Endocrine Journal 2017, **64** (2), 141-149

ORIGINAL

Comparison of two different regimens of intravenous methylprednisolone for patients with moderate to severe and active Graves' ophthalmopathy: a prospective, randomized controlled trial

Yebing He[1)], Ketao Mu[2)], Rong Liu[3)], Jing Zhang[2)] and Nan Xiang[3)]

[1)] *Department of Endocrinology and Metabolism, The Affiliated Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan 430030, P.R. China*
[2)] *Department of Radiology, The Affiliated Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan 430030, P.R. China*
[3)] *Department of Ophthalmology, The Affiliated Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan 430030, P.R. China*

Abstract. The intravenous methylprednisolone (iv MP) strategy for Graves' ophthalmopathy (GO) and evaluation of its activity against the disease warrants further exploration. A prospective randomized controlled trial for 3 months was performed in a tertiary referral teaching hospital to compare the efficacy and safety of two different regimens of iv MP, and determine the value of clinical activity score (CAS) and T2 relaxation time (T2RTs) and areas of extraocular muscles (EOMs) by magnetic resonance imaging for diagnosis of active GO. Forty patients with moderate to severe GO and CAS ≥ 3 or $1 \leq \text{CAS} < 3$ with prolonged T2RTs on EOMs were randomly assigned to a monthly (MR: 1.5 g iv MP monthly for 3 months) or weekly (WR: 0.5 g iv MP weekly for 6 weeks, followed by 0.25 g weekly for 6 weeks) regimen. Overall response based on ophthalmic symptoms, T2RTs, areas of EOMs and adverse effects were recorded at each visit. The total rate of response was 71.9%. Rates of improved, unchanged, deteriorated were similar between the MR and WR groups ($p>0.05$). The maximum T2RTs and areas significantly decreased at the end of intervention in both groups ($p<0.05$). Results show that both MR and WR are effective and safe in treatment of GO. T2RTs combined with CAS can sensitively detect active GO and predict the response to iv MP.

Key words: Graves' ophthalmopathy, Magnetic resonance imaging, Methylprednisolone

GRAVES' OPHTHALMOPATHY (GO) remains a clinical problem to be resolved [1, 2]. Randomized controlled trials (RCTs) [3, 4] have demonstrated the superior effectiveness and tolerance of intravenous methylprednisolone (iv MP) compared to oral glucocorticoids (GCs) for patients with moderate to severe and active GO. The efficacy and safety of the administration of different cumulative doses of iv MP have been investigated in nonrandomized and randomized trials [5]. The commonly recommended treatment is a course of 0.5 g of iv MP once a week for 6 weeks, followed by 0.25 g weekly for 6 weeks (cumulative dose of 4.5 g) [1]. Recently, comparative studies of different cumulative doses and administration methods of iv MP have been performed [6–8].

In previous RCTs, active GO was confirmed by clinical activity score (CAS) [7, 8]. However, CAS is a subjective and qualitative measure and is not sufficient to adequately detect active GO on its own [9, 10]. Recent studies [11–15] have demonstrated that a prolonged T2 relaxation time (T2RT) determined by orbital magnetic resonance imaging could sensitively detect edematous changes in acutely inflamed extraocular muscles (EOMs). Thus, T2RT may quantitatively and objectively reflect disease activity, and could be

Submitted Feb. 19, 2016; Accepted Sep. 17, 2016 as EJ16-0083
Released online in J-STAGE as advance publication Nov. 16, 2016
Correspondence to: Yebing He, Department of Endocrinology and Metabolism, The Affiliated Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, #1095 Jiefang Road, Wuhan 430030, P.R. China.
E-mail: ybhe@tjh.tjmu.edu.cn
Correspondence to: Jing Zhang, M.D., Department of Radiology, The Affiliated Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, #1095 Jiefang Road, Wuhan 430030, P.R. China. E-mail: hbclleo@163.com
Clinical Trial Registration Number: ChiCTR-IPR-15006848

useful marker for predicting the response to iv MP. We designed a prospective RCT to compare two regimens of cumulative doses and administration methods of iv MP, and to evaluate disease activity according to CAS in combination with T2RT.

Subjects and Methods

From March 2011 to July 2012, patients diagnosed with GO at Tongji Hospital were invited to participate in the study. The inclusion criteria were as follows: (1) patients, aged 18–70 years, were included if they had not received any immunosuppressive therapy, orbital radiotherapy, or surgery. (2) They could be either hyperthyroid or hypothyroid status upon enrollment, but only patients whose thyroid status was restored to normal levels within 2–4 weeks from antihyperthyroid drugs or levothyroxine with adjustment medication were included in the study. These patients remained in euthyroid status during the treatment. (3) Patients with intracranial diseases and other eye diseases were excluded. (4) Moderate to severe GO must meet at least one of the following three items, which was based on the European Group on Graves' Orbitopathy (EUGOGO) consensus statement [1]: lid retraction $\geq$ 2 mm, exophthalmos $\geq$ 3 mm above the upper normal limit of 18 mm, diplopia (inconstant or constant). (5) Active GO (CAS $\geq$ 3) [1] or inactive GO (1 $\leq$ CAS $<$ 3) coupled with a prolonged T2RT on at least one of EOMs [14]. The exclusion criteria were as follows: (1) contraindications to GCs, including active or unexplained gastric disease, active liver disease, a history of pulmonary tuberculosis, heart arrhythmias, uncontrolled hypertension, uncontrolled diabetes mellitus, or glaucoma, (2) having previously received iv MP, (3) symptoms and signs of other eye diseases, (4) recurrent GO, (5) dysthyroid optic neuropathy (DON). The criteria for premature termination were: (1) allergic to GCs or unable to tolerate iv MP, (2) aggravated symptoms after treatment, (3) unable to complete regular visits to the clinic.

Overall ophthalmic measures were evaluated at baseline and at each visit. The same ophthalmologist recorded soft tissue involvements, palpebral aperture (in millimetres), exophthalmos (in millimetres), duction (elevation, depression, adduction, abduction), diplopia, visual acuity (in decimals using the Snellen chart), intraocular pressure (mmHg), and CAS. The 7 point CAS [9, 10] and the modified NOSPECS classification results were evaluated as previously reported [16, 17]. Duction of eyeball was recorded as follows: 0, no restriction. 1, little restriction. 2, obvious restriction. 3, total restriction. The subjective diplopia score was assessed according to the EUGOGO consensus statement [1] as follows: 0, no diplopia. 1, intermittent. 2, inconstant. 3, constant. Adverse events were also recorded at each visit.

Blood samples were obtained at baseline and at each visit for assessment of thyroid function (serum FT4, FT3, and TSH), serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), blood glucose, and blood potassium. Blood pressure (BP, mmHg) and body weight (kg) were also monitored. Serum autoantibodies against thyroid peroxidase and TSH receptor (TsAb) were measured at baseline and at the end of intervention.

Volunteers with normal eyes (VNEs) were recruited if they met the following criteria: euthyroidism and no symptoms of eye diseases, negative for TsAb, no history of hyperthyroidism and family thyroid disease, no history of eye disease and other autoimmune diseases. The ethical committee of Tongji Hospital approved the study. All participants provided written informed consent.

Study procedures

Eligible patients meeting the inclusion criteria were randomly assigned to the monthly regimen (MR) [18] group or weekly regimen (WR) [1, 5] group according to a random number table. The MR was as follows: 0.5 g iv MP daily for three consecutive days (pulse therapy on days 1, 2, and 3 in weeks 1, 5, 9, and 13) over 4 cycles at 4 weekly intervals, for a total dose of 6.0 g over 3 months. The WR involved administration of 0.5 g MP weekly for 6 weeks, followed by 0.25 g weekly for 6 weeks, for a cumulative dose of 4.5 g MP over 12 consecutive weeks. Patients in the MR group were evaluated at weeks 1, 5, 9, and 13 within one week after administration of each cumulative dose of 1.5 g iv MP. Patients in the WR group were evaluated at weeks 3, 6, 9, and 12 within one week after administration of each cumulative dose of 1.5 g MP (former 6 weeks) or 0.75 g MP (later 6 weeks).

Efficacy evaluation

Evaluation of treatment outcomes included three major criteria and six minor criteria of response adapted from previous studies [18, 19]. The three

major criteria: (1) improvement in diplopia by at least 1 grade according to the adapted NOSPECS classification 4 [18], (2) improvement of eye movement in any direction of 8° or more, (3) improvement in CAS by at least 2 points. The six minor criteria: (1) reduction of proptosis by at least 2 mm, (2) improvement in the grade of soft tissue involvement, according to the modified NOSPECS class 2 [16], (3) improvement in visual acuity by at least 1.0 Snellen line, (4) decrease in CAS of 1 point, (5) decrease in palpebral aperture by at least 2 mm, (6) improvement in duction or diplopia failing to meet 1 grade. We found that contradictory outcomes could simultaneously be observed on the same eye by the end of intervention. For example, a decrease in CAS by 2 points could coexist with an increase in diplopia by 1 degree on the same eye. Therefore, all of the outcomes were considered together to determine an overall improvement or deterioration on the same eye. An improvement was defined as showing at least one major criterion or two minor criteria in any eye. Deterioration was defined as exhibiting one of the following: worsening of diplopia by 1 grade or more, decrease in eyeball duction of 8° or more, increase in palpebral aperture by at least 2 mm, increase in the modified NOSPECS class 2 by at least 2 grades, or increase in exophthalmos by at least 2 mm. No change was defined as no difference throughout the study period or changes smaller than any of the above mentioned criteria.

Orbital Magnetic Resonance imaging

Orbital Magnetic Resonance (OMR) imaging was performed in all participants with a 3.0 T MR system (Signa HDxt, GE Healthcare, Pennsylvania, USA). A head coil (8ch HD Brain Coil, GE Healthcare, Pennsylvania, USA) was used to image the orbit. The T2RTs of EOMs were measured using a multislice multispinecho pulse sequence with a TR of 1,500 ms and 7 TE values (22, 33, 44, 55, 66, 77, and 88 ms), 180 × 180 mm field of view, 3.0 mm slice thickness, 256 × 256 matrix, and 1 NEX. Color coded T2 calculated maps were generated using T2 mapping software (ADW4.4 workstation, GE, USA) with a monoexponential curve fit. T2RTs (ms) and areas (mm^2) of four EOMs (superior, inferior, medial, and lateral rectus) of the two eyes were measured. The largest value of the T2RT or area on a coronal section in each EOM was recorded as the final T2RT or area value. The mean of T2RTs (meanT2RTs) and areas (meanAreas) were calculated as the averages of the sums of T2RTs or areas of four EOMs of the two eyes. Maximum T2RTs (maxT2RTs) and areas (maxAreas) were recorded as the largest value of T2RTs or areas among the eight EOMs. The same radiologist recorded T2RT and area values at baseline and at each visit. The ophthalmologist and radiologist were both blinded to the grouping and did not consult each other with regards to the results and observations throughout the study period. The ophthalmologist provided ophthalmic data regularly to the same endocrinologist to ensure the safety of patients during the treatment.

Statistical analysis

All data were analysed using the statistical software package SPSS version 17.0. Continuous variables are expressed as means ± standard deviation. The *t* test, Wilcoxon Mann-Whitney test and Kruskal-Wallis test was used to compare continuous variables between groups. Comparisons of sample means between groups were analysed with one-way ANOVA. Categorical variables are expressed as the frequency or percentage. Fisher's exact test was used to examine differences in categorical variables between groups using 2 × 2 contingency tables. A $p<0.05$ was taken as the minimum level of statistical significance.

Results

Subjects

Up until July 2012, 40 eligible patients were randomly assigned to either the MR or WR group. Their baseline characteristics are shown in Table 1. Eight patients withdrew from the trial prematurely because they could not endure fatigue, abdominal pain, or insomnia, or became inconvenient to complete the regular visits, or came to believe that the treatment was invalid (only 1 patient). The remaining 32 patients appeared for at least one of all visits after receiving the first treatment, and the results of their last assessment were carried forward and evaluated as the last visit. Among them, 23 patients had active GO and 9 had inactive GO. There were 17 and 15 patients in the MR and WR groups, respectively, and the ratio of active/inactive patients was not significantly different (11/6 and 12/3, respectively, $p>0.05$, Table 2).

Overall, there were 24 VNEs, consisting of 11 men and 13 women, aged 19–62 years old (41.0 ± 14.3 years), who participated in the OMR survey.

Table 1 Baseline characteristics of the patients

	MR group	WR group	*p* value
n	22	18	
Age (yr)	42.3 ± 9.0	41.2 ± 12.4	0.882
Gender (male) (n, %)	6 (27.3)	8 (44.4)	0.327
Weight (kg)	63.3 ± 11.6	61.0 ± 10.6	0.575
Waist circumference (mean, cm)	84.5 ± 9.8	84.3 ± 9.7	0.878
Current smoker (n, %)	3 (13.6)	4 (22.2)	0.680
Duration of GD (month) (median) (Q_{25}, Q_{75})	6 (1.3, 24)	7 (2.3, 29)	0.607
Duration of GO (month) (median) (Q_{25}, Q_{75})	7 (3.3, 12)	6 (3.0, 12)	0.464
Current thyroid disease			
GO (n, %)	4 (18.2)	5 (27.8)	0.705
GD, GO (n, %)	17 (77.3)	13 (72.2)	0.731
Hypothyroidism, GO (n, %)	1 (4.6)	0 (0)	1.000
Previous or current treatments			
Radioiodine (n, %)	6 (27.3)	6 (33. 3)	0.738
Methimazole or propylthiouracil (n, %)	10 (45.5)	6 (33.3)	0.526
Thyroidectomy (n, %)	2 (9.1)	1 (5.6)	1.000
None (n, %)	4 (18.2)	5 (27.8)	0.705
Biochemical characteristics			
Blood glucose (mmol/L)	5.3 ± 0.6	5.3 ± 0.8	0.966
ALT (mmol/L)	23.2 ± 10.4	17.2 ± 9.7	0.077
AST (mmol/L)	19.6 ± 6.0	17.4 ± 3.6	0.261
Serum potassium (mmol/L)	4.2 ± 0.3	4.0 ± 0.3	0.984
TsAb (IU/L)	15.7 ± 13.8	18.4 ± 9.5	0.161
Ophthalmic symptoms and signs			
CAS (median) (Q_{25}, Q_{75})	4 (2, 4)	4 (3, 4)	0.758
Palpebral aperture (mean, mm)	10.0 ± 2.3	10.3 ± 1.9	0.989
Soft tissure involvements			0.732
Mild	6 (27.3)	7 (38.9)	1.000
Moderate	11 (50.0)	6 (33.3)	1.000
Severe	5 (22.7)	5 (27.8)	1.000
Diplopia			0.747
Absent (n, %)	9 (40.9)	6 (33.3)	1.000
Intermittent (n, %)	3 (13.6)	2 (11.1)	1.000
Inconstant (n, %)	5 (22.7)	6 (33.3)	1.000
Constant (n, %)	5 (22.7)	4 (22.2)	1.000
Exophthalmos (mean, mm)	17.9 ± 3.4	17.2 ± 2.1	0.258
IOP (mean, mmHg)	17.7 ± 3.2	18.3 ± 4.2	0.627
Visual acuity (mean)	0.73 ± 0.3	0.75 ± 0.3	0.777

Continuous variables were evaluated with the Student's *t*-test; categorical valuables or non-normally distributed variables were evaluated with the Mann-Whitney test; proportions were evaluated with the χ^2 test or Fisher's exact test. Median (Q_{25}, Q_{75}): median (25^{th} to 75^{th} percentiles).

Table 2 Ophthalmologic examination at baseline and the end of treatment

Variable	MR group		WR group	
	Baseline (n=22)	13 week (n=17)	Baseline (n=18)	12 week (n=15)
Diplopia	1.4 ± 1.2	1.3 ± 1.1	1.5 ± 1.4	1.3 ± 1.3
Improved (n, %)		8 (47.1)		4 (26.7)
Unchanged (n, %)		6 (35.3)		8 (53.3)
Deteriorated (n, %)		3 (17.7)		3 (20.0)
Ocular motility	1.7 ± 1.4	1.0 ± 1.0	1.92 ± 0.8	1.67 ± 1.1
Improved (n, %)		8 (47.1)		6 (40.0)
Unchanged (n, %)		5 (29.4)		5 (33.3)
Deteriorated (n, %)		4 (23.5)		4 (26.7)
CAS	3.2 ± 1.3	2.5 ± 1.1	3.6 ± 1.1	3.0 ± 1.5
Median (Q_{25}, Q_{75})	4 (2, 4)	2 (2, 3) [a]	4 (3, 4)	2 (2, 4) [a]
Improved (n, %)		8 (47.1)		6 (40.0)
Unchanged (n, %)		9 (52.94)		6 (40.0)
Deteriorated (n, %)		0		3 (20.0)
Visual acuity	0.7 ± 0.3	0.9 ± 0.3 [a]	0.8 ± 0.3	0.9 ± 0.3
Improved (n, %)		11 (64.7)		7 (46.7)
Unchanged (n, %)		5 (29.4)		5 (33.3)
Deteriorated (n, %)		1 (5.9)		3 (20.0)
IOP	17.7 ± 3.2	16.8 ± 3.4	18.3 ± 4.2	17.8 ± 4.1
Improved (n, %)		8 (47.1)		8 (53.3)
Unchanged (n, %)		1 (5.9)		1 (6.7)
Deteriorated (n, %)		8 (47.1)		6 (40.0)
Exophthalmos	17.9 ± 3.4	16.8 ± 3.4	17.2 ± 2.1	16.0 ± 3.1
Improved (n, %)		6 (35.3)		5 (33.3)
Unchanged (n, %)		6 (35.3)		5 (33.3)
Deteriorated (n, %)		5 (29.4)		5 (33.3)
Soft tissure involvement				
Mild	6 (27.3)	10 (58.8) [a]	7 (38.9)	8 (53.3)
Moderate	11 (50.0)	7 (41.2)	6 (33.3)	6 (40.0)
Severe	5 (22.7)	0 (0)	5 (27.8)	1 (6.7)
Palpebral aperture (mean, mm)	10.0 ± 2.3	9.8 ± 1.7	10.3 ± 1.9	9.9 ± 1.3
Improved (n, %)		4 (23.5)		4 (26.7)
Unchanged (n, %)		12 (70.6)		10 (66.7)
Deteriorated (n, %)		1 (5.9)		1 (6.7)
Overall response				
Improved (n, %)		13 (76.5)		10 (66.7)
Unchanged (n, %)		1 (5.9)		1 (6.7)
Deteriorated (n, %)		3 (17.6)		4 (26.7)
Total rate of response		71.9		

Wilcoxon Mann-Whitney test for categorical variables or non-normally distributed variables; χ^2 test or Fisher's exact test for proportions. Unless otherwise noted, there were no significant differences between the MR group and WR group. Median (Q_{25}, Q_{75}): median (25^{th} to 75^{th} percentiles). *p* value: *vs*. Baseline, $p<0.05$.

Treatment efficacy

Significant improvements of soft tissue involvement and visual acuity were found in the MR group. CAS decreased significantly in both the MR and WR groups. Among 8 observation items in our study, there were 6 items in the MR group, in which the rates of improved symptoms were greater than the WR group. In contrast, there were 5 and 7 items respectively in the WR group, in which the rates of unchanged or deteriorated symptoms were higher than MR group (Table 2). During treatment, the reappearance of ophthalmic symptoms occurred in 29.4% (5/17) patients in the MR group and 53.3% (8/15) patients in the WR group ($p>0.05$), which were characterized by either the reappearance or aggravation of original symptoms. The severity of symptoms was generally more serious in the MR group and the worse subjective feelings were particularly evident about 10 days after each pulse therapy. In the WR group, the worse subjective feelings occurred shortly after the doses of iv MP being halved at the seventh week or about 5 days after each pulse therapy. One patient ultimately underwent orbital decompression surgery and another developed retrobulbar pain at the seventh week, which did not previously exist.

OMR parameters

Difference of meanT2RTs and meanAreas between each two EOMs was statistically significant in the VNEs (Table 3). The meanT2RTs and meanAreas of EOMs in patients with GO were significantly higher than that in the VNEs [15]. The maxT2RTs levels were significantly reduced in both groups at the end of treatment, however, the process of which were slow and steady in the WR group and fluctuant in the MR group (Table 4). The maxAreas levels were significantly reduced at visit 3 and visit 4 in the WR group, but such a significant change was not seen in the MR group. All four EOMs were found to be involved in all 40 patients with active or inactive GO. Both prolonged T2RT and enlarged area often occurred on the identical EOM. The rates of four EOMs involvement on both sides of orbita were 65.2% (15/23) in active GO and 33.3% (3/9) in inactive GO patients ($p<0.05$). Inversely, the rates of three or less EOMs involvement were 34.8% (8/23) in active GO and 66.7% (6/9) in inactive GO patients ($p<0.05$).

Table 3 The meanT2RTs and meanAreas of 4 EOMs in VNEs

EOMs	VNEs (n=24)	Mean 95% CI Lower-upper Limit	95% CI Reference ranges
MeanT2RTs (ms)			
Superior	75.4 ± 5.0	73.2 ~ 77.5	65.5 ~ 85.2
Inferior	82.8 ± 4.3 [a]	81.0 ~ 84.6	74.4 ~ 91.2
Medial	68.7 ± 5.6 [ab]	66.3 ~ 71.1	57.6 ~ 79.7
Lateral	77.8 ± 4.0 [ab]	76.1 ~ 79.5	70.0 ~ 85.6
MeanAreas (mm^2)			
Superior	30.5 ± 5.0	28.4 ~ 35.6	20.8 ~ 40.2
Inferior	24.7 ± 6.4 [a]	22.0 ~ 27.4	12.2 ~ 37.2
Medial	33.6 ± 5.1 [b]	31.4 ~ 35.7	23.7 ~ 43.5
Lateral	32.2 ± 8.0 [b]	28.8 ~ 35.6	16.5 ~ 48.0

Data are means ± SD. VNEs, volunteers with normal eyes; CI, confidence interval. Kruskal-Wallis test for comparisons among VNEs, $p<0.017$ was considered as a significant difference. One-way ANOVA for comparisons between meanT2RTs and between meanAreas among 4 EOMs on both sides of eyes in VNEs ($F=59.986$, $p<0.001$ and $F=17.046$, $p<0.001$, respectively). [a] $p<0.001$: *vs*. superior rectus; [b] $p<0.001$: *vs*. inferior rectus.

Table 4 The maxT2RTs and maxAreas at baseline and after iv MP

	Monthly Regimen		Weekly Regimen	
	maxT2RTs	maxAreas	maxT2RTs	maxAreas
Baseline	108.5 ± 14.9	85.5 ± 31.4	106.9 ± 13.7	92.8 ± 25.2
Visit 1	98.0 ± 12.2 [a]	72.8 ± 27.1	101.3 ± 14.2	78.5 ± 22.8
Visit 2	99.3 ± 8.0	77.6 ± 29.2	97.5 ± 12.8	74.0 ± 16.4
Visit 3	96.3 ± 5.3	69.4 ± 21.1	96.1 ± 7.7	69.6 ± 21.0 [a]
Visit 4	94.6 ± 7.5 [a]	67.6 ± 25.3	95.2 ± 10.5 [a]	62.5 ± 14.2 [a]

Data are means ± SD. Mean values were compared with the Kruskal Wallis test. [a] $p < 0.05$: *vs*. baseline.

Adverse events

There were no significant differences between the two groups with respect to adverse events. One patient quit the study early owing to severe gastric aches and headache with hidrosis after the first treatment. Fortunately, she recovered one day after cessation of treatment. No severe adverse events, *e.g.*, relevant hepatotoxicity defined as a 3 fold or more increase in serum liver enzymes, occurred. Weight gain was the most common adverse event (Table 5).

Table 5 Adverse-events (n, %)

Adverse-events	MR group (n=22)	WR group (n=18)	*p* value
Insomnia	4 (18.2)	2 (11.1)	0.684
Fatigue	3 (13.6)	2 (11.1)	
Teeth weakness	2 (9.1)	0	0.498
Facial edema	1 (4.5)	1 (5.6)	1.000
Abdominal pain	2 (9.1)	2 (11.1)	1.000
Hidrosis	1 (4.5)	3 (15.0)	0.335
Weight gain	8 (36.4)	3 (15.0)	0.490
Glucose levels elevated	2 (9.1)	1 (5.6)	1.000
Blood pressure elevated	2 (9.1)	2 (11.1)	1.000
Hypokalemia	1 (4.5)	1 (5.6)	1.000
Liver enzyme elevated	0	0	
Adverse-events rate	14 (63.6)	9 (50.0)	0.793

Weight gain is about 4 kg during treatment. Proportions were evaluated with the χ^2 test or Fisher's exact test.

Discussion

We found that both MR and WR were effective, but the total rate of response (71.9%) was lower than that reported in previous RCTs [4, 8]. There were no significant differences in the rates of overall response (improved, unchanged, and deteriorated) and adverse events between the MR and WR groups. However, compared to the WR group, more improvements of ophthalmic symptoms, lower recurrence rates, and lower levels of maxT2RTs were found in the MR group.

Monthly regimen is not a commonly used strategy in the treatment of GO [5] and characterized by a larger dose at each cycle (1.5 g iv MP) and a long interval (4 weeks) between any two cycles [18]. Clinical trials [6, 7] have showed that with the same administration method, a better response was proportional to receiving larger doses of iv MP in the range of safe dosage. Our study shown similar results again (Table 2). However, there was a different result in a most recent RCT [8], in which a weekly protocol of 2.0 g MP and a daily protocol of 3.0 g MP led to almost the same response rate at the fourth week.

The times to reach the same cumulative doses (*e.g.*, 1.5 g, 3.0 g and 4.5 g) of iv MP and to achieve significant improvement in ophthalmic symptoms were earlier in patients receiving higher dose MP. The strongest feelings caused by improvement were most often reported after the first iv MP or about 1 week after each pulse therapy. The earliest reduction of maxT2RTs level occurred significantly after the first iv MP in the MR group (Table 4). Recent studies also [14, 20] suggest that early use of a higher dose of iv MP might be better way to improve ophthalmic symptoms and control ocular inflammation rapidly, which is worthy of further attention.

The process of improvement in ophthalmic symptoms were generally unstable. The duration of feeling well after infusion is generally limited, and that the severity of illness may fluctuate if interval between any two cycles is greater than 1 week. In order to achieve a better therapeutic effect and a quality of life, the early use (*e.g.*, in the first 6–8 weeks) of higher doses (*e.g.*, 0.5 g twice a week) of iv MP for a longer period of treatment (*e.g.* 4–6 weeks) might be a better strategy and the interval between any two cycles should not be longer than 2 weeks.

Comparisons of commonly recommended treatments among different RCTs showed that only improvements in parameters, such as diplopia and/or CAS [7, 8] were evidently better than those observed in the WR group (Table 2). As a subjective symptom, evaluation with CAS could also be affected by the professional's experience. Furthermore, the diagnostic sensitivity of CAS may be weakened when applied in Asian patients, because their facial shapes, unlike Europeans, is not as clear. Indeed, patients with $1 \leq \text{CAS} < 3$ were included in the present study, which could also affect the outcomes. Therefore, there is a need to establish a widely accepted and standardized evaluation system for comparing the advantages and disadvantages of different treatment schedules [1].

We preliminary established the normal ranges of meanT2RTs and meanAreas of four EOMs in 24 Asian VENs (Table 3), which will play a major role in identifying the prolonged T2RT and enlarged EOM. The T2RTs values are higher than those based on the data from 12 control females with either thyroid cancer or a benign thyroid nodule [13], but the normal range of

maxT2RT (80 ms or lower) employed in a subsequent RCT [14] was close to the highest meanT2RTs (82.8 ± 4.3 ms) in our study (Table 3). The prolonged TR2T and enlarged area existed often on the same EOM in almost all patients. The maxT2RTs and maxAreas levels showed a strong positive correlation each other [15] and were associated with the remissions and fluctuations of ophthalmic symptoms during treatment. Therefore, the coexistence of them is most likely a clear sign of acute inflamed EOMs. No correlations were found between CAS and meanT2RTs or meanAreas in patients with GO, which is different from previous results [14]. The reasons may be that T2RT and area value of every EOM is originally not the same, and EOM inflammation is an random events. Although the maxAraes levels were significantly reduced in the WR group at visit 3 and visit 4 (Table 4), the enlarged EOM may not be suitable for identifying active GO alone, because it could represent the presence of chronic fibrotic EOMs.

Ophthalmic symptoms could be also ameliorated by iv MP in some patients with $1 \leq CAS < 3$ and a prolonged TR2T of EOM, which was similar to previous studies [9, 14]. Our study suggest again that a prolonged T2RT may be a more sensitive and quantitative marker than CAS alone for the detection of patients who will be responsive to iv MP. Therefore, we think that in patients with mild GO with an unsatisfactory quality of life and a prolonged T2RT of EOMs, immunosuppressive treatments might be offered, if careful consideration of risks and benefits favours intervention.

The main limitations of the present study are relatively small sample size, and the fact that 20% (8/40) of patients withdrew from the study early. Furthermore, we have only explored a short period of observation (12–14 weeks) and did not perform follow up assessments. Finally, the responses in some patients were based on prolonged T2RTs on at least one of EOMs instead of $CAS \geq 3$, therefore, direct comparison of treatment outcomes between different RCTs is not possible.

In conclusion, both MR and WR are effective and safe in treatment of GO. The process of improvement in ophthalmic symptoms and maxT2RTs were fluctuant in the MR group and were slow and steady in the WR group. However, there was a higher recurrence rate of the disease in the WR group and the overall efficacy was slightly worse than MR. The maxT2RTs values appear to be useful for detecting and monitoring the changes of acute inflamed EOMs, and could predict the response to iv MP. Importantly, this marker could help to identify euthyroid patients with active GO or patients with GO and negative CAS.

Acknowledgements

This work was partly supported by grants from National Natural Science Foundation of China (NNSFC) (81301192).

Disclosure Statement

The authors have nothing to disclose.

References

1. Bartalena L, Baldeschi L, Dickinson A, Eckstein A, Kendall-Taylor P, *et al*. (2008) European Group on Graves' Orbitopathy (EUGOGO). Consensus statement of the European Groupon Graves' orbitopathy (EUGOGO) on management of GO. *Eur J Endocrinol* 158: 273-285.
2. Bartalena L (2011) The dilemma of how to manage Graves' hyperthyroidism in patients with associated orbitopathy. *J Clin Endocrinol Metab* 96: 592-599.
3. Marcocci C, Bartalena L, Tanda ML, Manetti L, Dell'Unto E, *et al*. (2001) Comparison of the effectiveness and tolerability of intravenous or oral glucocorticoids associated with orbital radiotherapy in the management of severe Graves' ophthalmopathy: results of a prospective, single-blind, randomized study. *J Clin Endocrinol Metab* 86: 3562-3567.
4. Kahaly GJ, Pitz S, Hommel G, Dittmar M (2005) Randomized, single-blind trial of intravenous versus oral steroid monotherapy in Graves' orbitopathy. *J Clin Endocrinol Metab* 90: 5234-5240.
5. Zang S, Ponto KA, Kahaly GJ (2011) Intravenous glucocorticoids for Graves' orbitopathy: efficacy and morbidity. *J Clin Endocrinol Metab* 96: 320-332.
6. Sánchez-Ortiga R, Moreno-Pérez O, González Sánchez V, Arias Mendoza N, Mauri Dot M, *et al*. (2009) Treatment of Graves' ophthalmopathy with high-dose intravenous methylprednisolone: a comparison of two dosing regimens. *Endocrinol Nutr* 56: 118-122.

7. Bartalena L, Krassas GE, Wiersinga W, Marcocci C, Salvi M, *et al*. (2012) European Group on Graves' Orbitopathy. Efficacy and safety of three different cumulative doses of intravenous methylprednisolone for moderate to severe and active Graves' orbitopathy. *J Clin Endocrinol Metab* 97: 4454-4463.
8. Zhu W, Ye L, Shen L, Jiao Q, Huang F, *et al*. (2014) A prospective, randomized trial of intravenous glucocorticoids therapy with different protocols for patients with graves' ophthalmopathy. *J Clin Endocrinol Metab* 99: 1999-2007.
9. Mourits MP, Koornneef L, Wiersinga WM, Prummel MF, Berghout A, *et al*. (1989) Clinical criteria for the assessment of disease activity in Graves' ophthalmopathy: a novel approach. *Br J Ophthalmol* 73: 639-644.
10. Mourits MP, Prummel MF, Wiersinga WM, Koornneef L (1997) Clinical activity score as a guide in the management of patients with Graves' ophthalmopathy. *Clin Endocrinol (Oxf)* 47: 9-14.
11. Hosten N, Sander B, Cordes M, Schubert CJ, Schörner W, *et al*. (1989) Graves ophthalmopathy: MR imaging of the orbits. *Radiology* 172: 759-762.
12. Hiromatsu Y, Kojima K, Ishisaka N, Tanaka K, Sato M, *et al*. (1992) Role of magnetic resonance imaging in thyroid-associated ophthalmopathy: its predictive value for therapeutic outcome of immunosuppressive therapy. *Thyroid* 2: 299-305.
13. Ohnishi T, Noguchi S, Murakami N, Tajiri J, Harao M, *et al*. (1994) Extraocular muscles in Graves ophthalmopathy: usefulness of T2 relaxation time measurements. *Radiology* 190: 857-862.
14. Tachibana S, Murakami T, Noguchi H, Noguchi Y, Nakashima A, *et al*. (2010) Orbital magnetic resonance imaging combined with clinical activity score can improve the sensitivity of detection of disease activity and prediction of response to immunosuppressive therapy for Graves' ophthalmopathy. *Endocr J* 57: 853-861.
15. He Yebing, Xiang Nan, Shi Shaomin, Mu Ketao, Zhang Jing, *et al*. (2015) The diagnostic value of orbital magnetic resonance imaging T2RT-Mapping in patients with active Graves' ophthalmopathy. *Chin J Endocrinol Metab* 31: 327-331 (In Chinese).
16. Werner SC (1977) Modification of the classification of the eye changes of Graves' disease: recommendations of the Ad Hoc Committee of the American Thyroid Association. *J Clin Endocrinol Metab* 44: 203-204.
17. Werner SC (1977) Modification of the classification of the eye changes of Graves' disease. *Am J Ophthalmol* 83: 725-727.
18. van Geest RJ, Sasim IV, Koppeschaar HP, Kalmann R, Stravers SN, *et al*. (2008) Methylprednisolone pulse therapy for patients with moderately severe Graves' orbitopathy: a prospective, randomized, placebo-controlled study. *Eur J Endocrinol* 158: 229-237.
19. Bartalena L, Marcocci C, Bogazzi F, Manetti L, Tanda ML, *et al*. (1998) Relation between therapy for hyperthyroidism and the course of Graves' ophthalmopathy. *N Engl J Med* 338: 73-78.
20. Hart RH, Kendall-Taylor P, Crombie A, Perros P (2005) Early response to intravenous glucocorticoids for severe thyroid-associated ophthalmopathy predicts treatment outcome. *J Ocul Pharmacol Ther* 21: 328-336.