Assessment of the Minimum Clinically Important Difference in the Timed Up and Go Test After Surgery for Lumbar Degenerative Disc Disease

Oliver P. Gautschi, MD*

Martin N. Stienen, MD*

Marco V. Corniola, MD*

Holger Joswig, MD‡

Karl Schaller, MD*

Gerhard Hildebrandt, MD‡

Nicolas R. Smoll, MBBS§

*Department of Neurosurgery and Faculty of Medicine, University Hospital Geneva, Geneva, Switzerland; ‡Department of Neurosurgery, Cantonal Hospital St. Gallen, St. Gallen, Switzerland; §School of Medicine and Public Health, University of Newcastle, Newcastle, New South Wales, Australia

Correspondence:
Nicolas Smoll, MBBS,
School of Medicine and Public Health,
University of Newcastle,
University Drive,
Callaghan NSW 2308,
Australia.
E-mail: nrsmoll@gmail.com

BACKGROUND: The Timed Up and Go Test (TUG Test) has previously been described as a reliable tool to evaluate objective functional impairment in patients with degenerative disc disease.

OBJECTIVE: The aim of this study was to assess the minimum clinically important difference (MCID) of the TUG Test.

METHODS: The TUG Test (measured in seconds) was correlated with validated patient-reported outcome measures (PROs) of pain intensity (Visual Analog Scale for back and leg pain), functional impairment (Oswestry Disability Index, Roland Morris Disability Index), and health-related quality of life measures (Short Form-12 and EuroQol 5D). Three established methods were used to establish anchor-based MCID values using responders of the following PROs (Visual Analog Scale back and leg pain, Oswestry Disability Index, Roland Morris Disability Index, EuroQol 5D index, and Short Form-12 Physical Component Summary) as anchors: (1) average change, (2) minimum detectable change, and (3) change difference approach.

RESULTS: One hundred patients with a mean ± SD age of 56.2 ± 16.1 years, 57 (57%) male, 45 patients undergoing microdiscectomy, 35 undergoing lumbar decompression, and 20 undergoing fusion surgery were studied. The 3 MCID computation methods revealed a range of MCID values according to the PRO used from 0.9 s (Oswestry Disability Index based on the change difference approach) to 6.0 s (EuroQol 5D index based on the minimum detectable change approach), with a mean MCID of 3.4 s for all measured PROs.

CONCLUSION: The MCID for the TUG Test time is highly variable depending on the computation technique used. The average TUG Test MCID was 3.4 s using all 3 methods and all anchors.

KEY WORDS: Degenerative disc disease, Low back pain, Lumbar spine surgery, MCID, Objective outcome measurement, Postoperative outcome, Timed Up and Go Test

P atient-reported outcome measures (PROs) have become the standard measure of treatment effectiveness in patients undergoing spine surgery for degenerative disc disease (DDD).[1] Outcome assessment in lumbar DDD does not only assist in monitoring safety and treatment efficacy of any surgical or nonsurgical treatment, but is also useful for comparing different treatment strategies.[2] One of the major drawbacks of well-established PROs (eg, Oswestry Disability Index [ODI] and Roland-Morris Disability Index [RMDI] for functional impairment and Short Form 36, Short Form 12 [SF-12], and EuroQol 5D [EQ5D] for health-related quality of life [HRQoL]) is their limited comparability. Patients may rate their subjective pain

---

ABBREVIATIONS: D3, day 3; **DDD,** degenerative disc disease; **EQ5D,** EuroQol 5D; **FU,** follow-up; **HRQoL,** health-related quality of life; **MCID,** minimum clinically important difference; **ODI,** Oswestry Disability Index; **OFI,** objective functional impairment; **PCS,** Physical Component Summary; **PRO,** patient-reported outcome measure; **RMDI,** Roland-Morris Disability Index; **SF-12,** Short Form 12; **TUG Test,** Time Up and Go Test; **VAS,** Visual Analog Scale

or functional disability differently for reasons of educational, cultural, and motivational interindividual differences.[3] Altogether, there is a dearth of objective, or at least less subjective, measures of functional disability to aid the process of clinical decision making in lumbar DDD. Currently, the only existing objective clinical outcome measures are the measurement of range of movement (eg, with a goniometer), the measurement of absolute muscle strength (eg, with a newton meter), or the measurement of walking speed and distance.[3,4] New objective outcome measures have been tested using an advanced tracking technology based on global positioning systems. Thus far, none of these measurements have been adopted into daily clinical routine.[5,6]

We recently presented the validation of a simple test, the Timed Up and Go Test (TUG Test), as a measurement tool of objective functional impairment (OFI) with an excellent intra- and inter-rater reliability.[4,7] The validity of the TUG Test was demonstrated with good correlation with the Visual Analog Scale (VAS) back and leg pain and functional impairment indexes (RMDI and ODI), as well as with HRQoL. The upper limit of "normal" was determined as 11.5 s, and the corresponding OFI could be categorized as mild, moderate, and severe OFI according to age, sex, and TUG Test time performance.[7]

Although all established PROs deliver specific numeric values of pain intensity, functional impairment, and HRQoL, the scores do not directly translate into clinically meaningful improvement.[1] Thus, the concept of the minimum clinically important difference (MCID) was established and later introduced as a critical threshold to define/determine treatment effectiveness.[1,8] The original conceptualization of MCID was an implicit ratio to weigh benefits and risks of treatment.[9] Currently, MCID research mainly focuses on the benefit side of this ratio and most commonly considers MCID to be "the smallest change that is important to patients."[10] So far, several studies have reported a range of different MCIDs for established PROs in different patient populations with various spinal pathologies.[1]

The aim of this study was to assess the MCID of the TUG Test as a measure of functional impairment for patients undergoing surgery for lumbar DDD.

METHODS

Patients were prospectively enrolled at the Departments of Neurosurgery of the University Hospital Geneva and the Cantonal Hospital St. Gallen in Switzerland between September 2013 and April 2015. Patients were scheduled for lumbar spine surgery for the following diagnoses: (1) lumbar disc herniation, (2) lumbar spinal stenosis, and (3) lumbar DDD with or without instability requiring lumbar fusion (transforaminal lumbar interbody fusion, posterior lumbar interbody fusion, or extreme lateral interbody fusion). Exclusion criteria were age younger than 18 years, pregnancy, severe neurological deficits, or other medical reasons interfering with the patients' ability to walk and perform the TUG Test (eg, Parkinson disease, wheelchair user, severe congestive heart failure). Patient demographic data (age, sex, size, weight, body mass index, and working status) as well as clinical data (British Medical

Research Council muscle strength of index muscles of the lower extremities, smoking status, and previous surgery) were obtained preoperatively. VAS back and leg pain scores, RMDI and ODI questionnaires, and the SF-12 (including both physical and mental component summary [PCS and MCS] scores), and EQ5D questionnaires were recorded, and a single TUG Test was performed at baseline, day 3 (D3), and at week 6 (W6) of follow-up (FU). There were 4 qualified raters who performed the TUG Test, including 3 residents in advanced training and 1 board-certified staff neurosurgeon.

Ethics Approval

The study was approved by the Institutional Review Board of the University of Geneva (14-079) and the Ethics Committee St. Gallen (14/049). Written informed consent was obtained from every study participant.

Statistical Analysis

TUG Test times and various subjective outcome measures at baseline, D3, and W6 were compared using t tests. Previously established PROs, including pain intensity (VAS back and leg pain scores), functional impairment (ODI and RMDI), and HRQoL (EQ5D index and SF-12 PCS score) were used as anchors according to the method of Parker et al.[1] Responder status to surgical treatment was defined at the W6 FU using previously published MCID values for an ODI of 12.8, RMDI of 5.0, PCS of 4.9, EQ5D index of 0.359, VAS leg pain score of 1.6, and VAS back pain score of 1.2 to generate a binary responder and nonresponder variable.[8,10-13] Three previously published anchor-based approaches were used to assess the MCID after lumbar spine surgery.[1] According to Parker et al,[1] we computed the MCID based on specified definitions using the previously mentioned anchors: Method (1) The average change approach defines the MCID as the average change in the study cohort defined as responders. Method (2) The minimum detectable change approach defines the MCID as the smallest change that can be considered above the measurement error. The MCID is thus equal to the upper value of the 95% confidence interval for the average change score seen in the study cohort defined as nonresponders. Method (3) The change difference approach defines the MCID as the difference in the average change score for responders and nonresponders. These methods produce a multitude of MCIDs, and although there is no consensus on the optimal method, we built the average of all MCIDs to determine the clinically used MCID.

Data were analyzed using Stata version 14.0 (StataCorp, College Station, Texas). Statistical significance was accepted at $P < .05$.

RESULTS

The patient cohort included 100 patients (43 female patients, 43.0%) with a mean $\pm$ SD age of 56.2 $\pm$ 16.1 years. Before hospitalization, 44 patients (44.0%) worked full or part time, 16 patients (16.0%) were not working, 3 patients (3.0%) were disabled, and 37 patients (37.0%) were retired. Forty-five patients (45.0%) underwent microdiscectomy for lumbar disc herniation or recurrent lumbar disc herniation, 35 patients (35.0%) underwent a uni- or bilateral decompression for lumbar spinal stenosis, and 20 patients (20.0%) underwent a lumbar fusion procedure. At the W6 FU, statistically significant improvement ($P < .001$) was observed for the following PROs assessed: VAS

TABLE 1. Pre- and Postoperative Subjective and Objective Measures at Baseline, 3 Days, and 6 Weeks[a,b]

Measure	Preoperative	D3	P	W6 FU	P	D3 vs W6
TUG Test time (seconds)	10.1 (5.0)	9.4 (4.5)	.3198	6.6 (3.2)	<.0001	<0.0001
VAS back pain score	3.8 (1.0)	2.1 (1.8)	<.0001	1.4 (1.7)	<.0001	0.0031
VAS leg pain score	5.5 (2.8)	0.9 (1.4)	<.0001	1.0 (1.7)	<.0001	0.7484
ODI	46 (18)	32 (18)	<.0001	21 (17)	<.0001	<0.0001
RMDI	12 (5)	9 (6)	.0014	5 (5)	<.0001	<0.0001
EQ5D index	0.440 (0.218)	0.674 (0.202)	<.0001	0.744 (0.187)	<.0001	0.0008
SF-12 PCS	31.3 (7.7)	35.0 (6.9)	.0008	40.5 (8.3)	<.0001	<0.0001
SF-12 MCS	41.9 (9.7)	42.5 (10.7)	.5177	48.1 (10.3)	<.0001	0.0003

[a]D3, day 3; EQ5D, EuroQol-5D; FU, follow-up; MCS, Mental Component Summary; ODI, Oswestry Disability Index; PCS, Physical Component Summary; RMDI, Roland-Morris Disability Index; SF-12, Short Form 12; TUG Test, Timed Up and Go Test; VAS, Visual Analog Scale; W6, week 6.
[b]The statistical analysis is in comparison with the preoperative baseline measures. The last row compares D3 and W6 follow-up. Values shown are mean ± SD.

back and leg pain scores (mean improvement, −2.5 ± 2.9 and −4.5 ± 3.3, respectively), ODI (−24.6 ± 17.4), RMDI (−6.4 ± 6.1), EQ5D index (+0.304 ± 0.249), SF-12 PCS (+9.2 ± 9.1), and SF-12 MCS (+6.2 ± 9.6). Also, the TUG Test time revealed a significant mean improvement (−3.5 ± 4.6 s, $P < .001$).

Subjective and objective measures (mean ± SD values) are shown in Table 1. All subjective outcome measures except SF-12 MCS significantly improved at D3 after the surgery. The mean TUG Test time did not improve significantly in the early postoperative period (10.1 ± 5.0 s preoperatively compared with 9.4 ± 4.5 s at D3 [$P = .32$]). From D3 to the W6 FU, all subjective and objective outcome measures significantly improved ($P < .003$) except VAS leg pain score, which already had improved to its maximum at D3 (Table 1).

Sixty-one patients (61%) and 82 patients (82%) were responders according to the MCID for VAS back and leg pain, respectively. Seventy-three patients (73%) and 61 patients (61%) were responders according to the MCID for the ODI and RMDI, respectively. Forty-three patients (43%), 29 patients (29%), and 38 patients (38%) were responders according to the MCID for the EQ5D index, the SF-12 PCS, and the SF-12 MCS, respectively. Using VAS back pain score as an anchor, the MCID threshold for

the TUG Test time ranged from 1.3 to 5.3 s depending on the computation method (Table 2). Using the VAS leg pain score as an anchor, the MCID threshold for the TUG Test time ranged from 1.3 to 4.8 s depending on the computation method. Using the ODI as an anchor, the MCID threshold for the TUG Test ranged from 0.9 to 4.8 s depending on the computation method. Using the RMDI as an anchor, the MCID threshold for the TUG Test time ranged from 1.5 to 5.1 s depending on the computation method. Using the EQ5D index as an anchor, the MCID threshold ranged from 1.7 to 6.0 s depending on the computation method. Finally, using the SF-12 PCS as an anchor, the MCID threshold for the TUG Test time ranged from 2.0 to 3.4 s depending on the computation method. For all anchors except the SF-12 PCS, the largest threshold was always derived from the minimum detectable change approach, whereas the smallest threshold was always derived from the change difference approach. The mean MCID for all methods and for all measured PROs was 3.4 s.

DISCUSSION

This study aimed to determine the MCID of the TUG Test time as a measure of OFI in lumbar DDD. An average of

TABLE 2. Three MCID Computation Methods Using 6 Different Anchors Including Pain Intensity (VAS Back and Leg Pain Scores), Functional Impairment (ODI and RMDI), and Health-Related Quality of Life (EQ5D Index, SF-12 PCS)[a]

Anchor-Based Approach	Back Pain	Leg Pain	ODI	RMDI	EQ5D Index	SF-12 PCS	Overall
Method (1) Average change (secs)	4.0	3.8	3.8	4.1	4.5	2.0	3.7
Method (2) MDC (95% CI) (secs)	5.3	4.8	4.8	5.1	6.0	3.4	4.9
Method (3) Change difference secs	1.3	1.3	0.9	1.5	1.7	2.1	1.5
Overall secs	3.5	3.3	3.2	3.6	4.1	2.5	**3.4**

[a]CI, confidence interval; EQ5D, EuroQol-5D; HRQoL, health-related quality of life; MCID, minimum clinically important difference; MCS, Mental Component Summary; MDC, minimum detectable change; ODI, Oswestry Disability Index; PCS, Physical Component Summary; RMDI, Roland-Morris Disability Index, SF-12, Short Form 12; VAS, Visual Analog Scale.

TABLE 3. Case Presentations Mentioned in the Discussion Section[a]			
	Preoperative	**D3 FU**	**W6 FU**
Patient 1 (44 y, female, microdiscectomy L4-5)	VAS back pain score 6; VAS leg pain score, 6; RMDI, 19; ODI, 52%; SF-12 PCS, 21.2; TUG Test, 12.53 s (z score), 2.75 (t score), 127.50, mild OFI	VAS back pain score, 0; VAS leg pain score, 0; RMDI, 2; ODI, 6%; SF-12 PCS, 42.4; TUG Test, 7.53 s (z score, 0.64; t score, 106.36; no OFI)	VAS back pain score, 0; VAS leg pain score, 0; RMDI, 2; ODI, 6%; SF-12 PCS, 42.4; TUG Test, 4.54 s (z score, −0.63; t score, 93.69; no OFI)
Patient 2 (70 y, male, decompression L3-4 with undercutting)	VAS back pain score, 2; VAS leg pain score, 4; RMDI, 10; ODI, 44%; SF-12 PCS, 33.9; TUG Test, 18.36 s (z score, 5.22; t score, 152.25; moderate OFI)	VAS back pain score, 2; VAS leg pain score, 0; RMDI, 17; ODI, 76%; SF-12 PCS, 35.2; TUG Test, 30 s (z score, 10.16; t score, 201.57; severe OFI)	VAS back pain score, 1; VAS leg pain score, 1; RMDI, 13; ODI, 50%; SF-12 PCS, 35.2; TUG Test, 7.83 s (z score, 0.76; t score, 107.63; no OFI)

[a]D3, day 3; FU, follow-up; ODI, Oswestry Disability Index; OFI, objective functional impairment; PCS, Physical Component Summary; RMDI, Roland-Morris Disability Index, SF-12, Short Form 12; TUG Test, Timed Up and Go Test; VAS, Visual Analog Scale; W6, week 6.

3 different computation methods, based on the MCIDs of established PROs including pain intensity, functional impairment, and HRQoL as anchors, was used. This approach revealed an average MCID for the TUG Test time of 3.4 s. In other words, an improvement of 3.4 s translates into clinically meaningful improvement in function in patients with lumbar DDD. The MCID is an important concept, as the absolute numeric results of PROs assessing pain intensity, functional impairment, and HRQoL may not correlate well with the patient's perception of clinical improvement. The MCID reflects the threshold of a clinically meaningful improvement (or worsening) for the patient. The change in t score of the TUG Test result in responders was −31.00, with an SD of 21.98 and a range of −87.50 to −14.41, indicating a significant improvement in function in patients who achieve an improvement by 3.4 s. The change in t score in nonresponders was −5.19, with an SD of 8.77 and a range of −14.32 to 47.07. We believe that the obtained MCID values confirm our personal experience from patient FU in clinics.

Two clinical examples from our collected data are presented to demonstrate the clinical use of the TUG Test and OFI (Table 3).

Case 1

A 44-year-old female patient is scheduled for a lumbar microdiscectomy at L4-5. The patient's clinical and neurological examination does not reveal a motor deficit, but does reveal significant disability in terms of VAS back and leg pain (both 6/10) scores and remarkable impairment in function and HRQoL. Her preoperative TUG Test time is 12.53 s (z score of 2.75 and t score of 127.50). Thus, she is classified as having mild OFI.[7] Surgery is uneventful, and on D3, her D3 TUG Test time is 7.53 s, which is 5 s faster than preoperatively. At D3, this patient can already be considered a responder to treatment.

Case 2

A 70-year-old male patient is scheduled to undergo selective decompression with undercutting at L3-4 for degenerative lumbar spinal stenosis. He presents without motor deficit, but with classic neurogenic claudication. His preoperative TUG Test time is 18.36 s (z score of 5.22 and t score of 152.25), which classifies him as having moderate OFI.[7] Subsequent surgery is performed without complications, but on D3, his TUG Test time is 30 s. At this early point in time/FU, the patient clearly is a nonresponder, if not a negative responder, in terms of TUG Test metrics. In addition, the PROs confirm increased disability (Table 3). We saw the patient again at the W6 FU in our outpatient clinic. His W6 TUG Test time was 7.83 s (z score of 0.76 and t score of 107.63), which is an improvement of 10.53 s compared with baseline. Using the MCID of 3.4 s, this patient can now be considered a treatment responder in terms of TUG Test metrics at W6, although the PROs still show some impairment.

Even though the mean MCID is 3.4 s, we noticed a large range of 0.9 to 6.0 s depending on the respective subjective scale and computation method used (Table 2). This is in line with the literature, which reports a large variability of reported MCIDs due to a variety of methods used to calculate MCIDs. This range of values stems from the fact that there is no gold standard measure of change, and thus many different theories exist on how to best estimate and measure the concept of change. Currently, there is no consensus on which is the superior method.[1] The previously mentioned range suggests that the true value of the change lies somewhere between 0.9 and 6 s, but the greatest likelihood is that it should lie at ~3.4 s. Given this range, it is feasible that, on further research, the MCID may change as estimates improve in accuracy or a single method becomes accepted. The application of confidence

intervals to the range of MCIDs presented here is not appropriate because the many MCIDs generated do not represent sampling distributions.

We would like to take the opportunity to emphasize that that the anchors used (gold standard) were based on PROs, and PROs measure different dimensions than the objective outcome measures used in this study. Other methods, such as using the SE of measurement, may be useful to validate the previously mentioned MCIDs, and further work on this aspect is ongoing.

TUG Test time measurements after surgery is an important issue that should be mentioned. The variation of ∆TUG 3 days after surgery was quite large for the current sample of 100 patients: the mean change in TUG was −0.66 s, and the range was −20.3 to 18.5 s. Applying the MCID of 3.4 s, only 14% of patients would be considered to be responders to surgery at D3 on the objective scale using the TUG Test. The responder rate early after surgery may be higher when subjective scales are used. Our current data show responder rates ranging from 8% (SF-12 PCS), 34% (RMDI), 49% (ODI), 54% (VAS back pain score) up to 79% (VAS leg pain score) as determined by the MCID of each scale, respectively. We concluded that on D3, patients clearly experience relief from preoperative irradiating pain and/or are already experiencing some extent of neurological recovery. The psychological effect of having undergone successful surgery may further factor into the patient's subjective perception of pain. The TUG Test here demonstrates what physicians and health care workers experience in the daily care of patients after lumbar spine surgery: despite feeling better, walking ability and mobility are still greatly constrained. Thus, a clinically meaningful change in the TUG Test metric is rather difficult to obtain in the early postoperative period and should be reserved for medium- and long-term FU. At the W6 FU assessment, the mean change in TUG Test time was −3.5 s, with a range of −20.6 to 11.1 s. According to the average MCID of 3.4 s, 38% of the current sample would be considered treatment responders on the TUG Test, whereas 62% of patients would be classified as nonresponders. Comparing this with the subjective scales, on which W6 FU responder rates ranged from 29% (SF-12 PCS) to 82% (VAS leg pain score), may indicate that recovery of measurable function is still in process at W6. Long-term data will shed more light on the dynamics of recovery of patients with DDD on the TUG Test metric.

CONCLUSION

The MCID for the TUG Test time is highly variable depending on the computation technique used. Using all 3 methods and all anchors, the suggested TUG MCID to be used in future studies is 3.4 s. The obtained MCID values confirm our personal experience of clinically meaningful recovery from patient FU in clinical daily routine.

Disclosures

REFERENCES

1. Parker SL, Godil SS, Shau DN, Mendenhall SK, McGirt MJ. Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion: clinical article. *J Neurosurg Spine*. 2013;18 (2):154-160.
2. Vavken P, Ganal-Antonio AK, Shen FH, Chapman JR, Samartzis D. Fundamentals of clinical outcomes assessment for spinal disorders: study designs, methodologies, and analyses. *Glob Spine J*. 2015;5(2):156-164.
3. 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*. 2014;14(10):2521-2522.
4. Gautschi OP, Corniola MV, Joswig H, et al. The timed up and go test for lumbar degenerative disc disease. *J Clin Neurosci*. 2015;22(12):1943-1948.
5. Barzilay Y, Noam S, Meir L, et al. Assessing the outcomes of spine surgery using global positioning systems. *Spine (Phila Pa 1976)*. 2011;36(4): E263-E267.
6. Mobbs RJ, Phan K, Maharaj M, Rao PJ. Physical activity measured with accelerometer and self-rated disability in lumbar spine surgery: a prospective study. *Glob Spine J*. 2015 [Epub ahead of print].
7. Gautschi OP, Smoll NR, Corniola MV, 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*. 2015 [Epub ahead of print].
8. 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*. 2008;8(6): 968-974.
9. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4): 407-415.
10. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris back pain questionnaire: part 1. *Phys Ther*. 1998;78(11): 1186-1196.
11. Ostelo RW, Deyo RA, Stratford P, 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)*. 2008;33(1): 90-94.
12. Parker SL, Mendenhall SK, Shau D, 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*. 2012;16 (1): 61-67.
13. Parker SL, Mendenhall SK, Shau DN, et al. Minimum clinically important difference 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*. 2012;16(5):471-478.

Acknowledgments

COMMENT

In this manuscript, the authors examine the minimum clinically significant difference (MCID) in the timed-up-and-go (TUG) test in lumbar spine surgery. One hundred patients undergoing a variety

of lumbar spine surgeries were examined. Standard patient-reported outcome (PRO) measures, as well as the TUG test, were gathered preoperatively and at intervals postoperatively. Using various PROs as anchors, and using 3 different methodologies, the MCID for the TUG test was calculated.

The results showed a mean MCID for the TUG test of 3.4 seconds. Values ranged from 0.9 seconds to 6.0 seconds, highlighting the spectrum of results that can be obtained with the various anchors and methodologies.

These results provide further insight into the potential clinical usefulness of the TUG test in lumbar spine surgery. The authors have also illustrated in their analysis the heterogeneity of calculated MCID values, even using sound statistical methods. This heterogeneity speaks to the need for an agreed-upon methodology for generating these important parameters, especially in the current era of value-based care.

Chris Wolfla
Milwaukee, Wisconsin