



Published in final edited form as:

Technol Innov. 2016 September ; 18(2-3): 193–201. doi:10.21300/18.2-3.2016.193.

PSYCHOMETRIC EVALUATION OF THE HILL ASSESSMENT INDEX (HAI) AND STAIR ASSESSMENT INDEX (SAI) IN HIGH-FUNCTIONING TRANSFEMORAL AMPUTEES

M. Jason Highsmith^{1,2,3}, Jason T. Kahle^{4,5}, Brian Kaluf⁶, Rebecca M. Miro¹, Larry J. Mengelkoch⁷, and Tyler D. Klenow⁸

¹School of Physical Therapy & Rehabilitation Sciences, University of South Florida, Tampa, FL, USA

²Extremity Trauma & Amputation Center of Excellence (EACE), U.S. Department of Veterans Affairs, Tampa, FL, USA

³319th Minimal Care Detachment, U.S. Army Reserves, Pinellas Park, FL, USA

⁴OP Solutions, Tampa, FL, USA

⁵Prosthetic Design + Research, Tampa, FL, USA

⁶Ability Prosthetics and Orthotics, Greenville, SC, USA

⁷Doctor of Physical Therapy Program, University of St. Augustine for Health Sciences, St. Augustine, FL, USA

⁸Prosthetics and Sensory Aids Service, James A. Haley Veterans' Hospital, Tampa, FL, USA

Abstract

The hill assessment index (HAI) and stair assessment index (SAI) were developed to objectively evaluate ramp and stair gait. This study's purpose was to determine the validity and reliability of these tests in a sample of persons with unilateral transfemoral amputation (TFA) using microprocessor prosthetic knee systems. All subjects were fit with a microprocessor knee system. After accommodation, subjects performed three trials ascending and descending a 5° ramp and a flight of stairs while being recorded on video. Sensitivity and specificity for the HAI was calculated against degree of asymmetry in step length using Dartfish video analysis software. Reliability was assessed using intraclass correlational coefficients calculated using Spearman's Rho (r_s). *A priori* significance level was set at $p = 0.05$. Twenty ($n = 20$) individuals with TFA completed the study protocol. Sensitivity and specificity of the HAI were calculated at 88.0% and 75.0% during ascending conditions and 94.0% and 67.0% during descending conditions, respectively. Significant correlations for the HAI included $r_s = 0.87$ and $r_s = 0.73$ within raters uphill and downhill, respectively. Corresponding coefficients of $r_s = 0.80$ and $r_s = 0.67$ were calculated between raters. For the SAI, significant correlations included $r_s = 1.00$ for both comparisons within raters and in the comparison between raters in the ascending condition. A correlation of $r_s = 0.89$ was calculated for the between-rater comparison in the descending

condition. The HAI showed moderate to excellent sensitivity and specificity but good to adequate reliability. The SAI showed excellent to good reliability.

Keywords

Gait evaluation; Observational gait assessment; Outcome measures; Physical therapy; Ramps; Rehabilitation; Step

INTRODUCTION

Observational gait analysis (OGA) serves an integral role in many rehabilitation fields and is a current clinical assessment standard for prosthetics. However, while OGA has been found to have approximately 60% to 85% reliability (1), it has been infrequently investigated on ramps and stairs (2). Further, few clinically useful outcome measures to objectively evaluate gait on ramps and stairs have been introduced. The hill assessment index (HAI) and stair assessment index (SAI) were developed for this reason (3,4). These assessments utilize an 11- and 13-point ordinal scale, respectively, to describe gross motor pattern implementation, degree of step length symmetry, and use of assistive devices. The inter-rater and intra-rater reliability of the HAI was found to be excellent in hill descent tasks for a sample of unilateral transfemoral amputation (TFA) patients using mechanical and C-Leg microprocessor knee (MPK) systems (2). However, no further evidence of psychometric properties for either outcome measure is available.

Video motion analysis has been used to determine criterion validity for other observational and functional outcome measures in the past and may allow for establishing evidence of the validity of the HAI and SAI (5–7). Therefore, the purpose of this study was to determine the criterion validity and inter-rater and intra-rater reliability of the HAI and SAI in a sample of persons with unilateral TFA using a prosthetic MPK system.

METHODS

All procedures were reviewed and approved by the University of South Florida's Institutional Review Board, and subjects gave informed consent prior to study participation.

Subjects were considered for study inclusion if they met the following criteria:

- Unilateral transfemoral or knee disarticulation amputation
- Use of a prosthesis with MPK for at least one year
- Independent level of community ambulation without use of an assistive device
- Ability to ascend and descend a ramp without human support
- Ability to ascend and descend stairs without human support

Subjects who met the inclusion criteria were enrolled in the study and evaluated by the study prosthetist to assure proper fit and function. The study prosthetist was state-licensed and

certified by the American Board for Certification in Orthotics, Prosthetics, and Pedorthics. Subjects' prosthetic sockets and suspension systems were not changed for the duration of the experiment to reduce confounding from fit and acclimation issues. All subjects were fit with the same model energy-storing prosthetic foot for use over the study duration. Manufacturer specifications were used to set componentry alignment and were verified using the LASAR alignment system (Ottobock Healthcare, Duderstadt, Germany). Each subject received training from the study physical therapist on the functions of the MPK and study foot for transitional movements, obstacle crossing, ramps, stairs, speed variation, and variable surfaces. Training techniques were used from prior publications (8,9). Subjects then accommodated with the new foot components until they reported and demonstrated confidence in walking unassisted on level ground, inclines, declines, and up and down stairs. Following this accommodation, subjects were scheduled for testing.

Subjects were asked to ascend and descend an Americans with Disabilities Act-compliant cement ramp, which was 0.9 m in width, 4.9 m in length, and sloped at an angle of 5°. Each subject completed three trials up and three trials down at a self-selected walking speed using handrails as they deemed necessary. These trials were timed using a stopwatch and recorded (sagittally) on video. The recordings were scored by two independent raters who reviewed the video recordings to determine the appropriate HAI value based on the criteria outlined in Figure 1 (3,4). The recordings were then assessed using Dartfish™ 2D (v7, Dartfish USA, Inc. Alpharetta, GA, USA) gait analysis software to determine prosthetic and sound side step length. Step length was measured relative to a 1.0 m reference object captured in the video background and measured using reflective markers placed on subjects' heels and toes. A previous investigation found inconsequential differences in marker movement tracked by 3D motion capture and Dartfish software on the order of 5 mm (10). Further, this program was previously found to be a valid (PCC = 0.95; p < 0.05) and reliable (ICC = 0.93; p < 0.05) motion analysis tool (11). Degree of asymmetry (DoA) was then calculated for step length using the following equation:

$$\text{Degree of Asymmetry} = \frac{\text{Sound Step Length} - \text{Prosthetic Step Length}}{\text{Sound Step Length} + \text{Prosthetic Step Length}}$$

Subjects then repeated the protocol on a flight of Americans with Disabilities Act-compliant stairs (12). The three trials ascending and three trials descending at a self-selected speed were timed using a handheld stopwatch and sagittally recorded on video. These recordings were again scored by two independent raters, this time using the criteria outlined in Figure 2 to determine the appropriate SAI score (3,4). Subjects repeated the study protocol on the ramp and stairs 90 d after initial testing to assess intra-rater reliability.

The HAI is scored on an ordinal scale from 0 to 11. In this sample, however, all subjects received a score of 10 or 11, resulting in a dichotomous data set. This allowed for step length, specifically DoA of step length (13), to become the reference standard for an assessment of sensitivity and specificity of the instrument at its higher scoring levels. The HAI indicates that a perfect score of 11 is an even step length without an assistive device determined visually by the rater. Asymmetrical movements in function and gait between 5%

and 10% have previously been described as typical (13,14). A perfect DoA score of 0, providing allowances for normal asymmetrical movement, can then be assumed to equate to a perfect HAI score. Therefore, absolute DoA values of ± 0.1 were determined to be representative of a symmetric step length and thus eligible for an HAI rating of 11/11. This was used to establish the reference standard and those participants with absolute DoA values $> \pm 0.1$ from the video motion analysis were determined to have an asymmetric step length and scored 10/11 on the HAI in this dichotomous set.

Sensitivity and specificity are statistical measures of the validity of an outcome measure (15). Sensitivity is the proportion of subjects who test positive for a condition to all those who truly have the condition. Specificity is the proportion of subjects who test negative for a condition to all those who truly do not have the condition. In this case, sensitivity and specificity of the HAI in identifying subjects who do and do not exhibit step length symmetry during ramp gait was assessed. These variables were calculated using a two-by-two configuration of HAI score and step length symmetry determined by the Dartfish software and the calculation of DoA in step length. Sensitivity and specificity were calculated as:

$$\begin{aligned} \text{Sensitivity} &= \left(n_{HAI\ score\ 10} / n_{|DoA| > 0.1} \right) * 100 \\ \text{Specificity} &= \left(n_{HAI\ score\ 11} / n_{|DoA| \leq 0.1} \right) * 100 \end{aligned}$$

In the above equation for sensitivity, $n_{HAI\ score\ 10}$ represents the number of subjects who scored 10 on the HAI because the rater found them to have an asymmetrical step length, and $n_{|DoA| > 0.1}$ represents the number of subjects who truly exhibited asymmetrical steps as assessed with video motion analysis and the DoA equation for step length symmetry. In the above equation for specificity, $n_{HAI\ score\ 11}$ represents the number of subjects who scored 11 on the HAI because the rater found them to have symmetrical step length, and $n_{|DoA| \leq 0.1}$ represents the number of subjects who truly did exhibit symmetrical steps as assessed with video motion analysis and the DoA equation for step length symmetry. Higher levels of sensitivity and specificity provide evidence to support the validity of the HAI to assess step length symmetry walking on ramps. To evaluate levels of sensitivity and specificity, we interpreted scores from 90% to 100% as excellent, 75% to 89% as good, 50% to 74% as adequate, and below 50% as poor sensitivity and specificity for the HAI uphill and downhill.

Positive and negative likelihood ratios can be calculated from the sensitivity and specificity values. A positive likelihood ratio indicates how many times more likely a positive test will be seen in those with more symmetric step lengths than in those with less symmetric step lengths (15). Interpretation of likelihood ratios vary (16,17). We interpreted a positive likelihood ratio of 10 or more as large, 5 to 9.94 as moderate, 2 to 4.94 as small, and 1 to 1.94 as very small positive likelihood. These positive likelihood ratios were used to determine if the HAI was a very useful, often useful, sometimes useful, or rarely useful test, respectively. A negative likelihood ratio, then, indicated how many times more likely a negative test will be seen in those with more step length symmetry compared to those with less step length symmetry (15). A negative likelihood ratio of less than 0.1 was interpreted as large, 0.11 to 0.2 as moderate, 0.21 to 0.5 as small, and 0.51 to 1.0 as very small negative

likelihood. These negative likelihood ratios were used to determine if the HAI was a very useful, often useful, sometimes useful, or rarely useful test, respectively.

Comparisons between rater one and rater two at initial testing and rater one at the repeat evaluation were made to assess inter- and intra-rater reliability of each instrument, respectively. Data were compared for both tests for ascending and descending conditions. Intraclass correlation coefficients (ICC) were determined using Spearman's Rho. Definitions for correlation strength also vary (15,18–20). We interpreted 0.9 to 1.0 as an excellent correlation, 0.75 to 0.89 a good correlation, 0.5 to 0.74 an adequate correlation, and below 0.5 a poor correlation. The level for statistical significance was determined *a priori* to be p 0.05.

RESULTS

Twenty ($n = 20$) individuals with TFA completed the study. The sample was 80% male with etiology of trauma (70%), tumor (20%), or vascular disease (10%). Mean (SD) age was 46.5 years of age (± 14.2). All subjects were unlimited community ambulators. Sensitivity of the HAI was calculated at 88.0% for uphill and 94.0% for downhill conditions. These results correspond to positive likelihood ratios of 3.5 and 2.8, respectively. Specificity for the instrument was determined to be 75.0% for uphill and 67.0% for downhill. These results correspond to negative likelihood ratios of 0.17 and 0.10, respectively. These results are summarized in Table 1.

For the HAI, Spearman Rho ICCs of $r_s = 0.80$ ($p < 0.001$) for uphill and 0.67 ($p = 0.001$) for downhill conditions between rater one and rater two were found. ICCs of $r_s = 0.87$ ($p < 0.001$) and 0.73 ($p < 0.001$) were calculated between initial and repeat testing scores for uphill and downhill conditions, respectively (Table 2). Mean (SD) uphill times (Table 3), in seconds, for trials one, two, and three were 4.71 (0.86), 4.71 (0.99), and 4.72 (1.03), respectively, resulting in a mean percent difference of 0.11% among trials. Mean and median scores were 10.8 and 11 for all trials. Mean downhill times, in seconds, for trials one, two, and three were 4.80 (1.14), 4.74 (1.19), and 4.78 (1.40), respectively, resulting in a mean percent difference of 1.1% among trials. Mean and median downhill scores were 10.9 and 11 for all trials.

For the SAI, ICCs of $r_s = 1.00$ ($p < 0.001$) and 0.89 ($p < 0.001$) were found between raters one and two for ascending and descending conditions, respectively. ICCs of $r_s = 1.00$ ($p < 0.001$) were calculated for both ascending and descending conditions between initial and repeat testing (Table 2). Mean (SD) ascent times (Table 3), in seconds, for trials one, two, and three were 4.95 (1.01), 5.15 (1.62), and 4.68 (1.33), respectively, resulting in a mean percent difference of 6.7% among trials. Corresponding mean trial scores were 9.7, 9.9, and 10.1, with a constant median of 11. Mean descent times, in seconds, for trials one, two, and three were 3.82 (0.97), 3.72 (0.80), and 3.61 (0.80), respectively, resulting in a mean percent difference of 2.76%. Corresponding mean descending scores were 12.0, 11.7, and 11.8, with a constant median of 11. Times and scores for the HAI and SAI are shown in Table 3.

DISCUSSION

The results of this study provide evidence of the validity and reliability of the HAI and SAI in a sample of persons with unilateral TFA using an MPK system and ambulating at an unlimited-community level.

The sensitivity of the HAI was determined to be good for uphill and excellent for downhill evaluations in this sample of community ambulating transfemoral amputees. The good sensitivity for uphill and excellent sensitivity for downhill evaluations resulted because the raters on the HAI correctly identified those who walked with step length DoA above the 0.1 reference standard. Especially on hill descent, the excellent sensitivity shows that the HAI will have a low rate of false negatives, meaning that if a subject is identified by the rater as having asymmetrical step length, it is highly likely that they actually do have asymmetrical step length. Both positive likelihood ratios indicate the HAI is a “sometimes useful” test for uphill and downhill assessments of hill gait in higher functioning TFA subjects.

The specificity of the HAI was determined to be good for uphill and adequate for downhill assessment. The lower specificity on ramp descent was a result of some subjects being scored by the raters as having asymmetrical step length although they actually had symmetrical step lengths based on having step length DoA below 0.1. This could mean that identifying step length symmetry during ramp descent may be more difficult and there is the potential for an increased rate of false positives (i.e., giving a lower HAI score for step symmetry). The negative likelihood ratios confirm that the HAI is a “sometimes useful” test for both uphill and downhill assessments of hill gait. In summary, the HAI was found to be an adequate to excellent assessment tool of step length symmetry during uphill and downhill gait of subjects in this sample. The assessment was found to be more sensitive for downhill assessment and more specific for uphill in high-functioning persons with unilateral TFA. Due to the lower specificity in ramp descent, low scores may be vulnerable to false results in this population. However, due to the good and excellent sensitivity of the HAI, lower scores should truly identify those with less symmetrical step length during ramp gait.

Assessment of intra-rater reliability for the HAI was found to be good for uphill but adequate for downhill gait in this sample based on the strength of correlation (Spearman’s Rho). These results were identical to the inter-rater reliability in this sample, with uphill being found as good and downhill being found as adequate. The downhill inter-rater reliability assessment was adequate. These results differ slightly from the results of previous work, which found excellent inter-rater reliability in a more heterogeneous sample of TFA subjects utilizing C-Leg MPKs (2). The modest differences could be multi-factorial and potentially include sample heterogeneity, rater experience and background, a lack of scoring instructions for the instrument, slight performance interpretation differences, and other factors.

The consistency of HAI times, with a percent change in time of 0.1% and 1.1% for ascent and descent, respectively, and the reliability of uphill HAI scores among trials provide some evidence of the instrument’s stability on repeat testing within a single visit. Based on the negligible differences in time to ambulate on the ramp and HAI scores, no learning or

fatigue effects were noted in this sample. This internal consistency commonly suggests no benefit between a single or repeated testing protocol. Thus, in the interest of valuable clinical time, it seems feasible that a single administration of the test may be sufficient to determine the patient's ability to ambulate ramps. Repeated testing of high-functioning TFA patients within a single clinic visit seems unlikely to yield different scores provided conditions (e.g., components, physical status) are not changed.

The SAI was found to be a very reliable assessment tool for high-functioning transfemoral prosthetic users. Inter-rater reliability was excellent for ascending and good for downhill conditions based on the strength of correlation. Excellent intra-rater reliability was found for both ascending and descending assessments. The percent differences among trial times was higher than the HAI, but the mean differences remain adequately low to preliminarily declare the instrument as stable in comparable samples. No clear trend in time or scoring for the SAI was found among trials, further suggesting no learning or fatigue effect. The results of this work also show no benefit or detriment to performing the test once or multiple times. For clinical simplicity, the test may only require a single administration to determine a high-functioning TFA subject's true stair ascent or descent capability. As with the HAI, this is provided that the subject has been trained and is confident in using the specific functional features (e.g., stair ascent mode) of the MPK. The SAI may provide more obvious differentiation between scores relative to the HAI, as there is less opportunity for subjectivity in the construct. The HAI requires the rater to determine step length ratios, which may have large variability between subjects, whereas the SAI requires identification of more discretely discernable stepping patterns. Future research should attempt to establish validity for the SAI, as this study demonstrated strong evidence of the inter-rater and intra-rater reliability for the SAI as a measure of stair ascent and descent. Replication research is needed to confirm the recommendation for a single test in comparable samples but also in TFA patients using mechanical knees or other types of knees and in those who ambulate at lower functional levels.

Limitations

This study utilized a group of high-functioning subjects with unilateral TFA who ambulated with MPKs. Therefore, results may not be applicable to individuals of differing amputation level(s), those with bilateral or upper extremity involvement, amputees of lower functional levels, or individuals of other diagnostic groups. Further, the dichotomous results of the HAI allowed for sensitivity and specificity calculation but prevented evaluation of the entire ordinal scale (i.e., at the lower end of the functional spectrum). A lack of low scores was also observed in SAI scores, resulting in left skewed data. Also, the reference standard adopted for this study, DoA in step length, while a more objective measure of step length symmetry than observational gait analysis, has not been thoroughly evaluated. The DoA cut-off score of ± 0.1 was chosen based on the assumption that 5% to 10% asymmetry in movement is typical in gait and other functional movement patterns (13,14). For this reason, it is difficult to make definitive conclusions of the reliability and validity of the entirety of the HAI and SAI. Further research is needed to confirm the results of this work in other amputation populations, including lower-functioning patients who would score lower in the range of the HAI and SAI measures.

CONCLUSION

The results of this work established validity and reliability of the upper end of the HAI and SAI in a population of high-functioning subjects with unilateral TFA using microprocessor knee systems. The HAI showed moderate sensitivity and specificity. Intra-rater reliability of the HAI was good for uphill and adequate for downhill assessment. Inter-rater reliability was found to be good for uphill and adequate for downhill assessment. Since other ramp gait assessment tools are scarce, the HAI was shown to be a viable assessment tool. The SAI showed excellent and good inter-rater reliability for ascending and descending conditions, respectively, and excellent intra-rater reliability for both. Both HAI and SAI were shown to be stable instruments for both ascending and descending assessments, resulting in a recom-200 mendation for one trial as a true assessment of hill and stair gait. More research is needed to determine validity of the SAI and to confirm all of these results in other amputation populations.

Acknowledgments

The authors declare no conflicts of interest and are solely responsible for the content of this manuscript. Contents of this manuscript represent the opinions of the authors and not necessarily those of the U.S. Department of Defense, U.S. Department of the Army, U.S. Department of Veterans Affairs, or any academic or health care institution.

This project was funded by:

1. The Florida High Tech Corridor/USF Connect. (Grant # FHT 10-26)
2. National Institutes of Health Scholars in Patient Oriented Research (SPOR) grant (1K30RR22270)

References

1. Krebs DE, Edelstein JE, Fishman S. Reliability of observational kinematic gait analysis. *Phys Ther.* 1985; 65:1027–33. [PubMed: 3892553]
2. Highsmith MJ, Kahle JT, Miro RM, Mengelkoch LJ. Ramp descent performance with the C-Leg and interrater reliability of the Hill Assessment Index. *Prosthet Orthot Int.* 2013; 37(5):362–8. [PubMed: 23327837]
3. Buell, NC.; Willingham, LL.; Allyn, KJ.; Hafner, BJ.; Smith, DG. Evaluation of gait style to ascend and descend stairs for lower limb amputees. In: Boone, D., editor. *Proceedings of the 11th World Congress of the International Society of Prosthetics and Orthotics*; 2004 Aug 1–6; Hong Kong. Hong Kong: Hong Kong National Society of the International Society for Prosthetics and Orthotics; 2004. p. 367
4. Hafner BJ, Willingham LL, Buell NC, Allyn KJ, Smith DG. Evaluation of function, performance, and preference as transfemoral amputees transition from mechanical to microprocessor control of the prosthetic knee. *Arch Phys Med Rehabil.* 2007; 88:207–17. [PubMed: 17270519]
5. Houdijk H, Appelman FM, Van Velzen JM, Van der Woude LH, Van Bennekom CA. Validity of DynaPort GaitMonitor for assessment of spatiotemporal parameters in amputee gait. *J Rehabil Res Dev.* 2008; 45:1335–42. [PubMed: 19319757]
6. McDonough AL, Batavia M, Chen FC, Kwon S, Ziai J. The validity and reliability of the GAITRite system's measurements: a preliminary evaluation. *Arch Phys Med Rehabil.* 2001; 2001:419–25.
7. van Loo MA, Moseley AM, Bosman JM, de Bie RA, Hassett L. Inter-rater reliability and concurrent validity of step length and step width measurement after traumatic brain injury. *Disabil Rehabil.* 2003; 25:1195–200. [PubMed: 14578058]
8. Highsmith MJ, Kahle JT, Lewandowski AL, Kim SH, Mengelkoch LJ. A method for training step-over-step stair descent gait with stance yielding prosthetic knees. *J Prosthet Orthot.* 2012; 24:10–5.

9. Highsmith MJ, Kahle JT, Lura DJ, Lewandowski AL, Quillen WS, Kim SH. Stair ascent and ramp gait training with the Genium knee. *Technol Innov.* 2014; 15(4):349–58.
10. Eltoukhy M, Asfour S, Thompson C, Latta L. Evaluation of the performance of digital video analysis of human motion: Dartfish tracking system. *Int J Sci Eng Res.* 2012; 3:1–6.
11. Norris BS, Olson SL. Concurrent validity and reliability of two-dimensional video analysis of hip and knee joint motion during mechanical lifting. *Physiother Theory Pract.* 2011; 27:521–30. [PubMed: 21568816]
12. US Department of Justice. ADA standards for accessible design. Washington (DC): US Department of Justice; 2010.
13. Highsmith MJ, Schulz BW, Hart-Hughes S, Latlief GA, Phillips SL. Differences in the spatiotemporal parameters of transtibial and transfemoral amputee gait. *J Prosthet Orthot.* 2010; 22:26–30.
14. Highsmith MJ, Kahle JT, Carey SL, Lura DJ, Dubey RV, Csavina KR, Quillen WS. Kinetic asymmetry in transfemoral amputees while performing sit to stand and stand to sit movements. *Gait Posture.* 2011; 34:86–91. [PubMed: 21524913]
15. Domholdt, E. Physical therapy research: principles and applications. 2. Philadelphia (PA): W.B. Saunders Co; 2000.
16. McGee S. Simplifying likelihood ratios. *J Gen Intern Med.* 2002; 17:647–50.
17. Anvari A, Halpern EF, Samir AE. Statistics 101 for radiologists. *Radiographics.* 2015; 35:1789–801. [PubMed: 26466186]
18. Highsmith MJ, Lura DJ, Carey SL, Mengelkoch LJ, Kim SH, Quillen WS, Kahle JT, Miro RM. Correlations between residual limb length and joint moments during sitting and standing movements in transfemoral amputees. *Prosthet Orthot Int.* 40(4):522–7. 20165. [PubMed: 25628379]
19. Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. *Radiology.* 2003; 227:617–22. [PubMed: 12773666]
20. Kubiszyn, T.; Borich, G. Educational testing and measurement: classroom application and practice. 7. Hoboken (NJ): John Wiley & Sons, Inc; 2003.

Score	Mobility Descriptor
0	Cannot do/ refuses to do
1	Needs assist
2	Side step, with assistive device
3	Side step, without assistive device
4	Step-to with assistive device
5	Step a little past ($\frac{1}{2}$ foot length), with assistive device
6	Step past (more than $\frac{1}{2}$ foot length), with assistive device
7	Even step with assistive device
8	Step-to, without assistive device
9	Step a little past ($\frac{1}{2}$ foot length), without assistive device
10	Step past (more than $\frac{1}{2}$ foot length), without assistive device
11	Even step, without assistive device

Adapted from Buell et al.

Figure 1.
Hill Assessment Index (HAI).

Score	Mobility Descriptor
0	Cannot do/ refuses to do
1	Needs assist
2	With rail and assistive device, step-to pattern
3	With rail, step-to pattern
4	With assistive device, step-to pattern
5	Without rail or assistive device, step-to pattern
6	With rail and assistive device, skipping step pattern
7	With rail, skipping step pattern
8	With assistive device, skipping step pattern
9	Without rail or assistive device, skipping step pattern
10	With rail and assistive device, step-over-step pattern
11	With rail, step-over-step pattern
12	With assistive device, step-over-step pattern
13	With rail and assistive device, step-over-step pattern

Adapted from Buell et al.

Figure 2.
Stair Assessment Index (SAI).

Table 1

Sensitivity, Specificity, and Likelihood Ratios

Condition	Uphill	Downhill
Sensitivity	88.0%	94.0%
Specificity	75.0%	67.0%
+ Likelihood Ratio'	3.50	2.80
- Likelihood Ratio'	0.17	0.10

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

HAI/SAI Inter/Intra-Rater Correlations

Hill Assessment Index	Inter-rater		Intra-rater	
Variable	Uphill	Downhill	Uphill	Downhill
Spearman's Rho (r_s)	0.80	0.67	0.87	0.73
ICC significance	$p < 0.001$	$p = 0.001$	$p < 0.001$	$p < 0.001$
Stair Assessment Index	Inter-rater		Intra-rater	
Variable	Ascent	Descent	Ascent	Descent
Spearman's Rho (r_s)	1.00	0.89	1.00	1.00
ICC significance	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Mean/Median Stair Performance

Instrument Condition	HAI				SAI			
	Uphill		Downhill		Ascent		Descent	
Time (sec)	mean	SD	mean	SD	mean	SD	mean	SD
Trial 1	4.71	0.86	4.80	1.14	4.95	1.01	3.82	0.97
Trial 2	4.71	0.99	4.74	1.19	5.15	1.62	3.72	0.80
Trial 3	4.72	1.03	4.78	1.40	4.68	1.33	3.61	0.80
% diff	0.11%	--	1.07%	--	6.72%	--	2.76%	--
Score	mean	median	mean	median	mean	median	mean	median
Trial 1	10.8	11.0	10.9	11.0	9.7	11.0	12.0	11.0
Trial 2	10.8	11.0	10.9	11.0	9.9	11.0	11.7	11.0
Trial 3	10.8	11.0	10.9	11.0	10.1	11.0	11.8	11.0