowing disorder, need a muddy diet 5=Can’t swallow
Other cranial nerve function		0 = normal 1 = only signs 2 = mild: clinically detectable, patients have complaints 3=moderate 4=severe

0. Normal
1. Only signs
2. Moderate nystagmus and / or moderate ocular dyskinesia and / or other mild cranial nerve disorders
3. Severe nystagmus and/or severe ocular dyskinesia and/or other moderate cranial nerve disorders
4. Severe enough dysfunction and / or other severe cranial nerve disorders
5. Can't swallow or talk

（3）Cone beam function

Reflection	L	R	note
Biceps reflex			0 = disappear 1 = weaken 2=Normal 3=accentuation 4=Do not continue to fight 5=Continuous clonus
Triceps reflex			
Periosteal reflex			
Knee reflex			
Ankle reflex			
Plantar reflex			0=bend， 1=neutral， 2=stretch
Abdominal reflex			0= normal, 1 = weakened, 2 = disappeared
Palmormental reflex			0 = none, 1 = yes
Muscle strength	L	R	note
Shoulder			0 = no muscle contraction 1 = visible muscle contraction, but can't drive joint activity 2 = horizontal movement, but can't resist gravity 3 = can resist gravity, but can't resist resistance (can lift out of bed) 4 = can resist resistance but limited 5 = normal The score of the weakest muscle strength in each group was the improvement score.
Elbow flexor			
Elbow extensor			
Flexors of hand/finger			
Extensors of hand/finger			
Flexor hip muscle			
Flexor muscles			
Extensor muscles of knee			
Flexors of foot/toe			
Extensor foot/toe muscle			
Functional testing	L	R	note
Upper limb paralysis test			0= Limbless decline， 1 = slight decrease, 2 = significant decrease

Lower limb paralysis test			0 = no drop, 1 = slight drop, 2 = significant drop, 3 = only one leg can be lifted at a time, 4 = no leg can be lifted.
Heel walking			0 = normal, 1 = restricted, 2 = impossible
Walking on tiptoe			
Voronin hop			0 = normal, 1 = 6 - 10, 2 = 1 - 5, 3 = impossible
Muscular tension	L	R	note
Upper limb			0 = normal 1 = mild: slight increase in muscle tension 2 = moderate: moderate increase in muscle tension, can be overcome, unlimited movement 3 = severe: severe increase in muscle tension, difficult to overcome, restricted movement 4 = contracture
The legs			
Gait			0 = normal 1 = almost imperceptible 2 = obvious: slightly affecting function 3 = continuous swing, seriously affecting function

Normal
Only physical signs
2. Mild disability: patients complained of fatigue weakness and/or group 1-2 muscle strength grade 4
3. Mild to moderate quadriplegia or hemiplegia; muscle strength grade 4 above group 2; or muscle strength grade 3 of group 1-2 (against gravity); or severe single paralysis (muscle strength less than grade 2 in group 1).
4. Severe quadriplegia or hemiplegia: 2 limbs of grade 2; or (> 3 limbs of grade 3); or a limb of grade 0-1 of grade
5. All lower limbs of muscle of grade 0-1; or (> 3 limbs of muscle strength (< 2); and/or hemiplegia of grade
6. Quadriplegia: all limbs of muscle strength of grade 0-1

（4）Cerebellar function

Ataxia		Fraction	note
Ataxia of limbs	Tremor		0 = No 1 = only signs 2 = mild: tremor or clumsiness, mild dysfunction 3 = moderate: tremor or clumsiness affects function in all respects 4 = severe: most severe dysfunction
	Poor distance discrimination		
	Rapid rotation		
Trunk ataxia			0 = none; 1 = only physical signs; 2 = mild: swing when closing eyes; 3 = moderate: swing when opening eyes; 4 = severe: unable to sit alone without assistance

Head tremor		0 = normal; 1 = mild; 2 = moderate; 3 = severe
Ataxia gait		0 = no; 1 = only physical signs; 2 = mild: patients have complaints; 3 = moderate: daily sitting or walking balance disorder; 4 = severe: can only walk a few steps or because of ataxia need to walk with help.
Straight line		0 = No problem, 1 = constrained, 2 = incomplete
Romberg sign		0 = normal; 1 = mild: mild eye closure instability ;2 = moderate: eye closure instability; 3 = severe: eye opening instability

0.Normal
1.Abnormal signs, no disability
2. Mild ataxia and/or moderate Romberg positive
3. Moderate trunk ataxia and/or moderate limb ataxia and/or moderate, severe trunk/gait ataxia
4. Severe gait/trunk ataxia and 3-4 limb weight ataxia
5. Unable to coordinate movement due to ataxia
X Impairment of pyramidal tract (limb muscle strength < grade 3) interferes with cerebellar function testing
FS=
Note: Only severe trunk/gait ataxia, FS = 3；
When evaluating cerebellar function due to limb weakness, add "X" after the actual score.

（5）Sensory function

Sensory system examination		L	R	note
Shallow sensation	Superficial sensation of upper limbs			0 = normal; 1 = only physical signs, no complaint 2 = Mild: Patient perceives mild tactile or pain disorders, but still distinguishes between sharp and dull 3 = Medium: Differentiating between sharp and blunt disorders 4 = Severity: unable to distinguish between sharp/blunt and/or light touch 5 = Total sensory loss
	Superficial sensation of trunk			
	Superficial sensation of lower limbs			
Vibration perception	Vibrational sensation of upper limbs			0= normal 1 = Mild: greater than 10s but less than the examiner 2 = Medium: greater than 2S but less than 10s 3 = Severity: total loss
	Vibrational sensation of lower limbs			
Sense of location	Upper limb position perception			0= normal； 1 = Mild: 1-2 wrong answers during examination, involving only distal joints 2 = Moderate: Many finger and toe movements are imperceptible and involve proximal joints. 3 = Severity: No movement at all, no standing.
	Lower extremity position perception			
Paresthesia	Abnormal sensation of upper limbs			0=no 1=yes
	Trunk			

	paresthesia			
	Lower extremity paresthesia			
Lhermitte				0=no；1=yes

normal
1.1-2 limbs with mild sense of vibration or pattern or hypothermia
2.The sensation of touch, pain or position decreased slightly or the sensation of vibration decreased moderately in 1-2 limbs.
3-4 limbs with slight decrease in single sense of vibration or temperature or figure
3. The sensation of touch, pain or position decreased moderately or vibration disappeared in 1-2 limbs.
Mild hypoesthesia of all sensations in 3-4 limbs
4. 1-2 severely impaired sensation of touch or pain;
> Moderate tactile, hypoalgesia and/or proprioceptive hypoesthesia in two limbs
5. 1-2 limb sensory loss；
Most of the body below the head has moderate impairment of touch, pain and/or proprioception
The sensation below the head basically disappeared.
FS=

（6）Bladder and rectal function

Bladder and rectum function	score	note
Urinary Waiting/Urinary Retention		0=no; 1=Mild: No significant impact on life 2=Moderate: urinary retention, frequent urinary tract infections 3=Severe: Need for catheterization 4=Loss of function
Urinary urgency/incontinence		0=no; 1=Mild: no significant impact on life; 2=Moderate: Rare urinary incontinence, once a week, with a pad 3=Severe: Frequent urinary incontinence occurs several times a week to several times a day, with a urine bag or pad 4=Loss of function and uncontrolled bladder
Catheterization		0=no； 1=Intermittent self-catheterization； 2=Continuous catheterization
Rectal dysfunction		0=no； 1=Mild: no fecal incontinence, no significant impact on life, mild constipation; 2=Medium: Must have a pad or be near the toilet. 3=Severe: Need enema or other means of cleaning rectum 4=Complete loss of function

normal
Mild urinary waiting, urgency and/or constipation
Moderate urinary waiting and/or urgency and/or rare urinary incontinence and/or severe constipation
Frequent urinary incontinence or intermittent self-catheterization; need manual cleaning of intestinal tract
Need for continuous catheterization
5. Loss of bladder or rectal function; need for catheterization
6. Loss of rectal bladder function
FS=

FS score of rectum and bladder	6	5	4	3	2	1
Rectal bystander FS score after conversion	5	4	3	3	2	1

（7）Brain Function Examination

Project	Score	Note
Depression and euphoria		0=no; 1=yes
Cognitive impairment		0=no; 1=There were only physical signs, but the patients and their families were not aware of them. 2=Mild: Patients and their families report mild cognitive changes (such as decreased ability to quickly associate and analyze complex problems; limited ability to quickly judge in specific emergencies; competent for daily work, but unable to withstand additional stress; intermittent symptoms at daily stress levels, occasional behavioral decline; and The tendency of fatigued children to forget things) 3=Moderate: Simple cognitive impairment test had definite abnormalities, but the orientation of time, place and character was normal. 4=Severity: There are 1 or 2 abnormalities in time, place and character orientation, and life is affected. 5=Dementia, confusion and/or loss of orientation
Fatigue		0=no; 1=Mild: Daily activities are unaffected 2=Moderate: Daily activities affected less than 50%. 3=Severe: Severe impact on daily activities (>50%)

normal
Only emotional changes (depression and/or euphoria) and/or mild fatigue and/or cognitive impairment
2. were found, but no complaint was made.
3 .2 = mild cognitive impairment; moderate or moderate fatigue
4. Moderate cognitive impairment
5. Severe cognitive impairment
6. dementia
FS=
Note: Fatigue is not usually counted in FS.

Only depression and/or euphoria, the brain FS = 1 point, at this time, does not count into EDSS; if the brain FS = 1 point is due to other reasons, then count into EDSS.

（8）Action ability

Walking without assistance or rest ≥500m(<4.0) ≥300m(4.5) ≥200m(5.0) ≥100m(5.5) 6.0 Unilateral or bilateral assisted walking (cane, crutch) > 100m; or unilateral assisted walking > 50m 6.5 Bilateral assisted walking (cane, crutch) > 20m without rest; or unilateral assisted walking < 50m 7.0 assisted still can not walk 5 meters, action is basically limited to wheelchair, can rely on wheelchair action; every day in the wheelchair for about 12 hours 7.5 It is still unable to walk with the assistance of wheelchair. Movements are basically confined to wheelchair. Shaking wheelchair and transferring body need help. 8.0 Actions are basically confined to beds, chairs or wheelchairs, but most of the time is spent under the bed every day; a lot of self-care ability is retained; both arms are basically functional. 8.5 Daily activities are basically limited to trauma; both arms have partial function; and some self-care ability is retained. 9.0 Patients are bedridden, able to communicate and eat 9.5 Patients are bedridden and can hardly communicate effectively or eat or swallow. 10 died in MS

note：
EDSS: 0-1.5: Patients should walk normally.
EDSS: > 2 points: Patients walk independently (> 500 meters), but their mobility is limited. At this time, the EDSS score is only determined by the FS score, and the FS of pyramidal tract function or vesicle function should be (> 2 points).
EDSS: > 4 points; patients walk < 500 meters independently
EDSS: > 6 points: patients need help walking
Another person's assisted walking is equivalent to bilateral assisted walking.

EDSS score：
0.0 All FS are 0.
1.0 1 item FS is 1
1.5 > 1 item FS is 1
2.0 1 FS is 2, the rest is 0 or 1
2.5 2 FS is 2, the rest is 0 or 1.
3.0 1 FS is 3, the rest is 0 or 1, 3-4 FS is 2, the rest is 0 or 1.
3.5 1 items FS were 3,1-2 items FS 2, the rest FS were 0 or 1; 2 items FS were 3, the rest were 0 or 1; 5 items FS were 2, and the rest were 0 or 1.
4.0 Walk independently (> 500 meters), 1 FS is 4, and the rest is 0 or 1.
4.5 Walking independently (> 300 meters), 1 FS is 4
5.0 Walking independently (> 200 meters) and (> 1 FS) 5
5.5 Walking independently (> 100m)
6.0-10.0; see Action Capability

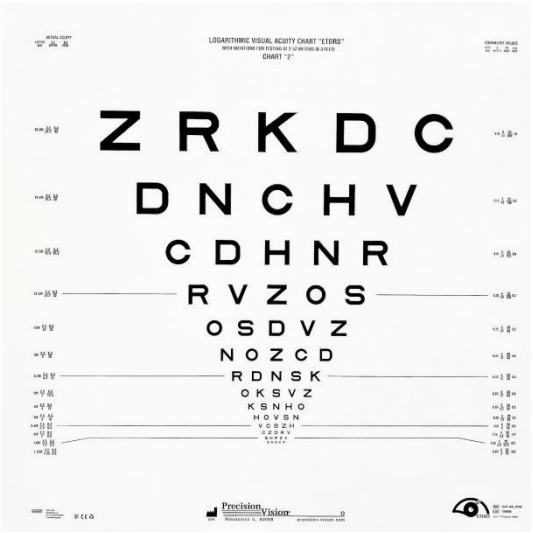
note：
EDSS < 4: Walking independently (> 500 meters), the exact score is determined by FS.
EDSS 4.0-5.0: Accurate scoring is determined by FS and walking range, and the more serious parameters are the scoring results of huge Eastern EDSS.
EDSS 5.5-8.0: Completely determined by operational capability

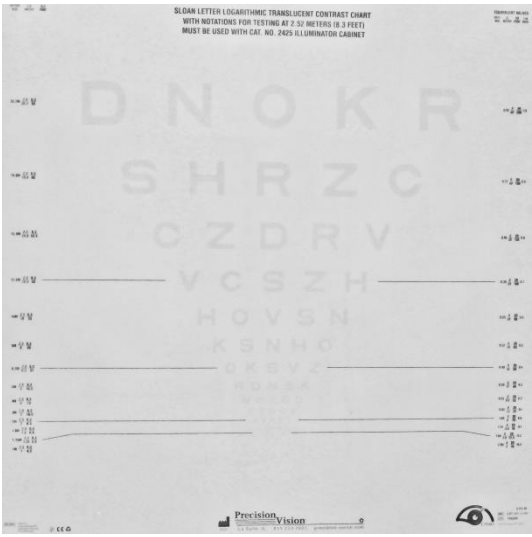
EDSS0-4.0; EDSS score should not change to 1 unless the FS score changes by 1.
EDSS score should be lower than any single FS score, except visual acuity and bladder and rectal function.

APPENDIX 4. VISUAL ACUITY TEST

The following charts are used under 160 cd/m^2 luminance level at 2.52 m chart distance (Precision Vision, Inc.).

ETDRS Chart **2.5% Low Contrast ETDRS Chart**





APPENDIX 5. CLINICAL LARORATORY TESTS

Chemistry Panal sodium Potassium calcium chloride Sodium bicarbonate Blood urea nitrogen Creatinine Uric acid glucose Alkaline phosphatase Alanine aminotransferase (ALT) Aspartate aminotransferase (AST) Total bilirubin Direct bilirubin Indirect bilirubin albumin Total protein Total cholesterol Triglyceride Low-density lipoprotein High density lipoprotein **Complete blood count (CBC)** White blood cell count (WBC) White blood cell identification Red blood cell count (RBC) RBC mean corpuscular volume (MCV) RBC distribution width Hemoglobin Platelet count Monocyte ratio Neutrophil ratio Lymphocyte ratio Neutrophil ratio Eosinophil ratio	**Human Chorionic Gonadotropin (β-HCG)** **Auto Aquaporin 4 antibody (AQP4-IgG)** **Flow cytometry** B cell ($CD19^{pos}$) B cell subsets Naive B Memory B Double Positive B Double Negative B Antibody Secrting Cells

APPENDIX 6. COMMON TERMINOLOGY CRITERIA FOR ADVERSE EVENTS (CTCAE 5.0)

NCI CTCAE=National Cancer Institute Common Terminology Criteria for Adverse Events.
Note: Based on the NCI CTCAE (v5.0), which can be found at:
https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_5.0/ or
https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf

INVESTIGATOR'S AGREEMENT

PROTOCOL TITLE: A RANDOMIZED, CONTROLLED, OPEN-LABEL, MULTI-CENTER TRIAL TO COMPARE THE SAFETY AND EFFICACY OF TOCILIZUMAB VERSUS AZATHIOPRINE IN PATIENTS WITH HIGHLY RELAPSING NEUROMYELITIS OPTICA SPECTRUM DISORDER (NMOSD): A HEAD-TO-HEAD COMPARATIVE STUDY

PROTOCOL NUMBER: 2017kylc005

I have received and read the Investigator's Brochure for the trial. I have read the 2017kylc005 study protocol and agree to conduct the study as outlined. I agree to maintain the confidentiality of all information received or developed in connection with this protocol. I agree to conduct the trial in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with International Conference on Harmonisation (ICH)/Good Clinical Practice (GCP) and applicable regulatory requirements.

Printed Name of Investigator

_____

Signature of Investigator

_____

Date

_____

SYNOPSIS

Supported by:
Grants from Tianjin Medical University Clinical Research Project (2017kylc005); the National Key Research and Development Program of China (2018YFC1312200); the Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing; National Science Foundation of China (91642205, 81830038 and 81601019)

The Investigational Medicinal Products (IMPs):
Tocilizumab and Azathioprine

Title of Trial:
A Randomized, Controlled, Open-label, Multi-Center Trial to Compare the Safety and Efficacy of Tocilizumab versus Azathioprine in Patients with Highly Relapsing Neuromyelitis Optica Spectrum Disorder (NMOSD): a Head-to-Head Comparative Study

Abbreviated Trial Name:
TANGO

Trial Rationale:
NMOSD is a severe, relapsing immune-mediated inflammatory disorder of the central nervous system (CNS), characterized by attacks of optic neuritis and longitudinally extensive transverse myelitis. The typical marker of NMOSD is the presence of serum autoantibody, termed NMO-IgG, directly against the extracellular domain of the water channel protein aquaporin-4 (AQP4) expressed on astrocytic endfeet. The AQP4-IgG is pathological and may cause astrocytic toxicity, thus rapid progression of disability. Relapse prevention is of paramount importance to reduce the risk of disability accumulation. Currently there are no approved therapies for the treatment of NMOSD. Azathioprine, as a first-line immunosuppressant in many other autoimmune diseases, is also widely used for the prevention of relapse in NMOSD patients. However, the efficacy and safety of azathioprine vary in different individuals. Though it is administered with oral corticosteroids, many patients with NMOSD still relapsed during long-term follow-ups. Recently, interleukin-6 (IL-6) is presumed to be critical in the pathogenesis of NMOSD because it is significantly elevated in the serum and cerebrospinal fluid (CSF) of NMOSD patients and promotes AQP4-IgG production by plasmablasts. Tocilizumab, a humanized IL-6 receptor monoclonal antibody, may provide therapeutic benefits through inhibition of IL-6 signaling pathway. In previous case series studies, tocilizumab reduced annualized relapse rate (ARR) and disability in highly active NMOSD patients. In this way, TANGO trial is intended to compare the safety and efficacy of tocilizumab with azathioprine in patients with highly relapsing NMOSD.

Trial Centers:
5 centers in North China and 1 center in South China

Investigators:
A list containing all Investigators will be provided when all the patients have been screened.

Study period:
Estimated date of the first patient enrolled: Nov, 2017
Estimated date of the last patient completed: Sep, 2019
The mean follow-up period was set 60 weeks.

Objectives:

Primary outcome measures:

- Time to first relapse from the baseline in a time-to-event analysis

Secondary outcomes measure:

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

Exploratory efficacy measures including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks

- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

Methodology:
This is an Investigator-initiated, randomized, parallel-group, multi-center, open-label, time-to-event trial to compare the safety and efficacy of tocilizumab with azathioprine in patients with highly relapsing NMOSD. Eligible patients will be randomized 1:1 to one of two parallel treatment arms: 1) tocilizumab or 2) azathioprine. Patients will receive tocilizumab at the dose of 8 mg/kg every 4 weeks or azathioprine 2-3 mg/kg every day. Tocilizumab will be administered intravenously over an approximate 60 minutes duration. Adjustment of the infusion rate and symptomatic treatments (prednisone and diphenhydramine) are permitted to manage infusion-related reactions. Azathioprine will be initiated at 25 mg/d orally and increased stepwise by 25 mg/d increments until the target dose is reached. Patients experiencing medication-related side effects during the loading period are allowed symptomatic treatments, with the exception of any new immunosuppressants. Patients receive their final, stable dosage of azathioprine, as daily maintenance until relapse, discontinuation, or the end of the trial.

Patients randomized to azathioprine receive concomitant immunosuppressants (oral corticosteroids, mycophenolate mofetil, cyclophosphamide, or methotrexate) for their initial 24 weeks of treatment based on the following schema: 1) patients without prior azathioprine treatment receive 24 weeks of concomitant immunosuppressants, 2) patients receiving azathioprine for < 24 weeks before randomization receive supplementary immunosuppressants until they reach 24 weeks of azathioprine therapy, 3) patients receiving azathioprine for ≥ 24 weeks before randomization receive no concomitant immunosuppressants. All azathioprine patients continue medication as monotherapy after 24 weeks of combined treatment. Patients in the tocilizumab group receive concomitant immunosuppressants for the first 12 weeks, and thereafter tocilizumab is used as monotherapy. The total duration of therapy is 60 weeks since randomization.

Based on the estimated baseline demographics and high relapse rate before the trial, the study is designed to continue until a mean follow-up of 60-week in the overall trial arrives. The overall trial duration is estimated to take approximately 2 years including enrollment. In this time-to-event trial, the trial duration for an individual patient will vary depending on when the patient enters the trial and on the patient's outcome. The course of the trial for an individual patient will consist of **Screening Period**, **Study Period**, and **Safety Follow-up Period**. The end of the study visit for an individual patient will take place when one of the following conditions is met, whichever comes first: (a) the patient experiences a definite relapse and early termination; or (b) when the last participating patient completes the last scheduled visit or when the Expert Panel decide to discontinue the study or development program. Patients were encouraged to receive sufficient adjusted rituximab or other therapeutic regimen (tocilizumab if possible) after completion of the end of study visit.

Screening Period (1-4 weeks)
At the Screening Visit, after informed consent form (ICF) is obtained from the patient, the patient will be screened for trial eligibility through medical history review, demographic data, and laboratory assessments. The medical history review will include confirmation of the diagnosis of

NMOSD. Relapses within the last 2 years prior to screening must be assessed by the Investigator to determine if they meet the inclusive criteria as specified by this protocol. Detailed information related to relapses within the 2 years prior to the Screening Visit will also be collected. This includes date of onset, clinical presentation for each relapse (e.g., optic neuritis [ON], transverse myelitis [TM]/longitudinally extensive transverse myelitis [LETM], acute brainstem syndrome, acute diencephalic clinical syndrome, symptomatic cerebral syndrome), and Expanded Disability Status Scale (EDSS) score at the baseline. For each historical relapse, the treatment for acute attacks and the regimen to prevent relapses will also be collected and recorded, including the dosage and the course of corticosteroids, the immunosuppressants, intravenous immunoglobulin (IVIG), plasma exchange (PE) or high-dose methylprednisolone (HDMP) treatment. The interval days between the occurrence of the last relapse and randomization will be calculated and recorded.

The baseline immunosuppressants for relapse prevention are allowed during the first 12 weeks for tocilizumab or 24 weeks for azathioprine of the trial under certain restrictions. No new immunosuppressants are permitted during the trial unless the patient experiences a relapse.

Immunosuppressants for the purpose of relapse prevention or treatment of a relapse prior to the screening visit and all other medications for symptomatic treatment within 30 days of screening will be reviewed and recorded on the case report form (CRF). Patients who experience a relapse during the Screening Period will be considered a screening failure. Such patients have the chance to be re-screened after receiving treatment for the relapse and when the patient is medically stable.

Deficient thiopurine methyltransferase (TPMT) activity may suggest an increase in the risk of azathioprine-related side effects.To reduce the risk of dropout for the patients in the azathioprine group, TPMT genetic polymorphisms including four variant TPMT alleles TPMT*2 (G238C), TPMT*3A (A719G/G460A), TPMT*3B (G460A), and TPMT*3C (A719G) for all the patients will be detected. Presence of TPMT mutation (homozygous or heterozygous) will exclude the patients from the trial.

Study Period (5 weeks - 60 weeks and later)
Randomization:
All patients who are cleared for randomization by the Investigator will be randomized on Day 1 on a 1:1 basis to the Tocilizumab Arm or the Azathioprine Arm. This is a completely randomized grouping experiment. The random numbers produced by a binomial distribution determine the group distribution of patients. The randomization will be across centers. Patients will receive either tocilizumab or azathioprine according to the randomization assignment.

The treatment duration for an individual patient varies in this time-to-event trial. All patients must remain on randomized treatment assignment until the end of study or early termination due to special reasons.

Relapse Evaluation:
Identification of potential relapse is critical for patient disability and for the integrity of the trial. Patients will receive a Patient Education Brochure that describes all the signs and symptoms of a definite attacks of NMOSD and the Package Inserts of the IMPs. The ICF contains instructions to contact the study site and the Investigator or the Investigator at the first sign or symptom of a potential relapse. The Investigator should review, in detail, this information and any additional warning signs of a relapse specific to that patient's clinical picture at each visit. Patients should be evaluated within 24 hours for a possible relapse by the Investigator. All reports of possible relapses and actions taken for dealing with a possible relapse must be recorded.

As part of the Relapse Evaluation, a blinded EDSS rater will perform the Kurtzke neurological assessment to determine the Functional System Scores (FSS) and EDSS score, and also finish the

assessments of low-contrast letter scores measured with retro-illuminated 2.5% Sloan letter chart and best corrected high-contrast logMAR visual acuity measured using retro-illuminated Early Treatment Diabetic Retinopathy Study (ETDRS) chart at 2.52 m, respectively.

High resolution SD-OCT images are acquired using identical protocols across centers. The technicians who conduct SD-OCT are unaware of the trial-group assignment. VEP images are also recorded by neuroelectrophysiologists who are also blinded to the trial-group assignment.

MRI is obtained on 3T scanner including enhancement with identical scanning protocols to identify new lesions associated with relapses. MRI scans are analyzed independently at a central MRI reading center (School of Radiology, Tianjin Medical University) by staff members who are also unaware of the trial-group assignment.

Blood will be collected for measurements of serum AQP4-IgG and B cell subsets by the blinded technician. If applicable, CSF samples will also be collected.

An Expert Panel is composed by the Principle Investigators (PI) across the centers. The Investigator will submit the medical record of the potential relapse in the patients and the Expert Panel will make the decision as to whether the clinical signs, symptoms and the change meet the definition of a relapse. The Expert Panel will decide whether each possible relapse meets the pre-defined objective criteria for a relapse. If the event is confirmed as a relapse, the patient must receive immediate treatments, according to the recommended standardized acute relapse treatment regimen.

Follow-up Relapse Evaluation Visits will be performed every 4 weeks after the onset of relapse. Additional unscheduled Follow-Up Relapse Evaluation Visits outside the specified time points are permitted at the discretion of the Investigator. As this trial is a time-to-event trial, patients who experience a relapse will be discontinued from this trial after completion of the week-24 Relapse Evaluation Visit, which also serves as the end of study visit. That is, patients who has a relapse will have a prolonged 24-week follow up.

Patients who complete this trial, either because of a relapse or because the trial is ended when a mean 60-week follow up is achieved, may continue to receive tocilizumab or azathioprine, if they had no relapses.

Safety Follow-up Period (12 weeks)

If a patient withdraws from the trial at any time after receiving any amount of tocilizumab or azathioprine, or does not wish to receive tocilizumab or azathioprine after completion of this trial, the patient will be required to complete an early termination or end of study visit at the time of withdrawal and a further follow-up visit 12 weeks after the last dose of tocilizumab or azathioprine for safety measures.

If a patient is discontinued due to an adverse event (AE), the event will be followed until it is resolved or in the opinion of the Investigator is determined medically stable.

Number of Patients (planned):

This is a randomized open-label parallel-group study, to evaluate the safety and efficacy of tocilizumab and azathioprine in patients with highly relapsing NMOSD. As such, the study is based on observing relapse events. With 1:1 randomization for the trial groups, we calculated that 118 patients would provide a power of 80% to determine the pre-specified between-group difference on the basis of a two-sided log-rank test at a 5% level of significance, assuming a 10% dropout rate.

Primary Efficacy Endpoint

The primary efficacy endpoint is time to first relapse from the baseline in a time-to-event analysis.

The trial will be considered to have met its primary efficacy objective if a statistically significant difference is observed between the tocilizumab treatment group and the azathioprine group. The comparison of the treatment groups for the primary endpoint will use a log-rank test. The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model.

Secondary Efficacy Endpoints

During the Study Period, Baseline is defined as the last available assessment prior to tocilizumab or azathioprine treatment for all patients regardless of their treatment group.

The comparison of the treatment groups for secondary efficacy endpoint - 12 weeks confirmed disability progression and serum AQP4-IgG titres will use a log-rank test.

Diagnosis and main criteria for inclusion:

Inclusion Criteria:

1. Male or female patients ≥ 18 years old
2. Diagnosis of NMOSD as defined by 2015 Criteria by Wingerchuk et al.
3. Historical Relapse (as defined by this protocol) of at least 2 relapses in the last 12 months or 3 relapses in the last 24 months, with at least 1 relapse in the 12 months prior to the screening
4. Able and willing to give written informed consent and comply with the requirements of the study protocol.
5. EDSS ≤ 7.5
6. Men and women of reproductive potential must agree to use a highly effective method of birth control from screening to 6 months after final dose of the experimental product.

Exclusion Criteria:

1. Current evidence or known history of clinically significant infection (Herpes simplex virus, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human immunodeficiency virus, Hepatitis viruses, Syphilis, et al)
2. Pregnant, breastfeeding, or child-bearing potential during the course of the study
3. Patients will not participate in any other clinical therapeutic study or will not have participated in any other experimental treatment study within 30 days of screening
4. Participation in another interventional trial within the last 3 months
5. Heart or kidney insufficiency
6. Tumor disease currently or within last 5 years
7. Clinically relevant liver, kidney or bone marrow function disorder
8. Receipt of rituximab or any experimental B-cell depleting agent within 6 months prior screening and the frequency of CD 19^+ B cells in peripheral blood mononuclear cells, as measured with flow cytometry, exceeded 1%.
9. Prior relapses or toxic intolerance on azathioprine treatment (adequate dosage and follow-up period observation)

- **Criteria for evaluation:**
- **Efficacy:**
- Duration of treatment commences with the randomization. The study period defines the time
- period for assessment of the trial endpoints. A 25% difference in relapse-free probability between the two groups is to be observed. Patients will continue to receive tocilizumab or azathioprine concomitant with corticosteroids or other immunosuppressants from this trial after completion of study. When the trial is stopped, all data from all patients will be collected, and the database cleaned, locked, and analyzed. Data from the study period will be used for efficacy analysis.
- Relapses will be monitored closely throughout the trial and evaluated.
- Disability will be assessed by the EDSS scores comparing the change from the baseline.
- The blinded EDSS Rater will perform the Kurtzke neurological assessment to determine the Functional System Scores (FSS) and EDSS score, and also finish the assessments of low-contrast letter scores measured with retro-illuminated 2.5% Sloan letter chart and best

corrected high-contrast logMAR visual acuity measured using retro-illuminated Early Treatment Diabetic Retinopathy Study (ETDRS) chart at 2.52 m, respectively.

- Serum AQP4-IgG titers will be performed by the blinded technician at the protocol specified time points.
- RNFL thickness and GCC volume by SD-OCT and P100 latency and amplitude in VEP at specified time points will also be collected by the blinded technician.
- MRI including enhancement at relapses will be conducted. MRI scans will be analyzed independently at a central MRI reading center by staff members who are also unaware of the trial-group assignment.
- Peripheral B cell subsets will be analyzed by the blinded flow cytometry technician at the protocol specified time points.

Safety:

The safety of tocilizumab or azathioprine will be assessed based on the all treatment-emergent adverse events (TEAEs), Common Terminology Criteria for Adverse Events (CTCAEs 5.0), serious adverse events (SAEs), and the changes from baseline through trial completion in vital signs, electrocardiogram (ECG), routine clinical laboratory tests, and pregnancy tests for female patients of childbearing potential.

Statistical methods:

Analyses will be produced for the study period in order to compare the tocilizumab group with the azathioprine group. The analyses will include efficacy and safety.

Efficacy:

Efficacy analyses will be performed on the Full Analysis Set (FAS) population as well as on the Per-Protocol (PP) population.

FAS Population:

The population on which primary, secondary and additional exploratory efficacy analyses will be performed. This set consists immunosuppressants of all patients who are randomized to treatment and who have received at least 1 dose of drug. Patients will be compared for efficacy according to the treatment they were randomized to receive, irrespective of the treatment they actually received.

PP Population:

The per-protocol population is a subset of the full analysis set population, excluding patients with major protocol deviations. The PP population will include all patients who:

- Have no major protocol deviations or key inclusion/exclusion criteria deviations that might potentially affect efficacy
- Patients who took at least 80% of the required treatment doses while they were in the study period
- Patients who received tocilizumab or azathioprine as monotherapy during the study period

Primary Efficacy Analysis for the Study Period:

Note: During the study period, baseline is defined as the last available assessment prior to tocilizumab or azathioprine treatment for all patients regardless of treatment group.

The primary efficacy endpoint is time to first relapse as defined in the protocol. The trial will be considered to have met its primary efficacy objective if a statistically significant difference is observed between the tocilizumab treatment group and the azathioprine group. The comparison of the treatment groups for the primary endpoint will use a log-rank test. Confidence intervals and p-values will be presented. Kaplan-Meier curves for both treatment groups will be produced. Hazard ratio and risk reduction will be summarized. A sensitivity analysis will be performed on time to first relapse (as identified by the Expert Panel) using a log-rank test including strata for the randomization stratification variables. In addition, a sensitivity comparison of the primary

endpoint will use a Cox proportional hazards regression model with treatment group indicator, and randomization stratification variables as the covariates in the model.

Secondary Efficacy Analysis for the Study Period:
The secondary efficacy analyses will use the available data from the study period.

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

Additional Exploratory Efficacy Analysis including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord MRI
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

The time to onset of confirmed disability progression (12-week confirmation [weeks]) is defined as the time from baseline to the onset of the first disability progression that is confirmed at the next regularly scheduled visit ≥ 12 weeks after the initial disability progression. Baseline for the time to onset of confirmed disability is the date of randomization. Disability progression is defined as an increase of ≥ 1.0 point from baseline EDSS score if the baseline EDSS value is ≤ 5.5 points (inclusive) or an increase of ≥ 0.5 points if the baseline EDSS value is > 5.5 points. Assessments within 30 days after a protocol-defined relapse will not be used for confirmation of confirmed disability progression. The non-confirmatory EDSS assessments (between the initial disability progression and the confirmation of disability progression should also fulfill the requirements of the progression.The comparison of the treatment groups for secondary efficacy endpoint--12 weeks confirmed disability progression will use a log-rank test including strata for the randomized intervention variable. The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model.

Cell based assay and Fluorescence immunoprecipitation assay will be used to dynamically evaluate the serum AQP4-IgG titers of patients in this study at the time of baseline, 12 weeks, 24 weeks, 36 weeks, 48 weeks and 60 weeks from the randomization. Wilcoxon rank sum test is used to compare the serum AQP4-IgG titers of two groups.

Other exploratory efficacy endpoints listed in the protocol will also be analyzed but will be reported in separate study reports or publications.

Safety:
Safety analyses include all patients who receive at least 1 dose of trial drug. Patients will be compared for safety according to the treatment they actually received. All adverse events and other safety information including untreated patients collected after the signing of informed consent will be reported in listings, as applicable.

Safety for the Study Period:

For the Study Period, adverse events will be summarized by incidence, system organ class (SOC), preferred term (PT), seriousness, severity, relationship to treatment, and by treatment group. Concomitant medications will be summarized by treatment group.

Changes from Baseline in vital signs and laboratory assessments (chemistry and hematology) will be summarized by treatment group. Likewise, shift tables (L [low], N [normal], H [high]) by treatment group will be produced for clinical laboratory tests and pregnancy tests will be summarized in patient listings.

TABLE OF CONTENTS

List of Tables

GLOSSARY OF ABBREVIATIONS

ADE	Adverse drug events
AE	Adverse event
ANC	Absolute neutrophil count
APRIL	A proliferation-inducing ligand
AQP4	Water channel protein aquaporin aquaporin-4
ARR	Annualized relapse rate
AZA	Azathioprine
BAFF	B-cell activating factor
BP	Blood pressure
CBC	Complete blood count
CDP	Confirmed disability progression
CFDA	China Food and Drug Administration
CNS	Central nervous system
CRF	Case report form
CSF	Cerebrospinal fluid
CTCAEs	Common Terminology Criteria for Adverse Events
Cz	Midline central
DMARDs	Disease-modifying anti-rheumatic drugs
DTI	Diffusion tensor imaging
ECG	Electrocardiogram
EDSS	Expanded Disability Status Scale
EFNS	European Federation of neurological Societies
EMA	European Medicines Agency
ETDRS	Early Treatment Diabetic Retinopathy Study
EULAR	European League Against Rheumatism
FAS	Full Analysis Set
fMRI	Resting-state functional MRI
FSS	Functional System Scores
FUs	Fluorescent units
Fz	Midline frontal
GCC	Ganglion cell complex
GCP	Good Clinical Practice
H	High
HEK293	Human embryonic kidney 293
HR	Heart rate
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
IEC	Independent Ethics Committee
IL-6	Interleukin-6
IPND	International Panel for NMOSD Diagnosis
IRB	Institutional Review Board
iTregs	Regulatory T cells
IV	Intravenous
IVIG	Intravenous immunoglobulin
IVMP	Intravenous methylprednisolone
L	Low
LCLA	Low-contrast letter acuity
LETM	Longitudinally extensive transverse myelitis
MedDRA	Medical Dictionary for Regulatory Activities

MMF	Mycophenolate mofetil
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MTX	Methotrexate
N	Normal
NMOSD	Neuromyelitis Optica Spectrum Disorder
ON	Optic neuritis
Oz	Occipital midline
PBs	Plasmablasts
PE	Plasma exchange
PH	Proportional hazards
PI	Principal Investigator
PP	Per-Protocol
PT	Preferred Term
RA	Rheumatic arthritis
RNFL	Retinal nerve fiber layer
RR	Respiration rate
RTX	Rituximab
SAEs	Serious adverse events
SAP	Statistical Analysis Plan
SD-OCT	Spectral-domain optical coherence tomography SD
sJIA	Systemic juvenile idiopathic arthritis
SOC	System organ class
SWI	Susceptibility weighted imaging
TCZ	Tocilizumab
TEAEs	Treatment-emergent adverse events
TGF	Transforming growth factor
TM	Transverse myelitis
TPMT	Thiopurine methyltransferase
VA	Visual acuity
VAS	Visual analogue scale
VEP	Visual evoked potentials
WBC	White blood cell
WHO Drug	World Health Organization Drug Dictionary

5. BACKGROUND

5.1 Neuromyelitis Optica Spectrum Disorder (NMOSD)

Neuromyelitis optica spectrum disorder (NMOSD), also known as neuromyelitis optica (NMO), is an autoimmune inflammatory disease of the central nervous system (CNS), which preferentially affects the optic nerves and spinal cord[1]. Though it was previously thought to be a subtype of multiple sclerosis, clinical features, neuroimaging, immunological, and histopathological characteristics distinguish NMOSD from classical MS since the presence of specific serum autoantibody aquaporin-4 (AQP4-IgG) was found in 2004[2]. More than 60%-80% of the patients with NMOSD have detectable AQP4-IgG (also known as NMO-IgG) in the serum. For the patients with negative AQP4-IgG in the serum, the diagnostic criteria have been refined in 2015 to increase the diagnostic accuracy to a large extent[3].

Unlike MS, most patients with NMOSD experience a severe relapsing course[4], thus risk of increased progression in disability makes NMOSD a devastating disease in terms of social and economic burden. NMOSD is a quite important cause of irreversible neurological disability in Chinese young adults. To date, there are few definite epidemic data worldwide for NMOSD patients[5]. But NMOSD is more common in certain regions of Asia, Africa and Latin America, where MS is relatively rare[6,7]. This is different from European countries. Similarly, till now there is no sufficient epidemic data of NMOSD in China. In some clinics specialized in NMOSD, the number of NMOSD patients is about 10 times as many as that of MS[8].

5.2 Clinical Features of NMOSD

The clinical hallmarks of NMOSD are acute optic neuritis and transverse myelitis that involves more than 3 vertebral levels, described as longitudinally extensive transverse myelitis (LETM). These clinical events can occur either simultaneously or in isolation. Signs and symptoms attributable to lesions beyond the optic nerves and spinal cord can also occur in patients with NMOSD and are reported in about 15% of patients. So, the revised diagnostic criteria were released by the International Panel for NMOSD Diagnosis (IPND) in 2015. NMOSD was stratified by AQP4-IgG serostatus, that is, AQP4-IgG-serosepositive NMOSD and AQP4-IgG-seronegative NMOSD (or unknown serostatus)[3]. In AQP4-IgG-seropositive NMOSD, if AQP4-IgG is reliably positive (cell-based assay is preferred) and alternative diagnoses are excluded, only core clinical characteristic (optic neuritis, acute myelitis or brain syndrome) is required for the diagnosis. Brain syndromes, such as area postrema syndrome manifested by intractable hiccup, nausea and vomiting, acute brainstem syndrome, symptomatic narcolepsy or acute diencephalic clinical syndrome with NMOSD-typical diencephalic MRI lesions and symptomatic cerebral syndrome with NMOSD-typical brain lesions. About 60%-80% of NMOSD patients are AQP4-IgG-seropositive. On the other hand, for the diagnosis of AQP4-IgG-seronegative NMOSD, more stringent criteria were set to exclude a variety of diseases mimicking NMOSD, although exclusion of alternative diagnoses was imperative (see *Appendix 1 and Appendix 2*).

Most frequently, NMOSD develops as a recurrent disorder, which predominantly affects women.The majority of patients suffer from a recurrent course (80 – 90%) whilst monophasic (10 – 20%) and primary or secondary progressive courses are rare. Clinical features are often severe involvements of optic nerve and/or spinal cord and also many other symptoms and signs. Acute to subacute onset of blindness in one or both eyes, preceded or followed by a severe transverse or ascending myelitis. The spinal cord lesions are often necrotizing, more likely to lead to permanent than those of typical demyelination in MS. If myelitis extends up into the brainstem causing respiratory failure and death. Uncommon features such as vomitting and hiccups reflect damage in the area postrema. If brainstem is involved, the patient may have vertigo, hearing loss, diplopia, ptosis or nystagmus. Hypothalamic involvement, could cause hypersomnolence, hypothermia or hyperthermia[9,10].

Once a relapsing course has been established in NMOSD, recurrent optic neuritis and myelitis attacks result in stepwise accumulation of neurological disability. Within 5 years, more than 50% of such

patients are functionally blind (visual acuity worse than 20/200) or have lost the ability to ambulate without assistance. In NMOSD patients, the disability accumulation is associated with relapses. Therefore, relapse prevention is paramount for successful treatment of relapsing NMOSD. As NMOSD has the potential to cause significant disability, the prognosis of relapsing NMOSD is poor. The 5-year mortality of NMOSD was reported to be 30%. 50% individuals sustain permanent severe disability, visual (blind in one or both eyes) or ambulatory (requiring a wheelchair). Most deaths result from neurogenic respiratory failure secondary to a high cervical cord or brainstem lesion. Frequent early relapses predict a poor prognosis[11].

5.3 Different Treatment Options in NMOSD

Acute NMOSD relapses are generally treated with high-dose intravenous (IV) steroids. Plasma exchange (PE) or intravenous immunoglobulin (IVIG) often used as a rescue therapy for those who do not respond. After acute phase, supportive treatments against relapse are necessary. But there is no therapy approved for the prevention of NMOSD. Of the off-label prescription of immunosuppressive agents, corticosteroids, azathioprine (AZA), mycophenolate mofetil (MMF) and rituximab (RTX) are commonly used for long-term prophylaxis[12-14]. However, a significant number of patients will continue to have attacks resulting in additional and permanent neurological deficits and disability when using these empirical drugs[15]. Empirically, AZA is the most frequently used to prevent relapses in NMOSD. However, a well-defined protocol for choosing the most optimal treatment is lacking[16]. Given the seriousness of the disease, limitations of currently available therapies, and the limited options for treatment, there remains a significant unmet medical need for an effective and safe treatment for NMOSD. There are no streamline disease modified drugs as MS in European or North American countries. Many patients with NMOSD in China received irregular treatments for prevention, or even received no preventive treatments. As a result, higher rate of relapses was common in these patients. To reduce relapse rate and improve the disability of NMOSD patients, clinical trials, especially that can provide high-level evidence, are urgent on comparing the efficacy and safety of the drugs available in China.

5.4 Efficacy and Tolerability of AZA for Preventing Relapses of NMOSD

A review published in 2016 and the guideline by the European Federation of neurological Societies (EFNS) issued in 2010 recommended AZA, a purine analog that blocks deoxyribonucleic acid synthesis particularly of rapidly proliferating B and T lymphocytes, as one of the first-line therapies for NMOSD[16,17]. However, these recommendations were only based on studies involving AZA with few comparator groups; and therefore, the reported measures of therapeutic effect may be remarkably biased. For instance, several NMOSD case series involving patients who took AZA had shown that post-treatment ARR was significantly reduced compared to pre-treatment ARR and there was a trend towards improved or stabilized EDSS scores at post-treatment compared to EDSS scores at pret-reatment[18,19]. However, these beneficial results of AZA should be brought to proper perspective since pre-treatment and post-treatment measurements are not independent of each other and these are affected by natural processes and patient characteristics. To control for these variables, between-group comparison must be performed so the effects of the intervention can be discerned. More studies suggested that AZA may be inferior to RTX in terms of reduction of ARR and reduction of EDSS while fewer other studies revealed no difference in terms of ARR and EDSS[20]. In the comparison between AZA and MMF groups, few studies did not find significant differences in ARR or and EDSS[21].

Adverse drug events (ADE) associated with AZA were seemingly frequent and may contribute to patient non-adherence to prescribed medication. In the pooled analyses of ADE, there was significant number of patients in the AZA group who had experienced any adverse event, including leukopenia, hair loss, elevated liver enzymes, hepatotoxicity or severe myelosuppression compared to MMF[22]. The increased occurrence of ADE in patients who took AZA may be due to possible genetic mutations at the TPMT*3C in those individuals which causes decreased levels of thiopurine methyltransferase (TPMT) leading to toxicity. Although the AZA treated individuals with TPMT*3C heterozygous or

homozygous genetic mutation may contribute to these adverse events, only 10% of the side effects may be elucidated by TPMT insufficiency[23].

Since AZA is often used in conjunction with corticosteroids, the efficacy of AZA as monotherapy, remains unclear. Due to the overall limited therapeutic evidence for AZA treatment, it is highly recommended that the implementation of well-conducted, randomized, controlled clinical trials is imperative to determine with certainty the role of AZA for patients with NMOSD.

5.5 Role of IL-6 Signaling Pathway in NMOSD

NMOSD is thought to be immunological distinct from MS. Abnormal levels of multiple cytokines in the serum and CSF have been reported in NMOSD patients[24]. Of the diversity of cytokines, only the IL-6 level showed significant elevation in NMOSD in serum analysis. The CSF level of IL-6 was significantly increased in NMOSD compared with MS and other neurological diseases. Importantly, the CSF IL-6 level had a significant correlation with the CSF glial fibrillary acidic protein (GFAP) level and CSF cells, and a weak correlation with AQP4-IgG titers[25]. The GFAP level is remarkably high in the CSF of AQP4-IgG-seropositive NMOSD patients during acute exacerbations whereas the CSF GFAP is not elevated at all in typical multiple sclerosis. CSF IL-6 levels are directly associated with CSF GFAP, reflecting IL-6-induced astrocytic damage. The amount of AQP4-IgG present in the CNS is associated with astrocyte injury and inflammatory responses during NMOSD attacks[26]. AQP4-IgG selectively induced IL-6 production by AQP4-positive astrocytes and that IL-6 signaling to endothelial cells decreases barrier function, increases chemokine production, and enhances leukocyte transmigration under flow. These effects were reversed after application of IL-6 neutralizing antibody[27]. In NMOSD, increased CSF matrix metalloproteinase-2, likely induced by interleukin-6 signaling, may disrupt the blood-brain barrier and enable serum AQP4-IgG migration into the CNS[28]. The crosstalk between IL-6 and AQP4-IgG plays important roles in pathogenesis of NMOSD.

Previous studies have proved that $CD19^{int}CD27^{high}CD38^{high}CD180^{-}$ plasmablasts (PBs), which are a subset of terminal stage of B cells, are capable of producing autoantibodies, including AQP4-IgG. The survival of PBs and the production of AQP4-IgG are both promoted by IL-6, but not B-cell activating factor (BAFF) and a proliferation-inducing ligand (APRIL)[29]. IL-6 was originally cloned as B-cell stimulatory factor-2. IL-6 activates a receptor complex consisting of the IL-6 receptor (IL-6R) and the signal-transducing receptor subunit gp130. IL-6 can bind to both a transmembrane form and a soluble form of IL-6R, which can interact with gp130 to trigger downstream signal transduction and gene expression. Apart from B cells, IL-6 exerts pleiotropic effects on other immune cells. Th1 cells can be differentiated from naïve T cells by IL-6 and transforming growth factor (TGF)-β, whereas IL-6 inhibits the differentiation into TGF-β induced regulatory T cells (iTregs). IL-6 dysregulation is involved in the development of many autoimmune diseases such as systemic lupus erythematosus (SLE), rheumatic arthritis (RA). In SLE, in vivo blockade of the IL-6 receptor by tocilizumab decreases lymphocyte activation and restores B and T cell homoeostasis by either blocking differentiation and/or trafficking in patients with SLE and leads to normalization of the abnormal B and T cell subsets seen at baseline[30]. Inhibition of IL-6 receptor IL-6 signal-blockade therapy may serve as a promising option to decrease disease activity in NMOSD.

5.6 Therapeutic Strategy of Tocilizumab (TCZ) as the First IL-6R Monoantibody Inhibitor

In rheumatic arthritis (RA), characterized by systemic and joint inflammation, TCZ showed protective efficacy against joint destruction and suppressed disease activity. TCZ is the sole biologic that, as monotherapy, shows greater efficacy than standard disease-modifying anti-rheumatic drugs (DMARDs) and methotrexate (MTX). TCZ add-on or switch to TCZ monotherapy was studied for patients with active RA despite MTX. According to the 2016 European League Against Rheumatism (EULAR) updated recommendations for the management of RA, IL-6 pathway inhibitors may have some advantages (i.e., safety) compared with other biological agents in patients who cannot use

conventional synthetic DMARDs as co-medications[31]. This provides favorable information of TCZ in the treatment of NMOSD. Based on the available safety data in the RA trials, the adverse effects of TCZ have
been shown to be manageable, reversible, and usually not treatment limiting. In addition, based on the RA safety data, it is justified to conduct a clinical study in adult NMOSD patients, using TCZ doses of 8 mg/kg every 4 weeks in hospitals.

Indeed, treatment with TCZ ameliorated AQP4-IgG production in the serum of NMOSD patients and suppressed relapse rate and neuropathic pain in patients with refractory NMOSD, suggesting its clinical efficacy[32-34]. In a retrospective observational study, TCZ significantly reduced EDSS, active MRI lesions and also serum AQP4-IgG titers. But attacks occurred when TCZ dosage was reduced[35]. Though TCZ may be a promising drug for preventing acute attacks in patients with NMOSD, the adjusted dosage and long-term efficacy and safety after TCZ is initiated warrant an investigation in large cohorts of patients with NMOSD.

6. STUDY RATIONAL

To date, there is no standard treatment for prevention of NMOSD relapses. Evidence for choice of drugs available still lacks. We aim to look for efficient regimen with high level of evidence. This study is a pivotal head-to-head clinical trial comparing the safety and efficacy of TCZ and AZA. All patients will be permitted to use immunosuppressive therapy for a washout period after randomization. Efficacy of monotherapy of TCZ or AZA afterwards will also be evaluated. Opinions vary widely among Investigators regarding ethics of placebo-controlled studies for maintenance treatment of NMOSD[36].Though US FDA strongly prefers a pivotal monotherapy, placebo-controlled arm, it is in direct contrast to rulings by the European Medicines Agency (EMA), China Food and Drug Administration (CFDA) and the viewpoints of academic and patient-advocacy group for such a devastating disease. AZA is a widely used therapy, even if it was not sufficiently supported by trial evidence and considered experimental rather than standard of care. To protect the patients from more relapses, AZA shows better efficacy than placebo. By combining the available data on first-line treatments AZA and RTX, risks of relapses could be calculated, with AZA 53% and RTX 18%[11]. As previous case series studies showed that TCZ reduced relapse rates of patients who are refractory to RTX treatment, we assume TCZ is no inferior or at least as effective as RTX.

6.1 Statement on the Use of AZA as a Comparator

NMOSD is a severe disabling disease. More than 85% of the patients will have a new relapse within one year if sufficient immunosuppressants is not maintained. To date, no FDA-approved disease-modifying treatments have shown effectiveness with long-term tolerance in NMOSD. AZA has been recommended as the first-line treatment in many countries. But its efficacy varies, especially in most cases it is used with concomitant corticosteroids. The maintenance corticosteroid therapy is of great importance at the beginning of NMOSD treatment, because of the gradual onset of the effects of AZA, which can take several months to exert its full activity. It was challenging to determine the solitary effects of AZA since this regimen was typically administered with oral corticosteroids. It should be noted that the sole administration of steroids may reduce relapses in NMOSD. Therefore, cautious interpretation of efficacy data should be performed due to variability of the interventions[37]. In addition, AZA is metabolized by the enzyme TPMT. Those who have low TPMT activity are more likely to experience side effects, including risk of irreversible myelosuppression. The mutation rate of TPMT in our population limited the bias against using AZA as a treatment modality. So we will detect TPMT gene single nucleotide polymorphisms at Screening Period to reduce the dropout rate for AZA during the trial (see *Appendix 3*). After AZA treatment, 54.5% patients reached a relapse-free state in at least two-year follow ups. But steroid-sparing effects of AZA therapy are limited. Overall, the safety issues of AZA may refine it to extensively usage. Its efficacy as monotherapy remained unclear. Head-to-head comparison studies are needed to define the patient groups that will profit most from AZA.

6.2 TCZ or AZA Dose for NMOSD

Empirical data from RA clinical trials support the now approved dose regimen of 8 mg/kg/4w. Baseline concomitant immunosuppressants would be discontinued at 12 weeks, when TCZ will be used as monotherapy completely. Because it may take at least 12 weeks for TCZ to reduce serum IgG1 level significantly compared to baseline in SLE patients receiving TCZ[30], we assume TCZ exert its efficacy at 12 weeks.

In the AZA group, patients will receive oral AZA 2-3 mg/kg every day as a maintenance. AZA will be given to patients while their concomitant corticosteroids or other immunosuppressants are being tapered down until withdrawn. Patients may start at 25 mg/day with the incremental dosages: 50 mg/day, 75 mg/day, 100 mg/day, 125 mg/day, 150 mg/day, to 2-3 mg/kg/day with good tolerability. They must remain on that dose for the duration of the study or until the patient experiences a relapse. It may take 3-6 months for AZA to become biologically active[17]. So patients in the AZA group took concomitant immunosuppressants for 24 weeks, during which period AZA achieved stable efficacy. Thereafter, AZA would be used as monotherapy. If the patients had used AZA for at least 24 weeks

before randomization, concomitant immunosuppressants would be discontinued at baseline.

6.3 Rationale for Early Concomitant Usage of Corticosteroids and Immunosuppressants

NMOSD is distinct from MS in the treatment of corticosteroids usage. In MS, disease modifying drugs will be given to the patients just after high-dose of corticosteroids treatments, with no tapering of corticosteroids. In NMOSD, initial or recurrent episodes are usually treated with high-dose intravenous methylprednisolone (1 g daily for three to five consecutive days). In many countries of the EU, the intravenous therapy with methylprednisolone is followed by an oral taper and needs to be performed slowly[16]. Some patients experience clinical worsening when prednisone is reduced below 5 -15 mg/day. If an immunosuppressant would be added, then corticosteroids will be used for rapid immunosuppression until the immunosuppressant exerts its full effect. Taking AZA as an example, as the treatment may only take full effect after 3 - 6 months, it should initially be combined with oral steroid therapy, as oral steroids have been beneficial to suppress disease activity in NMOSD[17].

7. TRIAL OUTCOMES

7.1 Primary Objective

The primary objective of this trial is to compare the safety and efficacy of TCZ treatment as compared to AZA in highly relapsing NMOSD patients based on time to first relapse and relapse risk reduction.

7.2 Secondary Objectives

The secondary objectives of this trial are as follows:

- Reduced risk reduction of confirmed disability progression (CDP) for at least 12 weeks by TCZ compared to AZA
- To determine whether TCZ or AZA could reduce serum AQP4-IgG titers from baseline to 60 weeks in AQP4-IgG positive patients
- To compare the overall safety and tolerability of TCZ to AZA

7.3 Efficacy Endpoints

7.3.1 Primary Efficacy Endpoint

Time to first relapse for TCZ compared with AZA.

7.3.2 Secondary Efficacy Endpoints (TCZ outcomes as compared with AZA outcomes):

The secondary efficacy analyses will use the available data from the study period.

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG from baseline to 60 weeks
- Overall safety and tolerability of TCZ or AZA

7.3.3 Additional Exploratory Efficacy Endpoints:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord MRI
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

8. OVERALL TRIAL DESIGN AND PLAN

This is an Investigator-initiated, randomized, controlled, parallel-group, multi-center, time-to-event trial to evaluate the safety and efficacy of TCZ as compared with AZA in patients with highly relapsing NMOSD. The study will consist of the following periods: Screening Period, Study Period and Safety Follow-up Period.

8.1 Screening Period

At the Screening Visit, after informed consent is obtained from the patient, the patient will be screened for trial eligibility through medical history review, demographic data, electrocardiogram (ECG) and laboratory assessments. The medical history review will include confirmation of the diagnosis of NMOSD. Detailed information on relapses within the last 2 years prior to screening must be assessed by the Investigator to determine if they meet the definition for historical relapses as specified by this protocol. Detailed information related to relapses within the last 2 years will be collected and recorded in the case report form (CRF), if available. This includes date of onset and its clinical presentation for each relapse (corresponding to 6 core symptoms defined by Wingerchuk et al;, 2015; see *Appendix 1* and *Appendix 2*); and EDSS score at the following time points: prior to relapse, at nadir and during recovery, for severity of relapse and recovery. Start/stop dates and dose regimen of all immunosuppressants and non-drug therapies taken for relapse prevention or treatment of a relapse will also be collected and recorded. If PE or IVIG was administered for treatment of a relapse, the number of cycles of PE or IVIG will also be collected. Information on all other previous historical relapses including relapse onset date and its clinical presentations and treatment received for the acute relapse and/or to prevent relapse will also be collected, if available. Only validated diagnostic tests performed by a qualified laboratory are acceptable.

TPMT genotyping will be performed in all the patients to detect mutant alleles (TPMT*2, TPMT*3A, TPMT*3B, and TPMT*3C) (see *Appendix 3*). If the patient has homozygous or heterozygous mutation in TPMT gene, then the patient will not be enrolled. If the patient had AZA-related toxicity which caused discontinuation or relapses when AZA was used, the patient will be considered a screening failure.

Supportive immunosuppressants for relapse prevention are allowed during the first 24 weeks in the AZA group or 12 weeks in the TCZ group. The following immunosuppressants are allowed either as mono-therapy or in combination such as corticosteroids, AZA, MMF, MTX, tacrolimus, cyclosporine and cyclophosphamide. No new immunosuppressants are permitted during the trial unless a patient experiences a relapse. Immunosuppressants and/or therapies for NMOSD relapses (either acute treatment or prevention) prior to screening and all other medications taken within 30 days of the Screening Visit will be reviewed and recorded on the CRF.

Patients who experience a relapse during the Screening Period will be considered a screening failure. Such patients may be re-screened for enrollment into the trial after receiving treatment for the relapse and when, in the opinion of the Investigator, the patient is medically stable. The patient must meet the enrollment criteria in order to enter the trial.

8.2 Study Period

All patients who are cleared for randomization by the Investigator will be randomized on Day 1 in a ratio of 1:1 to the TCZ Arm or the AZA Arm. A randomization worksheet will be provided by the School of Public Health of Tianjin Medical University who are to ensure proper randomization. The randomization will be across centers.

Patients will receive either TCZ or AZA according to the randomization assignment and the treatment described in Section 10. The treatment duration for an individual patient varies in this time-to-event trial. All patients must remain on randomized treatment assignment until the end of the Study Period visit.

All the patients we screen have never used TCZ before.

Patients randomized to azathioprine received concomitant immunosuppressants (oral corticosteroids, mycophenolate mofetil, cyclophosphamide, or methotrexate) for their initial 24 weeks of treatment based on the following schema: 1) patients without prior azathioprine treatment received 24 weeks of concomitant immunosuppressants, 2) patients receiving azathioprine for < 24 weeks before randomization received supplementary immunosuppressants until they had reached 24 weeks of azathioprine therapy, 3) patients receiving azathioprine for ≥ 24 weeks before randomization received no concomitant immunosuppressants. All azathioprine patients continued medication as monotherapy after 24 weeks of combined treatment. Patients in the tocilizumab group received concomitant immunosuppressants for the first 12 weeks, and thereafter tocilizumab was used as monotherapy.

The end of the study visit for an individual patient will take place when one of the following conditions is met, whichever comes first: (a) the patient experiences a definite relapse and early termination; or (b) when the last participating patient completes the last scheduled visit or when the Expert Panel decide to discontinue the study or development program. Identification of potential relapse is critical for patient safety and for the trial. Any potential relapse will be evaluated according to the information below.

Follow-up visits to monitor the course of the relapse and disability progression by EDSS will be performed at 4, 8, 12, 16, 20, 24 weeks after the onset of relapse. Unscheduled Follow-up Relapse Evaluation Visits are permitted and will be made at the discretion of the Investigator. All reports of possible relapses and actions taken for the possible relapse must be documented in the patient's medical chart or source documents and recorded in the CRF.

As this is a time-to-event trial, patients who experience a relapse will be discontinued from this trial after completion of the Week 24 Relapse Evaluation Visit. Thus, the Week 24 Relapse Evaluation Visit also serves as the end of the study visit for these patients. For patients who do not have relapses, the end of the study visit will be defined as when the last participating patient completes the last scheduled visit or when the Expert Panel decide to discontinue the study or development program. Patients who complete the trial either because of a relapse or because of the trial is ended may be encouraged receive RTX (or TCZ, if possible) treatment.

8.2.1 Relapse Evaluation

Patients will be educated and directed to contact the study site at the first sign or symptom of a potential relapse. Patients should be evaluated within 24 hours of notification of the Investigator of a possible relapse, and no later than 24 hours. All potential relapses must be evaluated by both the Investigator and the blinded EDSS Rater. All reports of possible relapses and actions taken for the possible relapse must be documented in the patient's CRF. At each Relapse Evaluation Visit, the blinded EDSS Rater will perform the Kurtzke neurological assessment and document the Functional System Scores (FSS) and the EDSS score, including 100% high contrast visual acuity and 2.5% low contrast letter acuity. The Investigator will perform a complete neurological examination. The blinded SD-OCT rater will perform SD-OCT was also performed for patients who had optic neuritis. The blinded VEP rater will perform VEP was also performed for patients who had optic neuritis. MRI images will be obtained according to this protocol. Blood will also be collected.

The Investigator determines if the clinical signs, symptoms and neurological change (objective findings on examination) meet the definition for a relapse as outlined in this protocol. After all specified relapse evaluation procedures are complete and the relapse is confirmed, the Investigator may initiate the recommended relapse treatment outlined in this protocol and change the supportive immunosuppressants if needed.

All relapses should be reported in the CRF at the Relapse Evaluation Visit. Relapses that meet the

criteria of a severe adverse event (SAE) should also be reported as a SAE.

To monitor the course of the relapse, Follow-Up Relapse Evaluation Visits will be performed at 4, 8, 12, 16, 20, 24 weeks after the onset of relapse. Unscheduled Follow-Up Relapse Evaluation Visits are permitted at the discretion of the Investigator and must be documented in the patient's CRF. During the Relapse Evaluation Period, high-dose intravenous methylprednisolone (IVMP) administration visits may overlap with the Relapse Evaluation Visit and/or Follow-Up Relapse Evaluation Visits.

8.3 Safety Follow-up Period (Post-Treatment)

If patients withdraw from this trial after receiving any amount of TCZ or AZA, a follow-up visit for safety assessments is required at 12 weeks after the last dose of TCZ or AZA. If a patient is discontinued due to an adverse event (AE), the event will be followed until it is resolved or, in the opinion of the Investigator, is determined medically stable.

8.4 Standard Protocol Definitions

8.4.1 Historical Relapse

Historical relapses are the relapses that occurred prior to the Screening Visit, including the onset of the first attack. They are defined as a new onset of neurological symptoms or worsening of existing neurological symptoms with an objective change on neurological examination (confirmed by EDSS, CSF examination, neuroelectrophysiology, MRI findings or all of them) that persisted for more than 24 hours that required treatment. Treatment for relapses includes IVMP, PE or IVIG. Events that occur within a 30-day interval are considered as one relapse.

8.4.2 Definitions for Relapses during the Trial

Relapses are acute attacks that occur during the trial, defined as a new onset of neurological symptoms or worsening of existing neurological symptoms with an objective change on neurological examination that persists for more than 24 hours as confirmed by the Investigator. The signs and symptoms must be attributed to NMOSD, 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. Investigator is not required to wait 24 hours prior to initiating treatment for the relapse.

There are 6 core clinical characteristics in NMOSD. The relapse of each will meet the following:

Optic neuritis:

- The patient has acute vision loss with or without eye pain and optic papillae edema
- Abrupt abnormal visual field associated with optic nerve damage
- At least 5 characters drop in 100% high contrast visual acuity
- New relative afferent pupillary defect
- Abnormal Visual Evoked Potential
- Confirmatory MRI finding in optic nerve
- Exclude anterior segment lesions, retinal lesions, macular lesions, ametropia, glaucoma
- Exclude optic neuropathy caused by ischemia, compression and infiltration, trauma, toxicity and nutritional metabolic agents

Acute Myelitis:

- At least 0.5-point worsening in EDSS if baseline EDSS >5.5 or at least 1 point worsening in EDSS if baseline EDSS $\leq$ 5.5
- Confirmatory MRI finding in spinal cord

Area postrema syndrome

- Unexplained hiccups or nausea and vomiting
- Confirmatory MRI finding in the area postrema

Acute brainstem syndrome

- Clinical presentations caused by brainstem lesions, i.e., syndrome of medial longitudinal fasciculus, limbic weakness, bulbar palsy
- Confirmatory MRI finding in brainstem

Acute diencephalic clinical syndrome or symptomatic narcolepsy

- Symptomatic narcolepsy, excessive sleepiness or other diencephalic symptoms
- Confirmatory MRI finding in diencephalon

Symptomatic cerebral syndrome

- Clinical presentations caused by cortical or subcortical lesions, i.e., hemiparalysis, aphasia
- Confirmatory MRI finding in the brain

All the confirmatory MRI findings need to be consistent with neuroimaging characteristic of NMOSD (See *Appendix 2*).

8.4.3 The Washout Period of Baseline Immunosuppressants after Randomization

After randomization, patients with AZA group will continue to use AZA if AZA was used for relapse prevention since the last relapse. If AZA has been used for more than 24 weeks when randomization, then AZA will be used as monotherapy, withdrawing any other concomitant immunosuppressants or steroids. If AZA has not been used before or less than 24 weeks, AZA will be used concomitant with the regimen at the time of randomization until concomitant period achieve 24 weeks. If the patient is randomized into TCZ group, then immunosuppressants or steroids at randomization will be gradually withdrawn in 12 weeks.

8.4.4 The Responsibility of the Investigator

The Investigator will be responsible for the overall patient management including patient eligibility evaluation, supervision of the experimental drug administration, recording and treating of AEs and monitoring of safety assessment. At the time of a relapse, the Investigator will perform a complete neurological examination and determine if a patient experiences a relapse and may treat the patient's relapse according to the recommended treatment in Section 10. Treatment for relapse and any changes in the immunosuppressants following a relapse is at the discretion of the Investigator.

8.4.5 The EDSS Rater (Blinded)

The EDSS Rater, an independent neurologist, will be blinded to the randomization and treatments throughout the trial including at the time of a relapse. The rater will perform a complete Kurtzke neurological exam and document the FSS and the EDSS score.

8.4.6 Independent Laboratory / SD-OCT / VEP Assessors

The laboratory (including serum AQP4-IgG, B cell subsets determination) / SD-OCT / VEP assessors roles are responsible for the specified examination of the corresponding index. They will be blinded to the randomization and treatments throughout the trial including at the time of a relapse.

8.4.7 MRI imaging

MRI including enhancement at relapses will be conducted. MRI scans will be analyzed independently at a central MRI reading center by staff members who are also unaware of the trial-group assignment.

8.5 Schedule of Assessments

8.5.1 Screening Period

Table 1. Screening Period (1-4w)

Trial visit	**V1**
Screening Period Duration	0-4w
Informed consent	×
Demography	×
Physical or surgical history	×
NMOSD history	×
Thiopurine methyltransferase genotyping[a]	×
Pregnancy test (serum) [b]	×
Inclusion/Exclusion criteria[c]	×
Vital signs	×
Physical examination[d]	×
Expanded Disability Status Scale (EDSS)	×
Concomitant diseases and medications[e]	×
Electrocardiogram	×
Clinical laboratory tests[f]	×
Serum AQP4-IgG	×
High-contrast visual acuity (100%)	
Low-contrast letter acuity (2.5%)	
Spectrum Domain-Optical Coherence Tomography (SD-OCT)[g]	
VEP[h]	
MRI[i]	

8.5.2 Study Period

Table 2.Treatment Period

Trail visit	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26
Trial Week	0w	4w	8w	12w	16w	20w	24w	28w	32w	36w	40w	44w	48w	52w	56w	60w	64w	68w	72w	76w	80w	84w	88w	92w	96w
Trial day	D1	28 ±7	56 ±7	84 ±7	112 ±7	140 ±7	168 ±7	196 ±7	224 ±7	252 ±7	280 ±7	308 ±7	336 ±7	364 ±7	392 ±7	420 ±7	448 ±7	476 ±7	504 ±7	532 ±7	560 ±7	588 ±7	616 ±7	644 ±7	672 ±7
Demography																									
Physical/Surgical history																									
Thiopurine Methyltransferase genotyping[a]																									
Randomization	×																								
Inclusion/Exclusion criteria[c]																									
Vital signs	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Physical examination[d]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Expanded Disability Status Scale(EDSS)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Concomitant diseases and medications[e]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Electrocardiogram	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Clinical laboratory tests[f]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Serum AQP4-IgG	×			×			×			×			×			×			×			×			×
High-contrast visual	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

acuity (100%)																								
Low-contrast visual acuity (2.5%)	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Spectrum Domain-Optical Coherence Tomography (SD-OCT)[g]	×			×			×			×			×			×			×			×			×
VEP[h]（If necessary）	×			×			×			×			×			×			×			×			×
MRI[i]	×	If necessary, according to clinical needs														×	If necessary, according to clinical needs							×
B cell subsets	×			×			×			×			×			×			×			×			×
Toxicity assessment/Adverse Events[j]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Patient Education Brochure [k]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
IMPs administration[l]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

8.5.3 Relapse Evaluation Period

Table 3. Relapse Evaluation Period

Assessment	**Relapse Visit**	**Follow-up Relapse Visit**						
Trial Week	Within 24 Hours	+W4	+W8	+W12	+W16	+W20	+W24	Unscheduled visit
Vital signs	×	×	×	×	×	×	×	×
Physical examination[d]	×	×	×	×	×	×	×	×
Expanded Disability Status Scale (EDSS)	×	×	×	×	×	×	×	×
Concomitant diseases and medications[e]	×	×	×	×	×	×	×	×
Electrocardiogram	×	×	×	×	×	×	×	×
Clinical laboratory tests[f]	×						×	×
Serum AQP4-IgG	×			×			×	×
High-contrast visual acuity (100%)	×	×	×	×	×	×	×	×
Low-contrast letter acuity (2.5%)	×	×	×	×	×	×	×	×
Spectrum Domain- Optical Coherence Tomography (SD-OCT)[g]	×			×			×	
VEP[h]（If necessary）	×			×			×	
MRI[i]	×						×	
B cell subsets	×			×			×	×
Toxicity assessment/Adverse Events[j]	×	×	×	×	×	×	×	×
Survival[m]	×	×	×	×	×	×	×	×

8.5.4 Safety Follow-up Period (Post-Treatment)

Table 4. Safety Follow-up Period (Post-Treatment)

Assessment	Follow-up Visit
Trial Week	+12 w
Vital signs	×
Physical examination[d]	×
Expanded Disability Status Scale (EDSS)	×
Concomitant diseases and medications[e]	×
Electrocardiogram	×
Clinical laboratory tests[f]	×
Pregnancy test (serum) [b]	×
Spectrum Domain-Optical Coherence Tomography (SD-OCT)[g]	
VEP[h]（If necessary）	
MRI[i]	
B cell subsets	
Toxicity assessment/Adverse Events[j]	×
Survival[m]	×

Footnotes：

a. The serum thiopurine methyltransferase activity mainly reflects the ability of the patient to metabolize azathioprine in the liver. If the activity of the enzyme is low, it indicates that the possibility of liver function damage after the application of azathioprine is significantly increased, which may result in the patient not being able to complete follow-up is completed. Therefore, this test is one of the basic conditions for screening patients.

b. Women who are fertile before menopause must have a negative urine or serum pregnancy test within 7 days prior to the first dose. If the result is positive, the subject will not be eligible to participate in the study. If a pregnancy condition is suspected during the test, the test should be repeated.

c. If the assessment is performed within 7 days prior to the first dose and the listed exclusion criteria are met (if applicable), it does not need to be repeated on the first dose, unless the Investigator believes that significant changes may occur.

d. Physical examination includes heart rate, blood pressure, respiratory rate, body temperature, height, weight, and neurological examination. If weight is measured at screening, there is no need to repeat unless there is a clinical indication.

e. The conbined diseases can be obtained from the history of the patient, including all the diseases, irrespective of autuoimmune diseases or other systemic diseases. The collected combined medication data, including drug dose, route of administration, dosing schedule, start date, indications, end date, and end cause, must be recorded from the time the study drug is used until one month after discontinuation. The Investigator gets this information from the patient's hospital follow-up.

f. Hematology examinations such as blood routine and blood biochemistry during the screening period within 7 days before the first administration, followed by blood tests such as blood routine and blood biochemistry should be studied within 7 days before and after the time point specified in the flow chart.

g. Retinal OCT examination must be carried out regularly according to the visiting cycle, recording the thickness of the retinal nerve fiber layer,ganglion cell complex, etc. If clinical recurrence is suspected during the non-visit period, or new lesions are suspected to occur in any part, OCT should be reviewed appropriately and in a timely manner, with the same regular visit.

h. FF-VEP checks include visually induced potential p100 wave and incubation period. If clinical recurrence is suspected during the non-visit period, or new lesions are suspected to occur in any part, FF-VEP should be reviewed appropriately and promptly, with the same regular visit.

i. Neuroimaging examination must be carried out regularly according to protocol, and the number and location of lesions must be recorded. If clinical recurrence is suspected during the non-visit period, or new

lesions are suspected to occur in any area, appropriate and timely imaging examination should be carried out. These imaging examinations include head, neck, thoracic MRI flat-sweep and gadolinium enhancement examination.
j. If any conscious abnormal symptoms, new or worsening neurological symptoms (e.g. limb numbness, weakness, pain, etc.) occurs, the patient should be given appropriate medical assistance immediately. Any symptoms are treated as clinically routine and, if defined, reported as adverse events or severe adverse events (SAEs). The collection of adverse events begins when the patient signs the informed consent form. After the study is terminated, all unhealed adverse events or serious adverse events need to be tracked until resolved, unless the researchers believe that the patient's own illness is unlikely to ease. All new adverse events and serious adverse events within 30 days after the last dose of the study were reported and tracked as described above. Record the most severe levels of toxic reactions in each course of treatment in the CRF.
k.Patient Education Brochure describes all the signs and symptoms of a definite attacks of NMOSD and the Package Inserts of the IMPs. Patients can make a record of all the discomfort during the trial. At each visit throughout the trial, the Investigator will ensure that the patient has the Patient Education Brochure.
l. Patients received oral azathioprine (2-3 mg/kg/d) or intravenous tocilizumab (8 mg/kg/4 weeks). Azathioprine was initiated at 25 mg/d and increased stepwise by 25 mg/d increments until the target dose was reached. Patients experiencing medication-related side effects during the loading period were allowed symptomatic treatments, with the exception of any new immunosuppressants. Patients received their final, stable dosage of azathioprine, as daily maintenance until relapse, discontinuation, or the end of the trial. Tocilizumab was administered intravenously over an approximate 60 minutes duration.
m. The Investigator should obtain information about the patient’s overall survival until the end of the study.

8.6 Schedules of Assessments and Procedures

8.6.1 Screening Examination for Eligibility (Screening Visit [Visit 1]) Occurs 1-4 weeks Prior to Baseline

Once a patient has written the informed consent, the following evaluations will be performed to make sure she or he will fulfill the entry criteria.The following examinations will be performed within 1-4 weeks prior to randomization to determine patient eligibility for participation in the trial:
The Investigator will perform the following:

- Signed informed consent form is indispensible. Note: the rial duration for an individual patient in this time-to-event trial will vary, i.e., the trial duration for in individual patient depends on the time of relapse or the time of trial closure
- Demography
- Physical or surgical history
- NMOSD history: Review the signs and symptoms of potential NMOSD relapse with the patient and instruct the patient to contact the study site at the first signs or symptoms of potential relapse
- AQP4-IgG test history confirmed by cell-based assay (CBA) method
- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS (See *Appendix 4*). Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity (See *Appendix 5*).
- Thiopurine methyltransferase genotyping
- Vital signs include assessments of systolic and diastolic blood pressure (BP), temperature, respiration rate (RR) and heart rate (HR)
- Electrocardiogram (ECG)
- Concomitant diseases and medications
- Clinical laboratory tests (hematology and chemistry) (See *Appendix 6*)
- Pregnancy test (serum human chorionic gonadotropin) must be performed on all women of child-bearing age. Note: if the patient is taking/using contraceptive medication/device, please be sure to record the medication or device in the CRF.
- Provide the brochure for the patients that describes the potential signs and symptoms of NMOSD relapse, package inserts of IMPs and the contact information of the study center.

8.6.2 Study Period

Visit intervals during the Study Period are every four weeks (every 28 days ± 7 days).

8.6.2.1 Baseline ([Visit 2/Day 1]

Once all of the Baseline visit procedures have been performed and the eligibility criteria have been confirmed, the patient will be randomized to one of the two treatment group: TCZ or AZA. The subject randomization numbers will be generated by the randomization center, School of Public Health, Tianjin Medical University and the numbers are to be allocated sequentially in the order in which the subjects are enrolled according to the specification document agreed with the randomization center. As confirmation, the site will be provided with a verification of each patient's randomization. the following tests and procedures will be completed at the Baseline visit (Visit 2/Day 1):

The Investigator will perform the following assessments:

- Review and assess the patient for potential signs or symptoms indicative of relapse
- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Bilateral monocular SD-OCT
- Bilateral monocular VEP
- Protocol-defined MRI
- Record the dosage of the IMPs
- Provide the brochure for the patients that describes the potential signs and symptoms of NMOSD relapse, package inserts of IMPs, and the contact information of the study center

8.6.2.2 Routine follow-up evaluation (From Visit 3 [Week 4] Through the end of the study or early termination)

The Investigator will review and assess the patient for any potential signs or symptoms indicative of relapse. The following tests and procedures will be completed at every visit until the end of the study or early termination:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Serum AQP4-IgG at protocol-defined visit
- Bilateral monocular SD-OCT at protocol-defined visit
- Bilateral monocular VEP at protocol-defined visit
- Protocol-defined MRI at protocol-defined visit
- Record the dosage of the IMPs
- Provide the brochure for the patients that describes the potential signs and symptoms of NMOSD relapse, package inserts of IMPs, and the contact information of the study center
- Body weight will be measured at 60 weeks or end of study visit or early termination visit

8.6.2.3 Washout Period

The patients will experience a washout period during which baseline immunosuppressants will taper until discontinuation.

- The patients in the TCZ group must taper immunosuppressants till withdrawal at 12 weeks.
- The patients in the AZA group will also discontinue concurrent immunosuppressants at 24 weeks. After withdrawal of concurrent immunosuppressants, patients in both groups will receive mono-therapy without other immunosuppressants. If the patients had used AZA for at least 24 weeks before randomization, concomitant immunosuppressants would be discontinued at baseline. If the patients had received AZA treatment with concomitant immunosuppressants for less than 24 weeks before randomization, the patients will continue to use concomitant immunosuppressants until concomitant period achieved 24 weeks. Then concomitant immunosuppressants will be discontinued.

After the washout period, TCZ or AZA will be administered as monotherapy. No other or new immunosuppressants will be given.

8.6.2.4 Missed Visits

Patients may fail to return for a scheduled visit due to specific factors. In this event, the Investigator must contact the patients by telephone. The reason why the patients cannot get to the study site must recorded in the CRF. If a potential relapse or an AE is considered, patients are strongly encouraged to ask the Investigator for help to assure the missing visit was or not due to a potential relapse or an AE. If the patients cannot come to the study site for examination, then the patients will be instructed to see the local neurologist. The Investigator must contact the local neurologist to obtain as much information as possible about the patient's medical and neurological condition, and provide clinical guidance. The Investigator must record all the relevant medical information.

If the patients decide to withdraw from the trial at any time, a further 12-week follow-up is needed.

8.6.3 Relapse Evaluation Visit (Within 24 Hours)

Patients will be educated and directed to contact the study site if they have new symptoms or worsening on the existing symptoms. The Investigator and the blinded EDSS Rater.will evaluate within 24 hours to determine a possible relapse.

The Investigator will perform the following assessments:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Serum AQP4-IgG at protocol-defined visit
- Bilateral monocular SD-OCT at protocol-defined visit
- Bilateral monocular VEP at protocol-defined visit
- Protocol-defined MRI at protocol-defined visit
- Record the dosage of the IMPs

The Investigator determines if the clinical signs, symptoms and neurological change (objective findings on the examination) meet the definition for relapses as outlined in this protocol. The Expert Panel will evaluate all the information done as above to assure the relapses. If the relapse is confirmed, the Investigator may initiate the recommended treatment regimen for confirmed relapse outlined in

this protocol (see *Section 10*).After acute attack, patients will be encouraged to receive potent treatments to prevent relapses.

8.6.4 Follow-up Relapse Evaluation Visits (Weeks 4, 8, 12, 16, 20, 24)

If the patients have confirmed relapses, then they must be discontinued from this trial and may experience a prolonged 24-week visit. Week 24 Follow-up Relapse Evaluation Visit will serve as the end of visit.

The following tests and procedures will be completed at these Relapse Evaluation Period visits:

The Investigator will perform the following:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Serum AQP4-IgG at protocol-defined visit
- Bilateral monocular SD-OCT at protocol-defined visit
- Bilateral monocular VEP at protocol-defined visit
- Protocol-defined MRI at protocol-defined visit

8.6.5. Unscheduled Follow-up Relapse Evaluation Visits

Unscheduled Follow-Up Relapse Evaluation Visits outside the specified visits are permitted at the discretion of the Investigator. Procedures, tests and assessments listed under the Relapse Evaluation Visit will be performed at the discretion of the Investigator. Any tests, procedures or assessments performed at the unscheduled visit must be recorded in the CRFs.

8.6.6. Safety Follow-up Period (Post-Treatment)

If a patient withdraws from the trial at any time during the Study Period after receiving any amount of experimental drugs, a follow-up visit for safety assessments is required at 12 weeks after the last dose of IMPs. The following tests and procedures will be completed at the Safety Follow-up Visit:

- A complete physical examination, including assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination
- EDSS. Note: The designated EDSS Rater will perform the Kurtzke neurological assessment to determine the FSS and EDSS score and 100% high contrast visual acuity and 2.5% low contrast letter acuity.
- Vital signs including assessments of systolic and diastolic BP, temperature, RR and HR
- ECG
- Concomitant diseases and medications
- Clinical laboratory tests (chemistry and hematology)
- Pregnancy test (serum human chorionic gonadotropin) will be re-performed on all women of child-bearing age. Note: if the patient is taking/using contraceptive medication/device, please be sure to record the medication or device in the CRF.
- Any new AEs or changes in AEs since the previous visit will be evaluated and recorded

If a patient discontinued from the study due to an AE, the event will be followed until it is resolved or, in the opinion of the Principle Investigator, is determined medically stable.

8.7 Number of Patients

A maximum of 118 NMO patients will be randomized in a ratio of 1:1 (TCZ: AZA) across 6 centers. Patients will be randomized to 1 of 2 treatment groups, 59 for each group.

8.8 Treatment Assignment

If all 118 patients are enrolled, approximately 59 patients will be randomized to TCZ and 59 patients will be randomized to AZA. Randomized patients who discontinue after initiation of study treatment will not be replaced. All patients will remain on their assigned treatment until the end of study visit.

8.9 Criteria for Trial Termination

8.9.1 End of Trial for an Individual Patient

The trial will be ended for a patient when one of the following conditions is met, whichever comes first:

- The patient experiences a definite relapse and early termination
- When the last participating patient completes the last scheduled visit or when the Expert Panel decide to discontinue the study or development program.

8.9.2 End of Trial for All Patients

The end of the trial has been defined as the date at which the last data point from the last patient is collected.or when the Expert Panel decide to discontinue the study or development program. The end of trial is defined as completion of the end of study by all patients. The end of study visit is to be completed as soon as possible, preferably within 1 month of the end of the trial notification, with the exception of patients who experience a relapse who will continue until completion of the Relapse Evaluation Period.

9. SELECTION AND WITHDRAWAL OF PATIENTS

9.1 Patient Inclusion Criteria

1. Male or female patients ≥ 18 years old
2. Diagnosis of NMOSD as defined by 2015 Criteria.
3. Historical Relapse (as defined by this protocol) of at least 2 relapses in the last 12 months or 3 relapses in the last 24 months, with at least 1 relapse in the 12 months prior to the screening
4. Able and willing to give written informed consent and comply with the requirements of the study protocol.
5. EDSS ≤ 7.5
6. Men and women of reproductive potential must agree to use a highly effective method of birth control from screening to 6 months after final dose of the experimental drugs.

9.2 Patient Exclusion Criteria

1. Current evidence or known history of clinically significant infection (Herpes simplex virus, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human immunodeficiency virus, Hepatitis viruses, Syphilis, etc.)
2. Pregnant, breastfeeding, or child-bearing potential during the course of the study
3. Patients will not participate in any other clinical therapeutic study or will not have participated in any other experimental treatment study within 30 days of screening
4. Participation in another interventional trial within the last 3 months
5. Heart or kidney insufficiency
6. Tumor disease currently or within last 5 years
7. Clinically relevant liver, kidney or bone marrow function disorder
8. Receipt of rituximab or any experimental B-cell depleting agent within 6 months prior screening and B-cells below the lower limit of normal measured by flow cytometry
9. Prior relapses or toxic intolerance on azathioprine treatment (adequate dosage and follow-up period observation)

9.3 Patient Discontinuation Criteria

The Investigator has the right to discontinue a patient from study drug or withdraw a patient from the study at any time. In addition, patients have the right to voluntarily discontinue study drug or withdraw from the study at any time for any reason. Reasons for discontinuation of study drug or withdrawal from the study may include, but are not limited to, the following:

- Patient withdrawal of consent at any time
- Any medical condition that the Investigator determines may jeopardize the patient's safety if he or she continues in the study
- Investigator determines it is in the best interest of the patient
- Patient non-compliance, defined as significant non-compliance with instructions for home administration of study drug

Patients who discontinue study drug prematurely will be asked to return to the clinic and may undergo follow-up assessments. A follow-up visit for safety assessment is required at 12 weeks after the last dose of experimental drug administration. The primary reason for premature study drug discontinuation should be documented on the CRF. Patients who discontinue study drug prematurely will not be replaced.

If a patient is discontinued due to an AE, the event will be followed until it is resolved or in the opinion of the PI the patient is determined to be medically stable. Every effort will be made to undertake protocol-specified safety follow-up procedures.

9.4 Withdrawal from Study

Every effort should be made to obtain information on patients who withdraw from the study. The primary reason for withdrawal from the study should be documented on the CRF. Patients will not be followed for any reason after consent has been withdrawn. Patients who withdraw from the study will not be replaced.

10. MEDICATIONS FOR THE PATIENTS

10.1 Investigational Midicinal Product (IMP)

10.1.1 Investigational Midicinal Product (IMP) Dosage and Administration

Each vial of TCZ contains TCZ 80 mg for IV administration. TCZ (8 mg/kg) will be administered to the patient intravenously over approximately 60 mins via an infusion pump. If an AE occurs during the administration of the TCZ, the infusion may be slowed or stopped at the discretion of the Investigator, depending upon the nature and severity of the event. The AE must be recorded in the patient's CRF.

Each tablet of AZA contains 50 mg for oral administration. AZA (2-3 mg/kg) will be administered orally to the patient. AZA was initiated at 25 mg/d and increased in stepwise by 25 mg/d increments until the target dose was reached.

10.1.2 IMP Dose Modifications, Interruptions and Delays

No dose modifications are foreseen.

10.1.2.1 AZA Administration Adjustment

If the patients had medication-related side effects during the loading period, symptomatic treatment was allowed with the exception of any new immunosuppressants. Patients received their maximal tolerated dose of AZA (1-2 mg/kg/d would be also permitted due to side effects), as daily maintenance until relapse, trial discontinuation, or the end of the trial.

10.1.2.2 TCZ administration adjustment

Slowing of the infusion rate or interruption of the infusion, may be necessary in the event of an infusion reaction. In rare cases, TCZ treatment may need to be Discontinued. Handling of infusion related reaction will depend on the intensity of symptoms. In the event that a patient experiences a mild to moderate (Common Terminology Criteria for Adverse Events [CTCAE] Grade 1 or 2; (see *Appendix 7*) non-allergic infusion-related event, the infusion rate should be reduced to half the rate being given at the time of onset of the event. Once the event has resolved, the Investigator should wait for 30 minutes while delivering the infusion at the reduced rate. If tolerated, the infusion rate may then be increased to the next closest rate on the patient's infusion schedule and the rate increments resumed.

Patients who experience a severe infusion-related event (CTCAE Grade 3) or a complex of flushing, fever and throat pain symptoms, should have their infusion interrupted immediately and should receive aggressive symptomatic treatment (such as i.v. diphenhydramine 20-40 mg or 5-10 mg of dexamethasone or both). The infusion should be re-started only after all the symptoms have disappeared. The initial infusion rate at restart should be half of the infusion rate that was in progress at the time of onset of the reaction.

Patients who experience a life threatening infusion-related event (CTCAE Grade 4) during an infusion should have their infusion immediately stopped and should receive appropriate treatment (including i.v. diphenhydramine 20-40 mg or 5-10 mg of dexamethasone or both, use of resuscitation medications and equipment that must be available and used as clinically indicated). These patients will be withdrawn from treatment and should enter the Safety Follow-up Period.

Patients in the TCZ group may prolong infusion interval if they are not tolerable to the dosage of 8 mg/kg, especially when the patients have active infecious diseases.

10.1.3 IMP Packaging and Labeling

TCZ is manufactured by Chugai Pharma Manufacturing Co., Ltd., affiliated by Roche Pharma (Schweiz) Ltd. and purchased from the third Pharmacy. AZA is manufactured by Excella GmbH & Co.KG and purchased from the third Pharmacy. They are stored in the Pharmacy of Tianjin Medical

University General Hospital. Both vials and tablets will be labeled according to the protocol and local regulatory requirements. The drugs will be transported to each participating trial center upon receipt of all required essential documents.

Table 5. Investigational Medicinal Product (IMP)

	IMPs	
Product Name	**Tocilizumab**	**Azathioprine**
Dosage Form	Concentrate solution for infusion	Solid tablet
Unit Dose	80 mg	50 mg
Route of Administration	Intravenous Infusion	Oral
Physical Description	4 mL vial	-
Manufacturer	Chugai Pharma Manufacturing Co., Ltd., affiliated by Roche Pharma (Schweiz) Ltd.	Excella GmbH & Co.KG

10.1.4 IMPs Storage

IMPs kit will have a booklet label describing the contents and a place for the pharmacist to record the patient number, patient initials and Investigator name.

Upon arrival at the center, TCZ should be promptly removed from the cooler and stored in refrigerated conditions at 2°C to 8°C with minimal light exposure. The pharmacist should immediately record the receipt of TCZ or AZA. AZA should be stored at room temperature.

10.2 Concomitant Therapy

Any concomitant medications (including prescription drugs, over-the-counter medications, herbal/homeopathic medications, preventive vaccines, vitamins, and nutritional supplements) must be recorded in the CRF. A description of the type of drug, amount, duration, and reason for administration of drug must be documented. Adverse events that are judged to be related to the administration of a concomitant medication must also be documented on the CRF.

10.2.1 Concomitant Immunosuppressants

Immunosuppressants are permitted to be used in the washout period before experimental drug is given as monotherapy. Immunosuppressants including AZA, MMF, MTX, methotrexate, tacrolimus, cyclosporine or cyclophosphamide are permitted.

If a patient enters the trial receiving steroids either as mono-therapy or in combination with another immunosuppressant, the daily steroid dose will not be more than prednisone 20 mg daily (or equivalent).

No new immunosuppressants or switch to another immunosuppressant is permitted during the trial unless the patient experiences a relapse. After a relapse there are no restrictions on adjustments or changes of immunosuppressants.

10.2.2 Glucocorticoid-Induced Osteopenia/Osteoporosis Prevention and Treatment

Patients should receive oral calcium and 25-hydroxy vitamin D supplementation unless contraindicated (calcium 1200−1500 mg and vitamin D 800−1000 IU daily in divided doses). Unless contraindicated, bisphosphonate therapy (e.g., alendronate 70 mg weekly) will also be administered at the discretion of the physician-investigator for the prevention of glucocorticoid-induced osteoporosis.

Participants with documented osteoporosis will be treated with approved drugs for osteoporosis according to local practice or clinical guidelines.

10.2.3 Other Permitted Medications

The patients will be permitted to continue treatments for chronic historical diseases. For example, subjects in both arms should be treated with anti-platelet therapy (aspirin or clopidogrel) according to the local practice at the discretion of the Investigator. Symptomatic treatments are permitted during the course of the trial for underlying conditions. For example, oxcarbazepine could be used for neuropathic pain.

10.2.4 Prohibited Medications

Treatment with any cell-depleting therapies, biologic agents (e.g., tumor necrosis factor antagonists), Janus kinase inhibitors, alkylating agents such as chlorambucil, bone marrow transplantation with total lymphoid irradiation, or extracorporeal photopheresis is prohibited during the study.

Immunization with a live or attenuated vaccine is prohibited within 4 weeks of baseline for the duration of the patient's study participation and 12 weeks after administration of the last dose of study drug.

10.3 Recommended Standardized Relapse Treatment

For this protocol, the treatment for relapse is at the discretion of the Investigator. The following standardized treatment regimen for a confirmed relapse is recommended.

1. One-gram IV methylprednisolone (IVMP) administered daily for 3-5 days followed by an oral prednisone tapering. If the patient improves, then continue the trial assessments as per the schedule of this protocol.
2. If the patient's condition does not sufficiently improve or the neurological symptoms worsen, therapeutic plasma exchange (PE, five to seven cycles; 1 - 1.5 plasma volume per exchange) can be performed.
3. If there is no or minimal response to methylprednisolone, Intravenous Immunoglobulin (IVIG) will be allowed at the discretion of the Investigator. IVIG 0.4 g/kg/d for 5 days is recommended for treatment of attacks that do not respond to IVMP.

11. ASSESSMENT OF EFFICACY

11.1 Efficacy Parameters

Duration of treatment commences with the first experimental drugs administration. The Study Period defines the time period for assessment of the trial endpoints. At the point when the trial is stopped, all data from all patients will be collected, database cleaned, locked, and analyzed. Data from the Study Period will be used for efficacy analysis.

11.1.1 Relapses

Accurate identification and evaluation of confirmed relapses is critical for the integrity of the study. The primary efficacy endpoint is time-to-first relapse. The secondary efficacy endpoints include Time to onset of confirmed disability progression (CDP) for at least 12 weeks, determination of serum AQP4-IgG titers from baseline to 60 weeks and overall safety and tolerability of experimental drugs.

Pre-treatment historical relapses will be reviewed by the Investigator to determine if they meet criteria for Historical Relapse as defined by this protocol. Relapses will be monitored throughout the trial. The Investigator will review, in detail, the signs and symptoms of a potential relapse with the patient at each visit. Patients will be instructed to contact the study site at the first sign or symptom of a relapse. Patients should be evaluated within 24 hours of notification of the Investigator.

All potential relapses must be evaluated by both the Investigator and EDSS Rater. The Investigator will make the decision as to whether the clinical signs, symptoms and neurological changes (objective findings on exam) meet the protocol definition of relapse and may treat the patient's relapse according to the Recommended Standardized Relapse Treatment. The relapse treatment is at the Investigator's discretion. All investigations/tests related to the relapse evaluation should be recorded in the CRF.

Follow-up Relapse Evaluation Visits to monitor the course of the relapse until stabilization will be made according at the Investigator's discretion. All reports of possible relapses and actions taken must be documented in the patient's source documents and recorded in the CRF. Relapses that do not meet the criteria for SAE should be reported as part of the Relapse Evaluation visits, and not as AEs.

The Investigator at each center are responsible for evaluating patient eligibility, supervising the administration of study medication, recording and managing adverse events, assessing relapses. The Investigator identify relapses according to the following criteria: a new onset of neurological symptoms or worsening of existing neurological symptoms with an objective change on neurological examination that persisted for more than 24 hours, signs and symptoms attributable to NMOSD rather than other causes, and onset preceded by at least 30 days of clinical stability. Patients will be evaluated within 24 hours after a possible relapse, which is confirmed by the Expert Panel and again following intervals of 4 weeks until 24 weeks by the Investigator and by EDSS raters. The EDSS raters, who are also unaware of trial-group assignments, are not involved in patient care. Confirmed relapses could be treated with IVMP or IVIG at the Investigator's discretion.

11.1.2 Disability

Disability will be assessed based on the EDSS scores comparing the change from baseline at the end of trial in the two treatment groups. An EDSS Rater who is blinded to all other trial and patient clinical data will be responsible for performing the EDSS assessments throughout the trial at the protocol specified time points as well as at visits during the Relapse Evaluation Period.

To investigate the efficacy of TCA compared with AZA in patients with relapsing NMOSD, as measured by the time to onset of sustained confirmed disability progression (CDP) over the treatment period, defined as an increase in EDSS that is sustained for at least 12 or 24 weeks, based on regularly scheduled visits. Disability progression is defined as an increase of ≥ 1.0 point from the baseline EDSS when the baseline score is 5.5 or less, and ≥ 0.5 when the baseline score is more than 5.5, that is not attributable to another etiology (e.g., fever, concurrent illness, or concomitant medication).

Confirmation of disability progression must occur at a regularly scheduled visit that is at least 12 or 24 weeks after the initial disease progression. The non-confirmatory EDSS assessments (if any) between the initial and confirmation of disability progression should be at least as high as the minimum change required for progression.

11.1.3 Neurological Functions

Neurological function will be assessed based on EDSS. To assess vision precisely, low-contrast letter scores are measured using retro-illuminated 2.5% Sloan letter chart (Precision Vision, La Salle, IL, USA) using best refractive correction for each eye at 2.52 m[38]. Best corrected high-contrast logMAR visual acuity is measured using retro-illuminated Early Treatment Diabetic Retinopathy Study chart (Precision Vision, La Salle, IL, USA) at 2.52 m. The patient can miss one letter per row to score that row. When no letters could be correctly identified, a score of 1.7 is assigned by the masked researcher[39].

11.1.4 Serum AQP4-IgG Titers

For AQP4-IgG positive NMOSD patients, serum AQP4-IgG titers will be determined by cell-based assay (CBA)[40,41] at scheduled visits.

11.1.4.1 Plasmid construction

We construct AQP4 M23 Plasmid DNA and purify from E. coli XL-2 grown in LB-medium using the Endo Free Plasmid Maxi Kit (QIAGEN). The detailed procedures follow the manufacturer's instructions.

11.1.4.2 Transfection

- Human embryonic kidney (HEK293) cells are seeded on cover glasses with a density of 1×10^5 cells in 12 well.
- When the adherent cells' confluence reaches 70%-80%, the cultured HEK293 cells will be transfected with AQP4 M23 plasmid using Lipofectamine 2000 (Invitrogen) the detailed transfection procedures are followed the manufacturer's protocol.
- After 48 h of incubation, the transfected cells are fixed with PBS and 4% paraformaldehyde. Coated cover glasses are then cut into millimeter-sized fragments. And store in -80℃.

11.1.4.3 Immunofluorescence assay

Each slide coated with AQP4 transfected cells and non-transfected cells, are incubated with the patients' blood samples.

- Incubate each slide with PBS-0.3%Triton for 5min at room temperature.
- Dilute sample with PBS-0.3% Triton into different concentration ratio:1:10,1:60,1:240,1:480,1:720,1:960,1:1440,1:1920, 1:3840, 1:7680.
- The slides are coated with 3%BSA for 30min.
- Incubate with PBS-0.3%Triton for 1h at room temperature.
- Wash with PBS 3 times for 30mins.
- Incubate with goat anti-human IgG (Invitrogen) for 30mins, and then washed as before.

Two independent assessors, who are unaware of patients' timeline classified every blood sample as positive or negative based on the intensity of surface immunofluorescence in direct comparison with non-transfected cells and control samples as positive or negative.

11.1.4.4 Fluorescence Immunoprecipitation Assay (FIPA) for serum AQP4-IgG titers

The changes of serum AQP4-IgG titers will also be confirmed by fluorescence immunoprecipitation assay (FIPA)[42,43]. The protocol for serum AQP4-IgG titers by FIPA is as follows:

- $2\text{-}3\times10^6$ Human embryonic kidney (HEK293) cells were planted in 60 cm^2 dish.
- When the adherent cells' confluence reached 70%, the cultured HEK293 cells were transfected with EGFP-tagged AQP4 M1/M23 plasmid using Lipofectamine 2000 (Invitrogen).

- The detailed transfection procedures are as followed:
  1）Dilute Lipofectamine 2000 60ul in 1.5ml Opti-MEM Medium.
  2）Dilute 15ug DNA in 1.5ml Opti-MEM Medium.
  3）Add diluted DNA to diluted Lipofectamine® 2000 Reagent (1:2 ratio).
  4）Incubate 10 min.
  5）Add DNA-lipid complex to cells.
- After 48 h of incubation, the transfected cells are lysed with extraction buffer (10 mM Tris–HCl pH 7.5,100 mM NaCl, 1 mM EDTA, 1% Triton X-100), adding 1ml buffer per flask, shaking the flask.
- The supernatant is acquired after centrifugation (4°C 15min, 13000 rpm). And the extracted EGFP-AQP-4 protein is assayed, the fluorescent units (FUs) are assayed and adjusted to 500–600 FU/200μL extract.
- Preparation of Protein A Sepharose: using 5% ultra-low fetal calf serum (Gibco) to block the sepharose (75μL/case).
- 20μL serum is incubated with 200μL extracted EGFP-AQP-4 protein at 4°C overnight.
- The antibody IgG is precipitated by the addition of 75μL Protein A Sepharose (IgG:sepharose=1:1).
- Washing beads with extraction buffer 3 times (1000 rpm, 3min).
- The beads with 150μL extraction buffer are removed to 96-well black microtiter plates and the FU values (Excitation 488 nm, Emission 507 nm) are obtained on a Microplate reader.

11.1.5 SD-OCT

High resolution spectral domain OCT images were acquired at baseline and the last follow-up using identical protocols across sites (RTVUE100-2, Optovue lnc, Fremont, CA, USA). Appropriate quality assurance was undertaken to ensure comparability, with acceptable inter-rater coefficients of variation for measurements of the RNFL (0.51%) and macular volume (0.45%). RNFL measurements used a 3.45 mm diameter circle scan. A fast macular volume scan (20×20° field, 25 horizontal B scans, ART9) was also done. Scans were excluded if they had a signal strength of less than 25 dB or violated international consensus quality control criteria. If there was a relapse of optic neuritis, OCT of the affected eye will be performed 3 months later.

11.1.6 VEP

Visual evoked potentials (VEP) to reversal achromatic checks were recorded at baseline, relapses and the last follow-up according to International Federation of Neurophysiology guidelines on a Synergy system (EMG & Evoked Potential Response Unit, Nicolet, NE, USA) in standard background office lighting. Responses were recorded from the occipital midline (Oz) using midline frontal (Fz) as reference and midline central (Cz) as ground. Latency and amplitude of the P100 component were measured to one decimal place in the replicates. Participants with absent VEP latencies or amplitudes were assigned a value of 200 and 0, respectively.

11.1.7 MRI

Brain and spinal cord MRI scans will be obtained in all patients at baseline and at week 60, as well as the relapse period. T1 weighted gadolinium-enhancing lesions or new/enlarging hyperintense T2 lesions related to relapses will also performed during relapse period. To evaluate the atrophy of CNS, additional sequences will be done. The scan range of brain includes the whole brain, and the scan range of spine includes C1~C7. The following sequences were performed in head. 3D-T1, 3D-T2 FLAIR, 2D-T2WI, Diffusion tensor imaging (DTI), Resting-state functional MRI (fMRI) , Arterial spin labeling (ASL), and Quantitative susceptibility mapping (QSM) or susceptibility weighted imaging (SWI) was optional sequence. Sagittal PD-T2WI, axial 3D T2*, 3D-T1 and DTI performed in cervical spinal cord. All the scans will be performed by trained and certified MRI technicians.
MRI scans will be analyzed by a centralized center for relapses and CNS atrophy. The centralized reading center is blinded to the treatment assignment and clinical information of the patients. At the

Investigator's site, only the local radiologist/technician assigned to the study may have access to the MRI scans, except at baseline and relapses, when the Investigator may view the MRI scan. If safety concern is considered, the Investigator is permitted to have access to the MRI results.

11.1.8 Flow Cytometry (FACS)

FACS will include (but is not limited to) the following cells:

- Total B cells ($CD19^{pos}$)
- B cell subsets, e.g. memory B cells, naïve B cells, double negative B cells, double positive B cells, and antibody secreting cells.

12. ASSESSMENT OF SAFETY

12.1 Safety Parameters and Definitions

Safety assessments will consist of monitoring and recording adverse events, including serious adverse events and non-serious adverse events of special interest; measurement of protocol-specified safety laboratory assessments; measurement of protocol-specified vital signs; and other protocol-specified tests that are deemed critical to the safety evaluation of the study.

12.1.1 Demographic/Medical History

At Visit 1 (Screening), patients' initials, date of birth, race or ethnic origin and sex will be collected, medical history will be reviewed, and data will be recorded. Medical history including relevant medical/surgical history and NMOSD history will be reviewed and recorded.

12.1.2 Vital Signs

Vital signs will be measured at every visit and will include assessments of systolic and diastolic BP, temperature, RR and HR. Vital signs will be obtained after the patient has been supine or seated for at least 5 minutes. Ideally, each patient's BP should be measured using the same arm. Systolic and diastolic BPs will be documented in mmHg. Temperature will be obtained in degrees Celsius. HR will be documented in beats per minute.

12.1.3 Weight and Height

Body weight will be measured in pounds or kilograms. Height will be measured in inches or centimeters.

12.1.4 Physical Examination

The complete physical examination will include assessments of the following organ/body systems: skin, head, ears, eyes, nose, throat, neck, lymph nodes, pulse, chest, heart, abdomen, extremities, musculoskeletal, and general neurological examination. For consistency, all efforts should be made to have the physical examination performed by the same qualified study staff at these visits.

12.1.5 Electrocardiogram (ECG)

A 12-lead ECG will be conducted. Additional ECG assessments are permitted, at the Investigator's discretion. The Investigator will be responsible for reviewing the ECG to assess whether the ECG is within normal limits and to determine the clinical significance of the results. These assessments will be indicated on the CRF.

12.1.6 Laboratory Assessments

Patients will have biologic samples collected for analysis of various parameters. Chemistry panel, complete blood count (CBC) and hepatic function measures, renal function measures, and serum pregnancy test will be prepared. In the NMOSD patient population that may also be treated with immunosuppressants that are known to affect WBC counts, close monitoring of cell counts is imperative.

Routine hematology laboratory assessment including CBC will be performed at various time points as specified by the protocol and should be reviewed as soon as the lab result is available. Additional assessments to monitor WBC counts can be performed at the discretion of the Investigator as medically indicated. Treatment of leukopenia is at the discretion of the Investigator and consultation with hematologist is encouraged as medically indicated.

AEs and events related to the patients' underlying disease that have occurred during the trial will be collected at every visit.

Any clinically significant, abnormal laboratory result is to be reported as an AE.

12.2 Adverse and Serious Adverse Events

12.2.1 Definition of Adverse Events

12.2.1.1 Adverse Event

According to the International Conference of Harmonisation (ICH) guideline for Good Clinical Practice, an adverse event is any untoward medical occurrence in a patient or clinical investigation patient administered a pharmaceutical product,which does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign, including an abnormal laboratory finding, symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the medicinal product. Pre-existing conditions which worsen during a study are to be reported as AEs.

All the events were recorded and classified according to Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0 (see *Appendix 7*). For AEs with a causal relationship to the experimental drugs, the Investigator must follow-up on the outcome of the event until the event or sequelae either resolve or stabilize.

12.2.1.2 Assessment of Severity of Adverse Events

Table 6 provides guidance for assessing adverse event severity.The adverse event severity grading scale for the NCI CTCAE (v5.0) will be used for assessing adverse event severity. Table 6 will be used for assessing severity for adverse events that are not specifically listed in the NCI CTCAE.

Table 6. Adverse Event Severity Grading Scale

Grade Severity	Description
Grade 1	Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; or intervention not indicated
Grade 2	Moderate; minimal, local, or non-invasive intervention indicated; or limiting age-appropriate instrumental activities of daily living*
Grade 3	Severe or medically significant, but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; or limiting self-care activities of daily living**
Grade 4	Life-threatening consequences or urgent intervention indicated
Grade 5	Death related to adverse event

Activities of Daily Living (ADL)
*Instrumental ADL refer to preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.
**Self care ADL refer to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden.

Adverse events not listed by the CTCAE will be graded using the following criteria:
Grade 1: Discomfort noticed but no disruption of normal daily activity
Grade 2: Discomfort sufficient to reduce or affect normal daily activity
Grade 3: Inability to work or perform normal daily activity
Grade 4: Represents an immediate threat to life (falling into the category of SAE).
In the CRF, adverse events will be reported at each visit.

12.2.1.3 Serious Adverse Events (Immediately Reportable to the Sponsor)

A Serious Adverse Event is any experience that suggests a significant hazard, contraindication, side effect or precaution. It is any AE that, at any dose, fulfils at least one of the following criteria:

- Fatal (i.e., the adverse event actually causes or leads to death)
- Life threatening (refers to an event in which the patient was at immediate risk of death at the time of the event; it does not refer to an event which could hypothetically have caused a death had it been more severe)
- Requires or prolongs inpatient hospitalization

- Results in persistent or significant disability/incapacity (i.e., the adverse event results in substantial disruption of the patient's ability to conduct normal life functions)
- Congenital anomaly/birth defect in a neonate/infant born to a mother exposed to study drug
- medically significant or requires intervention to prevent one or other of the outcomes listed above

When the NMOSD relapse results in hospitalization for any reason other than for routine treatment of the relapse (such as for a treatment course beyond the standard treatment described in or when hospitalization is prolonged, the NMOSD relapse should be considered a SAE.

The terms "severe" and "serious" are not synonymous. Severity refers to the intensity of an adverse event (rated as mild, moderate, or severe, or according to CTCAE criteria); the event itself may be of relatively minor medical significance (such as severe headache without any further findings). Severity and seriousness need to be independently assessed for each adverse event recorded on the CRF.

Serious adverse events are required to be reported by the Investigator within 24 hours after learning of the event.

12.2.2 Hospitalization

AEs that are associated with hospitalization or prolongation of hospitalization are considered SAEs. All admissions to a health care facility meet the criteria, even if for less than 24 h.
Criteria for seriousness are also met if transfer within the hospital is done to receive more intense medical/surgical care (e.g., from the medical floor to the Intensive Care Unit).
Hospitalization does not include the following:

- Rehabilitation facility
- Nursing facility
- Emergency Room
- Same day surgery

Hospitalization or prolongation of hospitalization not associated with an AE is not an SAE, examples include:

- Admission for a pre-existing condition not associated with either a new AE or with worsening of a pre-existing AE
- Protocol-specified admission
- Pre-planned admission

If a relapse of NMOSD result in hospitalization for the patient, then AE can be recorded. If prolonged hospitalization occurs, then a SAE can be recorded.

12.2.3 Causality Assessment

An Investigator causality assessment (Unrelated, Unlikely, Possible, Probable, or Definite) must be provided for all AEs (both serious and non-serious). This assessment must be recorded in the CRF and any additional SAE forms as appropriate. The definitions for the causality assessments appear below.

- Not related (unrelated): This relationship suggests that there is no association between the IMP and the reported event
- Unlikely related: This relationship suggests that the clinical picture is highly consistent with a cause other than the IMP but attribution cannot be made with absolute certainty and a relationship between the experimental drug and the AE cannot be excluded with complete confidence
- Possibly related: This relationship suggests that treatment with the IMP may have caused or contributed to the AE, i.e., the event follows a reasonable temporal sequence from the time of drug administration and/or follows a known response pattern to the experimental drug, but could also have been produced by other factors
- Probably related: This relationship suggests that a reasonable temporal sequence of the event with the IMP administration and the likely association of the event with the IMP. This will be based upon the known pharmacological action of the IMP, known or previously reported adverse

reactions to the IMP or class of drugs, or judgment based on the Investigator's clinical experience

- Definitely related: Temporal relationship to the IMP, other conditions (concurrent illness, concurrent medication reaction, or progression/expression of disease state) do not appear to explain event, corresponds with the known pharmaceutical profile, improvement on discontinuation, re-appearance on re-challenge

12.2.4 Lack of Efficacy or Worsening of NMOSD

Since TCZ or AZA treatment in relapsing NMOSD patients is not approved by Food and Drug Administration and CFDA, lack of efficacy is not definite if reported as an AE.The determination of relapses of NMOSD will be based on Investigator's assessment at every visit. If there is any uncertainty as to whether an event is due to disease progression, it should be reported as an adverse event.

12.2.5 Pregnancy

Female patients should take all appropriate precautions to avoid becoming pregnant during this study. In the event that a female patient becomes pregnant during the study or within 90 days after taking the last dose of study drug, she will be instructed to immediately inform the Investigator. The CRF should be completed by the Investigator within 24 hours after learning of the pregnancy. But pregnancy should not be recorded on the Adverse Event CRF. The Investigator should discontinue study drug and counsel the patient, discussing the risks of the pregnancy and the possible effects on the fetus. Monitoring of the patient should continue until conclusion of the pregnancy. Information on the health and well being of the baby will also be collected. Whether the drug is excreted in the semen is unknown. Therefore, pregnancy occurring in the partner of a male patient participating in the study should also be reported to the Investigator.

12.3 Reporting Requirements for Adverse Events

The Investigator must collect all AEs observed, obtained by direct questioning or volunteered from the trial patient.

To further evaluate the adverse events, detailed information about these events should be documented and reported that can be gathered within 24 hours. For non-serious AE, the reporting period starts following the first dose of the IMPs and continues through the last study visit including the Safety Follow-up Visit. AEs, particularly causally related, are to be followed until the event or sequelae resolve or are determined to be medically stable.

The Investigator should follow each adverse event until the event has resolved to baseline grade or better, the event is assessed as stable by the Investigator, the patient is lost to follow up, or the patient withdraws consent. Every effort should be made to follow all serious adverse events considered to be related to study drug or trial-related procedures until a final outcome can be reported.

During the study period, resolution of adverse events (with dates) should be documented on the Adverse Event CRF and in the patient's medical record to facilitate source data verification.

12.3.1 Reporting Requirements for Serious Adverse Events

For SAEs, the Investigator must be notified immediately or within 24 hours of the Investigator site becoming aware of the event, regardless of presumed relationship to the drugs. If the event meets criteria for a fatal or life-threatening SAE, the Investigator should notify Tianjin Medical University General Hospital immediately.

Emergency Medical Contacts
Center Number: [redacted]
Investigator's phone: [redacted]

The Investigator must complete, sign, and date the SAE pages, verify the accuracy of the information recorded on the SAE pages with the corresponding source documents, and send a copy via Email or fax to contact information provided below:
Email: [redacted]
Fax: [redacted]

Additional follow-up information, if required or available, should be emailed or faxed to Tianjin Medical University General Hospital within 24 hours of the Investigator becoming aware of this additional information. Follow-up information should be recorded on the SAE CRF and placed with the original SAE information and kept with the appropriate section of the original subject records and/or trial file.
For all SAEs the Investigator must provide the following:

- Appropriate and requested follow-up information in the time frame detailed above
- Causality of the serious event(s)
- Outcome of the serious event(s)
- Medical records and laboratory/diagnostic information

All the SAEs must be reported to the local Independent Ethics Committee by the Investigators.

12.4 Warnings and Precautions Management of Specific Adverse Events

12.4.1 Opportunistic Infections and Serious Infections

Physicians should exercise caution when considering the use of IMPs in patients with a history of recurring infections or with underlying conditions, which may predispose patients to infections.IMPs should not be administered to patients with active infection. The signs and symptoms of infection should be considered when evaluating a patient for a potential infection, especially for the possible reactivation of viral and other serious infections (e.g., Epstein−Barr virus or varicella-zoster virus). Patients must be instructed to contact the Investigator immediately when any symptoms suggesting infection appear, to ensure rapid evaluation and appropriate treatment. If a patient develops a serious infection, the IMPs will be interrupted until the infection is controlled. The investigator should consider the benefits and risks to the patient before resuming treatment.

12.4.2 Malignancies

Although no direct association of the IMPs with the malignancies was established, malignancies have been identified as a concern for IMPs in this trial. Once malignancies is suspected, diagnostic procedures will be started to determine if the patients have malignancies. IMPs must be discontinued for patients with malignancies.

12.4.3 Elevated Liver Enzymes

Elevated liver enzymes have been reported in the patients treated with tocilizumab or azathioprine in previous clinical trials and clinical practice. So monitoring liver enzymes is necessary for safety consideration. Treatment with tocilizumab or azathioprine in this trial will be adjusted or discontinuation according to Table 7.

Table 7 Risk Mitigation for Hepatic Enzyme Elevation

ALT or AST Values	Action for TCZ group	Action for AZA group
1-3 × ULN	Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. The dosage of TCZ will remain unchanged.	Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. The dosage of AZA will remain in the range of 2-3 mg/kg/day.
3-5 × ULN	Interrupt concomitant hepatotoxic drugs and	Interrupt concomitant hepatotoxic drugs and

	concomitant hepatoprotective drugs will be added. If ALT/AST levels return to baseline, the dosage of TCZ may be resumed to 8 mg/kg/4 weeks. For persistent increases in this range,the dosage of TCZ may be adjusted to 8 mg/kg/6 weeks.	concomitant hepatoprotective drugs will be added. If ALT/AST levels return to baseline, the dosage of AZA may be resumed to 2-3 mg/kg/day. For persistent increases in this range,the dosage of AZA may be adjusted to 1-2 mg/kg/day.
> 5 × ULN	Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. Laboratory tests should be repeated to confirm value. If confirmed, TCZ should be discontinued.	Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. Laboratory tests should be repeated to confirm value. If confirmed, AZA should be discontinued.

TCZ=tocilizumab; AZA=azathioprine; ALT=alanine amiotransferase; AST=aspartate aminotransferase; ULN= upper limit of normal range.

Patients withdrawn from the study because of elevated liver function test results must have repeat tests performed as clinically indicated until levels return to baseline values. If the patient's liver function test results have not returned to normal or to the patient's baseline level within 6 months (or sooner if deemed necessary by the Investigator), a specialist referral is recommended and an ultrasound should be considered.

12.4.4 Neutropenia

Decrease in neutrophil has been observed following treatment with TCZ or AZA in patients with other autoimmune diseases. The risk mitigation strategies for neutropenia is summarized in Table 8.

Table 8. Risk Mitigation for Neutropenia

ANC (cells/mm^3)	Action for TCZ group	Action for AZA group
> 1000	Maintain dose	Maintain dose
500− 1000	If neutropenia persists, prolong TCZ treatment from every 4 weeks to 6 weeks). When ANC increases to > 1000 cells/mm^3, resume IV TCZ every 4 weeks, as clinically appropriate	Interrupt concomitant suspicious toxic drugs. If ANC returns to baseline, the dosage of AZA may be resumed to 2-3 mg/kg/day. For persistent decreases in this range,the dosage of AZA may be adjusted to 1-2 mg/kg/day. Administration of Recombinant Human Granulocyte will be appropriate.
< 500	Discontinue TCZ permanently after repeat confirmation	Discontinue AZA permanently after repeat confirmation. Administration of Recombinant Human Granulocyte will be appropriate.

ANC= absolute neutrophil count; TCZ= tocilizumab; AZA=azathioprine.

12.5 Follow-up Period for Adverse Events

The Investigator must collect all AEs observed, obtained by direct questioning or volunteered from the trial patient. Withdrawal due to an AE or SAE must be clearly differentiated from withdrawal due to other reasons.

For non-serious AE, the reporting period starts following the first dose of the IMPs (Day 1, Visit 2) and continues through the last study visit including the safety Follow-Up Visit. AEs, particularly causally related, are to be followed until the event or sequelae resolve or are determined to be medically stable.

For SAEs the reporting begins following the patient's signing of the ICF (providing consent to participate in the trial) and continues through 12 weeks after the last IMP dose. No time limit on reporting SAEs that are thought to be possibly or probably or definitely related to the IMP.

13.STATISTICAL METHOD AND PLANNED ANALYSES

13.1 General considerations

TANGO is a randomized, multi-center, open-label, two-arm, time-to-event trial comparing tocilizumab and azathioprine in patients with NMOSD. At completion of the trial, the School of Public Health of Tianjin Medical University will analyze the trial data. Further elaboration of statistical issues is provided in the Statistical Analysis Plan (SAP).

The School of Public Health of Tianjin Medical University will be responsible for data collection and editing, reviewing and validating all the information in the CRFs, statistical analysis, and generation of the clinical report.

Prior to locking the database, all data editing will be complete and decisions regarding the evaluability of all patient data for inclusion in the statistical analysis for the Per-Protocol (PP) Set will be made. The rationale for excluding any data from the statistical analyses will be prospectively defined, and classification of all or part of a patient's data as non-evaluable will be completed and documented before the database is locked and before the statistical analysis is begun. The statistical analysis will not begin until the entire database is locked and signed off.

The School of Public Health of Tianjin Medical University will perform the statistical analysis of the data derived from this trial. The analysis will be performed using the SAS® statistical software system Version 9.4.

All summary statistics will be computed and displayed by treatment group and scheduled assessment time. Summary statistics for continuous variables will minimally include n, mean, standard deviation, minimum, median, and maximum. For categorical variables, frequencies and percentages will be presented. Graphical displays will be provided as appropriate.

13.2 Determination of Sample Size

This is a randomized, open-label, multi-center, parallel-controlled trial to evaluate tocilizumab in NMO patients with a primary endpoint of time to first relapse. As such, the trial is based on observing relapse events.

The sample size and power calculation assumptions for this time to first event trial are as follows:

- Log-rank test for comparison of tocilizumab to azathioprine
- 1:1 randomization (tocilizumab:azathioprine)
- Power 80%
- Two-sided 5% level of significance
- Drop-out rate 10%
- Accrual period of approximately 60 weeks
- Relapse-free rate of 85% for tocilizumab and 60% for azathioprine at 12 months (hazard ratio = 0.318)

13.3 Analyses Sets

13.3.1 Full Analysis Set (FAS)

The population on which primary, secondary and other efficacy analyses will include all patients who are randomized to treatment and who have received at least 1 dose of the experimental drug. Patients will be compared for efficacy according to the treatment they were randomized to receive, irrespective of the treatment they actually received.

13.3.2 Per-Protocol (PP) Set

The PP Set is a subset of the FAS, excluding patients with major protocol deviations. The PP population will include all patients who:

- Have no major protocol deviations or key inclusion/exclusion criteria deviations that might

potentially affect efficacy
- Patients who took at least 80% of the required treatment doses while they were in the Study Period (for patients who have relapses any dosing after the relapse will not be included in this calculation)
- Patients who received the treatment of IMP as monotherapy

The PP population will be fully described in the statistical analysis plan, and patients identified prior to database lock.

13.3.3 Safety Set

Safety analyses will be performed on the Safety Set Population. The Safety Population includes all patients who receive at least 1 dose of the experimental drug. Patients will be compared for safety according to the treatment they actually received.

Patients who have signed informed consent but are not treated in the trial are not in the Safety Population. However, if these patients report AEs or SAEs, these events will be summarized separately in tables and listings as appropriate.

13.4 Demographics and Baseline Characteristics

All demographic and baseline characteristics information will be summarized using the following sets: Full Analysis, PP, and Safety Sets. No formal hypothesis testing will be performed. Summary statistics will be presented by treatment group and overall.
Medical history and medical history related to NMOSD will be summarized by treatment group. Listings related to medical history will also be produced.

13.5 Subject Disposition and Treatment Compliance

The number of patients screened, randomized, treated, completing the trial, and included in the safety and efficacy analysis sets will be tabulated by counts and percentage of patients by treatment group and overall. Reasons for any patient withdrawals will be provided.
Treatment compliance with the experimental drug will be summarized using descriptive statistics. The extra usage of the experimental drug for patients who are treated with PE during the trial will be summarized and listings will be produced.

13.6 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by treatment group. Listings of prior and concomitant medications will be produced. Supportive immunosuppressants are allowed during the trial under certain restrictions. The following immunosuppressants are allowed either as mono-therapy or in combination: corticosteroids, AZA, MMF, methotrexate, tacrolimus, cyclosporine and cyclophosphamide. Thus, supportive IST will be summarized by treatment group. Listings of supportive immunosuppressants will be produced. Changes in immunosuppressants during the trial will be summarized.

For patients with relapses, the use of IVMP and IVIG will be summarized by treatment group. Listings of IVMP and IVIG usage by patients with relapses will be produced. Medications will be coded using the World Health Organization Drug Dictionary (WHO Drug).

13.7 Efficacy Analyses

Analyses will be produced for the Study Period in order to compare the tocilizumab group with azathioprine group. The analyses will include efficacy and safety analyses. Efficacy analyses will be performed on the FAS population as well as the PP population. The analysis are prespecified by two subgroups of the patients: with and without concomitant autoimmune diseases.

13.8 Primary Efficacy Endpoint

The primary efficacy endpoint is time to first relapse. The trial will be considered to have met its primary efficacy objective if a statistically significant difference is observed between the tocilizumab treatment group and the azathioprine group. Confidence intervals and p-values will be presented. Hazard ratio and risk reduction will be summarized. A sensitivity analysis will be performed on time to first relapse (as identified by the Investigator) using a log-rank test including strata for the randomization stratification variables. In addition, a sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator, and randomization stratification variables as the only covariates in the model. A second sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator, randomization stratification variables, and region as the only covariates in the model. Region will be defined based on the sites for the trial and will include North America, South America, Europe, Asia-Pacific, and Other as applicable. In the event that the number of patients in some regions is too small to permit modeling, the smaller regions will be pooled together.

An additional sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator and randomization stratification variables as covariates and will also include withdrawals due to AEs as outcomes events (i.e., relapses).
An additional sensitivity of the treatment groups for the primary endpoint will use a log-rank test including strata for the randomization stratification variables for the FAS patients with a follow-up assessment (i.e., FAS patients without a follow-up assessment will be excluded from the analysis).
An additional sensitivity comparison of the primary endpoint will use a Cox proportional hazards regression model with treatment group indicator and randomization stratification variables as covariates only including the relapses asse
ssed within 24 hours. Confirmed relapses assessed after 24 hours will be censoring events based on the patient's last date in the study.

In order to account for relapses that are not evaluated according to the protocol (i.e., within 24 hours of the onset of relapse sign and symptoms), a Cochran-Mantel-Haenszel (CMH) test will be used assessing tocilizumab and azathioprine treatment and patient evaluated within 24 hours (yes, no) for patients with confirmed relapses. Summaries of the severity of the confirmed relapses versus patient evaluated within 24 hours (yes, no) by treatment group will also be produced.
Kaplan-Meier curves for both treatment groups will be produced. Likewise, Kaplan-Meier curves for the strata within each treatment group will be produced.

We pre-specify that the patients in each group will be categorized into two subgroups, the patients with and without concomitant autoimmune diseases.

13.9 Secondary Efficacy Analysis

During the Study Period, Baseline is defined as the last available assessment prior to treatment for all patients regardless of their treatment group.
Unless otherwise specified, the secondary efficacy analyses will use the available data from the Study Period. Hypothesis testing comparing tocilizumab treatment with azathioprine treatment for the secondary efficacy analyses will be performed:

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

The time to onset of confirmed disability progression (12-week confirmation [days]) is defined as the time from baseline to the onset of the first disability progression that is confirmed at the next regularly scheduled visit ≥ 12 weeks after the initial disability progression. If the patient has an infection, dosing may not occur on the Day 1 visit.

Baseline for the time to onset of confirmed disability is the date of randomization, independent of the date of first dosing. For example, a subject with delayed dosing may receive the first dose 2 weeks after the baseline visit; if an EDSS score is recorded between the randomization date and the date of the first dose, this value will be considered for the date of initial disability progression. Disability progression is defined as an increase of ≥ 1.0 point from baseline EDSS score if the baseline EDSS value is ≤ 5.5 points (inclusive) or an increase of ≥ 0.5 points if the baseline EDSS value is > 5.5 points. Assessments within 30 days after a protocol-defined relapse will not be used for confirmation of confirmed disability progression. The non-confirmatory EDSS assessments (between the initial disability progression and the confirmation of disability progression should also fulfill the requirements of the progression. Otherwise, the initial disability progression is not confirmed.

The baseline EDSS value is the average score of the EDSS assessment at screening and baseline (Day 1 visit) up to and including the date of randomization. If one of the values are missing, the non-missing values will be used as baseline.

A blinded rater at each study site will assess EDSS for all patients at the site at screening, baseline, every 12 weeks (regularly scheduled visit) during the treatment period of the study, during the safety follow-up period, at any unscheduled visits, and at withdrawal-from-treatment and end-of-study visits. Additional EDSS assessments for individual patients may be requested between visits (i.e., during an NMOSD relapse).

13.10 Other Efficacy Endpoints

Additional exploratory efficacy measures including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

13.11 Safety Analyses

Safety analyses will be performed on the Safety Population. The Safety Population includes all patients who receive at least 1 dose of the experimental drug. Patients will be compared for safety according to the treatment they actually received.

During the Study Period, Baseline is defined as the last available assessment prior to treatment for all patients regardless of their treatment group.

All AEs and other safety information including untreated patients collected after the signing of informed consent will be reported in listings, as applicable.

AEs will be summarized by incidence, terms of Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, seriousness, severity, relationship to treatment, and by treatment group. Concomitant medications will be summarized by treatment group.
Changes from Baseline in vital signs, laboratory assessments (chemistry and hematology) will be

summarized by treatment group. Likewise, shift tables (L [low], N [normal], H [high]) by treatment group will be produced for clinical laboratory tests and pregnancy tests will be summarized in patient listings.

13.11.1 Physical Examinations and Vital Signs

Physical examinations will be summarized by visit and treatment group. Vital signs (systolic and diastolic BP, temperature, and sitting or supine HR), height, and weight and changes from baseline in vital signs (including height and weight) will be summarized by visit and by treatment group. Listings of physical exams and vital signs will be produced.

13.11.2 Laboratory Assessments

Changes from Baseline in laboratory assessments (chemistry and hematology,) will be summarized by visit and treatment group. Likewise, shift tables (L [low], N [normal], H [high]) by visit and treatment group will be produced for clinical laboratory tests. Listings of laboratory data will be produced.

13.11.3 Adverse Events

SAEs occurring from the signing of informed consent and prior to the initiation of tocilizumab or azathioprine treatment (pre-treatment SAEs) will be summarized by treatment group.

Treatment-emergent AEs (TEAEs) are AEs that onset after the start of treatment in the trial. TEAEs will be summarized by incidence, terms of Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, seriousness, severity, relationship to treatment, and by treatment group. SAEs will be summarized by treatment group. TEAEs and SAEs will be summarized by gender and treatment group, by race and treatment group, and by region (of the world) and treatment group.

AEs and general medical/surgical histories will be coded by Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0, corresponding to primary SOC of the Medical Dictionary for Regulatory Activities (MedDRA) (version 22.0).

13.11.4 Other Safety Endpoints

ECG results will be summarized in patient listings. Pregnancy tests will be summarized in patient listings.

13.12 Significance Levels

For all analyses, the tocilizumab treated group will be compared to the azathioprine group and all hypothesis testing will be two-sided and performed at the 0.05 level of significance, unless otherwise specified. Estimates of treatment effect on efficacy parameters will be accompanied by two-sided 95% confidence intervals for the effect size.

13.13 Missing or Invalid Data

For secondary and tertiary efficacy analyses, missing post-Baseline efficacy and safety data will not be imputed unless indicated in the described analysis in the SAP.

14. ETHICAL CONSIDERATIONS

14.1 Compliance with Laws and Regulations

This study will be conducted in full conformance with the ICH E6 guideline for Good Clinical Practice and the principles of the Declaration of Helsinki, or the laws and regulations of the country in which the research is conducted, whichever affords the greater protection to the individual. The study will comply with the requirements of the ICH E2A guideline (Clinical Safety Data Management: Definitions and Standards for Expedited Reporting). Studies conducted in China will comply with National Medical Products Administration regulations and applicable local laws.

14.2 Institutional Review Board (IRB) and Institutional Ethics Committee (IEC)

This protocol, the Informed Consent Forms (ICF), any information to be given to the patient, and relevant supporting information must be approved or given a favorable opinion in writing by an IRB or IEC as appropriate. The Investigator must submit a copy of the written approval to the Tianjin Medical University General Hospital before he or she can enroll any subject into the trial.

The Principal Investigator (PI) is responsible for providing written summaries of the status of the study to the IRB/IEC annually or more frequently in accordance with the requirements, policies, and procedures established by the IRB/EC. Investigators are also responsible for promptly informing the IRB/IEC of any protocol amendments.

Progress reports and notifications of serious adverse drug reactions will be provided to the IRB or IEC according to local regulations and guidelines.

14.3 Written Informed Consent

The PI(s) at each trial site will ensure that the patient is given full and adequate oral and written information about the nature, purpose, possible risk and benefit of the trial. The patient should be given the opportunity to ask detailed questions and allowed time to consider the benefits and risks. Patients must be informed that they are free to discontinue from the trial at any time.

The patient’s signed and dated informed consent must be obtained before conducting the trial. All signed and dated Consent Forms must remain in each patient’s study file or in the site file and must be available for verification by study monitors at any time.

The Consent Forms should be revised whenever there are changes to study procedures or when new information becomes available that may affect the willingness of the patient to participate.The final revised ICF must be approved by IRB/IEC. Patients must be re-consented to the most current version of the ICF (or to a significant new information/findings addendum in accordance with applicable laws and IRB/IEC policy) during their participation in the study.

15. RETENTION AND INSPECTION OF RECORDS

The IRB/IEC will be allowed to inspect facilities and records relevant to this study. The Investigator agrees to allow IRB/IEC to monitor the IMPs storage area, experimental drug stocks, experimental drug accountability records, patient charts and trial source documents, and other records relative to trial conduct.The Investigator must maintain adequate and accurate records to enable the conduct of the study to be fully documented and the study data to be subsequently verified.all documentation relating to the trial will be kept for at least 5 years following the discontinuance of the test article for investigation.

16. CONFIDENTIALITY

Patient medical information obtained by this study is confidential. The Investigators maintain confidentiality standards by coding each patient enrolled in the study through assignment of a unique patient identification number. This means that patient names are not included in data sets. Medical information may be given to a patient's personal physician or other appropriate medical personnel responsible for the patient's welfare, for treatment purposes.

Data generated by this study must be available for inspection upon request by representatives of the collaborators, and the IRB/EC for each study site, as appropriate.

17. REFERENCES

1. Wingerchuk DM, Lennon VA, Lucchinetti CF, Pittock SJ, Weinshenker BG. The spectrum of neuromyelitis optica. The Lancet Neurology 2007;6:805-15.
2. Lennon VA, Wingerchuk DM, Kryzer TJ, et al. A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. Lancet 2004;364:2106-12.
3. Wingerchuk DM, Banwell B, Bennett JL, et al. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. Neurology 2015;85:177-89.
4. Pandit L, Asgari N, Apiwattanakul M, et al. Demographic and clinical features of neuromyelitis optica: A review. Multiple sclerosis 2015;21:845-53.
5. Mealy MA, Wingerchuk DM, Greenberg BM, Levy M. Epidemiology of neuromyelitis optica in the United States: a multicenter analysis. Archives of neurology 2012;69:1176-80.
6. Bukhari W, Prain KM, Waters P, et al. Incidence and prevalence of NMOSD in Australia and New Zealand. Journal of neurology, neurosurgery, and psychiatry 2017;88:632-8.
7. Flanagan EP, Cabre P, Weinshenker BG, et al. Epidemiology of aquaporin-4 autoimmunity and neuromyelitis optica spectrum. Annals of neurology 2016;79:775-83.
8. ZhangBao J, Zhou L, Li X, et al. The clinical characteristics of AQP4 antibody positive NMO/SD in a large cohort of Chinese Han patients. Journal of neuroimmunology 2017;302:49-55.
9. Nagaishi A, Takagi M, Umemura A, et al. Clinical features of neuromyelitis optica in a large Japanese cohort: comparison between phenotypes. Journal of neurology, neurosurgery, and psychiatry 2011;82:1360-4.
10. Sato DK, Lana-Peixoto MA, Fujihara K, de Seze J. Clinical spectrum and treatment of neuromyelitis optica spectrum disorders: evolution and current status. Brain pathology 2013;23:647-60.
11. Levy M. The ethics of placebo controlled clinical trials in NMO - A balance of risks. Multiple sclerosis and related disorders 2015;4:512-4.
12. Papadopoulos MC, Bennett JL, Verkman AS. Treatment of neuromyelitis optica: state-of-the-art and emerging therapies. Nature reviews Neurology 2014;10:493-506.
13. Kimbrough DJ, Fujihara K, Jacob A, et al. Treatment of Neuromyelitis Optica: Review and Recommendations. Multiple sclerosis and related disorders 2012;1:180-7.
14. Stellmann JP, Krumbholz M, Friede T, et al. Immunotherapies in neuromyelitis optica spectrum disorder: efficacy and predictors of response. Journal of neurology, neurosurgery, and psychiatry 2017;88:639-47.
15. Mealy MA, Wingerchuk DM, Palace J, Greenberg BM, Levy M. Comparison of relapse and treatment failure rates among patients with neuromyelitis optica: multicenter study of treatment efficacy. JAMA neurology 2014;71:324-30.
16. Sellner J, Boggild M, Clanet M, et al. EFNS guidelines on diagnosis and management of neuromyelitis optica. European journal of neurology 2010;17:1019-32.
17. Trebst C, Jarius S, Berthele A, et al. Update on the diagnosis and treatment of neuromyelitis optica: recommendations of the Neuromyelitis Optica Study Group (NEMOS). Journal of neurology 2014;261:1-16.
18. Elsone L, Kitley J, Luppe S, et al. Long-term efficacy, tolerability and retention rate of azathioprine in 103 aquaporin-4 antibody-positive neuromyelitis optica spectrum disorder patients: a multicentre retrospective observational study from the UK. Multiple sclerosis 2014;20:1533-40.
19. Costanzi C, Matiello M, Lucchinetti CF, et al. Azathioprine: tolerability, efficacy, and predictors of benefit in neuromyelitis optica. Neurology 2011;77:659-66.
20. 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. Journal of neurology 2017;264:2003-9.
21. Xu Y, Wang Q, Ren HT, et al. Comparison of efficacy and tolerability of azathioprine, mycophenolate mofetil, and cyclophosphamide among patients with neuromyelitis optica spectrum disorder: A prospective cohort study. Journal of the neurological sciences 2016;370:224-8.
22. Chen H, Qiu W, Zhang Q, et al. Comparisons of the efficacy and tolerability of mycophenolate mofetil and azathioprine as treatments for neuromyelitis optica and neuromyelitis optica spectrum

disorder. European journal of neurology 2017;24:219-26.
23. Chen D, Lian F, Yuan S, et al. Association of thiopurine methyltransferase status with azathioprine side effects in Chinese patients with systemic lupus erythematosus. Clinical rheumatology 2014;33:499-503.
24. Uzawa A, Mori M, Kuwabara S. Cytokines and chemokines in neuromyelitis optica: pathogenetic and therapeutic implications. Brain pathology 2014;24:67-73.
25. Uzawa A, Mori M, Arai K, et al. Cytokine and chemokine profiles in neuromyelitis optica: significance of interleukin-6. Multiple sclerosis 2010;16:1443-52.
26. Sato DK, Callegaro D, de Haidar Jorge FM, et al. Cerebrospinal fluid aquaporin-4 antibody levels in neuromyelitis optica attacks. Annals of neurology 2014;76:305-9.
27. Takeshita Y, Obermeier B, Cotleur AC, et al. Effects of neuromyelitis optica-IgG at the blood-brain barrier in vitro. Neurology(R) neuroimmunology & neuroinflammation 2017;4:e311.
28. Uchida T, Mori M, Uzawa A, et al. Increased cerebrospinal fluid metalloproteinase-2 and interleukin-6 are associated with albumin quotient in neuromyelitis optica: Their possible role on blood-brain barrier disruption. Multiple sclerosis 2017;23:1072-84.
29. Chihara N, Aranami T, Sato W, et al. Interleukin 6 signaling promotes anti-aquaporin 4 autoantibody production from plasmablasts in neuromyelitis optica. Proceedings of the National Academy of Sciences of the United States of America 2011;108:3701-6.
30. Shirota Y, Yarboro C, Fischer R, Pham TH, Lipsky P, Illei GG. Impact of anti-interleukin-6 receptor blockade on circulating T and B cell subsets in patients with systemic lupus erythematosus. Annals of the rheumatic diseases 2013;72:118-28.
31. Smolen JS, Landewe R, Bijlsma J, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Annals of the rheumatic diseases 2017;76:960-77.
32. Kieseier BC, Stuve O, Dehmel T, et al. Disease amelioration with tocilizumab in a treatment-resistant patient with neuromyelitis optica: implication for cellular immune responses. JAMA neurology 2013;70:390-3.
33. Ayzenberg I, Kleiter I, Schroder A, et al. Interleukin 6 receptor blockade in patients with neuromyelitis optica nonresponsive to anti-CD20 therapy. JAMA neurology 2013;70:394-7.
34. Araki M, Matsuoka T, Miyamoto K, et al. Efficacy of the anti-IL-6 receptor antibody tocilizumab in neuromyelitis optica: a pilot study. Neurology 2014;82:1302-6.
35. Ringelstein M, Ayzenberg I, Harmel J, et al. Long-term Therapy With Interleukin 6 Receptor Blockade in Highly Active Neuromyelitis Optica Spectrum Disorder. JAMA neurology 2015;72:756-63.
36. Weinshenker BG, Barron G, Behne JM, et al. Challenges and opportunities in designing clinical trials for neuromyelitis optica. Neurology 2015;84:1805-15.
37. Watanabe S, Misu T, Miyazawa I, et al. Low-dose corticosteroids reduce relapses in neuromyelitis optica: a retrospective analysis. Multiple sclerosis 2007;13:968-74.
38. Balcer LJ, Raynowska J, Nolan R, et al. Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. Multiple sclerosis 2017;23:734-47.
39. Raftopoulos R, Hickman SJ, Toosy A, et al. Phenytoin for neuroprotection in patients with acute optic neuritis: a randomised, placebo-controlled, phase 2 trial. The Lancet Neurology 2016;15:259-69.
40. Pittock SJ, Lennon VA, McKeon A, et al. Eculizumab in AQP4-IgG-positive relapsing neuromyelitis optica spectrum disorders: an open-label pilot study. The Lancet Neurology 2013;12:554-62.
41. Zhang C, Tian DC, Yang CS, et al. Safety and Efficacy of Bortezomib in Patients With Highly Relapsing Neuromyelitis Optica Spectrum Disorder. JAMA neurology 2017;74:1010-2.
42. Waters PJ, McKeon A, Leite MI, et al. Serologic diagnosis of NMO: a multicenter comparison of aquaporin-4-IgG assays. Neurology 2012;78:665-71; discussion 9.
43. Wang J, Li X, Zhang DQ, et al. Quantitative analysis of aquaporin-4 antibody in longitudinally extensive transverse myelitis. Journal of neuroimmunology 2015;278:26-9.

18. APPENDICES

Appendix 1: 2015 DIAGNOSTIC CRITERIA FOR NMOSD
Appendix 2: NEUROIMAGING CHARACTERISTICS OF NMOSD
Appendix 3. TPMT SINGLE NUCLEOTIDE POLYMORPHISM
Appendix 4: KURTZKE EXPANDED DISABILITY STATUS SCALE (EDSS)
Appendix 5: VISUAL ACUITY TEST
Appendix 6: CLINICAL LARORATORY TESTS
Appendix 7. COMMON TERMINOLOGY CRITERIA FOR ADVERSE EVENTS (CTCAE 5.0)

APPENDIX 1. 2015 DIAGNOSTIC CRITERIA FOR NMOSD

NMOSD diagnostic criteria for adult patients
Diagnostic criteria for NMOSD with AQP4-IgG 1. At least 1 core clinical characteristic 2. Positive test for AQP4-IgG using best available detection method (cell-based assay strongly recommended) 3. Exclusion of alternative diagnoses
Diagnostic criteria for NMOSD without AQP4-IgG or NMOSD with unknown AQP4-IgG status 1. At least 2 core clinical characteristics occurring as a result of one or more clinical attacks and meeting all of the following requirements: a. At least 1 core clinical characteristic must be optic neuritis, acute myelitis with LETM, or area postrema syndrome b. Dissemination in space (2 or more different core clinical characteristics) c. Fulfillment of additional MRI requirements, as applicable 2. Negative tests for AQP4-IgG using best available detection method, or testing unavailable 3. Exclusion of alternative diagnoses
Core clinical characteristics 1. Optic neuritis 2. Acute myelitis 3. Area postrema syndrome: episode of otherwise unexplained hiccups or nausea and vomiting 4. Acute brainstem syndrome 5. Symptomatic narcolepsy or acute diencephalic clinical syndrome with NMOSD-typical diencephalic MRI lesions 6. Symptomatic cerebral syndrome with NMOSD-typical brain lesions
Additional MRI requirements for NMOSD without AQP4-IgG and NMOSD with unknown AQP4-IgG status 1. Acute optic neuritis: requires brain MRI showing normal findings or only nonspecific white matter lesions, OR optic nerve MRI with T2-hyperintense lesion or T1-weighted gadolinium-enhancing lesion extending over >1/2 optic nerve length or involving optic chiasm 2. Acute myelitis: requires associated intramedullary MRI lesion extending over ≥ 3 contiguous segments (LETM) OR ≥ 3 contiguous segments of focal spinal cord atrophy in patients with history compatible with acute myelitis 3. Area postrema syndrome: requires associated dorsal medulla/area postrema lesions 4. Acute brainstem syndrome: requires associated periependymal brainstem lesions

Abbreviations: AQP4 = aquaporin-4; IgG = immunoglobulin G; LETM = longitudinally extensive transverse myelitis lesions; NMOSD = neuromyelitis optica spectrum disorders.

APPENDIX 2. NEUROIMAGING CHARACTERISTICS OF NMOSD

Neuroimaging characteristics of NMOSD
Spinal cord MRI, acute
LETM lesion associated with acute TM
• Increased signal on sagittal T2-weighted (standard T2-weighted, proton density, or STIR sequences) extending over 3 or more complete vertebral segments • Central cord predominance (more than 70% of the lesion residing within the central gray matter) • Gadolinium enhancement of the lesion on T1-weighted sequences (no specific distribution or pattern of enhancement is required)
Other characteristic features that may be detected
• Rostral extension of the lesion into the brainstem • Cord expansion/swelling • Decreased signal on T1-weighted sequences corresponding to region of increased T2-weighted signal
Optic nerve MRI
Unilateral or bilateral increased T2 signal or T1 gadolinium enhancement within optic nerve or optic chiasm; relatively long lesions (e.g., those extending more than half the distance from orbit to chiasm) and those involving the posterior aspects of the optic nerves or the chiasm are associated with NMO
Cerebral MRI: NMOSD-typical brain lesion patterns (increased signal on T2-weighted MRI sequences unless otherwise noted)
• Lesions involving the dorsal medulla (especially the area postrema), either small and localized, often bilateral, or contiguous with an upper cervical spinal cord lesion • Periependymal surfaces of the fourth ventricle in the brainstem/cerebellum • Lesions involving the hypothalamus, thalamus, or periependymal surfaces of the third ventricle • Large, confluent, unilateral, or bilateral subcortical or deep white matter lesions • Long (1/2 of the length of the corpus callosum or greater), diffuse, heterogeneous, or edematous corpus callosum lesions • Long corticospinal tract lesions, unilateral or bilateral, contiguously involving internal capsule and cerebral peduncle • Extensive periependymal brain lesions, often with gadolinium enhancement

Abbreviations: LETM = longitudinally extensive transverse myelitis lesions; NMOSD = neuromyelitis optica spectrum disorders; STIR = short tau inversion recovery.

APPENDIX 3. TPMT SINGLE NUCLEOTIDE POLYMORPHISM

Testing	TPMT*2,*3A,*3B,*3C single nucleotide polymorphism				
Methods	PCR and sequencing				
Instrument	ABI 3500xL Dx				
Results	Gene	Polymorphism	Locus	Genotype	Mutation
	TPMT	*2	c.238	G/G	NO
		*3A	c.460/c.719	G/G AND A/A	NO
		*3B	c.460	G/G	NO
		*3C	c.719	A/A	NO

Results:
No TPMT mutation was detected in the allelic locus of this patient. According to reported references, the side effects may be lower after using azathioprine.

Note:
The three mutation locus of the 5th exon G238C, the 7th exon G460A and the 10th exon A7I9G of the common TPMT (thiopurine methyltransferase) gene are mononucleotide polymorphisms. Mutation of the 5th exon G238C results in a mutation in the allele TPMT*2.Mutation of the 7th exon G460A results in a mutation in the allele TPMT*3B, and mutation in the 10th exon A719G results in an allele mutation of the gene TPMT*3C, mutation of the 7th exon G460A and the 10th exon A719G resulted in a mutation of the allele TPMT*3A. Among the above alleles, according to the reported paper, the unmutated genotype, which is normal, has a lower toxic effect of azathioprine. And the heterozygous mutant adopts the toxic side effects of the steroids, and the homozygous mutant uses the ridges. The drug has high side effects. This test is used for those patients who lack the TPMT gene of the dredging methyltransferase. It provides reference for the dosage of sedative drugs at the time of treatment. The report results are only responsible for the specimens in this test and are used for medical research.

APPENDIX 4. KURTZKE EXPANDED DISABILITY STATUS SCALE (EDSS)

(1) Visual function examination

Visual function	L	R	note
Vision (after correction)			The visual acuity score is obtained by reading the Snellen eye chart at a distance of 5 meters from the patient standing. The recognition error cannot exceed one.
Field of view (coarse measurement)			0 = normal 1 = only physical signs, visual field defects only found during formal visual field examination 2 = moderate, the patient can feel the visual field defect, and the visual field examination reveals that the patient is not completely hemian 3=severe, completely ametropic or blind
Visual field defect (perimeter results)			0=none 1 = small, only found in formal inspections 2=large, the patient has a complaint
Optic papilla pale			0=none 1=Yes

0. Normal
1. The papillary pale and / or small blind spots and / or poor eye is less than 1.0 but better than 0.67
2. poor eye corrected visual acuity 0.67-0.34
3. The poor eye has a large blind spot and / or moderate visual field defect and / or 0.33-0.21
4. The poor eye has obvious visual field defect and / or visual acuity 0.2-0.1; 3 plus good eye is less than or equal to 0.33
5. The poor eye is less than 0.1; 4 plus the good eye is less than or equal to 0.33
6.5 plus good eye is less than or equal to 0.33

FS=

FS score correction:：

FS	6	5	4	3	2	1
After conversion FS	4	3	3	2	2	1

（2）brain stem function

Cranial nerve	score	note
Eye movement disorder		0=None 1 = only signs, no complaints (obscure, double vision) 2 = mild, the patient may have a slight EOM disorder that is aware of it, or a clinical examination reveals an incomplete dyskinesia of the eye, but the patient is unaware 3 = moderate, the patient can be aware of the obvious eye movement incomplete movement disorder, or any movement can not fully move when looking at a certain direction 4=Severe, any movement can not fully move when looking in more than one direction
Nystagmus		0=none 1 = only signs or mild: nystagmus when gazing, equivalent to brain stem FS =

		1 2 = moderate, continuous nystagmus at 30° horizontal or vertical gaze, but no nystagmus in situ, patients may or may not complain 4=Severity: In-situ continuous nystagmus, or gross continuous nystagmus in any direction with limited vision, complete internuclear ophthalmoplegia with continuous nystagmus during abduction, vibrating hallucinations
Trigeminal nerve injur		0=none 1 = only signs 2 = mild: clinical numbness can be detected and the patient has a complaint 3 = moderate: trigeminal nerve 1, 2 or 3 branch area tip / blunt discernment disorder, or trigeminal neuralgia (at least 1 episode in the past 24h) 4=Severe: One side or both sides of the face are completely pointed/bluntly dysfunctional
Facial paralysis		0 = no 1 = only signs 2 = mild: clinically detectable facial paralysis, and the patient has a complaint 3=Moderate: Incomplete facial paralysis, such as incomplete closure of the eyelids 4=Severe: unilateral or bilateral complete facial paralysis, resulting in increased ocular fissure, redness of the conjunctiva, difficulty in drinking water
Hearing disorder		0=none 1 = only physical signs: unilateral or bilateral ear hearing finger friction is weakened, Weber experiment is not centered, but the patient has no complaint 2 = mild: same as 1, but the patient has a complaint 3=Moderate: One-sided or two-sided ears cannot hear finger friction, and hear a few whispers 4=Severe: almost all whispering numbers are heard
Dysphonic disorder		0=none 1 = only signs 2 = mild: clinically detectable enough sound disorder, the patient has a complaint 3=Moderate: obvious enough obstacles in daily conversation, affecting understanding 4=Severity: a language that cannot be understood 5=Word can't
Swallowing disorder		0=none 1 = only signs 2 = mild: swallowing thin fluid barrier 3 = moderate: swallowing liquid and solid food barriers 4 = Severe: continuous swallowing disorder, need a muddy diet 5=Can’t swallow
Other cranial nerve function		0 = normal 1 = only signs 2 = mild: clinically detectable, patients have complaints 3=moderate 4=severe

0. Normal
1. Only signs
2. Moderate nystagmus and / or moderate ocular dyskinesia and / or other mild cranial nerve disorders
3. Severe nystagmus and/or severe ocular dyskinesia and/or other moderate cranial nerve disorders
4. Severe enough dysfunction and / or other severe cranial nerve disorders
5. Can't swallow or talk

（3）Cone beam function

Reflection	L	R	note
Biceps reflex			0 = disappear 1 = weaken 2=Normal 3=accentuation 4=Do not continue to fight 5=Continuous clonus
Triceps reflex			
Periosteal reflex			
Knee reflex			
Ankle reflex			
Plantar reflex			0=bend， 1=neutral， 2=stretch
Abdominal reflex			0= normal, 1 = weakened, 2 = disappeared
Palmormental reflex			0 = none, 1 = yes
Muscle strength	L	R	note
Shoulder			0 = no muscle contraction 1 = visible muscle contraction, but can't drive joint activity 2 = horizontal movement, but can't resist gravity 3 = can resist gravity, but can't resist resistance (can lift out of bed) 4 = can resist resistance but limited 5 = normal The score of the weakest muscle strength in each group was the improvement score.
Elbow flexor			
Elbow extensor			
Flexors of hand/finger			
Extensors of hand/finger			
Flexor hip muscle			
Flexor muscles			
Extensor muscles of knee			
Flexors of foot/toe			
Extensor foot/toe muscle			
Functional testing	L	R	note
Upper limb paralysis test			0= Limbless decline， 1 = slight decrease, 2 = significant decrease

Lower limb paralysis test			0 = no drop, 1 = slight drop, 2 = significant drop, 3 = only one leg can be lifted at a time, 4 = no leg can be lifted.
Heel walking			0 = normal, 1 = restricted, 2 = impossible
Walking on tiptoe			
Voronin hop			0 = normal, 1 = 6 - 10, 2 = 1 - 5, 3 = impossible
Muscular tension	L	R	note
Upper limb			0 = normal 1 = mild: slight increase in muscle tension 2 = moderate: moderate increase in muscle tension, can be overcome, unlimited movement 3 = severe: severe increase in muscle tension, difficult to overcome, restricted movement 4 = contracture
The legs			
Gait			0 = normal 1 = almost imperceptible 2 = obvious: slightly affecting function 3 = continuous swing, seriously affecting function

Normal
Only physical signs
2. Mild disability: patients complained of fatigue weakness and/or group 1-2 muscle strength grade 4
3. Mild to moderate quadriplegia or hemiplegia; muscle strength grade 4 above group 2; or muscle strength grade 3 of group 1-2 (against gravity); or severe single paralysis (muscle strength less than grade 2 in group 1).
4. Severe quadriplegia or hemiplegia: 2 limbs of grade 2; or (> 3 limbs of grade 3); or a limb of grade 0-1 of grade
5. All lower limbs of muscle of grade 0-1; or (> 3 limbs of muscle strength (< 2); and/or hemiplegia of grade
6. Quadriplegia: all limbs of muscle strength of grade 0-1

（4）Cerebellar function

Ataxia		Fraction	note
Ataxia of limbs	Tremor		0 = No 1 = only signs 2 = mild: tremor or clumsiness, mild dysfunction 3 = moderate: tremor or clumsiness affects function in all respects 4 = severe: most severe dysfunction
	Poor distance discrimination		
	Rapid rotation		
Trunk ataxia			0 = none; 1 = only physical signs; 2 = mild: swing when closing eyes; 3 = moderate: swing when opening eyes; 4 = severe: unable to sit alone without assistance

Head tremor		0 = normal; 1 = mild; 2 = moderate; 3 = severe
Ataxia gait		0 = no; 1 = only physical signs; 2 = mild: patients have complaints; 3 = moderate: daily sitting or walking balance disorder; 4 = severe: can only walk a few steps or because of ataxia need to walk with help.
Straight line		0 = No problem, 1 = constrained, 2 = incomplete
Romberg sign		0 = normal; 1 = mild: mild eye closure instability ;2 = moderate: eye closure instability; 3 = severe: eye opening instability

0.Normal
1.Abnormal signs, no disability
2. Mild ataxia and/or moderate Romberg positive
3. Moderate trunk ataxia and/or moderate limb ataxia and/or moderate, severe trunk/gait ataxia
4. Severe gait/trunk ataxia and 3-4 limb weight ataxia
5. Unable to coordinate movement due to ataxia
X Impairment of pyramidal tract (limb muscle strength < grade 3) interferes with cerebellar function testing
FS=
Note: Only severe trunk/gait ataxia, FS = 3；
When evaluating cerebellar function due to limb weakness, add "X" after the actual score.

（5）Sensory function

Sensory system examination		L	R	note
Shallow sensation	Superficial sensation of upper limbs			0 = normal; 1 = only physical signs, no complaint 2 = Mild: Patient perceives mild tactile or pain disorders, but still distinguishes between sharp and dull 3 = Medium: Differentiating between sharp and blunt disorders 4 = Severity: unable to distinguish between sharp/blunt and/or light touch 5 = Total sensory loss
	Superficial sensation of trunk			
	Superficial sensation of lower limbs			
Vibration perception	Vibrational sensation of upper limbs			0= normal 1 = Mild: greater than 10s but less than the examiner 2 = Medium: greater than 2S but less than 10s 3 = Severity: total loss
	Vibrational sensation of lower limbs			
Sense of location	Upper limb position perception			0= normal； 1 = Mild: 1-2 wrong answers during examination, involving only distal joints 2 = Moderate: Many finger and toe movements are imperceptible and involve proximal joints. 3 = Severity: No movement at all, no standing.
	Lower extremity position perception			
Paresthesia	Abnormal sensation of upper limbs			0=no 1=yes
	Trunk			

	paresthesia			
	Lower extremity paresthesia			
Lhermitte				0=no；1=yes

normal
1.1-2 limbs with mild sense of vibration or pattern or hypothermia
2.The sensation of touch, pain or position decreased slightly or the sensation of vibration decreased moderately in 1-2 limbs.
3-4 limbs with slight decrease in single sense of vibration or temperature or figure
3. The sensation of touch, pain or position decreased moderately or vibration disappeared in 1-2 limbs.
Mild hypoesthesia of all sensations in 3-4 limbs
4. 1-2 severely impaired sensation of touch or pain;
> Moderate tactile, hypoalgesia and/or proprioceptive hypoesthesia in two limbs
5. 1-2 limb sensory loss；
Most of the body below the head has moderate impairment of touch, pain and/or proprioception
The sensation below the head basically disappeared.
FS=

（6）Bladder and rectal function

Bladder and rectum function	score	note
Urinary Waiting/Urinary Retention		0=no; 1=Mild: No significant impact on life 2=Moderate: urinary retention, frequent urinary tract infections 3=Severe: Need for catheterization 4=Loss of function
Urinary urgency/incontinence		0=no; 1=Mild: no significant impact on life; 2=Moderate: Rare urinary incontinence, once a week, with a pad 3=Severe: Frequent urinary incontinence occurs several times a week to several times a day, with a urine bag or pad 4=Loss of function and uncontrolled bladder
Catheterization		0=no； 1=Intermittent self-catheterization； 2=Continuous catheterization
Rectal dysfunction		0=no； 1=Mild: no fecal incontinence, no significant impact on life, mild constipation; 2=Medium: Must have a pad or be near the toilet. 3=Severe: Need enema or other means of cleaning rectum 4=Complete loss of function

normal
Mild urinary waiting, urgency and/or constipation
Moderate urinary waiting and/or urgency and/or rare urinary incontinence and/or severe constipation
Frequent urinary incontinence or intermittent self-catheterization; need manual cleaning of intestinal tract
Need for continuous catheterization
5. Loss of bladder or rectal function; need for catheterization
6. Loss of rectal bladder function
FS=

FS score of rectum and bladder	6	5	4	3	2	1
Rectal bystander FS score after conversion	5	4	3	3	2	1

（7）Brain Function Examination

Project	Score	Note
Depression and euphoria		0=no; 1=yes
Cognitive impairment		0=no; 1=There were only physical signs, but the patients and their families were not aware of them. 2=Mild: Patients and their families report mild cognitive changes (such as decreased ability to quickly associate and analyze complex problems; limited ability to quickly judge in specific emergencies; competent for daily work, but unable to withstand additional stress; intermittent symptoms at daily stress levels, occasional behavioral decline; and The tendency of fatigued children to forget things) 3=Moderate: Simple cognitive impairment test had definite abnormalities, but the orientation of time, place and character was normal. 4=Severity: There are 1 or 2 abnormalities in time, place and character orientation, and life is affected. 5=Dementia, confusion and/or loss of orientation
Fatigue		0=no; 1=Mild: Daily activities are unaffected 2=Moderate: Daily activities affected less than 50%. 3=Severe: Severe impact on daily activities (>50%)

normal
Only emotional changes (depression and/or euphoria) and/or mild fatigue and/or cognitive impairment
2. were found, but no complaint was made.
3 .2 = mild cognitive impairment; moderate or moderate fatigue
4. Moderate cognitive impairment
5. Severe cognitive impairment
6. dementia
FS=
Note: Fatigue is not usually counted in FS.

Only depression and/or euphoria, the brain FS = 1 point, at this time, does not count into EDSS; if the brain FS = 1 point is due to other reasons, then count into EDSS.

（8）Action ability

Walking without assistance or rest
≥500m(<4.0)
≥300m(4.5)
≥200m(5.0)
≥100m(5.5)
6.0 Unilateral or bilateral assisted walking (cane, crutch) > 100m; or unilateral assisted walking > 50m
6.5 Bilateral assisted walking (cane, crutch) > 20m without rest; or unilateral assisted walking < 50m
7.0 assisted still can not walk 5 meters, action is basically limited to wheelchair, can rely on wheelchair action; every day in the wheelchair for about 12 hours
7.5 It is still unable to walk with the assistance of wheelchair. Movements are basically confined to wheelchair. Shaking wheelchair and transferring body need help.
8.0 Actions are basically confined to beds, chairs or wheelchairs, but most of the time is spent under the bed every day; a lot of self-care ability is retained; both arms are basically functional.
8.5 Daily activities are basically limited to trauma; both arms have partial function; and some self-care ability is retained.
9.0 Patients are bedridden, able to communicate and eat
9.5 Patients are bedridden and can hardly communicate effectively or eat or swallow.
10 died in MS

note：
EDSS: 0-1.5: Patients should walk normally.
EDSS: > 2 points: Patients walk independently (> 500 meters), but their mobility is limited. At this time, the EDSS score is only determined by the FS score, and the FS of pyramidal tract function or vesicle function should be (> 2 points).
EDSS: > 4 points; patients walk < 500 meters independently
EDSS: > 6 points: patients need help walking
Another person's assisted walking is equivalent to bilateral assisted walking.

EDSS score：
0.0 All FS are 0.
1.0 1 item FS is 1
1.5 > 1 item FS is 1
2.0 1 FS is 2, the rest is 0 or 1
2.5 2 FS is 2, the rest is 0 or 1.
3.0 1 FS is 3, the rest is 0 or 1, 3-4 FS is 2, the rest is 0 or 1.
3.5 1 items FS were 3,1-2 items FS 2, the rest FS were 0 or 1; 2 items FS were 3, the rest were 0 or 1; 5 items FS were 2, and the rest were 0 or 1.
4.0 Walk independently (> 500 meters), 1 FS is 4, and the rest is 0 or 1.
4.5 Walking independently (> 300 meters), 1 FS is 4
5.0 Walking independently (> 200 meters) and (> 1 FS) 5
5.5 Walking independently (> 100m)
6.0-10.0; see Action Capability

note：
EDSS < 4: Walking independently (> 500 meters), the exact score is determined by FS.
EDSS 4.0-5.0: Accurate scoring is determined by FS and walking range, and the more serious parameters are the scoring results of huge Eastern EDSS.
EDSS 5.5-8.0: Completely determined by operational capability

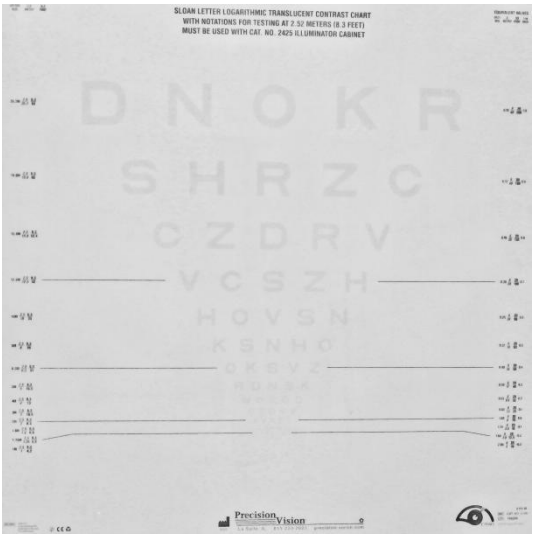
EDSS0-4.0; EDSS score should not change to 1 unless the FS score changes by 1.
EDSS score should be lower than any single FS score, except visual acuity and bladder and rectal function.

APPENDIX 5. VISUAL ACUITY TEST

The following charts are used under 160 cd/m^2 luminance level at 2.52 m chart distance (Precision Vision, Inc.).

ETDRS Chart

2.5% Low Contrast ETDRS Chart



APPENDIX 6. CLINICAL LARORATORY TESTS

Chemistry Panal sodium Potassium calcium chloride Sodium bicarbonate Blood urea nitrogen Creatinine Uric acid glucose Alkaline phosphatase Alanine aminotransferase (ALT) Aspartate aminotransferase (AST) Total bilirubin Direct bilirubin Indirect bilirubin albumin Total protein Total cholesterol Triglyceride Low-density lipoprotein High density lipoprotein **Complete blood count (CBC)** White blood cell count (WBC) White blood cell identification Red blood cell count (RBC) RBC mean corpuscular volume (MCV) RBC distribution width Hemoglobin Platelet count Monocyte ratio Neutrophil ratio Lymphocyte ratio Neutrophil ratio Eosinophil ratio	**Human Chorionic Gonadotropin (β-HCG)** **Auto Aquaporin 4 antibody (AQP4-IgG)** **Flow cytometry** B cell ($CD19^{pos}$) B cell subsets Naive B Memory B Double Positive B Double Negative B Antibody Secrting Cells

APPENDIX 7. COMMON TERMINOLOGY CRITERIA FOR ADVERSE EVENTS (CTCAE 5.0)

NCI CTCAE=National Cancer Institute Common Terminology Criteria for Adverse Events.
Note: Based on the NCI CTCAE (v5.0), which can be found at:
https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_5.0/ or
https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf

TABLE OF CONTENTS FOR

TANGO PROTOCOL AMENDMENT (VERSION 2.0)

TANGO PROTOCOL AMENDMENT, VERSION 2.0: RATIONALE

TANGO protocol has been amended for the following reasons:

- **To emphasize the reason for early concomitant usage of corticosteroids or other immunosuppressants**

  Neuromyelitis optica sepctrum disorder (NMOSD) is distinct from multiple sclerosis (MS） in the pathological mechanisms and prevention of treatments. For acute attacks of MS, there is no corticosteroids tapering after high-dose of corticosteroids is given. Disease modifying drugs (DMDs) will be given to the patients of MS just after relapses for prevention. But in NMOSD, corticosteroids will be tapered slowly after high-dose of corticosteroids treatment in acute relapses. From our experience, abrupt discontinuation of corticosteroids or other immunosuppressants may increase the risk of relapses in NMOSD patients. In TANGO trial, investigators insist that early concomitant usage of corticosteroids or other immunosuppressants be necessary before the trial agents exert maximal effects.
- **To add TPMT genetyping table in the Appendix**

  Intolerance of azathioprine has been reported frequently in Chinese patients in NMOSD and other autoimmune diseases. The screening examination of TPME gene are commonly seen in Chinese patients. However, to be noted, even though the patients have no mutations in TPMT genetyping, they may still experience adverse events as other variable factors may also affect the tolerability of azathioprine. In this investigator-initiated trial, we make every efforts to reduce the occurrence of serious adverse events and drop-out rates during the trial.
- **To establish an Expert Panel to determine the relapses of NMOSD**

  TANGO is an investigator-initiated trial and there are no authorized representatives of the Industrial Sponsor or the Sponsor's designee to perform audits or inspections for relapses. The diagnosis of relapses in NMOSD should be taken careful consideration. Each suspicious relapse reported from an investigator will be reviewed by all the Principle Investigators, that is, an Expert Panel. In addition, patients with NMOSD may have fluctuations of symptoms. It is important to discriminate whether such symptoms reflect new relapses. Determination of relapses will be confirmed only after consultation of all the Principle Investigators across centers.
- **To specify the roles of laboratory /SD-OCT/ VEP assessors**

  These assessors are independent of the trial randomization and treatments.They will perform the testing at protocol-defined time-points.
- **To clarity the "bilateral monocular" measurements of high-contrast visual acuity and low-contrast letter acuity**
- Previous studies show that RNFL thickness moderately correlated with 2.5% low-contrast letter acuity scores in multiple sclerosis patients when the scores of all demyelinating disease eyes were compared to controls. We aim to find associations of visual acuity with OCT/VEP parameters including RNFL, for each eye. As optica neuritis is common in NMOSD, the parameters of the affected and unaffected eyes will be investigated.
- **To exclude the patients who had prior relapses or toxic intolerance on azathioprine treatment (adequate dosage and follow-up period observation)**

  It is not ethical to assign the patients who had prior relapses or toxic intolerance on azathioprine treatment to resume azathioprine for prevention.

- **To clarity permitted concomitant usage of non-immunosuppressants**
  The patients will be permitted to continue treatments for chronic historical diseases. For example, subjects in both arms should be treated with anti-platelet therapy (aspirin or clopidogrel) according to the local practice at the discretion of the Investigator.Any concomitant medications (including prescription drugs, over-the-counter medications, herbal/homeopathic medications, preventive vaccines, vitamins, and nutritional supplements) must be recorded in the appropriate CRF. A description of the type of drug, amount, duration, and reason for administration of drug must be documented.
- **To clarify the adjustments of administration of tocilizumab or azathioprine during the trial**
  Treatment-emergent adverse events may affect routine administrations of the investigational medicinal products (IMP). The dose and interval of the IMP will have to be adjusted corresponding to the actual conditions in case of serious adverse events.

PROTOCOL AMENDMENT, VERSION 2: SUMMARY OF CHANGES

PROTOCOL SYNOPSIS

The protocol synopsis has been updated to reflect the changes to the protocol, where applicable.

SECTION 6.3 Rationale for Early Concomitant Usage of Corticosteroids and Immunosuppressants

NMOSD is distinct from MS in the treatment of corticosteroids usage. In MS, disease modifying drugs will be given to the patients just after high-dose of corticosteroids treatments, with no tapering of corticosteroids. In NMOSD,~~Usually,~~ initial or recurrent episodes are usually treated with high-dose intravenous methylprednisolone (1 g daily for three to five consecutive days). In many countries of the EU, the intravenous therapy with methylprednisolone is followed by an oral taper and needs to be performed slowly[16]. Some patients experience clinical worsening when prednisone is reduced below 5 -15 mg/day. ~~So~~ *If an* immunosuppressant would be added, *then corticosteroids will be used for rapid immunosuppression until the immunosuppressant exerts its full effect.* Taking AZA as an example, as the treatment may only take full effect after 3 - 6 months, it should initially be combined with oral steroid therapy, as oral steroids have been beneficial to ~~suppress~~ reduce disease activity in NMOSD[17].

SECTION 8.1 Screening Period

....

TPMT genotyping will be performed in all the patients *to detect mutant alleles (TPMT*2, TPMT*3A, TPMT*3B, and TPMT*3C) (see Appendix 3).* If the patient has homozygous or heterozygous mutation in TPMT gene, then the patient will not be enrolled.

....

SECTION 8.2 Study Period

....

The end of the study visit for an individual patient will take place when one of the following conditions is met, whichever comes first: (a) the patient experiences a definite relapse and early termination; or (b) when the last participating patient completes the last scheduled visit or when the ~~Investigator~~ *Expert Panel* decide to discontinue the study or development program. Identification of potential relapse is critical for patient safety and for the trial. Any potential relapse will be evaluated according to the information below.

....

As this is a time-to-event trial, patients who experience a relapse will be discontinued from this trial after completion of the Week 24 Relapse Evaluation Visit. Thus, the Week 24 Relapse Evaluation Visit also serves as the end of the study visit for these patients. For patients who do not have relapses, the end of the study visit will be defined as when the last participating patient completes the last scheduled visit or when the ~~Investigator~~ *Expert Panel* decide to discontinue the study or development program. Patients who complete the trial either because of a relapse or because of the trial is ended may be encouraged receive RTX (or TCZ, if possible) treatment.

SECTION 8.4.6 Independent Laboratory */ SD-OCT / VEP* Assessors

The laboratory *(including serum AQP4-IgG, B cell subsets determination) / SD-OCT / VEP* assessors roles are responsible for the specified examination of the corresponding index. They will will be blinded to the randomization and treatments throughout the trial including at the time of a relapse.

SECTION 8.6.2 Study Period

....

- ~~Bi-ocular~~ *Bilateral monocular* SD-OCT
- ~~Bi-ocular~~ *Bilateral monocular* VEP

....

SECTION 8.6.3 Relapse Evaluation Visit (Within 24 Hours)

....

The Investigator determines if the clinical signs, symptoms and neurological change (objective findings on the examination) meet the definition for relapses as outlined in this protocol. The ~~Investigator~~ *Expert Panel* will evaluate all the information done as above to assure the relapses.

....

SECTION 8.9.1 End of Trial for an Individual Patient

The trial will be ended for a patient when one of the following conditions is met, whichever comes first:

- The patient experiences a definite relapse and early termination
- When the last participating patient completes the last scheduled visit or when the ~~Investigators~~ *Expert Panel* decide to discontinue the study or development program.

SECTION 8.9.2 End of Trial for All Patients

The end of the trial has been defined as the date at which the last data point from the last patient is collected.or when the ~~Investigators~~ *Expert Panel* decide to discontinue the study or development program.

SECTION 9.2 Patient Exclusion Criteria

1. Current evidence or known history of clinically significant infection (Herpes simplex virus, varicella-zoster virus, cytomegalovirus, Epstein-Barr virus, human immunodeficiency virus, Hepatitis viruses, Syphilis, etc.)
2. Pregnant, breastfeeding, or child-bearing potential during the course of the study
3. Patients will not participate in any other clinical therapeutic study or will not have participated in any other experimental treatment study within 30 days of screening
4. Participation in another interventional trial within the last 3 months
5. Heart or kidney insufficiency
6. Tumor disease currently or within last 5 years
7. Clinically relevant liver, kidney or bone marrow function disorder
8. Receipt of rituximab or any experimental B-cell depleting agent within 6 months prior screening and B-cells below the lower limit of normal measured by flow cytometry

9. *Prior relapses or toxic intolerance on azathioprine treatment (adequate dosage and follow-up period observation)*

SECTION 10.2.3 Other Permitted Medications

The patients will be permitted to continue treatments for chronic historical diseases. For example, subjects in both arms should be treated with anti-platelet therapy (aspirin or clopidogrel) according to the local practice at the discretion of the Investigator. Symptomatic treatments are permitted during the course of the trial for underlying conditions. For example, oxcarbazepine could be used for neuropathic pain.

SECTION 11.1.1 Relapses

Patients will be evaluated within 24 hours after a possible relapse, which is confirmed by the Expert Panel and again following intervals of 4 weeks until 24 weeks by the ~~Investigator~~ *Expert Panel* and by EDSS raters. The EDSS raters, who are also unaware of trial-group assignments, are not involved in patient care. Confirmed relapses could be treated with IVMP or IVIG at the investigator's discretion.

SECTION 12.4.3 Elevated Liver Enzymes

Elevated liver enzymes have been reported in the patients treated with tocilizumab or azathioprine in previous clinical trials and clinical practice. So monitoring liver enzymes is necessary for safety consideration. *Treatment with tocilizumab or azathioprine in this trial will be adjusted or discontinuation according to Table 7.*

Table 7 Risk Mitigation for Hepatic Enzyme Elevation

ALT or AST Values	***Action for TCZ group***	***Action for AZA group***
1-3 × ULN	*Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. The dosage of TCZ will remain unchanged.*	*Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. The dosage of AZA will remain in the range of 2-3 mg/kg/day.*
3-5 × ULN	*Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs will be added. If ALT/AST levels return to baseline, the dosage of TCZ may be resumed to 8 mg/kg/4 weeks. For persistent increases in this range,the dosage of TCZ may be adjusted to 8 mg/kg/6 weeks.*	*Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs will be added. If ALT/AST levels return to baseline, the dosage of AZA may be resumed to 2-3 mg/kg/day. For persistent increases in this range,the dosage of AZA may be adjusted to 1-2 mg/kg/day.*
> 5 × ULN	*Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. Laboratory tests should be repeated to confirm value. If confirmed, TCZ should be discontinued.*	*Interrupt concomitant hepatotoxic drugs and concomitant hepatoprotective drugs is permitted to be added. Laboratory tests should be repeated to confirm value. If confirmed, AZA should be discontinued.*

TCZ=tocilizumab; AZA=azathioprine; ALT=alanine amiotransferase; AST=aspartate aminotransferase; ULN= upper limit of normal range.

Patients withdrawn from the study because of elevated liver function test results must have repeat tests performed as clinically indicated until levels return to baseline values. If the patient's liver function test results have not returned to normal or to the patient's baseline level within 6 months (or sooner if deemed necessary by the Investigator), a specialist referral is recommended and an ultrasound should be considered.

SECTION 12.4.4 Neutropenia

Decrease in neutrophil has been observed following treatment with TCZ or AZA in patients with other autoimmune diseases. ~~Administration of Recombinant Human Granulocyte will be appropriate, if necessary.~~ *The risk mitigation strategies for neutropenia is summarized in Table 8.*

Table 8. Risk Mitigation for Neutropenia

ANC (cells/mm^3)	***Action for TCZ group***	***Action for AZA group***
> 1000	*Maintain dose*	*Maintain dose*
500− 1000	*If neutropenia persists, prolong TCZ treatment from every 4 weeks to 6 weeks). When ANC increases to > 1000 cells/mm^3, resume IV TCZ every 4 weeks, as clinically appropriate*	*Interrupt concomitant suspicious toxic drugs. If ANC returns to baseline, the dosage of AZA may be resumed to 2-3 mg/kg/day. For persistent decreases in this range,the dosage of AZA may be adjusted to 1-2 mg/kg/day. Administration of Recombinant Human Granulocyte will be appropriate.*
< 500	*Discontinue TCZ permanently after repeat confirmation*	*Discontinue AZA permanently after repeat confirmation. Administration of Recombinant Human Granulocyte will be appropriate.*

ANC= absolute neutrophil count; TCZ= tocilizumab; AZA=azathioprine.

APPENDIX 3. TPMT SINGLE NUCLEOTIDE POLYMORPHISM

Testing	*TPMT*2,*3A,*3B,*3C single nucleotide polymorphism*				
Methods	*PCR and sequencing*				
Instrument	*ABI 3500xL Dx*				
Results	*Gene*	*Polymorphism*	*Locus*	*Genotype*	*Mutation*
	TPMT	**2*	*c.238*	*G/G*	*NO*
		**3A*	*c.460/c.719*	*G/G AND A/A*	*NO*
		**3B*	*c.460*	*G/G*	*NO*
		**3C*	*c.719*	*A/A*	*NO*

Results:

No TPMT mutation was detected in the allelic locus of this patient. According to reported references, the side effects may be lower after using azathioprine.

Note:

The three mutation locus of the 5th exon G238C, the 7th exon G460A and the 10th exon A7I9G of the common

*TPMT (thiopurine methyltransferase) gene are mononucleotide polymorphisms. Mutation of the 5th exon G238C results in a mutation in the allele TPMT*2.Mutation of the 7th exon G460A results in a mutation in the allele TPMT*3B, and mutation in the 10th exon A719G results in an allele mutation of the gene TPMT*3C, mutation of the 7th exon G460A and the 10th exon A719G resulted in a mutation of the allele TPMT*3A. Among the above alleles, according to the reported paper, the unmutated genotype, which is normal, has a lower toxic effect of azathioprine. And the heterozygous mutant adopts the toxic side effects of the steroids, and the homozygous mutant uses the ridges. The drug has high side effects. This test is used for those patients who lack the TPMT gene of the dredging methyltransferase. It provides reference for the dosage of sedative drugs at the time of treatment. The report results are only responsible for the specimens in this test and are used for medical research.*

STATISTICAL ANALYSIS PLAN

TITLE: A RANDOMIZED, CONTROLLED, OPEN-LABEL, MULTI-CENTER TRIAL TO COMPARE THE SAFETY AND EFFICACY OF TOCILIZUMAB VERSUS AZATHIOPRINE IN PATIENTS WITH HIGHLY RELAPSING NEUROMYELITIS OPTICA SPECTRUM DISORDER (NMOSD): A HEAD-TO-HEAD COMPARATIVE STUDY (TANGO)

PROTOCOL NUMBER: 2017kylc005
STUDY DRUG: Tocilizumab or Azathioprine
PLAN PREPARED BY: School of Public Health, Tianjin Medical University, Tianjin, China
DATE: 20 October 2017

TABLE OF CONTENT

1. BACKGROUND

The TANGO trial is an investigator-initiated, randomized, multi-center, parallel-group, controlled, open-label study to compare the safety and efficacy of tocilizumab versus azathioprine in patients with highly relapsing neuromyelitis optica spectrum disorder (NMOSD).

To date, no FDA-approved treatment has been demonstrated to significantly reduce the relapses and progression of disability in patients with NMOSD. Classical immunosuppressants such as azathioprine, mycophenolate, are being used for relapse prevention. Tocilizumab has been used for refractory NMOSD patients who did not well respond to immunosuppressants or rituximab. However, evidence of choice of these drugs for prevention of NMOSD still lacks. Thus, TANGO trial is designed as a head-to-head study ever conducted.

2. STUDY DESIGN

The TANGO trial will evaluate the safety and efficacy of intravenous tocilizumab, compared with oral azathioprine in adults with highly relapsing NMOSD. Patients will be treated for a minimum of 60 weeks.

This study consists of the following periods: a Screening Period, an open-label Study Period, and a Safety Follow-up Period.

The primary analysis of the study will be performed when the last patient enrolled has been treated for at least 60 weeks or has a new relapse, or the end of the study.

2.1 Outcome Measures

2.1.1 Efficacy Outcome Measures

Primary outcome measures:

- Time to first relapse from the baseline in a time-to-event analysis

Secondary outcome measures:

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

Exploratory efficacy measures including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials

(VEP)

- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

2.1.2 Safety Outcome Measures

The safety of tocilizumab or azathioprine will be assessed based on the all treatment-emergent adverse events (TEAEs), Common Terminology Criteria for Adverse Events (CTCAEs 5.0), serious adverse events (SAEs), and the changes from baseline through trial completion in vital signs, electrocardiogram (ECG), routine clinical laboratory tests, and pregnancy tests for female patients of childbearing potential.

2.2 Determination of Sample Size and Power

This is a randomized open-label parallel-group study, to evaluate the safety and efficacy of tocilizumab and azathioprine in patients with highly relapsing NMOSD. As such, the study is based on observing relapse events. With 1:1 randomization for the trial groups, we calculated that 118 patients would provide a power of 80% to determine the pre-specified between-group difference on the basis of a two-sided log-rank test at a 5% level of significance, assuming a 10% dropout rate. Randomized patients who discontinue after initiation of treatment will not be replaced.

The sample size and power calculation assumptions for this time to first event study are as follows:

- Log-rank test for comparison of tocilizumab and azathioprine
- 1:1 randomization (tocilizumab and azathioprine)
- Power 80%
- Two-sided 5% level of significance
- Dropout rate 10%
- Relapse-free rate of 85% for tocilizumab and 60% for azathioprine at 12 months (hazard ratio = 0.318)

The formula for calculating sample size is：

$$n_1 = n_2 = \frac{(\mu_\alpha+\mu_\beta)^2(1+h)^2}{[2-S_S(t)-S_N(t)](1-h)^2} \qquad \mathrm{h} = \frac{lnS_S(t)}{lnS_N(t)}$$

Parameters：$S_S(t)$ and $S_N(t)$ are the relapse-free rate of patients on the medication of tocilizumab and azathioprine, respectively. $\mu_{0.05/2}$=1.96, $\mu_{0.2}$=0.842.

3. STUDY CONDUCT

3.1 Randomizaiton

All patients who are evaluated eligible for randomization by the Investigator will be randomized on Day 1 on a 1:1 basis to the tocilizumab arm or the azathioprine arm. The randomization will be across centers. This is a completely randomized grouping experiment. The random numbers produced by a binomial distribution determine the group distribution of patients. The following SAS code had been used to do this:

```
data aa;
random=ranbin(11,1,0.5);
run;
```

Patients will be centrally randomized based randomization application.

3.2 Independent Review Facility

The primary outcome is the first relapse occurring after the baseline. The relapse will be confirmed by the Expert Panel. MRI scans will also be performed and read by the MRI radiologist and technicians, blinded to the treatment assignment and the clinical information.

The secondary efficacy endpoint will be derived from the EDSS values recorded at any visits. EDSS assessments are performed by an examining neurologist (not the treating neurologist or Principle Investigator). On serum AQP4-IgG titers, two independent assessors, who are unaware of clinical information, monitored and classified every blood sample.

4. STATISTICAL METHODS

4.1 Analysis Populations

One patient population will be defined for the purpose of the safety analysis, and two populations will be defined for the efficacy analysis. All efficacy endpoints will be analyzed using the intent-to-treat (ITT) population. The per-protocol population will be used for the primary efficacy endpoint only, to evaluate the influence of major protocol deviations and as a sensitivity check for the ITT analysis.

4.1.1 Intent-to-Treat Population

All randomized patients will be included in the ITT population. Patients who prematurely withdraw from the study for any reason or for whom an assessment is not performed for any reason will still be included in the ITT analysis.

4.1.2 Per-Protocol Population

The per-protocol population will include all patients in the ITT population who adhere to the protocol and will be summarized according to the randomization arm. Per-Protocol analyzes the patients receiving tocilizumab or azathioprine treatment as monotherapy after a washout period during which early concomitant immunosuppressants will be discontinued. Patients may be excluded if they violate the inclusion or exclusion criteria or deviate from the study plan. Only those patients with deviations that are deemed to potentially affect the efficacy of study treatment will be excluded from the per-protocol population.

4.1.3 Safety Population

The safety population will include all patients who received at least one dose of investigational medicinal products (tocilizumab or azathioprine). Randomized patients who receive the incorrect therapy will be summarized in the group according to the therapy actually received.

4.2 Analysis of Study Conduct

The following analyses will be conducted to evaluate the study conduct:

- Summary of protocol deviations
- Summaries of ITT, per protocol defined, and safety populations, including numbers of patients in each population, and reasons for exclusion from the per-protocol population
- Summary of subject disposition, including the number of treatment doses received, and the number of patients entering into safety follow-up

4.3 Analysis of Treatment Group Comparability

For continuous variables, the mean, median, standard deviation (SD), median (interquatile range), and minimum and maximum values will be calculated. For categorical variables, the number and percentage in each category will be displayed. For each item in the following lists, the units and categories to be used are indicated in parentheses and separated by commas. All durations are calculated with respect to the date of randomization, if not stated otherwise.

Demography and stratification factors based on Case Report Form (CRF) data:

- Age (years) at baseline
  Age stratification category (≤ 45, > 45 years)
- Sex (male patients, female patients)
- Weight (kg)
- Body mass index (measured in kg/m^2)

Baseline NMOSD disease history:

- AQP4-IgG seropositivity (%)
- Duration since NMOSD symptom onset (calculated in years; i.e., divide by 365.25)
  Duration stratification category (≤ 6, > 6 years)
- Total number of previous relapses
  Number of relapses stratification category (≤ 8, > 8)
- Number of relapses 1 year before randomization
  Number of relapses 1 year before randomization stratification category (≤ 1, > 1)
- Number of relapses 2 year before randomization
  Number of relapses 2 year before randomization stratification category (≤ 3, > 3)

Baseline disease characteristics:

- EDSS (continuous) ： EDSS stratification category (≤4.5, > 4.5 years)

Baseline serum AQP4-IgG titer data

- Serum AQP4-IgG titer

Baseline visual acuity data

- 100% high-contrast visual acuity
- 2.5% low-contrast letter acuity (LCLA)

Baseline OCT/VEP data

- Retinal nerve fiber layer (RNFL) thickness
- Retinal ganglion cell complex (GCC) volume
- P100 latency
- P100 amplitude

Baseline MRI data:

- Whole-brain volume at baseline
- Number of T2 lesions at baseline

4.4 Efficacy Analysis

All statistical hypotheses for the primary and secondary endpoints and treatment comparisons will be tested at the 5% significance level (α= 0.05) against two-sided alternatives. The analysis are pre-specified by two subgroups of the patients: with and without concomitant autoimmune diseases.

4.4.1 Primary Efficacy Endpoint

Significance Level

The null hypothesis will be tested at the α= 0.05 level (two-sided test).

The hypotheses to be tested are:

- H0 (null hypothesis): There is no difference in the probability of relapse-free between the tocilizumab and azathioprine groups.
- H1 (alternative hypothesis): There is a difference in the probability of relapse-free between the tocilizumab and azathioprine groups.

Definition

The time to first relapse is defined as the time from baseline to the onset of the first new relapse during the follow-ups.

Analysis Methods

The ITT population analysis will be presented. Patients who did not have initial relapse at the time of end of the trial, time of early discontinuation, or loss to follow-up will be censored at the date of their last EDSS assessment that occurred during the treatment period. The proportion of patients who do not have a new relapse during the trial at predefined time points (i.e., 1 year, 2 years, etc.) will be estimated using Kaplan-Meier methodology. The overall hazard ratio will be estimated using a stratified Cox regression model with the same stratification factors used in the aforementioned stratified log-rank test.

SAS code for primary endpoint

The primary endpoint is time to the first relapse. The comparison of the treatment groups for the primary endpoint will use a two-sided log-rank test including strata for the randomized intervention variable— different medication. The basic SAS code for this log-rank analysis is:

```
proc lifetest data=trial;
time ftgapw*relapse(0);
strata group;
run;
```

Where group is a variable that indicates patient's randomized treatment group, ftgapw is a variable for the patient's time in study period at time of relapse or censoring, relapse is the censoring variable (0 = no event (censored), 1= event).

The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model. The basic SAS code for this analysis is:

```
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
run;
```

SAS code for primary endpoint with main covariates

The comparison of the situation of concomitant autoimmune diseases for the primary endpoint will use a log-rank test in two randomized treatment group, respectively. The basic SAS code for this is:

```
data aza;
set trial(where=(group=0));
run;
proc lifetest data=aza;
time ftgapw*relapse(0);
strata concom;
run;
proc phreg data=aza;
model ftgapw*relapse(0)=concom/RL;
run;

data tcz;
set trial(where=(group=1));
run;
proc lifetest data=tcz;
time ftgapw*relapse(0);
strata concom;
run;
proc phreg data=tcz;
model ftgapw*relapse(0)=concom/RL;
run;
```

Where concom is a variable that indicates the patient's status of concomitant autoimmune diseases.

SAS code for primary endpoint by subset

Stratified Cox proportional-hazards models in different strata of EDSS at randomization, age at randomization and duration of disease are conducted to estimate the hazard ratio for the primary end point. The basic SAS code for this is:

```
proc sort data=trial;
by edssstra;
run;
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by edssstra;
run;
proc sort data=trial;
by relapstra;
run;
```

```
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by relapstra;
run;
proc sort data=trial;
by hisystra;
run;
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by hisystra;
run;
proc sort data=trial;
by agestra;
run;
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by agestra;
run;
```

Where edssstra is a variable that described the grade of EDSS status at randomization (1 = baseline EDSS score > 4.5, 2 = baseline EDSS score ≤ 4.5). Variable relapstra described the hierarchy of the number of relapse (1 = relapse number > 8, 2 = relapse number ≤ 8). Variable hisystra described the grade of disease duration (1 = disease duration > 6 years, 2 = disease duration ≤ 6 years). Variable agestra described the hierarchy of age at randomization (1 = age at randomization > 45 years, 2 = age at randomization ≤ 45 years).

4.4.2 Secondary Efficacy Endpoints

4.4.2.1 12-week Confirmed Disability Progression

Definition

The time to onset of confirmed disability progression (12-week confirmation [weeks]) is defined as the time from baseline to the onset of the first disability progression that is confirmed at the next regularly scheduled visit ≥ 12 weeks after the initial disability progression. If the patient has an infection, dosing may not occur on the Day 1 visit.

Baseline for the time to onset of confirmed disability is the date of randomization. Disability progression is defined as an increase of ≥ 1.0 point from baseline EDSS score if the baseline EDSS value is ≤ 5.5 points (inclusive) or an increase of ≥ 0.5 points if the baseline EDSS value is > 5.5 points. Assessments within 30 days after a protocol-defined relapse will not be used for confirmation of confirmed disability progression. The non-confirmatory EDSS assessments (between the initial disability progression and the confirmation of disability progression should also fulfill the requirements of the progression. Otherwise, the initial disability progression is not confirmed.

The baseline EDSS value is the average score of the EDSS assessment at screening and baseline

(Day 1 visit) up to and including the date of randomization. If one of the values is missing, the non-missing values will be used as baseline.

A blinded rater at each study site will assess EDSS for all patients at the site at screening, baseline, every 4 weeks (regularly scheduled visit) during the blinded treatment period of the study, during the safety follow-up period, at any unscheduled visits, and at withdrawal-from-treatment and end-of-study visits. Additional EDSS assessments for individual patients may be requested between visits (i.e., during an NMOSD relapse).

The examining investigator is not the physician responsible for the patient care (the treating investigator).

The EDSS is based on a standard neurological examination; the seven categories of the EDSS representing FS (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, and cerebral [or mental], plus "other") are rated and scored (collectively, FSs), in addition to an ambulation score (0−12). Each domain of the FS is an ordinal clinical rating scale from 0 to 5 or 6. These ratings are then used in conjunction with observations and information regarding ambulation and the use of assistive devices (which will also be scored) to determine the EDSS score. The EDSS is a disability scale that ranges in 0.5-point steps from 0 (normal) to 10.0 (death).

Analysis Methods

The ITT population analysis will be presented. Patients who did not have initial disability progression at the time of end of the trial, time of early discontinuation, or loss to follow-up will be censored at the date of their last EDSS assessment that occurred during the treatment period. There are various options to obtain EDSS assessments (confirmatory and non-confirmatory) and to record how the discontinuation from treatment and study can influence the confirmation of the disability progression. The following rules apply for all patients with initial disability progression:

- Initial disability progression must occur when the patient is still on treatment.
- All non-confirmatory EDSS assessments (if any) after initial disability progression and up to and including the confirmatory EDSS assessment should also fulfill the requirements of the progression as defined above. Otherwise an initial disability progression event cannot be confirmed.

The comparison of the treatment groups for secondary efficacy endpoint--12 weeks confirmed disability progression will use a log-rank test including strata for the randomized intervention variable. The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model. The basic SAS code for this log-rank analysis is:

```
proc lifetest data=trial;
time edssgapw*edsspro(0);
strata group;
run;
proc phreg data=trial;
model edssgapw*edsspro(0)=group/RL;
```

run;

4.4.2.2 Changes of serum AQP4-IgG titers by CBA and FIPA

Two methods are used to dynamically evaluate the serum AQP4-IgG titers of patients in this study at the time of baseline, 12 weeks, 24 weeks, 36 weeks, 48 weeks and 60 weeks from the randomization. Wilcoxon rank sum test is used to compare the serum AQP4-IgG titers of two groups. The basic SAS code for this is:

```
proc univariate normal data=trial normal;
class group;
var AQP40WC AQP412WC AQP424WC AQP436WC AQP448WC AQP460WC AQP40WF
AQP412WF AQP424WF AQP436WF AQP448WF AQP460WF;
run;
proc npar1way data=trial wilcoxon;
class group;
var AQP40WC AQP412WC AQP424WC AQP436WC AQP448WC AQP460WC AQP40WF
AQP412WF AQP424WF AQP436WF AQP448WF AQP460WF;
run;
data trialaqp;
set trial;
a5 = AQP460WC - AQP40WC;
b5 = a5/ AQP40WC*100;
run;
proc npar1way data=trialaqp HL(REFCLASS=2) wilcoxon;
class group;
var a5 b5;
run;
```

4.4.3 SAS Code for Demographic and Clinical Characteristics of the Patients at Baseline

Demographic and clinical characteristics of the patients at baseline are compared between two groups. Continuous Variables, (age, EDSS at randomization, disease duration), are analyzed using t-test or Wilcoxon rank sum test according to the distribution of different variables. Qualitative variables, (gender, AQP4-IgG seropositivity, the occurrence of first relapse), are analyzed using chi-square test. The basic SAS code for this is:

```
proc freq data=trial;
tables (gender AQP4 relapse )*group/chisq;
run;

proc univariate normal data=trial normal;
class group;
var age hisy edss0w;
run;
proc ttest cochran data=trial;
```

```
class group;
var hisy;
run;

proc npar1way data=trial wilcoxon;
class group;
var age edss0w;
run;
```

Where AQP4 is a variable that indicates AQP4-IgG seropositivity of patient at baseline, hisy indicates the duration of disease, edss0w indicates the EDSS score of patient at baseline, relapse is the censoring variable (0 = no event (censored), 1= event).

4.4.4. SAS Code and Analysis Method (PP)

There are 5 patients dropping out from this trial, 2 (1 had severe adverse event; 1died) of them in tocilizumab group, and 3 (2 had severe adverse event; 1died) of them in azathioprine group. Per-protocol (PP) analysis is conducted to evaluate the stability of the main results of intention-to-treat (ITT) analysis above. The basic SAS code for this is:

```
data trialpp;
set trial(where=(relaperiod^=0));
run;
data trialpp1;
set trialpp(where=(withdraw=1));
run;
proc lifetest data=trialpp1;
time ftgapwpp*relapse(0);
strata group;
run;
proc phreg data=trialpp1;
model ftgapwpp*relapse(0)=group/RL;
run;
data concom;
set trialpp1(where=(concom=1));
run;
data noconcom;
set trialpp1(where=(concom=2));
run;
proc lifetest data=concom;
time ftgapwpp*relapse(0);
strata group;
run;
proc phreg data=concom;
model ftgapwpp*relapse(0)=group/RL;
```

```
run;
proc lifetest data=noconcom;
time ftgapwpp*relapse(0);
strata group;
run;
proc phreg data=noconcom;
model ftgapwpp*relapse(0)=group/RL;
run;
```

Where withdraw is a variable that indicates whether patients withdraw from the trial or not (1=not withdraw; 0=withdraw). And relaperiod indicates the medication period of relapse (0=relapse happened in combination medication period; 1=relapse happened in individual medication period). Variable ftgapwpp is the time from individual medication to first relapse. Where concom is a variable that indicates the patient's status of concomitant autoimmune diseases.

4.4.5 Time to Onset of Confirmed Disability Progression Confirmed for ≥24 Weeks

In addition to the primary analysis, the time to onset of confirmed disability progression will be analyzed in the tocilizumab and azathioprine treatment arms with use of a 24-week confirmation window for disability progression. The analysis of EDSS progression will be conducted as described for the primary analysis, with the exception that the time to onset of confirmed disability progression (24-week confirmation) is defined as the time from baseline to the first disability progression that is confirmed at the next regularly scheduled visit ≥ 24 weeks (≥ 161 days) after the initial disability progression. The basic SAS code for this log-rank analysis is:

```
proc lifetest data=trial;
time edsspg2*edssp2(0);
strata group;
run;
proc phreg data=trial;
model edsspg2*edssp2(0)=group/RL;
run;
```

4.4.6 Exploratory Efficacy Endpoints

The following exploratory efficacy outcomes will be analyzed:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)

- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

Other exploratory efficacy endpoints listed in the protocol but not included in this SAP, will be reported in separate study reports or publications.

4.4.7 Sensitivity Analyses

The following sensitivity analyses of the primary endpoint will be conducted according to the analysis described in Section 4.4.1:

1. The primary analysis will be repeated using the per-protocol population as the analysis population. The below sensitivity analyses will be done with the ITT population.

2.

2. A sensitivity analysis using multiple imputations will be performed for the ITT population to explore the potential influence of informative censoring on the results of the primary efficacy analyses. Censoring for a reason that is not independent of that patient's prognosis is called "informative censoring." The influence of informative censoring needs to be explored, because for time-to-event endpoints, the intent-to-treat principle requires that essentially all randomly assigned patients be observed up to the endpoint or up to the end of the study. When a patient's follow-up is censored X number of months after random assignment, then, in the computation of the Kaplan-Meier estimates, log-rank or Cox regressions, that patient's outcome after X number of months is assumed to have the same outcome as the other patients in their treatment group who also are free of the outcome at X number of months and who remain under follow-up beyond X number of months.

Thus, unless the reason for being censored is independent of that patient's prognosis, failure to observe that patient until occurrence of his study endpoints could lead to significant bias as well as increased variability in the evaluation of treatment effects. To censor time to CDP at dropout due to premature withdrawal and lost to follow-up after an initial disability progression is potentially to favor the most toxic or less efficacious treatment and, consequently, should be avoided whenever possible. Thus the purpose of this sensitivity analysis is to determine whether informative censoring has any influence on the primary endpoint.

Multiple imputations will be used to impute the events for patients who had initial disability progression and then discontinued treatment with no confirmatory relapses.

Multiple imputation inference involves three distinct phases:

- The missing data are filled in m times to generate m complete data sets. Instead of filling in a single value for each missing value, multiple imputation replaces each missing value with a set of m plausible values that represent the uncertainty about the right value to impute.
- The m complete data sets are analyzed using standard statistical analyses.
- The results from the m complete data sets are combined to produce inferential results.

3. A sensitivity analysis, in which the patients who had initial disability progression and then discontinued the treatment with no confirmatory EDSS assessments will be considered to not have confirmed disability progression, will be performed for the ITT population. This sensitivity analysis assumes that all patients lost to follow up after initial disability progression events will not have reached CDP. Hence, it will give an estimate of the maximal effect of informative censoring on the parameter estimates.

4. The influence of early progression events on treatment effect will also be evaluated by omitting the EDSS assessments performed between randomization and the Week 12 visit (≤ 83 days after randomization) with use of the ITT population.

4.5 Safety Analysis

All safety parameters will be summarized and presented in tables on the basis of this safety population. Patients who are not randomized but who receive study drug will be included in the safety population and summarized according to the therapy actually received. The safety data will be listed and summarized at the time of the primary analyses with use of all safety data available at the primary database lock, as well as at the end of the follow-up period with use of all safety data available.

4.5.1 Exposure of Investigational Medicinal Products

The amount of tocilizumab or azathioprine will be listed and summarized using descriptive statistics.

Definition of dose: A dose of tocilizumab is given as two infusions administered 4 weeks apart. Patients will be considered to have received a dose of treatment if at least part of one infusion of that dose (either Day 1 or Day 28 for dual infusions) was given. If a dose is completely missed instead of delayed, the next dose number will be based on the number of previous doses received.

The duration of observation for a patient will be calculated as follows:
(Date of last contact*−Date of first infusion in the first dose) + 1
*Earliest 1) date of clinical cutoff date for the primary analysis reporting; 2) date of subject completed or discontinued early from the study completion end of study page; or 3) date of death.

The duration of observation, within a dose, is defined in a similar manner as follows:
(Day prior to first infusion in the n + 1^{th} dose*−Date of first infusion in the n^{th} dose) + 1
*With the exception of the last dose received by the patient where the date of last contact is used as defined above. If the last contact is after the date of the clinical cutoff, the date of last contact will be the clinical cutoff date.

4.5.2 Adverse Events

Adverse events will be coded and tabulated by system organ class (SOC) and/or preferred term from the Medical Dictionary for Regulatory Activities (MedDRA) (version 21.1 or higher).
For each recorded AE, the term entered by the Investigator describing the event (the “reported term”) will be assigned a standardized term from the Common Terminology Criteria for Adverse

Events (CTCAE) v5.0, PT from the MedDRA. The term will also be assigned to a superclass term on the basis of the MedDRA.

Summaries of AEs will be generated to summarize the incidence of treatment-emergent AEs only. Treatment-emergent events are defined as those AEs with an observed or imputed date of onset on or after the start date of trial treatment. If the onset date of the AE is prior to the day of first dose, the AE will be considered treatment-emergent only if the most extreme intensity is greater than the initial intensity (i.e., the intensity for a given AE increases and its end date is on or after the date of the first dose). An AE with a completely missing, non-imputed start date will be assumed to be treatment emergent unless the AE has a complete, non-imputed end date that is prior to the date of the first dose.

The number of patients who experienced a related AE will be summarized by SOC and PT terms. AEs will be summarized by SOC and PT terms by intensity grade. For AEs leading to death, the most extreme intensity will be overwritten by Grade 5 (death). Any AEs and the SOC overall rows of the summary table will count patients according to AEs by intensity (grade).

The number of patients who experienced an SAE will be summarized by SOC and PT. Related SAEs will be summarized by SOC and PT terms.

A patient may experience an AE that leads to the discontinuation of his/her study treatment. Discontinuation of study treatment for an AE may not necessarily lead to discontinuation from the study because the patient can enter the safety follow-up period of the protocol. Only AEs that led to the discontinuation of study treatment are of interest. Patients who withdraw early from the study because of AEs will be summarized under disposition. The number of patients who experienced an AE that led to discontinuation of study treatment will be summarized by SOC and PT terms. The number of patients who experienced an AE that led to modification or interruption of study drug will be summarized by SOC and PT terms.

For each treatment group, the incidence count for each AE PT term will be defined as the number of patients reporting at least one treatment-emergent occurrence of the event. The incidence rate will be calculated as the incidence count divided by the total number of patients in the population. Each table will also present the overall number of patients experiencing at least one AE and the total number of AEs reported (multiple occurrences of the same AE in 1 patient will be counted only once).

4.5.3 Laboratory Data

All laboratory data will be converted to Standard International units. Summary tables will detail the actual values and changes from baseline of the laboratory parameters over visits up Week 60. Marked abnormalities will be summarized. Summaries of the number of patients by worst CTC grade for hematology and hepatic lab parameters will be produced (for summaries referring to NCI CTCAE, v5.0 grading will be used). For liver laboratory tests, shift tables will summarize the number of patients by baseline and worst post-baseline result up to Week 60.

5. REFERNCES

Hodges, J. L., Jr. and Lehmann, E. L. (1983), “Hodges-Lehmann Estimators,” in S. Kotz, N. L. Johnson, and C. B. Read, eds., Encyclopedia of Statistical Sciences, volume 3, 642–645, New York: John Wiley & Sons.

Hollander, M. and Wolfe, D. A. (1999), Nonparametric Statistical Methods, 2nd Edition, New York: John Wiley & Sons.

Levy M. The ethics of placebo controlled clinical trials in NMO - A balance of risks. Multiple sclerosis and related disorders 2015;4:512-4.

Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.

Zink RC and Koch GG. NParCov3: A SAS/IML Macro for Nonparametric Randomization Based Analysis of Covariance. Journal of Statistical Software 2012; 50(3): 1-17.

STATISTICAL ANALYSIS PLAN

TITLE: A RANDOMIZED, CONTROLLED, OPEN-LABEL, MULTI-CENTER TRIAL TO COMPARE THE SAFETY AND EFFICACY OF TOCILIZUMAB VERSUS AZATHIOPRINE IN PATIENTS WITH HIGHLY RELAPSING NEUROMYELITIS OPTICA SPECTRUM DISORDER (NMOSD): A HEAD-TO-HEAD COMPARATIVE STUDY (TANGO)

PROTOCOL NUMBER: 2017kylc005
CLINICALTRIALS.GOV NUMBER: NCT03350633
STUDY DRUG: Tocilizumab or Azathioprine
PLAN PREPARED BY: School of Public Health, Tianjin Medical University, Tianjin, China
DATE FINAL: 21 September 2018

TABLE OF CONTENT

1. BACKGROUND

The TANGO trial is an investigator-initiated, randomized, multi-center, parallel-group, controlled, open-label study to compare the safety and efficacy of tocilizumab versus azathioprine in patients with highly relapsing neuromyelitis optica spectrum disorder (NMOSD).

To date, no FDA-approved treatment has been demonstrated to significantly reduce the relapses and progression of disability in patients with NMOSD. Classical immunosuppressants such as azathioprine, mycophenolate, are being used for relapse prevention. Tocilizumab has been used for refractory NMOSD patients who did not well respond to immunosuppressants or rituximab. However, evidence of choice of these drugs for prevention of NMOSD still lacks. Thus, TANGO trial is designed as a head-to-head study ever conducted.

Growing evidence shows that autoimmune disorders, both organ-specific and non-organ-specific, are concomitant with NMOSD in > 20% of the patients. In TANGO trial, the disease course is highly relapsing and many patients have concomitant autoimmune diseases based on the baseline demographic information. Recognizing these comorbidities is undoubtedly important for proper management and preventing neurological and systemic disability over time. We will analyze the primary endpoint by pre-specifying the patients into two subgroups: NMOSD with and without concomitant autoimmune diseases.

2. STUDY DESIGN

The TANGO trial will evaluate the safety and efficacy of intravenous tocilizumab, compared with oral azathioprine in adults with highly relapsing NMOSD. Patients will be treated for a minimum of 60 weeks.

This study consists of the following periods: a Screening Period, an open-label Study Period, and a Safety Follow-up Period.

The primary analysis of the study will be performed when the last patient enrolled has been treated for at least 60 weeks or has a new relapse, or the end of the study.

2.1 Outcome Measures

2.1.1 Efficacy Outcome Measures

Primary outcome measures:

- Time to first relapse from the baseline in a time-to-event analysis

Secondary outcome measures:

- Time to onset of confirmed disability progression (CDP) for at least 12 weeks
- Determination of serum AQP4-IgG titers from baseline to 60 weeks
- Overall safety and tolerability of tocilizumab or azathioprine

Exploratory efficacy measures including:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

2.1.2 Safety Outcome Measures

The safety of tocilizumab or azathioprine will be assessed based on the all treatment-emergent adverse events (TEAEs), Common Terminology Criteria for Adverse Events (CTCAEs 5.0), serious adverse events (SAEs), and the changes from baseline through trial completion in vital signs, electrocardiogram (ECG), routine clinical laboratory tests, and pregnancy tests for female patients of childbearing potential.

2.2 Determination of Sample Size and Power

This is a randomized open-label parallel-group study, to evaluate the safety and efficacy of tocilizumab and azathioprine in patients with highly relapsing NMOSD. As such, the study is based on observing relapse events. With 1:1 randomization for the trial groups, we calculated that 118 patients would provide a power of 80% to determine the pre-specified between-group difference on the basis of a two-sided log-rank test at a 5% level of significance, assuming a 10% dropout rate. Randomized patients who discontinue after initiation of treatment will not be replaced.

The sample size and power calculation assumptions for this time to first event study are as follows:

- Log-rank test for comparison of tocilizumab and azathioprine
- 1:1 randomization (tocilizumab and azathioprine)
- Power 80%
- Two-sided 5% level of significance
- Dropout rate 10%
- Relapse-free rate of 85% for tocilizumab and 60% for azathioprine at 12 months (hazard ratio = 0.318)

The formula for calculating sample size is：

$$n_1 = n_2 = \frac{(\mu_\alpha+\mu_\beta)^2(1+h)^2}{[2-S_S(t)-S_N(t)](1-h)^2} \qquad \mathrm{h} = \frac{lnS_S(t)}{lnS_N(t)}$$

Parameters： $S_S(t)$ and $S_N(t)$ are the relapse-free rate of patients on the medication of tocilizumab and azathioprine, respectively. $\mu_{0.05/2}$=1.96, $\mu_{0.2}$=0.842.

3. STUDY CONDUCT

3.1 Randomizaiton

All patients who are evaluated eligible for randomization by the Investigator will be randomized on Day 1 on a 1:1 basis to the tocilizumab arm or the azathioprine arm. The randomization will be across centers. This is a completely randomized grouping experiment. The random numbers produced by a binomial distribution determine the group distribution of patients. The following SAS code had been used to do this:

```
data aa;
random=ranbin(11,1,0.5);
run;
```

Patients will be centrally randomized based randomization application.

3.2 Independent Review Facility

The primary outcome is the first relapse occurring after the baseline. The relapse will be confirmed by the Expert Panel. MRI scans will also be performed and read by the MRI radiologist and technicians, blinded to the treatment assignment and the clinical information.

The secondary efficacy endpoint will be derived from the EDSS values recorded at any visits. EDSS assessments are performed by an examining neurologist (not the treating neurologist or Principle Investigator). On serum AQP4-IgG titers, two independent assessors, who are unaware of clinical information, monitored and classified every blood sample.

4. STATISTICAL METHODS

4.1 Analysis Populations

One patient population will be defined for the purpose of the safety analysis, and two populations will be defined for the efficacy analysis. All efficacy endpoints will be analyzed using the intent-to-treat (ITT) population. The per-protocol population will be used for the primary efficacy endpoint only, to evaluate the influence of major protocol deviations and as a sensitivity check for the ITT analysis.

4.1.1 Intent-to-Treat Population

All randomized patients will be included in the ITT population. Patients who prematurely withdraw from the study for any reason or for whom an assessment is not performed for any reason will still be included in the ITT analysis.

4.1.2 Per-Protocol Population

The per-protocol population will include all patients in the ITT population who adhere to the protocol and will be summarized according to the randomization arm. Per-Protocol analyzes the patients receiving tocilizumab or azathioprine treatment as monotherapy after a washout period during which early concomitant immunosuppressants will be discontinued. Patients may be excluded if they violate the inclusion or exclusion criteria or deviate from the study plan. Only those patients with deviations that are deemed to potentially affect the efficacy of study treatment will be excluded from the per-protocol population.

4.1.3 Safety Population

The safety population will include all patients who received at least one dose of investigational medicinal products (tocilizumab or azathioprine). Randomized patients who receive the incorrect therapy will be summarized in the group according to the therapy actually received.

4.2 Analysis of Study Conduct

The following analyses will be conducted to evaluate the study conduct:

- Summary of protocol deviations
- Summaries of ITT, per protocol defined, and safety populations, including numbers of patients in each population, and reasons for exclusion from the per-protocol population
- Summary of subject disposition, including the number of treatment doses received, and the number of patients entering into safety follow-up

4.3 Analysis of Treatment Group Comparability

For continuous variables, the mean, median, standard deviation (SD), median (interquatile range), and minimum and maximum values will be calculated. For categorical variables, the number and percentage in each category will be displayed. For each item in the following lists, the units and categories to be used are indicated in parentheses and separated by commas. All durations are calculated with respect to the date of randomization, if not stated otherwise.

Demography and stratification factors based on Case Report Form (CRF) data:

- Age (years) at baseline
  Age stratification category (≤ 45, > 45 years)
- Sex (male patients, female patients)
- Weight (kg)
- Body mass index (measured in kg/m^2)

Baseline NMOSD disease history:

- AQP4-IgG seropositivity (%)
- Duration since NMOSD symptom onset (calculated in years; i.e., divide by 365.25)
  Duration stratification category (≤ 6, > 6 years)
- Total number of previous relapses
  Number of relapses stratification category (≤ 8, > 8)
- Number of relapses 1 year before randomization
  Number of relapses 1 year before randomization stratification category (≤ 1, > 1)
- Number of relapses 2 year before randomization
  Number of relapses 2 year before randomization stratification category (≤ 3, > 3)

Baseline disease characteristics:

- EDSS (continuous) ： EDSS stratification category (≤4.5, > 4.5 years)

Baseline serum AQP4-IgG titer data

- Serum AQP4-IgG titer

Baseline visual acuity data

- 100% high-contrast visual acuity
- 2.5% low-contrast letter acuity (LCLA)

Baseline OCT/VEP data

- Retinal nerve fiber layer (RNFL) thickness
- Retinal ganglion cell complex (GCC) volume
- P100 latency
- P100 amplitude

Baseline MRI data:

- Whole-brain volume at baseline
- Number of T2 lesions at baseline

4.4 Efficacy Analysis

All statistical hypotheses for the primary and secondary endpoints and treatment comparisons will be tested at the 5% significance level (α= 0.05) against two-sided alternatives. The analysis are pre-specified by two subgroups of the patients: with and without concomitant autoimmune diseases.

4.4.1 Primary Efficacy Endpoint

Significance Level

The null hypothesis will be tested at the α= 0.05 level (two-sided test).

The hypotheses to be tested are:

- H0 (null hypothesis): There is no difference in the probability of relapse-free between the tocilizumab and azathioprine groups.
- H1 (alternative hypothesis): There is a difference in the probability of relapse-free between the tocilizumab and azathioprine groups.

Definition

The time to first relapse is defined as the time from baseline to the onset of the first new relapse during the follow-ups.

Analysis Methods

The ITT population analysis will be presented. Patients who did not have initial relapse at the time of end of the trial, time of early discontinuation, or loss to follow-up will be censored at the date of their last EDSS assessment that occurred during the treatment period. The proportion of patients who do not have a new relapse during the trial at predefined time points (i.e., 1 year, 2 years, etc.) will be estimated using Kaplan-Meier methodology. The overall hazard ratio will be estimated using a stratified Cox regression model with the same stratification factors used in the aforementioned stratified log-rank test.

SAS code for primary endpoint

The primary endpoint is time to the first relapse. The comparison of the treatment groups for the primary endpoint will use a two-sided log-rank test including strata for the randomized intervention variable— different medication. The basic SAS code for this log-rank analysis is:

```
proc lifetest data=trial;
time ftgapw*relapse(0);
strata group;
run;
```

Where group is a variable that indicates patient's randomized treatment group, ftgapw is a variable for the patient's time in study period at time of relapse or censoring, relapse is the censoring variable (0 = no event (censored), 1= event).

The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model. The basic SAS code for this analysis is:

```
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
run;
```

SAS code for primary endpoint with main covariates

The comparison of the situation of concomitant autoimmune diseases for the primary endpoint will use a log-rank test in two randomized treatment group, respectively. The basic SAS code for this is:

```
data aza;
set trial(where=(group=0));
run;
proc lifetest data=aza;
time ftgapw*relapse(0);
strata concom;
run;
proc phreg data=aza;
model ftgapw*relapse(0)=concom/RL;
run;

data tcz;
set trial(where=(group=1));
run;
proc lifetest data=tcz;
time ftgapw*relapse(0);
strata concom;
run;
proc phreg data=tcz;
model ftgapw*relapse(0)=concom/RL;
run;
```

Where concom is a variable that indicates the patient's status of concomitant autoimmune diseases.

SAS code for primary endpoint by subset

Stratified Cox proportional-hazards models in different strata of EDSS at randomization, age at randomization and duration of disease are conducted to estimate the hazard ratio for the primary end point. The basic SAS code for this is:

```
proc sort data=trial;
by edssstra;
run;
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by edssstra;
run;
proc sort data=trial;
by relapstra;
run;
```

```
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by relapstra;
run;
proc sort data=trial;
by hisystra;
run;
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by hisystra;
run;
proc sort data=trial;
by agestra;
run;
proc phreg data=trial;
model ftgapw*relapse(0)=group/RL;
by agestra;
run;
data trial11;
set trial(where=(rn1a>1));
run;
data trial12;
set trial(where=(rn1a<=1));
run;
proc lifetest data=trial11;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial11;
model ftgapw*relapse(0)=group/RL;
run;
proc lifetest data=trial12;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial12;
model ftgapw*relapse(0)=group/RL;
run;
data trial21;
set trial(where=(rn2a>3));
run;
data trial22;
set trial(where=(rn2a<=3));
run;
```

```
proc lifetest data=trial21;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial21;
model ftgapw*relapse(0)=group/RL;
run;
proc lifetest data=trial22;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial22;
model ftgapw*relapse(0)=group/RL;
run;
```

Where edssstra is a variable that described the grade of EDSS status at randomization (1 = baseline EDSS score > 4.5, 2 = baseline EDSS score ≤ 4.5). Variable relapstra described the hierarchy of the number of relapse (1 = relapse number > 8, 2 = relapse number ≤ 8). Variable hisystra described the grade of disease duration (1 = disease duration > 6 years, 2 = disease duration ≤ 6 years). Variable agestra described the hierarchy of age at randomization (1 = age at randomization > 45 years, 2 = age at randomization ≤ 45 years). Variable rn1a described the number of relapse 1 year before randomization. Variable rn2a described the number of relapse 2 year before randomization.

4.4.2 Secondary Efficacy Endpoints

4.4.2.1 12-week Confirmed Disability Progression

Definition

The time to onset of confirmed disability progression (12-week confirmation [weeks]) is defined as the time from baseline to the onset of the first disability progression that is confirmed at the next regularly scheduled visit ≥ 12 weeks after the initial disability progression. If the patient has an infection, dosing may not occur on the Day 1 visit.

Baseline for the time to onset of confirmed disability is the date of randomization. Disability progression is defined as an increase of ≥ 1.0 point from baseline EDSS score if the baseline EDSS value is ≤ 5.5 points (inclusive) or an increase of ≥ 0.5 points if the baseline EDSS value is > 5.5 points. Assessments within 30 days after a protocol-defined relapse will not be used for confirmation of confirmed disability progression. The non-confirmatory EDSS assessments (between the initial disability progression and the confirmation of disability progression should also fulfill the requirements of the progression. Otherwise, the initial disability progression is not confirmed.

The baseline EDSS value is the average score of the EDSS assessment at screening and baseline (Day 1 visit) up to and including the date of randomization. If one of the values is missing, the non-missing values will be used as baseline.

A blinded rater at each study site will assess EDSS for all patients at the site at screening, baseline, every 4 weeks (regularly scheduled visit) during the blinded treatment period of the study, during the safety follow-up period, at any unscheduled visits, and at withdrawal-from-treatment and end-of-study visits. Additional EDSS assessments for individual patients may be requested between visits (i.e., during an NMOSD relapse).

The examining investigator is not the physician responsible for the patient care (the treating investigator).

The EDSS is based on a standard neurological examination; the seven categories of the EDSS representing FS (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, and cerebral [or mental], plus "other") are rated and scored (collectively, FSs), in addition to an ambulation score (0−12). Each domain of the FS is an ordinal clinical rating scale from 0 to 5 or 6. These ratings are then used in conjunction with observations and information regarding ambulation and the use of assistive devices (which will also be scored) to determine the EDSS score. The EDSS is a disability scale that ranges in 0.5-point steps from 0 (normal) to 10.0 (death).

Analysis Methods

The ITT population analysis will be presented. Patients who did not have initial disability progression at the time of end of the trial, time of early discontinuation, or loss to follow-up will be censored at the date of their last EDSS assessment that occurred during the treatment period. There are various options to obtain EDSS assessments (confirmatory and non-confirmatory) and to record how the discontinuation from treatment and study can influence the confirmation of the disability progression. The following rules apply for all patients with initial disability progression:

- Initial disability progression must occur when the patient is still on treatment.
- All non-confirmatory EDSS assessments (if any) after initial disability progression and up to and including the confirmatory EDSS assessment should also fulfill the requirements of the progression as defined above. Otherwise an initial disability progression event cannot be confirmed.

The comparison of the treatment groups for secondary efficacy endpoint--12 weeks confirmed disability progression will use a log-rank test including strata for the randomized intervention variable. The hazard ratio and risk reduction will be estimated from a stratified Cox proportional hazards (PH) model. The basic SAS code for this log-rank analysis is:

```
proc lifetest data=trial;
time edssgapw*edsspro(0);
strata group;
run;
proc phreg data=trial;
model edssgapw*edsspro(0)=group/RL;
run;
```

The comparison of the EDSS score changes from baseline between two groups will use Wilcoxon rank sum test, *t* test and chi-square test. The basic SAS code for this is:

```
proc npar1way data=trial HL(REFCLASS=2) wilcoxon;
class group;
var edsschange;
run;
proc univariate data=trial normal;
class group;
var edsschange;
run;
proc ttest test=diff data=trial;
class group;
var edsschange;
run;
proc freq data=trial;
tables edsschange*group/chisq fisher;
run;
```

4.4.2.2 Changes of serum AQP4-IgG titers by CBA and FIPA

Two methods are used to dynamically evaluate the serum AQP4-IgG titers of patients in this study at the time of baseline, 12 weeks, 24 weeks, 36 weeks, 48 weeks and 60 weeks from the randomization. However, some patients' serum AQP4-IgG titers cannot be obtained because of serum AQP4-IgG seronegativity. The number of patients got tested data were shown on the corresponding figure. Wilcoxon rank sum test is used to compare the serum AQP4-IgG titers of two groups. The basic SAS code for this is:

```
proc univariate normal data=trial normal;
class group;
var AQP40WC AQP412WC AQP424WC AQP436WC AQP448WC AQP460WC AQP40WF
AQP412WF AQP424WF AQP436WF AQP448WF AQP460WF;
run;
proc npar1way data=trial wilcoxon;
class group;
var AQP40WC AQP412WC AQP424WC AQP436WC AQP448WC AQP460WC AQP40WF
AQP412WF AQP424WF AQP436WF AQP448WF AQP460WF;
run;
data trialaqp;
set trial;
a5 = AQP460WC - AQP40WC;
b5 = a5/ AQP40WC*100;
run;
proc npar1way data=trialaqp HL(REFCLASS=2) wilcoxon;
class group;
```

```
var a5 b5;
run;
```

4.4.3 SAS Code for Demographic and Clinical Characteristics of the Patients at Baseline

Demographic and clinical characteristics of the patients at baseline are compared between two groups. Continuous Variables, (age, EDSS at randomization, disease duration), are analyzed using t-test or Wilcoxon rank sum test according to the distribution of different variables. Qualitative variables, (gender, AQP4-IgG seropositivity, the occurrence of first relapse), are analyzed using chi-square test. The basic SAS code for this is:

```
proc freq data=trial;
tables (gender AQP4 relapse )*group/chisq;
run;

proc univariate normal data=trial normal;
class group;
var age hisy edss0w;
run;
proc ttest cochran data=trial;
class group;
var hisy;
run;

proc npar1way data=trial wilcoxon;
class group;
var age edss0w;
run;
```

Where AQP4 is a variable that indicates AQP4-IgG seropositivity of patient at baseline, hisy indicates the duration of disease, edss0w indicates the EDSS score of patient at baseline, relapse is the censoring variable (0 = no event (censored), 1= event).

4.4.4. SAS Code and Analysis Method (PP)

There are 5 patients dropping out from this trial, 2 (1 had severe adverse event; 1died) of them in tocilizumab group, and 3 (2 had severe adverse event; 1died) of them in azathioprine group. Per-protocol (PP) analysis is conducted to evaluate the stability of the main results of intention-to-treat (ITT) analysis above. The basic SAS code for this is:

```
data trialpp;
set trial(where=(relaperiod^=0));
run;
data trialpp1;
set trialpp(where=(withdraw=1));
run;
```

```
proc lifetest data=trialpp1;
time ftgapwpp*relapse(0);
strata group;
run;
proc phreg data=trialpp1;
model ftgapwpp*relapse(0)=group/RL;
run;
data concom;
set trialpp1(where=(concom=1));
run;
data noconcom;
set trialpp1(where=(concom=2));
run;
proc lifetest data=concom;
time ftgapwpp*relapse(0);
strata group;
run;
proc phreg data=concom;
model ftgapwpp*relapse(0)=group/RL;
run;
proc lifetest data=noconcom;
time ftgapwpp*relapse(0);
strata group;
run;
proc phreg data=noconcom;
model ftgapwpp*relapse(0)=group/RL;
run;
```

Where withdraw is a variable that indicates whether patients withdraw from the trial or not (1=not withdraw; 0=withdraw). And relaperiod indicates the medication period of relapse (0=relapse happened in combination medication period; 1=relapse happened in individual medication period). Variable ftgapwpp is the time from individual medication to first relapse. Where concom is a variable that indicates the patient's status of concomitant autoimmune diseases.

4.4.5 Time to Onset of Confirmed Disability Progression Confirmed for ≥24 Weeks

In addition to the primary analysis, the time to onset of confirmed disability progression will be analyzed in the tocilizumab and azathioprine treatment arms with use of a 24-week confirmation window for disability progression. The analysis of EDSS progression will be conducted as described for the primary analysis, with the exception that the time to onset of confirmed disability progression (24-week confirmation) is defined as the time from baseline to the first disability progression that is confirmed at the next regularly scheduled visit ≥ 24 weeks (≥ 161 days) after the initial disability progression. The basic SAS code for this log-rank analysis is:

```
proc lifetest data=trial;
time edsspg2*edssp2(0);
strata group;
run;
proc phreg data=trial;
model edsspg2*edssp2(0)=group/RL;
run;
```

4.4.6 Exploratory Efficacy Endpoints

The following exploratory efficacy outcomes will be analyzed:

- Time to onset of confirmed disability progression (CDP) for at least 24 weeks
- Change of high-contrast visual acuity (VA) from baseline to 60 weeks
- Change of low-contrast letter acuity (LCLA) from baseline to 60 weeks
- Change of average retinal nerve fiber layer (RNFL) thickness from baseline to 60 weeks measured by spectral-domain optical coherence tomography (SD-OCT)
- Change of average retinal ganglion cell complex (GCC) volume from baseline to 60 weeks measured by SD-OCT
- Change of P100 latency from baseline to 60 weeks measured in visual evoked potentials (VEP)
- Change of P100 amplitude from baseline to 60 weeks in VEP
- Number of new and/or enlarging T2 hyperintense lesions as detected by brain and spinal cord magnetic resonance imaging (MRI)
- Change of counts of peripheral blood B cell subsets measured by flow cytometry

Other exploratory efficacy endpoints listed in the protocol but not included in this SAP, will be reported in separate study reports or publications.

4.4.7 Sensitivity Analyses

The following sensitivity analyses of the primary endpoint will be conducted according to the analysis described in Section 4.4.1:

1. The primary analysis will be repeated using the per-protocol population as the analysis population. The below sensitivity analyses will be done with the ITT population.

2.

2. A sensitivity analysis using multiple imputations will be performed for the ITT population to explore the potential influence of informative censoring on the results of the primary efficacy analyses. Censoring for a reason that is not independent of that patient’s prognosis is called “informative censoring.” The influence of informative censoring needs to be explored, because for time-to-event endpoints, the intent-to-treat principle requires that essentially all randomly assigned patients be observed up to the endpoint or up to the end of the study. When a patient's follow-up is censored X number of months after random assignment, then, in the computation of the Kaplan-Meier estimates, log-rank or Cox regressions, that patient's outcome after X number of months is assumed to have the same outcome as the other patients in their treatment group who also are free of the outcome at X number of months and who remain under follow-up beyond X number of months.

Thus, unless the reason for being censored is independent of that patient's prognosis, failure to observe that patient until occurrence of his study endpoints could lead to significant bias as well as increased variability in the evaluation of treatment effects. To censor time to CDP at dropout due to premature withdrawal and lost to follow-up after an initial disability progression is potentially to favor the most toxic or less efficacious treatment and, consequently, should be avoided whenever possible. Thus the purpose of this sensitivity analysis is to determine whether informative censoring has any influence on the primary endpoint.

Multiple imputations will be used to impute the events for patients who had initial disability progression and then discontinued treatment with no confirmatory relapses.

Multiple imputation inference involves three distinct phases:

- The missing data are filled in m times to generate m complete data sets. Instead of filling in a single value for each missing value, multiple imputation replaces each missing value with a set of m plausible values that represent the uncertainty about the right value to impute.
- The m complete data sets are analyzed using standard statistical analyses.
- The results from the m complete data sets are combined to produce inferential results.

3. A sensitivity analysis, in which the patients who had initial disability progression and then discontinued the treatment with no confirmatory EDSS assessments will be considered to not have confirmed disability progression, will be performed for the ITT population. This sensitivity analysis assumes that all patients lost to follow up after initial disability progression events will not have reached CDP. Hence, it will give an estimate of the maximal effect of informative censoring on the parameter estimates.

4. The influence of early progression events on treatment effect will also be evaluated by omitting the EDSS assessments performed between randomization and the Week 12 visit (≤ 83 days after randomization) with use of the ITT population.

4.5 Safety Analysis

All safety parameters will be summarized and presented in tables on the basis of this safety population. Patients who are not randomized but who receive study drug will be included in the safety population and summarized according to the therapy actually received. The safety data will be listed and summarized at the time of the primary analyses with use of all safety data available at the primary database lock, as well as at the end of the follow-up period with use of all safety data available.

4.5.1 Exposure of Investigational Medicinal Products

The amount of tocilizumab or azathioprine will be listed and summarized using descriptive statistics.

Definition of dose: A dose of tocilizumab is given as two infusions administered 4 weeks apart. Patients will be considered to have received a dose of treatment if at least part of one infusion of

that dose (either Day 1 or Day 28 for dual infusions) was given. If a dose is completely missed instead of delayed, the next dose number will be based on the number of previous doses received.

The duration of observation for a patient will be calculated as follows:
(Date of last contact*−Date of first infusion in the first dose) + 1
*Earliest 1) date of clinical cutoff date for the primary analysis reporting; 2) date of subject completed or discontinued early from the study completion end of study page; or 3) date of death.

The duration of observation, within a dose, is defined in a similar manner as follows:
(Day prior to first infusion in the n + 1^{th} dose*−Date of first infusion in the n^{th} dose) + 1
*With the exception of the last dose received by the patient where the date of last contact is used as defined above. If the last contact is after the date of the clinical cutoff, the date of last contact will be the clinical cutoff date.

4.5.2 Adverse Events

Adverse events will be coded and tabulated by system organ class (SOC) and/or preferred term from the Medical Dictionary for Regulatory Activities (MedDRA) (version 21.1 or higher).
For each recorded AE, the term entered by the Investigator describing the event (the "reported term") will be assigned a standardized term from the Common Terminology Criteria for Adverse Events (CTCAE) v5.0, PT from the MedDRA. The term will also be assigned to a superclass term on the basis of the MedDRA.

Summaries of AEs will be generated to summarize the incidence of treatment-emergent AEs only. Treatment-emergent events are defined as those AEs with an observed or imputed date of onset on or after the start date of trial treatment. If the onset date of the AE is prior to the day of first dose, the AE will be considered treatment-emergent only if the most extreme intensity is greater than the initial intensity (i.e., the intensity for a given AE increases and its end date is on or after the date of the first dose). An AE with a completely missing, non-imputed start date will be assumed to be treatment emergent unless the AE has a complete, non-imputed end date that is prior to the date of the first dose.

The number of patients who experienced a related AE will be summarized by SOC and PT terms. AEs will be summarized by SOC and PT terms by intensity grade. For AEs leading to death, the most extreme intensity will be overwritten by Grade 5 (death). Any AEs and the SOC overall rows of the summary table will count patients according to AEs by intensity (grade).

The number of patients who experienced an SAE will be summarized by SOC and PT. Related SAEs will be summarized by SOC and PT terms.

A patient may experience an AE that leads to the discontinuation of his/her study treatment. Discontinuation of study treatment for an AE may not necessarily lead to discontinuation from the study because the patient can enter the safety follow-up period of the protocol. Only AEs that led to the discontinuation of study treatment are of interest. Patients who withdraw early from the study because of AEs will be summarized under disposition. The number of patients who

experienced an AE that led to discontinuation of study treatment will be summarized by SOC and PT terms. The number of patients who experienced an AE that led to modification or interruption of study drug will be summarized by SOC and PT terms.

For each treatment group, the incidence count for each AE PT term will be defined as the number of patients reporting at least one treatment-emergent occurrence of the event. The incidence rate will be calculated as the incidence count divided by the total number of patients in the population. Each table will also present the overall number of patients experiencing at least one AE and the total number of AEs reported (multiple occurrences of the same AE in 1 patient will be counted only once).

4.5.3 Laboratory Data

All laboratory data will be converted to Standard International units. Summary tables will detail the actual values and changes from baseline of the laboratory parameters over visits up Week 60. Marked abnormalities will be summarized. Summaries of the number of patients by worst CTC grade for hematology and hepatic lab parameters will be produced (for summaries referring to NCI CTCAE, v5.0 grading will be used). For liver laboratory tests, shift tables will summarize the number of patients by baseline and worst post-baseline result up to Week 60.

5. REFERNCES

Freitas E, Guimarães J. A review of the current literature and a guide to the early diagnosis of autoimmune disorders associated with neuromyelitis optica. Rheumatol Int. 2015;35:243-53.

Iyer A, Elsone L, Appleton R, Jacob A. A review of the current literature and a guide to the early diagnosis of autoimmune disorders associated with neuromyelitis optica. Autoimmunity. 2014;47:154-61.

Hodges, J. L., Jr. and Lehmann, E. L. (1983), “Hodges-Lehmann Estimators,” in S. Kotz, N. L. Johnson, and C. B. Read, eds., Encyclopedia of Statistical Sciences, volume 3, 642–645, New York: John Wiley & Sons.

Hollander, M. and Wolfe, D. A. (1999), Nonparametric Statistical Methods, 2nd Edition, New York: John Wiley & Sons.

Levy M. The ethics of placebo controlled clinical trials in NMO - A balance of risks. Multiple sclerosis and related disorders 2015;4:512-4.

Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.

Zink RC and Koch GG. NParCov3: A SAS/IML Macro for Nonparametric Randomization Based Analysis of Covariance. Journal of Statistical Software 2012; 50(3): 1-17.

TABLE OF CONTENTS FOR

TANGO STATISTICAL ANALYSIS PLAN (SAP) AMENDMENT

(VERSION 2.0)

TANGO STATISTICAL ANALYSIS PLAN (SAP) AMENDMENT, VERSION 2.0: RATIONALE

TANGO statistical analysis plan (SAP) has been amended for the following reasons:

- **To add Clinicaltrials.gov, Number on the cover page**
- **To pre-specify NMOSD patients into two subgroups: with and without concomitant autoimmune diseases in the primary outcome analysis**

  Neuromyelitis optica sepctrum disorder (NMOSD) is an immune-mediated neurological disorder characterized by recurrent episodes of optic neuritis and longitudinally extensive transverse myelitis. NMOSD may be associated with a variety of different types of autoimmune diseases, since growing evidence shows that autoimmune disorders, both organ-specific and non-organ-specific, have been concomitant with NMOSD in > 20% of the patients.The course of the disease may be affected when two autoimmune diseases coincide. Recognizing these comorbidities is undoubtedly important for proper management and preventing neurological and systemic disability over time. So far, there has been no study to evaluate the superiority of a particular drug group in treating the autoimmune diseases associated with NMOSD. In TANGO trial, the disease course is highly relapsing and many patients have concomitant autoimmune diseases based on the baseline demographic information. And these comorbidities may affect the efficacy of the treatment strategy.We will analyze the primary endpoint by pre-specifying the patients into two subgroups: NMOSD with and without concomitant autoimmune diseases.
- **To update the stratification variables in analysis of reduced disease activity between the two groups**

  The inclusion criteria “highly relapsing” refers to high annualized attack rate in two years before randomization. The patients may have just one relapse in one year before randomization. Comparing the efficacy of the drugs on reducing the disease activity, the patients will be stratified into two subgroups based on historical relapses in one year or in two years before randomization.
- **To add the analysis on the proportion of the patients in each group who has worsening in EDSS score from baseline at last visit**

  Disability of the patients may be improved for the patients who have no relapses in NMOSD since the interval between the most recent relapse before randomization and the time of randomization may be not long enough to result in irreversible disability. This change is to clarify the overall disability accumulation among the patients, who care much about the efficacy of the drugs in their ambulance.
- **To clarify the populations of determined serum AQP4-IgG titers**

  For patients with AQP4-IgG seronegative, AQP4-IgG will not be determined. Occasionally, there may be missing data for the patients with AQP4-IgG seropositive at each protocol-defined visits.
- **To add two related references**

  1) Freitas E, Guimarães J. A review of the current literature and a guide to the early diagnosis of autoimmune disorders associated with neuromyelitis optica. Rheumatol Int.

2015;35:243-53.
2) Iyer A, Elsone L, Appleton R, Jacob A. A review of the current literature and a guide to the early diagnosis of autoimmune disorders associated with neuromyelitis optica. Autoimmunity. 2014;47:154-61.

TANGO STATISTICAL ANALYSIS PLAN (SAP) AMENDMENT,

VERSION 2.0: SUMMARY OF CHANGES

COVER PAGE:
PROTOCOL NUMBER: 2017kylc005
CLINICALTRIALS.GOV NUMBER: NCT03350633
STUDY DRUG: Tocilizumab or Azathioprine

SECTION 1. BACKGROUND

Growing evidence shows that autoimmune disorders, both organ-specific and non-organ-specific, are concomitant with NMOSD in > 20% of the patients. In TANGO trial, the disease course is highly relapsing and many patients have concomitant autoimmune diseases based on the baseline demographic information. Recognizing these comorbidities is undoubtedly important for proper management and preventing neurological and systemic disability over time. We will analyze the primary endpoint by pre-specifying the patients into two subgroups: NMOSD with and without concomitant autoimmune diseases.

SECTION 4.4.1 Primary Efficacy Endpoint

SAS code for primary endpoint by subset

Stratified Cox proportional-hazards models in different strata of EDSS at randomization, age at randomization and duration of disease are conducted to estimate the hazard ratio for the primary end point. The basic SAS code for this is (added):

```
data trial11;
set trial(where=(rn1a>1));
run;
data trial12;
set trial(where=(rn1a<=1));
run;
proc lifetest data=trial11;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial11;
model ftgapw*relapse(0)=group/RL;
run;
proc lifetest data=trial12;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial12;
model ftgapw*relapse(0)=group/RL;
```

```
run;
data trial21;
set trial(where=(rn2a>3));
run;
data trial22;
set trial(where=(rn2a<=3));
run;
proc lifetest data=trial21;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial21;
model ftgapw*relapse(0)=group/RL;
run;
proc lifetest data=trial22;
time ftgapw*relapse(0);
strata group;
run;
proc phreg data=trial22;
model ftgapw*relapse(0)=group/RL;
run;
```

Where edssstra is a variable that described the grade of EDSS status at randomization (1 = baseline EDSS score > 4.5, 2 = baseline EDSS score ≤ 4.5). Variable relapstra described the hierarchy of the number of relapse (1 = relapse number > 8, 2 = relapse number ≤ 8). Variable hisystra described the grade of disease duration (1 = disease duration > 6 years, 2 = disease duration ≤ 6 years). Variable agestra described the hierarchy of age at randomization (1 = age at randomization > 45 years, 2 = age at randomization ≤ 45 years). *Variable rn1a described the number of relapse 1 year before randomization. Variable rn2a described the number of relapse 2 year before randomization.*

SECTION 4.4.2.1 12-week Confirmed Disability Progression

The comparison of the EDSS score changes from baseline between two groups will use Wilcoxon rank sum test, t test and chi-square test. The basic SAS code for this is:

```
proc npar1way data=trial HL(REFCLASS=2) wilcoxon;
class group;
var edsschange;
run;
proc univariate data=trial normal;
class group;
var edsschange;
run;
proc ttest test=diff data=trial;
```

class group;
var edsschange;
run;
proc freq data=trial;
*tables edsschange*group/chisq fisher;*
run;

SECTION 4.4.2.2 Changes of Serum AQP4-IgG Titers by CBA and FIPA

Two method had been used to dynamically evaluate the serum AQP4-IgG titers of patients in this study at the time of baseline, 12 weeks, 24 weeks, 36 weeks, 48 weeks and 60 weeks from the randomization. *However, some patients' serum AQP4-IgG titers cannot be obtained because of serum AQP4-IgG seronegativity. The number of patients got tested data were shown on the corresponding figure.* Wilcoxon rank sum test is used to compare the serum AQP4-IgG titers of two groups.

SECTION 5. REFERENCES

Freitas E, Guimarães J. A review of the current literature and a guide to the early diagnosis of autoimmune disorders associated with neuromyelitis optica. Rheumatol Int. 2015;35:243-53.

Iyer A, Elsone L, Appleton R, Jacob A. A review of the current literature and a guide to the early diagnosis of autoimmune disorders associated with neuromyelitis optica. Autoimmunity. 2014;47:154-61.