Utility of Stepping, Walking, and Head Impulses for Screening Patients for Vestibular Impairments

Helen S. Cohen, EdD, OTR§1, Haleh Sangi-Haghpeykar, PhD2, Natalia A. Ricci, PhD, PT3, June Kampangkaew4, and Robert A. Williamson, MD1

Abstract

Objective. To determine if some common screening tests predict scores on detailed, objective diagnostic tests of the vestibular system.

Study Design. Sixty patients with vestibular disorders were compared with 60 asymptomatic controls.

Setting. Vestibular diagnostic laboratory, tertiary care center.

Subjects and Methods. Subjects were screened with head impulse tests, Fukuda stepping tests while walking and marching in place, and tandem walking tests with eyes open and closed. All subjects underwent bithermal caloric tests and Dix-Hallpike maneuvers; patients underwent low-frequency sinusoidal tests of the vestibulo-ocular reflex in darkness and cervical vestibular evoked myogenic potentials.

Results. On tandem walking tests, patients differed significantly from controls, but receiver operating characteristic scores were < 0.80. On Fukuda tests, patients turned significantly more than controls for walking but not marching, but receiver operating characteristic values were considerably less than 0.80. On head impulse tests, patients with bithermal caloric weakness (≥20% and <60%) did not differ from controls, but patients with severe bithermal caloric weakness (≥60%) differed significantly from controls. Receiver operating characteristic values were >0.80 only for subjects with severe bithermal caloric weakness and were highest, at 0.88, for subjects with severe weakness and age ≥ 60 years.

Conclusions. The Fukuda test is a poor screening test because it does not correlate well with objective test findings. Tandem walking is best used for screening older patients for vestibular disorders. Positive findings on a head impulse test are probably consistent with severe peripheral vestibular impairment and may be most useful in older patients. In younger patients with vertigo, negative results on head impulse tests may not be informative.

Keywords
vestibular system, diagnosis, tandem walking, Fukuda stepping test, head impulse test

Received November 14, 2013; revised January 17, 2014; accepted February 21, 2014.

Introduction

Clinicians use a variety of screening tests in the office to assess patients who they suspect have vestibular disorders. Although some tests are established and widely used, they may have limited value for screening patients to detect vestibular impairments. In this study, we examined 3 such tests: the Fukuda stepping test, the tandem walking test, and the head impulse test. All of these tests have the value of being inexpensive and easy to administer, with minimal equipment, in <1 minute per test.

To perform the Fukuda stepping test, the patient walks in place with eyes closed. The dependent measure is the angle turned. Fukuda1 modified tests described by earlier investigators2 by having the blindfolded subject stand in the middle of a circle with 15° increments marked on the floor and having the subject walk at approximately 2 Hz for 50 to 100 steps. Dependent measures were postural sway, changes in the relative positions of the head and body, the angle of rotation, and the distance of rotation, used to calculate the angle of displacement from straight ahead. Fukuda reported that patients with vestibular impairments rotated and...
deviated more than controls. Toussaint et al\(^3\) showed that scores were significantly worse when the test was performed with the head pitched downward.

Several studies have indicated poor test-retest reliability of various versions of the test. Using Fukuda’s original paradigm, Bonanni and Newton\(^4\) found considerable variability and only moderate test-retest reliability. Several studies have shown that the test has poor sensitivity and specificity, does not indicate the side of the lesion reliably, and is variable even in normal controls.\(^5\)-\(^8\) Despite these problems, the Fukuda stepping test continues to be used in many clinics.

Similarly, the well-known tandem walking test has been used for many years.\(^9\),\(^10\) Originally developed for use on rails but performed with eyes open, tandem walking is now typically performed with eyes closed. Performance seems to decline slightly with age.\(^11\)-\(^14\) It is challenging for patients as well as for healthy controls.\(^15\) When performed with 10 steps and eyes closed, using patients and controls who had no joint limitations, the test had moderately high sensitivity for lower extremity peripheral neuropathy.\(^16\) In subjects without peripheral neuropathy or joint replacement, it had high specificity compared with healthy controls but poor sensitivity for unilateral vestibular weakness.\(^17\) The patients in that study, however, were fairly homogenous and may not have represented the general population of patients with vestibular impairments. Also, control subjects were screened with Dix-Hallpike maneuvers and head impulse tests but were not tested on objective diagnostic tests.

The head impulse test performed without instrumentation\(^18\),\(^19\) is newer than the other tests but has become well established and is now recommended for use by therapists as well as physicians.\(^20\),\(^21\) The test is performed by having the patient stare at a central focal point, such as the examiner’s nose, and then the examiner briskly rotates the patient’s head in yaw, either left or right, approximately 20°. The results are positive (abnormal) if the examiner observes a saccade during counter-rotation of the eye.

The purpose of this study was to determine the utility of these 3 tests in screening patients for vestibular disorders. Specifically, we examined the Fukuda stepping test performed while walking in place with small steps and marching in place with higher steps, the tandem walking test performed with eyes open and eyes closed, and the uninstrumented head impulse test. We included patients with peripheral neuropathy and joint replacements if they were ambulatory without gait aids, and we verified that healthy controls had no unilateral weaknesses on bithermal caloric testing.

**Methods**

**Subjects**

Two groups of ambulatory participants, 60 per group, were recruited. No subject was excluded because of peripheral neuropathy or joint replacements. The asymptomatic controls, including 33 women and 27 men, were screened with a health history and Dix-Hallpike maneuvers and were given bithermal caloric tests, on which they all had normal-range responses (ie, unilateral weakness of <20%). The patient subjects, including 28 women and 32 men, were recruited from the caseload of patients by the senior author and from patients who were tested in the laboratory with a complete battery of vestibular diagnostic tests, including bithermal caloric tests, Dix-Hallpike maneuvers, low-frequency sinusoidal tests of the vestibulo-ocular reflex in darkness in the rotatory chair, and cervical vestibular evoked myogenic potentials. No patients in this study had benign paroxysmal positional vertigo or abnormal responses on Dix-Hallpike tests; all patients had ≥20% unilateral weakness on bithermal caloric tests, and the neurootologist who read those tests (R.A.W.) determined that each patient’s responses were consistent with a peripheral vestibular weakness. Control subjects were slightly but not statistically significantly younger than patient subjects, with a mean age of 54.0 ± 14.5 years (range, 23-66 years) and 38 subjects aged <60 years compared with a mean age among patients of 58.7 ± 13.5 years (range, 35-86 years) and 32 subjects aged <60 years (\(P = .08\)).

All subjects gave informed consent before participation in the study. The study was approved by the Institutional Review Board for Baylor College of Medicine and Affiliated Hospitals.

**Procedure**

Subjects were tested on tandem walking, 10 steps per condition, with the arms crossed, under 2 conditions: eyes open and eyes closed. The test was scored by the number of consecutive heel-to-toe steps taken without stepping out of line, moving the arms, or opening the eyes during eyes-closed tests. Subjects were tested on the Fukuda stepping test on a plastic mat on which a large semicircle was marked in 10° increments. Subjects stood at the 0° mark, held their arms at their sides, closed their eyes, and performed 2 conditions for 20 steps each: walking in place and marching in place. For the walking condition, subjects were instructed to walk in place and lift their feet off the floor. For the marching condition, subjects were instructed to lift their knees as high as possible. The dependent measures were the angle of rotation away from the center line, the distance walked forward and the distance walked to the side.

For the head impulse test, the subject sat in a comfortable chair and the examiner crouched in front of the subject. The examiner held the subject’s head with both hands and, after briefly assessing passive range of motion in yaw to be sure that the participant had at least 30° of yaw rotation to either side, pitched the head downward 30° and then briskly moved the head 20° to one side while the subject stared at the examiner’s nose. Two trials per side were performed. The dependent measure was the presence or absence of a saccade. To determine interrater reliability, data were collected by 2 raters simultaneously for 84 subjects.

Examiners who performed head impulse tests knew which subjects were patients and controls but did not know...
the levels of unilateral weakness in patients. Thus, testing was partially blinded.

**Statistical Analysis**

Patients and controls were initially compared using t tests for continuous variables and $\chi^2$ tests for grouped variables. We examined the validity of the various tests using receiver operating characteristic (ROC) curves for continuous variables and sensitivity and specificity for discrete variables. These analyses were also stratified by age (<60 and ≥60 years) because standing balance decreases significantly around age 60 years, and vestibular function decreases with age. Interobserver reliability of the tests was assessed using $\kappa$ coefficients for discrete variables and correlation coefficients for continuous variables. All analyses were performed using SAS version 9.3 (SAS Institute, Inc, Cary, North Carolina). $P$ values < .05 were considered statistically significant.

**Results**

**Tandem Walking**

Tests for eyes open and eyes closed were performed separately. For both conditions, the total control and patient groups differed significantly ($P < .0001$). Patients aged <60 years and those aged ≥60 years also differed significantly from controls ($P < .0008$ to $P < .0001$; Table 1). More importantly for the value of the test for screening, however, with the age groups collapsed and also with the age groups divided at age 60 years, ROC values were mostly less than 0.80. As shown in Table 1, the ROC value was higher only for subjects aged ≥60 years, at 0.83. Sensitivity was 75% and specificity 78%. The level of agreement between the 2 observers was high for eyes open ($r = 0.99$) and for eyes closed ($r = 0.93$).

**Fukuda Stepping Test**

Similarly, for the Fukuda stepping test, we used the entire cohort and then divided the groups at age 60 years. Patients turned significantly more than controls when walking ($P = .002$) but not when marching ($P = .47$), for the entire cohort. No differences were found between the groups in the number of steps taken, the distance walked forward, or the distance walked to the side. When the groups were divided by age, for both younger and older groups in the walking condition, but not the marching condition, patients turned significantly more than controls (younger, $P = .02$; older, $P = .03$). No other significant differences were found between patients and controls in younger and older groups (Table 2).

As shown in Table 2, ROC values were well below 0.80 for all measures. Therefore, sensitivity and specificity tests were not performed. Interobserver reliability was high for all subtests ($r = 0.95-1.0$), except the distance walked forward in the walking, as opposed to the marching, condition ($r = 0.44$).

**Head Impulse Test**

For this test, we divided the age range and also divided patients into mild unilateral weakness (≥20% and <60%) and severe unilateral weakness (≥60%). For the total cohort and when the groups were broken down by age, significantly fewer controls than patients with mild unilateral weakness and significantly fewer controls than patients with severe unilateral weakness had no response to the test (total cohort, $P = .0003$; younger, $P < .0001$; older, $P = .0006$), but the number of positive responses did not differ significantly between groups (Table 3).

As shown in Table 4, with the total cohort and with groups broken out at age 60 years, ROC values were > 0.80 only for patients with severe disease. Not surprisingly, the best ROC value was found for older patients with severe disease, at 0.88, with 88.9% sensitivity and 86.4% specificity. The level of interobserver agreement was high, with $\kappa$ coefficients ranging from 0.90 to 0.95.

**Discussion**

The significant differences between groups were expected and are also misleading. On the basis of those significant
differences, clinicians may believe that these tests are useful for screening. As indicated by the weak ROC values, however, none of these tests is excellent for screening patients suspected of having vestibular disorders. The Fukuda stepping test is particularly poor, as no ROC values were \( \geq 0.80 \). One way to think about these findings is that tests show that groups differ at least somewhat. ROC values, however, indicate whether groups are different enough to be really useful for testing.

Because the ROC values were so low for the Fukuda stepping test, we recommend that this test not be used. The ROC values for tandem walking were also poor. It is most useful in the eyes-closed condition with older subjects. Clinicians may learn something further but should consider the possibility that patients may have peripheral neuropathy, central ataxia, or some other disorder causing a balance problem. For older patient who are able to perform the test, clinicians should not abandon the potential diagnosis of vestibular disorder, because the patient might have mild vestibular disease and still be able to perform the test. Thus, the ability to perform the test might lead clinicians who rely on the test too strongly to consider a false-negative.

The head impulse test is somewhat more complicated. Ideally, a test should have sensitivity and specificity \( \geq 95\% \). No test here meets that criterion. To be at least adequate for an initial screening, specificity and sensitivity

### Table 2. Fukuda Stepping Test.

<table>
<thead>
<tr>
<th>Group</th>
<th>Degrees Turned</th>
<th>Mean ± SD</th>
<th>Number of Steps</th>
<th>ROC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>UW</td>
<td>Controls</td>
<td>ROC Value</td>
</tr>
<tr>
<td>Walking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total group</td>
<td>8.7 ± 12.5</td>
<td>2.8 ± 6.0</td>
<td>0.66</td>
<td>20 (2-20)</td>
</tr>
<tr>
<td>Age &lt; 60 y</td>
<td>7.5 ± 9.0</td>
<td>2.8 ± 6.3</td>
<td>0.68</td>
<td>20 (14-20)</td>
</tr>
<tr>
<td>Age ≥ 60 y</td>
<td>10.0 ± 15.6</td>
<td>3.0 ± 5.5</td>
<td>0.63</td>
<td>20 (2-20)</td>
</tr>
<tr>
<td>Marching</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total group</td>
<td>12.2 ± 16.0</td>
<td>7.2 ± 10.7</td>
<td>0.58</td>
<td>20 (1-20)</td>
</tr>
<tr>
<td>Age &lt; 60 y</td>
<td>8.8 ± 12.0</td>
<td>5.0 ± 7.3</td>
<td>0.57</td>
<td>20 (3-20)</td>
</tr>
<tr>
<td>Age ≥ 60 y</td>
<td>16.0 ± 19.1</td>
<td>11.1 ± 14.2</td>
<td>0.58</td>
<td>20 (1-20)</td>
</tr>
</tbody>
</table>

### Table 3. Head Impulse Test: Number of Negative and Positive Responses by Age Group and Severity of Disease.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number Negative</th>
<th>Number Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>59</td>
<td>1</td>
</tr>
<tr>
<td>Younger a</td>
<td>38</td>
<td>0</td>
</tr>
<tr>
<td>Older b</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>UW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>All younger</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>All younger, mild c</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>All younger, severe d</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>All older</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>All older, mild</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>All older, severe</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

Abbreviations: ROC, receiver operating characteristic; UW, patients with unilateral weakness.

aROC values are presented for all tests. Sensitivity and specificity values are not shown, because no ROC values were \( \geq 0.80 \). All control subjects walked for 20 steps, and 19 control subjects marched for 20 steps.

differences, clinicians may believe that these tests are useful for screening. As indicated by the weak ROC values,
should be ≥85%. On the basis of that criterion, the head impulse test may be useful in detecting a severe impairment in an older patient, if the patient has adequate cervical range of motion, is able to relax enough to allow the clinician to perform the test, and is able to follow instructions for the test. Some older patients may not meet these criteria.24 Positive results may indicate severe peripheral vestibular impairment. Recent evidence suggests that positive results on an uninstrumented head impulse test in elderly patients may be associated with increased falls.25 Therefore, a positive finding suggests the need for further testing. Given the less than ideal sensitivity and specificity of the test, however, if the results are negative, the physician should not necessarily abandon the potential diagnosis of vestibular disease, because the patient may have a weakness too mild to be detected by the test. Therefore, relying on a negative result too strongly might lead the physician to consider a false-negative. The physician should consider the history and other aspects of the office screening when making the initial diagnosis.

This study had some limitations. Some inherent subjectivity on the part of the examiner is involved in determining the exact number of degrees turned in the Fukuda stepping test. We did not take into account the speed of stepping in either the Fukuda stepping test or the tandem walking test. We did not perform careful sensory examinations of subjects’ lower extremities for peripheral neuropathy or proprioceptive dysfunction. This omission may have skewed the data, but the effect would be expected to be mild because patients were screened by history of significant peripheral neuropathy. Head impulse testing can also vary by patient, related to age, speed of head movement, and range of motion, as well as comfort with the examination. Finally, a larger sample size would have allowed more thorough stratification of the subjects by age and by caloric weakness, rather than just the division between those older and younger than 60 years. Additional age-specific data would be useful clinically in this population.

Acknowledgment

We thank the staff of the Center for Balance Disorders for their invaluable assistance.

Author Contributions

Helen S. Cohen, design, data collection, interpreting data, manuscript writing and revision, final approval, agreement to be accountable; Hach Sangi-Haghpeykar, statistician, design, data analysis and interpretation, manuscript writing and revision, final approval, agreement to be accountable; Natalia A. Ricci, design, data collection, manuscript revision, final approval, agreement to be accountable; June Kampangkaew, design, data collection, manuscript revision, final approval, agreement to be accountable; Robert A. Williamson, design, data interpretation, manuscript revision, final approval, agreement to be accountable.

Disclosures

Competing interests: None.
Sponsorships: None.
Funding source: This work was supported by grant R01DC009031 from the National Institutes of Health.

References


