European Spine Journal (2019) 28:665–673
https://doi.org/10.1007/s00586-019-05897-3

ORIGINAL ARTICLE

Can objective functional impairment in lumbar degenerative disease be reliably assessed at home using the five-repetition sit-to-stand test? A prospective study

Victor E. Staartjes[1,2,3] · **Femke Beusekamp**[1] · **Marc L. Schröder**[1]

Received: 30 October 2018 / Accepted: 17 January 2019 / Published online: 24 January 2019

Abstract

Purpose Objective functional tests like the five-repetition sit-to-stand test (5R-STS) can supplement an objective dimension to conventional patient-reported outcome measures. The reliability of unsupervised obtainment of 5R-STS performance is currently unknown.

Methods We included patients with degenerative pathologies of the lumbar spine. Patients performed the 5R-STS during the initial clinical visit (supervised), as well as at home after instruction by a physiotherapist. At home, patients were first timed by a relative (unsupervised) and subsequently produced a video recording of themselves performing the 5R-STS for digital measurement (telesupervised). Two raters independently assessed the recordings.

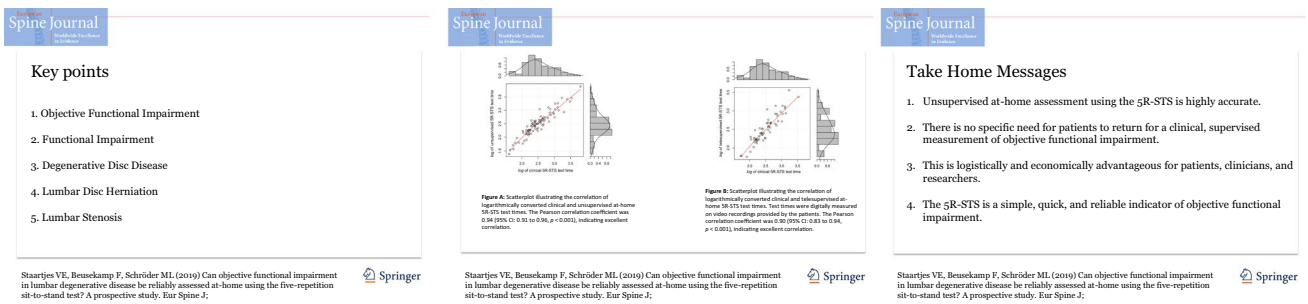
Results One hundred and twenty-one patients were recruited, of which 100 were eligible. Eighty-eight reported unsupervised results. Sixty-four returned recordings, of which 61 were ratable. Both unsupervised (r: 0.94, 95% CI 0.91–0.96, $p < 0.001$) and telesupervised (r: 0.90, 95% CI 0.83–0.94, $p < 0.001$) measurements demonstrated excellent correlation with clinical test times. Patients did not perform more slowly at home ($p > 0.05$). The interrater agreement for digital judgement of the telesupervised recording was excellent (ICC: 0.996, 95% CI 0.993–0.998, $p < 0.001$). We confirmed convergent validity with self-reported disability, back pain, and quality of life (all $p < 0.05$), but not with leg pain ($p = 0.189$).

Conclusions Unsupervised at-home assessment using the 5R-STS is highly reliable. There does not appear to be a specific need for patients to return for a supervised 5R-STS follow-up. Rather, instructions can be provided, and the test performed and rated by a partner or family member at home. This is logistically and economically advantageous for patients, clinicians, and researchers.

Trial registry number ClinicalTrials.gov Identifier: NCT03321357.

Graphical abstract

These slides can be retrieved under Electronic Supplementary Material.



Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00586-019-05897-3) contains supplementary material, which is available to authorized users.

Extended author information available on the last page of the article

Keywords Degenerative disc disease · Lumbar stenosis · Lumbar disc herniation · Functional impairment · Objective functional impairment

Introduction

Outcomes of interest have been shifting from radiographic indicators of success, such as fusion rates, to patient-oriented and value-based indicators of treatment success [1]. Meanwhile, the use of patient-reported outcome measures (PROMs) has become standard practice for the majority of clinicians and forms the basis of clinical research in spinal patient care [2–9]. These validated questionnaires provide accurate evaluations of pain severity, functional impairment, and health-related quality of life (HRQOL) [1, 2, 5, 10].

In the past decade, objective measures of functional impairment have attracted attention [11–17]. Some objective functional tests have been introduced and have found their way into clinical practice, such as the Six-Minute Walk Test (6MWT), Timed Up and Go (TUG) test, gait analysis and other biomechanical measurements, as well as accelerometer-based tools [12, 13, 15–20]. Their advantages include quick execution, repeatability, straightforward interpretation, and supplementation of an objective dimension different from traditional PROMs [18]. Patients also frequently prefer an objective functional test over questionnaires [21].

The five-repetition sit-to-stand test (5R-STS) has seen use for chronic obstructive pulmonary disease and Parkinson's disease, among others [22, 23]. Recently, Staartjes et al. [14] performed a formal validation for degenerative pathologies of the lumbar spine. Still, little is known on the 5R-STS in this population (Table 1) [24, 25]. While objective functional tests are simple to perform with instructions during a clinical visit, little is known about their remote and unsupervised use. PROM questionnaires are often mailed to patients for the initial assessment or follow-up. Only few studies have evaluated the possibility of remote sit-to-stand assessment, but employed motion tracking or sensor-based supervision [19, 20]. While highly accurate, these instruments are not routinely available to patients at home and to clinicians in daily practice.

Our aim was to assess the reliability of completely unsupervised and telesupervised at-home objective performance assessment in patients with degenerative spinal disorders using the 5R-STS in a prospective study.

Table 1 Published evidence on the 5R-STS test in patients with degenerative lumbar spinal pathologies

Parameter	Published evidence	References
Content validity		
Back pain [VAS]	$r = 0.250$ to 0.31	[14, 24], current report
Leg pain [VAS]	$r = 0.10$ to 0.189	[14], current report
Pain affect [VAS]	$r = 0.332$	[24]
Functional disability [RMDQ]	$r = 0.453$ to 0.49	[14, 24], current report
Functional disability [ODI]	$r = 0.40$ to 0.44	[14], current report
Health-related quality of life [EQ-5D index]	$r = -0.41$ to -0.24	[14], current report
Health-related quality of life [EQ-VAS]	$r = -0.30$ to -0.14	[14], current report
Reliability		
Interrater reliability [clinical]	ICC = 0.99	[24]
Interrater reliability [telesupervised]	ICC = 0.996	Current report
Test–retest reliability	ICC = 0.45 to 0.97	[14, 24]
Day-to-day reliability	ICC = 0.89	[24]
Upper limit of normal	10.4 s	[14]
Severity stratification of OFI [prevalence]	Not significant: ≤ 10.4 s [50%] Mild: 10.5 to 15.2 s [25%] Moderate: 15.3 to 22.0 s [15%] Severe > 22.0 s [10%]	[14]
Responsiveness	Unknown	–

r correlation coefficient, *ICC* intraclass correlation coefficient, *OFI* objective functional impairment, *RMDQ* Roland–Morris disability questionnaire, *ODI* Oswestry Disability Index, *VAS* visual analogue scale

Materials and methods

Population

All enrolled patients were scheduled for surgery and were assessed during outpatient consultations. Inclusion criteria were the presence of lumbar disc herniation (LDH), lumbar stenosis, lumbar spondylolisthesis, degenerative disc disease (DDD), or synovial facet cysts, requiring surgical treatment. Patients with hip or knee prosthetics and those requiring walking aides were excluded to eliminate these confounders.

Study course

Between December 2017 and June 2018, patients were seen at a specialized outpatient spine surgery clinic. The study was compiled according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [26]. This prospective study (ClinicalTrials.gov Identifier: NCT03321357*)* was approved by the local institutional review board (Medical Research Ethics Committees United, Registration Number: W17.134) and was conducted according to the Declaration of Helsinki. Informed consent was obtained from all participants.

Outcome measures

The 5R-STS was performed according to the protocol described by Jones et al. [14, 22]. The first measurement was obtained during the initial clinical visit (supervised). Subsequently, patients received instructions for at-home assessment from a licensed physiotherapist. Patients were asked to independently perform two measurements during the 2 days following the initial clinical visit.

For the first (unsupervised) measurement, patients performed the 5R-STS as instructed and were timed by a partner or relative. For the second (telesupervised) measurement, patients were asked to again perform the test, but to produce a video recording of this performance, in place of being timed. Video recordings were obtained using the proprietary video recording software on patient´s smartphones, and the resulting video files were sent per email to a secure server managed by the research nurse and digitally timed. Patients gave written informed consent for these procedures.

The telesupervised measurements were timed independently by two raters (V.E.S. and M.L.S.) for assessment of interrater reliability. For all other analyses, we used the mean of the two raters' test times. Patients were asked to wait approximately 1 h between the unsupervised and telesupervised measurements to avoid fatigue or test-induced pain.

A range of PROMs were additionally used. Patients were asked to complete questionnaires containing baseline sociodemographic data, as well as numeric rating scales (NRS) for back and leg pain severity, and validated Dutch versions of the Oswestry Disability Index (ODI), Roland–Morris Disability Questionnaire (RMDQ), and EuroQOL-5D-3L (EQ-5D) to capture subjective functional impairment as well as HRQOL. Participants filled in the questionnaires right after initially performing the test during the clinical visit. We assessed convergent validity of the 5R-STS by correlation with these secondary outcome measures.

Statistical analysis

Data were reported as mean ± standard deviation for continuous and numbers (percentages) for categorical data. Linear regression was used to evaluate the effect of decade of life on clinical 5R-STS performance. For correlation analysis, 5R-STS test times were $\log_{10}$ transformed to achieve a normal distribution. Pearson's product–moment correlation was used to assess the correlation between $\log_{10}$-transformed 5R-STS test times obtained during clinical visits and those obtained at home. A post hoc power calculation was performed to detect a Pearson correlation coefficient (PCC) of 0.30 or greater. A PCC of 0.80 or greater was considered as "excellent" correlation according to the Landis–Koch criteria [27]. To determine interrater agreement of the telesupervised method, we calculated intraclass correlation coefficients (ICC) for the two raters in a two-way random-effects model set for absolute agreement, along with their 95% confidence intervals (CI). We additionally performed Bland–Altman analysis of interrater agreement. Wilcoxon's signed-rank test was used to compare intraindividual performances to identify any learning effect between measurements. We also quantified convergent validity by correlating PROMs with $\log_{10}$-transformed 5R-STS test times obtained during clinical visits. Analyses were carried out using R version 3.5.1 (the R Foundation for Statistical Computing, Vienna, Austria) [28]. A $p \leq 0.05$ on two-sided tests was considered significant.

Results

Population

One hundred and twenty-one patients were initially enrolled (Fig. 1). Of these, 119 (98%) performed the initial 5R-STS measurement, and 100 (83%) completed at least one of the at-home measurements and were eligible for the quantitative analysis. Table 2 provides an overview of the patient cohort. The majority of patients suffered from LDH (65%), followed by stenosis (23%), DDD (9%), and spondylolisthesis (3%).

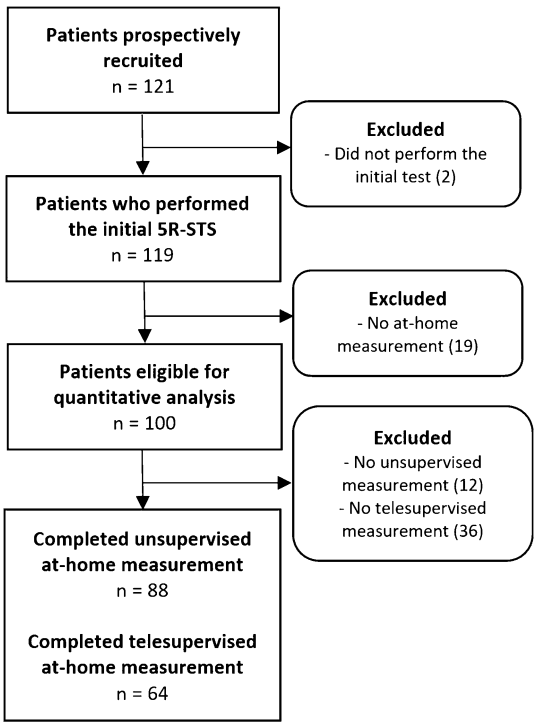


Fig. 1 Flow chart demonstrating the flow of patients at each stage of the study

No adverse events were reported during testing. Our post hoc power analysis indicated a power of $1-\beta=0.81$ for the analysis of our primary endpoint. Among the included patients, none were known to have other concomitant comorbidities that may have significant influence on 5R-STS performance, such as cervical or thoracic myelopathy, Parkinson's disease, or other diffuse neuropathies. An analysis per decade of life (Fig. 2) indicated that age did not significantly influence clinical 5R-STS performance (β: 0.49, 95% CI −0.74 to 1.72, $p=0.433$).

Correlation of clinical and at-home test times

Unsupervised measurement

Eighty-eight patients (88%) reported an unsupervised test time (Fig. 3). Unsupervised measurements showed excellent correlation with the clinical measurement (r: 0.94, 95% CI 0.91–0.96, $p<0.001$).

Telesupervised measurement

Overall, we received video recordings from 64 patients (64%). Three videos (5%) were not ratable due to recording errors. In the 61 remaining recordings (Fig. 4), the correlation among clinical and telesupervised at-home measurements was excellent (r: 0.90, 95% CI 0.83–0.94, $p<0.001$).

Table 2 Baseline characteristics of the patient cohort

Parameter	Value $n=100$
Male gender	44 (44)
Age (yrs)	45.25 ± 11.90
Height (cm)	176.65 ± 9.23
Weight (kg)	79.52 ± 12.17
Body mass index (kg/m^2)	25.49 ± 3.39
Smoking status	
Active smoker	27 (27)
Ceased smoking	22 (22)
Never smoked	51 (51)
Ability to work	
Full	21 (21)
Limited	26 (26)
Unable	53 (53)
Prior spine surgery	25 (25)
History of pain	
None–6 weeks	2 (2)
6 weeks–6 months	14 (14)
6 months–1 year	25 (25)
> 1 year	59 (59)
Analgesic drug use	
Daily	79 (79)
Weekly	8 (8)
Not regularly	13 (13)
Indication	
Disc herniation	65 (65)
Stenosis	23 (23)
DDD	9 (9)
Spondylolisthesis	3 (3)
Index level	
L2–L3	4 (4)
L3–L4	7 (7)
L4–L5	33 (33)
L5–S1	56 (56)
PROM	
RMDQ	12.75 ± 5.48
ODI	47.85 ± 16.00
VAS back pain	5.87 ± 2.62
VAS leg pain	7.54 ± 1.88
EQ-5D index	0.33 ± 0.27
EQ-5D VAS	48.52 ± 16.07

Continuous variables are presented as mean ± standard deviation and categorical variables as frequency (percentage)

DDD degenerative disc disease, *PROM* patient-reported outcome measure, *RMDQ* Roland–Morris Disability Questionnaire, *ODI* Oswestry Disability Index, *VAS* visual analogue scale

Interrater reliability of telesupervised measurement

Measurement of test times using the telesupervised method appeared to be highly reliable, based on absolute interrater

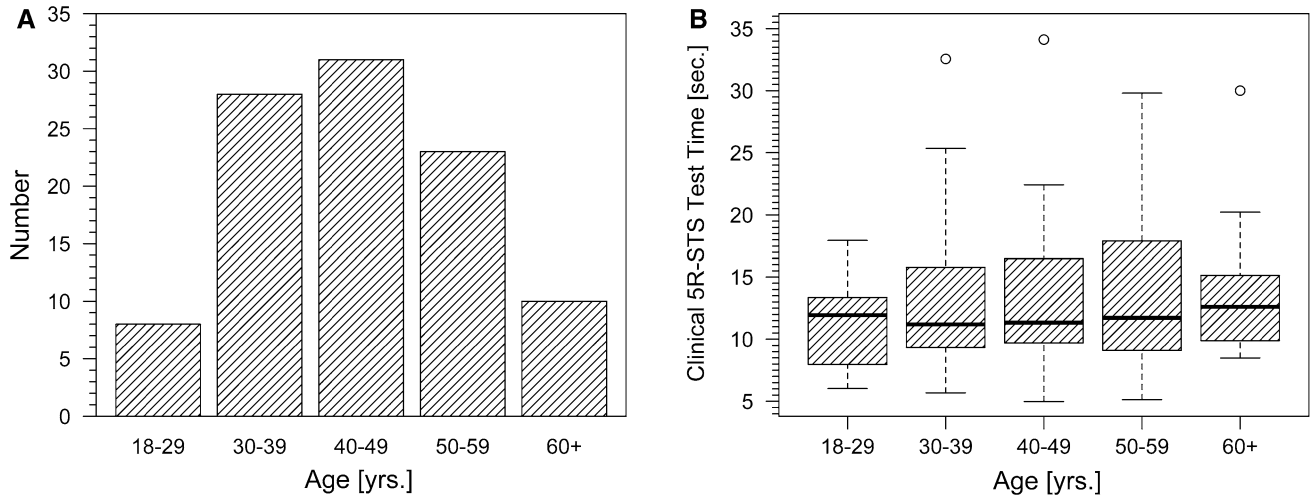


Fig. 2 Analysis of clinical 5R-STS performance per decade of life. The age distribution (**a**) as well as its effect on test performance (**b**) is provided

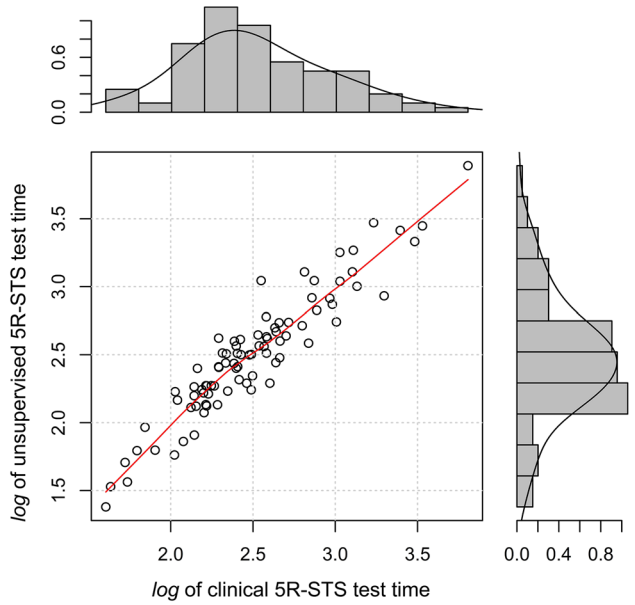


Fig. 3 Scatterplot illustrating the correlation of logarithmically converted clinical and unsupervised at-home 5R-STS test times. The Pearson correlation coefficient was 0.94 (95% CI 0.91–0.96, $p < 0.001$), indicating excellent correlation

agreement (ICC: 0.996, 95% CI 0.993–0.998, $p < 0.001$). A Bland–Altman plot (Fig. 5) illustrates the mean interrater bias of −0.15 s, with a 95% limit of agreement of −0.81 to 0.51 s.

Intraindividual performance differences

Unsupervised measurement

Patients did not perform more slowly at home (Table 3), and no learning effects were observed (Δ: −0.17 s, 95% CI −0.68 to 0.34 s, $p = 0.565$).

Telesupervised measurement

There were no significant intraindividual differences between clinical and telesupervised measurements (Δ: −0.17 s, 95% CI −0.82 to 0.48 s, $p = 0.897$).

Convergent validity

We observed direct correlation of test times and functional impairment as measured by RMDQ ($n = 100$, r: 0.48, 95% CI 0.31–0.63, $p < 0.001$) and ODI (r: 0.40, 95% CI 0.21–0.55, $p < 0.001$), as well as with VAS back pain severity (r: 0.25, 95% CI 0.05–0.43, $p = 0.017$). There was no correlation with VAS leg pain severity ($p = 0.189$). The 5R-STS also demonstrated indirect correlation with HRQOL as measured by EQ-5D index (r: −0.24, 95% CI −0.42 to −0.04, $p = 0.021$) and EQ-VAS (r: −0.30, 95% CI −0.47 to −0.10, $p = 0.004$).

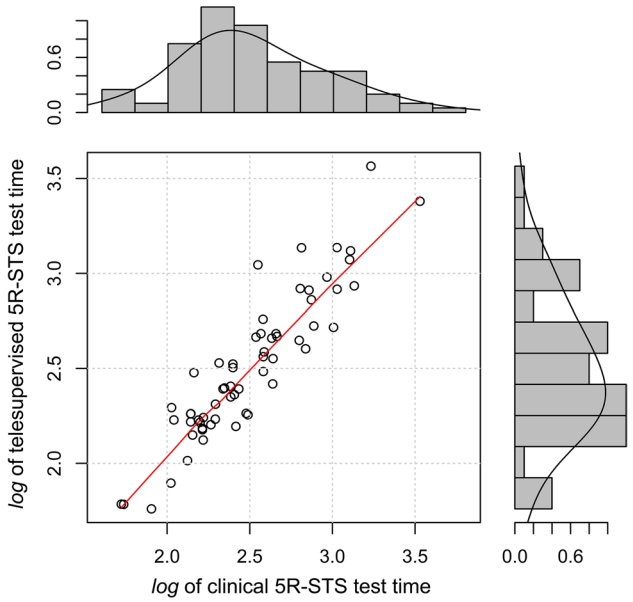


Fig. 4 Scatterplot illustrating the correlation of logarithmically converted clinical and telesupervised at-home 5R-STS test times. Test times were digitally measured on video recordings provided by the patients. The Pearson correlation coefficient was 0.90 (95% CI 0.83–0.94, $p < 0.001$), indicating excellent correlation

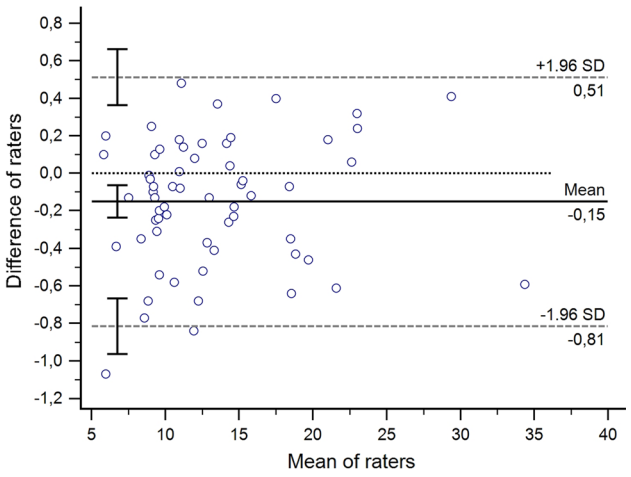


Fig. 5 Bland–Altman analysis of interrater agreement of telesupervised measurements. Two raters independently measured test times on the video recordings. The mean rating among the two raters is plotted against the difference between raters. The plot demonstrates a mean interrater bias of −0.15 s, with a 95% limit of agreement of −0.81 to 0.51 s

Discussion

In this prospective study, we demonstrated that the 5R-STS can, without supervision, objectively assess functional impairment in patients with degenerative pathologies of the lumbar spine. Using either the unsupervised or telesupervised method, excellent reliability can be achieved. Remote patient assessment using the 5R-STS appears to be an accurate and safe option for the initial or follow-up evaluation of objective functional impairment (OFI).

Objective functional tests have the capability of adding a new dimension to patient assessment, namely, OFI. Next to simplicity of interpretation for the patient, test–retest reliability, patient preference, and language independence, OFI has the advantage of capturing deficits that are not assessed by questionnaires, such as extensor paresis and proprioceptive deficits [14, 18, 21, 22]. The presence and degree of OFI can be graded using validated severity stratifications [14, 29]. Although it is still unclear if the preoperative presence of OFI influences outcomes after surgery, they effectively communicate severity of disease and provide a complete overview of a patient's clinical status if complemented by conventional PROMs [7–9].

Not much is known about unsupervised or remote objective functional assessment. Regterschot et al. [19] evaluated

Table 3 5R-STS performance measured during the initial clinical visit and at home using the unsupervised and telesupervised methods

Group	Clinical $n = 100$		Unsupervised $n = 88$			Telesupervised $n = 61$		
	Mean	SD	Mean	SD	p	Mean	SD	p
Overall	13.75	6.84	13.54	7.11	0.565	13.25	5.66	0.897
Age group								
< 50	13.64	7.04	13.37	7.71	–	13.21	6.21	–
≥ 50	13.99	6.54	13.91	5.76	–	13.36	4.49	–
Gender								
Male	15.04	8.42	14.62	9.32	–	14.32	6.93	–
Female	12.91	5.31	12.78	4.67	–	12.72	4.09	–

Values are given in seconds as mean ± standard deviation. p values for intraindividual differences from the initial clinical measurement were obtained using Wilcoxon's signed-rank test

5R-STS five-repetition sit-to-stand test, *SD* standard deviation

sensor-based assessment of sit-to-stand manoeuvres and found that reliability was high. Ejupi et al. [20] used motion tracking for clinical and at-home 5R-STS assessment in a geriatric population. With both unsupervised and supervised at-home motion tracking-based testings, they were able to accurately measure 5R-STS performance and to develop an algorithm that discriminates fallers from nonfallers. They observed a high correlation of clinical and at-home tracking-based performance. These studies achieved very high accuracy by the use of biomechanical analysis of sit-to-stand movements. However, this is impractical in daily practice, and most patients do neither have access to the required equipment, nor are they able to correctly apply the equipment by themselves. For these reasons, we aimed to assess if simple 5R-STS test time measurement is feasible. Nonetheless, the findings of the aforementioned studies support the notion that home-based 5R-STS testing without supervision could be viable [19, 20].

An integral part of the 5R-STS testing procedure is that the patient aims to perform the manoeuver as swiftly as possible. This allows testing for peak muscle strength, balance, and to some extent, endurance [14, 18, 19, 22, 30, 31]. Therefore, motivation and clear instructions on how to perform the 5R-STS are the key. It was conceivable that patients would perform somewhat more slowly at home, without direct supervision, than during the doctor's presence. However, our findings indicate that there is no difference in performance, rather with a slight trend for faster performance at home. The latter is possibly explained by selection bias, since patients who returned at-home measurements are perhaps more motivated.

In contrast to our initial hypothesis, telesupervision using mailed videos did not lead to increased reliability compared to unsupervised assessment. Although we were able to demonstrate that interrater agreement was high for this method, the lack of superiority is probably explained by methodological artefacts. Three patients were unable to record the video correctly. It is also conceivable that older patients, who may be somewhat less technically experienced, would achieve worse results, although this was not evident from our results.

In a recent study, Staartjes et al. [14] demonstrated the convergent validity of the 5R-STS in 157 patients. Because the 5R-STS is still a relatively novel test in spinal patient care, we aimed to assess its convergent validity in an independent cohort. We found good correlation of 5R-STS scores and validated PROMs for pain and disability, except for leg pain severity. Our findings corroborate the results previously published [14]. Here, it is important to stress again that the goal of objective functional tests is not primarily to converge with conventional PROMs like leg pain severity or to replace them, but to supplement them by adding an objective dimension, namely, OFI. Capturing OFI has been demonstrated to provide valuable additional information about patients—such as motor deficits or dependency—while being independent of linguistic restrictions and generally preferred over questionnaires by patients [7–9, 11, 18, 21]. As of yet, the responsiveness of the 5R-STS has not been evaluated [6, 25, 32, 33]. Thus, the postoperative validity and sensitivity in this patient population, and concomitantly its usefulness as a tool for follow-up, instead of baseline assessment only, remain unknown [14, 32]. Consequently, there has also not been any calculation of the minimum clinically important difference (MCID) [10, 33]. A systematic review on objective functional tests in the spine may be warranted to summarize the available evidence [6]. In addition, we cannot judge the accuracy of the 5R-STS for any of the specific subgroups of included degenerative pathologies, since we included a somewhat heterogenous population. However, this approach has clinical relevance, since the population closely resembles that seen by spine surgeons in clinical practice.

It is conceivable that the 5R-STS will be clinically used in a similar fashion to other tests such as the TUG and 6MWT, including but not limited to severity stratification, tracking the postoperative course, and setting treatment goals [11]. While its responsiveness has not been demonstrated as of yet, the 5R-STS may be used by clinicians to gauge the degree of improvement in OFI compared to preoperative baseline testing. In addition, an Upper Limit of Normal (ULN) of 10.4 s has been suggested, with which the presence of OFI can be evaluated [14]. The proportion of patients with baseline OFI returning to "normal" values postoperatively could be used as a measure of surgical success.

Overall, the results of our study indicate that the 5R-STS can be reliably and safely performed at home, even without video-based supervision. Although this conclusion warrants validation in other prospective cohorts, it is reasonable to apply the 5R-STS as an outcome measure for OFI in clinical trials. As PROMs are often captured using mailed or web-based questionnaires at follow-up, remote assessment without supervision is logistically and economically superior for patient, clinician, and the health system in general, compared to each patient having to return for a clinical 5R-STS measurement. Unsupervised 5R-STS performance allows clinicians to objectively evaluate patients for impairment at follow-up in daily practice, apart from its application in research.

Limitations

While the present sample size reached sufficient power, our study may be limited by its sample size. Since the estimation of correlations relies on rather large samples, further, larger studies may be warranted. Since we included patients with various degenerative spinal diseases, we are unable to

demonstrate the validity of at-home testing for any particular diagnosis. However, particularly for LDH and stenosis, which comprised the majority of patients, the results of this study indicate that reliability is high. We cannot make any claims towards the reliability of at-home 5R-STS testing without prior instructions. Confounders which may not have been captured and accounted for in this study may have influenced performance. However, all included patients had American Society of Anesthesiologists (ASA) Scores of I or II, and none were known to have comorbidities with potential confounding effect on 5R-STS performance, such as myelopathies or neuropathies. Lastly, it is currently unknown if unsupervised performance of other objective functional tests, such as the 6MWT and TUG, is a reliable indicator of OFI.

Conclusions

Unsupervised at-home assessment using the 5R-STS is highly accurate and reliable. There does not appear to be a specific need for patients to return for a clinical, supervised follow-up measurement of objective functional impairment. Rather, instructions can be provided, and the test performed and rated by a partner or family member at home. This is logistically and economically advantageous for patients, clinicians, and researchers. The 5R-STS is a simple, quick, and reliable indicator of objective functional impairment.

Acknowledgements We thank Marlies P. de Wispelaere, MSc, for her efforts in clinical informatics.

Compliance with ethical standards

Conflict of interest The authors declare that the article and its content were composed in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Falavigna A, Dozza DC, Teles AR et al (2017) Current status of worldwide use of patient-reported outcome measures (PROMs) in spine care. World Neurosurg 108:328–335. https://doi.org/10.1016/j.wneu.2017.09.002
2. Deyo RAM, Battie M, Beurskens AJHM et al (1998) Outcome measures for low back pain research: a proposal for standardized use. Spine 23:2003–2013
3. Fairbank JC, Couper J, Davies JB, O'Brien JP (1980) The Oswestry low back pain disability questionnaire. Physiotherapy 66:271–273
4. Roland M, Morris R (1983) A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 8:141–144
5. Rabin R, de Charro F (2001) EQ-5D: a measure of health status from the EuroQol Group. Ann Med 33:337–343
6. Prinsen CAC, Mokkink LB, Bouter LM et al (2018) COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 27:1147–1157. https://doi.org/10.1007/s11136-018-1798-3
7. Guzman JZ, Cutler HS, Connolly J et al (2016) Patient-reported outcome instruments in spine surgery. Spine 41:429–437. https://doi.org/10.1097/BRS.0000000000001211
8. Parai C, Hägg O, Lind B, Brisby H (2018) The value of patient global assessment in lumbar spine surgery: an evaluation based on more than 90,000 patients. Eur Spine J 27:554–563. https://doi.org/10.1007/s00586-017-5331-0
9. Algattas H, Cohen J, Agarwal N, Hamilton DK (2017) Trends in the use of patient-reported outcome instruments in neurosurgical adult thoracolumbar deformity and degenerative disease literature. J Craniovertebral Junction Spine 8:103–107. https://doi.org/10.4103/jcvjs.JCVJS_29_17
10. Ostelo RWJG, Deyo RA, Stratford P et al (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine 33:90–94. https://doi.org/10.1097/BRS.0b013e31815e3a10
11. Gautschi OP, Corniola MV, Schaller K et al (2014) The need for an objective outcome measurement in spine surgery—the timed-up-and-go test. Spine J 14:2521–2522. https://doi.org/10.1016/j.spinee.2014.05.004
12. Guyatt GH, Sullivan MJ, Thompson PJ et al (1985) The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. Can Med Assoc J 132:919–923
13. Mobbs RJ, Phan K, Maharaj M, Rao PJ (2016) Physical activity measured with accelerometer and self-rated disability in lumbar spine surgery: a prospective study. Glob Spine J 6:459–464. https://doi.org/10.1055/s-0035-1565259
14. Staartjes VE, Schröder ML (2018) The five-repetition sit-to-stand test: evaluation of a simple and objective tool for the assessment of degenerative pathologies of the lumbar spine. J Neurosurg Spine 29:380–387. https://doi.org/10.3171/2018.2.SPINE171416
15. Moke L, Severijns P, Schelfaut S et al (2018) Performance on balance evaluation systems test (BESTest) impacts health-related quality of life in adult spinal deformity patients. Spine 43:637. https://doi.org/10.1097/BRS.0000000000002390
16. Haddas R, Lieberman IH, Block A (2018) The relationship between fear-avoidance and objective biomechanical measures of function in patients with adult degenerative scoliosis. Spine 43:647. https://doi.org/10.1097/BRS.0000000000002381
17. Ferguson SA, Marras WS, Burr DL et al (2009) Quantification of a meaningful change in low back functional impairment. Spine 34:2060. https://doi.org/10.1097/BRS.0b013e3181b34764
18. Gautschi OP, Smoll NR, Corniola MV et al (2016) 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. https://doi.org/10.1227/NEU.0000000000001195
19. Regterschot GRH, Zhang W, Baldus H et al (2014) Test–retest reliability of sensor-based sit-to-stand measures in young and older adults. Gait Posture 40:220–224. https://doi.org/10.1016/j.gaitpost.2014.03.193
20. Ejupi A, Brodie M, Gschwind YJ et al (2015) Kinect-based five-times-sit-to-stand test for clinical and in-home assessment of fall risk in older people. Gerontology 62:118–124. https://doi.org/10.1159/000381804
21. Joswig H, Stienen MN, Smoll NR et al (2017) Patients' preference of the timed up and go test or patient-reported outcome measures before and after surgery for lumbar degenerative disk

disease. World Neurosurg 99:26–30. https://doi.org/10.1016/j.wneu.2016.11.039
22. Jones SE, Kon SSC, Canavan JL et al (2013) The five-repetition sit-to-stand test as a functional outcome measure in COPD. Thorax 68:1015–1020. https://doi.org/10.1136/thoraxjnl-2013-203576
23. Duncan RP, Leddy AL, Earhart GM (2011) Five times sit to stand test performance in Parkinson disease. Arch Phys Med Rehabil 92:1431–1436. https://doi.org/10.1016/j.apmr.2011.04.008
24. Simmonds MJ, Olson SL, Jones S et al (1998) Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 23:2412–2421
25. Mokkink LB, Terwee CB, Patrick DL et al (2010) The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 63:737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006
26. Von Elm E, Altman DG, Egger M et al (2007) Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 335:806–808. https://doi.org/10.1136/bmj.39335.541782.AD
27. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
28. Core Team R (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
29. Stienen MN, Smoll NR, Joswig H et al (2017) Validation of the baseline severity stratification of objective functional impairment in lumbar degenerative disc disease. J Neurosurg Spine. https://doi.org/10.3171/2016.11.SPINE16683
30. Shum GLK, Crosbie J, Lee RYW (2007) Three-dimensional kinetics of the lumbar spine and hips in low back pain patients during sit-to-stand and stand-to-sit. Spine 32:E211–E219. https://doi.org/10.1097/01.brs.0000259204.05598.10
31. Shum GLK, Crosbie J, Lee RYW (2005) Effect of low back pain on the kinematics and joint coordination of the lumbar spine and hip during sit-to-stand and stand-to-sit. Spine 30:1998–2004
32. Gautschi OP, Joswig H, Corniola MV et al (2016) 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. https://doi.org/10.1007/s00701-016-2899-9
33. Gautschi OP, Stienen MN, Corniola MV et al (2016) Assessment of the minimum clinically important difference in the timed up and go test after surgery for lumbar degenerative disc disease. Neurosurgery. https://doi.org/10.1227/NEU.0000000000001320

Affiliations

Victor E. Staartjes[1,2,3] · Femke Beusekamp[1] · Marc L. Schröder[1]

✉ Victor E. Staartjes
victor.staartjes@gmail.com

[1] Department of Neurosurgery, Bergman Clinics Amsterdam, Rijksweg 69, 1411 GE Naarden, The Netherlands

[2] Department of Neurosurgery, Clinical Neuroscience Center, University Hospital Zurich, University of Zurich, Zurich, Switzerland

[3] Department of Neurosurgery, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, Amsterdam, The Netherlands