

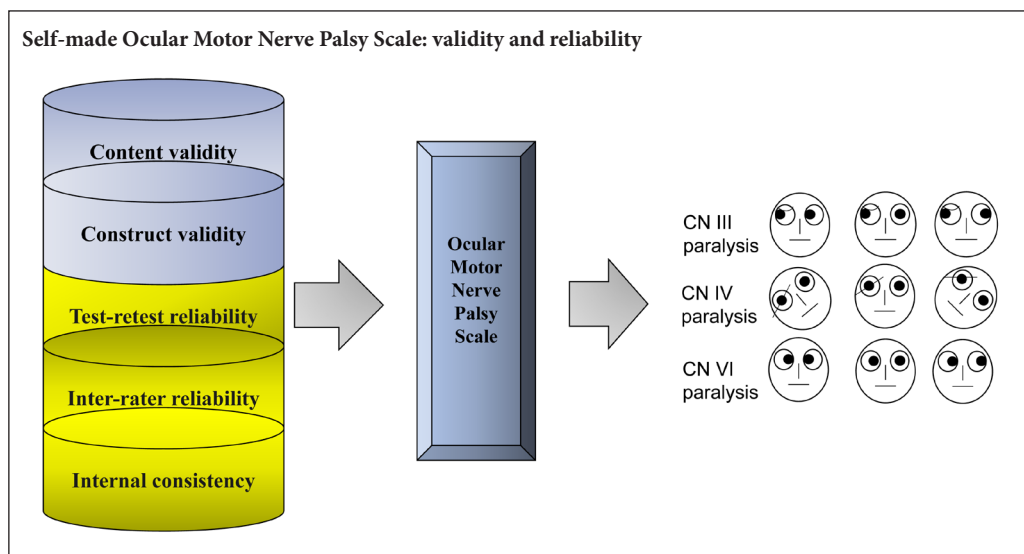
# Validity and reliability of the Ocular Motor Nerve Palsy Scale

Ling-Yun Zhou<sup>\*</sup>, Chang Su, Tie-Juan Liu, Xue-Mei Li

Ocular Motility Disorder Treatment & Rehabilitation Center, Department of Acupuncture, First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang Province, China

**Funding:** This study was supported by the National Natural Science Foundation of China, No. 81674052.

## Graphical Abstract



\*Correspondence to:

Ling-Yun Zhou, PhD, MD,  
no1zhly@163.com.

orcid:

0000-0001-6659-3839  
(Ling-Yun Zhou)

doi: 10.4103/1673-5374.238716

Accepted: 2018-06-25

## Abstract

Objective and accurate assessment of the degree of ocular motor nerve palsy is helpful not only in the evaluation of prognosis, but also for the screening of treatment methods. However, there is currently no comprehensive measure of its severity. In this study, we designed the Ocular Motor Nerve Palsy Scale and investigated its validity and reliability. Six experts were invited to grade and evaluate the scale. The study recruited 106 patients with a definite diagnosis of unilateral isolated ocular motor nerve palsy. Three physicians evaluated the patients using the scale. One of the three physicians evaluated the patients again after 24 hours. The content validity index (CVI) and factor analysis were used to analyze the scale's construct validity. The intraclass correlation coefficient and Cronbach's alpha were used to evaluate the inter-rater and test-retest reliability and the internal consistency. The CVI results (I-CVI = 1.0, S-CVI = 0.9,  $P_c = 0.016$ ,  $K^* = 1$ ) indicated good content validity. Factor analysis extracted two common factors that accounted for 85.2% of the variance. Furthermore, the load value of each component was above 0.8, indicating good construct validity. The Ocular Motor Nerve Palsy Scale was found to be highly reliable, with an inter-rater reliability intraclass correlation coefficient of 0.965 ( $P < 0.01$ ), a test-retest reliability intraclass correlation coefficient of 0.976 ( $P < 0.01$ ), and Cronbach's alpha values of 0.63–0.70. In conclusion, the Ocular Motor Nerve Palsy Scale with good validity and reliability can be used to quantify the severity of ocular motor nerve palsy. This study was registered at Chinese Clinical Trial Registry (registration number: ChiCTR-OOC-17010702).

**Key Words:** nerve regeneration; Ocular Motor Nerve Palsy Scale; oculomotor nerve paralysis; abducens nerve paralysis; trochlear nerve paralysis; validity; reliability; internal consistency; diplopia; eyeball movement disorder; ptosis; neural regeneration

## Introduction

The ocular motor nerves (cranial nerves III, IV, and VI) innervate different extraocular muscles. Damage to these nerves results in complex clinical symptoms and signs, and it is difficult to rate the severity of damage. At present, the severity of damage to these nerves is usually demonstrated by describing a single symptom. For example, the Hirshberg cover-uncover test with prim bars (Rucker et al., 2011) and the Maddox rod and red glass test are used to detect ocular deviation (Newman-Toker and Rizzo, 2010). The Hess

screen test and Lancaster screen test are used to test diplopia (Christoff and Guyton, 2006; Roper-Hall, 2006). Scott and Kraft (1985) developed a grading method to describe eye movement limitations. However, there is no system or scale to synthetically evaluate the severity of ocular motor nerve palsy (OMNP). To address the urgent need for a useful clinical tool, this study designed a scale based on the symptoms and signs of OMNP. The Ocular Motor Nerve Palsy Scale (OMNPS) evaluates severity by assessing diplopia at 1 meter, eye movement, palpebral fissure, pupil size, and light reflex.

The reliability and validity of the scale were investigated to verify its practicability and feasibility.

## Participants and Methods

### Participants

One hundred and six patients with unilateral OMNP were enrolled through advertisement and clinical practice at the First Affiliated Hospital of Harbin Medical University, Department of Ocular Motility Disorder Treatment & Rehabilitation Center, China, from March 2016 to February 2017.

### Inclusion criteria

Patients presenting with all of the following criteria were considered for inclusion in the study: voluntary participation in the evaluation and diagnosis of unilateral isolated CN-III IV VI paralysis, a binocular visual acuity difference of less than 0.1, and able to complete a computerized diplopia test independently.

### Exclusion criteria

Patients with one or more of the following conditions were excluded from the study: color blindness, color weakness, abnormal retinal correspondence, or any other severe condition.

All patients gave signed informed consent, and the study was approved by the Ethics Committee of the First Affiliated Hospital of Harbin Medical University of China (approval number: 201452).

### Scale design

#### Item selection

The five main items, diplopia at 1 meter, eye movement condition, palpebral fissure, pupil size, and light reflex, were assessed according to the clinical symptoms and signs of OMNP (Danchaivijitr, 2004; Brazis, 2009; Prasad and Volpe, 2010; Galtrey et al., 2015). The scale is displayed in **Table 1**.

#### Testing methods

Diplopia at 1 meter (Zhou et al., 2016, 2017): After head fixation, patients completed a computerized diplopia test at 1 meter wearing red and blue glasses. The direction and maximum angle of deviation of diplopia were measured and recorded automatically by the computer.

Eye movement (Corbett, 2003): After head fixation, patients were asked to move the eyeball toward the left and right, and straight up and down. The functional eye movement in four directions was observed and graded according to the scoring criteria.

The palpebral fissure: The frontal muscle above the eyebrow was fixed and did not affect the activity of the upper eyelid. Patients were asked to open their eyelids as wide as possible. The distance between the midpoint of the upper and lower eyelids (accurate to the millimeter) was measured.

Pupil size (Jacobson, 1998): Pupil diameter was measured using a ruler under stationary indoor lights and all testing rooms had the same level of light.

Light reflex (Cornblath, 2014): The physician directly tested the patient's light reflex with a flash light.

### OMNPS scoring criteria

The scale is divided into three sub-scales that reflect the type of cranial nerve damage. PART CN-IV is designed for trochlear nerve paralysis patients, and includes diplopia at 1 meter only. PART CN-VI is designed for abducens nerve paralysis, and includes diplopia at 1 meter and eye movement. PART CN-III is designed for oculomotor nerve paralysis patients and includes all five items. (1) Diplopia at 1 meter: the maximum angle of diplopia is graded on five levels, with each level representing 5°. If the maximum deviation angle is horizontal, it only needs grading; if it is vertical, the angle is multiplied by 3 before grading. (2) Eye movement: This test consists of two parts. Horizontal movement is assessed on five levels with reference to the midline and two quarter lines. Vertical movement is assessed on four levels with reference to the line of the inner and temporal canthus. (3-4) Palpebral fissure and pupil size: affected eyes are compared with healthy eyes and there are four levels for palpebral fissure and four levels for pupil size. (5) Light reflex: The direct light reflex and grading are observed. Higher scores represent more severe symptoms. The details of the scoring criteria are listed in **Table 1**.

### OMNPS grading by six experts

Our scale was sent to six experts by email. The experts were invited to grade the scale in relation to OMNP. The scale is shown in **Table 2**. The experts were asked to choose 1 (not relevant) to 4 (highly relevant) according to the item's relevance to OMNP. Content validity indices were calculated from the experts' scores, including item-level content validity index ( $I-CVI = \text{number of experts who gave a grade of 3 or 4} / \text{number of experts}$ ), scale-level content validity index/universal agreement ( $S-CVI/UA = \text{number of items graded as 3 or 4} / \text{number of items}$ ), and  $S-CVI/Ave = \text{average } I-CVI \text{ scores of five items}$ .

### Clinical examination procedure

#### Physicians

The physicians familiarized themselves with the scale before administering the tests. In a pilot study, the physicians administered the test to five patients (not including the 106 patients recruited for this study) until they could run the test smoothly and accurately. The physicians were blind to the patients' medical histories. The patients received the examination in random order.

#### Testing process

Three physicians tested the 106 consecutive patients. The computerized diplopia test was performed under the physicians' instruction, and data were automatically generated by computer and then graded by the physicians. To prevent bias, we applied the following procedures. The testing procedure was performed in three independent clinical consulting rooms. Communication between physicians was forbidden during the process, and the results were compared by physicians who were not involved with the research. One of the three physicians measured the patients again after 24 hours

Table 1 Details of Ocular Motor Nerve Palsy Scale

Center: <input type="checkbox"/>		Number <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>		Name :		Affected eye:							
PART CN-IV = $\alpha$ <input type="checkbox"/>		PART CN-VI = $\alpha + \beta$ <input type="checkbox"/>											
PART CN-III = $\alpha + \beta + \gamma + \delta + \epsilon$ <input type="checkbox"/>													
$\alpha$ diplopia in 1 mm (max angel)		$\beta$ eye movement			$\gamma$ palpebral fissure		$\delta$ pupil	$\epsilon$ light reflex					
<p>※Max deviation angel multiplied by 3 if it was vertical</p> <p>0=Without diplopia, the max deviation &lt; 5°</p> <p>1=Diplopia, the max deviation between 5° and 10°(including 5°)</p> <p>2=Diplopia, the max deviation between 10° and 15°(including 10°)</p> <p>3=Diplopia, the max deviation between 15° and 20°(including 15°)</p> <p>4=Diplopia, the max deviation ≥ 20°</p>		<p><b>Horizontal: Force on the eyeball and move to the adduction and abduction, score respectively</b></p> <p>0=No difference compared with the healthy eye movement</p> <p>1=Activity is approximate to the healthy eye, reach or exceed 1/4 line which is the same side of the paralyzed direction</p> <p>2=Exceed the midline, cannot reach 1/4 line which is the same side of the paralyzed direction</p> <p>3=reach or exceed 1/4 line which is the opposite side of the paralyzed direction, cannot reach or exceed midline</p> <p>4=Can not move the eye to the paralyzed direction/can not exceed 1/4 line which is the opposite side of the paralyzed direction</p> <p><b>Vertical: Force on the eyeball and move to the upward and downward, score respectively</b></p> <p>0=No difference compared with the healthy eye movement</p> <p>1=Affected eye pupil can exceed line of inner and temporal canthus</p> <p>2=Affected eye pupil on the line of inner and temporal canthus</p> <p>3=Affected eye pupil cannot exceed line of inner and temporal canthus/Cannot move the eye to the paralyzed direction</p>			<p>0=Activity is normal, palpebral fissure no difference compared with the healthy eyes</p> <p>1=Palpebral fissure reach to the healthy eye 1/2 or more</p> <p>2=Palpebral fissure cannot reach the healthy eye 1/2</p> <p>3=Force on the eyebrow could see a small fissure</p> <p>4=Eyelids drooping completely even force on the eyebrow</p>		<p>0=Bilateral pupils diameter difference &lt; 1.0 mm</p> <p>1=Bilateral pupils diameter difference in 1.0-2.0 mm (including 1.0 mm)</p> <p>2=Bilateral pupils diameter difference in 2.0-3.0 mm (including 2.0 mm)</p> <p>3=Bilateral pupils diameter difference &gt; 3.0 mm (including 3.0 mm)</p>	<p>0=sensitive</p> <p>1=slow</p> <p>2=disappear</p>					
date:		max angel		H <input type="checkbox"/> V <input type="checkbox"/>		U D I O TS		score		L R		L R	
score		L R		score		distance (mm)		score		diameter		light reflex	

L: Left; R: right; TS: total score; U: upside; D: downside; I: inner side; O: outside; H: horizontal; V: vertical.

without any intervention. The data were independently cross-checked by the other two researchers.

### Statistical analysis

All normally distributed data are expressed as the mean ± SD. The data were analyzed using Statistical 19.0 software (IBM SPSS Statistics, IBM Co., Somers, NY, USA). The CVI was used to assess content validity. Factor analysis was utilized to evaluate the construct validity of PART CN-III. The intraclass correlation coefficient (ICC) with a 95% confidence interval was calculated to assess the inter-rater reliability and test-retest reliability. After comparison of the total score, the scores for each item were compared except for diplopia at 1 meter. Cronbach's Alpha was used to evaluate internal consistency.

## Results

### General information

The 106 patients with OMNP consisted of 67 males and 39 females ranging in age from 11-79 (54.72 ± 14.42) years. **Table 3** presents patients' scores grouped according to the affected cranial nerve.

### Validity analysis

#### Content validity

As displayed in **Table 4**, the I-CVI was 0.83-1, demonstrating that the five items were strongly correlated with OMNP ( $P_c = 0.016$ ,  $K^* = 0.82-1$ ) (Polit et al., 2007). The S-CVI/UA was 0.8 (> 0.8 indicates good content validity; Davis, 1992). S-CVI/Ave was 0.966 (> 0.9 indicates good content validity; Waltz et al., 2005). The results indicate that the content va-

**Table 2 Content validity grading form**

Do you think weather the scale has correlation with ocular motor nerve palsy, and how much is it?				
Max angle of diplopia	1= not relevant <input type="checkbox"/>	2=somewhat relevant <input type="checkbox"/>	3= quite relevant <input type="checkbox"/>	4= highly relevant <input type="checkbox"/>
Eye movement condition	1= not relevant <input type="checkbox"/>	2= somewhat relevant <input type="checkbox"/>	3= quite relevant <input type="checkbox"/>	4= highly relevant <input type="checkbox"/>
Palpebral fissure	1= not relevant <input type="checkbox"/>	2= somewhat relevant <input type="checkbox"/>	3= quite relevant <input type="checkbox"/>	4= highly relevant <input type="checkbox"/>
Pupil size	1= not relevant <input type="checkbox"/>	2= somewhat relevant <input type="checkbox"/>	3= quite relevant <input type="checkbox"/>	4= highly relevant <input type="checkbox"/>
Light reflex	1= not relevant <input type="checkbox"/>	2= somewhat relevant <input type="checkbox"/>	3= quite relevant <input type="checkbox"/>	4= highly relevant <input type="checkbox"/>

Single choice, please choose the best answer.

**Table 4 Results of content validity**

Item	Experts grade (1-4)						Number of experts who scored 3 or 4	I-CVI	Pc	K*
	A	B	C	D	E	F				
Eyeball movement	4	4	4	4	4	4	6	1.0	0.016	1.0
Diplopia in 1 m	4	4	4	4	4	4	6	1.0	0.016	1.0
Palpebral fissure	4	4	3	4	4	3	6	1.0	0.016	1.0
Pupil size	3	3	3	4	3	3	6	1.0	0.016	1.0
Light reflex	3	3	3	2	3	3	5	0.83	0.094	0.81

1 means not relevant, 2 means somewhat relevant, 3 means quite relevant, and 4 means highly relevant. The content validity index includes I-CVI, S-CVI/UA, and S-CVI/Ave. The item-level content validity index (I-CVI, 0.83-1) is the number of experts who gave a grade of 3 or 4/total number of experts. The scale-level content validity index/universal agreement (S-CVI/UA) = 0.8 (number of items with experts grading 3 or 4, n = 4/number of items, n = 5) S-CVI/Ave = 0.966 (average I-CVI for the 5 items). Pc (probability of a chance occurrence) was computed using the formula for a binomial random variable, with one specific outcome:  $Pc = [N!/A!(N-A)!]^{0.5N}$ , where N represents the number of experts and A represents the number who agreed that the item has high relevance; K\*=kappa, indicating agreement on relevance:  $K^* = (I-CVI-pc)/(1-pc)$ .

**Table 3 Ocular motor nerve palsy score in 106 patients**

Affected cranial nerve	n	Highest score	Lowest score	Mean±SD
III	41	22	3	9.29±4.03
IV	30	4	1	2.83±1.34
VI	35	8	1	5.11±2.27

III, IV and VI represent the affected cranial nerve.

**Table 6 Intraclass correlation coefficient (ICC) results of each item**

	Eyeball movement	Palpebral fissure	Pupil size	Light reflex	Total score
Test-retest	0.962	0.913	0.930	0.968	0.976
Inter-rater	0.928	0.941	0.844	0.941	0.965

Test-retest represents the ICC of test-retest reliability. Inter-rater represents the ICC of Inter-rater reliability. ICC > 0.75 indicates good reliability. All coefficients are significant (P < 0.01).

**Table 5 Results of construct validity (variance, cumulative contribution rate and load values of main components)**

Common factor	Variance	Cumulative contribution rate (%)	Load values on factor 1 (c)	Load values on factor 2 (c)
Diplopia in 1 m	1.013	85.221	-	0.990
Eyeball movement	3.248	64.953	0.863	-
Palpebral fissure	0.332	91.866	0.887	-
Pupil size	0.292	97.700	0.927	-
Light reflex	0.115	100	0.899	-

c: Coefficient of load values on factor 1 or 2.

lidity of the OMNPS is good.

**Construct validity**

First, the Kaiser-Meyer-Olkin test and Bartlett’s spherical test were performed. The Kaiser-Meyer-Olkin value was 0.792 and Bartlett’s spherical test was significant at P < 0.01, confirming that the data were suitable for factor analysis.

Principal component analysis was performed. Two common factors were extracted and the contribution rate was 85.2%. After matrix rotation, eye movement condition, palpebral fissure, pupil size, and light reflex loaded on factor 1 with values of 0.867, 0.887, 0.927, and 0.899, respectively. Maximum angle of diplopia at 1 meter loaded on factor 2 with a value of 0.990. The variance, contribution rate, and cumulative contribution rate are listed in **Table 5**. Two common factors were extracted, demonstrating that all five items reflected symptoms and signs of OMNP. This was in agreement with the theoretical design of the scale. All of the items loaded above 0.8 on their own factor. These results show that the OMNPS PART CN-III has good construct validity.

**Reliability analysis**

The results of the reliability analysis are shown in **Table 6**. In terms of the test-retest reliability, the ICC of each individual item was 0.913–0.968 (P < 0.01); and the ICC of the total score was 0.976 (P < 0.01). The ICC for inter-rater reliability was 0.844–0.941 (P < 0.01), and the total score ICC was 0.965

( $P < 0.01$ ). A 95% confidence interval was used for all ICCs. ICC values less than 0.4 indicate poor reliability, and values greater than 0.75 indicate good reliability (Pan and Ni, 1999).

Internal consistency measured by Cronbach's alpha was 0.636–0.701 ( $P < 0.01$ ), indicating high reliability.

## Discussion

Rating scales are commonly used to evaluate the severity of disease and efficacy of treatment, especially in neurological diseases and dyskinesia (Hobart et al., 2001; Middleton, 2012; Giang et al., 2013). OMNP is a common type of dyskinesia that affects patients' quality of life (Kuo et al., 2010). However, due to the complex manifestations of OMNP, it is difficult to assess its severity. To date, no study or scale has examined this issue and no system or scale is available to evaluate the severity of OMNP in neurology and ophthalmology, or to document the progression and degree of recovery. Therefore, to address this clinical need, this study designed a measurement scale and evaluated its reliability and validity.

### Content validity

In the present study, I-CVI and S-CVI were used to assess the content validity. After correcting the random consistency ( $P_c = 0.016$ ,  $K^* = 1$ ), the results still indicate good content validity (I-CVI = 0.83–1 S-CVI/UA = 0.83 S-CVI/Ave = 0.966) according to previous studies (Davis, 1992; Waltz et al., 2005). Good content validity not only demonstrates the strong association between scale scores and OMNP, but also indicates that the score reflects the severity of OMNP in some respects. All five items measure the symptoms and signs of OMNP and are strongly associated with its pathophysiology (Brazis, 2009; Cornblath, 2014). The testing method generally follows the accepted method of neuro-ophthalmology examination (Danchaivijitr, 2004; Brazis, 2009; Prasad and Volpe, 2010; Galtrey et al., 2015). We made some corrections and modifications to the scale to make it suitable for daily practice. For the eye movement measure, we graded horizontal movement on five levels and vertical movement on four levels according to the horizontal and vertical palpebral fissure size in healthy subjects. The reference values of the computerized diplopia test were horizontal  $2.55^\circ$  and vertical  $0.76^\circ$  (Zhou et al., 2016, 2017). The horizontal value is about three times the vertical value, which is why we multiplied the vertical measure by three before grading.

### Construct validity

Factor analysis is generally used to assess construct validity. In this study, two common factors were extracted and their total contribution was 85.2%, which is significant factor analysis (Nia et al., 2017; Oh and Kim, 2017). After matrix rotation, eye movement, palpebral fissure, pupil size, and light reflex loaded on factor 1 and maximum angle of diplopia at 1 meter loaded on factor 2. We concluded that factor one measures the signs of OMNP and factor 2 reflects the symptoms of OMNP. The extraction of these two factors

demonstrates that all five items reflect the symptoms and signs of OMNP. All items had loadings above 0.8 on their own factor, indicating agreement with the intended theoretical design of the scoring and confirming that the scale has good construct validity.

### Reliability

The ICCs of inter-rater reliability and test-retest reliability were all above 0.75 ( $P < 0.01$ ), indicating very good reliability (Pan and Ni, 1999). The inter-rater and test-retest reliability results are similar to those of other scale development studies (Ardali et al., 2017; Cavalini et al., 2017; De Roo et al., 2017). Cronbach's alpha values were in the acceptable range, from 0.636–0.701 ( $P < 0.01$ ). Overall, the OMNPS shows good reliability. We excluded patients' scores for the angle of diplopia from the reliability and test-retest analysis; it would be meaningless for the patients to repeat the computerized diplopia test three times a day, because the data were automatically generated by computer and the error was very small.

The OMNPS includes three sub-scales for different types of OMNP (Newman and Bruce, 2007; Cornblath, 2014). For instance, abducens nerve paralysis only involves eye movement disorder and diplopia, and the other three items are irrelevant for these patients. Hence, we developed three sub-scales for different situations. PART CN-IV is scored from 0–4, PART CN-VI 0–8, and PART CN-III 0–27; higher scores indicate greater severity.

### Bias

To minimize bias, we used various control processes (Kottner et al., 2011). The test order was randomized, and the tests were conducted in the consulting room and resembled the daily situation as far as possible. Communication between physicians was forbidden during the process (Stochkendahl et al., 2006). To avoid the Hawthorne effect, physicians who were not involved in the research compared the results (Wickstrom and Bendix, 2000). The single blind method was used for the physicians (Kottner et al., 2009). The data were cross-checked and consolidated by two other researchers.

### Limitations

We did not report criterion validity, because there was no suitable standard method for comparison. The scale's application is limited to patients with unilateral mononeuropathy, because three items compare the affected eye with the healthy eye. The scale requires further modification to enable a wide range of application in the clinic. Rotation diplopia could not be quantified because the computerized diplopia test was based on the Hess principle. The Hess test does not assess rotation diplopia, but we believe that computerized diplopia testing will undergo further development. The sensitivity and responsiveness of the scale need to be investigated in the future.

### Conclusion

The OMNPS shows high content validity, construct validity,

inter-rater reliability, test-retest reliability, and internal consistency. It is therefore suitable for use as a clinical scale to evaluate the severity of OMNP.

**Acknowledgments:** We are very grateful to De-Sheng Wang, Cui Hao, Li-Qiong Zhang, Zhong-Ren Sun, Dong-Yan Wang and Zhi-Qiang Su for evaluating and grading this scale, and giving valuable advice on this scale.

**Author contributions:** LYZ designed and supervised the study. CS performed the study. XML and TJL analyzed data. TJL and CS wrote the paper. All authors approved the final version of the paper.

**Conflicts of interest:** The authors have no conflicts of interest to declare.

**Financial support:** This study was supported by the National Natural Science Foundation of China, No. 81674052. The funding body played no role in the study design, in the collection, analysis and interpretation of data, in the writing of the paper, and in the decision to submit the paper for publication.

**Institutional review board statement:** The study followed the request of the Declaration of Helsinki and was approved by Ethics Committee of The First Affiliated Hospital of Harbin Medical University of China (approval number: 201452).

**Declaration of patient consent:** The authors certify that they have obtained all appropriate patient consent forms. In the form the patients have given their consent for their images and other clinical information to be reported in the journal. The patients understand that their names and initials will not be published and due efforts will be made to conceal their identity, but anonymity cannot be guaranteed.

**Reporting statement:** This study follows the Strengthening the reporting of observational studies in epidemiology (STROBE) statement.

**Biostatistics statement:** The statistical methods of this study were reviewed by the biostatistician of The First Affiliated Hospital of Harbin Medical University, China.

**Copyright license agreement:** The Copyright License Agreement has been signed by all authors before publication.

**Data sharing statement:** Anonymized trial data will be available indefinitely at [www.figshare.com](http://www.figshare.com).

**Plagiarism check:** Checked twice by iThenticate.

**Peer review:** Externally peer reviewed.

**Open access statement:** This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non-Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

**Open peer reviewer:** Tom Johnson, Johns Hopkins University, Baltimore, MD, USA.

**Additional file:** Open peer review report 1.

## References

Ardali G, Brody LT, States RA, Godwin EM (2017) Reliability and validity of the floor transfer test as a measure of readiness for independent living among older adults. *J Geriatr Phys Ther* doi: 10.1519/JPT.000000000000142.

Brazis PW (2009) Isolated palsies of cranial nerves. *Semin Neurol* 29:14-28.

Cavalini MA, Berduszek RJ, van der Sluis CK (2017) Construct validity and test-retest reliability of the revised upper extremity work demands (UEWD-R) scale. *J Occup Environ Med* 74:763-768.

Christoff A, Guyton DL (2006) The lancaster red-green test. *Am Orthopt J* 56:157-165.

Corbett JJ (2003) The bedside and office neuro-ophthalmology examination. *Semin Neurol* 23:63-76.

Cornblath WT (2014) Diplopia due to ocular motor cranial neuropathies. *Continuum (Minneapolis)* 20:966-980.

Danchaivijitr C (2004) Diplopia and eye movement disorders. *J Neurol Neurosurg Psychiatry* 75:iv24-iv31.

Davis LL (1992) Instrument review: getting the most from a panel of experts. *App Nursing Res* 5:194-197.

De Roo ML, Tanghe MF, Van Den Noortgate NJ, Piers RD (2017) Development and validation of the symptom assessment to improve symptom control for institutionalized elderly scale. *J Am Med Dir Assoc* 19:148-153.

Galtrey CM, Schon F, Nitkunan A (2015) Microvascular non-arteritic ocular motor nerve palsies-what we know and how should we treat? *Neuroophthalmology* 39:1-11.

Giang P, Williams A, Argyros L (2013) Automated extraction of the Barthel Index from clinical texts. *AMIA Annu Symp Proc* 2013:486-495.

Hobart JC, Lamping DL, Freeman JA, Langdon DW, McLellan DL, Greenwood RJ, Thompson AJ (2001) Evidence-based measurement: which disability scale for neurologic rehabilitation? *Neurology* 57:639-644.

Jacobson DM (1998) Pupil involvement in patients with diabetes-associated oculomotor nerve palsy. *Arch Ophthalmol* 116:723-727.

Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, Roberts C, Shoukri M, Streiner DL (2011) Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Int J Nurs Stud* 48:661-671.

Kottner J, Dassen T, Tannen A (2009) Inter- and intrarater reliability of the Waterlow pressure sore risk scale: A systematic review. *Int J Nurs Stud* 46:369-379.

Kuo LT, Huang APH, Yang CC, Tsai SY, Tu YK, Huang SJ (2010) Clinical outcome of mild head injury with isolated oculomotor nerve palsy. *J Neurotrauma* 27:1959-1964.

Middleton PM (2012) Practical use of the Glasgow Coma Scale; a comprehensive narrative review of GCS methodology. *Australas Emerg Nurs J* 15:170-183.

Newman NJ, Bruce BB (2007) Third nerve palsies. *Semin Neurol* 27:257-268.

Newman-Toker DE, Rizzo JF (2010) Subjectively quantified Maddox rod testing improves diagnostic yield over alternate cover testing alone in patients with diplopia. *J Clin Neurosci* 17:727-730.

Nia HS, Sharif SP, Lehto RH, Allen KA, Goudarzian AH, Yaghoobzadeh A, Soleimani MA (2017) Psychometric properties of the persian version of death depression scale-revised in Iranian patients with acute myocardial infarction. *Iran J Psychiatry* 12:172-181.

Oh J, Kim JA (2017) Factor analysis of the Zarit Burden Interview in family caregivers of patients with amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener* 19:50-56.

Pan X, Ni Z (1999) Application of intraclass correlation coefficient to reliability assessment. *Hua Xi Yi Ke Da Xue Xue Bao* 30:62-67.

Polit DF, Beck CT, Owen S V. (2007) Focus on research methods: Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health* 30:459-467.

Prasad S, Volpe NJ (2010) Paralytic strabismus: third, fourth, and sixth nerve palsy. *Neurol Clin* 28:803-833.

Roper-Hall G (2006) The hess screen test. *Am Orthopt J* 56:166-174.

Rucker JC, Kennard C, Leigh RJ (2011) The neuro-ophthalmological examination, 1st ed. Elsevier.

Scott AB, Kraft SP (1985) Botulinum toxin injection in the management of lateral rectus paresis. *Ophthalmology* 92:676-683.

Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, Adams A, Bronfort G (2006) Manual examination of the spine: a systematic critical literature review of reproducibility. *J Manipulative Physiol Ther* 29:475-485.

Waltz CF, Strickland OL, Lenz ER (2005) Measurement in nursing and health research. *Nephrol Nurs J* 33:76-76.

Wickstrom G, Bendix T (2000) The "Hawthorne effect"-What did the original Hawthorne studies actually show? *Scand J Work Environ Health* 26:363-367.

Zhou L, Liu T, Li X, Su C, Ji X, Zhao M (2017) A new interpretation and quantitative method for diplopia test: 304 cases of ocular motor nerve palsy for clinical test and verify. *Int J Ophthalmol* 10:1768-1770.

Zhou LY, Liu TJ, Li XM (2016) Adult reference values of the computerized diplopia test. *Int J Ophthalmol* 9:1646-1650.

(Copyedited by Baron R, Stow A, Wang J, Li CH, Qiu Y, Song LP, Zhao M)