別紙様式１（論文博士申請者用）

博 士 学 位 論 文

Development of the Japanese version of the stroke stigma scale: a validity and reliability assessment
（日本語版 Stroke Stigma Scale の開発：信頼性・妥当性の検証）

（西暦）2024 年　7 月　4 日　提出

東京都立大学大学院
人間健康科学研究科 人間健康科学専攻 作業療法科学域
満期退学時学修番号：18996701
研究員受入通知書の文書番号：5 都立大荒管第 522 号
氏　名：北村　新

（指導/紹介教員名：宮本　礼子）

TOPICS IN STROKE REHABILITATION
https://doi.org/10.1080/10749357.2024.2318097

Development of the Japanese version of the stroke stigma scale: a validity and reliability assessment

Shin Kitamura [a,b], Reiko Miyamoto [c], Shota Watanabe [d], Taiki Yoshida [b], and Yoshikazu Ishii [e]

[a]Department of Occupational Therapy, Graduate School of Human Health Sciences, Tokyo Metropolitan University, Tokyo, Japan; [b]Faculty of Rehabilitation, School of Health Sciences, Fujita Health University, Aichi, Japan; [c]Department of Occupational Therapy, School of Health Science, Tokyo Metropolitan University, Tokyo, Japan; [d]Department of Rehabilitation Medicine, Tokyo Bay Rehabilitation Hospital, Chiba, Japan; [e]Department of Occupational Therapy, Faculty of Rehabilitation, Gunma Paz University, Gunma, Japan

ABSTRACT

Background: The stigma perceived by many post-stroke persons hinders their social lives. A scale to measure stigma is needed to identify social problems related to stigma, and to evaluate effectiveness of interventions.

Objectives: This study aimed to develop a Japanese version of the Stroke Stigma Scale (SSS-J), and confirm its utility by examining reliability and validity.

Methods: Eighty community-dwelling post-stroke individuals were enrolled at six sites. After translating the scale into Japanese using back translation methods, psychometric properties of the rating scale, internal scale validity, and reliability were examined to fit the Rasch model. Criterion-related validity, construct validity, and test-retest reliability were examined using total scores transformed to logit. For test-retest reliability, 30 participants completed the SSS-J twice, one week apart.

Results: Rasch analysis showed that the SSS-J had the best fit with 15 items on a 3-category rating scale. Item difficulty logits were -2.01 to 2.21, person ability logits were -4.69 to 0.62 (mean, -1.41), person reliability coefficient was 0.71 (separation index, 1.58), and item reliability coefficient was 0.96 (separation index, 5.04). For criterion validity, Spearman's rank correlation coefficient with the Center for Epidemiologic Studies Depression Scale was 0.51 ($p < 0.001$). For construct validity, Spearman's rank correlation coefficients with each subscale of the Stroke Impact Scale ranged from -0.36 to -0.16 ($p = 0.002$–0.126). For test-retest reliability, the intra-class correlation coefficient was 0.64 ($p < 0.001$).

Conclusions: The SSS-J adapted to the Rasch model was reliable and valid. This scale can be used to quantitatively measure stigma among community-dwelling post-stroke persons in Japan.

ARTICLE HISTORY
Received 22 June 2023
Accepted 4 February 2024

KEYWORDS
Cerebrovascular disorders; community; social stigma; Rasch analysis; rehabilitation; validity; reliability

Introduction

For post-stroke persons, stigma has been identified as a major factor in the inhibition of social participation and functional improvement through rehabilitation.[1] Stigma is defined as "an attribute that is deeply discrediting" and that reduces the bearer "from a whole and usual person to a tainted, discounted, and inferior one."[2] People with stroke experience are labeled by society or feel shame and inferiority because of their illness, which impairs their opportunities for social participation, such as going out, interacting with others, leisure activities, and returning to work.[3–6] These represent stroke-related stigma and are important issues that need to be addressed to enable post-stroke persons to rebuild their social lives.

Recently, instruments to assess stroke-related stigma have been developed[7–11] and quantitative research has expanded. Previous evidence indicates that many people experience stigma after stroke.[10,12–14] The experience of stigma is influenced by individual demographic characteristics (e.g. age, income) as well as clinical characteristics (e.g. level of functional independence after the stroke).[12,14] Stigma negatively affects people's social lives and leads to feelings of isolation.[15–17]

CONTACT Reiko Miyamoto miyamoto@tmu.ac.jp Department of Occupational Therapy, School of Health Science, Tokyo Metropolitan University, 7-2-10 Higashiogu, Arakawa-ku, Tokyo 116-8551, Japan

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10749357.2024.2318097

However, in Japan, the experience of stigma and its associated factors in post-stroke persons have not been quantitatively clarified, and tools to measure such stigma are lacking. Studies in Japan interviewing community-dwelling people after a stroke show that post-stroke persons feel shame and a sense of inferiority in their social lives because of their illness, and tend to avoid interacting with others and going out.[18,19] These findings indicate that post-stroke persons in Japan also experience stigma similar to that reported in other countries.[10,14,16,20,21] To understand the implications of stroke-related stigma in Japan and to examine the effects of interventions, a scale measuring stigma that can be used in Japan is needed.

The Stroke Stigma Scale (SSS) is an instrument developed to assess perceived stigma specific to post-stroke persons.[7] The feelings of inferiority and shame associated with the illness experienced by Japanese post-stroke persons, as identified in previous studies,[18,19] can be regarded as perceived stigma. Therefore, translating the SSS into Japanese would allow us to measure the stigma specific to stroke in Japan. Because the SSS is an ordinal scale, the intervals between scores are not uniform and mathematical operations such as calculating total scores should not be performed.[22] By adapting Rasch analysis,[23] which transforms ordinal data into interval data expressed in linear log establishment units, scores can be mathematically manipulated[24] and intra- and inter-individual comparisons can be made from scores.

This study aimed to develop a Japanese version of the Stroke Stigma Scale (SSS-J) that measures stigma as an interval scale, and to confirm its usefulness for community-dwelling post-stroke persons in Japan. To this end, the SSS was translated into Japanese, transformed into an interval scale by Rasch analysis, and its test-retest reliability, criterion validity, and construct validity were examined.

Materials and methods

Study outline

This was a two-part study. First, the SSS was translated from Chinese to Japanese, and next, the reliability and validity of the SSS-J were examined in community-dwelling, post-stroke persons. This study conformed to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) – a guideline for scale development.[25]

The SSS is a Chinese instrument that assesses perceived stigma specific to post-stroke persons, and has been confirmed to be reliable and valid.[7] The perceived stigma assessed by the SSS indicates actual or potentially aware physical impairment, social isolation, discrimination experience, and internalized stigma associated with stroke.[7] The SSS consists of 16 items, and each item is answered on a 5-category rating scale regarding the frequency of stigma-related experiences in the past week. The total score ranges from 16 to 80, with higher scores indicating stronger stigma.

The study protocol was approved by the Ethics Committee of Tokyo Metropolitan University (Reference: 19107, 20075), Fujita Health University (HM22–105), and Tokyo Bay Rehabilitation Hospital (250, 280). The study was conducted in accordance with the Declaration of Helsinki of 1964. All participants provided written informed consent prior to participation.

Translation of the scale

The SSS was translated into Japanese by back-translation methods[25] with two forward translators, one back-translator, and the principal investigator (SK). When translating from Chinese to Japanese (forward translation), efforts were made to ensure that the item content reflected Japanese culture and the actual situation of post-stroke persons. The forward translation was performed by two native Japanese speakers who were also familiar with Chinese and were experts in stroke rehabilitation, in consultation with the principal investigator (SK) when necessary. In addition, the quality of the translation was evaluated by the original author each time the translation was completed, and the translation process was repeated until the original author fully agreed with the translation. We then interviewed eight stroke patients who were outpatients at a rehabilitation hospital about their understanding of the items. The patient characteristics are presented in Table 1. The translators and authors reviewed the participants' comments from the interviews, and modified the

Table 1. Participant characteristics.

	Interview sample	Psychometric sample	Retest reliability sample
Total number	8	80	30
Sex, male/female, n	6/2	52/28	15/15
Age, year (SD)	50.6 (11.4)	69.3 (11.9)	72.8 (12.3)
Duration after stroke onset, year (SD)	3.0 (4.9)	6.2 (6.2)	4.9 (3.4)
Type of stroke, n			
Ischemic	2	44	18
Hemorrhagic	5	33	12
Subarachnoid hemorrhage	1	3	0
Independence level of activities of daily living, n			
Independent living outdoors	8	27	8
Independent in indoor living but needs assistance in outdoor living	0	9	4
Partial assistance with indoor living	0	43	17
Total assistance with indoor living	0	1	1

SD: standard deviation; n, number of participants.

wording of the SSS-J items as needed. After revisions were made, the original author checked the validity of the revised wording. For detailed translation procedures, please refer to Supplementary Table s1.

Reliability and validity of the Japanese scale

Participants

Participants were recruited between September 2021 and January 2023 by convenience sampling at six facilities (outpatients at one hospital, clients at four day rehabilitation centers, and clients at one home-based rehabilitation center) in the Tokyo, Chiba, Kanagawa, and Aichi prefectures. The inclusion criteria were post-stroke persons living in the community, those aged 20 years or older, and those who were able to understand the text when answering the questionnaire. Persons whose answers to the SSS-J were determined by the medical staff to be mentally burdensome were excluded. The sample size was determined based on the duration of the study and on the reference value of a minimum of 50 participants that are required to perform Rasch analysis, as previously determined.[25] After receiving an explanation regarding the study design from a representative of each facility, the participants completed the questionnaire at their convenience. To avoid influencing responses, prior to responding, participants were informed that the results of their responses to the scale would not be shared with staff at their facilities. Those who were able to provide a second set of responses to the SSS-J, one week after the first set of responses, were invited to participate in the test-retest reliability examination. The sample size was 30, which is the minimum number required for test-retest reliability, as previously determined.[25] The authors did not treat participants' stigma or related symptoms based on the results of the responses (e.g. providing counseling to highly stigmatized persons).

Instruments

The center for epidemiologic studies depression scale (CES-D).. Depression has been reported to be correlated with stigma,[14,20,21] and depression scales have been used as a reference for examining criterion validity in previous studies that developed instruments to assess stigma.[7,26] Since there is no stigma scale available in Japan, the CES-D was used in this study to examine criterion validity. The CES-D is a self-rating screening tool for depression consisting of 20 items.[27] The frequency of symptoms in the previous week is answered on a 4-category rating scale; scores range from 0 to 60, with higher scores indicating a greater degree of depression. The CES-D has been translated into Japanese, and its reliability and validity have been confirmed.[28]

Stroke impact scale ver.3 (SIS).. To examine the construct validity of the SSS-J, we used the SIS – a tool used to assess multidimensional health-related quality of life. The SIS is a patient-reported questionnaire that assesses stroke-specific health status and consists of the following nine domains: Strength, Memory, Emotion, Communication, Activities of daily living (ADL)/ Instrumental activities of daily living (IADL), Mobility, Hand function, Social participation, and Recovery.[29] The difficulty of completing each item

experienced within the previous 1 to 4 weeks is rated on a 5-category rating scale, with only the "Recovery" domain being answered on a visual analog scale from 0 to 100. The SIS has been translated into Japanese, and its reliability and validity have been confirmed.[30]

Data analyses

For each participant, if the majority of responses were missing for each instrument or SIS domain, the participant was excluded from the analysis for that instrument. Other missing values were interpolated using the expectation-maximization algorithm. After confirming the fit to the Rasch model using Rasch analysis, the criterion validity, construct validity, and test-retest reliability were examined using logit-transformed total scores. Winsteps version 5.3.0 was used for Rasch analysis, and other statistical analyses were performed using IBM SPSS Statistics version 28 (IBM Corp., Armonk, NY, USA). A p value of < 0.05 was considered statistically significant.

In the Rasch analysis, the psychometric properties of the rating scale were analyzed first. The criteria were a calibration threshold of at least 1.4 logits between the rating scale categories, and an outfit mean square (MnSq) of less than 2.0.[31,32] If these criteria were not met, the categories were collapsed, reorganized, and analyzed to determine the most suitable category. The internal validity of the scale was examined. Items that did not meet both criteria of an infit MnSq of 0.6–1.4 and standardized z goodness-of-fit statistics (Zstd) of less than 2.0 were removed.[32,33] Item and person measures were calculated, and ceiling and floor effects were confirmed. Finally, we examined the "person separation reliability," which indicates the reproducibility of the relative rankings of the study participants, and the "item separation reliability," which indicates the accuracy of the item difficulty estimates.

For criterion validity, the relationship between SSS-J and CES-D total scores was examined. Because the CES-D is an ordinal scale, the Spearman's rank correlation coefficient was calculated to demonstrate the relationship.

For test-retest reliability, the degree of agreement between the results of the two answers to the SSS-J by the same participant was examined using the intraclass correlation coefficient (ICC) (1,1).

For construct validity, the relationship between the SSS-J total score and the SIS score for each domain was examined to determine which aspects of health status were reflected by the SSS-J. Since the SIS is an ordinal scale, Spearman's rank correlation coefficient was used to examine relationships. Correlations between stigma and depression,[7,14,16,20,21,26] social participation,[16] ADL,[7] and stroke symptom severity[7] have been previously reported. Therefore, we predicted that the SIS domains of Emotion, Communication, ADL/IADL, Mobility, Social participation, and Recovery would correlate with the SSS-J, and therefore, these domains were used to examine convergent validity. The other domains (Strength, Memory, and Hand function) were predicted to be unlikely to be correlated and were used to examine discriminant validity.

Results

Translation of the scale

After two rounds of translation, the original authors agreed on the appropriateness of the wording. In the interview, four items (corresponding to items 11, 14, 15, and 16 in the original SSS) were identified as "difficult to understand phrasing." Based on the comments, the translators and principal investigator (SK) discussed and revised the wording of these items, and the translation process was completed after the original author confirmed that the translation was equivalent to the original version.

Reliability and validity

Eighty participants from six sites were enrolled. The participants' characteristics are shown in Table 1. Mean participant age was 69.3 years (SD, 11.9), 65% were male, 20 were hospital outpatients, 57 were day rehabilitation center clients (9, 10, 15, and 23 in 4 sites, respectively), and 3 were home-based rehabilitation center clients. In the rating scale analysis, the calibration threshold between categories was less than 1.4 on a 5-category rating scale (Table 2). As a result of reorganizing the

Table 2. Rating scale statistics.

Score	Frequency (%)	Outfit MnSq	Calibration threshold	Category measure
5-Category scale (12345)				
1	450 (35)	0.92	–	-2.17
2	301 (24)	0.98	-0.65	-0.85
3	293 (23)	0.91	-0.54	0.00
4	148 (12)	0.89	0.52	0.85
5	88 (7)	1.33	0.67	2.17
4-Category scale (12234)				
1	450 (35)	0.94	–	-2.80
2	594 (46)	0.88	-1.65	-0.57
3	148 (12)	0.83	0.96	0.90
4	88 (7)	1.46	0.68	2.28
4-Category scale (12334)				
1	450 (35)	0.92	–	-2.52
2	301 (24)	0.88	-1.01	-1.04
3	441 (34)	1.07	-0.97	0.66
4	88 (7)	1.22	1.98	3.12
3-Category scale (12223)				
1	450 (35)	0.91	–	-3.25
2	742 (58)	0.85	-2.14	0.00
3	88 (7)	1.24	2.14	3.25

MnSq: mean square. Parentheses indicate scoring models.

rating scale categories and examining the optimal categories, the calibration threshold between categories was greater than 1.4 logits when the second to fourth categories were combined (i.e. the 3-category rating scale), and the outfit MnSq met the criterion in all categories (Figure 1). Therefore, the SSS-J adopted a 3-category rating scale. In the examination of the internal scale validity, item 1 ("I can accept the influence of stroke on my body and appearance") did not meet the criteria (MnSq, 1.58; and Zstd, 2.92) and was removed. The SSS-J with 15 items, the modified category probability curves, and the distribution of persons and items are shown in Figure 2. The item measurement report is shown in Table 3. Item difficulty logits were -2.01 to 2.21, and person ability logits were -4.69 to 0.62 (mean -1.41). None of the participants had the lowest or highest total SSS-J scores, and neither ceiling nor floor effects were observed. The person separation index was 1.58 with a reliability coefficient of 0.71, indicating fair reliability.[34] The item separation index was 5.04 with a reliability coefficient of 0.96, indicating excellent reliability.[34] The Supplement file provides the Japanese and English notation for the SSS-J (Supplementary Tables S2 and S3) and the logit transformation table (Supplementary Table S4).

For criterion validity, Spearman's rank correlation coefficient between the SSS-J and CES-D total score (median 12, interquartile 6) was $r = 0.51$ ($p < 0.001$), indicating a moderate correlation.[35]

For test-retest reliability, the ICC (1,1) for the first and second SSS-J total scores was 0.64 (95% confidence interval 0.38–0.81, $p < 0.001$), indicating moderate reliability.[36]

For construct validity, the SIS domains of Memory, Emotion, Communication, ADL/IADL, Mobility, Social participation, and Recovery showed significant and low correlations, with no significant correlations in the domains of Strength and Hand function (Table 4).

Discussion

In this study, the SSS-J was translated into Japanese, and its reliability and validity were examined in community-dwelling post-stroke persons in Japan. We confirmed that the SSS-J adapted to the Rasch model and ensured its reliability and validity. The SSS-J is the first scale in Japan that can measure post-stroke stigma.

Analysis of the rating scale and item difficulty showed that the SSS-J fits the Rasch model with 15 items on a 3-category rating scale. One factor that may have contributed to the adoption of the three-category rating scale is that the participants in this study had more diverse life experiences than those in a previous study of hospitalized patients,[10] and their responses may have been more dependent on events in their daily lives. Since the SSS-J assesses stigma-related experiences within a short timeframe (the previous week), responses may be influenced by

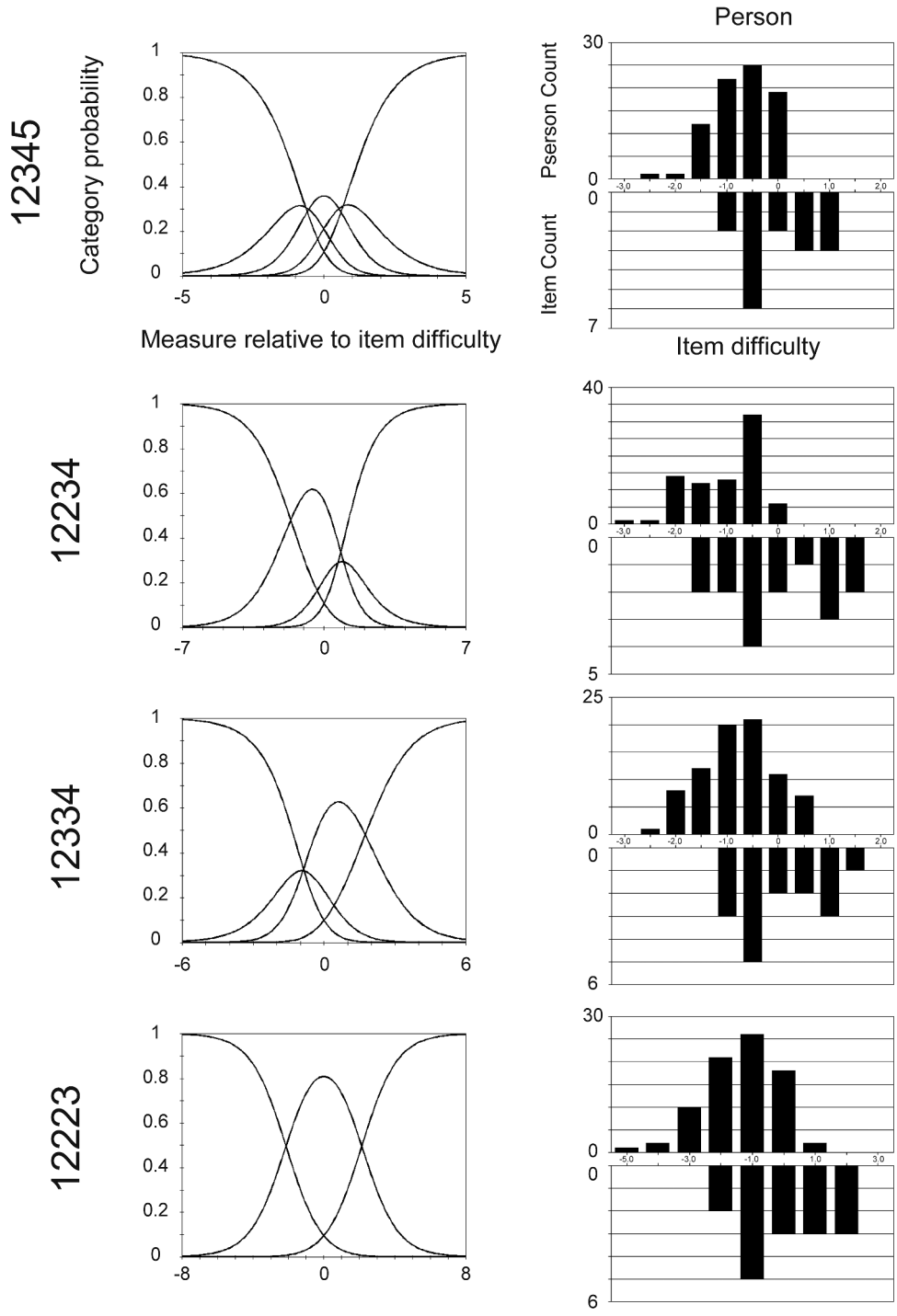

Figure 1. Category probability curves and person and item distributions for each rating category scale. The numbers to the left of the Y-axis of graphics for each category probability curve indicate the correspondence with the five rating scales in the original version. In the graphs on the right, person and item distributions are presented; the vertical axis of person count indicates that the farther the upper limit is from the horizontal axis, the higher is the number of persons, and the vertical axis of item count indicates that the farther the upper limit is from the horizontal axis, the larger is the number of items.

events during that period. These findings suggest that asking about the frequency of stigma-related experiences in three categories could reflect the degree of stigma without relying too much on the number of experiences. Regarding the examination of internal scale validity, item 1 demonstrated a misfit to the model and was removed. This item examines the acceptability of one's body and

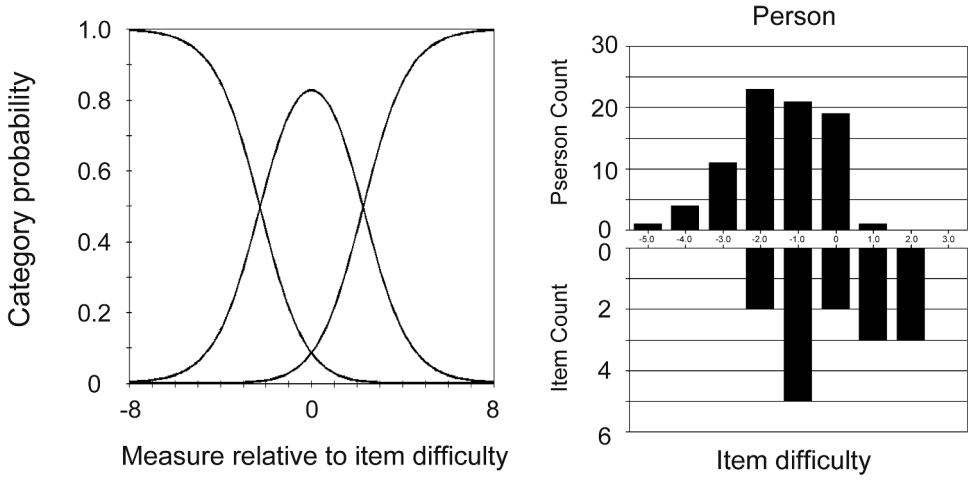

Figure 2. Category probability curves and person and item distributions for the final version of SSS-J. Category probability curves and person and item distributions for the new model with a 3-category rating scale, rather than the 5-category rating scale of the original model using Rasch analysis, are shown. In the graphs on the right, person and item distributions are presented; the vertical axis of person count indicates that the farther the upper limit is from the horizontal axis, the greater the number of persons, and the vertical axis of item count indicates that the farther the upper limit is from the horizontal axis, the larger is the number of items. SSS-J, Japanese version of the Stroke Stigma scale

Table 3. Item measurement report.

				Infit	
Item (English version)		Item difficulty (logits)	SE	MnSq	Zst
10	I had discrimination experience.	2.21	0.28	0.70	-2.31
9	Someone thinks little of my opinion because of stroke.	1.64	0.26	0.75	-2.11
8	I had been treated unfairly because of stroke.	1.58	0.26	0.87	-1.00
12	I feel I am losing face after stroke.	1.39	0.25	0.83	-1.39
6	I feel uncomfortable that the way someone treats me.	0.90	0.24	0.93	-0.46
15	I feel little respect after stroke.	0.90	0.24	0.90	-0.75
2	I have difficulty in speaking and it makes me feel embarrassed.	0.01	0.24	1.04	0.30
13	I blame myself because of stroke.	-0.11	0.24	1.07	0.46
3*	I have a meaningful life though stroke hit me.	-0.58	0.25	0.87	-0.74
7	I have smaller communication scopes after stroke.	-0.76	0.25	1.07	0.44
4*	I can persist in rehabilitation training to recover as soon as possible.	-0.82	0.25	1.08	0.48
16	I feel unhappy when I see someone doing what I cannot do.	-1.01	0.25	0.92	-0.41
11*	Someone still helps me patiently after stroke.	-1.43	0.24	1.27	1.52
14*	I still feel I am useful person after stroke.	-1.90	0.24	1.26	1.52
5	I worry to be the burden of others because of stroke.	-2.01	0.24	1.16	1.03

*reverse-scored items, SE: standard error, MnSq: mean square, Zst: t-standardized fit statistics, The item number corresponds to the original version (i.e. the number before item 1 "I can accept the influence of stroke on my body and appearance" was removed by the examination of the Internal scale validity).

Table 4. Construct validity of the SSS-J.

SIS domains	Median score (Interquartile)	Correlation coefficient*	P values
Strength	9.5 (3.3)	-0.16	0.163
Memory	32.0 (4.9)	-0.25	0.032†
Emotion	34.1 (5.0)	-0.32	0.004†
Communication	34.0 (3.8)	-0.25	0.033†
ADL/IADL	41.0 (7.7)	-0.32	0.005†
Mobility	37.0 (8.6)	-0.30	0.008†
Hand function	14.0 (7.1)	-0.18	0.126
Social participation	30.0 (9.7)	-0.36	0.002†
Recovery	50 (19.2)	-0.25	0.035†

SIS: Stroke Impact Scale, * Spearman's rank correlation coefficient, †significant correlation at $p < 0.05$.

appearance, and the reason for misfit may be the variation in "acceptability" among individuals. As the distinction between "acceptable and unacceptable" is not clear and varies individually, it is possible that the responses to item 1 varied independently of the degree of stigma. Item difficulty

logits were -2.01 to 2.21, and person ability logits were -4.69 to 0.62. The distribution of items was on the strong side of stigma compared to the distribution of participants, indicating that the SSS-J is useful for measuring persons who experience relatively strong stigma after stroke. However, no participant had the lowest total score on the SSS-J (i.e. no floor effect), indicating that a degree of measurement is possible, even for those who do not strongly experience stigma.

For the reliability of the SSS-J, the ICC was 0.64, indicating moderate test-retest reliability. Compared to the ICC of 0.92 in the original SSS,[7] the coefficient obtained in this study was relatively low. Community-dwelling participants of this study may have experienced more diverse events in their social lives than hospitalized patients who participated in the previous study.[7] This may be attributed to the greater dependence on the events of the different time periods referenced in the first and second sets of responses, resulting in a relatively lower agreement in the SSS-J responses. Expanding the period referenced in the responses may reduce the dependence on events and increase the reproducibility of answers in the SSS-J. Other reliability indices were fair, at 0.71, for person separation reliability, and excellent, at 0.96, for item separation reliability. These results indicate that the SSS-J is a reliable scale.

To confirm the criterion and structural validity of the SSS-J, we examined its relationship with other instruments. The correlation coefficient between the SSS-J and the CES-D was 0.51, similar to previous studies that found criterion validity of scales using depression screening tools,[7,26] indicating that the SSS-J can capture psychological distress due to stigma. A correlation coefficient of criterion validity of 0.70 or higher is considered desirable in comparison with the gold standard,[37] and this study showed a slightly lower correlation. Stigma and depression are strongly related but they are not identical concepts. The moderate correlations shown in this study may indicate that the SSS-J is capable of specifically measuring concepts (i.e. perceived stigma) that reflect depression but are somewhat independent. Regarding construct validity, correlations between the SIS and the SSS-J were as hypothesized in most domains, indicating that the SSS-J is a comprehensive measure that reflects each concept. However, only the memory area contradicted the hypothesis and was significantly correlated with the SSS-J. People with memory impairment after stroke fear they will be subjected to social stigma by being diagnosed with memory impairment,[38] similar to dementia patients, where the problem of social stigma has been identified.[39] In this study, participants who were aware of their memory deficits perceived stigma more strongly, indicating that the SSS-J reflects the stigma resulting from cognitive dysfunction.

The strength of the SSS-J for clinical use is that it is adapted to the Rasch model. By using this scale to measure stigma among post-stroke persons and logit-transforming the scores, the total score can be treated numerically as an interval scale. This allows for intra- and inter-individual comparisons and contributes to an accurate assessment of the degree of stigma experienced by post-stroke persons. Note that the SSS-J has been changed to a different rating scale category and number of items from the original SSS, and should be used with caution when comparing stigma internationally using both scales. In addition, responses to items on the scale may be influenced by cultural characteristics. Therefore, it should also be noted that the constructs of the scale may need to be modified for use with persons with different characteristics and in different cultures, and that reliability and validity may need to be reexamined.

This study has several limitations. All participants in this study were recruited through convenience sampling due to the psychological burden of responding to the scale, and the majority of participants attended rehabilitation facilities and had a certain amount and range of activities. Therefore, the distribution of participants enrolled in this study may be biased, and generalization of the results should be performed with caution. This distributional bias may have resulted in a poor model fit. Thus, future validation with post-stroke persons from more diverse backgrounds, such as those with lower activity levels, is required to develop a more accurate and stable scale.

Conclusions

We developed the SSS-J by back translating the original version of the scale, and examined its fit

with the Rasch model. Additionally, we evaluated its test-retest reliability, criterion validity, and construct validity. This scale was able to quantitatively measure stigma among community-dwelling, post-stroke persons in Japan. We expect that SSS-J will be a useful outcome indicator in clinical practice.

Acknowledgments

We thank the research participants and collaborators for their valuable contributions to this study. We would also like to thank Editage (www.editage.com) for English language editing.

Disclosure statement

The authors report that there are no competing interests to declare.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Shin Kitamura http://orcid.org/0000-0002-4442-0854
Reiko Miyamoto http://orcid.org/0000-0002-6058-7435
Taiki Yoshida http://orcid.org/0000-0002-5778-7102

Data availability statement

The data that support the findings of this study are available from the corresponding author, RM, upon reasonable request.

References

1. Hyman MD. The stigma of stroke. Its effects on performance during and after rehabilitation. *Geriatrics*. 1971;26:132–141.
2. Goffman E. Stigma: notes on the management of spoiled identity. London: Penguin; 1963.
3. Chang FH, Lin YN, Liou TH. Recovery after stroke: perspectives of young stroke survivors in Taiwan. *Disabil Rehabil*. 2022;44(10):1976–1983. doi:10.1080/09638288.2020.1817157.
4. Omu O, Reynolds F. Health professionals' perceptions of cultural influences on stroke experiences and rehabilitation in Kuwait. *Disabil Rehabil*. 2012;34(2):119–127. doi:10.3109/09638288.2011.591883.
5. Anderson S, Whitfield K. Social identity and stroke: 'they don't make me feel like, there's something wrong with me'. *Scand J Caring Sci*. 2013;27(4):820–830. doi:10.1111/j.1471-6712.2012.01086.x.
6. Yang K, Armstrong N, Diamond C, Lane AR, Dunne S. The meaning of loneliness to stroke survivors: a qualitative study in Northeast England. *J Health Psychol*. 2022;27(11):2539–2548. doi:10.1177/13591053211017198.
7. Zhu M, Zhou H, Zhang W, et al. The stroke stigma scale: a reliable and valid stigma measure in patients with stroke. *Clin Rehabil*. 2019;33(11):1800–1809. doi:10.1177/0269215519862329.
8. Rao D, Choi SW, Victorson D, et al. Measuring stigma across neurological conditions: the development of the stigma scale for chronic illness (SSCI). *Qual Life Res*. 2009;18(5):585–595. doi:10.1007/s11136-009-9475-1.
9. Molina Y, Choi SW, Cella D, Rao D. The stigma scale for chronic illnesses 8-item version (SSCI-8): development, validation and use across neurological conditions. *Int J Behav Med*. 2013;20(3):450–460. doi:10.1007/s12529-012-9243-4.
10. Sarfo FS, Nichols M, Qanungo S, et al. Stroke-related stigma among West Africans: patterns and predictors. *J Neurol Sci*. 2017;375:270–274. doi:10.1016/j.jns.2017.02.018.
11. Wan M, Tan Y, Huang Y, et al. Development and psychometric evaluation of public stigma of stroke scale (PSSS). *Sci Rep*. 2023;13(1):545. doi:10.1038/s41598-023-27504-8.
12. Zheng Z, Song R, Zhao Y, Lv H, Wang Y, Yu C. An investigation of the level of stigma and the factors influencing it in the rehabilitation of young and middle-aged stroke patients-a cross-sectional study. *BMC Neurol*. 2023;23(1):139. doi:10.1186/s12883-023-03189-4.
13. Li C, Hu M, Yang T, Shao X, Zheng D. Correlates of stigma for poststroke patients: a meta-analysis. *J Clin Nurs*. 2023;32(9–10):1952–1962. doi:10.1111/jocn.16250.
14. Zhu M, Zhou H, Zhang W, et al. Stigma experienced by Chinese patients with stroke during inpatient rehabilitation and its correlated factors: a cross-sectional study. *Top Stroke Rehabil*. 2019;26(5):342–348. doi:10.1080/10749357.2019.1605759.
15. Wu Y, Yan Z, Fornah L, Zhao J, Wu S. The mediation effect of social support between stigma and social alienation in patients with stroke. *Front Public Heal*. 2023;11:1290177. doi:10.3389/fpubh.2023.1290177.
16. Lu Q, Wang D, Fu L, et al. The effect of stigma on social participation in community-dwelling Chinese patients with stroke sequelae: a cross-sectional study. *Clin Rehabil*. 2022;36(3):407–414. doi:10.1177/02692155211050558.
17. Fan W, Ma KK, Yang CX, Guo YL. The mediating effect of stigma between self-perceived burden and loneliness in stroke patients. *Front Psychiatry*. 2023;14:1219805. doi:10.3389/fpsyt.2023.1219805.

18. Kitamura S, Miyamoto R. The process to the use or non-use of the paretic hand in daily life in persons with post-stroke hemiplegia. *Japanese Occupat Therap Res*. 2019;38:45–54. (in Japanese).
19. Takashima R, Murata W, Saeki K. The meaning of movement-related experiences of post-stroke hemiplegia patients in the maintenance period: using hermeneutic phenomenological research. *Japanese Occupat Therap Res*. 2011;30:602–611. (in Japanese).
20. Lu Q, Deng C, Fu L, et al. Reliability and validity of a Chinese version of the stigma scale for chronic illness (SSCI) in patients with stroke. *Top Stroke Rehabil*. 2019;26(4):312–317. doi:10.1080/10749357.2019.1592306.
21. Hu R, Wang X, Liu Z, et al. Stigma, depression, and post-traumatic growth among Chinese stroke survivors: a longitudinal study examining patterns and correlations. *Top Stroke Rehabil*. 2022;29(1):16–29. doi:10.1080/10749357.2020.1864965.
22. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil*. 1989;70:308–312.
23. Rasch G. *Studies in Mathematical Psychology: I. Probabilistic Models for Some Intelligence and Attainment Tests*. Oxford, England: Nielsen & Lydiche; 1960.
24. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989;70:857–860.
25. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1171–1179. doi:10.1007/s11136-017-1765-4 .
26. Yoo SH, Kim SR, So HS, et al. The validity and reliability of the Korean version of the stigma scale for chronic illness 8-items (SSCI-8) in patients with neurological disorders. *Int J Behav Med*. 2017;24(2):288–293. doi:10.1007/s12529-016-9593-4.
27. Nudo RJ, Wise BM, SiFuentes F, Milliken GW. Neural substrates for the effects of rehabilitative training on motor recovery after ischemic infarct. *Sci*. 1996;272(5269):1791–1794. doi:10.1126/science.272.5269.1791.
28. Shima S, Shikano T, Kitamura T, Asai M. New self-rating scales for depression. *Seishin Igaku*. 1985;27:717–723. (in Japanese).
29. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The stroke impact scale version 2.0: evaluation of reliability, validity, and sensitivity to change. *Stroke*. 1999;30(10):2131–2140. doi:10.1161/01.STR.30.10.2131.
30. Ochi M, Ohashi H, Hachisuka K, Saeki S. The reliability and validity of the Japanese version of the stroke impact scale version 3.0. *J Uoeh*. 2017;39(3):215–221. (in Japanese). doi:10.7888/juoeh.39.215.
31. Linacre JM. Investigating rating scale category utility. *J Outcome Meas*. 1999;3:103–122.
32. Bond T, Yan Z, Heene M. *Applying the Rasch Model Fundamental: Measurement in the Human Sciences*. 4th ed. London: Routledge; 2021.
33. Wright B, Linacre M, Gustafsson JE, Martin-Loff P. Reasonable mean-square fit values. *Rasch Meas Trans*. 1994;8:370.
34. Fisher W. Rating scale instrument quality criteria. *Rasch Meas Trans*. 2007;21:1095.
35. Guilford JP. *Fundamental Statistics in Psychology and Education*. 3rd ed. New York: McGraw Hill; 1956.
36. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. doi:10.1016/j.jcm.2016.02.012.
37. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42. doi:10.1016/j.jclinepi.2006.03.012.
38. Tang EYH, Price C, Stephan BCM, Robinson L, Exley C. Gaps in care for patients with memory deficits after stroke: views of healthcare providers. *BMC Health Serv Res*. 2017;17(1):634. doi:10.1186/s12913-017-2569-5.
39. Vernooij-Dassen MJFJ, Moniz-Cook ED, Woods RT, et al. Factors affecting timely recognition and diagnosis of dementia across Europe: from awareness to stigma. *Int J Geriat Psychiatry*. 2005;20(4):377–386. doi:10.1002/gps.1302.

Supplementary table 1. Translation procedures

Phase	Description of the process
1. Forward Translation	Two translators independently translated the original SSS into Japanese. Both were occupational therapists with experience in the rehabilitation of persons post-stroke and were native Japanese speakers. They had resided in China for two years, had clinical experience in Chinese hospitals, and had the ability to converse in fluent Chinese with native Chinese speakers, understand complex text, and engage in technical discussions in their areas of expertise. One of the two translators (Translator 1, T1) understood the structure and details of the scale, whereas the other (Translator 2, T2) was only informed that the scale measured stroke-related stigma.
2. Reconciliation	The first author (SK), a native Japanese speaker, reviewed the two translations, had discussions with the two translators regarding any discrepancies, and combined the two translations into one.
3. Back Translation	The Japanese SSS was translated back into Chinese by a native Chinese speaker (Translator 3, T3). T3 was an occupational therapist with experience in the rehabilitation of persons post-stroke in Japan, a native Chinese speaker who had resided in Japan and had the ability to express herself naturally and fluently in Japanese in social situations, including academic and professional life. T3 was informed that the scale measured stroke-related stigma and was not given any other details.
4. Review of the Back Translation	The original author of the SSS evaluated the back-translated version. Items that were flagged by the original author as needing modifications were re-translated into Japanese by SK and T1, and then T3 translated them into Chinese; later, the original author checked them again. This procedure was repeated until the original author agreed with the translation results.
5. Cognitive Debriefing	Community-dwelling, post-stroke persons were interviewed regarding their ease of understanding of the items. The participants were recruited through convenience sampling of outpatients at a rehabilitation hospital. Participants were invited to a semi-structured interview after completing the SSS-J. The interviews were conducted by an occupational therapist (SW) who understood the details of the scale. Data collection continued until no new opinions were obtained during the interviews.
6. Review of Cognitive Debriefing Results and Finalization	SK, SW, and T1 reviewed the participants' comments obtained from the interviews and modified the phrasing of the SSS-J items as necessary. After modification, the items were translated into Chinese by T3, and the appropriateness of the modifications was verified by the original author. This procedure was repeated until the original author's approval was obtained.

Supplementary table 2. Japanese version of the Stroke Stigma Scale

以下の 15 項目の内容に関して、この 1 週間の生活のなかで感じた頻度に
当てはまる番号を選択してください。

No	質問項目	解答欄		
		全く当てはまらない	ときどき当てはまる	いつも当てはまる
1	私は話をする時に喋りにくく気まずさを感じる.	1	2	3
2※	脳卒中になったことで不便さはあるものの，私の生活は充実している.	1	2	3
3※	私はリハビリテーションを継続することで，できる限り早く回復する自信がある.	1	2	3
4	脳卒中になってから私は周囲の人に負担をかけていると感じる.	1	2	3
5	脳卒中になってから周囲の人が私を手助けしてくれる際，方法によっては不快な思いをする.	1	2	3
6	脳卒中になってから私は人との交流が減った.	1	2	3
7	脳卒中になってから私は不公平な扱いを受けた.	1	2	3
8	脳卒中になってから周囲の人に私の意見を聞き入れてもらえない感じがする.	1	2	3
9	脳卒中になってから私は差別をされることがある.	1	2	3
10※	脳卒中になってから周囲の人は私に対して寛容である.	1	2	3
11	脳卒中になったことで私は尊厳を失ったと感じる.	1	2	3
12	脳卒中になってから私は自分自身をよく責める.	1	2	3
13※	脳卒中になっても私はまだ人の役にたてる存在だと思う.	1	2	3
14	脳卒中になってから私は人から尊敬されていないと感じる.	1	2	3
15	できないことがあり私は悲しく感じる.	1	2	3

※反転項目

Supplementary table 3. Japanese version of the Stroke Stigma Scale (English notation)

For each of the following 15 items, please select the number that corresponds to the frequency with which you perceived the item in your daily life during the past week.

No	Items	Responses		
		Never	Sometimes	Always
1	I have difficulty in speaking and it makes me feel embarrassed.	1	2	3
2*	I have a meaningful life though stroke hit me.	1	2	3
3*	I can persist in rehabilitation training to recover as soon as possible.	1	2	3
4	I worry to be the burden of others because of stroke.	1	2	3
5	I feel uncomfortable that the way someone treats me.	1	2	3
6	I have smaller communication scopes after stroke.	1	2	3
7	I had been treated unfairly because of stroke.	1	2	3
8	Someone thinks little of my opinion because of stroke.	1	2	3
9	I had discrimination experience.	1	2	3
10*	Someone still helps me patiently after stroke.	1	2	3
11	I feel I am losing face after stroke.	1	2	3
12	I blame myself because of stroke.	1	2	3
13*	I still feel I am useful person after stroke.	1	2	3
14	I feel little respect after stroke.	1	2	3
15	I feel unhappy when I see someone doing what I cannot do.	1	2	3

* reverse-scored items

Supplementary table 4. Translation table

Raw score	Logit	SD
15	-6.86	1.87
16	-5.55	1.07
17	-4.7	0.81
18	-4.13	0.71
19	-3.67	0.65
20	-3.27	0.62
21	-2.9	0.6
22	-2.55	0.59
23	-2.21	0.58
24	-1.88	0.57
25	-1.55	0.57
26	-1.23	0.56
27	-0.92	0.56
28	-0.61	0.56
29	-0.3	0.55
30	0.01	0.55
31	0.31	0.55
32	0.62	0.55
33	0.92	0.55
34	1.23	0.56
35	1.54	0.56
36	1.86	0.57
37	2.18	0.58
38	2.52	0.59
39	2.88	0.61
40	3.26	0.63
41	3.67	0.66
42	4.14	0.72
43	4.72	0.82
44	5.58	1.08
45	6.89	1.87

By converting raw scores to logits, numerical values can be treated as interval scales.