Title: Inter- and intra-rater reliability of the GAITRite system among individuals with sub-acute stroke

Authors: Jennifer S. Wong;[1,2] Hardika Jasani;[2,3] Vivien Poon;[1,2] Elizabeth L. Inness;[1,2] William E. McIlroy;[1-4] Avril Mansfield[1,2,4]

Affiliations: [1]University of Toronto, Toronto, ON, Canada; [2]Toronto Rehabilitation Institute - UHN, Toronto, ON, Canada; [3]University of Waterloo, Waterloo, ON, Canada; [4]Heart and Stroke Foundation Canadian Partnership for Stroke Recovery, ON, Canada

Study conducted at: Toronto Rehabilitation Institute - UHN**;** 550 University Ave. Toronto, ON, M5G 2A2, Canada

Acknowledgements: This study was supported by the Heart and Stroke Foundation Canadian Partnership for Stroke Recovery and the Canadian Institutes of Health Research. The authors acknowledge the support of the Toronto Rehabilitation Institute. Equipment and space have been funded with grants from the Canada Foundation for Innovation, Ontario Innovation Trust and the Ministry of Research and Innovation. The views expressed do not necessarily reflect those of the Ministry.

Corresponding author: Avril Mansfield; 550 University Ave. Toronto, ON, M5G 2A2
Tel: 416-597-3422 ext. 7831; e-mail: avril.mansfield@uhn.ca

Word count: 1,198

Keywords: Stroke; Reliability of Results; Gait

Conflict of interest: The authors declare no conflict of interest.

ABSTRACT

Technology-based assessment tools with semi-automated processing, such as pressure-sensitive mats used for gait assessment, may be considered to be objective; therefore it may be assumed that rater reliability is not a concern. However, user input is often required and rater reliability must be determined. The purpose of this study was to assess the inter- and intra-rater reliability of spatial and temporal characteristics of gait in stroke patients using the GAITRite system. Forty-six individuals with stroke attending in-patient rehabilitation walked across the pressure-sensitive mat 2-4 times at preferred walking speeds, with or without a gait aid. Five raters independently processed gait data. Three raters re-processed the data after a delay of at least 1 month. The intraclass correlation coefficients (ICC) and 95% confidence intervals of the ICC were determined for velocity, step time, step length, and step width. Inter-rater reliability for velocity, step time, and step length were high (ICC >0.90). Intra-rater reliability was generally greater than inter-rater reliability (from 0.81 to >0.99 for inter-rater versus 0.77 to >0.99 for intra-rater reliability). Overall, this study suggests that GAITRite is a reliable assessment tool; however, there still remains subjectivity in processing the data, resulting in no patients with perfect agreement between raters. Additional logic checking within the processing software or standardization of training could help to reduce potential errors in processing.

RESEARCH HIGHLIGHTS

- Inter-rater reliability for velocity, step time, and step length were high.
- Intra-rater reliability was generally greater than inter-rater reliability.
- Logic checking or standardization of training may reduce potential processing errors.
- Overall, this study suggests that GAITRite is a reliable assessment tool of gait post-stroke.

INTRODUCTION

Walking assessment post-stroke is important to inform physiotherapy intervention.[1] Clinical gait assessment can be facilitated by technology, such as pressure-sensitive mats that detect footfall location and timing during walking.[2,3] This technology objectively calculates measures such as walking speed (a measure of overall gait function),[4] stride-to-stride variability (a measure of dynamic stability control),[5] and symmetry (a measure of gait quality).[6] The GAITRite walkway is one such pressure-sensitive mat with good test-retest reliability at comfortable walking speeds among individuals with stroke (intraclass correlation coefficients (ICCs): 0.72-0.98).[7,8] Since GAITRite is a computerized assessment tool with semi-automated processing it may be assumed that rater reliability is not a concern. However, many individuals with impaired gait, such as those with stroke, exhibit atypical gait patterns. Use of a gait aid, shuffling steps, and heel or toe drag often requires input by the user for data processing (e.g. deleting a portion of or an entire footfall), which may influence reliability. This study aimed to determine inter- and intra-rater reliability of GAITRite for evaluating spatio-temporal gait characteristics post-stroke.

METHODS

Participants

This study involved secondary analysis of data collected during routine care at Toronto Rehabilitation Institute. We randomly selected 50 patients from a pool of 195 who completed a walking assessment on admission to in-patient stroke rehabilitation. Three patients were subsequently removed due to an insufficient number of footfalls (<18), and 1 patient was removed because, upon chart review, diagnosis of recent stroke was not confirmed by neuroimaging. Thus, gait data of 46 patients were re-analyzed for this study. Table 1 shows patient characteristics.

Gait assessment

The assessment was performed using the GAITRite walkway ('Platinum' model, CIR Systems Inc., Clifton, New Jersey, USA), which records the location and timing of each footfall during ambulation. Patients walked at their preferred speed using their usual gait aid, if applicable. A physiotherapist walked beside the patient to ensure safety. Each participant completed enough passes to allow for at least 18 footfalls to be analyzed (2-4 passes).

Raters

Five raters independently processed gait data using the GAITRite application software (version 4.0). Two had less than 2 months' experience, two had 1-2 years' experience, and one had >5 years' experience with GAITRite. Raters read a set of instructions for data processing and were knowledgeable of gait mechanics and stroke rehabilitation. Three raters with varying levels of experience with GAITRite were available to re-process the data after a delay of at least 1 month.

Data analysis

Basic spatio-temporal measures were calculated for each patient using the GAITRite software: velocity, step time, step length, and step width. Step time, length, and width were averaged for the left and right limbs separately. Values were averaged across all steps for each patient. Inter- and intra-rater reliability were assessed using intraclass correlation coefficients ($ICC_{2,1}$).[9] To determine inter-rater reliability, the ICCs were calculated separately for each pair of raters; the resulting 10 ICC values were used to calculate the mean and 95% confidence interval of the ICC for each gait measure. For intra-rater reliability, ICCs were calculated separately for each rater. We determined that a measure had 'acceptable' reliability if the 95% lower confidence limit was >0.90.[10] Patients requiring manual processing were isolated and ICCs were re-calculated. To identify patients with large variability between raters, we calculated the coefficient of variation of each measure across raters. Patients with a coefficient of variation >5% were examined in greater detail to determine sources of disagreement.

Table 1: Patient information. Values presented are means (standard deviations; ranges) or counts for the 46 patients included in the study. Spatio-temporal gait characteristics are from rater B's first attempt at processing the data.

Age (years)	68 (14; 35 – 94)
Sex (number)	
Men	32
Women	14
Time post-stroke (days)	27 (15; 8 – 90)
Type of stroke (number)	
Ischemic	30
Hemorrhagic	11
Transforming to hemorrhagic	5
Affected hemisphere (number)	
Right	12
Left	22
Both	12
Affected side (number)	
Right	27
Left	13
Both	6
Berg Balance Scale[11] (score)	37.0 (13.6; 4 - 56)
Clinical Outcome Variables Scale[12] (score)	67.6 (11.3; 39 – 91)
Gait aid (number)	
None	14
Single-point cane	4
Quad cane	2
Two-wheeled walker	12
Rollator	14
Walking speed (m/s)	0.56 (0.27; 0.15 – 1.23)
Left step time (s)	0.84 (0.32; 0.51 – 2.18)
Right step time (s)	0.82 (0.26; 0.55 – 2.08)
Left step length (cm)	41.9 (13.0; 13.7 – 71.6)
Right step length (cm)	41.3 (13.1; 11.9 – 73.5)
Left step width (cm)	11.2 (3.9; 3.3 – 20.2)
Right step width (cm)	11.2 (3.9; 3.2 – 20.2)

RESULTS

Inter-rater and intra-rater reliability

For inter-rater reliability, the mean ICCs for all gait measures were ≥0.94, except step width, which had ICCs of 0.81 and 0.84 for left and right step width respectively (Table 2). Velocity, step time and step length had acceptable reliability (ICC lower confidence limit>0.90). ICCs for intra-rater reliability ranged from 0.77 to >0.99. Velocity, step time and step length displayed excellent intra-rater reliability (>0.90). Intra-rater reliability was lower for left and right step width, with ICCs ranging from 0.77 to 0.97.

It was possible to apply automatic processing to data of 13 patients; ICCs for inter- and intra-rater reliability for these patients were >0.99. Data for 33 patients required manual processing (i.e. footfall identification algorithm could not be automatically applied); ICCs were slightly lower for this sub-set than for the whole sample (Table 3).

Sources of disagreement

No trial had a coefficient of variation of 0; thus, at least one rater over-wrote the automatic processing for each patient. Thirteen patients had a coefficient of variation >5% for at least one measure. For 6 patients, disagreement between raters was due to rater judgement (i.e., some raters chose to remove footfalls from analysis that other raters chose to keep). For the remaining patients, disagreement was

due to failure to correct for errors in automatic processing. Typically, this was because the software misidentified the left and right feet (5 patients), but in two cases illogical values (e.g. negative step lengths) were calculated for at least one footfall.

Table 2: Inter-rater and intra-rater reliability for each gait measure (all trials; n=46). Values presented are mean ICCs and 95% confidence intervals. Measures with 'acceptable' reliability (i.e. lower 95% confidence limit >0.90) are marked with an asterisk. Prior to the start of the study, Rater A had had less than 2 months' experience, Rater B had >5 years' experience, and Rater C had 1-2 years' experience with the GAITRite system.

Gait measure	Inter-rater reliability	Intra-rater reliability		
		Rater A	Rater B	Rater C
Velocity	0.998 [0.997-0.999]*	>0.99	>0.99	>0.99
Left step time	0.94 [0.92-0.97]*	0.99	>0.99	0.93
Right step time	0.95 [0.92-0.98]*	>0.99	>0.99	0.94
Left step length	0.96 [0.94-0.98]*	0.93	>0.99	>0.99
Right step length	0.98 [0.97-0.99]*	>0.99	>0.99	>0.99
Left step width	0.81 [0.72-0.90]	0.88	0.97	0.77
Right step width	0.84 [0.76-0.91]	0.88	0.97	0.78

Table 3: Inter-rater and intra-rater reliability for each gait measure (trials requiring manual processing; n=33). Values presented are mean ICCs and 95% confidence intervals. Measures with 'acceptable' reliability (i.e. lower 95% confidence limit >0.90) are marked with an asterisk. Prior to the start of the study, Rater A had had less than 2 months' experience, Rater B had >5 years' experience, and Rater C had 1-2 years' experience with the GAITRite system.

Gait measure	Inter-rater reliability	Intra-rater reliability		
		Rater A	Rater B	Rater C
Velocity	0.996 [0.994-0.998]*	0.99	>0.99	>0.99
Left step time	0.93 [0.90-0.96]	0.99	>0.99	0.92
Right step time	0.93 [0.90-0.96]	>0.99	>0.99	0.92
Left step length	0.93 [0.91-0.96]*	0.90	>0.99	>0.99
Right step length	0.97 [0.96-0.99]*	0.99	0.99	>0.99
Left step width	0.76 [0.67-0.85]	0.87	0.97	0.74
Right step width	0.79 [0.72-0.87]	0.86	0.96	0.74

DISCUSSION

This study suggests that GAITRite demonstrates good inter- and intra-rater reliability for spatio-temporal characteristics of gait post-stroke. Nearly all variables met the reliability criterion for application to patient care (ICC >0.90), except for step width. Previous studies of test-retest reliability of GAITRite within healthy individuals also found that step width was less reliable than other spatio-temporal characteristics of walking.[13,14] The authors assumed lower step width reliability was due to the spatial resolution of the GAITRite mat.[14] However, the current study raises the possibility that test-retest reliability of step width may have been confounded by combined rater and software error.

While no trial had perfect agreement between raters, disagreement was due to differences in decision making (e.g. to remove or keep certain footfalls) and errors in automatic processing not detected by the raters (e.g. the software misidentified the left and right footfalls). In other cases, the software miscalculated values for individual steps despite the fact that all footfalls appeared to be identified correctly on visual inspection. When raters must intervene and manually process gait data, there may be disagreement when handling atypical gait patterns. An additional algorithm can identify potential errors in processing (e.g. negative step length or time) prior to generating a clinical report.

This work has implications for application of technology to clinical assessment. It may be assumed that extraction of features of movement with semi-automated processing is more accurate and precise than human observation. Several studies have demonstrated good test-retest reliability[7,8,13,14] and concurrent validity[15] of GAITRite, but none previously determined rater reliability. However, this study demonstrates that errors and disagreements in data processing can exist when semi-automated technology is used, and that rater reliability must be still determined.

Conclusions

Overall, the results suggest that GAITRite is a reliable assessment tool for most gait measures among individuals with sub-acute stroke. Wherever possible, we recommend that the same rater process all GAITRite data. For clinical applications or large research studies this may not be possible; therefore, we recommend that investigators determine inter-rater reliability for individuals who will be processing gait data and for the specific variables of interest prior to generating clinical reports or starting the study. When patients or study participants are followed over time, we recommend that the same individual process data.

REFERENCES

1. Latham NK, Jette DU, Slavin M, et al. Physical therapy during stroke rehabilitation for people with different walking abilities. Arch Phys Med Rehabil. 2005;86:S41-50.

2. 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;82:419-425.

3. Bilney B, Morris M, Webster K. Concurrent related validity of the GAITRite® walkway system for quantification of the spatial and temporal parameters of gait. Gait Posture. 2003;17:68-74.

4. Schmid A, Duncan PW, Studenski S, et al. Improvements in speed-based gait classifications are meaningful. Stroke. 2007;37:2096-2100.

5. Balasubramanian CK, Neptune RR, Kautz SA. Variability in spatiotemporal step characteristics and its relationship to walking performance post-stroke. Gait Posture. 2009;29:408-414.

6. Patterson KK, Parafianowicz I, Danells CJ, et al. Gait asymmetry in community-ambulating stroke survivors. Arch Phys Med Rehabil. 2008;89:304-310.

7. Kuys SS, Brauer SG, Ada L. Test-retest reliability of the GAITRite system in people with stroke undergoing rehabilitation. Disabil Rehabil. 2011;33:1848-1853.

8. Lewek MD, Randall EP. Reliability of spatiotemporal asymmetry during overground walking for individuals following chronic stroke. J Neurol Phys Ther. 2011;35:116-121.

9. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86:420-428.

10. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. 3rd ed: Prentice Hall; 2007.

11. Berg K, Wood-Dauphinée S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. Physiother Can. 1989;41:304-311.

12. Seaby L, Torrance G. Reliability of a physiotherapy functional assessment used in a rehabilitation setting. Physiother Can. 1989;41:264-270.

13. Menz HB, Latt MD, Tiedemann A, Kwan MMS, Lord SR. Reliability of the GAITRite walkway system for the quantification of temporo-spatial parameters of gait in young and older people. Gait Posture. 2004;20:20-25.

14. van Uden C, Besser M. Test-retest reliability of temporal and spatial gait characteristics measured with an instrumented walkway system (GAITRite(R)). BMC Musculoskeletal Disorders. 2004;5:13.

15. Webster KE, Wittwer JE, Feller JA. Validity of the GAITRite walkway system for the measurement of averaged and individual step parameters of gait. Gait Posture. 2005;22:317-321.