Published in final edited form as:

Arch Phys Med Rehabil. 2008 December ; 89(12): 2293–2296. doi:10.1016/j.apmr.2008.06.010.

The Reliability and Validity of Measures of Gait Variability in Community-Dwelling Older Adults

Jennifer S. Brach, PhD, PT, Subashan Perera, PhD, Stephanie Studenski, MD, MPH, and Anne B. Newman, MD, MPH

Department of Physical Therapy (Brach), Division of Geriatric Medicine, Department of Medicine (Perera, Studenski, Newman), and Department of Epidemiology (Newman), University of Pittsburgh, Pittsburgh, PA

Abstract

Objective—To examine the test-retest reliability and concurrent validity of variability of gait characteristics.

Design—Cross-sectional study.

Setting—Research laboratory.

Participants—Older adults (N=558) from the Cardiovascular Health Study.

Interventions—Not applicable.

Main Outcome Measures—Gait characteristics were measured using a 4-m computerized walkway. SD determined from the steps recorded were used as the measures of variability. Intraclass correlation coefficients (ICC) were calculated to examine test-retest reliability of a 4-m walk and two 4-m walks. To establish concurrent validity, the measures of gait variability were compared across levels of health, functional status, and physical activity using independent *t* tests and analysis of variances.

Results—Gait variability measures from the two 4-m walks demonstrated greater test-retest reliability than those from the single 4-m walk (ICC=.22–.48 and ICC=.40–.63, respectively). Greater step length and stance time variability were associated with poorer health, functional status and physical activity ($P<.05$).

Conclusions—Gait variability calculated from a limited number of steps has fair to good test-retest reliability and concurrent validity. Reliability of gait variability calculated from a greater number of steps should be assessed to determine if the consistency can be improved.

Correspondence to Jennifer S. Brach, PhD, PT, Department of Physical Therapy, University of Pittsburgh, 6035 Forbes Tower, Pittsburgh, PA 15260, e-mail: E-mail: jbrach@pitt.edu. Reprints are not available from the author.

Presented at the International Society for Posture and Gait Research, Burlington, VT, July 14–18, 2007.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Publisher's Disclaimer: This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues. Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit: <http://www.elsevier.com/copyright>

Keywords

Gait; Rehabilitation; Reliability and validity

Gait variability is emerging as an important indicator of impaired mobility in community-dwelling older adults, among whom it has been shown to predict future falls and incident mobility disability.¹⁻⁴ An important quality of any measure is test-retest reliability, the consistency of repeated measurements for a subject. Measurements that do not have good test-retest reliability will have high measurement error and are likely to be imprecise, making it difficult to assess associations or change over time.⁵ Although the test-retest reliability of gait speed and mean gait characteristics have been extensively assessed for reliability,⁶⁻¹⁰ the test-retest reliability of gait variability measures is unknown.

Naturally, the number of data points or steps used in the calculation of gait variability are important in the consistency of the measure, with longer walks giving more stable variability estimates.^{11,12} However, the optimal number of consecutive steps or data points needed to calculate a consistent measure of gait variability has yet to be determined, and such a recommendation must consider tradeoffs between maximizing consistency and the potential constraints of subject fatigue and space limitations.

Accuracy or validity of a measure is also essential for research. A measure is valid to the extent that it measures what it purports to measure. The predictive validity of gait variability for future falls^{1,2} and incident mobility disability⁴ has been reported. However, reports of concurrent validity have been limited by modest sample size and focus on only temporal aspects of gait variability.¹³

The purpose of this study was to examine reliability and validity of measures of temporal and spatial gait variability in a large sample of community-dwelling older adults. Specifically, we report test-retest reliability of measures of gait variability, and estimate the impact of the length of the walk on the consistency of the measure. The concurrent validity of temporal and spatial measures of gait variability is also examined by comparison to measures of health, functional status, and physical activity.

Methods

Study Sample

Gait characteristics were assessed in ambulatory older adults from the Pittsburgh site of the CHS at the tenth follow-up visit (between 1998 through 1999). CHS is a population-based, ongoing longitudinal multicenter study of coronary heart disease and stroke risk in community-dwelling older adults age 65 years and older.^{14,15} At the initiation of the CHS in 1989 through 1990, individuals were identified from the Health Care Financing Administration sampling frame. Individuals who were 65 years or older, noninstitutionalized, expected to remain in the area for 3 years, and able to give informed consent were included in the study. Individuals who were wheelchair-bound in the home or were receiving hospice care, radiation therapy, or chemotherapy for cancer were excluded.^{14,15} In 1989 to 1990 an original cohort of 5201 predominately (>95%) white men and women were enrolled, and in 1992 through 1993 a cohort of 687 black men and women was added.

Participants in the current study included men and women who attended the tenth clinic visit in 1998 through 1999 at the Pittsburgh site, who could walk without the assistance of another person, and who could follow directions to complete the gait assessment (N=558). The

University of Pittsburgh Institutional Review Board approved this study, and all study participants provided written informed consent prior to data collection.

Gait Characteristics

The GaitMat II system^a was used for the gait analysis.¹⁶ The GaitMat II consists of a 4-m long walkway on which the subject walks and a computer system that controls the GaitMat II and analyzes the data. In addition to the 4-m long walkway, there are initial and final 1-m inactive sections to allow for acceleration and deceleration of the participant. The GaitMat II is an automated gait analysis system based on the opening and closing of pressure-sensitive switches, which are represented on the computer screen as footprints when the participant walks on the walkway. Participants completed 2 practice passes on the GaitMat II followed by 4 passes at their self-selected walking speed for data collection.

We were primarily interested in variability of step length, step width, and stance time. Step length and width represent spatial characteristics in 2 different planes. Stance time was selected as the temporal gait characteristic. Step length, step width, and stance time were also specifically selected because they have been studied by other investigators.^{1,2,17,18} Gait speed was determined by dividing the distance traversed by the time between the first and last step (eg, switch closure) and was recorded in meters/second. Step length was defined as the distance between 2 consecutive footprints, measured from the heel of 1 footprint to the heel of the next footprint and was recorded in meters. Step width was defined as the distance between the outermost borders of 2 consecutive footprints and was recorded in meters. Stance time was defined as the length of time that 1 foot was in contact with the floor (ie, from initial foot-floor contact until final foot-floor contact) and was recorded in seconds. The SDs of step length, step width, and stance time determined from all of the steps recorded over the passes of interest (1 pass for the 4-m analyses and 2 passes for the two, 4m analyses) were used as the measures of variability.

Other Measures

Measures used for the concurrent validity analyses included a measure of general health perception, self-reported difficulty with ADLs and IADLs, difficulty walking a half mile, and physical activity level.¹⁹ Perception of general health was self-reported as excellent, very good, good, fair, or poor. The ADLs assessed were bathing, dressing, eating, using the toilet, walking around the home, or getting out of a bed or chair. The IADLs were heavy housework, light housework, shopping for personal items, preparing own meals, paying bills or managing money, or using the telephone. Self-reported difficulty walking a half mile was also recorded. Physical activity was measured as the self-reported number of blocks walked in the past week and the participants were classified as physically active (≥ 7 blocks) or physically inactive (< 7 blocks).

Statistical Analysis

Step length, stance time, and step width variability calculated from the four individual passes and an ICC was calculated to determine test-retest reliability of a 4-m walk. Step length, stance time, and step width variability calculated from the first 2 passes and the last 2 passes were used to obtain an ICC to determine the test-retest reliability of two 4-m walks. ICCs were interpreted as follows: less than 0.4, poor; 0.4 to 0.75, fair to good; and more than 0.75, excellent.²⁰ Similar assessment of test-retest reliability was made for mean gait characteristics and gait speed to examine the consistency of the methodology for measuring gait characteristics in our sample compared with the methodology of other studies.

To establish the concurrent validity, the measures of step length variability, stance time variability, and step width variability (dependent variables) were compared across individuals

with and without ADLs and IADLs difficulty, with and without difficulty walking a half mile, who were physically active (reported ≥ 7 blocks walked previous week) and inactive (reported < 7 blocks walked previous week), who could and could not tandem stand for 10 seconds, and with different levels of health status (poor, fair, good, very good, and excellent) using independent samples *t* tests for comparison of 2 means (ADLs difficulty, IADLs difficulty, difficulty walking one-half mile, tandem stand, and physical activity) and 1-way analysis of variance for comparison of more than 2 means (health status).

Results

Table 1 provides the participant characteristics. The mean age of the sample was 79.4 years. Approximately 18% of the sample reported having fallen in the previous year and only a small percentage of the participants ($< 10\%$) used a cane for ambulation. On average the participants demonstrated a mean gait speed that was slightly less than the desired gait speed of 1.2m.

The gait variability measures from the two 4-meter walks demonstrated greater test-retest reliability than those from the single 4-m walk (table 2). When calculated from the single 4-m walk, step width and stance time variability demonstrated poor reliability (ICC=0.22 and 0.37, respectively) and step length variability demonstrated fair (ICC=0.48) test-retest reliability. When calculated from the longer two, 4-m walks, step width variability continued to demonstrate marginal (ICC=0.40) reliability, whereas step length and stance time variability had fair to good (ICC=0.50 and 0.63, respectively) test-retest reliability. In general, step length and stance time variability were more reliable measures than the measure of step width variability.

Gait speed and the mean gait characteristics demonstrated excellent test-retest reliability (ICC ≥ 0.80). Gait speed, mean step length, and mean stance time were the most reliable (ICC > 0.92), whereas mean step width was somewhat less reliable (ICC=0.80–0.89) but still excellent.

The concurrent validity of the measures of gait variability is presented in table 3. The largest effects across all measures were found for stance time variability (see table 3). Step length variability significantly varied across levels of IADLs difficulty, difficulty walking a half mile, tandem stand, and physical activity level (see table 3). Step width variability significantly varied across levels of IADLs difficulty and difficulty walking a half mile. Greater stance time and step length variability were associated with poorer levels of health, functional status, and physical activity. Lesser step width variability was associated with individuals reporting difficulty with IADLs or walking a half mile.

Discussion

Gait variability calculated from a limited number of steps measured using a computerized walkway has poor to good test-retest reliability. Step length and stance time variability seem to be more consistent measures than step width variability. This is consistent with a model proposed by Gabell and Nayak.²⁵ They hypothesize that the variability in step width, an indicator of balance control, is a reflection of the adaptive power of the balance system, necessary for maintaining balance. This model is also supported by the work of Donelan et al,²⁶ who reported that body lateral motion is partially stabilized by medio-lateral foot placement or step width. However, this should not necessarily be an issue if step width variability is measured in an environment that does not challenge balance, such as walking on a GaitMat at usual walking speed. Stance time and step length are thought to be regulated by an automatic stepping mechanism, thus potentially explaining the consistency of the measures.²⁵

The number of steps used in the calculation of the measures of gait variability plays an important role in the consistency of the measure. This finding is similar to those for other gait analysis techniques, which report improved reliability with a greater number of testing trials.¹² In our study, the measures of gait variability calculated from the steps recorded during two 4-m walks (≈ 10 – 12 steps) were more reliable than those from the single 4-m walk (≈ 5 – 6 steps). Measures of dispersion (such as the SD), when applied to most bell shaped distributions, have an inherent tendency toward more frequent underestimation, unlike the measures of central tendency (such as the mean). This phenomenon is characterized by the well-known right skewness of the sampling distributions of the measures of dispersion. Dispersion measures seek to measure the width of a distribution. The first few observations (ie, steps) most likely represent the center of the distribution and therefore underestimate its width. As the sample size (ie, number of steps) increases, the likelihood increases that a few extreme observations (ie, steps) will represent the tails of the distribution, producing a more accurate measure of the actual width of the distribution. Therefore, any variability estimate will first increase with sample size and then gradually stabilize in most subjects. Further study is needed to determine whether an even longer walk length would yield greater consistency among measures of gait variability. Also to be determined is the point at which constraints of fatigue and space begin to produce missing data, increasingly unreliable data, or other problems.

The excellent test-retest reliability of gait speed and mean gait characteristics in our sample is consistent with the literature,⁶⁻¹⁰ thus providing some evidence that our methodology (computerized walkway) for measuring gait characteristics is sound. The excellent reliability of the measures of gait speed and mean gait characteristics in the sample also indicates that the subjects do not vary their walking dramatically from trial to trial. Therefore, the modest reliability of gait variability measures cannot be attributed to true changes in gait speed or mean characteristics between trials.

All measures of gait variability demonstrated concurrent validity against measures of general health, functional status, and physical activity level. Greater amounts of step length and stance time variability and lesser amounts of step width variability are associated with poorer health status, impaired functional status, and physical inactivity. This is consistent with other reports that step length and stance time variability are increased and step width is decreased in individuals who fall compared with those who do not fall.^{1,2,17}

A small number of participants who used a cane during ambulation (7.6%) were included in this study. We chose not to exclude them in order to increase the generalizability of the results. Individuals who use a cane likely have poorer gait, health, functional status, and physical activity than individuals who do not use a cane during ambulation, which may have influenced the concurrent validity findings. However, by including these individuals in the analyses we are able to describe the association between health, function, physical activity, and gait variability across a greater range of abilities.

This study has important strengths. The study included a relatively large sample of diverse community-dwelling older adults. The participants completed several passes on a computerized walkway thus we were able to examine both spatial and temporal gait characteristics. In addition, because the study participants were part of the CHS, several measures of health, physical function, and physical activity were available to characterize the individuals for the concurrent validity analyses. This study also has an important limitation of the length of the computerized walkway to record the gait characteristics used to calculate the measure of gait variability. To achieve the longer walk the gait characteristics had to be combined from two 4-m walks.

Because gait variability measures recorded from limited walk lengths seem to have good concurrent validity along with modest reliability, they must provide more signal than noise. In other words, the signal-to-noise ratio for measures of gait variability is certainly detectable, but is still imperfect. If the noise in the measures of gait variability can be reduced, reliability could improve, as would the signal-to-noise ratio, and ultimately, the ability of gait variability measures to be used to study mobility problems.

Acknowledgments

Supported by the National Institutes of Health Public Health Service (grant no. TG32 AG00181) and the National Institutes of Health contracts (grant no. N01-HC-75150, N01-HC-45133, N01-HC-85079 through 85085, and HL 87079 through 87086); and the University of Pittsburgh Older American's Independence Center (grant no. P30 AG024827); the National Institutes of Health Public Health Service (grant no. TG32 AG00181); a National Institutes on Aging and American Federation of Aging Research Paul Beeson Career Development Award (grant no. K23 AG026766). The sponsor had no direct role in the design, methods, subject recruitment, data collection, analysis, or preparation of the manuscript.

References

1. Maki BE. Gait changes in older adults: predictors of falls or indicators of fear? *J Am Geriatr Soc* 1997;45:313–20. [PubMed: 9063277]
2. Hausdorff JM, Rios DA, Edelberg HK. Gait variability and fall risk in community-living older adults: a 1-year prospective study. *Arch Phys Med Rehabil* 2001;82:1050–6. [PubMed: 11494184]
3. Brach JS, Berlin JE, VanSwearingen JM, Newman AB, Studenski SA. Too much or too little step width variability is associated with a fall history in older persons who walk at or near normal gait speed. *J Neuroeng Rehabil* 2005;2:21. [PubMed: 16042812]
4. Brach JS, Studenski S, Perera S, VanSwearingen JM, Newman AB. Gait variability and the risk of incident mobility disability. *J Gerontol Med Sci* 2007;62A:983–8.
5. Rothstein, JM. Measurement and Clinical Practice: Theory and application. In: Rothstein, JM., editor. *Measurement in physical therapy*. New York: Churchill Livingstone; 1985. p. 1-46.
6. Jette AM, Jette DU, Ng J, Plotkin DJ, Bach MA, The Musculoskeletal Impairment (MSI) Study Group. Are performance-based measures sufficiently reliable for use in multicenter trials? *J Gerontol* 1999;54A:M3.
7. Bohannon RW. Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants. *Age Aging* 1997;26:15–9.
8. Tager IB, Swanson A, Satariano WA. Reliability of physical performance and self-reported functional measures in an older population. *J Gerontol A Biol Sci Med Sci* 1998;53A:M295–300. [PubMed: 18314569]
9. 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;82:419–25. [PubMed: 11245768]
10. van Uden CJ, Besser MP. Test-retest reliability of temporal and spatial gait characteristics measured with an instrumented walkway system (GAITRite). *BMC Musculoskelet Disord* 2004;17:5–13.
11. Owings TM, Grabiner MD. Measuring step kinematic variability on an instrumented treadmill: how many steps are enough? *J Biomechanics* 2003;36:1215–8.
12. Monaghan K, Delahunt E, Caulfield B. Increasing the number of gait trial recordings maximises intra-rater reliability of the CODA motion analysis system. *Gait Posture* 2007;25:303–15. [PubMed: 16730177]
13. Hausdorff JM, Nelson ME, Kaliton D, et al. Etiology and modification of gait instability in older adults: a randomized controlled trial of exercise. *J Appl Physiol* 2001;90:2117–29. [PubMed: 11356774]
14. Tell GS, Fried LP, Hermanson B, Manolio TA, Newman AB, Borhani NO. Recruitment of adults 65 years and older as participants in the cardiovascular health study. *Ann Epidemiol* 1993;3:358–66. [PubMed: 8275211]

15. Fried LP, Borhani NO, Enright PL, et al. The cardiovascular health study: design and rationale. *Ann Epidemiol* 1991;1:263–76. [PubMed: 1669507]
16. Walsh, JP. Foot fall measurement Technology. In: Craik, RL.; Oatis, CA., editors. *Gait Analysis: theory and application*. St. Louis: Mosby-Year Book, Inc; 1995. p. 125-42.
17. Hausdorff JM, Edelberg HK, Mitchell SL, Goldberger AL, Wei JY. Increased gait unsteadiness in community-dwelling elderly fallers. *Arch Phys Med Rehabil* 1997;78:278–83. [PubMed: 9084350]
18. Owings TM, Grabiner MD. Step width variability, but not step length variability or step time variability, discriminates gait of healthy young and older adults during treadmill locomotion. *J Biomech* 2004;37:935–8. [PubMed: 15111081]
19. Fitti JE, Kovar MG. National Health Interview Survey. *Vital Health Stat* 1987;21:1–115. Publication DHHS (PHS) 87-1323
20. Rosner, B. *Fundamentals of Biostatistics*. Vol. 4th. Belmont: Duxbury Pr; 1995.
21. Blanke DJ, Hageman P. Comparison of gait of young and elderly men. *Phys Ther* 1989;69:144–8. [PubMed: 2913584]
22. Hageman PA, Blanke DJ. Comparison of gait of young women and elderly women. *Phys Ther* 1986;66:1382–7. [PubMed: 3749270]
23. Imms FJ, Edholm OG. Studies of gait and mobility in the elderly. *Age Ageing* 1981;10:147–56. [PubMed: 7270322]
24. Lerner-Frankiel MB, Vargas S, Brown M, Krusel L, Schoneberger W. Functional community ambulation: what are your criteria? *Clinical Management* 1986;6:12–5.
25. Gabell A, Nayak US. The effect of age and variability in gait. *J Gerontol* 1984;39:662–6. [PubMed: 6491179]
26. Donelan JM, Shipman DW, Kram R, Kuo AD. Mechanical and metabolic requirements for active lateral stabilization in human walking. *J Biomech* 2004;37:827–35. [PubMed: 15111070]

List of Abbreviations

ADLs	activities of daily living
CHS	Cardiovascular Health Study
IADLs	instrumental activities of daily living
ICC	intraclass correlation coefficient

Table 1**Participant Characteristics (N=558)**

Characteristic	
Demographics and health status	
Age (y)	79.4±4.1
	n (%)
Women	339 (60.5)
Black	127 (22.7)
Use assistive device	42 (7.6)
Fallen past year	101 (18.3)
Gait characteristics	
Gait speed (m/s)	1.00±0.23
Stance time (s)	0.40±0.23
Step length (m)	0.36±0.17
Step width (m)	0.35±0.15

Note. Values are mean ± SD unless otherwise noted.

Table 2
ICCs for Test-Retest Reliability of Gait Variability, Mean Gait Characteristics, and Gait Speed

Gait Characteristic	4-m Walk N=558	Two 4-m Walks N=558
Gait variability		
Step length SD	0.48	0.63
Step width SD	0.22	0.40
Stance time SD	0.37	0.50
Gait speed and mean gait characteristics		
Gait speed	0.97	0.98
Step length mean	0.97	0.99
Step width mean	0.80	0.89
Stance time mean	0.96	0.98

Table 3
Concurrent Validity of Measures of Gait Variability for Health, Functional Status, and Physical Activity

	Step Length SD Mean $\pm$ SD (m)	Step Width SD Mean $\pm$ SD (m)	Stance Time SD Mean $\pm$ SD (s)
Health perception			
Excellent (n=26)	.037 $\pm$.16	.041 $\pm$.18	.036 $\pm$.15
Very good (n=133)	.035 $\pm$.19	.038 $\pm$.20	.037 $\pm$.22
Good (n=273)	.035 $\pm$.13	.036 $\pm$.14	.040 $\pm$.22
Fair (n=116)	.038 $\pm$.14	.034 $\pm$.11	.047 $\pm$.29
Poor (n=7)	.042 $\pm$.18	.031 $\pm$.13	.078 $\pm$.36
	<i>P</i> =.21	<i>P</i> =.07	<i>P</i> <.0001
ADLs difficulty			
No (n=469)	.036 $\pm$.15	.037 $\pm$.15	.039 $\pm$.22
Yes (n=87)	.037 $\pm$.15	.033 $\pm$.14	.051 $\pm$.31
	<i>P</i> =.30	<i>P</i> =.06	<i>P</i> <.001
IADLs difficulty			
No (n=407)	.034 $\pm$.13	.037 $\pm$.16	.036 $\pm$.17
Yes (n=149)	.040 $\pm$.19	.033 $\pm$.13	.053 $\pm$.35
	<i>P</i> =.002	<i>P</i> =.009	<i>P</i> <.0001
Difficulty walking one-half mile			
No (n=421)	.035 $\pm$.15	.037 $\pm$.16	.037 $\pm$.21
Yes (n=133)	.040 $\pm$.15	.033 $\pm$.12	.053 $\pm$.29
	<i>P</i> =.0005	<i>P</i> =.003	<i>P</i> <.0001
Tandem stand 10 seconds			
No (n=163)	.039 $\pm$.16	.037 $\pm$.17	.046 $\pm$.23
Yes (n=372)	.034 $\pm$.14	.036 $\pm$.14	.036 $\pm$.17
	<i>P</i> =.0008	<i>P</i> =.50	<i>P</i> <.0001
Physically active *			
No (n=176)	.038 $\pm$.15	.035 $\pm$.15	.050 $\pm$.29
Yes (n=371)	.035 $\pm$.15	.037 $\pm$.15	.037 $\pm$.20
	<i>P</i> =.04	<i>P</i> =0.12	<i>P</i> <.0001

* Physically active defined as walking at least 7 blocks per week.