Journal of Current Science and Technology, July - September 2025 Vol. 15 No. 3, Article 127
ISSN 2630-0656 (Online)
Copyright ©2018-2025, Rangsit University

Cite this article: Chinsongkram, B., Churdchomjan, W., Pleumpitiwiriyawej, S., Hongthong, S., Kaewcum, N., Kongwattanakul, K., …., Nithisup, P. (2025). Reliability and validity of the modified O'Sullivan functional balance (mOFB) test in individuals with stroke. *Journal of Current Science and Technology*, *15*(3), Article 127. https://doi.org/10.59796/jcst.V15N3.2025.127

Journal of Current Science and Technology

Journal homepage: https://jcst.rsu.ac.th

# Reliability and Validity of the Modified O'Sullivan Functional Balance (mOFB) Test in Individuals with Stroke

Butsara Chinsongkram[1,*], Worachat Churdchomjan[1], Sutisa Pleumpitiwiriyawej[1], Somchanok Hongthong[1], Nattakarn Kaewcum[1], Kalaya Kongwattanakul[1], Duangruedee Dissanguan[1], Kanyarat Kiadtiwanit[2], and Peeraporn Nithisup[2]

[1]Faculty of Physical Therapy and Sports Medicine, Rangsit University, Pathum Thani 12000, Thailand
[2]Department of Physical Therapy, Faculty of Allied Health Science, Nakhon Ratchasima College, Nakhon Ratchasima 30000, Thailand

*Corresponding author; E-mail: butsara.c@rsu.ac.th

Received 2 May 2025; Revised 19 May 2025; Accepted 27 May 2025; Published online 15 June 2025

## Abstract

The modified O'Sullivan Functional Balance (mOFB) test is a refined version of the original OFB, developed to assess balance in sitting and standing positions in individuals with stroke using standardized perturbations and updated scoring criteria. Despite its clinical utility, the mOFB lacked standardized administration procedures and had not been validated in individuals with stroke. This study aimed to examine the intra-rater and inter-rater reliability, as well as the convergent validity of the mOFB following the development of standardized instructions and scoring criteria. The mOFB comprises four tasks that assess static and dynamic balance in both sitting and standing positions, each rated on a 5-point ordinal scale (where 0 indicates an inability to maintain balance and 4 indicates normal balance). Seventy-five individuals with a first-time stroke (aged 25–84 years) participated. For inter-rater reliability, assessment sessions were video recorded and scored independently by four raters. Intra-rater reliability was assessed by having the same rater evaluate the recordings twice. Convergent validity was examined using the Berg Balance Scale (BBS) and the Trunk Impairment Scale (TIS). Statistical analyses included intraclass correlations coefficients (ICCs) and Spearman's rank correlation. The mOFB demonstrated excellent intra-rater (ICC = 0.97; 95% CI: 0.96–0.98) and inter-rater reliability (ICC = 0.91; 95% CI: 0.85–0.95). Strong correlations with the BBS ($r = 0.82$, $p < 0.001$) and moderate correlations with the TIS ($r = 0.60$, $p < 0.001$) supported its convergent validity. No floor or ceiling effects were observed in the total scores. These findings support the mOFB as a reliable and valid tool for assessing balance in individuals with stroke. Its simplicity, brief administration time, and appropriate difficulty across stroke stages make it suitable as a clinical screening tool in high-volume clinical settings.

***Keywords***: *modified O'Sullivan functional balance (mOFB); Reliability; Validity; Stroke; Balance assessment*

## 1. Introduction

Stroke is a major global health concern, constituting the second leading cause of mortality and the third most prevalent cause of long-term disability worldwide (Katan, & Luft, 2018). Balance deficits are among the most common impairments following stroke and have a substantial impact on functional mobility, fall risk, and quality of life (Park, & Kim, 2019; Vincent-Onabajo et al., 2018; Xu et al., 2018). These deficits are primarily due to disruptions in postural control mechanisms, including muscle weakness, impaired trunk control, and abnormal movement patterns, most evident during the early stages of recovery. Therefore, balance rehabilitation is essential and should be prioritized early in the recovery process to ensure sufficient stability for initiating movement training and engaging in functional activities. Evidence suggests that early and targeted balance training can significantly improve balance control, facilitate mobility and functional performance in daily life, and enhance long-term quality of life in individuals recovering from stroke (Huh et al., 2015; Cabrera-Martos et al., 2020).

Accurate assessment of balance is essential for guiding rehabilitation and predicting outcomes. Over the years, numerous balance assessment tools have been developed and validated for individuals with stroke. It can be broadly categorized into laboratory-based assessments and clinically-based assessments. Laboratory-based assessments typically utilize advanced technologies such as force platforms, motion capture systems, accelerometers, and computerized posturography (Mancini, & Horak, 2010; Banyam, & Rakpongsiri, 2024; Rakpongsiri et al., 2023). While these assessment tools demonstrate high accuracy and reliability, their cost and complexity often limit their applicability in routine clinical settings. In contrast, low-cost, clinically based assessments such as the Berg Balance Scale (BBS) and the Balance Evaluation Systems Test (BESTest) are more accessible and have demonstrated good reliability and validity in individuals with stroke (Chinsongkram et al., 2014). However, the lengthy administration time of these assessments reduces their practicality in high-demand clinical environments, with a high patient-to-provider ratio.

The O'Sullivan Functional Balance (OFB) grade is a clinical tool designed to evaluate static and dynamic balance in individuals with neurological impairments, such as stroke or spinal cord injury. This test assesses a patient’s ability to maintain postural stability under both static and dynamic conditions in various positions, such as sitting and standing. It consists of four main components: static sitting balance, dynamic sitting balance, static standing balance, and dynamic standing balance. Each component is scored based on the patient's ability to perform specific tasks, such as turning the head or trunk, reaching for objects, and shifting body weight, with a total possible score of 16 points (O’Sullivan, & Schmitz, 2007; Shumway-Cook, & Woollacott, 2017). However, this assessment lacks a standardized evaluation protocol, leading to variations in its administration. In Thailand, the OFB assessment is widely used as a clinical tool for evaluating balance due to its ease of use, brief administration time, and lack of associated costs (Chinsongkram et al., 2020). However, it has been adapted and applied using a variety of assessment methods and scoring approaches, such as the addition of external resistance to disrupt static balance, the use of reach tasks to perturb dynamic balance, and the application of both original and modified grading criteria (Chinsongkram et al., 2020). Previous studies have found that variability in assessment methods and scoring approaches used in the modified O'Sullivan Functional Balance (mOFB) version has resulted in low to moderate both inter-rater and intra-rater reliability. These findings highlight the need to develop clear and standardized test instructions for the mOFB, as well as to establish the reliability and validity of the modified version (Chinsongkram et al., 2020).

Therefore, this study was conducted to develop standardized test instructions, clearer grading criteria, and refined test items to address inconsistent administration and interpretation based on the findings from a previous study (Chinsongkram et al., 2020), and examine the reliability and validity of the modified version in individuals with strokes. The aim of these modifications was to create a balanced assessment tool that clinicians could use confidently, that could be widely adopted for individuals with stroke, and that would be recognized and accepted as standard practice in rehabilitation. This foundational work laid the groundwork for further validation of the mOFB as a practical and reliable tool for assessing balance in individuals with stroke.

## 2. Objectives

This study aimed to (1) evaluate the intra-rater and inter-rater reliability of the mOFB test following the implementation of standardized test instructions and clearly defined scoring criteria, and (2) investigate the convergent validity of the mOFB test by comparing it with common clinical balance assessments in individuals with stroke.

## 3. Materials and Methods

This study was an observational investigation of the reliability and validity of the mOFB in individuals with stroke.

### 3.1 Participants

The participants were individuals with subacute or chronic stroke, referred to physical therapy services at Bueng Yitho Medical and Rehabilitation Center and the Thai Red Cross Rehabilitation Center in Thailand between October 1, 2018, and March 31, 2019. A six-month period following stroke onset was used as the cutoff to distinguish between the subacute and chronic phases (Bernhardt et al., 2017). Participants were eligible for inclusion if they met the following criteria: a diagnosis of unilateral ischemic or hemorrhagic stroke with a stable medical condition, were aged between 18 and 90 years, and had the ability to follow instructions required for assessment. Exclusion criteria included cognitive impairment indicated by a Mini-Mental State Examination (MMSE) score below 24 (Bour et al., 2010), brainstem or cerebellar lesions, presence of a neurological disorder other than stroke, or significant peripheral neuropathy or musculoskeletal issues that could interfere with balance. All participants provided written informed consent, and the study was approved by the Research Ethics Committee of Rangsit University, Thailand (Certificate Number: RSUERB2017-33).

### 3.2 Outcome Measures

Five balance evaluation scales were administered to each participant, including the O'Sullivan Functional Balance (OFB) test, the modified O'Sullivan Functional Balance (mOFB) test, the Berg Balance Scale (BBS), the sitting balance sub-item of the Motor Assessment Scale (MAS), and the Trunk Impairment Scale (TIS) 2.0.

#### *3.2.1 The O'Sullivan Function Balance Test (OFB)*

The OFB is a scoring tool used to assess both static and dynamic balance in various positions. It emphasizes the individual's ability to maintain postural stability and perform postural adjustments in response to voluntary movements, such as turning the head or trunk, reaching for an object from the floor, and shifting body weight. The OFB test uses a 5-level ordinal grading scale: zero, poor, fair, good, and normal. A score of zero indicates an inability to maintain balance, while a score of "normal" reflects normal balance performance (O'Sullivan, & Schmitz, 2007). This study converted the grading scale into ordinal scores ranging from 0 to 4, where 0 indicates an inability to maintain balance independently, and 4 indicates normal balance and assessed in four components: static sitting balance, dynamic sitting balance, static standing balance, and dynamic standing balance. The OFB has a total score of 16 points, and sub-scores can be assigned for individual test positions. For example, the static sitting balance sub-score ranges from 0 to 4 points. The combined score for sitting balance (static sitting and dynamic sitting sub-scores) totals 8 points.

#### *3.2.2 The modified O'Sullivan Functional Balance (mOFB) Test*

The mOFB is a modified version of the OFB that includes external resistance to challenge static balance and weight shifting to challenge dynamic balance, along with adjusted grading criteria (Chinsongkram et al., 2020). A previous study found this modified test had low to moderate reliability due to a lack of clear testing instructions and grading criteria (Chinsongkram et al., 2020). To address these limitations, we developed standardized test instructions and scoring criteria, as outlined in Supplementary Material and Table 4, prior to evaluating the psychometric properties of the mOFB. The assessment takes approximately 5 minutes to administer and can be performed on a stable bed or chair using only a stopwatch. Test condition consists of static sitting balance, dynamic sitting balance, static standing balance, and dynamic standing balance. Scoring is based on observation of the patient's ability to maintain balance while quietly sitting or standing for 60 seconds, respond to external perturbations, and reach beyond arm's length in various directions. Scoring details are provided in Table 4.

#### *3.2.3 The Berg Balance Scale (BBS)*

The BBS is commonly used and currently considered a reference standard for assessing balance in individuals with stroke. The BBS consists of 14 functional balance tasks that focus on the ability to maintain a position and postural adjustments to voluntary movements (Berg et al., 1992). It is simple to administer and requires minimal equipment and time to complete. Each item is scored from 0 to 4, where 0 indicates the inability to perform the task and 4 indicates optimal performance (Blum, & Korner-Bitensky, 2008). The maximum score is 56, with scores below 45 indicating a high risk of falling. The BBS is considered a reliable and valid tool for assessing functional balance in individuals with stroke. Internal consistency was excellent (Cronbach alpha=0.92-0.98) as was inter-rater reliability (ICC=0.95-0.98), intra-rater reliability (ICC=0.97), and test-retest reliability (ICC= 0.98). The BBS shows excellent correlations with the Barthel Index, the Postural Assessment Scale for Stroke Patients, the Functional Reach Test, the balance subscale of the Fugl-Meyer Assessment, the Functional Independence Measure, gait speed (Blum, & Korner-Bitensky, 2008), and BESTest (Chinsongkram et al., 2014).

*3.2.4 The Motor Assessment Scale (MAS): Balance Sitting Sub-item*

The MAS consists of 8 items to assess areas of motor function. This study used the balance sitting sub-item as the criterion measure of sitting balance. The balance sitting sub-item of the MAS comprises six progressively challenging tasks designed to assess postural control while sitting. The activities include: sitting unsupported, maintaining a static sitting position with feet together for 10 seconds, sitting upright with equal weight distribution on both legs, sitting with feet together and flat on the floor while turning the head to look behind, reaching 10 centimeters forward to touch the floor, and reaching to touch the floor on both the left and right sides while sitting with feet together. This sub-item is scored on a scale from 0 to 6, with each score reflecting the highest level of performance completed correctly according to standardized criteria. The MAS demonstrates excellent test-retest reliability and inter-rater reliability (Carr et al., 1985) and excellent concurrent validity with Fugl-Meyer (FMA) total scores in individuals with stroke (Malouin et al., 1994).

*3.2.5 The Trunk Impairment Scale (TIS) 2.0 Version*

The Trunk Impairment Scale (TIS) is designed to assess motor impairment of the trunk following stroke by evaluating static and dynamic sitting balance, as well as trunk coordination (Verheyden et al., 2004). The original TIS evaluates three components: static sitting balance (maximum score = 7), dynamic sitting balance (maximum score = 10), and trunk coordination (maximum score = 6), with a total possible score of 23. In this study, the TIS 2.0 version was used, which includes only the dynamic balance and coordination subscales. The static sitting balance subscale was excluded due to ceiling effects (Verheyden, & Kersten, 2010), resulting in a maximum total score of 16. The TIS demonstrates strong psychometric properties. It shows excellent test–retest reliability (ICC = 0.99) and good inter-rater reliability (ICC = 0.96) (Verheyden et al., 2004). Moreover, it demonstrates acceptable concurrent validity with the Berg Balance Scale (r = 0.72), the Barthel Index (r = 0.86), and the Trunk Control Test (r = 0.83) (Verheyden et al., 2004). A key advantage of the TIS is its minimal equipment requirement, making it highly applicable in clinical settings.

### 3.3 Procedure

*3.3.1 Reliability*

Inter-rater and intra-rater reliability were assessed by four physical therapists. The raters included a convenience sample of one physical therapist from a stroke rehabilitation center with 10 years of stroke rehabilitation experience, along with one neurological physical therapy lecturer and two senior physical therapy students from Rangsit University. All raters completed two training workshops on administering the OFB and mOFB tests. The training included demonstrations of the OFB and mOFB tests and grading in healthy volunteers, followed by detailed discussions of the scoring criteria. Raters also independently scored video-recorded performances and participated in consensus discussions to resolve discrepancies. In cases of disagreement, raters rewatched the videos with the trainer to reach a unified scoring agreement.

Following training, a reliability trial was conducted with 75 individuals with stroke. The OFB and mOFB tests were administered by one rater (BC), and all procedures were video-recorded in a standardized laboratory setting at Bueng Yitho Medical and Rehabilitation Center and Thai Red Cross Rehabilitation Center. Participants were positioned identically for each assessment, and standardized verbal instructions were provided. Vital signs were assessed before testing to ensure medical stability, and rest breaks were allowed as needed. Each rater independently scored the videotaped performances on two separate occasions, spaced seven days apart. Intra-rater reliability was evaluated by comparing scores between the two occasions for each rater. Inter-rater reliability was assessed by comparing scores from the first scoring occasion across all four raters. Raters used separate scoring sheets for each occasion and did not confer with each other at any point during the rating process.

*3.3.2 Validity*

Before initiating the validity study, rater BC received additional training in administering the BBS, MAS (sitting balance sub-item), and TIS. The sample size for the convergent validity analysis was determined using a null correlation coefficient of 0.50 and an expected correlation of 0.80, requiring a minimum of 29 participants per group. A total of 75 participants (29 in the subacute phase and 46 in the chronic phase) were enrolled after providing informed consent. Demographic and clinical data were collected by rater SH. Balance assessments were conducted by rater BC, who was blinded to participants' baseline characteristics and phase of stroke. The order of the balance tests was randomized, progressing from sitting to standing positions. Tasks that overlapped across tests were performed only once and scored according to the respective criteria for each assessment. All evaluations were carried out in the same laboratory environment, with standardized verbal instructions and rest periods provided to minimize fatigue. If assessments could not be completed in a single session, they were continued on the following day. To ensure scoring accuracy, rater BC reviewed all test performances via video recordings.

### 3.4 Statistical Analyses

Data were analyzed using IBM SPSS Statistics version 21.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were used to analyze baseline data, including mean, standard deviation, frequency, maximum score, and minimum score. Histograms and frequency distribution analyses were used to examine the score distributions. If more than 20% of participants obtained the minimum possible total score or a sub-score of grade 0 on the mOFB or OFB, this was interpreted as a floor effect, suggesting that the test may be too difficult for individuals with stroke. Conversely, if more than 20% of participants achieved the maximum possible total score or a sub-score of grade 4, this indicated a ceiling effect, suggesting that the test may be too easy for this population (McHorney et al., 1994). The intraclass correlation coefficient (ICC) with 95% confidence intervals was used to analyze reliability. Model 2,4 was used to assess inter-rater reliability, and model 3,4 was used to evaluate internal consistency. An ICC value (r) between 0.80 and 1.00 was interpreted as excellent reliability, between 0.50 and 0.79 as moderate reliability, and ≤0.50 as poor reliability (Portney, 2020). Spearman's rank correlation was employed to evaluate convergent validity by examining the relationships between the scores of mOFB and the total scores of the OFB, BBS, MAS sitting sub-item, and TIS. A correlation coefficient (r) between 0.80 and 1.00 indicated excellent validity, between 0.50 and 0.79 indicated moderate validity, and ≤0.50 indicated low validity (Portney, 2020). Both reliability and validity analyses were performed for all participations and separately for subgroups with subacute and chronic stroke to examine the measurement properties of the mOFB across different stages of stroke recovery. A significance level of $p < 0.05$ was used for all statistical tests.

## 4. Results

### 4.1 Participants

A total of 75 individuals with stroke from Bueng Yitho Medical and Rehabilitation Center and the Thai Red Cross Rehabilitation Center, who met the inclusion criteria and did not meet any exclusion criteria, were enrolled in the study between October 2018 and March 2019. There were no missing data in this study; all participants completed the assessments in full, and all data points were included in the analysis. The participants had a mean age of 61 ± 11 years, with 53 males and 22 females. The duration since stroke onset ranged from 1 month to 20 years (mean = 26.12 ± 38.87 months). Among all participants, 72% had ischemic stroke and 28% had hemorrhagic stroke. The majority of participants presented with left-sided impairments. The mean MMSE score was 24.72 (SD = 3.53), indicating that participants were cognitively intact and unlikely to have impairments that would affect balance performance. The mean scores of the mOFB were 9.80 $\pm$ 2.45, OFB was 11.35 $\pm$ 2.58, BBS was 33.51 $\pm$ 16.08, MAS was 3.71 $\pm$ 1.28, and TIS score was 4.20 $\pm$ 2.72.

For subgroup analysis, participants were divided into two groups based on stroke duration: subacute phase (≤ 6 months post-stroke; n = 29) and chronic phase (> 6 months post-stroke; n = 46). An independent t-test revealed no significant differences in baseline characteristics between groups including age and scores on the mOFB, OFB, BBS, MAS, sitting balance subscale and TIS except for MMSE scores, which were significantly higher in the subacute group ($p < 0.05$) as shown in Table 1.

### 4.2 Score Distribution of the mOFB and the OFB

The total scores of the mOFB were well distributed across participants, as shown in Figure 1. Only 1.3% scored the lowest possible score (0 points), and no participant reached the maximum score (16 points), indicating that the mOFB did not exhibit floor or ceiling effects when assessing both sitting and standing balance. Similarly, for the OFB, no participant scored 0, and only 2.7% achieved the maximum score of 16 points, indicating an appropriate score range for patients across all stroke stages.

**Table 1** Demographic and clinical characteristics of participants, classified by time after stroke

| Characteristic | All Participants (n=75) | Stage of Stroke | | P value |
|---|---|---|---|---|
| | | Subacute (n=29) | Chronic (n= 46) | |
| Age: years | 61.00 (11.61) | 57.03 (11.48) | 62.11 (11.37) | 0.065 |
| Sex: male/female, n (%) | 53/22 | 22/7 | 31/15 | - |
| Time after stroke: months | 26.12 (38.87) | 3.31 (1.80) | 40.50 (43.99) | 0.001 |
| Ischemic/hemorrhagic, n | 54/21 | 20/9 | 34/12 | - |
| Paresis side: Left/Right, n | 42/33 | 13/16 | 29/17 | - |
| MMSE score (/30) | 24.72 (3.53) | 25.72 (2.80) | 24.09 (3.81) | 0.050 |
| OFB score (/16) | 11.35 (2.58) | 11.21 (2.53) | 11.43 (2.63) | 0.712 |
| mOFB score (/16) | 9.80 (2.45) | 9.90 (2.57) | 9.74 (2.40) | 0.788 |
| BBS score (/56) | 33.51 (16.08) | 32.86 (17.21) | 33.91 (15.50) | 0.785 |
| MAS score (/6) | 3.71 (1.28) | 3.93 (1.19) | 3.57 (1.32) | 0.231 |
| TIS score (/16) | 4.20 (2.72) | 4.07 (2.80) | 4.28 (2.70) | 0.743 |

All values are presented as mean (SD). The stage of stroke is classified by time after stroke as subacute (0-6 months) and chronic (>6 months). MMSE=Mini-Mental State Exam, OFB=O'Sullivan Functional Balance Test, mOFB=Modified O'Sullivan Functional Balance Test, BBS=Berg Balance Scale, MAS= sitting balance subscale of Motor Assessment Scale, TIS=Trunk Impairment Scale

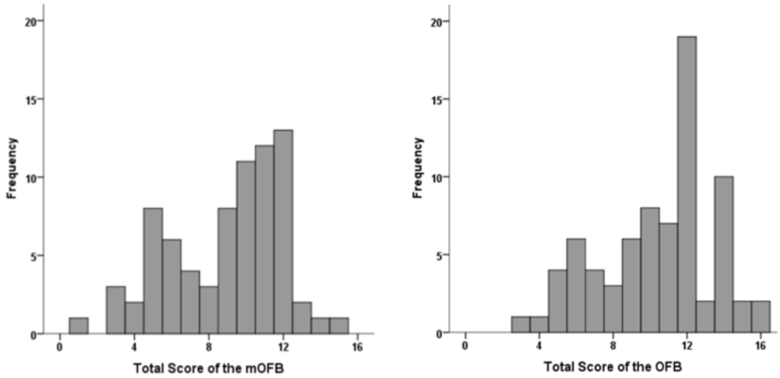


**Figure 1** Distribution of all participants' total scores from the mOFB and the OFB

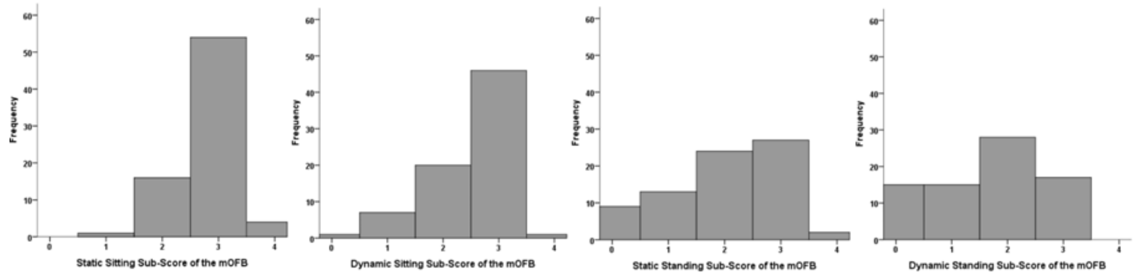


**Figure 2** Distribution of all participants' sub-score scores from the mOFB

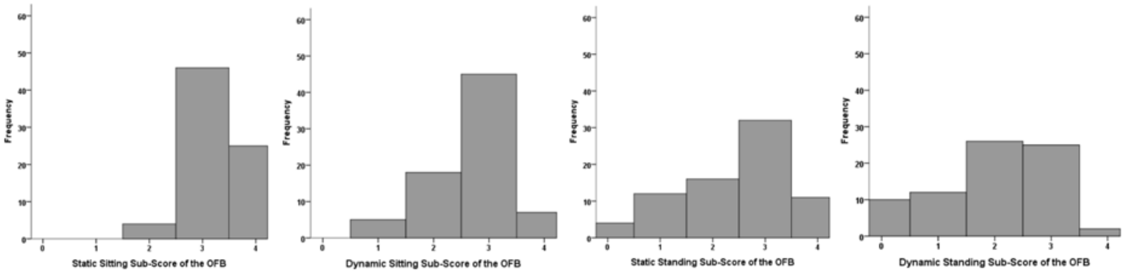


**Figure 3** Distribution of all participants' sub-score scores from the OFB

Figure 2 shows the sub-score distribution of the mOFB, revealing that 20% of all participants scored 0 on the dynamic standing sub-score, indicating a floor effect among individuals with stroke. Additionally, 20.7% of participants in the subacute group scored 0 on the static standing sub-score, also suggesting a floor effect among subacute stroke patients. Most of the sitting balance sub-scores were graded as 2 or 3, while the standing balance sub-scores were more broadly distributed, ranging from 0 to 4. Notably, the number of participants who achieved the maximum score was the lowest.

The distribution of the OFB sub-scores, as shown in Figure 3, indicated that 33.3% of all participants achieved the maximum score on the static sitting sub-score, suggesting a ceiling effect among stroke patients. Furthermore, 20.7% and 41.3% of participants in the subacute and chronic groups, respectively, attained the maximum score, indicating a ceiling effect in both subacute and chronic stroke populations. Most sitting balance sub-scores were graded as good (score 3), normal (score 4), and fair (score 2), respectively, while standing balance sub-scores showed a wider distribution from 0 to 3. Similar to mOFB, the number of participants achieving the maximum score on the standing sub-score was the lowest.

### 4.3 Reliability

As shown in Table 2, the mOFB demonstrated excellent intra-rater (ICC = 0.97, $p < 0.001$) and inter-rater reliability (ICC = 0.91, $p < 0.001$) for the total score. Sub-scores for sitting balance showed moderate to good reliability (intra-rater ICC = 0.91, $p < 0.001$; inter-rater ICC = 0.75, $p < 0.001$), while standing balance sub-scores showed excellent reliability (intra-rater ICC = 0.98, $p < 0.001$; inter-rater ICC = 0.93, $p < 0.001$). Subgroup analyses for subacute and chronic groups revealed reliability patterns similar to those of the full sample. The inter-rater reliability for static sitting sub-scores in both groups remained in the moderate range.

### 4.4 Validity

The Kolmogorov–Smirnov test showed that total scores for the mOFB, OFB, BBS, MAS sitting balance subscale, and TIS were not normally distributed. Therefore, Spearman's rank correlation was used to assess the relationships between test scores. The mOFB showed a strong positive correlation with the BBS ($r = 0.82$, $p < 0.001$) and moderate correlations with both the Sitting Balance (MAS) ($r = 0.67$, $p < 0.001$) and TIS score ($r = 0.60$, $p < 0.001$) across all participants. These results indicate good convergent validity. Similar trends were observed in both the subacute and chronic stroke groups (Table 3).

**Table 2** Inter-rater and intra-rater reliability of the modified O'Sullivan functional balance test and O'Sullivan functional balance test

| Outcome measure | All Participants (n=75) | | Stage of Stroke | | | |
|---|---|---|---|---|---|---|
| | | | Subacute (n=29) | | Chronic (n= 46) | |
| | Inter-rater reliability | Intra-rater reliability | Inter-rater reliability | Intra-rater reliability | Inter-rater reliability | Intra-rater reliability |
| **mOFB:** | | | | | | |
| Total score | 0.91* (0.845-0.946) | 0.97* (0.960-0.980) | 0.93* (0.861-0.966) | 0.98* (0.960-0.987) | 0.90* (0.801-0.945) | 0.97* (0.952-0.981) |
| Sitting score | 0.75* (0.632-0.837) | 0.91* (0.875-0.938) | 0.81* (0.669-0.904) | 0.92* (0.870-0.958) | 0.71* (0.541-0.830) | 0.90* (0.855-0.941) |
| Standing score | 0.93* (0.887-0.958) | 0.98* (0.966-0.983) | 0.94* (0.890-0.971) | 0.98* (0.962-0.988) | 0.92* (0.854-0.959) | 0.98* (0.963-0.985) |
| **OFB:** | | | | | | |
| Total score | 0.94* (0.868-0.965) | 0.98* (0.975-0.988) | 0.94* (0.856-0.973) | 0.98* (0.970-0.991) | 0.93* (0.860-0.967) | 0.98* (0.973-0.989) |
| Sitting score | 0.79* (0.662-0.869) | 0.93* (0.898-0.949) | 0.79* (0.599-0.897) | 0.93* (0.883-0.962) | 0.79* (0.651-0.880) | 0.93* (0.887-0.954) |
| Standing score | 0.96* (0.922-0.976) | 0.99* (0.981-0.991) | 0.96* (0.915-0.981) | 0.99* (0.978-0.993) | 0.96* (0.912-0.977) | 0.99* (0.980-0.992) |

All values are presented as ICC (95%CI). * Significant level of ICC at $p < 0.001$

**Table 3** Spearman's correlation coefficient of the modified O'Sullivan functional balance test, O'Sullivan functional balance test, and other criterion tests

| Criterion tests | All Participants (n=75) | | Stage of Stroke | | | |
|---|---|---|---|---|---|---|
| | | | Subacute (n=29) | | Chronic (n= 46) | |
| | **mOFB** score | **OFB** score | **mOFB** score | **OFB** score | **mOFB** score | **OFB** score |
| OFB score | 0.84* | 1.000 | 0.88* | 1.000 | 0.84* | 1.000 |
| BBS score | 0.82* | 0.80* | 0.89* | 0.86* | 0.77* | 0.77* |
| MAS score | 0.67* | 0.65* | 0.75* | 0.69* | 0.63* | 0.67* |
| TIS score | 0.60* | 0.70* | 0.45* | 0.59* | 0.70 * | 0.72* |

* Significant level of ICC at $p < 0.001$

## 5. Discussion

This study is the first to evaluate the reliability and validity of the modified O'Sullivan Functional Balance (mOFB) test following the implementation of standardized test instructions and clearly defined scoring criteria for use in individuals with stroke. The results demonstrate that the mOFB is a reliable and valid tool for assessing both sitting and standing balance in this population. The mOFB showed excellent intra-rater and inter-rater reliability, no evidence of floor or ceiling effects, and strong convergent validity when compared with established clinical balance measures such as the BBS.

Although the O'Sullivan Functional Balance (OFB) grading system has been widely utilized in clinical practice for individuals with neurological conditions and older adults (O'Sullivan, & Schmitz, 2007; University of Missouri, n.d.), a review of the literature indicates that its psychometric properties have not been formally investigated. The lack of standardized assessment procedures has resulted in considerable variability in its clinical application, including differences in initial positioning, the duration of static sitting or standing tasks, and the methods used to apply balance perturbations. This inconsistency in implementation may directly compromise the reliability of the assessment (Chinsongkram et al., 2020). The use of standardized test instructions and clearly defined, consensus-based scoring criteria has been shown to minimize scoring variability and enhance measurement reliability (Portney, 2020; Iansek, & Morris, 2013). This is consistent with the findings of the present study, which demonstrated a substantial improvement in the inter-rater reliability of the mOFB test following the implementation of standardized administration protocols and explicit scoring guidelines. These results align with those of Lim, & Chai, (2020), who examined the effects of standardized versus spontaneously translated instructions on the reliability of the Jebsen-Taylor Hand Function Test. Their findings indicated that standardized instructions significantly enhanced both test-retest and inter-rater reliability, underscoring the critical importance of consistent and clearly articulated procedures in ensuring the reliability and reproducibility of clinical assessments. The inter-rater reliability of the mOFB reported in this study was substantially higher than that observed in previous research by Chinsongkram et al., (2020), who assessed the mOFB in individuals with spinal cord injury (SCI). Their study reported low reliability for the sitting balance component (ICC = 0.21; 95% CI: 0.03–0.53) but relatively high reliability for the standing balance component (ICC = 0.92; 95% CI: 0.79–0.97). Although the earlier study involved a different clinical population, the current findings highlight the critical role of refining scoring procedures and providing clear performance criteria to improve reliability across diverse diagnostic groups. Importantly, this study extends the application of the mOFB to individuals with stroke, who exhibit unique postural control impairments compared to those with SCI. The incorporation of standardized procedures and explicit scoring criteria contributed to a consistent understanding among raters regarding test administration and interpretation, thereby enhancing reliability, promoting effective communication, and supporting continuity of care. This is particularly valuable in clinical settings involving multiple therapists or when patients transition between healthcare facilities such as returning to their primary hospital or continuing rehabilitation in community-based programs. Standardized assessment ensures a shared understanding of the patient's functional status, which is essential for planning, coordinating, and monitoring ongoing rehabilitation interventions.

The BBS is one of the most commonly used clinical tools for assessing balance, followed by the single-leg stance test (Albalwi et al., 2025), and is widely accepted as a criterion measure for evaluating functional balance in individuals with stroke (Pollock et al., 2007; Blum, & Korner-Bitensky, 2008). The excellent convergent validity demonstrated by the mOFB in this study suggests that it may serve as a viable alternative to the BBS in appropriate clinical contexts. However, the selection of a balance assessment tool should be aligned with the specific clinical objectives and care setting. The mOFB offers a simple and time-efficient assessment approach, making it particularly suitable for high-volume clinical environments with limited rehabilitation staff. Its ease of administration reduces the burden on clinicians while promoting patient comfort and compliance, particularly in individuals experiencing cognitive impairment or physical fatigue. Additionally, the clarity and structure of its scoring system facilitate consistent interpretation and effective communication among physical therapists and across interdisciplinary teams. Nevertheless, the mOFB provides only limited insight into the specific subsystems contributing to postural control deficits. Therefore, when balance impairments are detected using the mOFB, more comprehensive assessments such as the BBS or the BESTest (Horak et al., 2009) may be warranted to support more detailed clinical reasoning and to inform individualized rehabilitation planning.

For enhancing generalizability, this study employed broad inclusion criteria that did not restrict participants based on functional ability levels. Data were collected from two distinct rehabilitation settings to capture a wide range of patient characteristics: an inpatient subacute program at the Thai Red Cross rehabilitation center and a community-based outpatient program at the Bueng Yitho Medical and rehabilitation center, which included both subacute and chronic community-dwellings stroke survivors. Most participants were able to sit or stand with minimal to moderate assistance, corresponding to modified Rankin Scale (mRS) scores of 2 to 4, thereby supporting the external validity of the study findings. Subgroup analyses further reinforced the applicability of the results across different phases of stroke recovery. The mOFB demonstrated consistent reliability and validity in both subacute and chronic stroke groups, underscoring its potential utility throughout the rehabilitation continuum. The absence of floor and ceiling effects in total scores indicates that the test provides an appropriate level of challenge for detecting balance impairments across a broad functional spectrum. However, the presence of floor effects in the static and dynamic standing subscales among subacute participants suggests limited sensitivity for individuals with more severe impairments.

**Study Limitations and Future Research**

Our study population comprised individuals with stroke who were able to sit or stand with assistance and had an mRS score ranging from 2 to 4. Therefore, when generalizing these findings to more severely impaired populations or bedridden, as floor and ceiling effects may differ in such groups. Particularly, relying solely on subscale scores may introduce challenges related to floor or ceiling effects that differ from those observed in total scores. These effects may limit the sensitivity of the mOFB in detecting meaningful changes in postural control over time (Portney, 2020). Therefore, future refinement of the subscale items such as adjusting task difficulty or enhancing scoring granularity may be warranted to improve the instrument's responsiveness, particularly when used to monitor progress throughout the rehabilitation process. Additionally, future studies are warranted to examine the concurrent validity of the mOFB against laboratory-based balance assessments such as force plates or accelerometers that measure center of mass displacement. Further investigation is also needed to evaluate its utility as a screening tool, including its sensitivity, specificity, and predictive validity in identifying individuals at risk of balance impairments or falls.

## 6. Conclusion

The mOFB demonstrates strong reliability and validity in assessing balance in individuals with stroke. The test is appropriately challenging across all stages of stroke recovery, with minimal floor or ceiling effects, and is suitable for use in clinical screening. Its brief administration time and simple protocol make it particularly useful in clinical settings with high patient volumes. The mOFB is therefore recommended for use as an initial balance screening tool in stroke rehabilitation, with follow-up assessments as needed for treatment planning.

## 7. Acknowledgements

This study was financially supported by the Research Institute of Rangsit University, Thailand. The authors wish to thank Bueng Yitho Medical and Rehabilitation Center, Thai Red Cross Rehabilitation Center, and Rangsit University. We appreciate the participants' contributions to this study.

## 8. References

Albalwi, A. A., Alharbi, A. A., Al Amer, H. S., Alamrani, S. A., Albalawi, H. F., Alatawi, M. K., ... & Javid, J. (2025). Utilizing balance assessment tools by physical therapists for patients with balance disorders. *Healthcare*, *13*(8), Article 928. https://doi.org/10.3390/healthcare13080928

Banyam, T., & Rakpongsiri, K. (2024). Designing a low-cost force plate for postural assessment. *Journal of Current Science and Technology*,14(3), Article60. https://doi.org/10.59796/jcst.V14N3.2024.60

Berg, K. O., Maki, B. E., Williams, J. I., Holliday, P. J., & Wood-Dauphinee, S. L. (1992). Clinical and laboratory measures of postural balance in an elderly population. *Archives of Physical Medicine and Rehabilitation*, *73*(11), 1073-1080.

Bernhardt, J., Hayward, K. S., Kwakkel, G., Ward, N. S., Wolf, S. L., Borschmann, K., ... & Cramer, S. C. (2017). Agreed definitions and a shared vision for new standards in stroke recovery research: the stroke recovery and rehabilitation roundtable taskforce. *International Journal of Stroke*, *12*(5), 444-450. https://doi.org/10.1177/1747493017711816

Blum, L., & Korner-Bitensky, N. (2008). Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Physical Therapy*, *88*(5), 559-566. https://doi.org/10.2522/ptj.20070205

Bour, A., Rasquin, S., Boreas, A., Limburg, M., & Verhey, F. (2010). How predictive is the MMSE for cognitive performance after stroke?. *Journal of Neurology*, *257*, 630-637. https://doi.org/10.1007/s00415-009-5387-9

Cabrera-Martos, I., Ortiz-Rubio, A., Torres-Sánchez, I., López-López, L., Jarrar, M., & Valenza, M. C. (2020). The effectiveness of core exercising for postural control in patients with stroke: A systematic review and meta-analysis. *PM&R*, *12*(11), 1157-1168. https://doi.org/10.1002/pmrj.12330

Carr, J. H., Shepherd, R. B., Nordholm, L., & Lynne, D. (1985). Investigation of a new motor assessment scale for stroke patients. *Physical Therapy*, *65*(2), 175-180. https://doi.org/10.1093/ptj/65.2.175

Chinsongkram, B., Chaikeeree, N., Saengsirisuwan, V., Viriyatharakij, N., Horak, F. B., & Boonsinsukh, R. (2014). Reliability and validity of the balance evaluation systems test (BESTest) in people with subacute stroke. *Physical Therapy*, *94*(11), 1632-1643. https://doi.org/10.2522/ptj.20130558

Chinsongkram, B., Pluempitiwiriyawej, S., Hongthong, S., & Boonsinsukh, R. (2020). Reliability of the Modified O'Sullivan Functional Balance Test in Person with Spinal Cord Injury. *Journal of Food Health and Bioenvironmental Science*, *13*(2), 53-58.

Horak, F. B., Wrisley, D. M., & Frank, J. (2009). The balance evaluation systems test (BESTest) to differentiate balance deficits. *Physical Therapy*, *89*(5), 484-498. https://doi.org/10.2522/ptj.20080071

Huh, J. S., Lee, Y. S., Kim, C. H., Min, Y. S., Kang, M. G., & Jung, T. D. (2015). Effects of balance control training on functional outcomes in subacute hemiparetic stroke patients. *Annals of Rehabilitation Medicine*, *39*(6), 995-1001. https://doi.org/10.5535/arm.2015.39.6.995

Iansek, R., & Morris, M.E. (2013). *Rehabilitation in movement disorders*. Cambridge: Cambridge University Press

Katan, M., & Luft, A. (2018). Global burden of stroke. *Seminars in Neurology*, 38(2), 208-211. https://doi.org/10.1055/s-0038-1649503

Lim, Y. X., & Chai, S. C. (2020). Standardized translated instruction versus spontaneously translated instruction: Test-retest and interrater reliability of a hand function test. *Journal of Hand Therapy*, *33*(4), 553-561. https://doi.org/10.1016/j.jht.2019.09.001

Malouin, F., Pichard, L., Bonneau, C., Durand, A., & Corriveau, D. (1994). Evaluating motor recovery early after stroke: comparison of the Fugl-Meyer Assessment and the Motor Assessment Scale. *Archives of Physical Medicine and Rehabilitation*, *75*(11), 1206-1212. https://doi.org/10.1016/0003-9993(94)90006-X

Mancini, M., & Horak, F. B. (2010). The relevance of clinical balance assessment tools to differentiate balance deficits. *European Journal of Physical and Rehabilitation Medicine*, *46*(2), 239-248.

McHorney, C. A., Ware Jr, J. E., Lu, J. R., & Sherbourne, C. D. (1994). The MOS 36-ltem short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care*, *32*(1), 40-66. https://doi.org/10.1097/00005650-199401000-00004

O'Sullivan, S. B., & Schmitz, T. J. (2007). *Physical rehabilitation: Assessment and treatment* (5th ed.). Philadelphia, US: F. A. Davis Company.

Park, J., & Kim, T. H. (2019). The effects of balance and gait function on quality of life of stroke patients. *NeuroRehabilitation*, *44*(1), 37-41. https://doi.org/10.3233/NRE-182467

Pollock, A., Baer, G., Pomeroy, V. M., & Langhorne, P. (2007). Physiotherapy treatment approaches for the recovery of postural control and lower limb function following stroke. *Cochrane Database of Systematic Reviews*, *2007*(1). https://doi.org/10.1002/14651858.CD001920.pub2

Portney, L. G. (2020). *Foundations of clinical research: Applications to evidence-based practice* (4th ed.). Philadelphia, US: F. A. Davis Company.

Rakpongsiri, K., Srisurak, K., & Rakpongsiri, P. (2023). Predicting elderly falls using the roll factor in the "Smart Walk Trainer" equipment for body weight-balance analysis. *Journal of Current Science and Technology*, 13(3), 609-618. https://doi.org/10.59796/jcst.V13N3.2023.978

Shumway-Cook, A., & Woollacott, M. H. (2017). *Motor control: Translating research into clinical practice* (5th ed.). Philadelphia, US: Lippincott Williams & Wilkins.

University of Missouri. (n.d.). *Functional balance grades - static and dynamic*. Geriatric Assessment Tool Kit. Retrieved from https://geriatrictoolkit.missouri.edu/Balance_functional_grades--static--dynamic.doc

Verheyden, G., Nieuwboer, A., Mertin, J., Preger, R., Kiekens, C., & De Weerdt, W. (2004). The Trunk Impairment Scale: a new tool to measure motor impairment of the trunk after stroke. *Clinical Rehabilitation*, *18*(3), 326-334. https://doi.org/10.1191/0269215504cr733oa

Verheyden, G., & Kersten, P. (2010). Investigating the internal validity of the Trunk Impairment Scale (TIS) using Rasch analysis: the TIS 2.0. *Disability and Rehabilitation*, *32*(25), 2127-2137. https://doi.org/10.3109/09638288.2010.483038

Vincent-Onabajo, G., Joseph, E., & Musa, H. Y. (2018). Impact of balance on functional independence after stroke: A cross-sectional study at rehabilitation settings in Nigeria. *NeuroRehabilitation*, *42*(4), 499-504. https://doi.org/10.3233/NRE-172305

Xu, T., Clemson, L., O'Loughlin, K., Lannin, N. A., Dean, C., & Koh, G. (2018). Risk factors for falls in community stroke survivors: a systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation*, *99*(3), 563-573. https://doi.org/10.1016/j.apmr.2017.06.032

## Supplement Material
## Modified O'Sullivan Functional Balance Test (mOFB) Instructions

**Purpose**

The mOFB is an adaptation of the original O'Sullivan Functional Balance Grade. This test consists of a 4-condition performance test aimed to assess both static and dynamic balance in sitting and standing positions in individuals with neurological conditions who have significant balance impairments.

**Equipment Required**

Stopwatch, stable chair, or bed and reach targets (e.g., 3 points cane, cones, markers) positioned beyond the patient's arm length in six directions: anterior, posterior, left, right, upward, and downward

**Instructions and Procedures**

**Sitting balance test**

**Preparing step:**

- The patient is seated on a stable bed or chair without back support, with both feet placed flat on the floor. The examiner needs to adjust the subject's postural alignment before starting the test to ensure that the patient sits up straight and should stand beside the subject to give support if necessary.

**Static sitting balance test**

- Instruct the patient to sit in a normal and comfortable posture **for 60 seconds**. Try not to use your hands unless you must. (If the patient cannot maintain balance independently, minimal hand support or examiner assistance is permitted.)
- If the patient can maintain independent sitting for 60 seconds, in the next step, the examiner instructs the patient to maintain an upright sitting posture when they **apply external resistance** in 4 directions including forward, backward, left, and right sideways. The amount of external resistance given to the patient should be minimal just enough to trigger isometric contraction of trunk muscle.

**Dynamic sitting balance test**

- If the patient can maintain independent sitting for 60 seconds, in the next step, the examiner instructs the patient to reach out to touch targets placed just **beyond arm's length** in the following directions: forward, backward, left sideway, right sideway, upward, and downward.

**Standing balance test**

**Preparing step:**

- The patient is standing on a stable surface without back support, with both feet placed flat on the floor. The examiner needs to adjust the subject's postural alignment before starting the test to ensure that the patient stands straight and should stand beside the subject to give support if necessary.

**Static standing balance test**

- Instruct the patient to stand in a normal and comfortable posture **for 60 seconds**. Try not to use your hand support unless you must. (If the patient cannot maintain balance independently, minimal hand support or examiner assistance is permitted.)
- If the patient can maintain independent standing for 60 seconds, in the next step, the examiner instructs the patient to maintain an upright sitting posture when they **apply external resistance** in 4 directions including forward, backward, left, and right sideways. The amount of external resistance given to the patient should be minimal just enough to trigger isometric contraction of trunk muscle.

**Dynamic standing balance test**

- If the patient can maintain independent standing for 60 seconds, in the next step, the examiner instructs the patient to reach out to touch targets placed just **beyond arm's length** in the following directions: forward, backward, left sideway, right sideway, upward, and downward.

**Scoring**

- Performance is rated by **observation** of balance control during each task.
- Scoring criteria are based on the ability to maintain a steady balance while performing each task.

**Table 4** Scoring Criteria of the modified O'Sullivan functional balance test

| | **Scoring Criteria** | |
|---|---|---|
| **Grade** | **Static Balance** | **Dynamic Balance** |
| 4 (Normal) | Maintains balance without support and maintains steady balance against the external resistance for disturbed balance in all directions. | Maintains balance without support and maintains steady balance during reaching over-arm range in all directions. |
| 3 (Good) | Maintains balance without the support and maintains a steady balance against the external resistance for disturbed balance in some directions. | Maintains balance without support and maintains steady balance during reaching over-arm range in some directions. |
| 2 (Fair) | Maintains balance without support but is unable to maintain steady balance against the external resistance | Maintains balance without support but is unable to maintain steady balance during reaching over. |
| 1 (Poor) | Requires continuous hand support or significant assistance to maintain posture | Requires continuous hand support or significant assistance to maintain posture |
| 0 (Zero) | Unable to maintain posture independently | Unable to maintain posture independently |