



Original Article

A new two square agility test for workplace health—reliability, validity and minimal detectable change

RICHARD J. WICKSTROM, PT, DPT¹*, YING-CHIH WANG, OTR/L, PhD²,
NELL E. WICKSTROM, MS¹, ROSE L. SMITH, PT, DPT³, KARI K. DUNNING, PT, PhD³

¹) *WorkAbility Systems: 7665 Monarch Court, Suite 109, West Chester, OH 45069, USA*

²) *College of Health Sciences, Department of Occupational Science & Technology, University of Wisconsin-Milwaukee, USA*

³) *College of Allied Health Sciences, Department of Rehabilitation, Exercise and Nutrition Science, University of Cincinnati, USA*

Abstract. [Purpose] One promising strategy for workplace wellness programs is to emphasize functional mobility screening and coaching to promote suitable physical activity and reduce musculoskeletal risks. This study examined intra-rater reliability, test-retest reliability, concurrent validity, known-groups validity and minimal detectable change for a new Two Square Agility Test (TSAT) designed as a functional mobility measure to promote workplace health. [Participants and Methods] Two hundred forty eight non-disabled participants (ages 18–69) were measured for body size, physical activity and 3 trials of the Two Square Agility Test. 78 participants were tested a week later on the Two Square Agility Test and other functional mobility tests. [Results] Intra-rater reliability was excellent (ICC=0.94) and test-retest reliability was good (ICC=0.87). Two Square Agility Test correlated moderately with Timed Up and Go ($r=0.63$), Five Times Sit to Stand ($r=0.62$), and Maximum Step Length ($r=-0.54$), supporting its concurrent validity. Performances for Two Square Agility Test were better in males, younger age, higher physical activity, and non-obese groups. The minimal detectable change at a 95% confidence level (MDC_{95}) was 1.37 s. [Conclusion] Preliminary results supported reliability and validity of Two Square Agility Test as a functional mobility measure to promote workplace health.

Key words: Fitness, Functional mobility

(This article was submitted Apr. 27, 2019, and was accepted Jul. 27, 2019)

INTRODUCTION

In the past few decades, workplace wellness programs have grown in popularity and often include biometric screening to quantify health risks and motivate accountability for healthy lifestyles¹. The return on investment, however, is questionable and controversial^{2–5}. Baxter et al. conducted a systematic review that reported a negative return on investment for randomized control trials, after excluding early return-to-work and workplace injury prevention studies².

The traditional approach to health screening has been focused on identifying cardiovascular risks that may be improved through lifestyle behavior modification. This health screening for risk factors commonly uses health risk assessment surveys and biometric screening of blood labs, blood pressure, waist girth, and body mass index¹. Wellness programs that emphasize biometrics and interventions to reduce cardiovascular risk factors have not been shown to reduce healthcare utilization or cost^{4, 5}.

We believe that one promising strategy to improve return on investment for workplace wellness is to put more emphasis

*Corresponding author. Richard J. Wickstrom (E-mail: rick@workability.us)

©2019 The Society of Physical Therapy Science. Published by IPEC Inc.



This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License. (CC-BY-NC-ND 4.0: <https://creativecommons.org/licenses/by-nc-nd/4.0/>)

on functional mobility screening and individualized coaching to promote suitable physical activity for participants with musculoskeletal risks and obesity. There are several reasons. First, musculoskeletal disorders are the most commonly reported medical conditions in adults under 65, resulting in an estimated annual cost of \$980 billion for direct treatment and indirect disability in the United States⁶. Second, studies have shown that reduced functional mobility is a key risk factor associated with lower perceived quality of life^{7, 8}, obesity⁷⁻¹⁰, lower bone mineral density¹¹, more limitations on daily activities¹²⁻¹⁵, higher risk of fractures and greater fall risk^{11, 15-19}. Further, functional mobility screening supports triage to suitable physical activity interventions for costly comorbidities such as musculoskeletal disorders⁶ and obesity^{20, 21} that directly affect workplace safety and productivity.

Currently, there are only a few established functional mobility tests with normative data to interpret the results for working adults. These tests include Timed Up and Go (TUG)¹⁹, Usual Gait Speed (UGS)²³, and Fast Gait Speed (FGS)²³. Practical constraints that limit use of these tests at the workplace setting or health fair event include large space requirements, privacy concerns, and ceiling effects with higher functioning workers. Maximum Step Length (MSL)^{16, 24} and Five Times Sit to Stand (5TSTS)^{24, 25} are two other established tests of functional mobility that lack normative data for younger, working age populations. All of the above tests were designed primarily to screen for fall risk or assess intervention outcomes in more impaired elderly or clinical populations.

The Two Square Agility Test (TSAT) is a new performance test of functional mobility that was originated by the lead author for use in occupational health applications such as fitness-for-duty screening of job candidates, workplace wellness screening and functional capacity evaluation. It requires minimal equipment and is easy to administer within a small space in less than 5 minutes. It was designed to provide a simple, reciprocal stepping sequence that would be safe, practical and acceptable for workers that have a very diverse range of physical or cognitive fitness. One examiner is able to guard the participant from stable, spotting position during performance of the entire movement sequence.

Therefore, the main purpose of this study was to preliminarily assess intra-rater reliability, test-retest reliability, concurrent validity, known-groups validity and minimal detectable change (MDC) for the TSAT in non-disabled working adults.

PARTICIPANTS AND METHODS

A convenience sample of participants between the ages of 18–69 years old were recruited with flyers that were circulated at a local university, a local work-site and two outpatient physical therapy clinics. Participants were excluded if they were disabled from work, had uncontrolled blood pressure (BP \geq 160/100 mmHg), were unable to get up from a chair without use of arm rests, were unable to walk 30 feet without an assistive device, were restricted in physical activities by a doctor, fell within the past two months due to dizziness or loss of consciousness, were pregnant, reported having severe pain during the prior month, were legally blind, or were unable to follow simple instructions in English. This study was approved by the University of Cincinnati Institutional Review Board and participants were informed of the benefits and risks prior to signing the consent form.

At the start of data collection, we collected demographic information (age, gender, race) and determined the physical activity and body size measurements of the participants. Physical activity was determined using the Physical Activity Rating (PAR) scale, where each participant reported their level of recent physical activity on a scale from 0 (avoids walking or exercise) to 7 (runs more than 10 miles per week or spends more than 3 hours per week doing vigorous physical activity)²⁷. Body size measures for validity analysis included: height (cm); weight (kg); body mass index (BMI); waist (cm) defined as the shortest distance around at the level of the umbilicus with a non-stretch tape²⁸; and waist to height ratio (WtH)²⁹. Subsequently, participants completed 3 trials of the TSAT.

All participants were given the option to return for a 2nd testing session (one week apart) to retest their performance on the TSAT. Prior to testing at the 2nd session, participants were asked if any changes had occurred (e.g. symptoms, pain, injury, illness, doctor visit, physical activity changes) since the 1st session. At the 2nd testing session, participants completed the TSAT and five additional functional mobility measures (described below) with the same examiner from the 1st testing session. The order for functional mobility tests was randomized to minimize the effects of fatigue.

Two Square Agility Test (TSAT): This movement sequence involves stepping forward and backward between the two squares as fast as safely possible for a total of 5 complete cycles, while leading with the preferred leg (Fig. 1). To set-up for the TSAT, a metal tape measure was extended at least one meter along the floor about 0.5 meters parallel to a sturdy table and secured with painter's tape. The participant began facing the table with the both feet behind the tape measure.

The examiner initially demonstrated the movement sequence at a comfortable pace, while providing the following instructions: "This is a test of your ability to step forward and back across the tape. When I say begin, you will practice this sequence at a comfortable pace for about 5 cycles until you hear me say stop. You may lead with your preferred leg." After the participant demonstrated a satisfactory practice trial at a comfortable pace, the examiner provided the following instructions: "Next, you will practice this sequence as fast as you safely can." While the participant performed the movement sequences at a fast pace, the examiner stood behind and to one side in a stable, spotting position to guard the participant in event of any observed unsteadiness during the movement sequence. The stop watch was started when the lead foot of the participant first landed on the floor (heel or toe strike) in front of the marked tape and stopped after the lead foot landed again on the floor in front of the tape after 5 complete stepping cycles. Time was recorded with a stopwatch to the nearest 0.01 seconds. The best

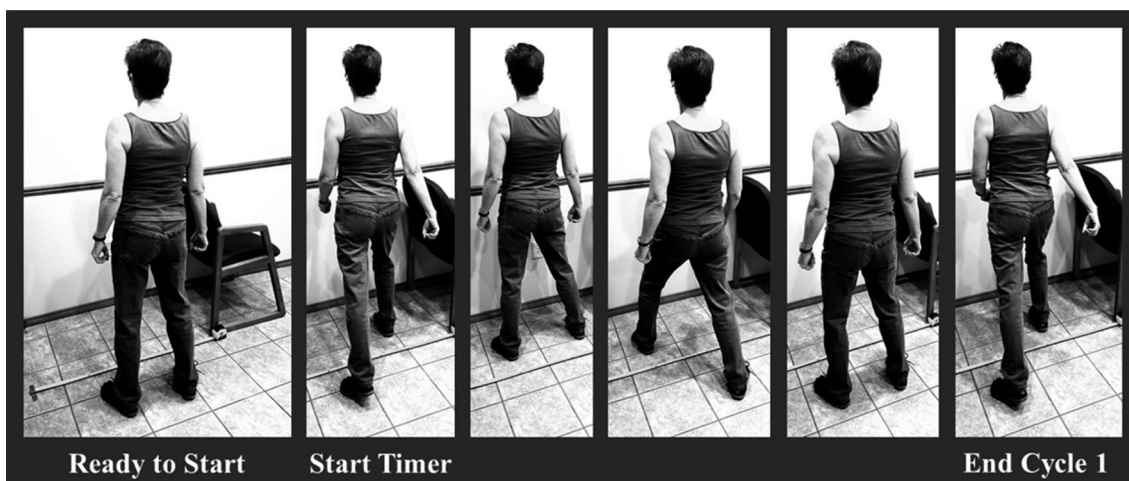


Fig. 1. Two Square Agility Test (TSAT) movements for 1st cycle with right leg lead.

(lowest completion time) of the 3 timed trials was used for analysis of overall performance.

Timed Up and Go (TUG): The “as quickly and as safely as possible” method for performing the TUG was selected for validity comparison because this method has been shown to be predictive of fall risk and was most similar to participant instructions for TSAT to step as fast or as far as safely possible¹⁵. The participant’s time was recorded (sec) for a movement sequence that consisted of standing up from a chair without using arms to assist, walking 3 m as fast as safely possible to cross a line marked on the floor, turn around, and walking back to sit down. Participants were given a practice trial and then performed two timed trials. The best of the 2 timed trials (lowest time) was used for analysis of concurrent validity.

Usual Gait Speed (UGS) and Fast Gait Speed (FGS): UGS and FGS were included for validity comparison because these tests are established measures of functional mobility that have normative data for working age adults²³. Performance time in seconds was measured for UGS “at a comfortable pace” and for FGS “as quickly as you safely can” over a 6-meter distance from a static start. This protocol by Cesari et al. was chosen because the entire stepping movements could be easily visualized by a single examiner without need for special equipment³⁰. Two trials of UGS and two trials of FGS were recorded to the nearest 0.01 second and speed (m/s) was calculated by dividing 6 m by the best (lowest) trial for analysis of concurrent validity.

Maximum Step Length (MSL): MSL is the maximum distance that a person can step forward from standing with feet even and successfully return without loss of balance. This test was chosen as a comparative functional mobility measure because it has been found to be reliable, valid and acceptable for participants with a range of ages and fitness levels (healthy to frail adults)^{16, 24, 31}. MSL was shown to be a predictor of mobility performance, frequent falls, self-reported function and balance confidence in older adults¹⁶. Prior to administration, a metal tape measure was extended at least 2 m along the floor a half meter parallel to a sturdy table and secured at the 0 start mark with a piece of painter’s tape. First, the examiner stood with both feet lined up behind the 0-start mark of the tape to demonstrate a comfortable step forward and returned behind the marked line. This was done separately with the right leg and then with the left leg. Then the participant was asked to complete a comfortable step practice trial with the right and then with the left leg. Next, the participant was asked to complete a practice trial for maximum step length for each leg. After successfully completing comfortable and maximum step length practice trials, a total of 3 trials of the MSL were administered for the right and left legs. The examiner stood in a spotting position to one side to guard each participant in event of unsteadiness and to observe maximum step length distance at the front of the shoe to the nearest cm. Trials were repeated if: participant lost balance, took more than one step in front of the line, was unable to return to behind the line in a single return step or moved one or both hands away from their sides for balance. The best (longest step length) in cm of 3 acceptable trials with the right leg was used for analysis of concurrent validity.

Five Times Sit to Stand (5TSTS): 5TSTS was included as a comparative measure because it has been found to be a reliable test of lower limb functional strength²⁶, acceptable for clinical use³² and often included in studies across an age-span that includes working age adults^{11, 26}. A standard chair without armrests was used and participants were instructed to keep the arms folded across the chest when standing up fully and sitting down for 5 repetitions as quickly as safely possible. Time was recorded to the nearest 0.01 second. Each participant performed one practice trial and was then timed for one trial to minimize fatigue effects.

Intra-rater reliability was assessed for TSAT assessed at the same session using intraclass correlation coefficient (ICC) (2,1)^{33, 34}. Test-retest reliability of TSAT was determined using ICC (2,1) to compare the results from the 1st and 2nd sessions^{33, 34}. Concurrent validity was assessed using Pearson correlation to correlate TSAT with the five functional measures:

Table 1. Characteristics by participation level

	Overall sample of participants who attended 1st testing session (N=248)	Subsample of participants who attended both 1st and 2nd testing sessions (N=78)
Demographic characteristics		
Age (years)	38.2 ± 14.3	35.8 ± 14.0
Gender: males/females	138 (55.6%)/110 (44.4%)	33 (42.3%)/45 (57.7%)
Race: Caucasian/all others	219 (88.3%)/29 (11.7%)	72 (92.3%)/6 (7.7%)
Physical activity rating:		
Vigorous	140 (56.5%)	50 (64.1%)
Moderate	71 (28.6%)	19 (24.4%)
Inactive	37 (14.9%)	9 (11.5%)
Body Mass Index (BMI) (kg/m ²)	27.9 ± 5.6	26.5 ± 5.6
Two Square Agility Test (TSAT)		
TSAT (s)	5.00 ± 1.34	4.90 ± 1.39

Data shown as mean (standard deviation) or number (%).

TUG, UGS, FGS, MSL and 5TSTS. To assess known-groups validity, TSAT was compared by gender (male/female), age (<35, 35–55, >55 years), PAR categories (inactive PAR score 0–1, moderate PAR score 2–3, vigorous PAR score 4–7) and BMI (<30, ≥30) using analysis of variance (ANOVA) followed by post hoc comparisons. The MDC at 95% confidence interval was calculated using standard error of measurement (SEM) using the following formulas: $MDC_{95} = 1.96 \times \sqrt{2} \times SEM$, where $SEM = (SD) \times \sqrt{(1-ICC)}$, where ICC is the test-retest reliability coefficient³⁵.

SPSS statistical package version 23 (SPSS Inc., Chicago, IL, USA) was used for all statistical analysis. The Kolmogorov-Smirnov test was used to test the data for normality. Statistical significance was defined as $p < 0.01$. ICC reliability values were considered poor if less than 0.5, moderate if between 0.5 to less than 0.75, good if between 0.75 to less than 0.9, and excellent if 0.9 to 1.00³³. Strength of correlations (positive or negative) for concurrent validity was defined by Pearson r values and considered to be a negligible if less than 0.3, low if 0.30 to less than 0.5, moderate if 0.50 to less than 0.70, high if 0.7 to less than 0.9, and very high if 0.9 to 1.00³⁶.

RESULTS

Two hundred and forty-eight adults participated in this study, of whom, 78 elected to return for a 2nd testing session. The demographic information was presented in Table 1. The majority of the 248 participants were males (55.6%), Caucasian (88.3%), and reported vigorous physical activity (56.6%). The average age was 38.2 years (range 18–69) and average BMI was 27.9. None of the participants that returned for a 2nd session reported having a change in symptoms or activity limitations between the sessions. The mean (SD) performances on the comparative functional mobility tests were TUG 4.5 (0.77) s, UGS 144.5 (19.1) cm/s, 288.5 (31.9) cm/s, 5TSTS 6.9 (1.8) s, and 114.1 (25.6) s.

Results supported that intra-rater reliability of TSAT was excellent with $ICC = 0.94$ (95% $CI = (0.92, 0.95)$). Some improvements in performance were noticed from trial 1 (mean performance = 5.72 s) to 2 (5.29 s) to trial 3 (5.06 s).

For the 78 participants that completed both sessions, the mean (SD) TSAT score was 4.90 (1.39) s in the 1st session and was 4.55 (1.15) s in the 2nd session. The test-retest reliability was good for TSAT ($ICC = 0.87$).

Table 2 presents the correlation relationships between TSAT and other functional mobility tests (concurrent validity) and several demographic variables of interests. Overall, results supported that participants who had better functional mobility/fitness as measured by TSAT, also completed TUG ($r = 0.63$) and 5TSTS ($r = 0.62$) faster, had longer step lengths ($r = -0.54$), had faster gait speed (correlation with FGS was -0.48 , and with UGS was negligible -0.06), and reported higher physical activity ratings (i.e., PAR scale). Meanwhile, these participants (who performed better in TSAT) tended to have lower body weight, BMI, and WtH.

The known-groups validity for TSAT is presented in Table 3. Overall, performances for TSAT were better in males, younger age, higher physical activity, and non-obese groups. With the SEM equaled to 0.50, the estimated MDC_{95} was 1.37 s.

DISCUSSION

This study supported the preliminary reliability and validity of TSAT. The traditional approach to health screening in workplace wellness programs fails to address functional mobility and other musculoskeletal risk factors that lead to workplace injuries and lost productivity. Screening for functional mobility is inexpensive, time-efficient and relevant to promote physical fitness from hire to retire. Combining functional mobility screening with individualized activity coaching is a practi-

Table 2. Concurrent validity of Two Square Agility Test (TSAT)

	Functional mobility tests						Body size measures			
	TSAT	UGS	FGS	TUG	FTSTS	MSL	Height	Weight	BMI	WtH
UGS	-0.06									
FGS	-0.48	0.28								
TUG	0.63	-0.25	-0.64							
5TSTS	0.62	-0.18	-0.52	0.48						
MSL	-0.54	0.05	0.32	-0.40	-0.25					
Height	-0.16	-0.06	0.16	-0.22	0.03	0.64				
Weight	0.31	-0.13	-0.27	0.25	0.32	0.18	0.48			
BMI	0.43	-0.11	-0.38	0.40	0.34	-0.14	-0.01	0.87		
WtH	0.52	-0.09	-0.45	0.44	0.40	-0.18	-0.03	0.78	0.91	
PAR	-0.45	0.09	0.20	-0.27	-0.39	0.18	0.12	-0.25	-0.35	-0.43

Bolded values indicate that correlation is significant at $p < 0.01$ (2-tailed).

UGS: usual gait speed; FGS: fast gait speed; TUG: timed up and go; FTSTS: five times sit to stand; MSL: maximum step length; BMI: body mass index; WtH: waist to height ratio.

Table 3. Known-groups validity (by age, gender, PAR and BMI)

Age (years)	TSAT (s)
(1) ≤ 35 (n=125)	4.47 \pm 1.07
(2) 35–55 (n=86)	5.21 \pm 1.39
(3) ≥ 55 (n=37)	6.29 \pm 1.36
p-value (1) vs. (2)	<0.0001
p-value (2) vs. (3)	<0.0001
p-value (1) vs. (3)	<0.0001
Gender	TSAT (s)
(1) Male (n138)	4.76 \pm 1.32
(2) Female (n110)	5.31 \pm 1.32
p-value (1) vs. (2)	0.001
PAR (Physical Activity Rating)	TSAT (s)
(1) Vigorously active (PAR 4-7) (n=140)	4.48 \pm 0.94
(2) Moderately active (PAR 2-3) (n=71)	5.27 \pm 1.40
(3) Inactive (PAR 0-1) (n=37)	6.44 \pm 1.38
p-value (1) vs. (2)	<0.0001
p-value (2) vs. (3)	<0.0001
p-value (1) vs. (3)	<0.0001
BMI (Body Mass Index)	TSAT (s)
(1) < 30 (n167)	4.65 \pm 1.06
(2) ≥ 30 (n81)	5.72 \pm 1.57
p-value (1) vs. (2)	<0.0001

cal approach that could be readily implemented by large or small employers.

The TSAT is a new method for assessing functional mobility that offers a number of practical advantages compared to other established tests of functional mobility. The TSAT was designed with a simple, reciprocal sequence of dynamic stepping movements to enhance its acceptability for assessment of functional mobility with low and high functioning persons. Persons with impaired physical or cognitive functioning are more likely to have difficulty following instructions or be unable to complete a complex task. For example, the Four Square Step Test (FSST)³⁷, which requires a more complex movement sequence that does not follow a reciprocal motor pattern, was found to be too difficult to complete by a significant percentage of persons in clinical subgroups such as poststroke³⁸ or vestibular disorders³⁹. Keeping the mobility movements simple and reciprocal supports inclusiveness and participation in workplace wellness programs by workers with more impaired cognitive or physical functioning. Another positive feature of the TSAT is that it can be administered by one examiner standing in place within a small area such as an exam room or booth during in a wellness fair. This represents both a safety and practical

advantage compared to other functional mobility tests such as UGS, FGS, and TUG that require more space or mobility challenges for an examiner to guard high or low functioning participants for safety in event of unsteadiness.

Among 248 participants who participated in the 1st testing session, a subsample of 78 participants agreed to come one week later to the 2nd testing session. The subsample was similar to the overall sample of 248 participants in age (35.8 vs. 38.2 years), BMI (26.5 vs. 27.9), TSAT (4.90 vs. 5.00 s). Gender differed slightly as the subsample was 57.7% female and the overall sample was 44.4% female, although this is unlikely to affect generalizability. The subsample of 78 participants, although smaller than the overall sample of 248 participants, is still adequate for test-retest reliability calculations³⁴). Although intra-rater reliability was excellent for TSAT during the 1st testing session, performances showed continual improvement from the 1st to 3rd trial, with means differing significantly ($p < 0.01$) between trials. This may result from motor learning effects and justifies our recommendation that clinicians use the “best” of 3 trials.

This was the first study to investigate the concurrent validity of the TSAT. TSAT demonstrated significant validity correlations as expected with TUG, FGS, MSL and 5TSTS measures that instructed the participant to move as fast or far as safely possible. The non-significant correlation for the TSAT with UGS implies that high functioning, non-disabled persons are likely to perform very differently when instructed to move as fast or far as safely as possible, compared to participants instructed to move at a submaximal or comfortable pace. Previous studies found FGS to be more relevant than UGS for independent community mobility, where the ability to walk fast is an important determinant of safety and mobility^{12, 14, 40}). This was consistent with a study of TUG performed at a “usual” speed that concluded that TUG may have more value to identify fall risk in less-healthy, lower-functioning older adults⁴¹). Studies investigating the TUG among older adults also identified large differences in cut-off times to predict fall risk when subjects are instructed to perform the TUG at their “usual” speed¹³) versus as “quickly as safely possible”¹⁶). The negligible relationship we found for UGS with the TSAT and other comparative mobility measures suggests that functional mobility tests that instruct participants to move at a usual or comfortable pace would not be useful with non-disabled adults that participate in a workplace wellness screening. Further investigation to compare usual versus fast paced instructions may be warranted for participants with impaired mobility that report a history of recent falls, dizziness, balance concerns, high pain levels, or demonstrate fear avoidance behavior. It may be helpful for clinicians to compare performance differences at a “usual” speed versus “as fast as safely possible” on functional mobility tasks such as the TSAT, TUG and Gait Speed when reduced balance confidence or symptoms may prompt self-limited performance.

TSAT time was more highly correlated with BMI and WtH measures of obesity than other comparative measures of functional mobility. TSAT was moderately correlated with WtH ($r = 0.52$) and correlated somewhat lower ($r = 0.43$) with BMI. This was consistent with findings by Ashwell et al. that WtH is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors²⁹). This has important implications for workplace health promotion, because it justifies use of both TSAT and WtH to demonstrate improvements in response to physical activity interventions that increase muscle mass and agility. Our findings of significant correlations for TSAT, FGS, TUG, and FTSTS with obesity were consistent with other studies⁷⁻¹⁰). Loss of functional mobility is an important life-relevant consequence to track for persons who present with significant obesity. Loss of mobility may explain findings by Ostbye et al. that employees in Obesity Class III (BMI ≥ 40) have more work-related slip and fall injuries, lower extremity injuries, overexertion injuries, lost work days, treatment costs and indemnity costs²⁰). Including TSAT and WtH as outcome measures for workers who participate in diet management and physical training interventions to address obesity would help reinforce accountability to motivate lifestyle behavior and further our understanding of population health statistics.

An alternate way of scoring the TSAT is to convert TSAT seconds to a speed variable by using an estimate of 11 meters for the distance covered during 5 complete cycles of stepping (i.e., m/s). This approach is similar to how gait speed is calculated and offers two important practical benefits compared with other mobility tests such as the TUG and 5TSTS because: 1) Participants may be more receptive and understanding of their results when presented in terms of TSAT speed (m/s) (where improvement occurs in a positive direction as the TSAT speed results), and 2) persons with severe physical or cognitive impairments that are unable to perform the TSAT task can be assigned a score of 0.0 meters per second as a reference point to compare with prior results or to include of their agility results as members of clinical or normative groups.

There are several limitations to the study. This study included healthier participants by excluding those with a recent history of severe pain and persons with insufficient mobility to stand up from sitting to perform the 5TSTS and TUG comparative validity tasks. Therefore, our results may not apply to the general working age population, because our volunteer sample that contained a higher mix of workers that were younger, non-obese and engaged in vigorous physical activity. It explains why the mean TUG score of 5.1 s for our sample was 40% faster than reported for participants in a similar age group for primary care patients that performed TUG at a normal or usual pace²²). Selection bias that results from recruiting healthier volunteers is a challenge that complicates the design of normative investigations as well as return on investment analyses of healthcare costs and utilization by participants versus non-participants in workplace wellness programs. The Illinois Workplace Wellness Study found that voluntary wellness programs engage participants with healthier behaviors and lower medical expenditures than non-participants and concluded that the primary value of wellness programs to employers may be to help recruit and retain workers that have low health care costs⁴). The TSAT is simple test of functional mobility that can be monitored from hire to retire to promote workplace safety, productivity and health management.

Funding and Conflict of interest

This study was partly funded by a grant by the Ohio Physical Therapy Association Research Committee. The funding agency had no role in the study design, writing the manuscript, or the decision to submit for publication. No manufacturer's aid was received in conjunction with the present study, and no conflicts of interest are declared.

ACKNOWLEDGEMENT

The authors gratefully acknowledge assistance from students in the University of Cincinnati Doctor of Physical Therapy (DPT) Program: Mark Geise, Samuel Meyer, Kyle Nordrum, Brooke Norbeck, Thomas Powers, Christopher Reis, Christine Stier, Thomas Toerpe.

REFERENCES

- 1) Mattke S, Liu HH, Caloyeras JP, et al.: Workplace wellness programs study: final report. *Rand Health Q*, 2013, 3: 7. [[Medline](#)]
- 2) Baxter S, Sanderson K, Venn AJ, et al.: The relationship between return on investment and quality of study methodology in workplace health promotion programs. *Am J Health Promot*, 2014, 28: 347–363. [[Medline](#)] [[CrossRef](#)]
- 3) Goetzel RZ, Henke RM, Tabrizi M, et al.: Do workplace health promotion (wellness) programs work? *J Occup Environ Med*, 2014, 56: 927–934. [[Medline](#)] [[CrossRef](#)]
- 4) Jones D, Molitor D, Reif J: What do workplace wellness programs do? Evidence from the Illinois Workplace Wellness Study. 2018. <https://www.nber.org/papers/w24229>
- 5) Mattke S, Kapinos K, Caloyeras JP, et al.: Workplace wellness programs: services offered, participation, and incentives. *Rand Health Q*, 2015, 5: 7. [[Medline](#)]
- 6) United States Bone and Joint Initiative: The burden of musculoskeletal diseases in the United States: prevalence, societal and economic cost, 4th ed. Executive Summary. 2018. https://www.boneandjointburden.org/docs/BMUS%20Impact%20of%20MSK%20on%20Americans%20booklet_4th%20Edition%20%282018%29.pdf
- 7) Fjeldstad C, Fjeldstad AS, Acree LS, et al.: The influence of obesity on falls and quality of life. *Dyn Med*, 2008, 7: 4. [[Medline](#)] [[CrossRef](#)]
- 8) Forhan M, Gill SV: Obesity, functional mobility and quality of life. *Best Pract Res Clin Endocrinol Metab*, 2013, 27: 129–137. [[Medline](#)] [[CrossRef](#)]
- 9) Hergenroeder AL, Wert DM, Hile ES, et al.: Association of body mass index with self-report and performance-based measures of balance and mobility. *Phys Ther*, 2011, 91: 1223–1234. [[Medline](#)] [[CrossRef](#)]
- 10) Vincent HK, Vincent KR, Lamb KM: Obesity and mobility disability in the older adult. *Obes Rev*, 2010, 11: 568–579. [[Medline](#)] [[CrossRef](#)]
- 11) Khazzani H, Allali F, Bennani L, et al.: The relationship between physical performance measures, bone mineral density, falls, and the risk of peripheral fracture: a cross-sectional analysis. *BMC Public Health*, 2009, 9: 297. [[Medline](#)] [[CrossRef](#)]
- 12) Artaud F, Singh-Manoux A, Dugravot A, et al.: Decline in fast gait speed as a predictor of disability in older adults. *J Am Geriatr Soc*, 2015, 63: 1129–1136. [[Medline](#)] [[CrossRef](#)]
- 13) Podsiadlo D, Richardson S: The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc*, 1991, 39: 142–148. [[Medline](#)] [[CrossRef](#)]
- 14) Salbach NM, O’Brien K, Brooks D, et al.: Speed and distance requirements for community ambulation: a systematic review. *Arch Phys Med Rehabil*, 2014, 95: 117–128.e11. [[Medline](#)] [[CrossRef](#)]
- 15) Shumway-Cook A, Brauer S, Woollacott M: Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther*, 2000, 80: 896–903. [[Medline](#)] [[CrossRef](#)]
- 16) Cho BL, Scarpace D, Alexander NB: Tests of stepping as indicators of mobility, balance, and fall risk in balance-impaired older adults. *J Am Geriatr Soc*, 2004, 52: 1168–1173. [[Medline](#)] [[CrossRef](#)]
- 17) Muir SW, Berg K, Chesworth B, et al.: Quantifying the magnitude of risk for balance impairment on falls in community-dwelling older adults: a systematic review and meta-analysis. *J Clin Epidemiol*, 2010, 63: 389–406. [[Medline](#)] [[CrossRef](#)]
- 18) Rubenstein LZ: Falls in older people: epidemiology, risk factors and strategies for prevention. *Age Ageing*, 2006, 35: ii37–ii41. [[Medline](#)] [[CrossRef](#)]
- 19) Verghese J, Holtzer R, Lipton RB, et al.: Quantitative gait markers and incident fall risk in older adults. *J Gerontol A Biol Sci Med Sci*, 2009, 64: 896–901. [[Medline](#)] [[CrossRef](#)]
- 20) Ostbye T, Dement JM, Krause KM: Obesity and workers’ compensation: results from the Duke Health and Safety Surveillance System. *Arch Intern Med*, 2007, 167: 766–773. [[Medline](#)] [[CrossRef](#)]
- 21) Yarborough CM 3rd, Brethauer S, Burton WN, et al.: Obesity in the workplace: impact, outcomes, and recommendations. *J Occup Environ Med*, 2018, 60: 97–107. [[Medline](#)] [[CrossRef](#)]
- 22) Kear BM, Guck TP, McGaha AL: Timed Up and Go (TUG) Test: Normative reference values for ages 20 to 59 years and relationships with physical and mental health risk factors. *J Prim Care Community Health*, 2017, 8: 9–13. [[Medline](#)] [[CrossRef](#)]
- 23) Bohannon RW: Comfortable and maximum walking speed of adults aged 20–79 years: reference values and determinants. *Age Ageing*, 1997, 26: 15–19. [[Medline](#)] [[CrossRef](#)]
- 24) Goldberg A, Schepens S, Wallace M: Concurrent validity and reliability of the maximum step length test in older adults. *J Geriatr Phys Ther*, 2010, 33: 122–127. [[Medline](#)] [[CrossRef](#)]
- 25) Bohannon RW: Test-retest reliability of the five-repetition sit-to-stand test: a systematic review of the literature involving adults. *J Strength Cond Res*, 2011, 25: 3205–3207. [[Medline](#)] [[CrossRef](#)]
- 26) Bohannon RW, Bubela DJ, Magasi SR, et al.: Sit-to-stand test: performance and determinants across the age-span. *Isokinet Exerc Sci*, 2010, 18: 235–240.

[\[Medline\]](#) [\[CrossRef\]](#)

- 27) Jackson AS, Blair SN, Mahar MT, et al.: Prediction of functional aerobic capacity without exercise testing. *Med Sci Sports Exerc*, 1990, 22: 863–870. [\[Medline\]](#) [\[CrossRef\]](#)
- 28) Mason C, Katzmarzyk PT: Variability in waist circumference measurements according to anatomic measurement site. *Obesity (Silver Spring)*, 2009, 17: 1789–1795. [\[Medline\]](#) [\[CrossRef\]](#)
- 29) Ashwell M, Gunn P, Gibson S: Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. *Obes Rev*, 2012, 13: 275–286. [\[Medline\]](#) [\[CrossRef\]](#)
- 30) Cesari M, Kritchevsky SB, Penninx BW, et al.: Prognostic value of usual gait speed in well-functioning older people--results from the Health, Aging and Body Composition Study. *J Am Geriatr Soc*, 2005, 53: 1675–1680. [\[Medline\]](#) [\[CrossRef\]](#)
- 31) Medell JL, Alexander NB: A clinical measure of maximal and rapid stepping in older women. *J Gerontol A Biol Sci Med Sci*, 2000, 55: M429–M433. [\[Medline\]](#) [\[CrossRef\]](#)
- 32) Silva PF, Quintino LF, Franco J, et al.: Measurement properties and feasibility of clinical tests to assess sit-to-stand/stand-to-sit tasks in subjects with neurological disease: a systematic review. *Braz J Phys Ther*, 2014, 18: 99–110. [\[Medline\]](#) [\[CrossRef\]](#)
- 33) Koo TK, Li MY: A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*, 2016, 15: 155–163 Erratum in: *J Chiropr Med*, 2017, 16: 346. [\[Medline\]](#) [\[CrossRef\]](#)
- 34) Mokkink LB, Terwee CB, Patrick DL, et al.: The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*, 2010, 63: 737–745. [\[Medline\]](#) [\[CrossRef\]](#)
- 35) Haley SM, Fragala-Pinkham MA: Interpreting change scores of tests and measures used in physical therapy. *Phys Ther*, 2006, 86: 735–743. [\[Medline\]](#) [\[CrossRef\]](#)
- 36) Mukaka MM: Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*, 2012, 24: 69–71. [\[Medline\]](#) [\[CrossRef\]](#)
- 37) Dite W, Temple VA: A clinical test of stepping and change of direction to identify multiple falling older adults. *Arch Phys Med Rehabil*, 2002, 83: 1566–1571. [\[Medline\]](#) [\[CrossRef\]](#)
- 38) Blennerhassett JM, Jayalath VM: The Four Square Step Test is a feasible and valid clinical test of dynamic standing balance for use in ambulant people post-stroke. *Arch Phys Med Rehabil*, 2008, 89: 2156–2161. [\[Medline\]](#) [\[CrossRef\]](#)
- 39) Whitney SL, Marchetti GF, Morris LO, et al.: The reliability and validity of the Four Square Step Test for people with balance deficits secondary to a vestibular disorder. *Arch Phys Med Rehabil*, 2007, 88: 99–104. [\[Medline\]](#) [\[CrossRef\]](#)
- 40) Dobkin BH: Short-distance walking speed and timed walking distance: redundant measures for clinical trials? *Neurology*, 2006, 66: 584–586. [\[Medline\]](#) [\[CrossRef\]](#)
- 41) Schoene D, Wu SM, Mikolaizak AS, et al.: Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: systematic review and meta-analysis. *J Am Geriatr Soc*, 2013, 61: 202–208. [\[Medline\]](#) [\[CrossRef\]](#)