



## TEST–RETEST RELIABILITY OF THE PARTICIPATION DOMAIN OF THE STROKE IMPACT SCALE IN PERSONS WITH CHRONIC STROKE

Elisabeth EKSTRAND, RPT, PhD<sup>1,2</sup>, Jan LEXELL, MD, PhD<sup>1,4</sup> and Christina BROGÅRDH, RPT, PhD<sup>1,3</sup>

From the <sup>1</sup>Department of Health Sciences, Lund University, Lund, <sup>2</sup>Department of Hand Surgery, Skåne University Hospital, Malmö,

<sup>3</sup>Department of Neurology and Rehabilitation Medicine, Skåne University Hospital, Lund, and <sup>4</sup>Department of Neuroscience, Rehabilitation Medicine, Uppsala University, Uppsala, Sweden

**Objective:** To evaluate the test–retest reliability and variability of the Participation domain of the Stroke Impact Scale (SIS Participation) in persons with stroke as it is widely used to assess perceived participation in rehabilitation after stroke.

**Design:** A test–retest design.

**Subjects:** Forty-five persons (mean age 65 years) with mild to moderate disability at least 6 months post-stroke.

**Methods:** The SIS Participation domain was rated on 2 occasions, 1 week apart. The test–retest reliability of the total score was evaluated using Kappa statistics. The 8 item scores were evaluated by the proportion of participants who rated the same score (percentage agreement, PA) or  $\pm 1$  point (PA  $\leq 1$  point) at T1 and T2. The Svensson method was used to evaluate systematic and random disagreement.

**Results:** The test–retest reliability of the total score showed excellent agreement (Kappa coefficient = 0.79). The items showed high PA  $\leq 1$  point (> 82%). No items, except 2, showed a systematic disagreement, and no items showed a random disagreement according to the Svensson method.

**Conclusion:** The SIS Participation domain is reliable in persons with chronic stroke and mild to moderate disability and can be used to assess perceived participation in this population.

**Key words:** outcome assessment; psychometrics; rehabilitation; reproducibility of results; self-report; social participation; stroke.

Accepted Jun 12, 2018; Epub ahead of print Aug 8, 2018

J Rehabil Med 2018; 50: 00–00

Correspondence address: Elisabeth Ekstrand, Department of Health Sciences, Physiotherapy Research Group, Lund University, Box 157, SE-221 00 Lund, Sweden. E-mail: elisabeth.ekstrand@med.lu.se

Stroke is one of the main causes of disability worldwide (1). It often leads to a variety of symptoms, such as sensorimotor deficits, cognitive dysfunction and communication difficulties, and affects the ability to perform daily activities and participate in society (1). The consequences of stroke can be classified according to the International Classification of Functioning Disability and Health (ICF), which describes functioning and disability as a dynamic interaction between body functions and structures, activities and participation (2).

### LAY ABSTRACT

After stroke, many persons perceive difficulties in resuming meaningful activities, social relations and being in control of their own lives. To address participation restrictions is therefore important in stroke rehabilitation. The Participation domain of the Stroke Impact Scale, SIS Participation, is commonly used to assess perceived restrictions in participation after stroke. The domain includes 8 items: impact of stroke on work; social activities; quiet recreations; active recreations; role as a family member and/or friend; religious or spiritual activities; ability of life control and ability to help others. This study shows that the SIS Participation domain is reliable and can be used to assess perceived participation after stroke.

According to the ICF, participation is defined as involvement in life situations, which includes being able to take part in meaningful activities, social relations and being in control of one's own life (2). Many persons after stroke perceive participation restrictions due to difficulties resuming previous social activities and roles. To assess and improve participation after stroke is therefore an important goal in rehabilitation (3).

Perceived participation after stroke is commonly assessed with the participation domain of the Stroke Impact Scale (SIS) 3.0 (4, 5). The SIS is a Rasch analysed comprehensive interview-based measure of self-perceived health after stroke. It includes 8 domains, 1 of which, SIS Participation, covers participation and can be used as a separate scale (4).

In order to evaluate changes in perceived participation a measure must be reliable. Two previous studies (6, 7) have evaluated the psychometric properties of the SIS 3.0 (current version) and found that the test–retest reliability is acceptable to high. However, these authors did not specifically investigate and discuss the reliability of the SIS Participation domain and did not comprehensively evaluate the test–retest reliability in terms of agreement and systematic and random disagreements (measurement variability).

The aim of this study was to perform a comprehensive evaluation of the test–retest reliability of the SIS Participation domain as a measure of perceived participation in persons with chronic stroke.

## METHODS

## Participants

Forty-five community-dwelling persons with ischaemic or haemorrhagic stroke, who had been treated at the Skåne University Hospital in the southern part of Sweden, were recruited from April to December 2013. The participants had mild to moderate disability (i.e. 1 to 3 according to the modified Rankin Scale (8)) and were at least 6 months post-stroke. The exclusion criterion was an inability to understand and follow test instructions due to cognitive impairments or communication difficulties.

## Ethics

Prior to inclusion, information about the study was provided and the individuals gave their written informed consent to participate. The principles of the Declaration of Helsinki were followed and the study was approved by the Regional Ethical Review Board, Lund, Sweden (Dnr 2012/591).

## Stroke Impact Scale Participation domain

The SIS Participation domain (SIS 3.0) (4) includes 8 items: impact of stroke on work; social activities; quiet recreations; active recreations; role as a family member and/or friend; religious or spiritual activities; ability to control one's life; and ability to help others. The items are scored on a 5-point scale according to the person's self-perceived restriction during the past 4 weeks, where 5=limited none of the time; 4=limited a little of the time; 3=limited some of the time; 2=limited most of the time; and 1=limited all of the time. The mean for the items is calculated as a composite score and converted into an index scale ranging from 0 to 100. A higher score indicates a greater degree of perceived participation. The SIS has been translated by MAPI Research Institute into several languages (<http://www.mapi-institute.com>) including a Swedish version, which was used in this study.

## Procedure

Participants were interviewed about their perceived participation on 2 occasions (referred to as test occasion 1, T1, and test occasions 2, T2), 1 week apart. The interviews were performed under similar conditions according to the COSMIN standards (9): at the same place in a quiet and separate room of the hospital, by the same interviewer (first author) and at the same time of the day. The SIS Participation domain took approximately 5–10 min to complete.

Prior to the ratings of SIS Participation the participants had rated their perceived strength in both upper and lower extremities according to SIS domain 1. They were also asked about their walking ability, if they used walking aids (yes/no), if they perceived shoulder pain (yes/no) or fatigue (yes/no). Their level of disability was rated according to the modified Rankin Scale (8).

## Statistical analysis

Descriptive statistics, such as frequencies and means and standard deviations (SD), were calculated for demographic data and clinical characteristics of the participants. The total scores and item scores of SIS Participation were presented as medians and quartiles (i.e. 25<sup>th</sup> and 75<sup>th</sup> percentiles).

The test–retest reliability of SIS Participation was evaluated for the total score (ranging from 0–100) with Kappa statistics (quadratic weights) to assess the proportion of agreement

observed beyond the agreement expected by chance (10). The strength of the Kappa coefficient was interpreted as <0.40 poor, 0.40–0.75 fair to good, and >0.75 excellent (11).

The test–retest reliability of the item scores (ranging from 0 to 5) was evaluated with the percentage agreement (PA) and the Svensson rank-invariant method (12). The PA was used to evaluate the agreement as the proportion of participants who rated the same score or  $\pm 1$  point ( $PA \leq 1$  point) at T1 and T2. The Svensson method was used to evaluate systematic and random disagreement. A systematic disagreement occurs when the ratings are systematically higher or lower at the retest occasion and can be due to a learning effect, change in behaviour or fatigue. Random disagreement is the variability that cannot be explained by the systematic disagreement. The systematic disagreement, according to the Svensson method, is evaluated as the relative position, RP. Possible RP values range from  $-1$  to  $1$  and zero values indicate a lack of systematic disagreement. A positive RP value indicates that the participants had higher scores on T2 than on T1. The random disagreement is evaluated by the relative rank variance, RV. RV ranges from 0 to 1 and a value larger than zero indicates the presence of random disagreement. The higher the RV value the more dispersed are the test–retest ratings. The RP and RV values are calculated together with a 95% confidence interval (95% CI) and a statistically significant disagreement is indicated by a 95% CI not covering zero.

Descriptive data were analysed with the IBM SPSS Statistics version 23 (IBM Corporation, Armonk, NY, USA), the Kappa coefficients with the MedCalc version 15 ([www.medcalc.org](http://www.medcalc.org)), and systematic and random disagreement with the Svensson rank-invariant method ([www.oru.se/esi/svensson](http://www.oru.se/esi/svensson)). Probability values less than 0.05 were considered statistically significant.

## RESULTS

## Participants

Data on the 45 participants (37 men; mean age 65 years and mean time from stroke onset 44 months) are presented in Table I. All participants had mild to

**Table I.** Characteristics of the 45 participants with chronic stroke

| Characteristics                                                   |                 |
|-------------------------------------------------------------------|-----------------|
| Sex (male), <i>n</i> (%)                                          | 37 (82)         |
| Age, years, mean (SD; range)                                      | 65 (7; 44–76)   |
| Type of stroke (ischaemic/haemorrhagic), <i>n</i> (%)             | 32 (71)/13 (29) |
| Months from stroke onset to first test occasion, mean (SD; range) | 44 (28; 10–116) |
| Side of paresis (right), <i>n</i> (%)                             | 25 (56)         |
| Modified Rankin scale level, <i>n</i> (%)                         |                 |
| 1. No significant disability despite symptoms                     | 6 (13)          |
| 2. Slight disability                                              | 26 (58)         |
| 3. Moderate disability                                            | 13 (29)         |
| Impaired strength in UE (SIS domain 1), <i>n</i> (%)              | 42 (93)         |
| Impaired strength in LE (SIS domain 1), <i>n</i> (%)              | 40 (89)         |
| Walking ability > 300 m outdoors, <i>n</i> (%)                    | 45 (100)        |
| Use of walking aids outdoors, <i>n</i> (%)                        | 14 (39)         |
| Hemiplegic shoulder pain, <i>n</i> (%)                            | 9 (20)          |
| Fatigue, <i>n</i> (%)                                             | 25 (56)         |
| Social situation (living alone), <i>n</i> (%)                     | 6 (13)          |
| Vocational situation, <i>n</i> (%)                                |                 |
| Working                                                           | 10 (22)         |
| Not working (unemployed, on sick leave)                           | 4 (9)           |
| Retired                                                           | 31 (69)         |

SD: standard deviation; UE: upper extremity; LE: lower extremity; SIS: Stroke Impact Scale.

**Table II.** Scores of SIS Participation from the 2 test occasions ( $n = 45$ )

| SIS Participation                 | Test occasion 1<br>Median (IQR) | Test occasion 2<br>Median (IQR) |
|-----------------------------------|---------------------------------|---------------------------------|
| Total score                       | 78 (63–86)                      | 75 (63–99)                      |
| Work                              | 4 (3–5)                         | 5 (3–5)                         |
| Social activities                 | 3 (3–5)                         | 4 (3–5)                         |
| Quiet recreation                  | 5 (4–5)                         | 5 (4–5)                         |
| Active recreation                 | 3 (3–5)                         | 3 (3–5)                         |
| Role as a family member           | 4 (3–5)                         | 5 (4–5)                         |
| Spiritual or religious activities | 5 (5–5)                         | 5 (5–5)                         |
| Life control                      | 4 (3–5)                         | 4 (3–5)                         |
| Ability to help others            | 4 (3–5)                         | 4 (3–5)                         |

Response options for each item range from 1 to 5.  
IQR: interquartile range.

moderate disability (modified Rankin Scale 1–3) and were able to walk independently, but approximately one-third of the subjects used a walking aid outdoors. Approximately half of the participants experienced fatigue. A majority lived together with another person and most were retired.

### *Test-retest reliability of Stroke Impact Scale Participation*

Table II presents the medians and quartiles of the SIS Participation total score and item scores of the 2 test occasions. The median for the total score was 78 at T1 and 75 at T2. The medians for the item scores ranged from 3 to 5 at both test occasions. The maximum score of 100 was rated by 4.4% of the participants at T1 and by 15.6% at T2.

The Kappa coefficient of the total score was 0.78 (95% CI 0.74–0.83).

Table III presents the test-retest reliability of the item scores. The PA ranged from 47% to 98% and the PA  $\leq 1$  point from 82% to 100%. According to the Svensson method, most items showed a positive RP, i.e. the participants rated higher scores at T2, but only 2 items "Active recreations" and "Role as a family member" showed a significant systematic disagreement. No random disagreements (RV) were found for any of the items.

**Table III.** Test-retest reliability for the 8 items of the SIS Participation ( $n=45$ )

| SIS Participation                 | PA, $\leq 1$ p, % | PA, $\leq 1$ p, % | Systematic disagreement RP (95% CI) | Random disagreement RV (95% CI) |
|-----------------------------------|-------------------|-------------------|-------------------------------------|---------------------------------|
| Work                              | 62                | 82                | 0.05 (–0.08–0.18)                   | 0.07 (0.00–0.14)                |
| Social activities                 | 47                | 82                | 0.15 (–0.01–0.31)                   | 0.10 (0.00–0.20)                |
| Quiet recreation                  | 64                | 93                | –0.03 (–0.16–0.09)                  | 0.05 (0.00–0.11)                |
| Active recreation                 | 53                | 84                | 0.24 (0.09–0.38)                    | 0.08 (0.00–0.17)                |
| Role as a family member           | 69                | 84                | 0.17 (0.05–0.28)                    | 0.02 (0.00–0.05)                |
| Spiritual or religious activities | 98                | 100               | 0.02 (–0.02–0.06)                   | 0.00 (0.00–0.00)                |
| Life control                      | 62                | 87                | 0.08 (–0.05–0.21)                   | 0.08 (0.00–0.16)                |
| Ability to help others            | 53                | 89                | –0.01 (–0.01–0.15)                  | 0.11 (0.00–0.23)                |

CI: confidence interval; PA: percentage agreement; PA  $\leq 1$  p: percentage agreement  $\leq 1$  points difference; RP: relative position (systematic disagreement); RV: relative rank variance (random disagreement).

## DISCUSSION

This study shows that the Participation domain of the SIS is reliable; the total score had an excellent agreement according to the Kappa coefficient and the item scores had high PA  $\leq 1$  point ( $>82\%$ ). No items, except 2, showed systematic disagreement and no items showed random disagreement according to the Svensson method.

The Kappa coefficient of 0.78 of the total scores of SIS Participation was excellent. This result is in agreement with previous studies reporting intraclass correlation coefficients (ICC) of 0.79–0.87 (6, 7). However, as these studies (6, 7) used parametric statistics to evaluate the total score for SIS Participation it is difficult to fully compare their results with those of the current study.

Reliability analyses of the item scores showed that the PA ranged from 47% to 69% for all items, except for the item "Spiritual or religious activities". This item had an agreement of 98%, but as most participants did not engage in religious activities they rated the item as "limited none of the time". The PA  $\leq 1$  point showed high agreement ( $>82\%$ ) for all items, indicating that, in most cases, the rating differed at most by 1 point. In the analyses of disagreements according to the Svensson rank-invariant method, most items showed a positive RP, indicating that the participants rated higher scores at T2. Only 2 items, "Active recreations" and "Role as a family member", showed significant systematic disagreements and no items showed any random disagreements. It is difficult to compare these results with previous reliability studies, as they did not evaluate systematic or random disagreements according to the Svensson method. However, the findings indicate that the SIS Participation domain is reliable even if some variability in ratings between 2 test occasions occurs.

Several factors might have influenced the participants' ratings of SIS Participation. Firstly, the term "participation" could be interpreted differently by different people. According to the ICF, participation is defined as involvement in life situations (2). However, that basically includes all aspects of being a human and there is no clear distinction between the ICF components participation and activity. It has therefore been suggested that the definition of participation should be limited to social participation, i.e. societal involvement and social roles (13). Secondly, SIS was developed in the USA by input from patients, caregivers and stroke experts (4), but it has no theoretical underpinning. Cultural differences in a person's activity repertoire might influence the ratings and, in fact, few participants in our study engaged in spiritual or religious activities. Thirdly, the SIS participation domain is constructed so that less restriction in participation is rated as higher scores, which could have been confusing for partici-



pants. It can also be difficult to distinguish between response option 4 (limited a little of the time) and 3 (limited some of the time), which might have resulted in the varied scoring at T1 and T2 in the present study. Moreover, most participants used the higher scores (i.e. 3–5) of the SIS Participation scale. This might be expected, as they all had a mild to moderate disability and were in a chronic phase (i.e. at least 6 months post-stroke), and had probably made adaptations in their lives. The somewhat higher ratings at T2 could have been due to a behavioural effect as the participants were more familiar and confident with the questions and the scale at the second test occasion. One way to reduce the systematic variability of the SIS Participation domain might be to perform more than 1 baseline assessment.

Taken together, as participation is not distinctly defined it might be difficult to measure. The broad aspects of the items of SIS Participation domain and the configuration of the scale might have caused variations in the participants' ratings. Further refinement of the scale is therefore suggested in order to reduce variability. It would be interesting to develop an outcome measure of participation by using a bottom-up method (14) based on the narrower construct of social participation.

A strength of the current study is that all participants were in a stable phase post-stroke and the procedure was conducted according to the COSMIN standards (9). Care was taken to standardize the test situation, and the SIS Participation domain was rated at the same location, at the same time of the day, and with the same time interval between test occasions. Furthermore, 45 participants were included, which can be considered sufficiently large when reliability of an outcome measure is evaluated (9, 15). Persons with major cognitive impairments or difficulties communicating were excluded, and more men than women volunteered to participate. Therefore, the results cannot be generalized to the entire stroke population.

In conclusion, this study has shown that the SIS Participation domain is reliable in persons with chronic stroke and mild to moderate disability and can be used to assess perceived participation in this population.

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

## REFERENCES

1. World Health Organization. Stroke/cerebrovascular accident. [Accessed 2018 Mar 26]. Available from: [http://www.who.int/topics/cerebrovascular\\_accident/en](http://www.who.int/topics/cerebrovascular_accident/en).
2. World Health Organization. International Classification of Functioning, Disability and Health 2001. [Accessed 2018 Mar 26]. Available from: <http://www.who.int/classifications/icf/en>.
3. Perenboom RJ, Chorus AM. Measuring participation according to the International Classification of Functioning, Disability and Health (ICF). *Disabil Rehabil* 2003; 25: 577–587.
4. Duncan PW, Bode RK, Min Lai S, Perera S. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Arch Phys Med Rehabil* 2003; 84: 950–963.
5. Tse T, Douglas J, Lentin P, Carey L. Measuring participation after stroke: a review of frequently used tools. *Arch Phys Med Rehabil* 2013; 94: 177–192.
6. Vellone E, Savini S, Fida R, Dickson VV, Melkus GD, Carod-Artal FJ, et al. Psychometric evaluation of the Stroke Impact Scale 3.0. *J Cardiovasc Nurs* 2015; 30: 229–241.
7. Carod-Artal FJ, Coral LF, Trizotto DS, Moreira CM. The stroke impact scale 3.0: evaluation of acceptability, reliability, and validity of the Brazilian version. *Stroke* 2008; 39: 2477–2484.
8. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988; 19: 604–607.
9. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010; 19: 539–549.
10. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull* 1968; 70: 213.
11. Fleiss JL. Statistical methods for rates and proportions. 2nd edn. New York: John Wiley; 1981.
12. Svensson E. Different ranking approaches defining association and agreement measures of paired ordinal data. *Stat Med* 2012; 31: 3104–3117.
13. Piskur B. Social participation: redesign of education, research, and practice in occupational therapy. *Scand J Occup Ther* 2014; 21: 89–95.
14. Wolfe EW, Smith EV. Instrument development tools and activities for measure validation using Rasch models: part II-validation activities. *J Appl Meas* 2007; 8: 204–234.
15. Hobart JC, Cano SJ, Warner TT, Thompson AJ. What sample sizes for reliability and validity studies in neurology? *J Neurol* 2012; 259: 2681–2694.