Published in final edited form as:

Laryngoscope. 2018 July ; 128(7): 1687–1691. doi:10.1002/lary.27022.

Tandem walking as a quick screening test for vestibular disorders

Helen S. Cohen, EdD¹, Jasmine Stitz, MS², Haleh Sangi-Haghpeykar, PhD³, Susan P. Williams, MD⁴, Ajitkumar P. Mulavara, PhD⁵, Brian T. Peters, PhD⁵, and Jacob J. Bloomberg, PhD⁶

¹Bobby R Alford Department of Otolaryngology – Head and Neck Surgery, Baylor College of Medicine, Houston, TX

²University of Applied Sciences/ Upper Austria, Linz, Austria

³Department of Obstetrics and Gynecology, Baylor College of Medicine, Houston, TX

⁴Department of Medicine, Baylor College of Medicine, Houston, TX

⁵KBRwyle, Houston, TX

⁶Neuroscience Laboratories, NASA/ Johnson Space Center

Abstract

Objectives—Although many screening tests of balance are available few of them have been well-validated for clinical or research uses. The goal of this study was to test an updated version of an old test, Tandem Walking, to determine how useful it is for screening patients with vestibular disorders.

Methods—Subjects were 90 adult patients with vestibular disorders and 292 healthy adult controls. They were tested on the number of correct tandem steps they could perform with arms crossed and eyes closed in a series of 10 steps. Correct steps could be nonconsecutive. Subjects were given one practice trial with eyes open and three experimental trials with eyes closed.

Results—Receiver operator characteristics (ROC), and sensitivity and specificity were calculated. ROC values, sensitivity, and specificity were, at best, only moderate, no matter how the age range was cut. Even for subjects in the age group with the highest ROV value, i.e. age less than 50 years, ROC = 0.8, sensitivity = 0.77 and specificity = 0.72.

Conclusion—These results indicate that 23% of patients will not be identified. Therefore we recommend that if this test is used for screening patients in the clinic or healthy volunteers the result should be interpreted with care.

Level of evidence—3b case-control study

Corresponding author: Helen S Cohen, EdD, Dept. of Otolaryngology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, hcohen@bcm.edu. Telephone: 713-7998-7702. Fax: 713-798-8658.

Conflicts of Interest: None

Meetings: Preliminary data were presented at the Mid-Winter Meeting of the Association for Research in Otolaryngology, Baltimore, MD, February 2017, and the World Congress of the International Society for Posture and Gait Research, Fort Lauderdale, FL, June 2017

Introduction

Tandem walking is widely used to screen patients for neurologic and vestibular disorders, and to screen automobile drivers for suspected intoxication. A fairly old test (1, 2), it is easy to administer. Normal people can perform it well with eyes open (3), although performance declines with age (4). Performing it with eyes closed is so challenging for patients with vestibular impairments that Longridge and Mallinson recommended against using it (5). It has been shown to be useful for screening patients with lower extremity peripheral neuropathy when performed with eyes closed and when the dependent measure is the number of consecutive tandem steps out of 10 (6) but the same test has been reported to have poor sensitivity when used to screen patients with vestibular disorders (7). Nevertheless, the test remains popular among clinicians.

We hypothesized that sensitivity would be improved when the test was given with eyes closed and the dependent measure was the total number of tandem steps out of ten, instead of the number of consecutive tandem steps out of ten. Counting the total number of tandem steps out of ten has been used previously to characterize changes in dynamic postural stability after spaceflight (8). The goal of the present study was to test that hypothesis.

Methods

Subjects

The 292 healthy control subjects with no otologic, neurologic or significant musculoskeletal impairments were compared to 90 patient subjects who had a variety of vestibular impairments, shown in Tables 1 and 2. Patients also had no significant musculoskeletal impairments or other neurologic disorders. Control subjects were recruited from the geriatric medicine caseload after screening by geriatricians, and from staff and visitors to our institution. Patient subjects were recruited from the caseloads of the senior author and other clinicians in the otology section of the department of otolaryngology. Patients had been diagnosed with vestibular disorders by board-certified physicians, mostly otolaryngologists and neurologists, based on clinical examinations, and findings from objective diagnostic testing including cervical vestibular evoked myogenic potentials, low frequency sinusoidal tests of the vestibulo-ocular reflex in darkness, Dix-Hallpike and other positional tests, and bi-thermal caloric testing (> 20% unilateral weakness), recorded with infra-red video-oculography. Patient subjects may also have had other laboratory tests based on the physician's judgement. Patients who had benign paroxysmal positional vertigo were symptomatic at the time of testing, and were treated subsequent to participation in this study. All subjects gave informed consent prior to participating. The study was approved by the Institutional Review Board at the senior author's institution.

Methodology

The test was performed in a well-lighted room with a vinyl tile floor. To standardize footwear subjects wore socks but no shoes. We loaned athletic socks to subjects who arrived without socks. To collect data on kinematic measures of head and torso motions subjects wore two inertial motion units (IMUs, Xsens North America Inc., Los Angeles), 5.25×3.75

× 2 cm, weight 28.3 g. One was centered on a plastic band atop the head and one was on a lightweight nylon vest at the mid-thoracic level of the back (6).

Subjects were asked to walk 10 steps, heel-to-toe, without spaces between the steps, with their arms crossed. Each subject performed a practice trial with 3 to 5 steps to get the idea of the movement and then performed a single trial of 10 steps with eyes open. Then the subject performed three trials of 10 steps with eyes closed. The observer recorded the number of correct tandem steps, whether or not they were consecutive. An error was defined as taking a side step, leaving more than 2 cm between the feet, and opening the eyes.

Statistical methods

Patients (N=90) and controls (N=292) were compared by t-tests for continuous variables, chi-square test for grouped data, and Wilcoxon sum rank for some pairwise comparisons. Logistic regression was performed for analysis of receiver operative characteristic (ROC) and to determine cut-points of performing correct tandem steps with eyes open or closed for the best pairs of sensitivity/specificity. From look-up tables generated by the analysis software the optimal pairs of sensitivity and 1-specificity were selected, using the rule that sensitivity and specificity should be as close to equal as possible. If they could not be equal then the pair selected had the closest possible match but greater sensitivity was preferred over lesser 1-specificity. The cut-points for the selected pairs of sensitivity/ 1-specificity were obtained from look-up tables generated by the analyses. Analyses were performed for different age groups. All analysis were done in SAS statistical software (9.4, Cary, NC).

Results

We tested the data from the two largest patient groups, with all ages combined, to determine if they differed from controls. On the eyes open trial as well as on all three eyes closed trials, the UW patients and the BPPV patients had significantly fewer correct tandem steps than controls, $p < 0.05$. See Table 3.

The number of correct tandem steps for the three trials with eyes closed were examined separately, with all ages combined for each trial and then with the age ranges broken into groups. The data are shown separately for the first, second and third trials with eyes closed. See Tables 4, 6 and 7, respectively. ROC values were calculated with the entire sample and without the BPPV subjects. Sample sizes per age group are given in Table 4. Note that even for age ranges with moderately good ROC values we found no combinations of high sensitivity and high specificity for any trials. Removing the BPPV patients from the analyses reduced the sample size and did not improve the ROC values. Another way to consider these data, based on the same analyses, is to look at the percentages of true and false positives and negatives, i.e. the percentages of patients who are correctly or incorrectly identified (true and false positives, respectively) and the percentages of controls who are correctly identified (true and false negatives, respectively). Clinicians may be more familiar with that type of analysis. Therefore, we have included Table 5, which shows the data from Table 4 as true and false negatives and positives. See, also Figures 1a and 1B.

Subjects walked significantly faster on succeeding trials, $p=0.001$, suggesting that they learned the task. Overall, the trial duration was slightly but significantly shorter for controls (mean 14.9 sec, SE 0.3) than for patients (mean 17.1 sec, SE 0.5), $p<0.001$. Figures 2a and b illustrate the changes in trial duration by age for controls and for patients, respectively. Although the groups are statistically significantly different this difference is not sufficient to be used as a variable for clinical screening for vestibular disorders.

For kinematic analyses the following root mean square values of the IMU variables for the trunk segment were quantified and used for further analysis: resultant acceleration (TAR), angular velocity about the roll axis (TRV), angular velocity about the pitch axis (TPV), angular velocity about the yaw axis (TYV). ROC values were calculated by age group for each measure. Then ROC values were calculated for each measure combined with the number of steps. For all age groups, ROC values of individual kinematic measures were < 0.75 . When kinematics were combined with the number of steps, ROC values were 0.78 to 0.79 for subjects younger than age 50; 0.79 for subjects younger than age 60; 0.75 to 0.77 for subjects aged 50 to 60; 0.73 to 0.74 for subjects aged 60 or older; 0.78 to 0.80 for subjects aged 70 to 79; and 0.69 to 0.70 for subjects aged 80 and older. Thus no individual kinematic measures or combinations with the number of steps improved the ROC values obtained by using just the number of steps.

Discussion

The two largest groups, unilateral weakness and BPPV, had significantly fewer correct tandem steps than controls. The finding about the UW patients was expected. The finding about BPPV patients, however, is new and unexpected. It suggests that these patients have impaired dynamic balance skills.

The results from the ROC analyses indicate that tandem walking, alone, is not an optimal screening measure to determine if someone has a vestibular disorder when that person is not acutely ill. We found moderate ROC values, sensitivity and specificity. This finding supports and extends previous work by examining a different breakdown by age groups, trials and the way in which steps were counted (5, 7). The decrease in performance with older age was expected, based on work done on standing balance in other studies. The sensitivity and specificity of testing decreased over trials for the oldest group, probably due to fatigue. The same norms used for young adults should not be used for middle-aged and older adults. This finding is in contrast to previous findings with peripheral neuropathy patients who performed this test, when the dependent measure was the number of consecutive tandem steps (6).

Removing the BPPV patients from the analyses did not result in improved ROC values. If BPPV patients had had normal balance then we would have seen an improvement in ROC values. These patients have been shown to have impaired performance on tests of standing balance and subjective visual vertical (9-11). These changes are probably mediated by both loading of the posterior semicircular canal with otoconia and unloading of the utricles, with concomitant changes in signals to the central vestibular nuclei. This idea is supported by our

new finding of decreased scores on this test by BPPV patients. Thus, benign paroxysmal positional vertigo is probably not so benign.

The results from the kinematic analyses suggested that using those measures does not improve the ability to separate healthy controls from patients with vestibular disorders. Although vestibularly impaired patients are known to have balance disorders such a fine-grained analysis of their movements did not appreciably augment the ability to screen for them in this context, which was similar to previous work with peripheral neuropathy patients (6).

We did not test patients who were acutely ill, in either an emergency room or a hospital. Thus we have no information about the value of this test when screening for acute illness. The acute stage of recovery from a vestibular insult may be similar to acutely post-flight astronauts when they have returned from space flight. Tandem walking when tested using the measure of total correct steps, as we did in this study, is useful for detecting change in dynamic postural stability in post-flight astronauts (8). The test might be useful for examining change over time, for examining acutely ill patients, or for assessing performance of patients with other types of motor disorders. This test might be useful for primary care physicians for assessing the influence of medications or rehabilitation interventions on gait and balance. A version of tandem walking has been used for a long time in roadside sobriety tests. It is not clear if the norms used during such tests should be updated to account for age. Those questions are beyond the scope of this study and were not addressed here.

Several problems may affect use of this test. By using two training trials with eyes open and three test trials with eyes closed we eliminated the influence of learning and the fear of possibly falling. We also provided close guarding so that subjects felt secure. Common musculoskeletal problems such as arthritis and foot deformities may affect balance enough to influence test performance. We did not observe any such deformities in our subjects and we excluded subjects with severe musculoskeletal problems.

Peripheral neuropathy affects performance on this test (6). Neurologic problems, attentional deficits and probably psychiatric problems affect performance. We had no subjects who had those problems. Despite having excluded potentially problematic subjects and eliminating other confounding variables, even with the younger group of subjects tandem walking without any other screening test was not a strong test to screen for clinically relevant vestibular disorders. Also, clinicians should be aware that the test screens for decreased dynamic balance, which might be consistent with impaired vestibular function via the vestibulospinal tracts. The level of information from this test, using the dependent measure of interest in this study, i.e., the number of total correct tandem steps, would not help the clinician to distinguish between peripheral and central vestibular disorders. The quality of motor performance might provide a neurologist with additional information that was not tested in this study.

Acknowledgments

The authors thank Chris Miller, KBRwyle, and the staffs of the Center for Balance Disorders and the Geriatric Medicine Clinic, Baylor College of Medicine, for their assistance.

Funding: NIH grant 2R01DC009031 (HSC), National Space Biomedical Research Institute through NASA NCC 9-58 (APM, JJB), and a fellowship from the Austria Marshall Plan Foundation (JS).

References

1. Fregly AR, Graybiel A. Labyrinthine defects as shown by ataxia and caloric tests. *Acta Otolaryngol.* 1970; 69:216–22. [PubMed: 4911180]
2. Graybiel A, Fregly AR. A new quantitative ataxia test battery. *Acta Otolaryngol.* 1966; 61:292–312. [PubMed: 5956097]
3. Kammerlind AS, Larsson PB, Ledin TEA, Skargren EIB. Reliability of clinical balance tests and subjective ratings in dizziness and disequilibrium. *Adv Physiother.* 2005; 7:96–107.
4. Vereeck L, Wuyts F, Truijzen S, Van de Heyning PH. Clinical assessment of balance: normative data, and gender and age effects. *Int J Audiol.* 2008; 47(2):67–75. [PubMed: 18236239]
5. Longridge NS, Mallinson AI. Clinical Romberg testing does not detect vestibular disease. *Otol Neurotol.* 2010; 31(5):803–10. [PubMed: 20502374]
6. Cohen HS, Mulavara AP, Peters BT, Sangi-Haghpeykar H, Kung DH, Mosier DR, et al. Sharpening the tandem walking test for screening peripheral neuropathy. *South Med J.* 2013; 106:565–9. [PubMed: 24096950]
7. Cohen HS, Mulavara AP, Peters BT, Sangi-Haghpeykar H, Bloomberg JJ. Tests of walking balance for screening vestibular disorders. *J Vestib Res.* 2012; 22:95–104. [PubMed: 23000609]
8. Bloomberg JJ. A comparison of tandem walk performance between bed rest subjects and astronauts. 39th Annual Meeting of the American Society of Biomechanics; August 5-8, 2015; Columbus, OH. 2015.
9. Mulavara AP, Cohen HS, Peters BT, Sangi-Haghpeykar H, Bloomberg JJ. New analyses of the Sensory Organization Test compared to the Clinical Test of Sensory Integration and Balance in patients with benign paroxysmal positional vertigo. *Laryngoscope.* 2013; 123:2276–80. [PubMed: 23553110]
10. Cohen HS, Kimball KT. Effectiveness of treatments for benign paroxysmal positional vertigo of the posterior canal. *Otol Neurotol.* 2005; 26:1034–40. [PubMed: 16151355]
11. Agarwal K, Bronstein AM, Faldon ME, Mandala M, Murray K, Silove Y. Visual dependence and BPPV. *J Neurol.* 2012; 259:1117–24. [PubMed: 22113702]

Figure 1.
Number of correct tandem steps out of ten, by age. A. Controls. B. Patients.

Figure 2.
Duration of the first trial of tandem walking, by age. A. Controls. B. Patients.

Table 1
Demographic data

	Age (mean years, SD, ranges)	Sex (males, females)
Controls (n=292)	55.1 ± 18.9, 21-95	111 M, 181 F
Patients (n=90)	59.7 ± 14.4, 24-88	36 M, 54 F

Table 2

Diagnoses of patient subjects. Benign paroxysmal positional vertigo (BPPV). Unilateral, peripheral vestibular weakness on bi-thermal caloric testing > 20% (UW), consistent with labyrinthitis or vestibular neuronitis. Meniere's disease (Meniere's). Acoustic neuroma (AN).

Disorder	Sample size
UW	33
BPPV	28
Dual BPPV and UW	8
Meniere's	7
Vestibular Migraine	3
Pre-operative AN	5
Post-operative AN	2
Bilateral vestibular weakness	2
Dual BPPV and migraine	2

Table 3

Comparisons between UW and BPPV group vs. controls. Mean number of correct tandem steps, rounded to the nearest integer (SD).

Group	Eyes open	Eyes closed Trial 1	Eyes closed Trial 2	Eyes closed Trial 3.
Controls	9 (2.2)	7 (3.3)	10 (5.2)	7 (3.4)
UW	7 (3.7)	4 (3.6)	4 (3.5)	4 (3.7)
BPPV	7 (3.5)	4 (3.2)	4 (3.6)	4 (3.6)

Table 4

Eyes closed trial 1: all subjects (all S's), ROC, sensitivity and specificity by age groups for the first eyes-closed trial. Recommended cut-points are given as the number of correct tandem steps. The sample sizes of controls and patients per age group are given in parentheses (C/P). Without BPPV the sample size for patients is smaller and ROC values are different. See ROC no BPPV for comparison.

Age	ROC all S's	ROC no BPPV	Sensitivity	Specificity	Cut-point
all ages (292 C/90P)	0.74	0.75	0.67	0.71	5
< 50 (116 C/ 22P)	0.79	0.82	0.77	0.65	8
< 60 (154 C/ 43P)	0.79	0.81	0.70	0.75	7
50 to 69 (133 C/ 47P)	0.74	0.76	0.74	0.6	6
60 (133C/47P)	0.72	0.72	0.68	0.62	4
70 to 79 (49C/ 18P)	0.77	0.73	0.78	0.61	3
80 (32C/ 7P)	0.72	0.80	0.72	0.57	2

Table 5

For eyes closed trial 1, all subjects, percentage of true and false positives (+) and negatives (-). Note that the True + and True- are the same as the sensitivity and specificity in Table 3.

Age	Patients		Controls	
	True +	False -	True -	False +
all ages	67	33	71	29
< 50	77	23	65	35
< 60	70	30	75	25
50 to 69	74	26	60	40
60	68	32	62	38
70 to 79	78	22	61	39
80	71	29	56	44

Table 6

Eyes closed trial 2: all subjects (all S's) ROC, sensitivity and specificity by age groups for the second eyes-closed trial. Recommended cut-points are given as the number of correct tandem steps. Without BPPV the sample size for patients is smaller and ROC values are different. See ROC no BPPV for comparison.

Age	ROC All S's	ROC No BPPV	Sensitivity	Specificity	Cut-point
all ages	0.725	0.721	0.69	0.64	6
< 50	0.80	0.78	0.77	0.72	8
< 60	0.75	0.76	0.73	0.62	8
50 to 69	0.71	0.74	0.71	0.64	6
60	0.72	0.71	0.68	0.65	3
70 to 79	0.76	0.74	0.72	0.61	3
80	0.70	0.74	0.57	0.61	1

Table 7

Eyes closed trial 1: all subjects ROC, sensitivity and specificity by age groups for the third eyes-closed trial. Recommended cut-points are given as the number of correct tandem steps. Without BPPV the sample size for patients is smaller and ROC values are different. See ROC no BPPV for comparison.

Age	ROC All S's	ROC no BPPV	Sensitivity	Specificity	Cut-point
all ages	0.71	0.70	0.61	0.64	6
< 50	0.71	0.67	0.73	0.64	9
< 60	0.74	0.67	0.67	0.67	8
50 to 69	0.72	0.74	0.76	0.52	7
60	0.68	0.66	0.65	0.63	3
70 to 79	0.74	0.70	0.67	0.61	2
80	0.63	0.71	0.57	0.5	1