Abstract

Objective: The unified multiple system atrophy rating scale (UMSARS) was used to evaluate various symptoms of multiple system atrophy (MSA). And UMSARS part 1 was originally developed for use in interviews, but the need for telemedicine is increasing in COVID-19 pandemic. The purpose of this study is to evaluate the reliability of the UMSARS part 1 telephone survey.

Methods: Thirty-two MSA patients took the UMSARS part 1 face-to-face, followed by two more telephone evaluations. Intraclass correlation coefficients (ICC) and Cronbach's alpha (α) coefficients were calculated, and the inter-rater reliability was determined. At the same time, we asked about the problems in COVID-19 pandemic.

Results: The study participants included 15 men and 17 women with mean age of 67.1 years (SD, 8.3). For the total UMSARS part 1 score, the inter-rater ICC and Cronbach's α coefficient were 0.89 to 0.92, and 0.84 to 0.87, respectively. More than half of the items had a relatively high ICC. Cronbach's α coefficients were more than 0.7 for all items. Changes that occurred in COVID-19 pandemic included reduced outings and lack of rehabilitation in about half of the cases.

Conclusion: The UMSARS part 1 has high inter-rater reliability and internal consistency. Evaluation of subjective symptoms showed that some variability could occur. In addition,

there was concern about the influence of lack of rehabilitation due to COVID-19 pandemic.

Key words: multiple system atrophy, unified multiple system atrophy rating scale, intraclass correlation coefficient, Cronbach's α coefficient, telemedicine

Abbreviations: MSA, multiple system atrophy; UMSARS, unified multiple system atrophy rating scale; COVID-19, coronavirus disease 2019; SARA, scale for the assessment and rating of ataxia; SCOPA-AUT, scales for outcomes in Parkinson's disease-autonomic questionnaire; CRC, clinical research coordinator; ICC, intraclass correlation coefficients; MMSE, mini mental state examination

1. Introduction

Multiple system atrophy (MSA) is an adult-onset, refractory, progressive neurodegenerative disease characterized by cerebellar ataxia, parkinsonism, and autonomic dysfunction [1,2]. MSA is classified as MSA-P with predominant Parkinsonism or MSA-C with predominant cerebellar ataxia. MSA-C has been reported to be more common in Japan [3]. It has also been discovered that appropriate evaluative indicators are not in place for MSA. This has led to the failure of MSA clinical trials [4]. Appropriate evaluative indicators are essential when considering future clinical trials for MSA.

The Unified Multiple System Atrophy Rating Scale (UMSARS) [5], published in 2004, is a widely used clinical assessment for evaluating MSA. It consists of a 12-item interview (Part 1), 14-item neurological examination (Part 2), orthostatic hypotension assessment (Part 3), and global disability scale (Part 4; five-level evaluation of activities of daily living). In Parts 1 and 2, each item is scored from 0 to 4, with 0 denoting no difficulty and 4 denoting severe difficulty. Thus, the minimum score that can be achieved is 0, while the maximum score is 48 in Part 1 and 56 in Part 2.

On the other hand, after 2020, the need for telemedicine increased due to the COVID-19 pandemic [6]. Originally, the evaluation of MSA was preferably performed in a face-

to-face interview. However, in a pandemic situation, this evaluation should be also performed without face-to-face contact, such as with a telephone. UMSARS Part 1 was developed on the premise of an evaluation by interview. Its' reliability has been confirmed in previous reports [7]. But, the reliability of the evaluation of UMSARS Part 1 using a method such as the telephone has not been examined. This study was primarily aimed at evaluating the reliability of the UMSARS Part 1 via the telephone. We also investigated the problems facing patients with MSA during the COVID-19 pandemic.

2. Patients and methods

This prospective, multicenter, observational study included outpatients at Hokkaido University Hospital, Japanese Red Cross Asahikawa Hospital, Hakodate Municipal Hospital, Obihiro Kosei Hospital, and Date Red Cross Hospital in Hokkaido, Japan. between October 2020 and November 2020. The enrolled patients were diagnosed with possible or probable MSA according to the second consensus statement [8]. Patients were divided into MSA-C and others, namely MSA-P and MSAP=C (Parkinsonism and cerebellar ataxia equally dominated), and their patient backgrounds are described in Table 1. This study was approved by the Institutional Review Board of Hokkaido University Hospital (019-0040). Written informed consent was obtained from all

participants prior to the start of the study. Patients with severe cognitive impairment who could not answer the questions properly were excluded from the study.

A published Japanese translation of the UMSARS [9] was used for the telephone interviews. This study was conducted with reference to the methods of previous studies. These studies investigated the reliability of symptom assessment scales such as the Scale for the Assessment and Rating of Ataxia (SARA), Berg balance scale, and Scales for Outcomes in Parkinson's Disease-Autonomic questionnaire (SCOPA-AUT) [10–12]. First, Rater A conducted a face-to-face interview and evaluated the interview using the Japanese version of the UMSARS Part 1. One week later, Rater B reassessed the UMSARS Part 1 interview using the telephone. Finally, after another week, Rater C again evaluated the UMSARS Part 1 interview using the telephone. Rater B interviewed the patient using UMSARS Part 1. Rater B also asked questions regarding the problems patients faced concerning the COVID-19 pandemic. If the patient had difficulty speaking due to dysarthria or a tracheostomy, a family member was asked to answer on the patient's behalf. Raters A and B were neurologists, and Rater C was a well-trained clinical research coordinator (CRC). Each trial was blinded. There were no clinical interventions during the study period. However, there were prior treatments. The collected data were anonymized and statistical analyses were performed.

2.1 Statistical analysis

The target number of cases was set at 32. The validation study of SARA was performed in a smaller group (n = 27), whereas the SCOPA-AUT study was conducted with a similar number of participants (n = 31) and the validity was confirmed [10,12]. JMP Pro 16 (SAS Institute Inc., Cary, NC, USA) was used for statistical analysis. Inter-rater reliability between Raters A and B, A and C, and B and C was also assessed. The total UMSARS Part 1 score and the scores of each item were analyzed based on Cronbach's α (whether the same construct can be measured) and intraclass correlation coefficients (ICCs). As a sub-analysis, we divided the subtotal of items 1-8 related to motor function and the subtotal of items 9-12 related to autonomic dysfunction. Items were considered to have high internal consistency if the Cronbach's α coefficient was greater than 0.8. ICC values were interpreted as follows: slight (0.000–0.200), fair (0.201–0.400), moderate (0.401–0.600), substantial (0.601–0.800), and approximately perfect (0.801–1.000) [13]. Continuous variables are presented as 'mean' ± 'standard deviation' (SD).

3. Results

Thirty-two patients (15 men and 17 women; age range, 49–83 years; mean age, 67.1 ±

8.3 years) were enrolled in this study. The demographic characteristics of the 32 patients are shown in Table 1. They included 21 MSA-C cases, 9 MSA-P cases, and 2 MSA-C=P cases (Parkinsonism and cerebellar ataxia equally dominated).

The total UMSARS Part 1 scores and the scores for individual items are shown in Table 2. The mean total scores for the first, second, and third assessments were 18.75 ± 9.64 (n = 32, range 6–46), 19.41 ± 9.28 (n = 32, range 5–42), and 18.66 ± 9.40 (n = 32, range 5–44), respectively. There were no statistically significant differences between the assessments. The ICCs and Cronbach's α coefficients are presented in Table 3. The ICC of the total UMSARS Part 1 score between Raters A and B was 0.89, and Cronbach's α coefficient was 0.87. In a sub-analysis divided into items 1-8 related to motor function and items 9-12 related to autonomic dysfunction, items 1-8 generally had higher ICC and Cronbach's α coefficients than items 9-12. ICCs and Cronbach's α coefficients between Raters A and C and Raters B and C showed similar trends. However, there were certain differences between the items. For the ICC between Raters A and B, items 5 (dressing) and 6 (hygiene) were evaluated as approximately perfect. Items 2 (swallowing), 3 (handwriting), 4 (cutting food and handling utensils), 8 (falling), and 10 (urinary function) had substantial scores. Items 1 (speech), 7 (walking), 9 (orthostatic symptoms), 11 (sexual function), and 12 (bowel function) had moderate scores. No item was judged

as fair or slight in a comparison of Raters A and B. Inter-rater ICC values for 7 out of 12 items were over 0.6. The Cronbach's α coefficients for all items were greater than 0.7.

When asked what had changed from their previous activities due to the COVID-19 pandemic, 13 patients mentioned that they went out less than before. Three patients said they had to refrain from visiting a doctor or physical therapist because they were worried about infection. Five patients said they were strengthening their infection control methods. Three patients stated they were experiencing trouble due to violations of visitation restrictions while they were in a hospital or medical facility. On the other hand, eight patients answered that they were rehabilitating, even during the COVID-19 pandemic. The patients' suggestions for public institutions included the expansion of visitation systems using tools such as online interviews, thorough infection control measures, and dissemination of methods to easily confirm the presence or absence of COVID-19 infection.

4. Discussion

The participants in this study experienced more MSA-C than MSA-P. The UMSARS Part 4 score of MSA-P cases was higher than that of MSA-C cases. These trends are in line with the results of previous studies on the natural history of MSA [3,14].

The average total scores of UMSARS Part 1 in this study were 18.75 ± 9.64, 19.41 ± 9.28, and 18.66 ± 9.40 for the first, the second, and the third assessment, respectively. These scores were lower than those of similar study subjects. This reflects the inclusion of patients with relatively early MSA symptoms in the study [14–16].

The total score of the Japanese version of the UMSARS Part 1 using the telephone showed typically high ICC and Cronbach's α coefficients. This indicated good test-retest reliability and internal consistency. More than half of the individual items exhibited high ICCs and all items showed relatively high Cronbach's α coefficients. The results of sub-analyses divided into items 1-8 related to motor function and items 9-12 related to autonomic dysfunction suggested that there may be more variation in evaluation for autonomic symptoms. There was little difference between the reliability of the neurologists (Raters A and B) and the reliability between neurologists and a trained CRC (Raters A and C, or Raters B and C). This indicates that UMSARS Part 1 using the telephone and conducted by a well-trained CRC is as viable as a telephone interview conducted by neurologists.

In addition, UMSARS Part 1 was originally developed on the premise of face-to-face contact. However, this study demonstrated that the interview could be performed using telemedicine without major problems. But, there were variations in the reliability among

some items. Stable reliability was confirmed for items 5 (dressing) and 6 (hygiene). Items 2 (swallowing), 3 (handwriting), 4 (cutting food and handling utensils), 7 (walking), 8 (falling), 9 (orthostatic symptoms), and 12 (bowel function) were the second-most reliable. Items 1 (speech) and 11 (sexual function) were relatively unreliable. Items with a low ICC might lack the accuracy of the answer because of unclear lines, such as the mild, moderate, and severe evaluation criteria. This was highlighted in other reports as a matter to be improved in UMSARS [17]. Regarding the large variation in speech scores, it was possible that patients were more likely to notice that it was difficult to speak on the phone. It was also likely that it was easier for patients to discuss changes in their physical condition at each in-person evaluation. Regarding sexual function, there was a possibility that the answer was easy to obscure as there are many Japanese people who tend to be embarrassed regarding the discussion of sexual matters. Similar considerations were made in a study that examined the reliability of SCOPA-AUT.

The limitations of this study are as follows. 1) During this study we used a phone without a screen. We could not use a video phone. If we had been able to use a video phone, it is possible that the UMSARS Part 1 scores may have been slightly different from this result. 2) Cognitive function scores, such as the Mini Mental State Examination (MMSE), were not evaluated in this study. Participants who did have deficits in daily

conversation were recruited. No patient had severe cognitive impairments to the extent that they could not answer a question. However, strictly speaking, it would be desirable to be able to confirm that cognitive function was within the normal range by using cognitive function tests such as the MMSE. 3) Family members answering on behalf of the patient may be insufficient. In this study, patients who had difficulty speaking because of MSA symptoms or a tracheostomy were permitted to be represented by their accompanying family members. Undeniably, the patient's verbatim answers might have been slightly different from those of family members who answered on the patient's behalf. For greater accuracy, it would be desirable to obtain all answers from the patients themselves.

5. Conclusion

The Japanese version of UMSARS Part 1 has high inter-rater reliability and high internal consistency. The UMSARS Part 1 is available for use via telephone. It can also be performed by a well-trained CRC in place of a neurologist. However, the telephone scores may vary more than a face-to-face interview. Due to the impact of the COVID-19 pandemic, leaving one's home was restricted for situations such as rehabilitation appointments. There was concern that the activities of daily living of patients with MSA

would decline.

References

1. Graham JG, Oppenheimer DR. Orthostatic hypotension and nicotine sensitivity in a case of multiple system atrophy. J Neurol Neurosurg Psychiatry. 32 (1969) 28-34.

2. Quinn N. Multiple system atrophy - The nature of the beast. J Neurol Neurosurg Psychiatry. Suppl (1989) 78-89. doi:10.1136/jnnp.52.Suppl.78

3. Sakushima K, Nishimoto N, Nojima M, et al. Epidemiology of Multiple System Atrophy in Hokkaido, the Northernmost Island of Japan. Cerebellum.14(6) (2015) 682-687. doi:10.1007/s12311-015-0668-6

4. Bitan G. The recent failure of the PROMESA clinical trial for multiple system atrophy raises the question—are polyphenols a viable therapeutic option against proteinopathies? Ann Transl Med. 8(11) (2020) 719-719. doi:10.21037/ATM.2020.01.117

5. Wenning GK, Tison F, Seppi K, et al. Development and validation of the Unified Multiple System Atrophy Rating Scale (UMSARS). Mov Disord. 19(12) (2004) 1391-1402. doi:10.1002/mds.20255

6. Greenough MC, Sajjadi NB, Rucker B, Vassar M, Hartwell M. The use of telecommunication and virtualization among ongoing and discontinued COVID-19 clinical trials: A cross-sectional analysis. Contemp Clin Trials. 114 (2022) 106681.

doi:10.1016/j.cct.2022.106681

7. Krismer F, Seppi K, Ois Tison F, et al. The Unified Multiple System Atrophy Rating Scale: Intrarater Reliability on behalf of the European Multiple System Atrophy Study Group. Mov Disord. 27(13) (2012) 1683–1685. doi:10.1002/mds.25181

8. Gilman S, Wenning FG, Low P, et al. Second consensus statement on the diagnosis of multiple system atrophy Background: A consensus conference on multiple system atrophy (MSA) in 1998 established. Neurology. 71 (2008) 670-676.

9. Chikada A, Mitsui J, Matsukawa T, et al. Reliability and validity of Japanese version of Unified Multiple System Atrophy Rating Scale. Neurol Clin Neurosci. 9(2) (2021) 171-180. doi:10.1111/ncn3.12477

10. Sato K, Yabe I, Soma H, et al. [Reliability of the Japanese version of the Scale for the Assessment and Rating of Ataxia (SARA)]. Brain Nerve. 61(5) (2009) 591-595.

11. Matsushima M, Yabe I, Uwatoko H, Shirai S, Hirotani M, Sasaki H. Reliability of the Japanese version of the berg balance scale. Intern Med. 53(15) (2014) 1621-1624. doi:10.2169/internalmedicine.53.2662

12. Matsushima M, Yabe I, Hirotani M, Kano T, Sasaki H. Reliability of the Japanese version of the scales for outcomes in Parkinson's disease-autonomic questionnaire. Clin Neurol Neurosurg. 124 (2014) 182-184. doi:10.1016/j.clineuro.2014.07.007

13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 33(1) (1977)159-174.

14. Matsushima M, Yabe I, Sakushima K, et al. Multiple system atrophy in Hokkaido, Japan: A prospective registry study of natural history and symptom assessment scales followed for 5 years. BMJ Open. 11(2) (2021) e045100. doi:10.1136/bmjopen-2020-045100

15. Wenning GK, Geser F, Krismer F, et al. The natural history of multiple system atrophy: A prospective European cohort study. Lancet Neurol. 12(3) (2013) 264-274. doi:10.1016/S1474-4422(12)70327-7

16. Low PA, Reich SG, Jankovic J, et al. Natural history of multiple system atrophy in the USA: A prospective cohort study. Lancet Neurol. 14(7) (2015) 710-719. doi:10.1016/S1474-4422(15)00058-7

17. Palma J-A, Vernetti PM, Perez MA, et al. Limitations of the Unified Multiple System Atrophy Rating Scale as outcome measure for clinical trials and a roadmap for improvement. Clin Auton Res. 31(2) (2021) 157-164. doi:10.1007/s10286-021-00782-w

Legends

Table 1: The demographic data of study patients

MSA, multiple system atrophy; SD, standard deviation; OPCA, olivopontocerebellar atrophy; SND, striatonigral degeneration; SDS, Shy-Drager syndrome; UMSARS, unified multiple system atrophy rating scale; DaT, dopamine transporter; CPAP, continuous positive airway pressure; TPPV, tracheotomy positive pressure ventilation

Table 2: The mean scores ± standard deviation (SD) for each item and the total UMSARS part 1 score

The first assessment was conducted by rater A, and the second assessment was performed by rater B, and third was by rater C.

Table 3: Inter-rater ICCs and Cronbach's α coefficients for the UMSARS part 1

ICC, intraclass correlation coefficient

Acknowledgements

We thank all patients and families, and all concerned with the Hokkaido Rare disease Consortium for MSA (HoRC-MSA) project for their active cooperation.

Author contribution

Drafting of the manuscript: M-M.

Patient medical treatment: M-M, A-N, A-K, K-E, M-W, S-S, I-I, K-H, T-M, S-U, H-H.

Calling the patients: M-M, A-N, R-N.

Data analysis: M-M, R-N.

Review of the manuscript for intellectual content: A-N, A-K, K-E, M-W, S-S, I-I, K-H, T-M, S-U, H-H, I-Y.

Funding: Ichiro Yabe; a Grant-in-Aid for the Research Committee for Ataxic Diseases of the Research on Measures for Intractable Diseases from the Ministry of Health, Welfare and Labor, Japan (20FC1041), and The Practical Research Project for Rare/Intractable Diseases from The Japan Agency for Medical Research and Development (AMED) (21ek0109527h0001).

Declaration of Interests: None. All authors have no conflict of interest with this study.

Table 1. The demographic data of the study patients.

	Total	MSA-C	MSA-P and others
Patients, n	32	21	11
Sex, male/female	15/17	12/9	3/8
Present age, years, mean ± SD	67.1 ± 8.3	68.8 ± 7.2	63.9 ± 9.4
Onset age, years, mean ± SD	63.8 ± 8.7	65.3 ± 7.8	60.8 ± 9.8
Disease duration, years, mean ± SD	3.4 ± 2.1	3.5 ± 2.3	3.1 ± 1.6
Onset symptoms, Cerebellar / Parkinsonism / Neurogenic orthostatic hypotension, n	20/9/3	18/1/2	2/8/1
UMSARS part 4, mean ± SD	2.8 ± 1.1	2.7 ± 1.2	2.9 ± 1.0
Signs and symptoms			
Dysarthria, n	28	20	8
Saccadic eye movement, n	30	20	10
Lateral gaze disturbance, n	11	6	5
Vocal cord movement disturbance, n	10	5	5
Dysphagia, n	18	12	6
Hyperreflexia, n	8	6	2
Limb ataxia, n	29	21	8
Rest tremor, n	1	1	0
Rigidity, n	23	12	11
Orthostatic hypotension, n	12	5	7
Lower urinary tract symptoms, n	25	16	9
MRI findings			
Cerebellar atrophy, n	26	20	6
Brainstem atrophy, n	25	20	5
Hot cross bun sign, n	22	18	4
Striatum abnormality, n	10	3	7
DaT density decrease, n	15	7	8
Treatment			
Taltirelin, n	20	16	4
Levodopa, n	15	6	9
Average levodopa dose, mg/day	310.0 ± 152.6	283.3 ± 175.1	327.8 ± 143.9
Vasopressors, n	5	2	3
Dysuria drug, n	9	5	4
Gastrostomy, n	3	2	1
Tracheostomy, n	2	0	2

CPAP, n	1	1	0
TPPV, n	0	0	0

Table 2. The mean scores ± standard deviation (SD) for each item and the total UMSARS part 1 score

	First assessment (A)	Second assessment (B)	Third assessment (C)
Item	mean ± SD	mean ± SD	mean ± SD
1. Speech	1.66 ± 1.07	1.81 ± 1.12	1.16 ± 1.25
2. Swallowing	1.22 ± 1.13	1.03 ± 1.15	1.19 ± 1.23
3. Handwriting	1.72 ± 1.05	1.84 ± 1.08	2.06 ± 1.34
4. Cutting food and handling utensils	1.47 ± 1.05	1.34 ± 1.15	1.19 ± 1.09
5. Dressing	1.56 ± 1.13	1.53 ± 1.22	1.28 ± 1.25
6. Hygiene	1.69 ± 1.06	1.66 ± 1.18	1.72 ± 1.46
7. Walking	2.25 ± 1.08	2.63 ± 0.79	2.69 ± 1.15
8. Falling (rate the past month)	1.31 ± 1.47	1.06 ± 1.39	0.91 ± 1.38
9. Orthostatic symptoms	1.13 ± 1.18	1.13 ± 1.45	1.22 ± 1.43
10. Urinary function	1.34 ± 1.10	1.59 ± 1.16	1.63 ± 1.52
11. Sexual function	1.59 ± 1.76	2.06 ± 1.84	2.06 ± 1.91
12. Bowel function	1.81 ± 0.97	1.78 ± 0.83	1.63 ± 1.10
Total	18.75 ± 9.64	19.41 ± 9.28	18.66 ± 9.40

Table 3. Inter-rater ICCs and Cronbach's α coefficients for the UMSARS part 1

	Inter-rater (A and B)		Inter-rater (A and C)		Inter-rater (B and C)	
Item	ICC	Cronbach α	ICC	Cronbach α	ICC	Cronbach α
1. Speech	0.49	0.74	0.42	0.73	0.20	0.74
2. Swallowing	0.76	0.74	0.67	0.73	0.70	0.74
3. Handwriting	0.78	0.73	0.57	0.73	0.61	0.73
4. Cutting food and handling utensils	0.77	0.74	0.67	0.73	0.84	0.73
5. Dressing	0.83	0.73	0.86	0.73	0.84	0.73
6. Hygiene	0.81	0.74	0.84	0.73	0.84	0.73
7. Walking	0.57	0.74	0.57	0.74	0.71	0.74
8. Falling (rate the past month)	0.73	0.73	0.54	0.74	0.55	0.74
9. Orthostatic symptoms	0.53	0.75	0.68	0.75	0.45	0.75
10. Urinary function	0.66	0.74	0.55	0.74	0.82	0.73
11. Sexual function	0.46	0.74	0.56	0.74	0.59	0.74
12. Bowel function	0.48	0.75	0.71	0.74	0.56	0.74
Items 1 - 8	0.94	0.87	0.90	0.86	0.94	0.85
Items 9 – 12	0.65	0.81	0.78	0.80	0.79	0.80
Total	0.89	0.87	0.89	0.85	0.92	0.84