Objective functional assessment using the “Timed Up and Go” test in patients with lumbar spinal stenosis

Martin N. Stienen, MD, FEBNS,¹ Nicolai Maldaner, MD,² Holger Joswig, MD, FEBNS,³
Marco V. Corniola, MD,³ David Bellut, MD,¹ Peter Prömmel, MD,² Luca Regli, MD,¹
Astrid Weyerbrock, MD,² Karl Schaller, MD,³ and Oliver P. Gautschi, MD^{3,4}

¹Department of Neurosurgery, University Hospital Zurich and Clinical Neuroscience Center, University of Zurich; ²Department of Neurosurgery, Cantonal Hospital St. Gallen; ³Department of Clinical Neuroscience, Division of Neurosurgery, Geneva University Hospitals and Faculty of Medicine, Geneva; and ⁴Neuro- and Spine Center, Hirslanden Clinic St. Anna, Lucerne, Switzerland

OBJECTIVE Patient-reported outcome measures (PROMs) are standard of care for the assessment of functional impairment. Subjective outcome measures are increasingly complemented by objective ones, such as the “Timed Up and Go” (TUG) test. Currently, only a few studies report pre- and postoperative TUG test assessments in patients with lumbar spinal stenosis (LSS).

METHODS A prospective two-center database was reviewed to identify patients with LSS who underwent lumbar decompression with or without fusion. The subjective functional status was estimated using PROMs for pain (visual analog scale [VAS]), disability (Roland-Morris Disability Index [RMDI] and Oswestry Disability Index [ODI]), and health-related quality of life (HRQoL; 12-Item Short-Form Physical Component Summary [SF-12 PCS] and the EQ-5D) preoperatively, as well as on postoperative day 3 (D3) and week 6 (W6). Objective functional impairment (OFI) was measured using age- and sex-standardized TUG test results.

RESULTS Sixty-four patients ($n = 32$ [50%] male, mean age 66.8 ± 11.7 years) were included. Preoperatively, they reported a mean VAS back pain score of 4.1 ± 2.7 , VAS leg pain score of 5.4 ± 2.7 , RMDI of 10.4 ± 5.3 , ODI of 41.9 ± 16.2 , SF-12 PCS score of 32.7 ± 8.3 , and an EQ-5D index of 0.517 ± 0.226 . The preoperative rates of severe, moderate, and mild OFI were 4.7% ($n = 3$), 12.5% ($n = 8$), and 7.8% ($n = 5$), respectively, and the mean OFI T-score was 116.3 ± 23.7 . At W6, 60 (93.8%) of 64 patients had a TUG test result within the normal population range (no OFI); 3 patients (4.7%) had mild and 1 patient (1.6%) severe OFI. The mean W6 OFI T-score was significantly decreased (103.1 ± 13.6 ; $p < 0.001$). Correspondingly, the PROMs showed a decrease in subjective VAS back pain (1.6 ± 1.7 , $p < 0.001$) and leg pain (1.0 ± 1.8 , $p < 0.001$) scores, disability (RMDI 5.3 ± 4.7 , $p < 0.001$; ODI 21.3 ± 16.1 , $p < 0.001$), and increase in HRQoL (SF-12 PCS 40.1 ± 8.3 , $p < 0.001$; EQ-5D 0.737 ± 0.192 , $p < 0.001$) at W6. The W6 responder status (clinically meaningful improvement) ranged between 81.3% (VAS leg pain) and 29.7% (EQ-5D index) of patients.

CONCLUSIONS The TUG test is a quick and easily applicable tool that reliably measures OFI in patients with LSS. Objective tests incorporating longer walking time should be considered if OFI is suspected but fails to be proven by the TUG test, taking into account that neurogenic claudication may not clinically manifest during the brief TUG examination. Objective tests do not replace the subjective PROM-based assessment, but add valuable information to a comprehensive patient evaluation.

<https://thejns.org/doi/abs/10.3171/2019.2.FOCUS18618>

KEYWORDS lumbar spinal stenosis; objective measure of function; outcome assessment; Timed Up and Go test; TUG; objective functional impairment

ABBREVIATIONS 6WD = 6-minute walking distance; 6WT = 6-minute walking test; CCI = Charlson Comorbidity Index; D3 = day 3; DDD = degenerative disc disease; HRQoL = health-related quality of life; LSS = lumbar spinal stenosis; MCID = minimum clinically important difference; MCS = Mental Component Summary; MTT = motorized treadmill test; ODI = Oswestry Disability Index; OFI = objective functional impairment; PCS = Physical Component Summary; PROM = patient-reported outcome measure; RMDI = Roland-Morris Disability Index; SF-12 = 12-Item Short-Form; SPWT = self-paced walking test; SWT = shuttle walking test; TUG = Timed Up and Go; VAS = visual analog scale; W6 = week 6.

SUBMITTED November 13, 2018. **ACCEPTED** February 18, 2019.

INCLUDE WHEN CITING DOI: 10.3171/2019.2.FOCUS18618.

LUMBAR spinal stenosis (LSS) is a highly disabling, degenerative condition that typically manifests in elderly patients and progressively interferes with ambulation and subsequent activities of daily living.²⁴ Multifactorial in origin, bulging discs, overgrown facet joints, and hypertrophied ligaments may eventually cause posture- and motion-dependent impingement and ischemia of lumbar nerve roots that is commonly referred to as neurogenic claudication.^{20,47} LSS is the most common indication for spinal surgery today,^{3,8} and constitutes a growing medical and economic concern facing an aging Western population.^{3,24}

Despite its relatively characteristic clinical picture, previous landmark studies varied in terms of diagnostic criteria and outcome measures,^{1,10,48} which limits the comparability and generalizability of their findings.^{20,21} The accurate measurement of an individual patient's clinical condition at first presentation and follow-up is essential to guide therapeutic decision-making and to determine treatment efficacy. In today's evidence- and value-based medicine environment, subjective patient-reported outcome measures (PROMs) are frequently incorporated into research and clinical healthcare. In parallel, the use of objective outcome measures increasingly complements the comprehensive evaluation, which is also true of patients with LSS and degenerative disc disease (DDD).^{13,39}

A number of objective outcome measures have been proposed for patients with LSS, including motorized treadmill tests (MTTs),^{7,36,44,45} self-paced walking tests (SPWTs),^{36,45,46} the shuttle walking test (SWT),³⁶ and the 6-minute walking test (6MWT),^{11,29} as well as accelerometric and/or gait analyses using body wearable devices. Although previously applied to this indication,^{26–28,31} data on the objective functional evaluation by means of the “Timed Up and Go” (TUG) test in homogenous cohorts of patients with LSS are currently limited. The aim of this study was to report data on the objective functional evaluation by means of the TUG test in patients with LSS, and to summarize the current literature.

Methods

Patient Identification

Consecutive adult patients with lumbar DDD were recruited between December 2013 and 2015 in a prospective two-center study aiming to quantify functional impairment by means of the TUG test.^{38,40} The original protocol included patients with lumbar disc herniation scheduled for microdiscectomy, LSS scheduled for microsurgical decompression, and both LSS and DDD with associated instability scheduled for spinal fusion surgery. In this study we retrospectively analyzed data sets of LSS patients with and without associated instability pertaining to patient characteristics, baseline functional status, and day 3 (D3) as well as week 6 (W6) postoperative outcome. Pregnant women as well as patients with walking impairment due to any other known neurological, orthopedic, or rheumatic disease (e.g., paraparesis, hemiparesis, hip/knee osteoarthritis, vertigo) were excluded a priori.

Data Collection

We recorded demographic patient characteristics (age,

sex, BMI, smoking, and working status) and information on comorbidity (Charlson Comorbidity Index and American Society of Anesthesiologists risk score),²⁵ as well as disease-specific and surgical parameters. Each patient's clinical condition was assessed by applying a set of generic and disease-specific PROMs, including back pain and leg pain on the visual analog scale (VAS; score range 0–10), functional impairment (Roland-Morris Disability Index [RMDI]; 24 items, score range from 0 [no disability] to 24 [severe disability]),³⁷ Oswestry Disability Index (ODI; 10 items, score range from 0 [no disability] to 100 [severe disability]),¹⁰ and health-related quality of life (HRQoL; 12-Item Short-Form [SF-12] Physical Component Summary [PCS] and Mental Component Summary [MCS; 12 items, results standardized to a mean of 50]; EQ-5D index: 5 items, score range from –0.074 [worst HRQoL] to 1.0 [best HRQoL], using European norms).⁹

The TUG Test

The TUG test was performed in a standardized manner,^{12,17,34} asking patients to stand up from a chair on the word “Go” and walk as fast as possible (no running) to a marked line on the floor at 3 meters distance. At that line, patients would turn around by 180°, return to the chair, and sit down again as quickly as possible. The TUG test result is the time interval between getting up from the chair and sitting down again in seconds and can be measured using a stopwatch or the free smartphone TUG app (see Appendix).

Patient Management and Surgical Treatment

Patients underwent a minimum of 6 weeks of conservative treatment before being considered for surgery. In patients with degenerative spondylolisthesis or enlarged zygapophysial intraarticular spaces on MRI, or in those reporting significant mechanical back pain, flexion-extension radiographs were requested. Decompression and instrumented fusion was recommended to patients with LSS and associated instability.

Microsurgical decompression was performed while patients were under general anesthesia and lying prone. The correct levels were marked using lateral radiographs. Skin and fascia were opened, and the paravertebral musculature was scraped off with a Cobb elevator. In fusion procedures, transpedicular screws were first inserted. The inferior part of the superior lamina was then drilled off under microscopic control. The ligamentum flavum was resected, and the interlaminar window widened. At the discretion of the surgeon, a hemilaminectomy or laminectomy was performed in some instances. In cases of bilateral stenosis, decompression was performed either via a unilateral fenestration with “over-the-top” decompression, or bilateral fenestrations. In fusion procedures, facetectomy and intersomatic cage implementation via a transforaminal lumbar interbody fusion trajectory was preferably performed. Dural tears were fixed with 5-0 microsutures. Drains were inserted only in cases of insufficient hemostasis.

Statistical Analyses

Raw TUG test times were transformed into age- and

sex-adjusted T-scores (with T-score > 123 indicating objective functional impairment [OFI]). The severity of OFI was categorically determined using previously determined cutoff values generated from 110 healthy subjects.^{16,40} The W6 response to surgical treatment was determined in terms of improvement by at least the minimum clinically important difference (MCID), using previously published values for VAS back pain (1.2), VAS leg pain (1.6), RMDI (5.0), ODI (12.8), EQ-5D index (0.359), SF-12 PCS (4.9), and the TUG test (3.4 seconds).^{5,19,30,32,33,43} A patient not improving by at least the MCID was considered a “non-responder” on the respective outcome measure. Logistic regression was used to estimate the relationship between presence of preoperative OFI and the W6 responder status.

Categorical variables were presented as count (percentage) and nominal variables as group mean (SD). Longitudinal outcomes were compared using t-tests. The software used for the statistical analysis was Stata (version 14.2, StataCorp LP). A p value < 0.05 was considered statistically significant.

Ethical Considerations

The study was approved by the local IRBs (University Hospital of Geneva and Cantonal Hospital St. Gallen). All patients provided written informed consent.

Results

We included a total sample of 64 patients (n = 32 male, 50%) with a mean age of 66.8 ± 11.7 years. Detailed baseline demographic data are outlined in Table 1.

Preoperative Status

The preoperative clinical status of the patients is summarized by the following mean scores (Fig. 1): VAS back pain 4.1 ± 2.7, VAS leg pain 5.4 ± 2.7, RMDI 10.4 ± 5.3, ODI 41.9 ± 16.2, SF-12 PCS 32.7 ± 8.3, EQ-5D index 0.517 ± 0.226, and SF-12 MCS 44.4 ± 10.8. A motor deficit was noted in 5 patients (7.8%; 1 patient with British Medical Research Council [BMRC] paresis grade 2/5 and 4 patients with BMRC paresis grade 4/5). Six patients (9.4%) were receiving regular opioid pain medication.

In the objective assessment, patients had a mean raw TUG test time of 10.2 ± 5.0 seconds (range 5.3–27.2 seconds). This corresponded to age- and sex-adjusted TUG test values within the range of the normal population in 48 patients (75.0%). The remaining 16 patients were classified as mild (n = 5, 7.8%), moderate (n = 8, 12.5%), or severe OFI (n = 3, 4.7%; Table 2). The mean OFI T-score was 116.3 ± 23.7 (range 94.9–222.4; Fig. 2).

All patients underwent microsurgical decompression of 1 (n = 49, 76.6%), 2 (n = 10, 15.6%), 3 (n = 4, 6.3%), or 4 segments (n = 1, 1.6%). Decompressions were performed unilaterally (n = 45, 70.3%) or bilaterally (n = 19, 29.7%). Surgical fusion in addition to decompression was performed in 17 (26.6%) of 64 cases.

Status on Postoperative D3

The patients' status on D3 was as follows (Fig. 1): mean

TABLE 1. Baseline characteristics of 64 patients with LSS

Variable	Value (%)
Mean age ± SD, yrs	66.8 ± 11.7
Sex	
Male	32 (50.0)
Female	32 (50.0)
Mean BMI ± SD, kg/m ²	27.5 ± 4.7
Smoking status	
Nonsmoker	52 (81.3)
Smoker	12 (18.7)
CCI	
0–1	58 (90.6)
2–4	6 (9.4)
ASA grade	
1	3 (4.7)
2	51 (79.7)
3	10 (15.6)
Working status	
Working	20 (31.3)
Retired	39 (60.9)
Invalidity pension	1 (1.6)
No occupation	4 (6.2)
Multilevel surgery	
No	49 (76.6)
Yes	15 (23.4)
Side of decompression	
Rt	22 (34.4)
Lt	23 (35.9)
Bilat	19 (29.7)
Preop motor deficit	
No	59 (92.2)
Yes	5 (7.8)

ASA = American Society of Anesthesiologists; CCI = Charlson Comorbidity Index.

VAS back pain score 2.3 ± 1.9 (p < 0.001), VAS leg pain 0.9 ± 1.3 (p < 0.001), RMDI 9.0 ± 5.8 (p = 0.168), ODI 33.2 ± 18.0 (p = 0.006), SF-12 PCS 33.2 ± 6.9 (p = 0.695), EQ-5D index 0.665 ± 0.215 (p = 0.002), and SF-12 MCS 45.0 ± 10.3 (p = 0.736).

The objective D3 assessment was available for 58 patients (90.6%). These patients had a mean TUG test time of 10.4 ± 5.4 seconds (range 4.5–30.0 seconds). This corresponded to age- and sex-adjusted TUG test values within the range of the normal population in 40 of 58 patients (69.0%). The remaining 18 patients were classified as mild (n = 9, 15.5%), moderate (n = 6, 10.3%), or severe OFI (n = 3, 5.2%). The mean OFI T-score was 116.9 ± 22.3 (range 93.2–211.7, p = 0.886; Fig. 2).

Status on Postoperative W6

Six weeks postoperatively, patients reported mean scores of: VAS back pain 1.6 ± 1.7 (p < 0.001), VAS leg pain 1.0 ± 1.8 (p < 0.001), RMDI 5.3 ± 4.7 (p < 0.001),

FIG. 1. Boxplots displaying the median with the 25th–75th percentile (*box*), the upper and lower adjacent values (*whiskers*), and outliers (*dots*) of each outcome measure at baseline (*blue*), postoperative D3 (*red*), and postoperative W6 (*green*). **A:** VAS back pain. **B:** VAS leg pain. **C:** RMDI. **D:** ODI. **E:** SF-12 PCS. **F:** EQ-5D index.

ODI 21.3 ± 16.1 ($p < 0.001$), SF-12 PCS 40.1 ± 8.3 ($p < 0.001$), EQ-5D index 0.737 ± 0.192 ($p < 0.001$), and SF-12 MCS 48.5 ± 10.4 ($p = 0.030$) (Fig. 1). Three patients (4.7%) were receiving daily opioid pain medication ($p = 0.492$).

The W6 objective assessment was available in all 64 patients. These patients had a mean TUG test time of $7.2 \pm$

3.9 seconds (range 4.0–33.5 seconds). This corresponded to age- and sex-adjusted TUG test values within the range of the normal population in 60 patients (93.8%). Of the remaining 4 patients, 3 were classified as mild (4.7%) and 1 as severe OFI (1.6%). The mean OFI T-score was 103.1 ± 13.6 (range 88.6–182.5, $p < 0.001$; Fig. 2).

Apart from higher preoperative VAS back pain in the

patient group with LSS and instability scheduled for decompression and surgical fusion (5.4 ± 2.7 vs 3.6 ± 2.3 , $p = 0.016$), the results of all subjective (PROM-based) and objective assessments (TUG test) were similar for patients with LSS undergoing decompression only at baseline, D3, and W6 (all $p > 0.05$).

Longitudinal Analysis of Patients With Preoperative OFI

Table 2 outlines case data pertaining to those patients who were found to have any degree of OFI in the preoperative evaluation. Some patients showed transient worsening of their functional status at D3 but distinct improvement at W6 (cases 1, 6, 10, and 11), whereas other patients already improved at D3 (cases 3–5, 7–9, 12, and 15). Case 9 showed some improvement at D3, but was functionally worse at W6, compared to the preoperative objective functional status.

Analysis of the W6 Responder Status

Table 3 outlines the W6 responder status to the surgical treatment. The metric on which most patients responded favorably was VAS leg pain (81.3%), followed by the ODI (65.6%), VAS back pain (62.5%), SF-12 PCS (57.8%), RMDI (48.4%), and the TUG test (31.3%). Fewer than 1 in 3 patients was considered a W6 responder on the EQ-5D index (29.7%).

Patients with any preoperative degree of OFI were as likely as patients without OFI to be W6 treatment responders on each of the PROM metrics, namely VAS back pain (odds ratio [OR] 1.00, 95% confidence interval [CI] 0.31–3.22, $p > 0.99$) and VAS leg pain (OR 0.60, 95% CI 0.15–2.34, $p = 0.462$), RMDI (OR 1.52, 95% CI 0.49–4.75, $p = 0.472$), ODI (OR 0.83, 95% CI 0.26–2.70, $p = 0.761$), EQ-5D index (OR 0.73, 95% CI 0.20–2.65, $p = 0.636$), and SF-12 PCS (OR 1.30, 95% CI 0.41–4.14, $p = 0.662$).

Discussion

The aim of this work was to provide more insight into the applicability of the TUG test in a cohort of patients with LSS before and after microsurgical decompression with or without fusion. Retrospectively analyzing prospectively collected data of 64 patients, several interesting notions emerged. First, we found that—despite evaluating selected patients with considerable degrees of suffering in terms of subjective PROMs—the preoperative objective functional status was within the range of the healthy normal population in three-quarters of the cohort. Second, at W6, the objective functional status was within the range of the healthy normal population in 94% of patients and, correspondingly, we observed significant improvements on each of the applied subjective PROMs. And third, the presence or absence of OFI at baseline was not associated with a favorable response to surgical treatment on any of the applied subjective outcome metrics.

In this study, the TUG test was applied, as it is easy to employ and tests a great range of activities, in which patients with degenerative diseases of the lumbar spine are limited: standing up, accelerating, walking, decelerating, turning around, and sitting down. Its reliability and convergent validity with commonly accepted PROMs for

TABLE 2. Longitudinal functional analysis of 16 patients undergoing lumbar microsurgical decompression with or without fusion and preoperative OFI of any degree

Case No.	Age (yrs), Sex	Objective Measure	Time Point		
			Preop	D3	W6
1	60, M	TUG test (sec)	16.2	16.8	10.9
		OFI	Moderate	Moderate	No
		T-score	131.4	133.1	114.3
2	40, M	TUG test (sec)	8.3	8.3	7.2
		OFI	Mild	Mild	No
		T-score	129.2	129.2	119.9
3	50, M	TUG test (sec)	16.6	10.6	6.5
		OFI	Severe	Moderate	No
		T-score	197.4	147.8	114.6
4	43, M	TUG test (sec)	10.0	5.4	5.9
		OFI	Moderate	No	No
		T-score	143.1	105.0	109.8
5	67, F	TUG test (sec)	24.5	13.7	6.5
		OFI	Moderate	No	No
		T-score	154.0	119.6	96.5
6	55, F	TUG test (sec)	9.9	24.2	5.4
		OFI	Mild	Severe	No
		T-score	123.8	211.7	96.5
7	52, M	TUG test (sec)	10.1	9.2	6.9
		OFI	Moderate	Moderate	No
		T-score	143.7	136.7	117.9
8	68, M	TUG test (sec)	14.3	11.1	6.33
		OFI	Mild	No	No
		T-score	125.4	114.8	99.6
9	80, F	TUG test (sec)	22.4	14.4	33.5
		OFI	Moderate	No	Severe
		T-score	147.1	121.8	182.5
10	43, F	TUG test (sec)	11.4	12.3	6.3
		OFI	Mild	Mild	No
		T-score	133.3	138.8	101.5
11	70, M	TUG test (sec)	18.4	30.0	7.8
		OFI	Moderate	Severe	No
		T-score	138.3	175.7	104.4
12	50, M	TUG test (sec)	11.5	8.8	5.4
		OFI	Moderate	Mild	No
		T-score	155.5	132.8	105.7
13	88, F	TUG test (sec)	20.5	NA	10.5
		OFI	Moderate	NA	Mild
		T-score	141.1	NA	127.8
14	58, F	TUG test (sec)	26.0	NA	10.5
		OFI	Severe	NA	Mild
		T-score	222.4	NA	127.8
15	74, F	TUG test (sec)	27.2	14.9	6.8
		OFI	Severe	Mild	No
		T-score	162.5	123.3	97.5
16	81, M	TUG test (sec)	15.5	NA	11.5
		OFI	Mild	NA	No
		T-score	129.1	NA	116.2

NA = not available.

FIG. 2. Boxplots displaying the median OFI T-score with the 25th–75th percentile (box), the upper and lower adjacent values (whiskers), and outliers (dots) at baseline (blue), postoperative D3 (red), and postoperative W6 (green).

lumbar DDD has been demonstrated.^{16,38,40,41} In addition, the TUG test has been shown to be a robust assessment tool, relatively independent of factors known to influence the PROM-based patient evaluation such as obesity,⁴⁰ smoking,²³ and mental health status.⁴² Its cutoff value, representing a clinically meaningful change in function (or MCID), was determined to be 3.4 seconds.¹⁹

So far, the TUG test has only been applied in a limited number of studies explicitly dealing with LSS that are summarized below.^{26–28,31,38}

In 2005, Lin and Lin evaluated how sensorimotor function, balance, and physical performance relate to disability and walking capacity.²⁸ They included 50 patients with LSS and a mean age of 60.7 ± 7.3 years, with a mean VAS pain score of 5.6 ± 1.9 , mean ODI of 32.0 ± 16.0 , and mean Swiss Spinal Stenosis Questionnaire (SSSQ) physical function scale score of 2.96 ± 0.68 . The mean raw TUG test result was 6.4 ± 2.3 seconds (range 3.6–15.0 seconds). The authors reported a significant correlation between the TUG test result and both the ODI ($r = 0.446$, $p < 0.01$) and SSSQ physical function scale score ($r = -0.530$, $p < 0.01$). The TUG test stood out as the variable with the highest correlation to disability, its results explaining approximately 20% of the observed variance. Because the TUG test is simple, reliable, and requires only a chair and 3 meters of walking space, the authors recommended its use as a screening tool for disability in the assessment of LSS patients.²⁸

In 2011, Kim et al.²⁶ applied the TUG test in 80 patients enrolled in a case-control study to compare the estimated risk of falling between patients with LSS and those with bilateral knee osteoarthritis ($n = 40$, each group). The LSS cohort consisted of 11 males (27.5%) with a mean age of 62.8 ± 7.3 years and a relatively low mean ODI score of 18.9 (SD not indicated), considering this was a cohort of surgical candidates. The mean raw TUG test result in the LSS cohort was 14.1 ± 1.8 seconds, considerably higher than in the report by Lin and Lin²⁸ and the raw TUG test results of the present study. The authors reported no sig-

TABLE 3. Analysis of the W6 responder status to lumbar microsurgical decompression with or without fusion using various subjective PROMs and the objective TUG test

Outcome Metric	Responders	Nonresponders
VAS back pain	40 (62.5)	24 (37.5)
VAS leg pain	52 (81.3)	12 (18.7)
RMDI	31 (48.4)	33 (51.6)
ODI	42 (65.6)	22 (34.4)
SF-12 PCS	37 (57.8)	27 (42.2)
EQ-5D index	19 (29.7)	45 (70.3)
TUG test	20 (31.3)	44 (68.7)

Values given as number of patients (%).

nificant correlation between the TUG test and the ODI score ($r = -0.447$, $p = 0.374$). They concluded that PROMs such as the ODI were an addition—but no substitute—for objective outcome measures, which is in agreement with the experience of our group in subsequent studies.^{12,38–40}

In 2014, Lee et al. published a follow-up report to compare the effect of conservative versus surgical treatment of LSS on the estimated risk of falling.²⁷ The surgical cohort consisted of 76 patients (mean age 62.4 ± 11.0 years; $n = 30$ [39.5%] males; mean ODI of 25.8 ± 9.7 ; mean raw TUG test 15.4 ± 2.3 seconds) and the conservative cohort of 50 patients (mean age 64.6 ± 6.2 years; $n = 18$ [36.0%] males; mean ODI of 22.3 ± 6.2 ; mean raw TUG test 16.9 ± 2.3 seconds). Despite gradual improvements on both the ODI and TUG metric at 3 and 12 months postoperatively, those group differences were statistically insignificant when a conservative Bonferroni method was applied to adjust for multiple testing. The 3-month ODI (mean 19.2 ± 8.2) and TUG test results (10.5 ± 1.8 seconds) in the surgically treated cohort are comparable (within one SD) to the results of the W6 outcome obtained in our patients. Interestingly, the ODI, EQ-5D, VAS, and TUG test scores were similar in the conservatively and surgically managed cohorts at their 3- and 12-month follow-ups.²⁷

In 2017, Park et al.³¹ examined female patients with LSS and healthy control subjects in the age range between 65 and 85 years. The authors applied PROMs (EQ-5D index, ODI) and the TUG test, which differed significantly between the healthy (mean age 70.6 ± 4.0 years; EQ-5D index 0.77 ± 0.15 ; ODI 13.6 ± 12.9) and diseased group (mean age 70.6 ± 4.7 years; EQ-5D index 0.57 ± 0.13 ; ODI 32.8 ± 15.3). There was a significant correlation between the TUG test and the ODI ($r = 0.54$, $p < 0.01$) and between the TUG test and the EQ-5D index by trend ($r = -0.32$; $p = 0.08$) in patients with LSS. The authors did not provide details on any posttherapeutic changes in subjective and objective measures of function.

The body of literature in conjunction with our experience of applying the TUG test in 110 healthy subjects and 375 patients with degenerative diseases of the lumbar spine (including 135 with LSS)^{38,40} show that raw results of the TUG test differ between healthy subjects and diseased LSS patients.^{31,40} Longer TUG test times correlate positively with back/leg pain and subsequent disability, and negatively with HRQoL.^{28,31,38,40} Correlations are weak to

moderate, however, suggesting that the objective functional evaluation adds a different dimension to the comprehensive patient evaluation, which can supplement but not replace the subjective PROM-based assessment.^{12,26,38,40} The TUG test is sensitive to change, to a similar extent as commonly used PROMs (Tables 2 and 3). The use of validated OFI categories/T-scores offers distinct strengths over the use of raw TUG test values as it allows an appreciation of an individual patient's deviation from the population norm, thereby minimizing bias introduced by the variables age and sex.^{14,18} In the first postoperative days, the TUG test result appears to be influenced by wound pain,¹⁵ which can be appreciated by the increase in patients with OFI on D3, compared to preoperatively. This should be taken into consideration when interpreting early postoperative objective assessments.

The fact that only about 1 in 4 surgical LSS candidates was classified to have OFI when applying the TUG test merits some further explanation. Because the threshold to classify a patient as "objectively impaired" was intentionally chosen high (99th percentile of the normal population TUG test results),⁴⁰ the specificity of the TUG test to detect functional impairment is high, which comes at the cost of a low sensitivity, however. It thus implies that even when a patient is classified with "mild" OFI, LSS symptoms are already pronounced. It should be borne in mind that clinically relevant neurogenic claudication may not appear over the short course of the TUG test (walking 3 meters twice). Therefore, other objective tests that challenge patients over a longer walking distance and duration are likely more sensitive and appropriate in the setting of patients with LSS.

The SPWT has previously been described as the "gold standard" of objective outcome measurement in patients with LSS.^{22,45} Patients are instructed to walk continuously at their own pace around an indoor 200-meter track until they have to stop or until a maximum time of 30 minutes. Its test-retest reliability was excellent (intraclass correlation coefficient = 0.98).⁴⁵ However, its results varied highly, ranging from 60 to 2065 meters (mean 776 ± 726 meters) and 67–1800 seconds (mean 840 ± 690 seconds).⁴ The MCID of the SPWT was found to be 363 meters in a small sample of 26 patients with LSS.⁴⁶ The convergent validity with the MTT, self-estimated walking time/distance, as well as with symptoms of neurogenic claudication were moderate to high.^{4,36} The SPWT outperformed the MTT in terms of internal responsiveness, whereas external responsiveness of both tests was rather poor.^{4,36} While the SPWT appears to be a good choice to accurately measure LSS-related disability, the need for a special facility and personnel/staff to conduct the measurement render this test cumbersome to apply in daily patient care. The same problem applies to the MTT, SWT, and other sophisticated laboratory-based gait analyses.

In this regard, more flexible solutions such as the 6WT, for example, are preferable. According to the American Thoracic Society guidelines, this test is traditionally performed in a well-illuminated flat hallway.⁶ Patients are instructed to walk back and forth as often as possible for 6 minutes. The test result is the 6-minute walking distance (6WD, in meters).^{11,29} Recently, smartphone applications have been programmed that measure the 6WD, as well as

time to first symptoms (in seconds) and distance to first symptoms (in meters) in the patient's home environment using global positioning system (GPS) technology (see Appendix).³⁸ The 6WT is less explored than the SPWT for LSS, but it has been applied as an outcome measure in recent high-quality studies.^{11,29} The MCID in the 6WD ranged from 14 to 30.5 meters in various populations with chronic cardiopulmonary conditions,² but still remains to be determined for LSS. The ease of administration using a smartphone in particular makes the 6WT a promising candidate for both future LSS studies and daily patient care.³⁸

Strengths and Limitations

Strengths of this study include the two-center design that applied the TUG test to patients with different linguistic and cultural backgrounds. This allows generalizing the present findings to different clinical settings. In addition, a comprehensive set of validated generic and disease-specific scales and health surveys was included. The most important limitations are inherent to the retrospective design and to its relatively small sample size. The latter was within the range of previous related research, however.^{4,26–29,35,36,45,46} The analysis included both LSS patients undergoing microsurgical decompression and those undergoing a fusion procedure, and differences in the recovery after each procedure type may have influenced the analysis of the postoperative functional level. As no additional follow-up data were available, we currently can describe neither the objective postoperative rehabilitation process of LSS patients in more detail, nor the value of the TUG test for long-term outcome evaluation.

Conclusions

The TUG test is a quick and easily applicable tool that reliably measures OFI in patients with LSS. Objective tests incorporating longer walking time such as the SPWT, MTT, or 6WT should be considered if OFI is suspected but fails to be proven by the TUG test, taking into account that neurogenic claudication may not clinically manifest during the relatively brief examination. Preferably, objective test results should be analyzed in a standardized fashion. They do not replace the subjective PROM-based assessment, but add valuable information to a comprehensive patient evaluation.

Acknowledgments

We thank Cornelia Lüthi, Dario Jucker, and Ivan Chau for their important contribution in the collection of data that are reported in this article. We thank Nicolas Smoll in particular, who provided essential input when establishing the concept of OFI measures, as well as Prof. Gerhard Hildebrandt for his incredibly kind support throughout the conduction of the study.

Appendix

The TUG app is available for smartphones and can be downloaded free of charge in the iTunes or Google Play app store in multiple languages, including English, German, French, Italian, Spanish, Portuguese, Turkish, Romanian, Hungarian, Dutch, Croatian, Arab, Chinese, Russian, and Albanian.

The 6WT app is available for smartphones and can be down-

loaded free of charge in the iTunes or Google Play app store in multiple languages, including English, German, and French.

References

1. Atlas SJ, Deyo RA, Keller RB, Chapin AM, Patrick DL, Long JM, et al: The Maine Lumbar Spine Study, Part III. 1-year outcomes of surgical and nonsurgical management of lumbar spinal stenosis. **Spine (Phila Pa 1976)** **21**:1787–1795, 1996
2. Bohannon RW, Crouch R: Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: a systematic review. **J Eval Clin Pract** **23**:377–381, 2017
3. Ciol MA, Deyo RA, Howell E, Kreif S: An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. **J Am Geriatr Soc** **44**:285–290, 1996
4. Conway J, Tomkins CC, Haig AJ: Walking assessment in people with lumbar spinal stenosis: capacity, performance, and self-report measures. **Spine J** **11**:816–823, 2011
5. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY: Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study Questionnaire Short Form 36, and pain scales. **Spine J** **8**:968–974, 2008
6. Crapo RO, Casaburi R, Coates AL, Enright PL, MacIntyre NR, McKay RT, et al: ATS statement: guidelines for the six-minute walk test. **Am J Respir Crit Care Med** **166**:111–117, 2002 (Erratum in **Am J Respir Crit Care Med** **193**:1185, 2016)
7. Deen HG Jr, Zimmerman RS, Lyons MK, McPhee MC, Verheijde JL, Lemens SM: Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. **J Neurosurg** **83**:27–30, 1995
8. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI: United States trends in lumbar fusion surgery for degenerative conditions. **Spine (Phila Pa 1976)** **30**:1441–1447, 2005
9. EuroQol Group: EuroQol—a new facility for the measurement of health-related quality of life. **Health Policy** **16**:199–208, 1990
10. Fairbank JC, Couper J, Davies JB, O'Brien JP: The Oswestry low back pain disability questionnaire. **Physiotherapy** **66**:271–273, 1980
11. Försth P, Ólafsson G, Carlsson T, Frost A, Borgström F, Fritzell P, et al: A randomized, controlled trial of fusion surgery for lumbar spinal stenosis. **N Engl J Med** **374**:1413–1423, 2016
12. Gautschi OP, Corniola MV, Joswig H, Smoll NR, Chau I, Jucker D, et al: The timed up and go test for lumbar degenerative disc disease. **J Clin Neurosci** **22**:1943–1948, 2015
13. Gautschi OP, Corniola MV, Schaller K, Smoll NR, Stienen MN: The need for an objective outcome measurement in spine surgery—the timed-up-and-go test. **Spine J** **14**:2521–2522, 2014
14. Gautschi OP, Corniola MV, Smoll NR, Joswig H, Schaller K, Hildebrandt G, et al: Sex differences in subjective and objective measures of pain, functional impairment, and health-related quality of life in patients with lumbar degenerative disc disease. **Pain** **157**:1065–1071, 2016
15. Gautschi OP, Joswig H, Corniola MV, Smoll NR, Schaller K, Hildebrandt G, et al: Pre- and postoperative correlation of patient-reported outcome measures with standardized Timed Up and Go (TUG) test results in lumbar degenerative disc disease. **Acta Neurochir (Wien)** **158**:1875–1881, 2016
16. Gautschi OP, Smoll NR, Corniola MV, Joswig H, Chau I, Hildebrandt G, et al: Validity and reliability of a measurement of objective functional impairment in lumbar degenerative disc disease: the Timed Up and Go (TUG) test. **Neurosurgery** **79**:270–278, 2016
17. Gautschi OP, Smoll NR, Corniola MV, Joswig H, Schaller K, Hildebrandt G, et al: Sex differences in lumbar degenerative disc disease. **Clin Neurol Neurosurg** **145**:52–57, 2016
18. Gautschi OP, Smoll NR, Joswig H, Corniola MV, Schaller K, Hildebrandt G, et al: Influence of age on pain intensity, functional impairment and health-related quality of life before and after surgery for lumbar degenerative disc disease. **Clin Neurol Neurosurg** **150**:33–39, 2016
19. Gautschi OP, Stienen MN, Corniola MV, Joswig H, Schaller K, Hildebrandt G, et al: Assessment of the minimum clinically important difference in the Timed Up and Go Test after surgery for lumbar degenerative disc disease. **Neurosurgery** **80**:380–385, 2017
20. Genevay S, Atlas SJ: Lumbar spinal stenosis. **Best Pract Res Clin Rheumatol** **24**:253–265, 2010
21. Genevay S, Atlas SJ, Katz JN: Variation in eligibility criteria from studies of radiculopathy due to a herniated disc and of neurogenic claudication due to lumbar spinal stenosis: a structured literature review. **Spine (Phila Pa 1976)** **35**:803–811, 2010
22. Jespersen AB, Gustafsson MEAK: Correlation between the Oswestry Disability Index and objective measurements of walking capacity and performance in patients with lumbar spinal stenosis: a systematic literature review. **Eur Spine J** **27**:1604–1613, 2018
23. Joswig H, Stienen MN, Smoll NR, Corniola MV, Chau I, Schaller K, et al: Effects of smoking on subjective and objective measures of pain intensity, functional impairment, and health-related quality of life in lumbar degenerative disc disease. **World Neurosurg** **99**:6–13, 2017
24. Kalichman L, Cole R, Kim DH, Li L, Suri P, Guermazi A, et al: Spinal stenosis prevalence and association with symptoms: the Framingham Study. **Spine J** **9**:545–550, 2009
25. Keats AS: The ASA classification of physical status—a recapitulation. **Anesthesiology** **49**:233–236, 1978
26. Kim HJ, Chun HJ, Han CD, Moon SH, Kang KT, Kim HS, et al: The risk assessment of a fall in patients with lumbar spinal stenosis. **Spine (Phila Pa 1976)** **36**:E588–E592, 2011
27. Lee BH, Kim TH, Park MS, Lim S, Park JO, Kim HS, et al: Comparison of effects of nonoperative treatment and decompression surgery on risk of patients with lumbar spinal stenosis falling: evaluation with functional mobility tests. **J Bone Joint Surg Am** **96**:e110, 2014
28. Lin SI, Lin RM: Disability and walking capacity in patients with lumbar spinal stenosis: association with sensorimotor function, balance, and functional performance. **J Orthop Sports Phys Ther** **35**:220–226, 2005
29. Loske S, Nüesch C, Byrnes KS, Fiebig O, Schären S, Münderrmann A, et al: Decompression surgery improves gait quality in patients with symptomatic lumbar spinal stenosis. **Spine J** **18**:2195–2204, 2018
30. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korf M, et al: Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. **Spine (Phila Pa 1976)** **33**:90–94, 2008
31. Park S, Han HS, Kim GU, Kang SS, Kim HJ, Lee M, et al: Relationships among disability, quality of life, and physical fitness in lumbar spinal stenosis: an investigation of elderly Korean women. **Asian Spine J** **11**:256–263, 2017
32. Parker SL, Mendenhall SK, Shau D, Adogwa O, Cheng JS, Anderson WN, et al: Determination of minimum clinically important difference in pain, disability, and quality of life after extension of fusion for adjacent-segment disease. **J Neurosurg Spine** **16**:61–67, 2012
33. Parker SL, Mendenhall SK, Shau DN, Adogwa O, Anderson WN, Devin CJ, et al: Minimum clinically important dif-

- ference in pain, disability, and quality of life after neural decompression and fusion for same-level recurrent lumbar stenosis: understanding clinical versus statistical significance. **J Neurosurg Spine** 16:471–478, 2012
34. Podsiadlo D, Richardson S: The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. **J Am Geriatr Soc** 39:142–148, 1991
 35. Pratt RK, Fairbank JC, Virr A: The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. **Spine (Phila Pa 1976)** 27:84–91, 2002
 36. Rainville J, Childs LA, Peña EB, Suri P, Limke JC, Jouve C, et al: Quantification of walking ability in subjects with neurogenic claudication from lumbar spinal stenosis—a comparative study. **Spine J** 12:101–109, 2012
 37. Roland M, Fairbank J: The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. **Spine (Phila Pa 1976)** 25:3115–3124, 2000
 38. Stienen MN, Bellut D, Regli L, Hausmann ON, Gautschi OP: [Functional assessment of patients with lumbar degenerative disc disease: who is right – the doctor, the patient or the objective test?] **Praxis (Bern)** 106:1041–1052, 2017 (Ger)
 39. Stienen MN, Joswig H, Chau I, Neidert MC, Bellut D, Wälchli T, et al: Efficacy of intraoperative epidural triamcinolone application in lumbar microdiscectomy: a matched-control study. **J Neurosurg Spine** 28:291–299, 2018
 40. Stienen MN, Joswig H, Smoll NR, Corniola MV, Schaller K, Hildebrandt G, et al: Influence of body mass index on subjective and objective measures of pain, functional impairment, and health-related quality of life in lumbar degenerative disc disease. **World Neurosurg** 96:570–577.e1, 2016
 41. Stienen MN, Smoll NR, Joswig H, Corniola MV, Schaller K, Hildebrandt G, et al: Validation of the baseline severity stratification of objective functional impairment in lumbar degenerative disc disease. **J Neurosurg Spine** 26:598–604, 2017
 42. Stienen MN, Smoll NR, Joswig H, Snagowski J, Corniola MV, Schaller K, et al: Influence of the mental health status on a new measure of objective functional impairment in lumbar degenerative disc disease. **Spine J** 17:807–813, 2017
 43. Stratford PW, Binkley JM, Riddle DL, Guyatt GH: Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. **Phys Ther** 78:1186–1196, 1998
 44. Tenhula J, Lenke LG, Bridwell KH, Gupta P, Riew D: Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. **J Spinal Disord** 13:276–282, 2000
 45. Tomkins CC, Battié MC, Rogers T, Jiang H, Petersen S: A criterion measure of walking capacity in lumbar spinal stenosis and its comparison with a treadmill protocol. **Spine (Phila Pa 1976)** 34:2444–2449, 2009
 46. Tomkins-Lane CC, Battié MC, Macedo LG: Longitudinal construct validity and responsiveness of measures of walking capacity in individuals with lumbar spinal stenosis. **Spine J** 14:1936–1943, 2014
 47. Truumees E: Spinal stenosis: pathophysiology, clinical and radiologic classification. **Instr Course Lect** 54:287–302, 2005
 48. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, et al: Surgical versus nonsurgical therapy for lumbar spinal stenosis. **N Engl J Med** 358:794–810, 2008

Disclosures

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions

Conception and design: Stienen, Gautschi. Acquisition of data: Stienen, Joswig, Corniola, Gautschi. Analysis and interpretation of data: Stienen, Joswig, Corniola, Gautschi. Drafting the article: Stienen. Critically revising the article: Maldaner, Joswig, Corniola, Bellut, Prömmel, Regli, Weyerbrock, Schaller, Gautschi. Reviewed submitted version of manuscript: Stienen. Approved the final version of the manuscript on behalf of all authors: Stienen. Statistical analysis: Stienen. Administrative/technical/material support: Schaller, Gautschi. Study supervision: Schaller, Gautschi.

Correspondence

Martin N. Stienen: University Hospital Zurich and Clinical Neuroscience Center, University of Zurich, Switzerland. mnstienen@gmail.com.