Validity of the StepWatch

Criterion validity of the StepWatch Activity Monitor as a measure of walking activity in individuals following stroke.

Criterion validity of the StepWatch Activity Monitor as a measure of walking activity in individuals following stroke.

Abstract

Objectives: The StepWatch Activity Monitor (SAM) is an accelerometer-based microprocessor designed for use in long-term ambulatory monitoring. The primary goal of this study was to test its validity in subjects with stroke against two criterion standards, 3-dimensional gait analysis (3-DGA) and footswitches in a variety of indoor and outdoor walking conditions, including different speeds and different terrains. A secondary aim was to test accuracy of the SAM when worn on the paretic limb.

Design: Criterion standard validation study

Setting: Gait laboratory and outside course

Participants: 25 participants with physical disability following stroke

Interventions: Not applicable.

Main Outcome Measures: Total step count measured simultaneously by SAM and either 3-DGA or footswitches for both paretic and non-paretic limbs

Results: Total step count measured by the SAM and 3-DGA was highly correlated (non-paretic limb, r=0.9585; paretic limb, r=0.8960). 95% limits of agreement (derived from Bland Altman analyses) between the SAM and 3-DGA were within ±10 steps for SAMs worn on either the non-paretic or paretic limb.

Total step count measured simultaneously by the SAM and footswitches was also highly correlated for each limb (non-paretic, r=0.9989; paretic, r=0.9631). 95% limits of agreement between the SAM and footswitches were ±9 steps on the non-paretic limb but higher at ±57 steps on the paretic limb. Further analysis showed that the measurement differences occurred during the outdoor component of the combined walk. 95% limits of agreement between footswitches on both limbs were not more than ±9 steps for walking, suggesting that the error was accounted for by the SAM on the paretic limb, which both over and under-read total step count in the outdoor walking conditions.

Conclusions: Criterion validity of the SAM to measure steps in both clinical and natural environments has been established when used on the non paretic limb. However more errors are apparent when the SAM is worn on the paretic limb whilst walking over a variety of outdoor terrains. Validation is recommended prior to use in patients with neurological conditions affecting bilateral legs as there may be more error, particularly in outdoor activities.

Key Words: Cerebrovascular accident; Gait; Rehabilitation; Motor activity

Stroke is the most common cause of severe disability in adults,[1] with persistent physical disability reported in 50-65% of individuals who survive stroke.[1-3] Although as many as 70% are able to walk independently,[3] it appears that only a small percentage of these individuals are able to walk functionally in the community.[4, 5]

There are a wide range of outcome measures available to assess walking ability following stroke.[6-8] The majority of these measures are directly observed tests administered in a standardized clinical or laboratory setting. These tests are thus more likely to measure capacity, which can be defined as the highest probable level of functioning that a person may reach in a given domain at a given moment.[9] Recently accelerometer based technology has been introduced as a way to measure ambulatory activity in an individual's usual environment. Activity monitors are small unobtrusive microprocessors worn for continuous monitoring usually at one body site. There are many such devices which are able to record data for extended periods, thus providing an objective measure of a person's performance rather than their capacity as might be measured in clinical or laboratory settings.[9]

The StepWatch 3 Activity Monitor (SAM)[a] is an example of an accelerometer based activity monitor that has been used widely in different population groups. The SAM is small (75 X 50 X 20 mm) and lightweight (38 g) and is worn at the ankle. The monitor contains a custom sensor that uses a combination of acceleration, position, and timing to detect steps. The SAM is calibrated based on each individual's height and gait pattern and the threshold can be adjusted for individuals with altered gait patterns.

Criterion validity of the SAM has been assessed by comparison to a hand held counter in healthy children,[10] adults with diabetes or lower limb amputation,[11] healthy and obese adults,[12, 13] adults with hip or knee arthroplasty,[13] healthy adults with total contact casts[14] and adults with stroke.[15] These studies have mainly tested validity in controlled settings over short distances with small variations in conditions, including different walking speeds,[12, 13, 15] two footwear conditions (athletic shoes, total contact cast),[14] stairs[11, 13, 14] and slopes (9% gradient).[11] Only one study of subjects with diabetes or lower limb amputation specifically mentions walking outdoors during the accuracy testing[11] and conditions of walking on uneven ground and grass have not been previously reported. Thus, there is little information on accuracy in outdoor conditions.

The SAM has also undergone reliability,[15, 16] concurrent validity[16, 17] and sensitivity[18] testing in participants with stroke, however testing has been carried out exclusively on the non-paretic leg. No studies have yet looked at criterion validity on the paretic limb. Foster et al report almost perfect agreement between monitors worn on each leg of healthy subjects (at worst 99.82%

accuracy)[12], however it is not known whether there is a difference in accuracy between the paretic and non-paretic side of individuals with stroke. Although the gait pattern is altered bilaterally after stroke, more deficits are usually seen on the paretic side. It is possible that biomechanical changes might result in less accurate calibration and hence, measurement on the paretic limb.

In addition, the SAM has not been tested for criterion validity against 3-dimensional gait analysis (3-DGA) or against other laboratory measures, such as footswitches. Although laboratory measures do not always reflect community ambulation, they are good criterion standards for assessment of the SAM, as they are known for their accuracy in step counting.

Thus, the aims of this study were to test the criterion validity of the SAM compared to two criterion standards (3-DGA and footswitches) over a variety of indoor (3-DGA, footswitches) and outdoor surfaces (footswitches) and between non-paretic and paretic legs.

Methods

Participants

A convenience sample of 25 individuals with chronic stroke was recruited from the hospital stroke service and local newspaper advertising. This sample size was chosen to provide sufficient numbers for analysis, based on previous sample sizes of between 10 and 16 in similar validation studies [11, 15]. Participants were eligible for inclusion if they were at least six months post stroke, were aged between 30-80, had not had more than two falls in the previous six months and had not had any lower limb surgery or botox treatment for their walking in the previous year. All participants were able to walk independently but with some residual difficulty confirmed by less than the full score on the physical functioning scale of the SF-36 (scored out of 30, with higher scores indicating better physical functioning). All participants gave written informed consent, and the study was approved by the Northern Y Regional Ethics Committee.

Testing Protocol

All participants attended the Gait Laboratory for testing. The Functional Walking Category[19] and the Rivermead Mobility Index (RMI)[20-22] were administered for descriptive purposes. The Functional Walking Category is a single scale with six levels of walking disability, ranging from physiological walker (1) to community walker (6). Levels 4-6 are community walkers with limitations noted at levels 4 and 5. The RMI is a scale to capture self-reported mobility and is scored out of a total of 15, with higher scores indicating better mobility. The items are hierarchical and the most difficult item relates to running. SAMs were calibrated to record data at three second intervals and were strapped to the lateral side of the ankle of both the non paretic and paretic legs. The sensitivity and cadence settings were adjusted for each participant so that the monitor recognised every step during fast, slow and self selected walking speeds.

Twenty-three retro-reflective markers were placed on pre-determined anatomical landmarks of the trunk, upper and lower limbs. Three-dimensional kinematic data were concurrently collected by an eight camera Vicon system[b] (sampling rate, 100 Hz). Foot events were identified in Workstation 5.2.4[b] and total left and right steps for each walk were counted separately.

Each participant was instructed to walk at a self selected pace on the six metre walkway with a five second pause before and after each turn. The pause was to ensure the SAM could differentiate each walking trial. Each participant completed six trials without shoes.

The second stage of the study involved the simultaneous collection of data with the footswitches (flexible on/off event switch)[c] and the SAM. Footswitches were chosen for the outdoor condition as 3-DGA cannot be used outside a laboratory environment. The footswitches were taped to the first metatarsal head of each foot and connected to a datalogger[d] which was worn in a small bag around the waist. The participants wore their usual shoes and orthotics for this part of the study.

Each participant initially walked for eight metres at a self-selected pace followed by a further eight metres at maximal pace following the instructions, 'Walk as fast as you safely can." Both of these walks were separated by five second pauses. The participant then walked over a predetermined outside course of approximately 200 metres, which included ascending and descending nine steps, walking on concrete, grass and negotiating a 16% incline and a 14% decline. Participants walked at a self-selected pace and were able to rest if required. Participants had the option to avoid part of the course if they perceived it was too difficult.

Statistical analyses

Pearson's correlation coefficients were calculated to assess the level of association between the SAM and the two criterion standards (3-DGA, footswitches) for both the non-paretic and paretic lower limbs. A Pearson's correlation coefficient (r) of above 0.85 was considered to be an acceptable correlation[23]. The levels of agreement between the SAM and the criterion standards (3-DGA or footswitches) were also calculated using methodology described by Bland and Altman[24, 25]. Bland and Altman advocate plotting the differences between the two methods against the mean of the two measures to give both an indication of bias between the two methods of measurement and also a 95% confidence interval, based on the calculated standard deviation of the differences. All calculations were performed using GraphPad Prism[e].

The percentage error of the SAM compared to the two criterion standards was calculated as (SAM count-criterion count)/criterion count X 100. A positive value indicates overcounting by the SAM and a negative value indicates undercounting.

Results

Twenty-five participants with a median age of 69 years (range 42 to 79) were enrolled in the study. There were 17 men and 8 women. Ten participants had right sided paresis. The median score on the physical functioning index of the SF-36 was 19 (range 11 to 29). The median gait speed was 0.5 m/s (range 0.1 to 0.9). All participants walked independently with a median score on the RMI of 14 (range 10 to 15). Twenty-two participants reported independent walking outside over pavements (item 9 of the RMI) and 20 participants reported independent ability over uneven surfaces (item 12 of the RMI); the remaining participants were not independent for these items. The median score of 6 (range 3 to 6) on the functional walking category indicates that the majority of participants rated themselves as community ambulators.

Agreement between 3-DGA and SAM

Participants took between 55 and 133 steps during the repeated six metre walks of the 3-DGA. Pearson's correlation between SAM measured 'total step count' and the 3-DGA measured 'total step count' were high, both for the non-paretic limb (r=0.959) and the paretic limb (r=0.896) (Fig 1A-B). 95% limits of agreement with 3-DGA (derived from Bland Altman analyses) were ±7 steps for the SAM on the non-paretic limb (Fig 1C) and ±10 steps for the SAM on the paretic limb (Fig 1D). There was a positive bias of one step for the non-paretic limb and three steps for the paretic limb, indicating that the SAM undercounted steps compared to 3-DGA on both sides. The mean error for the non-paretic side was -2.6% (range -26% to 16%) and was less than the mean error for the paretic side, which was -7.3% (range -36% to 5.7%). This was not a significant difference ($p=0.342$, $\alpha=0.05$).

Agreement between footswitches and SAM

Twenty-one participants completed the combined indoor and outdoor walking trials, with four unable to fully complete the outdoor walking trial due to limitations in physical ability or confidence.

The correlation between footswitch measured 'total step count' and SAM measured 'total step count' was high for both the non-paretic limb (r=0.999) and the paretic limb (r=0.963) over the combined indoor and outdoor walks. However, the 95% limits of agreement derived from Bland-Altman plots showed much wider limits of agreement between the SAM and footswitches for the paretic limb of ±57 steps, compared to ±9 steps on the non-paretic limb (Table 1). The mean errors were -1.3% (range -4.5% to 2.5%) and -4.2% (range -42% to 16%) for the non-paretic and paretic limbs respectively indicating that the SAM both under and over counted steps. The mean errors for each limb were not significantly different (p=0.220, α=0.05) but a wider range of errors was noted for the paretic limb.

When the walking conditions were analysed separately, the 95% limits of agreement between the SAM and footswitches for the combined indoor walk and the separated indoor walks at self selected and fast speeds were similar for the non-paretic and paretic limbs (Table 1). However the outdoor course revealed higher 95% limits of agreement of ±55 steps between SAM and footswitch measured total step count for the paretic limb (Table 1), which was not apparent from the correlation coefficient (r=0.999 for non-paretic limb, r=0.963 for paretic limb) or the mean error (non-paretic side, -1.1% (range: -4.7% to 2.6%); paretic side, -4.9% (range: -66% to 17%)).

There was high agreement between the footswitch measured 'total step count' for both non-paretic and paretic limbs under all conditions shown by the Pearson's correlation coefficients ranging from 0.983 to 0.999 (Fig 2A) and the narrow 95% limits of agreement (Fig 2C, Table 2). The Pearson's correlation coefficients were also high between the SAM measured 'total step count' for the paretic limb and the SAM measured total step count for the non-paretic limb (Fig 2B) (range: 0.940 to 0.973). However, the 95% limits of agreement (Fig 2C-D) showed that there was more error accounted for by the SAM on the paretic side during the outdoor course (Table 2). In particular, there were two participants with steps on the paretic side outside the limits of agreement. The SAM undercounted on the paretic limb by 110 steps for one participant and overcounted by 69 steps for another.

No correlation was found between gait velocity and either the absolute percentage error (non-paretic, r=-0.176; paretic, r=0.023) or the difference in total steps counted by the two measurement devices (non-paretic, r=-0.122; paretic, r=-0.159). Similarly, the SF 36 score did not correlate with either the absolute percentage error (non-paretic, r=0.138; paretic, r=0.011) or the difference in total steps counted by the two measurement devices (non-paretic, r=0.081; paretic, r=-0.051).

Discussion

In this study, we show that the SAM has good criterion validity for adults with stroke compared to 3-DGA and footswitches. This extends previous work with handheld counters[15]. Our study also extends previous work by using different environments and conditions, which we selected for their relevance to community mobility[26]. Therefore, a range of commonly encountered outdoor terrains was included; uneven surfaces, concrete, grass, inclines, declines and stairs. As the SAM is intended to be a measure of performance rather than capacity, it is important that it is validated in similar environments to its intended use, rather than a laboratory.

Our results confirm that the SAM is accurate to ± 9 steps when used on the non-paretic leg over a range of outdoor terrains. The 98.6% accuracy of the SAM in this study is consistent with previous studies that have reported percentage accuracies of 92.7% to 99.7% , when the SAM is compared to a handheld counter[10-15]. It is also encouraging to report that the accuracy of the SAM was similar in this study for self selected and fast speeds. This is relevant to community mobility where a range of speeds may be employed depending upon the task and context.

However, our results have identified reduced accuracy when the SAM is used on the paretic limb, which has not previously been reported. Although the mean error on the paretic limb is -4.9% and is consistent with previously reported mean error for stairs and slopes[11, 13, 14], Bland-Altman analyses reveal wide limits of agreement of ±55 steps for the walking outdoors test condition. The discrepancy between these two statistical tests is due to the averaging employed by the mean error score effectively cancelling out positive and negative values, thus masking the range of the error. In contrast, error is revealed by Bland and Altman plots, as the difference of the individual means is plotted against the average of the individual means[24]. The finding of increased error when the SAM is used on the paretic limb is an important finding and has implications for use of the SAM in bilateral neurological conditions, such as incomplete spinal cord injury or multiple sclerosis.

There are several possible explanations for the wide limits of agreement attributable to the SAM on the paretic limb. The SAM may not have been calibrated correctly. This risk was minimised by careful checking of the step counting at a range of speeds when first applied. The SAM was calibrated until it was detecting steps correctly in line with previous protocols[15]. It is also unlikely that calibration was incorrect as the limits of agreement for indoor walking conditions were similar to the non-paretic side and the SAM was not recalibrated between conditions.

The SAM is a microprocessor linked to an accelerometer which detects motion particularly at the hip and knee. It is possible that some participants altered their gait pattern during the outdoors condition in a number of ways that might affect acceleration. Firstly, a change of gait speed will result in altered step length and an associated change of acceleration of the leg. However, if this occurred, we would have expected to see errors occurring in both legs rather than just the paretic side.

It is also possible that the paretic leg changed the amount of motion during walking, and therefore acceleration. It is feasible to think of conditions where this might occur. Negotiating stairs or inclines possibly results in an increased degree of motion at the hip and knee, which may account for the mean positive error, indicating step overcounting, as found in two previous studies[11, 13]. However, this hypothesis is not supported by a third study, which reported a negative error during step negotiation, implying step undercounting[14]. The application of a total contact cast in the latter study may have contributed to this difference.

It is also feasible that movement of the paretic limb might be restricted by spasticity, which may be increased in certain situations. It was not possible to test this theory by assessing changes in the gait pattern when the participant was walking outdoors, and this is a limitation of this study.

The SAM exhibited a small but consistent bias of undercounting compared to both 3-DGA and footswitches. This is most likely due to differences between step counting definitions. As previously discussed, the SAM identifies a step at a threshold of acceleration of the leg. However 3-DGA identifies a step based on foot events, namely initial contact and toe off. So 3-DGA will count a very small step, which may not be detected by the SAM. This was seen to be the case during the first indoor walking trial comparing SAM with 3-DGA, when most participants took a step of less than the normal step length at the end of each walk to bring their feet together as they stopped. Footswitches define steps in another way as they are a pressure system, which counts each step from one pressure on to the next. So it is feasible that shifting the body weight without lifting the foot could be detected as a step.

It is understandable then that both the 3-DGA and footswitches will define steps that are not identified by the SAM. The different definitions of steps highlight some considerations for use of activity monitors. If an individual is engaged in a lot of interrupted walking, it is possible that the SAM may not identify all steps and therefore under-represent activity. It is likely that the SAM will be a more accurate representation of continuous walking.

Limitations of this study include the moderately small sample size, which may not be representative of the entire population. The study also excluded individuals who were not able to walk independently, thus limiting the generalizability of the results to this particular group. This study is also limited by the lack of available criterion standards for community ambulation. We chose 3-DGA and footswitches, which are laboratory measures, as criterion standards for step counts. However, the difference in step definition between the criterion standards and the SAM may have contributed to some of the step undercounting detected in the use of the SAM.

Conclusion

This study has shown that the SAM has criterion validity when used on the non-paretic limb to measure steps in both clinical and natural environments. However more errors are apparent when the SAM is worn on the paretic limb whilst walking over a variety of terrains. Validation is recommended prior to use of the SAM in patients with neurological conditions affecting bilateral legs as there may be more error, particularly in outdoor environments.

Acknowledgements

We gratefully acknowledge the contributions of Pat Bennett and Alan Barber, MD, PhD, of the Auckland District Health Board for subject recruitment.

References

1. Adamson J, Beswick A, Ebrahim S. Is stroke the most common cause of disability? J Stroke Cerebrovasc Dis 2004;13(4):171-7.
2. Bonita R, Solomon N, Broad JB. Prevalence of stroke and stroke-related disability. Estimates from the Auckland stroke studies. Stroke 1997;28(10):1898-902.
3. Kelly-Hayes M, Beiser A, Kase CS, Scaramucci A, D'Agostino RB, Wolf PA. The influence of gender and age on disability following ischemic stroke: The Framingham study. J Stroke Cerebrovasc Dis 2003;12(3):119-26.
4. Goldie PA, Matyas TA, Evans OM. Deficit and change in gait velocity during rehabilitation after stroke. Arch Phys Med Rehabil 1996;77(10):1074-82.
5. Lord SE, McPherson K, McNaughton HK, Rochester L, Weatherall M. Community ambulation after stroke: how important and obtainable is it and what measures appear predictive? Arch Phys Med Rehabil 2004;85(2):234-9.
6. Douglas H, Swanson C, Gee T, Bellamy N. Outcome measurement in Australian rehabilitation environments. J Rehabil Med 2005;37(5):325-9.
7. Haigh R, Tennant A, Biering-Sorensen F, Grimby G, Marincek C, Phillips S et al. The use of outcome measures in physical medicine and rehabilitation within Europe. J Rehabil Med 2001;33(6):273-8.
8. Mudge S, Stott S. Outcome measures to assess walking ability following stroke - a systematic review of the literature. Physiotherapy in press.

9. International Classification of Functioning, Disability and Health: ICF. Geneva: World Health Organization; 2001.

10. McDonald CM, Widman L, Abresch RT, Walsh SA, Walsh DD. Utility of a step activity monitor for the measurement of daily ambulatory activity in children. Arch Phys Med Rehabil 2005;86(4):793-801.

11. Coleman KL, Smith DG, Boone DA, Joseph AW, del Aguila MA. Step activity monitor: long-term, continuous recording of ambulatory function. J Rehabil Res Dev 1999;36(1):8-18.

12. Foster RC, Lanningham-Foster LM, Manohar C, McCrady SK, Nysse LJ, Kaufman KR et al. Precision and accuracy of an ankle-worn accelerometer-based pedometer in step counting and energy expenditure. Prev Med 2005;41(3-4):778-83.

13. Shepherd EF, Toloza E, McClung CD, Schmalzried TP. Step activity monitor: increased accuracy in quantifying ambulatory activity. J Orthop Res 1999;17(5):703-8.

14. Hartsell H, Fitzpatrick D, Brand R, Frantz R, Saltzman C. Accuracy of a custom-designed activity monitor: implications for diabetic foot ulcer healing. J Rehabil Res Dev 2002;39(3):395-400.

15. Macko RF, Haeuber E, Shaughnessy M, Coleman KL, Boone DA, Smith GV et al. Microprocessor-based ambulatory activity monitoring in stroke patients. Med Sci Sports Exerc 2002;34(3):394-9.

16. Haeuber E, Shaughnessy M, Forrester LW, Coleman KL, Macko RF. Accelerometer monitoring of home- and community-based ambulatory activity after stroke. Arch Phys Med Rehabil 2004;85(12):1997-2001.

17. Michael KM, Allen JK, Macko RF. Reduced ambulatory activity after stroke: the role of balance, gait, and cardiovascular fitness. Arch Phys Med Rehabil 2005;86(8):1552-6.

18. Shaughnessy M, Michael KM, Sorkin JD, Macko RF. Steps after stroke: capturing ambulatory recovery. Stroke 2005;36(6):1305-7.

19. Perry J, Garrett M, Gronley JK, Mulroy SJ. Classification of walking handicap in the stroke population. Stroke 1995;26(6):982-9.

20. Collen FM, Wade DT, Robb GF, Bradshaw CM. The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. Int Disabil Stud 1991;13(2):50-4.

21. Green J, Forster A, Young J. A test-retest reliability study of the Barthel Index, the Rivermead Mobility Index, the Nottingham Extended Activities of Daily Living Scale and the Frenchay Activities Index in stroke patients. Disabil Rehabil 2001;23(15):670-6.

22. Antonucci G, Aprile T, Paolucci S. Rasch analysis of the Rivermead Mobility Index: a study using mobility measures of first-stroke inpatients. Arch Phys Med Rehabil 2002;83(10):1442-9.

23. McDowell I. Measuring health:a guide to rating scales and questionnaires. Third ed. New York: Oxford University Press; 2006.

24. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307-10.

25. Bland JM, Altman DG. Measurement error. Bmj 1996;312(7047):1654.

26. Patla A, Shumway-Cook A. Dimensions of mobility: defining the complexity and difficulty associated with community walking. J Aging Phys Activ 1999;7:7-19.

Suppliers

a. Cyma Corporation, 6405 218th St SW, Suite 100, Mountlake Tce, WA 98043-2180, US

b. Oxford Metrics Ltd, 14 Minns Business Park, West Way, Oxford, OX2 0JB, UK

c. Motion Lab Systems, 15045 Old Hammond Highway, Baton Rouge, LA 70816, US

d. Plab vers 1, run on Compaq IPAQ Pocket PC, Hewlett-Packard Company, 3000 Hanover Street, Palo Alto, CA 94304-1185 US

e. Version 4.03; GraphPad Software Inc, 11452 El Camino Real, #215 San Diego, CA 92130, US

Fig. 1A-D. Comparison of concurrent measures of total step count by SAM and by 3-DGA: Scatterplots of total step count by SAM and by 3-DGA for both the non-paretic limb (A) and the paretic limb (B) show high correlations between SAM and 3-DGA for both limbs. Bland Altman plots graphed as the average and difference of the total step count measured by SAM and 3-DGA show smaller 95% limits of agreement for the non-paretic limb (C) compared to the paretic limb (D).

Fig. 2A-D. Comparison of concurrent measures of total step count for paretic and non-paretic limbs: Scatterplots of total step count for non-paretic and paretic limbs measured by footswitches (A) and SAM (B). Bland Altman plots graphed as the average and difference of the total step counts of the non-paretic and paretic limbs show smaller 95% limits of agreement for footswitches (C) compared to the SAM (D).

Table 1. Total step count: 95% limits of agreement between footswitches and SAM for each limb under different walking conditions

Table 2. Total step count: 95% limits of agreement between non-paretic and paretic limbs for both devices (SAM & footswitches) under different walking conditions