Predicting Home and Community Walking Activity Poststroke

George D. Fulk, PhD; Ying He, PhD; Pierce Boyne, DPT; Kari Dunning, PhD

Background and Purpose—Walking ability poststroke is commonly assessed using gait speed categories developed by Perry et al. The purpose of this study was to reexamine factors that predict home and community ambulators determined from real-world walking activity data using activity monitors.

Methods—Secondary analyses of real-world walking activity from 2 stroke trials. Home (100–2499 steps/d), most limited community (2500–4499 steps/d), least limited community (5000–74999), and full community (≥75000 steps/d) walking categories were developed based on normative data. Independent variables to predict walking categories were comfortable and fast gait speed, 6-minute walk test, Berg Balance Scale, Fugl Meyer, and Stroke Impact Scale. Data were analyzed using multivariate analyses to identify significant variables associated with walking categories, bootstrap method to select the most stable model and receiver-operating characteristic to identify cutoff values.

Results—Data from 441 individuals poststroke were analyzed. The 6-minute walk test, Fugl Meyer, and Berg Balance Scale combined were the strongest predictors of home versus community and limited versus unlimited community ambulators. The 6-minute walk test was the strongest individual variable in predicting home versus community (receiver-operating characteristic area under curve=0.82) and limited versus full community ambulators (receiver-operating characteristic area under curve=0.76). A comfortable gait speed of 0.49 m/s discriminated between home and community and a comfortable gait speed of 0.93 m/s discriminated between limited community and full community ambulators.

Conclusions—The 6-minute walk test was better able to discriminate among home, limited community, and full community ambulators than comfortable gait speed. Gait speed values commonly used to distinguish between home and community walkers may overestimate walking activity.

Key Words: discriminant analysis ■ gait ■ goal ■ stroke ■ walking

Recovering walking ability is a primary goal of people with stroke.1,2 Researchers and clinicians commonly use gait speed to assess walking ability poststroke. Gait speed is a valid and reliable measure of walking ability across the continuum of recovery after stroke.3–7 In 1995, Perry et al8 published a seminal study in which they demonstrated that gait speed could discriminate among different levels of walking handicap. Using functional walking categories adapted from Hoffer et al,9 participants were placed into 1 of 6 functional walking categories based on expert clinician opinion: physiological, limited household, unlimited household, most limited community, least limited community, and unlimited community ambulator. Perry et al10 found that there was a significant difference in gait speed between the participants categorized as unlimited community ambulators and the other categories. However, there was no significant difference in gait speed among those participants categorized as most limited community, unlimited household, limited household, and physiological ambulators. Using discriminant analysis, they found that gait speed was the only significant predictor of placement into the 6 walking categories, agreement of 44%. A gait speed of ≥0.42 m/s could distinguish between household and community ambulation.

These results have lead to the widely used gait speed values for categorization as home (<0.40 m/s), limited community (0.40–0.80 m/s), and full community (>0.80 m/s) ambulators.10–12 Although these values may be useful for categorizing people poststroke, gait speed may not reflect actual walking performance in people’s homes and community.13,14 Current advances in technologies allow clinicians and researchers to accurately measure walking activity in people with stroke in their homes and communities using activity monitors.15,16 The purpose of this study was to was to reexamine factors that predict walking activity determined from real-world walking activity data using activity monitors.

Methods

This study was a cross-sectional analysis of data that was obtained from 2 stroke rehabilitation trials, the LEAPS (Locomotor Experience Applied Post-Stroke)12 and FASTEST (Functional Ambulation: Standard Treatment vs Electrical Stimulation)13 trials. Inclusion criteria for both studies were diagnosis of stroke, able to walk ≥10 m with at most maximal assistance, and gait speed <0.80 m/s at start of the...
intervention. More detail on the purpose, inclusion/exclusion criteria, and methods of these trials has been previously published. This study does not report on the outcomes of either trial. Neither trial found a significant difference between groups for the primary or secondary outcomes. This secondary analysis was reviewed by the institutional review board at Clarkson University and was exempt. In the original studies, all participants provided informed consent, and the study protocols were approved by the appropriate institutional review board.

Both trials collected the following data at 12 months poststroke in the LEAPS trial and at ≥10.5 months poststroke in the FASTEST trial: comfortable gait speed (CGS), fast gait speed, lower extremity Fugl Meyer (FM), Berg Balance scale (BBS), SIS-mobility, SIS-participation, functional ambulation category, mini mental state exam, age, sex, marital status, and average steps per day (over a 7-day period in the FASTEST trial and a 2-day period in the LEAPS trial) using an activity monitor (StepWatch Activity Monitor; Orthocare Innovations, Oklahoma City, OK).

Instead of using the descriptors of functional walking categories and assigning participants to one of these categories based on expert opinion as was done in the study by Perry et al, we developed analogous functional walking categories based on population daily walking activity from the study by Tudor-Locke and Bassett and Tudor-Locke et al. Walking activity of 100 to 2499 steps/d was categorized as a household ambulator (sedentary basal walking activity by Tudor-Locke et al), walking activity of 2500 to 4999 steps/d was categorized as a most limited community ambulator (sedentary limited physical activity by Tudor-Locke and Bassett and Tudor-Locke et al), walking activity of 5000 to 7499 steps/d was categorized as a least limited community ambulator (low active by Tudor-Locke and Bassett and Tudor-Locke et al), and walking activity ≥7500 steps/d was categorized as an unlimited community ambulator (somewhat active by Tudor-Locke and Bassett and Tudor-Locke et al).

**Data Analysis**

On the basis of average steps taken per day from the activity monitor, data subjects were divided into 4 groups: home, most limited community, least limited community, and unlimited community ambulators using the steps criteria outlined above. We investigated the prediction performance of the independent variables described above (CGS, fast gait speed, 6MWT, FM, BBS, SIS-mobility, SIS-participation, functional ambulation category, mini mental state exam, age, sex, and marital status) to classify subjects into one of the 4 functional walking categories based on average steps per day. Univariate analysis was initially conducted to see whether there was any difference among the 4 categories for each independent variable with the use of analysis of variance for continuous variables and a χ² test or a Fisher exact test for categorical variables. Tukey comparison was then used to compare the 4 walking categories for those independent variables with significant differences.

Next, we used multivariate analysis to identify predictors for classifying community ambulators versus home ambulators and unlimited community ambulators versus most/least limited community ambulators, respectively. Our goal was to identify the most stable and efficient (ie, a minimum number of necessary independent variables) model. Because unstable models can be found using automated variable selection methods, such as step-wise selection, Bootstrap method (using 500 Bootstrap samples), combined with all-subset methods and Bayesian information criterion as the selection criteria were used in model selection. The sampling and selection process was as follows. First, the stratification technique was used to generate Bootstrap samples: Bootstrap samples with the same number of subjects were taken from each walking category, respectively, and then were merged together to get an overall Bootstrap sample. Next, for each overall Bootstrap sample, the logistic regression was applied to all possible subsets of variables, which were ranked from the highest to lowest order according to Bayesian information criterion from the regression models. Finally, among the 500 overall Bootstrap samples, the subset that was ranked number one most frequently was selected as the best combination of independent variables.

Based on the selected model, the predicted values were generated using a linear combination of the selected subset weighted by the regression coefficients. Receiver-operating characteristic (ROC) curves were then used to plot the value for sensitivity against the false-positive rate (1 specificity). The area under ROC curve (AUC) reflected the predictive ability of the selected subset (a higher AUC indicating a stronger prediction), and an optimal cutoff value for the linear combination was determined by Youden index (ie, the maximum sum of the sensitivity and the specificity).

For comparison, logistic regression models were also fitted when considering each variable in the selected subset as a single predictor. χ² Test was used to compare AUCs between the subset and each variable individually. Youden index was used to define the optimal cutoff point for each selected variable. Similar analysis was applied to CGS if it was not selected by Bootstrap method, in order to compare to Perry et al categories.

The best stable model identified using the bootstrap method from the combined data sets described above was further validated by applying the model to data from each trial (LEAPS and FASTEST) separately using the same ROC curve analyses described above. This was done only for the model that classified community ambulators versus home ambulators. It was not done for the model that classified limited versus unlimited community ambulators. We did not do this for the model developed to classify most/least limited versus full community as there was only one subject from the FASTEST data set that was identified as a full community ambulator (≥7500 steps/d).

To examine the stability of the 2 days of stepping data from the LEAPS data, we performed a paired t test to compare the mean steps taken on both days and an ICC, to examine the agreement in steps taken between the 2 days. All of the above statistical analyses were 2 sided, where P < 0.05 was considered statistically significant.

**Results**

A total of 441 participants were included in our analysis, 298 from the LEAPS trial and 143 from the FASTEST trial. Overall, the mean age was 61.4±12.4 years; mean CGS was 0.6±0.3 m/s; 49% were female; and 80% were independent ambulators. Among 441 participants, 43.08% were classified as household ambulators, 30.39% as most limited community ambulators, 14.29% as least limited community ambulators, and 12.24% as unlimited community ambulators (Table 1).

Table 1 shows the results of the univariate analysis and post hoc multiple comparisons. CGS, fast gait speed, SIS-mobility, SIS-participation, 6MWT, BBS, age, FM, and functional ambulation category were significantly different between the functional walking categories. Community ambulators, especially the unlimited community ambulators, had significantly higher values in these variables than household ambulators. Post hoc multiple comparisons found that for CGS, fast gait speed, 6MWT, BBS, and FM, there was a significant difference between all functional walking categories, except there was no significant difference between most limited and least limited community ambulators.

In the multivariate analysis for predicting community ambulators versus home ambulators, the combination of 6MWT, BBS, and FM were ranked at the top most frequently (241 out of 500 Bootstrap samples, 48.2%), which was suggested as the best subset of independent variables. Figure 1 plots the ROC curves generated from the combined variables model and each selected variable individually. Table 2 summarizes the ROC statistics and the prediction performances based on the corresponding cutoff values (for convenience of comparison, CGS is also included in Figure 1 and Table 2 although CGS was not selected as an independent variable in the best model). Overall, the AUC for 3 variables (6MWT, BBS, and FM) combined.
was 0.836, indicating the combination of these 3 variables had strong prediction. Furthermore, the AUC for 3 variables combined was found to be significantly greater than for each variable alone (P<0.05), showing that using the 3 variables together provided higher predictive power than only using each variable individually as the predictor. When considering the 3 variables individually, the 6WMT had the highest AUC (0.819). A cutoff of 205 m to distinguish between home and community ambulators yielded a slightly lower prediction than using the 3 variables together (sensitivity 71%, specificity 79%, overall accuracy 74%, + likelihood ratio [LR] 3.35, and −LR 0.37).

For predicting unlimited community ambulators versus most/least limited community ambulators, the combination of 6MWT and FM was selected (130/500 Bootstrap samples, 26%). It had an AUC of 0.795, suggesting a strong prediction. Figure 2 plots the ROC curves generated from the combined variables model and each selected variable individually, and the results are summarized in Table 3 (again for convenience of comparison, analysis outputs for CGS are also included in Figure 2 and Table 3 although CGS was not selected as an independent variable in the best model). The ROC shows that AUC for the combination did not significantly differ with the AUC for 6MWT only (P=0.09), indicating that only using 6MWT may provide the prediction as good as the 2 variables combined. A cutoff of 288 m can be used to distinguish between unlimited and limited community ambulators, which achieves similar accuracy as the combined model (sensitivity 68%, specificity 77%, overall accuracy 75%, +LR 2.91, and −LR 0.42).

Although CGS was not selected as a significant independent variable by the Bootstrap method, as noted in the Data Analysis section, ROC analysis was performed for CGS to identify cutoff values for comparison to values proposed by Perry et al.8 When classifying between the community ambulators and home ambulators, AUC of 0.799 for CGS...
method from the combined data sets to just the LEAPS data set to predict between home versus community ambulators, the ROC analysis found an AUC of 0.836, a cutoff of 6.24, a sensitivity of 75%, specificity of 82%, overall accuracy of 77%, +LR of 4.14, and a −LR of 0.30.

When comparing steps taken per day between day 1 and day 2 in the LEAPS data set, the \( t \) test was not significant (\( P>0.05 \)) and the ICC\(_{2,1} \) was 0.79.

**Discussion**

Based on our model selection, a combination of walking endurance, balance, and motor function are the strongest predictors of community walking activity. Walking endurance and motor function were common factors in discriminating between both home and community walking activity and limited and unlimited community walking ambulators, whereas balance was an important factor for discriminating between home and community ambulators. To ambulate in the community individuals poststroke require a minimal level of walking endurance, motor function, and balance. These findings are similar to other studies that have found that walking endurance, balance, and motor function play a role in community walking activity.8,13,28,29

Walking endurance as measured by the 6MWT was the strongest individual predictor of community walking activity. A 6MWT distance \( \geq 205 \) m discriminated between home and community ambulators, whereas a 6MWT distance \( \geq 288 \) m discriminated between limited and unlimited community ambulators. Without this minimal level of walking endurance, individuals poststroke may be confined to their homes and have limited ability to walk in their community. The 6MWT is likely a strong predictor of home versus community walking activity because in addition to measuring functional walking endurance it is related to a variety of other body structure/function, activity, and participation constructs.6,17,30–33 The cutoff values may be useful for clinicians to predict home versus community walking behavior in people with stroke.

Based on our ROC findings, CGS could accurately discriminate between home and community walking activity; however, it was not selected as a significant predictor of home and community walking activity in our models. Although other studies2,8,34 have found CGS to be a significant predictor of community walking as categorized by clinician opinion or self report, none of these studies included the 6MWT as a possible

### Table 2. Receiver-Operating Characteristic Curve Statistics and Prediction Performance for Variables of Interest Used to Discriminate Between Home and Community Ambulators

<table>
<thead>
<tr>
<th>Predictors</th>
<th>ROC Statistics</th>
<th>( \chi^2 ) Test *</th>
<th>Cutoff</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Overall Accuracy, %</th>
<th>LR+</th>
<th>LR−</th>
</tr>
</thead>
<tbody>
<tr>
<td>Six-min walk test, Berg balance scale, Fugl Meyer combined</td>
<td>0.836</td>
<td>...</td>
<td>6.81</td>
<td>70</td>
<td>85</td>
<td>76</td>
<td>4.54</td>
<td>0.36</td>
</tr>
<tr>
<td>Six-min walk test, m</td>
<td>0.819</td>
<td>0.04</td>
<td>205</td>
<td>71</td>
<td>79</td>
<td>74</td>
<td>3.35</td>
<td>0.37</td>
</tr>
<tr>
<td>Berg balance scale, points</td>
<td>0.786</td>
<td>&lt;0.001</td>
<td>48</td>
<td>66</td>
<td>77</td>
<td>71</td>
<td>2.90</td>
<td>0.44</td>
</tr>
<tr>
<td>Fugl Meyer, points</td>
<td>0.747</td>
<td>&lt;0.0001</td>
<td>27</td>
<td>61</td>
<td>77</td>
<td>67</td>
<td>2.59</td>
<td>0.51</td>
</tr>
<tr>
<td>Comfortable gait speed, m/s</td>
<td>0.799</td>
<td>&lt;0.0001</td>
<td>0.49</td>
<td>87</td>
<td>61</td>
<td>76</td>
<td>2.22</td>
<td>0.22</td>
</tr>
</tbody>
</table>

AUC indicates area under the curve; LR, likelihood ratio; and ROC, receiver-operating characteristic.

*\( \chi^2 \) test for comparing AUCs between combination and each variable individually.
predictor variable. The 6MWT and CGS measure similar but slightly different constructs. Because of the distance requirements for community walking, the 6MWT seems to be a better predictor of home and community walking activity than CGS. Fulk et al. also found that although both the 6MWT and CGS were strongly related to community walking activity, only the 6MWT was a significant predictor of community walking activity. Our findings also indicate that the 6MWT has a better balance between sensitivity and specificity than CGS.

A CGS ≥ 0.49 m/s discriminated between home and community ambulators, whereas a CGS ≥ 0.93 m/s discriminated between limited community and unlimited community ambulators. These values are faster than the CGS values proposed by Perry et al. for distinguishing between home (<0.42 m/s), limited community (0.42–0.80 m/s), and unlimited community functional walking categories (>0.80 m/s). The functional walking category CGS values developed by Perry et al. may overestimate actual walking activity. The CGS cutoff of 0.93 m/s found in our study for discriminating between limited and unlimited community walking ambulators based on real-world stepping identified in this study is similar to the CGS value (0.97 m/s) that can discriminate among older adults who walk >8000 steps/d.66

Walking in one’s home and community is meaningful activity for people poststroke. However, it is a complex behavior involving individual, environmental, and contextual factors that make it challenging to measure. In addition to walking capacity, balance, motor function, and other factors such as self-efficacy and psychosocial factors likely play a role in community mobility, which we did not take into account in our model. Another limitation of our study is that it was a secondary analysis of data collected from 2 different studies with different purposes. The bootstrap method develops a best fit model by randomly sampling subjects from both data sets (LEAPS and FASTEST) so that variables derived from the models are based on 500 random sample selections of the 2 data sets. When the selected model generated from both data sets was applied to each data set separately to predict home versus community ambulators, the results were similar to each other (LEAPS data set AUC=0.836; cutoff=6.24; and overall accuracy of 77%; FASTEST data set AUC=0.793; cutoff=6.16; and overall accuracy of 77%) and the overall model generated from both data sets (AUC=0.836; cut off=6.81; and overall accuracy of 76%). The similar trends in these data sets when examined together and separately coupled with the bootstrap methodology support combining the data sets to develop the overall models. Combining these 2 data sets provided daily walking activity from a larger number of people poststroke than have been used in other studies and enhances the external validity of the findings. Stepping activity in the LEAPS trial was only collected over a 2-day period, which may not be sufficiently stable to characterize habitual walking activity.41 However, the ICC2,1 value of 0.79 indicates good agreement between the stepping activity on the 2 days. Although the daily stepping activity from the SAM provided objective data, we cannot determine where participants were walking and their purpose. Future studies that combine activity monitors, the global positioning system, and self-report may provide further insight into home and community walking behavior and shed further light on the barriers to community ambulation in people with stroke.

Conclusions

Walking endurance, motor function, and balance play an important role in home and community walking activity.
poststroke. Rehabilitation interventions that target these areas may be beneficial for people with stroke in order to improve their ability to walk in their community. Although CGS can predict home and community ambulators, cutoff values commonly used to discriminate between home and community ambulators may overestimate actual walking activity.

Disclosures

None.

References

Predicting Home and Community Walking Activity Poststroke
George D. Fulk, Ying He, Pierce Boyne and Kari Dunning

Stroke. 2017;48:406-411; originally published online January 5, 2017;
doi: 10.1161/STROKEAHA.116.015309
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2017 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://stroke.ahajournals.org/content/48/2/406

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org//subscriptions/