Test–Retest Reliability and Practice Effect of the Oral-format Symbol Digit Modalities Test in Patients with Stroke

Chia-Lin Koh1, Wen-Shian Lu2,3, Hui-Chun Chen4, I-Ping Hsueh1,5,*, Jia-Ji Hsieh6, Ching-Lin Hsieh1,5

1School of Occupational Therapy, College of Medicine, National Taiwan University, Taipei, Taiwan
2School of Occupational Therapy, Chung Shan Medical University, Taichung, Taiwan
3Occupational Therapy Room, Chung Shan Medical University Hospital, Taichung, Taiwan
4Department of Kinesiology-Therapeutic Science Track, University of Wisconsin-Madison, Madison, WI, USA
5Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei, Taiwan
6Taipei City Hospital Ren-Ai Branch, Taipei, Taiwan

*Corresponding author at: School of Occupational Therapy, College of Medicine, National Taiwan University, Taipei 100, Taiwan.
Tel.: +886-2-33668174; fax: +886-2-23511331.
E-mail address: iping@ntu.edu.tw (I-P. Hsueh).

Accepted 3 April 2011

Abstract

The Symbol Digit Modalities Test (SDMT) is commonly used to evaluate an individual’s switching attention and processing speed. However, its test–retest reliability and practice effect are not well known in patients with stroke, limiting its utility in both clinical and research settings. The present study examined the two aforementioned psychometric properties of the oral-format SDMT on a group of 30 outpatients with stroke. The oral-format SDMT demonstrated excellent test–retest reliability (ICC = 0.89) and a small practice effect (Cohen’s d = 0.26) within a 1-week interval. A practice effect-corrected reliable change index [-5.29, 10.89] was also provided to help clinicians and researchers interpret their clients’ test results. Patients’ characteristics and the test–retest interval should be considered before applying the findings of the present study to clinical settings.

Keywords: Attention; Cerebrovascular disease/accident and stroke; Practice effects/reliable change

Introduction

Switching attention and information processing speed are crucial cognitive functions for stroke patients to live safely and independently (McDowd, Filion, Pohl, Richards, & Stiers, 2003). Research has shown that people who have suffered from stroke have reduced switching attention ability and longer information processing time than do healthy adults (McDowd et al., 2003). Impairments in these attention abilities are related to poor outcomes following stroke in the areas of physical function, memory, and social participation (McDowd et al., 2003). Moreover, attention is a key cognitive component in learning new motor skills (McDowd et al., 2003; Robertson, Ridgeway, Greenfield, & Parr, 1997). Early deficit of attention has been shown to interfere with the efficacy of interventions in stroke patients (Michel & Mateer, 2006). Thus, evaluating switching attention and information processing speed is important for stroke rehabilitation.

The Symbol Digit Modalities Test (SDMT) has become a widely used measure of switching attention and information processing speed in healthy adults (Rabin, Barr, & Burton, 2005; Smith, 1982) and people with neurological impairments, such as multiple sclerosis (Huijbregts, Kalkers, de Sonneville, de Groot, & Polman, 2006; Sheridan et al., 2006). The SDMT has several advantages that contribute to its widespread use. First, it is “inexpensive” and easy to administer in clinical settings. Second, its content is brief, so the assessment time is short. The SDMT has two formats, written and oral. In the present study, the oral format of the SDMT is considered to be more usable than the written format in subacute to chronic stroke patients. In such patients, the prevalence of aphasia is lower; in addition, most patients have recovered from aphasia within
3 months (Desmond, Moroney, Sano, & Stern, 1996), whereas up to 85% stroke patients remain impaired in motor function at 3 months post-event (Hendricks, van Limbeek, Geurts, & Zwarts, 2002). Furthermore, it is difficult to compare the performance of patients with the dominant hand affected and unaffected on the same basis. Writing with an affected dominant hand or with an unaffected non-dominant hand can confound a patient’s performance on a test. In accordance with the recommendation of a panel on neuropsychological assessment of multiple sclerosis patients (Benedict et al., 2002), the present study focused on examining only the oral-format SDMT.

An instrument has to be stable for repeated use over time to obtain meaningful and explainable results. Based on classical test theory, changes in score between test sessions can be partitioned into two components: (i) “true” change in the individual’s condition and (ii) measurement error (Portney & Watkins, 2009). Here, “true” means “the average score that would be obtained if the scale were given an infinite number of times” (Mokkink et al., 2010, p. 742). An individual’s performance on attention tests can be confounded by many factors, such as depression, sensory/motor impairment, and medication (Benedict et al., 2002; Salthouse, 2010). Controlling those related factors can help researchers and clinicians explain testing results.

On the other hand, “measurement error” is defined as “the systematic and random error of a patient’s score that is not attributed to true changes in the construct to be measured” (Mokkink et al., 2010, p. 743). Systematic errors, also called systematic bias, are predictable errors of measurement (Portney & Watkins, 2009). Examples of systematic bias are differences in judgments between raters (e.g., one rater consistently gives higher scores than another rater), the rater’s memory effect, and the practice effect. Random errors refer to factors that affect a patient’s score in an unpredictable way from trial to trial, such as fatigue or simple mistakes (Portney & Watkins, 2009). In order to understand how much of a measured change in score is attributable to measurement error and how much represents a true condition, examining the reliability of the measure is essential.

Reliability is defined as “the proportion of the total variance in the measurements which is because of ‘true’ differences among patients” (Mokkink et al., 2010, p. 743). Within this concept, test–retest reliability represents the extent to which a measure is consistent and free from error over time (Portney & Watkins, 2009). Without understanding of the test–retest reliability of a measure, it is hard to explain whether the observed score changes across serial evaluations are due to a change in the individual’s condition or to measurement error. Previous studies have shown that the SDMT has acceptable test–retest reliability ($r = 0.70–0.80$) in healthy adults measured at 1-month (Smith, 1982) or 2-week intervals (Hinton-Bayre & Geffen, 2005). However, the results of those studies are not “applicable to” the stroke population, since the test–retest reliability is a sample-dependent index (Scientific Advisory Committee of the Medical Outcomes Trust, 2002). The test–retest reliability is calculated based on the score variance of a particular sample. If a group of people perform differently on a test, the test–retest reliability of the test is potentially different. In pathological and longitudinal studies, stroke survivors have shown patterns of recovery that differ from those of other neural diseases (Selzer, 2006). To our knowledge, no study has examined the test–retest reliability of the oral-format SDMT (nor the written format) specifically in stroke patients. The manual of the SDMT provides the reliability based on normal adults, not the stroke population. Therefore, it is necessary to examine this psychometric property, particularly in stroke patients, to help interpret the retest results.

When interpreting score changes in neuropsychology tests, the practice effect is an important aspect to be addressed. The practice effect, which occurs when an individual repeatedly performs the same or a similar test, usually results in improvement of the individual’s performance (McCaffrey, Duff, & Westervelt, 2000). For example, mild traumatic brain injury patients have been found to demonstrate statistically significant improvement on several neuropsychological assessments across time (De Monte, Geffen, & Kwapil, 2005). McCaffrey and colleagues (2000) have suggested that distinguishing the practice effect from other factors is important for both clinicians and researchers. In the research setting, it is important to reduce the variance attributable to the practice effect in order to accurately evaluate changes in patients’ performance (McCaffrey et al., 2000). Although the practice effect may provide useful information about the underlying brain function to clinicians, clinicians should also consider the practice effect when interpreting the test results of their clients. That is, any difference in scores needs to be greater than the practice effect induced by the measure for true change to be determined. Knowledge of the practice effect is essential and should be established with various temporal intervals, specific measures, and particular patient populations (McCaffrey et al., 2000). So far, to our knowledge, the practice effect of the SDMT that may occur when it is used with stroke patients has not yet been examined.

Alternatively, to avoid the practice effect, many cognitive tests have been developed with one or more alternative forms. Hinton-Bayre and Geffen (2005) examined the comparability of four alternate forms of the written-format SDMT in 112 male contact-sport athletes. However, their results demonstrated that the use of alternative forms of the SDMT may not be sufficient to resolve the practice effect problem at a short retest interval (within 2 weeks). A significant practice effect still existed among different alternative forms (effect size 0.38–1.00). That study suggested that retest scores are still
somehow inaccurate, even when an alternative form has been used. We believe that estimating the practice effect of a test quantitatively would be more practical and helpful for clinicians and researchers in interpreting test results.

Therefore, the aim of the present study was to investigate the test–retest reliability and the practice effect of the SDMT in stroke patients. The results of the study should help researchers and clinicians decide whether the oral-format SDMT provides stable results across 1 week in patients who are at least 3 months post-event. In addition, the amount of the practice effect is also estimated in this study to help researchers and clinicians interpret the retest results.

Method

Participants

A group of outpatients with stroke was recruited from three rehabilitation units in northern Taiwan between November 2007 and May 2008. Stroke patients were recruited if they met the following criteria: (i) over 18 years of age; (ii) diagnosed as either ischemic stroke or cerebral hemorrhage; (iii) the onset of stroke was at least 3 months before the first evaluation to minimize the influence of spontaneous neurological recovery (Wade, Wood, & Hewer, 1988); (iv) medically stable without any other major diseases (e.g., schizophrenia, depression, or dementia) that influence cognitive function; (v) do not exhibit obvious cognitive impairment, as assessed with the Mini-Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975), and able to follow commands to perform the tests; (vi) free from visual spatial neglect, as assessed with the Behavioral Inattention Test (BIT; Wilson, Cockburn, & Halligan, 1987a); and (vii) able to see the stimuli on the testing sheet or materials. The participants were excluded if they had recurrent stroke or were unable to give oral responses to the oral-format SDMT due to severe aphasia or tracheostomy. The participants were required to provide informed consent at the first evaluation session. The study was approved by the ethical committees of the hospitals. All participants’ demographic and stroke-related information was obtained from medical records.

Procedure

All eligible participants were assessed using the oral-format SDMT by a well-trained research assistant in two sessions of evaluation. The examiner had undergone 8 h of training, including reading the manual, going through the instructions, and 10 trials of SDMT practice on stroke patients under the supervision of a therapist who was familiar with the SDMT. The accuracy rate of scoring of the examiner was 100% during and after the training. The time interval between the two sessions was 7 days. The interval for examining test–retest reliability should be close enough to avoid genuine changes in the measured variable (Portney & Watkins, 2009). Since spontaneous recovery achieves a plateau state at 3 months after stroke onset (Wade et al., 1988) and all included participants did not receive any attention training, we believed that limited genuine changes would occur during this short period. All participants were assessed in a quiet room to avoid any distractions that would influence their performance. Each participant’s condition during evaluation sessions was observed and recorded, including arousal, sleep, and emotional status, as well as cooperation status. These variables were judged by the examiner subjectively, using a 3-point Likert-like scale (1–3) wherein 1 represented good condition and 3 represented bad condition.

Instruments

Symbol Digit Modalities Test. The SDMT is a timed assessment (Smith, 1968, 1982). In this study, only the oral format was used. During the test, each participant is required to substitute a number (1–9) orally for a total of 120 geometric figures randomly presented on the sheet. Each number is paired with a particular geometric symbol given at the top of the assessment sheet. Testing time is limited to 90 s, and the number of correct verbal responses is recorded. A larger number of correct answers within the time limit represent better switching attention and information processing speeds of an individual.

Mini-Mental Status Examination. The MMSE is a screening test for general cognitive impairment. It consists of five domains of cognition: orientation, attention, memory, language, and construction skills. The MMSE is used to exclude participants with obvious cognitive impairment post-stroke. The cutoff point for stroke patients is 23/24 (Grace et al., 1995). People who obtain ≥24 are considered to have adequate cognitive function. The reliability and validity of the MMSE are acceptable in stroke populations (Agrell & Dehlin, 2000).

Behavioral Inattention Test. The BIT is a standardized battery for assessing unilateral visual neglect. The BIT consists of nine behavioral subtests and six conventional subtests, including line crossing, letter cancellation, star cancellation, figure and shape
copying, line bisection, and representational drawing. To rule out participants with unilateral visual neglect, five of the six conventional subtests were used. The letter cancellation test was not used because many elderly people in Taiwan cannot read the English alphabet. Any participant scoring below the cutoff score for one or more of the five subtests was excluded from this study (Wilson et al., 1987a). The BIT has excellent inter-rater and test–retest reliability, good concurrent validity, and moderate predictive validity (Wilson, Cockburn, & Halligan, 1987b).

Data Analysis

The Spearman correlation coefficient was analyzed to investigate the correlation between age and the change in scores, as well as the education level and the change in scores. The test–retest reliability of the SDMT was determined through the calculation of the Intra-class correlation coefficients (ICCs) based on the two-way random effects of analysis of variance. ICC values <0.39 indicate poor reliability; 0.40–0.59, moderate reliability; 0.60–0.79, good reliability; and 0.80–1.00, excellent reliability (Bushnell, Johnston, & Goldstein, 2001; Fleiss, 1986). The agreements of the two test results for each participant were plotted in a Bland–Altman plot (Bland & Altman, 1986). The 95% limit of agreement (LOA) was also illustrated to present the distribution of difference.

The practice effect between test and retest assessments was examined using one-tail paired t-test. The size of the practice effect was calculated and determined by Cohen’s criteria (Cohen, 1988): 0.21–0.49 indicating small; 0.50–0.79, medium; and over 0.80, large effect size.

A reliable change index modified for practice (RCIp) was calculated to provide criteria for meaningful change based on the calculations of measurement error for each score (Raymond, Hinton-Bayre, Radel, Ray, & Marsh, 2006). The 90% confidence interval (90% CI) RCIp was calculated using the following formulae:

\[
90\% \text{ CI RCIp} = \text{mean practice effect} \pm 1.645 \times \text{SE_{diff}}
\]

\[
\text{SE_{diff}} = \sqrt{2\left(\text{SE}_{m}\right)^2}
\]

\[
\text{SE}_{m} = SD_1 \left[ \sqrt{1 - \gamma_{xx}} \right]
\]

where mean practice effect is the mean of the difference between the first and second test scores; SE_{diff} the standard error of the differences; SE_{m} the standard measurement error; SD_1 the standard deviation of the first test scores; and \( \gamma_{xx} \) the ICC value of test–retest reliability. 90% CI were used for detecting the significant change, indicating that 5% of patients may be expected to fall above and 5% of patients below, the cutoff due to chance rather than true change beyond the practice effect.

Results

Participants

A total number of 158 patients who were confirmed to have stroke during the study period were approached. One hundred and five stroke patients were excluded due to the following reasons: 46 were excluded for unwillingness to participate; 22 for inability to follow instructions for assessments or for having MMSE scores <24; 18 for inability to give oral responses due to severe aphasia or tracheostomy based on medical records; 8 for having comorbidities (one each for Down syndrome, schizophrenia, and depression; and 5 for fever); 6 for being <3 months post-event when recruited; and 5 for having visual neglect. No patient was found to have recurrent stroke.

A total of 33 (62%) of the 53 eligible participants were recruited and completed all evaluation sessions. Twenty patients who were not able to complete the second test within the 1-week interval were excluded. Another three participants, who were found to have mild aphasia, corrected their answers by gestures in the first assessment session but could make the corrections orally in the second session. Therefore, they were excluded from further analysis due to the inconsistency in mode of assessment of the SDMT. The mean age of the 30 remaining participants was 59.2 years; male (83.3%) and ischemic stroke (56.7%) represented the majority of the sample. The demographic, stroke-related, and clinical information of the included participants are provided in Table 1.
In general, the participants had good arousal, sleeping condition, emotional status, and cooperation at the first test session (mean scores ranging from 1.0 to 1.3 and SD ranging from 0.0 to 0.6) and at the second test session (mean scores ranging from 1.0 to 1.4 and SD ranging from 0.0 to 0.7). No statistically significant difference was found for any of these factors between the two test sessions. No significant correlation was found between age and score change ($r = 2.20$, $p = .28$). The education level and the change in score was also found not to be correlated ($r = 2.05$, $p = .78$). Two subgroups, those who improved ($n = 20$) and those that did not ($n = 10$), were further analyzed using the $t$-test. No significant differences of age and education level were found between the two subgroups. The average numbers of errors made in the first and the second tests of the participants were 1.5 (SD = 1.7) and 1.3 (SD = 2.6), respectively. There was no significant difference in error rates between the two test sessions ($p = .47$).

The Test–Retest Reliability

The mean scores of the two evaluation sessions and the mean change scores between tests are shown in Table 2. The ICC of the SDMT was 0.89 (95% CI = 0.71–0.95), indicating excellent test–retest reliability. The Bland–Altman plots (Fig. 1) showed that the points were distributed randomly. The 95% LOA was −6.8 to 12.4. A total of 20 points (67%) lay above zero, and the mean of the difference scores was 2.8.

Table 1. Demographic, stroke-related, and clinical information of the participants ($n = 30$)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (men/women)</td>
<td>25/5</td>
</tr>
<tr>
<td>Age (years; mean [SD])</td>
<td>59.2 (13.7)</td>
</tr>
<tr>
<td>Education level ($n [%]$)</td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td>2 (6.7)</td>
</tr>
<tr>
<td>Middle school</td>
<td>2 (6.7)</td>
</tr>
<tr>
<td>High school</td>
<td>8 (26.7)</td>
</tr>
<tr>
<td>Vocational school</td>
<td>4 (13.3)</td>
</tr>
<tr>
<td>University</td>
<td>13 (43.3)</td>
</tr>
<tr>
<td>Post-graduate</td>
<td>1 (3.3%)</td>
</tr>
<tr>
<td>Stroke type ($n [%]$)</td>
<td></td>
</tr>
<tr>
<td>Hemorrhagic</td>
<td>13 (43.3)</td>
</tr>
<tr>
<td>Ischemic</td>
<td>17 (56.7)</td>
</tr>
<tr>
<td>Affected side ($n [%]$)</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>11 (36.6)</td>
</tr>
<tr>
<td>Right</td>
<td>18 (60.0)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>1 (3.3)</td>
</tr>
<tr>
<td>Time since onset (month; mean [SD])</td>
<td>24.9 (21.8)</td>
</tr>
<tr>
<td>Mini-Mental State Examination (mean [SD])</td>
<td>28.1 (1.6)</td>
</tr>
<tr>
<td>Behavioral Inattention Test ($a$ (mean [SD])</td>
<td></td>
</tr>
<tr>
<td>Line Crossing (0–36)</td>
<td>35.6 (1.2)</td>
</tr>
<tr>
<td>Star Cancellation (0–54)</td>
<td>52.9 (2.4)</td>
</tr>
<tr>
<td>Figure and Shape Copy (0–4)</td>
<td>3.9 (0.3)</td>
</tr>
<tr>
<td>Line Bisection (0–9)</td>
<td>8.8 (0.6)</td>
</tr>
<tr>
<td>Representational drawing (0–3)</td>
<td>3.0 (0.0)</td>
</tr>
</tbody>
</table>

Note: SD = standard deviation.

$a$ The bracket indicated the total scores of the sub-test.

Table 2. Parameters of the test–retest reliability of the oral-format Symbol Digit Modalities Test ($n = 30$)

<table>
<thead>
<tr>
<th>1st testing (mean [SD])</th>
<th>2nd testing (mean [SD])</th>
<th>Change scores (mean [SD])</th>
<th>Agreement (ICC)</th>
<th>$SE_{in}$</th>
<th>$SE_{diff}$</th>
<th>$RCI_{p-}$</th>
<th>$RCI_{p+}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.5 (10.5)</td>
<td>37.2 (12.4)</td>
<td>2.8 (4.9)</td>
<td>0.89</td>
<td>3.48</td>
<td>4.92</td>
<td>−5.29</td>
<td>10.89</td>
</tr>
</tbody>
</table>

Notes: SD = standard deviation; ICC = intra-class correlation coefficient; $SE_{in}$ = standard measurement error; $SE_{diff}$ = standard error of the differences; $RCI_{p}$ = reliable change index modified for practice (minus for the lower border and plus for the upper border).
The Practice Effect

The SDMT demonstrated a statistically significant practice effect at retest ($p = .004$). According to Cohen’s effect size criteria, there was a small practice effect of the SDMT ($d = 0.26$). Using the reliability coefficient mentioned above ($ICC = 0.89$), the 90% CI RCIp of the SDMT was $[-5.29, 10.89]$.

Discussion

This study examined the test–retest reliability and highlighted the practice effect of the oral-format SDMT in patients who were at least 3 months post-event. In addition, the error rate of the participants was provided. The small error rate could serve as reference to help clinicians and researchers judge their clients’ performance from this aspect related to inattention. The results of the present study show that the SDMT is a reliable measure for repeatedly evaluating stroke patients’ abilities in information processing speed and switching attention. However, a small practice effect was found between the two test sessions with a 1-week interval. Since most confounding factors were controlled and the potential spontaneous or treatment-induced recovery at 3 months is limited, the estimated range of score change was considered to have resulted mainly from the practice effect plus random measurement error. Based on the RCIp $[-5.29, 10.89]$ given in the present study, only score changes that are beyond the estimated range should be considered to reflect a statistically significant improvement or decline on the ability measured.

Excellent test–retest reliability for the oral-format SDMT was found at a 1-week interval in stroke patients. The excellent ICC value implies that the agreement of the two test results was very high. This result was better than previous findings on healthy adults (Hinton-Bayre & Geffen, 2005; Smith, 1982). Therefore, the oral-format SDMT can be considered stable for stroke patients when administered by the same examiner at a 1-week interval.

In addition, we compared the raw scores of the oral-format SDMT obtained by the participants in the present study with age appropriate norms in the original manual of the SDMT (Smith, 1982). The observed scores of our participants were lower than those of the norm to $1.5–2.0 \text{SD}$ (which means moderately low scores to very low scores). The substantially lower scores of the participants who had had a stroke indicated brain dysfunction. This result is similar to the report in the manual that patients with acute stroke obtained lower scores at least $1.5–2.0 \text{SD}$ below the mean of the norm group. Therefore, the oral-format SDMT is considered to be able to detect deficits of attention and information processing and is stable for use over a 1-week interval in subacute to chronic stroke patients.

According to the Bland–Altman plot, participants tended to obtain higher scores on the second test than on the first test (i.e., a positive difference score). In addition, a small practice effect was found between the two test sessions in the present study. Since the spontaneous recovery achieves a plateau state at 3 months after stroke onset (Wade et al., 1988) and all included participants did not receive any attention training, we believed limited genuine changes would occur during this short period. Moreover, the measurement error from the examiner was considered to be very limited. The scoring of the SDMT is based mainly on participants’ performance and does not require judgment by the examiner. In addition, all participants
were assessed by the same well-trained examiner in the present study. The consistency of the examiner in introducing and implementing the test was confirmed by a therapist who was familiar with the SDMT. No scoring errors were made after the training. Thus, the measurement error from the examiner was limited. Therefore, we consider the practice effect plus random measurement error to be the major sources of participants’ score improvements.

The test–retest interval in the present study was 1 week, and a small practice effect of the oral-format SDMT was found within this short period. A small practice effect (effect size = 0.32) of the SDMT within a 1-week interval was also found for community-dwelling older adults with or without cognitive impairment (Duff et al., 2010). The mean change in scores of the community-dwelling elderly people was 1.7 (Duff et al., 2010). The lesser improvement of their study when compared with our results may have occurred because nearly half (42%) of the participants in that study were classified with amnestic mild cognitive impairment, and thus had less learning ability than healthy adults. None of the participants in the present study exhibited obvious cognitive impairment, although their brain function for switching attention and information processing was not as good as that of healthy adults. As Duff and colleagues (2010) demonstrated in their study, it would be worthwhile to further study the impact of cognitive impairment severity on the short-term practice effect and the prediction of the practice effect for future health-related function in stroke patients.

Although the estimated RCIp provided in the present study was obtained from stroke patients who did not exhibit obvious cognitive impairment, this RCIp may also be helpful in the interpretation of test results for patients who do have cognitive impairment. People with cognitive impairment have less of a practice effect on performing cognitive assessments than people without cognitive impairment (Calero & Navarro, 2004). Therefore, despite the substantial value of the practice effect, the RCIp index can serve as a conservative criterion to explain score changes in stroke patients who do have cognitive impairment.

To help researchers and clinicians interpret the results of the oral-format SDMT, we calculated the 90% CI RCIp. The RCIp index in the present study was found to be similar to the RCIp in a study of healthy adults, who had a follow-up interval of between 4 and 24 months (Levine, Miller, Becker, Selnes, & Cohen, 2004). Thus, subacute to chronic stroke patients who do not exhibit obvious cognitive impairment may have a practice effect comparable with that of healthy adults. This result emphasizes the necessity of considering the amount of practice effect when interpreting changes of score in stroke patients. However, it is noted that the test–retest interval is longer in Levine and colleagues’ (2004) study than in the present study. The practice effect might be even larger for healthy adults if they were retested within 1 week. Given the same test–retest interval, whether a difference of practice effect exists between healthy adults and stroke patients needs further studies.

Several limitations of the present study should be noted. First, our results might not be generalized to the whole stroke population. Although we recruited a heterogeneous sample of stroke survivors, they were free of severe cognitive impairment, severe aphasia, or visual neglect. The RCIp index provided might not be so useful on patients with the aforementioned deficits. In addition, the sample size was small. Moreover, we examined only the oral-format SDMT in the present study, since the written format is not considered suitable for stroke patients with dominant hand impairment. However, further research may be worthwhile to confirm the test–retest reliability of the written-format SDMT, which can be used for patients with aphasia. Furthermore, we did not measure the incidental memory of the participants. Future research could focus on investigating incidental memory in stroke patients on the SDMT, as well whether incidental memory may influence the practice effect of the SDMT. Finally, numbers may not take the same amount of time to say in different languages. Further cross-cultural validations of our findings are warranted before application of our findings in different countries.

Conclusion

The oral-format SDMT was found to have excellent test–retest reliability in stroke patients who were at least 3 months post-event. However, a small practice effect exists between two test sessions at a 1-week interval. Therefore, an improvement of scores of an individual on the second test may be influenced by the practice effect (i.e., familiarity with the procedure or test items) instead of a true change in condition. Considering the practice effect, the users of the oral-format SDMT are advised to interpret their participants’ results according to the RCIp value given in the present study. Only those people whose scores on the oral-format SDMT improve by 11 or more points over a 1-week interval can be considered to have a statistically significant improvement in switching attention and processing speed. On the other hand, a decrease in a second test score of 6 or more points should be considered to indicate an adverse event.

Funding

This study was supported by research grants from the National Science Council (NSC96-2628-B-002-034-MY3) and the National Health Research Institute (NHRI-EX99-9512PI).
Conflict of Interest

None declared.

References

Agrell, B., & Dehlin, O. (2000). Mini Mental State Examination in geriatric stroke patients. Validity, differences between subgroups of patients, and relationships to somatic and mental variables. Aging (Milan, Italy), 12 (6), 439–444.


