Distance to first symptoms measured by the 6-min walking test differentiates between treatment success and failure in patients with degenerative lumbar disorders

Anna M. Zeitlberger¹ · Marketa Sosnova¹ · Michal Ziga¹ · Oliver P. Gautschi³ · Luca Regli⁴ · Oliver Bozinov¹ · Astrid Weyerbrock¹ · Martin N. Stienen^{1,4} · Nicolai Maldaner^{1,2,4}

Received: 21 February 2021 / Revised: 12 December 2021 / Accepted: 21 December 2021 / Published online: 11 January 2022
© The Author(s) 2022

Abstract

Purpose The smartphone-based 6-min walking test (6WT) is an established digital outcome measure in patients undergoing surgery for degenerative lumbar disorders (DLD). In addition to the 6WTs primary outcome measure, the 6-min walking distance (6WD), the patient's distance to first symptoms (DTFS) and time to first symptoms (TTFS) can be recorded. This is the first study to analyse the psychometric properties of the DTFS and TTFS.

Methods Forty-nine consecutive patients (55 ± 15.8 years) completed the 6WT pre- and 6 weeks (W6) postoperative. DTFS and TTFS were assessed for reliability and content validity using disease-specific patient-reported outcome measures. The Zurich Claudication Questionnaire patient satisfaction subscale was used as external criterion for treatment success. Internal and external responsiveness for both measures at W6 was evaluated.

Results There was a significant improvement in DTFS and TTFS from baseline to W6 ($p < 0.001$). Both measures demonstrated a good test–retest reliability ($\beta = 0.86$, 95% CI 0.81–0.90 and $\beta = 0.83$, 95% CI 0.76–0.87, both $p < 0.001$). The DTFS exceeded the 6WD capability to differentiate between satisfied (82%) and unsatisfied patients (18%) with an AUC of 0.75 (95% CI 0.53–0.98) vs. 0.70 (95% CI 0.52–0.90). The TTFS did not demonstrate meaningful discriminative abilities.

Conclusion Change in DTFS can differentiate between satisfied and unsatisfied patients after spine surgery. Digital outcome measures on the 6WT metric provide spine surgeons and researchers with a mean to assess their patient's functional disability and response to surgical treatment in DLD.

Keywords 6-min walking test · 6WT · Functional self-assessment · Objective functional impairment · Distance to first symptoms

Abbreviations

6WT 6-minute walking test
6WD 6-minute walking distance
app Application

CI	Confidence interval
COMI	Core outcome measures index
DLD	Degenerative lumbar disorders
hrQoL	Health-related quality of life
ICC	Intraclass correlation coefficient
LDH	Lumbar disc herniation
LSS	Lumbar spinal stenosis
MCID	Minimum clinically important difference
m	Metres
PROMs	Patient-reported outcome measures
SPWT	Self-paced walking test
SRM	Standardized responsive mean
VAS	Visual analogue scale
ZCQ	Zurich Claudication Questionnaire

✉ Nicolai Maldaner
nicolai.maldaner@usz.ch

¹ Department of Neurosurgery, Clinical Neuroscience Center, Kantonsspital St. Gallen, St. Gallen, Switzerland

² University Hospital Zurich, University of Zurich, Frauenklinikstrasse 10, 8091 Zurich, Switzerland

³ Neuro- and Spine Center, Hirslanden Clinic St. Anna, Lucerne, Switzerland

⁴ Department of Neurosurgery, Clinical Neuroscience Center, University Hospital Zurich, University of Zurich, Zurich, Switzerland

Introduction

The management of patient with degenerative lumbar disorders (DLD) requires reliable measures of functional impairment. Today, patient-reported outcome measures (PROMs) are used as a gold standard for the outcome assessment in spine surgery [1–3]. Apart from subjective PROMs, objective measures of function find increasing attention in research and clinical practice as they help to monitor and compare treatment results over time and across populations [4].

The 6-min walking test (6WT) is increasingly applied as an objective outcome measure in patients with lumbar DLD [5]. We recently developed a smartphone app-based version of the 6WT, which demonstrated excellent reliability [6, 7]. The 6WT can be self-performed by the patient in his/her home environment, and results are standardized with respect to age and sex [6, 8, 9]. By providing the ability to monitor patients from afar, digital outcome measures are invaluable tools for physicians and patients. In a recent study, three out of four patients favoured the smartphone-based 6WT over traditional paper-based PROMs for the assessment of spine-related symptoms [10]. This is a trend that will only accelerate in a time when a global pandemic hampers avoidable physical “face-to-face” consultations as those might endanger elderly or particularly fragile patients [11, 12].

The 6WT's primary outcome measure is the maximum distance a patient can walk within 6 min (6WD = 6-min walking distance, measured in metres) [6]. In addition, the 6WT-app provides patients with the possibility to push a “flash button” and records both the time (TTFS = time to first symptoms, measured in seconds) and distance (DTFS = distance to first symptoms, measured in metres) when first symptoms of neurogenic claudication appear. While the 6WD expresses the result of walking restrictions that occur over the whole duration of the test time, TTFS and DTFS may give more granular information about the severity or urgency of patient's symptoms. Studies have proven the 6WD to be a reliable, valid, and responsive measure of functional impairment [6]. The added value of both TTFS and DTFS in addition to the 6WD, however, is yet unclear.

This study aims to analyse the psychometric properties of DTFS and TTFS as determined by the 6WT. We hypothesize that the pre- to postoperative change in both measures may help to differentiate between treatment successes in patients with lumbar DLD and compare the responsiveness to the traditional 6WD outcome.

Methods

All adult patients with lumbar DLD scheduled for elective spine surgery between May 2019 and March 2020 with one of the following diagnosis (1) lumbar disc herniation (LDH), (2) lumbar spinal stenosis lumbar (LSS) or (3) DLD with or without instability requiring lumbar fusion were prospectively screened for study enrolment at the XX, XX, XX. A detailed prescription of the app-based outcome measures and PROMs used in this study is provided as Online Resource 1.

Inclusion and exclusion criteria for the study cohort

Patients fulfilling all of the following inclusion criteria were considered for this study:

- Male or female subject ≥ 18 years;
- Written informed consent.

Patients were not enrolled if any of the following exclusion criteria were met:

- Pregnancy;
- Inability to walk (extreme pain or severe neurological deficits);
- Severe chronic obstructive lung disease (COPD) corresponding to $\geq$ Gold III;
- Severe heart failure corresponding to $\geq$ NYHA III;
- Lung cancer and diffuse parenchymal lung disease;
- Other medical reasons interfering with the patient's ability to walk and perform the 6WT (e.g. osteoarthritis disease of the lower extremities, Parkinson's disease, heart failure, hip or knee prosthesis, peripheral artery disease causing intermittent claudication, etc.);
- Unavailability for follow-up and/or inability to complete assessment (planning to move, no smartphone, etc.).

The 6WT-app

The 6WT-App measures the maximum distance (in m) walked within six minutes (6WD) using global positioning system (GPS) coordinates, which is the primary test result [13]. Both distance walked and time elapsed are continuously displayed on the screen, while the 6WT is conducted. Patients were instructed to press a “flash” button on the app's user interface in case of appearance and/or first-time significant aggravation in leg or back pain during the test. This marks their time (= TTFS in sec) and walking distance (= DTFS in m) to first symptoms (Fig. 1). Patients were instructed to continue walking until the six minutes have elapsed, whenever possible. Completed measurements

Fig. 1 The 6-min walking test (6WT) smartphone application starts measuring the walking distance based on GPS coordinates once the “start” button is pressed. Patients are instructed to press a “flash” button on the app’s user interface in case of appearance and/or first-time significant aggravation in leg or back pain during the test. This records their time (= TTFS in sec) and walking distance (= DTFS in m) to first symptoms. Patients are instructed to continue walking until the six minutes have elapsed, whenever possible. Completed measurements are saved on the patient’s smartphone device with a date and time stamp and may be transferred to a secure online database

are saved on the patient’s smartphone device with a date and time stamp and may be transferred to a secure online database.

Data collection and PROMs

All patients underwent detailed both subjective and objective (6WT) assessments before surgery, as well as six weeks postoperatively (W6). The subjective assessment included the following PROMs:

1. The Visual analogue scale (VAS) for pain
2. The Zurich Claudication Questionnaire (ZCQ), with its two main scores:
 - a. ZCQ symptom severity (ZCQ SS)
 - b. ZCQ physical function (ZCQ PF)
3. The Core Outcome Measures Index (COMI) Back

Ethical considerations

The study was approved by the local ethic committee (XX, EKOS–2019–01,209) and was registered (<http://clinicaltrials.gov> identifier: XX). All patients provided written informed consent prior to initiation of the data collection.

Statistical considerations

Data are presented as mean $\pm$ standard deviation (SD) for continuous and count (percentage) for categorical variables. 6WT results are the raw 6WD (in m) as well as the DTFS (in m) and the TTFS (in s). We present the percentage (%) of patients who experienced the appearance and/or significant aggravation in leg or back pain during their pre- as well as postoperative 6WT. In case a patient did not indicate symptoms in the app, the DTFS was defined as corresponding to the 6WD of the same run and TTFS as the maximum walking time which is 360 s = 6 min.

According to Shrout and Fleiss, the intraclass correlation coefficient (ICC) was used to determine the agreement of repeated pre- and postoperative 6WT measurements [14]. ICC was deemed good (ICC between 0.75–0.9) or excellent (ICC > 0.9) in accordance with prior research [15]. Standard error of measurement (SEM) was calculated as the SD multiplied by the square root of 1 minus the intrarater ICC. The SEM represents a “grey zone” of uncertainty between two patient scores demonstrated by Stratford & Goldsmith [16].

Paired-sample *t*-tests were calculated to evaluate the changes between pre- and postoperative outcomes. Pearson correlation coefficients (*r*) were used to define the relationship between pre- and postoperative 6WT results and subjective outcome measures (PROMs).

The internal responsiveness of the 6WT results was assessed using standardized effect size (standardized response mean (SRM) = mean score of change from baseline to follow-up, divided by the SD of the score change). In accordance with prior research, SRM values were deemed as small (> 0.20), moderate (> 0.50), or large (> 0.80).

The external responsiveness of 6WT results was evaluated using receiver operating characteristics (ROC) curves. A reference standard indicating successful versus unsuccessful treatment was created by grouping results of the ZCQ patient satisfaction subscale (range of 1–4) into a binary variable of satisfied (combined scores ≤ 2 , including 1 = “completely satisfied” to 2 = “somewhat satisfied”) versus dissatisfied (combined score > 2, including 3 = “somewhat dissatisfied” to 4 = “completely dissatisfied”). External responsiveness determined the probability that the pre- to postoperative change in 6WT result correctly classified patients who were satisfied or unsatisfied with the treatment result. An area under the curve (AUC) of 0.5 indicates no

discrimination (no better than chance), whereas an AUC of 1.0 indicates perfect discrimination [17].

Analyses were carried out using “R version 3.6.3” for Mac (R Core Team, 2020, RStudio: Integrated Development for R. RStudio, Inc., Boston, Massachusetts, <http://www.rstudio.com/>). *P* values < 0.05 were considered significant.

Results

Study cohort

A total of 50 consecutive patients undergoing surgery for DLD were enrolled in this study. One patient dropped out due to incomplete follow-up assessments. The final analysis therefore included 49 patients (41% female) with a mean age of 55.5 ± 15.8 years. Table 1 summarizes demographic and clinical variables of the study population.

Indication of symptoms

Table 2 displays the number of patients who experienced the appearance and/or significant aggravation in leg or back pain during their pre- as well as postoperatively 6WT. Out of 35 patients who experienced symptoms preoperative, 21 (60%) no longer indicated symptoms in the 6WT postoperatively.

Test–retest reliability

For pre- and postoperative 6WD values, ICC was good ($\beta = 0.82$, 95% CI 0.75–0.87, $p < 0.001$), with a SEM of 58 m. ICC was similar for DTFS values ($\beta = 0.83$, 95% CI 0.77–0.88, $p < 0.001$, SEM = 85 m) and for TTFS values ($\beta = 0.79$, 95% CI 0.72–0.85, $p < 0.001$, SEM = 59 s).

Pre- and 6 weeks postoperative results

Table 3 contains the mean scores for each subjective and objective outcome measure at time points preoperative and postoperative W6. There was a significant ($p < 0.001$) improvement in each subjective and objective outcome measure from baseline to W6. The 6WD improved by 94 m (SD 109 m), DTFS improved by 205 m (SD 218 m) and TTFS improved by 112 s (SD 134 s).

Convergent validity

Correlation coefficients between pre- or postoperative 6WT values and PROMs are outlined in Table 4. Changes in the 6WD and DTFS inversely correlated with changes in PROMs. Changes in the TTFS showed a generally weaker inverse correlation with PROMs. 6WD shows a stronger correlation with DTFS compared to TTFS.

Table 1 Patient baseline characteristics

	Study group
Age in years	55.5 (15.8)
Sex	
Male	29 (59%)
Female	20 (41%)
Body dimensions	
Height in cm	169.8 (9.7)
Weight in kg	80.6 (22.0)
BMI in kg/m ²	27.8 (6.8)
Working status	
Full-time	15 (31%)
Part-time	6 (12%)
Retired	18 (37%)
Disabled	10 (20%)
Prior spine surgery	
Yes	15 (31%)
No	34 (69%)
Smoking status	
Smoker	11 (22%)
Non-smoker	38 (78%)
Operation	
Microdiscectomy	21 (43%)
Uni- or bilateral decompression	20 (41%)
Decompression and fusion	8 (16%)
ASA risk scale	
1	23 (46.9%)
2	22 (44.9%)
3	4 (8.2%)
CCI	
0	32 (65.3%)
1	13 (26.5%)
2	4 (8.2%)
	N = 49 (100%)

Results are presented as count (%) or mean (standard deviation). ASA American Society of Anesthesiologists; BMI body mass index; CCI Charlson comorbidity index

Table 2 Indication of pre- and 6 weeks postoperative (W6) DTFS/TTFS during the 6WT

		Postoperative		Total
		No	Yes	
Preoperative	No	13	1	14 (29)
	Yes	21	14	35 (71)
Total		34 (69)	15 (31)	49 (100)

Results are expressed as count (%) of patients indicating the appearance and/or first-time significant aggravation in leg or back pain during their pre- or postoperative 6WT

Table 3 Pre- and 6 weeks postoperative (W6) subjective and objective outcome measures

Outcome measure	Preoperative	Postoperative W6	P value
6WT			
6WD (in m)	401 (129)	495 (129)	<0.001
DTFS (in m)	227 (161)	431 (194)	<0.001
TTFS (in s)	192 (127)	304 (102)	<0.001
ZCQ			
ZCQ SS	3.15 (0.62)	2.20 (0.72)	<0.001
ZCQ PF	2.38 (0.60)	1.48 (0.50)	<0.001
ZCQ SAT	–	1.53 (0.52)	–
COMI			
COMI back	7.2 (1.8)	4.0 (2.5)	<0.001
VAS			
Back pain	4.6 (3.0)	2.5 (2.0)	<0.001
Leg pain	6.8 (2.7)	2.2 (2.0)	<0.001

Results are expressed as mean (SD). 6WD 6-min walking distance; COMI Core outcome measures index; VAS Visual analogue scale; ZCQ Zurich Claudication Questionnaire; ZCQ SAT, ZCQ patient satisfaction subscale; ZCQ PF, ZCQ physical function; ZCQ SS, ZCQ symptom severity

Table 4 Convergent validity: Pearson correlation [95% CI] of pre- and postoperative 6WD, DTFS and TTFS measurements

Variable	6WD	DTFS	TTFS
6WD			
DTFS	0.64** [0.51, 0.75]		
TTFS	0.25* [0.06, 0.43]	0.84** [0.76, 0.89]	
ZCQ PF	–0.57** [–0.69, –0.42]	–0.64** [–0.74, –0.51]	–0.51** [–0.65, –0.35]
ZCQ SS	–0.49** [–0.63, –0.33]	–0.52** [–0.65, –0.36]	–0.39** [–0.54, –0.20]
COMI	–0.51** [–0.64, –0.35]	–0.59** [–0.71, –0.44]	–0.44** [–0.59, –0.26]
VAS back	–0.52** [–0.65, –0.36]	–0.48** [–0.62, –0.31]	–0.33** [–0.50, –0.14]
VAS leg	–0.51** [–0.64, –0.34]	–0.58** [–0.69, –0.43]	–0.43** [–0.58, –0.25]

6WD 6-min walking distance; COMI Core outcome measures index; DTFS Distance to first symptoms; n/a Not applicable; TTFS Time to first symptoms; VAS Visual analogue scale; ZCQ Zurich Claudication Questionnaire; ZCQ PF, ZCQ physical function; ZCQ SS, ZCQ symptom severity; *Indicates $p < 0.05$. **Indicates $p < 0.01$

Responsiveness

Internal responsiveness analysis showed the highest SRM for DTFS (0.94) followed by 6WD (0.86) and TTFS (0.84). Based on the ZCQ patient satisfaction subscale, 40 (82%)

patients in our cohort were identified as responders to surgery at W6. Evaluation of external responsiveness revealed that the change in DTFS differentiated better between satisfied (82%) and unsatisfied patients (18%) than the change in 6WD with an AUC of 0.75 (95% CI 0.53–0.98) vs. AUC of 0.70 (95% CI 0.52–0.90; Fig. 2). Change in TTFS did not demonstrated meaningful differential capabilities (AUC = 0.59, 95% CI 0.34–0.83).

In a subgroup analysis examining the 14 patients who indicated symptoms both during their pre- as well as postoperative 6WT, external responsiveness revealed that the change in DTFS similarly had a greater capability to differentiate between satisfied (10 patients) and unsatisfied patients (4 patients) than the change in 6WD with an even greater AUC of 0.88 (95% CI 0.68–1.00) vs. AUC of 0.58 (95% CI 0.14–1.00; Online Resource 2: Suppl. Figure 1).

Discussion

This study analysed whether two sub-scores of the smart-phone-based 6WT, namely DTFS and TTFS, provide additional information to the main outcome score (6WD). The rationale for the study was that the 6WD as main score is somewhat insensitive towards symptoms onset and symptom severity. As long as patients can “tough it out”, they might be able to reach a high 6WD despite an early onset of back and leg pain, whereas both DTFS and TTFS account better for these aspects. Several interesting findings emerged. First, the study population demonstrated significant improvements not only in the overall 6WD, but also with respect to DTFS and TTFS at W6 postoperative. Secondly, both DTFS and TTFS showed strong convergent validity with each other and the 6WD, but only weak to moderate correlation with PROMs. Lastly, DTFS exceeded the capability of the main outcome score (6WD) to differentiate between satisfied and unsatisfied patients after surgery.

Restricted walking distance due to pain and/or neurological deficits is one of the most reported and disabling symptoms related to lumbar DLD. While the 6WD is an already established and thoroughly validated outcome measure [5, 6], we here analyse two new measures that might account for different aspects of a patient’s functional impairment. How far a patient can walk until symptoms impede continuation is a question routinely posed during spinal consultation with the intention to quantify impairment as much as possible. There is a solid body of evidence demonstrating that patients have difficulties estimating this distance and hence answering this question [18, 19]. Disabling symptoms may start to aggravate considerably earlier during the physical activity and may not necessarily result in a discontinuation or even slower walking in every patient. While in the 6WD, the overall walking distance covered is used as an absolute

Fig. 2 ROC curves for 6WD change (solid lines) and DTFS changes (dotted lines) in the 6WT. The DTFS exceeded the 6WD capability to differentiate between satisfied (82%) and unsatisfied patients (18%) with an 0.75 (95% CI 0.53–0.98) vs. AUC of 0.70 (95% CI 0.52–0.90)

measure of functional impairment, the DTFS and TTFS are measures that indicate the distance and time of their first symptoms which can be of significant relevance in the daily life of a patient.

As shown in this study, not every patient may experience the appearance or significant aggravation of symptoms during their 6WT. This is especially true postoperatively in case of a successful surgery. In fact, 60% of patients in our study cohort who did experience symptoms preoperatively no longer indicated symptoms postoperative. In order to be able to also analyze patients who did not (longer) experience symptoms, we opted for a pragmatic approach to define the DTFS as corresponding to the 6WD of the same run and TTFS as 360 s in case of missing symptoms. While this may generate a ceiling effect for the TTFS, this effect is similar to most PROM questionnaires [4].

Reliability and validity

The reliability of repeated DTFS and TTFS self-measurements proved to be good. Overall, the results showed a moderate inverse correlation of changes in walking distances (6WD and DTFS) to subjective measures, indicating that patients with more pain and/or disability walked shorter distances without relevant symptoms. In line with

our findings, previous studies have shown a weak to moderate correlation of different objective outcome measures, like the timed up-and-go (TUG) test or the motorized treadmill test (MTT), with PROMs [20, 21]. Subjective and objective assessments in lumbar DLD patients do not seem to always align in a linear fashion, indicating that the 6WT is not a mere objectification of PROM questionnaires and should be considered a separate dimension in the outcome assessment of spine patients [6]. Interestingly, the TTFS did demonstrate a generally weaker correlation with PROMs which might be explained by the fact that the TTFS does not consider walking speed. A patient with high functional impairment may, as a result of the disability, walk slower but trigger the app button for TTFS later. The findings suggest that both 6WD and DTFS may be more relevant and directly related to a patient's disability than the time span until symptoms exaggerate.

External responsiveness and prediction of treatment success

External responsiveness reflects on the relevance of a detected change compared to the overall change in a patient's clinical status. Changes in the DTFS exceeded the capability of the 6WD to differentiate between satisfied

and unsatisfied patients with AUCs of 0.75 vs 0.70, respectively. On the contrary, in line with a weaker correlation with PROMs, change in TTFS did not demonstrate a meaningful capability to differentiate between patient response and surgical treatment. This indicates that the distance a patient can ambulate without a noticeable aggravation in symptoms may be more relevant for treatment satisfaction than the overall distance he/she can ultimately walk within a certain time frame.

Interestingly, similar objective tests quantifying the patients walking distance such as the MTT or the self-paced walking tests (SPWT) previously failed to demonstrate external responsiveness in patients with lumbar spinal stenosis [22]. Contrary to the 6WT, both measures are designed with no time limit to test a patient's maximum physical capability. Given this lack of responsiveness to perceived treatment success, one can speculate that the patient's maximum physical capacity does not represent a measure most relevant to the patient.

Two additional advantages of the smartphone-based 6WT are that it does not require additional equipment, such as a treadmill, and that the patient may walk in a familiar setting without being accompanied by a health-care professional. Both factors have previously been reported to influence the walking distance [22–24]. In turn, the smartphone-based 6WT may lead to a more accurate measurement of the individual's disability in a real-world scenario.

Future prospects

Our study indicates that the DTFS, in contrast to the TTFS, might aid in the discrimination of satisfied and unsatisfied DLD patients after surgery. However, the cohort in our study included individuals with different spinal pathologies. While these pathologies share common characteristics, usually a form of mobility restriction resulting from back and/or leg pain, the predominant complaint (back pain vs. radicular leg pain vs. neurogenic claudication) often differs for specific spinal pathologies. An approach like the present one where we included a range of degenerative spinal pathologies may neglect these differences. In future studies, we therefore aim to include larger patient cohorts with specific diseases, which will allow to analyse fine differences in test responsiveness and validity for spinal pathologies separately. The TTFS, for instance, may show a greater responsiveness in LSS than in LDH as symptoms of neurogenic claudication generally do not start at the beginning of an exercise, whereas patients with lumbar radicular pain typically report pain immediately after mobilization. We are confident that the growing availability and usage of smart devices even among the elderly will help us to increasingly apply digital objective outcome assessments in spine patients.

Strengths and limitations

The strengths of this study lie in its prospective design and the comprehensive patient evaluation using the 6WT, the two innovative sub-scores DTFS and TTFS, as well as several well-validated PROMs. The main limitation is the relatively small sample size which might limit the implication for other patient cohorts. However, existing studies on objective outcome measures mostly analyze small patient cohort consisting of as few as < 20 patients which we exceed by far [5]. As mentioned before, our results also need to be interpreted in light of a heterogeneous patient cohort with various degenerative spinal pathologies. Secondly, we determined the ability of the 6WT's DTFS and TTFS to change over the prespecified time frame of 6 weeks. Recovery may still continue 6 weeks after surgery, especially for some patients who underwent multi-level surgery or fusion procedures. Further long-term data will have to shed more light on the dynamics of postoperative recovery as measured by an objective test like the 6WT and its effect on the responsiveness of different 6WT metrics.

Conclusions

The DTFS demonstrated both a higher external responsiveness and a better correlation with subjective outcome measures than the TTFS. Change in DTFS can differentiate between satisfied and unsatisfied patients after spine surgery. Digital outcome measures on the 6WT metric provide spine surgeons and researchers with a mean to assess their patient's functional disability and response to surgical treatment in DLD.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00586-021-07103-9>.

Funding Open access funding provided by University of Zurich. Dr. Maldaner was supported by the Forschungsförderung Kantonsspital St. Gallen, CTU 19/24.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability The “6WT” app is available as a free download from the Apple App Store and the Google Play Store (link provided in the supplemental material).

Declarations

Conflict of interest None of the authors has any potential conflict of interest.

Ethical approval The study was approved by the local ethic committee (Kantonale Ethikkommission, EKOS—2019–01209, <http://clinicaltrials.gov> identifier: NCT03977961).

Consent to participate All patients provided written informed consent prior to initiation of the data collection.

Previous presentation No parts of this manuscript have been presented previously.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Arnold PM, Robbins S, Paullus W et al (2009) Clinical outcomes of lumbar degenerative disc disease treated with posterior lumbar interbody fusion allograft spacer: a prospective, multicenter trial with 2-year follow-up. *Am J Orthop Belle Mead* 38(7):E115–122
2. McCormick JD, Werner BC, Shimer AL (2013) Patient-reported outcome measures in spine surgery. *JAAOS* 21(2):99
3. Lurie JD, Tosteson TD, Tosteson ANA et al (2014) Surgical versus non-operative treatment for lumbar disc herniation: eight-year results for the spine patient outcomes research trial (SPORT). *Spine* 39(1):3–16
4. Maldaner N, Stienen MN (2020) Subjective and objective measures of symptoms, function, and outcome in patients with degenerative spine disease. *Arthritis Care Res* 72(Suppl 10):183–199
5. Stienen MN, Ho AL, Staartjes VE et al (2019) Objective measures of functional impairment for degenerative diseases of the lumbar spine: a systematic review of the literature. *Spine J Off J North Am Spine Soc* 19(7):1276–1293
6. Maldaner N, Sosnova M, Zeitlberger AM et al (2020) Evaluation of the 6-minute walking test as a smartphone app-based self-measurement of objective functional impairment in patients with lumbar degenerative disc disease. *J Neurosurg Spine* 7:1–10
7. Tomic L, Goldberger E, Maldaner N et al (2020) Normative data of a smartphone app-based 6-minute walking test, test-retest reliability, and content validity with patient-reported outcome measures. *J Neurosurg Spine* 29:1–10
8. Zeitlberger AM, Sosnova M, Ziga M et al (2020) Smartphone-based self-assessment of objective functional impairment (6-minute walking test) in patients undergoing epidural steroid injection. *Neurospine* 17(1):136–142
9. Sosnova M, Zeitlberger AM, Ziga M et al (2020) Longitudinal smartphone-based self-assessment of objective functional impairment in patients undergoing surgery for lumbar degenerative disc disease: initial experience. *Acta Neurochir (Wien)* 162(9):2061–2068
10. Sosnova M, Zeitlberger AM, Ziga M et al (2020) Patients undergoing surgery for lumbar degenerative spinal disorders favor smartphone-based objective self-assessment over paper-based patient-reported outcome measures. *Spine J Off J North Am Spine Soc* 21:610–617
11. Debono B, Bousquet P, Sabatier P et al (2016) Postoperative monitoring with a mobile application after ambulatory lumbar discectomy: an effective tool for spine surgeons. *Eur Spine J* 25(11):3536–3542
12. Maldaner N, Tomkins-Lane C, Desai A et al (2020) Digital transformation in spine research and outcome assessment. *Spine J Off J North Am Spine Soc* 20(2):310–311
13. Stienen MN, Gautschi OP, Staartjes VE et al (2019) Reliability of the 6-minute walking test smartphone application. *J Neurosurg Spine* 13:1–8
14. Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86(2):420–428
15. Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 15(2):155–163
16. Stratford PW, Goldsmith CH (1997) Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. *Phys Ther* 77(7):745–750
17. Terwee CB, Bot SDM, de Boer MR et al (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60(1):34–42
18. Hurri H, Sainio P, Kinnunen H et al (2009) Walking distance as a measure of disability in lumbar spinal stenosis. *Orthop Proc* 91-B(SUPP_II):285–285
19. Tomkins-Lane CC, Battié MC (2010) Validity and reproducibility of self-report measures of walking capacity in lumbar spinal stenosis. *Spine* 35(23):2097–2102
20. Barz T, Melloh M, Staub L et al (2008) The diagnostic value of a treadmill test in predicting lumbar spinal stenosis. *Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc* 17(5):686–690
21. Stienen MN, Maldaner N, Sosnova M et al (2020) External validation of the timed up and go test as measure of objective functional impairment in patients with lumbar degenerative disc disease. *Neurosurgery* 88:E142–E149
22. Rainville J, Childs LA, Peña EB et al (2021) Quantification of walking ability in subjects with neurogenic claudication from lumbar spinal stenosis—a comparative study. *Spine J Off J North Am Spine Soc* 12(2):101–109
23. Swerts PM, Mostert R, Wouters EF (1990) Comparison of corridor and treadmill walking in patients with severe chronic obstructive pulmonary disease. *Phys Ther* 70(7):439–442
24. Moon E-S, Kim H-S, Park J-O et al (2005) Comparison of the predictive value of myelography, computed tomography and MRI on the treadmill test in lumbar spinal stenosis. *Yonsei Med J* 46(6):806–811

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.