Modified Emory Functional Ambulation Profile
An Outcome Measure for the Rehabilitation of Poststroke Gait Dysfunction

Heather R. Baer, MD; Steven L. Wolf, PhD, PT, FAPTA

Background and Purpose—The modified Emory Functional Ambulation Profile (mEFAP) is an easily administered test that measures the time to ambulate through 5 common environmental terrains with or without an assistive device or manual assistance. The mEFAP was evaluated for its interrater reliability, test-retest reliability, concurrent validity, and sensitivity to change during outpatient rehabilitation for poststroke gait dysfunction.

Methods—Twenty-six poststroke patients were followed up prospectively in a rehabilitation day-treatment program. The mEFAP, Berg Balance Test (BBT), and 7-item mobility subsection of the Functional Independence Measure + Functional Assessment Measure (FAMm) were completed at admission and discharge.

Results—mEFAP interrater reliability (intraclass coefficient [ICC] 0.999) and test-retest reliability (ICC 0.998) were high. The BBT demonstrated high interrater (ICC 0.992) but poor test-retest (ICC 0.605) reliability. Initial and final scores comparing the mEFAP with the BBT ($r = -0.735, r = -0.703$) and the mEFAP with the FAMm ($r = 0.685, r = -0.775$) were strongly correlated. Improvement on the mEFAP correlated with improved BBT performance ($r = -0.524$). There was no correlation between overall change observed on the FAMm and change on the mEFAP ($r = -0.145$). Total mEFAP and all mEFAP subtask scores improved over time ($P < 0.0001$).

Conclusions—The mEFAP is a reliable gait-assessment tool for patients with stroke and is sensitive to change in ambulation speed. (Stroke. 2001;32:973-979.)

Key Words: gait ■ movement disorders ■ rehabilitation ■ stroke outcome

Gait dysfunction is a particularly prevalent and important consequence of stroke for which task-specific functional outcome measurement tools are needed.1 Currently available gait-assessment tools range from complex and expensive laboratory techniques involving detailed analyses of kinematic and kinetic variables to simple measures used in the clinical setting, such as the timed 10-meter walk.2–7 High cost and problems with accessibility make laboratory gait analysis impractical for widespread clinical use.5,8 Clinically based measures have the possible advantages of immediate availability, ease of administration, ease of interpretation, and low cost.5

The Functional Ambulation Profile (FAP) is a timed walking test that was specifically designed to track the progress of patients with neurological impairments throughout their participation in a comprehensive outpatient rehabilitation program.9 In its most recent form, this test (the Emory Functional Ambulation Profile, or EFAP) requires an individual to negotiate 5 common environmental challenges, and it incorporates the use of orthotics or assistive devices (AD).5,9 Wolf et al5 demonstrated that the EFAP had high interrater reliability, construct validity, and concurrent validity in a group of 28 chronic poststroke patients.

The present study extends the assessment of the EFAP by incorporating the important component of manual assistance. The main objective of this study was to determine whether a modified version of the EFAP, which we call the mEFAP, is sensitive to changes that occur in the gait of poststroke patients during outpatient rehabilitation. We also established its interrater and test-retest reliability. The concurrent validity of the mEFAP was assessed by establishing a correlation between the mEFAP and 2 other measures that are often used in the assessment of ambulation for poststroke patients in our clinic: the mobility subsection of the Functional Independence Measure + Functional Assessment Measure (FIM+FAM) and the Berg Balance Test (BBT).

Materials and Methods

Subjects
Data were collected on 27 subjects on entry into the Rehabilitation Day Program (RDP) at the Center for Rehabilitation Medicine at Emory University. Complete data were obtained for 26 subjects (13 men and 13 women). One subject could not return for final testing and was excluded from the analyses. Subjects were 54.5±12.7 years old (range 34 to 81 years). At the time of enrollment, subjects were 32.2±13.7 (range 16 to 66) days poststroke. Inclusion criteria were...
first documented stroke, stroke onset within the past 2 years, newly admitted to the RDP, in need of rehabilitation for poststroke gait dysfunction, and at least 18 years of age. Subjects had to comprehend all elements of the testing protocol with no more than moderate underlying dysfunction than use of a narrow-based quad cane. The underlying dysfunction that mandates their use. For example, use of a straight cane is considered to reflect less perceived severity of the gait-related dysfunction than use of a narrow-based quad cane. The underlying dysfunction predominated in this subject pool.

**Measurements and Instrumentation**

The mEFAP comprises 5 individually timed tasks performed over different environmental terrains. The subtasks include (1) a 5-meter walk on a hard floor; (2) a 5-meter walk on a carpeted surface; (3) rising from a chair, a 3-meter walk, and return to a seated position (the "timed up-and-go" test); (4) traversing a standardized obstacle course; and (5) ascending and descending 5 stairs. The mEFAP is performed with or without the use of an orthotic device or an AD. Manual assistance (MA) is provided as necessary. The subject can perform the task, with or without the use of an orthotic device or an AD. Assistance (I/F) is provided as necessary. The subject can perform at least 50% to 75% of the test elements and required assistance for the remaining 25% to 50% with or without the use of an AD). Exclusion criteria included any major medical contraindication to participation, lack of approval from the managing physician, or active infection. The Human Investigations Committee at Emory University approved this protocol. All subjects gave full informed consent.

The majority of subjects had right-sided brain lesions (n=18, 70%). Left-sided and bilateral lesions were equally represented (n=4, 15% for each group). There were more subjects with cortically based lesions (n=11, 42%) than with subcortical (n=9, 35%), brain stem (n=3, 11%), mixed cortical-subcortical (n=2, 8%), or mixed subcortical-brain stem (n=1, 4%) lesions. Ischemic strokes (n=20, 77%) predominated in this subject pool.

**TABLE 1. mEFAP Data Sheet**

<table>
<thead>
<tr>
<th>Time, s</th>
<th>Floor</th>
<th>Carpet</th>
<th>Up &amp; Go</th>
<th>Obstacles</th>
<th>Stairs</th>
<th>Total Time, s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD/Orthotic (I/F)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFO</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KAFO</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Straight cane</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBQC</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBQC</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemiwalker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assistance (I/F)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified independ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervision</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact guard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimal assist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate assist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I indicates initial; F, final; AFO, ankle-foot orthotic; KAFO, knee-ankle-foot orthotic; NBQC, narrow-based quad cane; LBQC, large-based quad cane; Independ, completely independent; Modified independ, requires extra time or devices; Supervision, requires cueing or prompting; Contact guard, requires close contact support; Minimal assist, performs at least 75% of task; and Moderate assist, performs 50–74% of task.

The BBT is a 14-item assessment tool initially developed to help identify geriatric subjects at risk for falls. More recently, the BBT scores of poststroke patients have been shown to be strongly associated with performance indices on other validated measures of motor function and functional ambulation. Strong interrater reliability (intraclass coefficient [ICC] 0.98), test-retest reliability (ICC=0.97), and responsiveness to change have been observed when the BBT has been administered to poststroke subjects. The 14 items on the BBT are ordered according to increasing difficulty. Performance on each item is ranked on an ordinal scale from 0 to 4 (0 reflecting the need for assistance to even minimally perform the requirements of the task and 4 reflecting independence in maximal performance of the task), with a maximal total score of 56 points. A score of less than 45 predicts the need for an AD or supervision during ambulation.

The FIM + FAM is the primary outcome measure presently used in our RDP. The FIM + FAM, an expanded version of the FIM, has increased sensitivity to functional changes that are observed in patients with traumatic brain injury in the subacute and chronic stages of recovery. The FIM + FAM consists of 30 items that are each scored on an ordinal scale from 1 to 7 (1 reflecting complete dependence and 7 affirming total independence). The majority of the 12 FAM items address cognitive and psychological functioning; however, 2 of the FAM items assess mobility. The 7-item mobility subscale of the FIM + FAM (the FAMm) includes assessments of transfers, walking and/or wheelchair locomotion, stair climbing, and community mobility. When administered as a series of clinical vignettes, a group of trained raters demonstrated agreement of 86% to 92% (κ=0.85) on FAMm items and 89% (κ=0.87) on the FIM items. Specific testing of the FIM + FAM in a poststroke population has not been reported. The maximal FIM + FAM score is 210. The maximal FAMm subscore is 49. Only the FAMm data from the FIM + FAM were assessed in the present study.

**Procedure**

All testing took place in the same setting. The mEFAP and the BBT were administered and graded by members of the study team. Subjects underwent initial testing within 2 days of enrollment in and
discharge from the RDP. The mEFAP and the BBT were administered in random order during the same physical therapy session. Reliability assessments occurred during the initial testing sessions. Interrater reliability was assessed via simultaneous ratings of subject performance on 27% (n=7) of the initial administrations of the mEFAP. Interrater reliability for the BBT was assessed for 23% (n=6) of the subjects in the same fashion. Test-retest reliability was addressed by repeat testing during the same session with either the mEFAP or the BBT in 19% (n=5) and 15% (n=4) of the subjects, respectively. Adequate rest periods were provided between the test administrations. The FAMm was scored by the subjects’ assigned physical therapist (who was a member of the study team) and by the subjects’ occupational therapist (who was not part of the team). The FAMm scores obtained closest to the time of admission and discharge were included in the analyses. Each subject used the orthotic device or AD that was most recently recommended by the subjects’ treating physical therapist or physiatrist.

### Treatment
There was no controlled intervention. Each subject received pregait and gait-related therapy at the discretion of the treating physical therapist. Because the therapy time fluctuated daily, the total therapy time was divided by the duration of a typical session (45 minutes) to derive the number of sessions completed by each subject. The mean number of 45-minute sessions attended was 16.03±4.68 (range 8 to 25.88).

### Statistical Analyses
SAS 6.12 statistical software (SAS Institute Inc) was used for all analyses. The mean, median, SD, and range were determined for age, days since stroke onset, test scores, and number of sessions attended. Gender was evaluated by frequency of occurrence. All measured variables were analyzed to determine the possibility of nonrandom effects on the outcome data by t test, ANOVA, or correlations as appropriate. Interrater and test-retest reliability were assessed for the mEFAP and the BBT with the ICC. The mean and SD of the interrater and test-retest differences were calculated. With the exception of reliability testing, all analyses were conducted with the measurements obtained only by the primary observer. Continuous data were tested for normality with the Shapiro-Wilkes assessment. Because of a lack of normality in some of the timed mEFAP data, nonparametric statistics (eg, Spearman correlation coefficient, Wilcoxon rank sum, Friedman test, or Kruskal-Wallis test) were used for the remainder of the analyses. No attempt was made to normalize the data before the final analysis. Only the timed mEFAP data were used for the correlational analyses with the BBT and the FAMm. Nonparametric statistics were used to analyze the relationship between the use of ADs (eg, Kruskal-Wallis test and Wilcoxon 2-sample test) and the amount of required MA (eg, Spearman correlation coefficient and Wilcoxon rank sum) to the initial, final, and change scores on the 3 outcome measures. An α-level of P<.05 was selected as the minimum criterion for statistical significance on all tests.

There were no preexisting data from interventional studies to predict the magnitude of an effect size. Power analysis was based on the range of subject performance data from a reliability and validity study of the EFAP (the predecessor to the mEFAP) in which poststroke subjects and age-matched healthy controls were tested with the EFAP one time. Review of those data indicated that a sample size of 25 subjects would be required in the present study to detect a change in initial-to-final mEFAP scores of at least 40% to yield a statistical power of 0.9. Post hoc power analysis was performed to assess mEFAP effect size and power. The power analysis was done with nonparametric adjustments. With an SD of change on the mEFAP (n=26, SD=43). There was a power of 0.9 to detect a change of 24 seconds (61.29±37.35 seconds change observed in the present study). For all correlational analyses, there was a power of 0.9 to detect a correlation of 0.59.

### Results

#### Normality
Initial and final average mEFAP, BBT, and FAMm group scores are shown in Table 2. Although the initial time scores on the total mEFAP were normally distributed, much of the initial, final, and change subtask data, as well as the final total mEFAP time scores, did not meet the criteria for normality. Lack of normality in the mEFAP final scores was due to a skewing of the data toward subjects with faster final times, and lack of normality on the change scores was caused by skewness toward subjects with unusually large amounts of improvement.

#### Reliability
Interrater reliability and test-retest reliability (Table 3) were high for the total (summed) mEFAP and all mEFAP subtasks (all ICCs ≥0.985). For the BBT, interrater reliability was high, but test-retest reliability was less favorable. The mean and SDs of the interrater and test-retest differences are presented in Table 3.

#### Sensitivity to Change
Table 4 shows data for the mEFAP final-minus-initial scores, along with the results of nonparametric (Wilcoxon) analyses. All subjects showed a significant decrease in the time taken to complete the total mEFAP (Wilcoxon signed rank, P<0.0001). A slight increase in time to completion was seen in 1 subject on the mEFAP floor subtask (0.97 seconds), 1 subject on the carpet subtask (0.93 seconds), 1 subject on the up-and-go subtask (1.3 seconds), and 2 subjects on the stairs (2.88 and 10.35 seconds). When initial mEFAP scores were compared with final mEFAP scores (Spearman correlation coefficient, r=0.885, P<0.0001) and absolute mEFAP change (r=0.688, P<0.0001), more substantial improve-

### Table 2. Initial and Final mEFAP, BBT, and FAMm Group Scores

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>mEFAP initial</td>
<td>168.80</td>
<td>71.50</td>
<td>63.07</td>
<td>301.43</td>
</tr>
<tr>
<td>mEFAP final</td>
<td>103.51</td>
<td>52.95</td>
<td>39.89</td>
<td>217.65</td>
</tr>
<tr>
<td>BBT initial</td>
<td>37.50</td>
<td>9.36</td>
<td>23</td>
<td>51</td>
</tr>
<tr>
<td>BBT final</td>
<td>47.15</td>
<td>7.78</td>
<td>34</td>
<td>56</td>
</tr>
<tr>
<td>FAMm initial</td>
<td>27.69</td>
<td>3.16</td>
<td>22</td>
<td>35</td>
</tr>
<tr>
<td>FAMm final</td>
<td>37.35</td>
<td>4.28</td>
<td>28</td>
<td>43</td>
</tr>
</tbody>
</table>

mEFAP scores are given in seconds.

### Table 3. Interrater and Test-Retest Reliability for the mEFAP and the BBT

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>ICC</th>
<th>X*</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>mEFAP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrater</td>
<td>7</td>
<td>0.999</td>
<td>0.094</td>
<td>1.189</td>
</tr>
<tr>
<td>Test-retest</td>
<td>5</td>
<td>0.998</td>
<td>1.792</td>
<td>5.437</td>
</tr>
<tr>
<td>BBT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interrater</td>
<td>6</td>
<td>0.992</td>
<td>0.167</td>
<td>1.169</td>
</tr>
<tr>
<td>Test-retest</td>
<td>4</td>
<td>0.605</td>
<td>−4.000</td>
<td>0.816</td>
</tr>
</tbody>
</table>

*Mean differences between raters or between sequential test administrations.
The initial time to completion of mEFAP (in seconds) compared with the change in mEFAP score (in seconds) from the initial to the final testing session. Subjects who took the longest time to complete the first testing demonstrated the greatest improvement in time to completion of the mEFAP ($r = 0.688$, $P < 0.0001$).

Spearman correlation coefficients were calculated for comparison of subject performance on the mEFAP with performance on the BBT (Table 4). The magnitude of absolute change on the FAMm did not correlate with the amount of overall change in total mEFAP scores ($r = -0.145$, $P < 0.479$). Change on the mEFAP carpet subtask correlated marginally ($r = -0.407$, $P < 0.039$) with change on the FAMm.

A weak correlation was observed between the BBT and the FAMm initial test scores ($r = 0.494$, $P < 0.010$). Final scores on the BBT and the FAMm were highly correlated ($r = 0.852$, $P < 0.0001$). No correlation was found between the absolute change in scores on the BBT and the FAMm over time ($r = 0.244$, $P < 0.231$).

Orthotic Device and AD Usage
Only 1 subject had a change in the use of an orthotic device (discontinued use of an ankle-foot orthotic). No subject required a transition to a more supportive AD (eg, from a straight cane to a narrow-based quad cane) during the study. Five subjects (19%) had an improvement (ie, required a less elaborate AD) of 1 rank, 5 (19%) changed by 2 ranks, and 16 (62%) had no change in the use of an AD. AD usage had no significant effect on the initial, final, or change scores on the mEFAP (Kruskal-Wallis test, $P < 0.311$), the FAMm (Kruskal-Wallis test, $P < 0.461$), or the BBT (Kruskal-Wallis test, $P < 0.420$). The absolute change scores on the outcome measures in the 16 subjects who improved by 2 or more AD ranks were not different from the remaining 10 subjects (Wilcoxon 2-sample test, $P < 0.133$ for mEFAP, $P < 0.413$ for BBT, and $P < 0.233$ for FAMm).

**Manual Assistance**
Two (8%) of the subjects had no improvement in the maximal amount of MA required during mEFAP testing. In contrast, 17 (65%) of the subjects improved (ie, required less MA) by similar association was observed between the absolute change in the group scores seen over time on the total mEFAP and the BBT ($r = -0.524$, $P < 0.006$). Change on the BBT was correlated with change on 3 of the 5 mEFAP subtasks (carpet, $r = -0.519$; get-up-and-go, $r = -0.449$; and obstacles, $r = -0.522$). Visual analysis of the change scores revealed 2 consistent outliers, both of whom appeared to have had more substantial changes on the BBT than the mEFAP. No unique features were identified that would justify removal of these subjects from post hoc analysis.

Nonparametric testing indicated a significant correlation between the mEFAP and FAMm scores before ($r = -0.685$, $P < 0.0001$) and after ($r = -0.775$, $P < 0.0001$) rehabilitation (Table 5). The magnitude of absolute change on the FAMm did not correlate with the amount of overall change in total mEFAP scores ($r = -0.479$). Change on the mEFAP carpet subtask correlated marginally ($r = -0.407$, $P < 0.039$) with change on the FAMm.

**Concurrent Validity: Comparison of Measures**
Spearman correlation coefficients were calculated for comparison of subject performance on the mEFAP with performance on the BBT (Table 5). The inverse relationship seen between the mEFAP and the BBT scores reflects the improved performance over time as a lower value on the BBT testing, but 5 (19%) subjects did after treatment. Most subjects demonstrated significant improvement on the mEFAP ($r = -0.735$, $P < 0.0001$) vs the BBT ($r = -0.524$, $P < 0.006$). Change on the BBT was correlated with change on 3 of the 5 mEFAP subtasks (carpet, $r = -0.519$; get-up-and-go, $r = -0.449$; and obstacles, $r = -0.522$).

Visual analysis of the change scores revealed 2 consistent outliers, both of whom appeared to have had more substantial changes on the BBT than the mEFAP. No unique features were identified that would justify removal of these subjects from post hoc analysis.

Nonparametric testing indicated a significant correlation between the mEFAP and FAMm scores before ($r = -0.685$, $P < 0.0001$) and after ($r = -0.775$, $P < 0.0001$) rehabilitation (Table 5). The magnitude of absolute change on the FAMm did not correlate with the amount of overall change in total mEFAP scores ($r = -0.479$). Change on the mEFAP carpet subtask correlated marginally ($r = -0.407$, $P < 0.039$) with change on the FAMm.

**TABLE 4. mEFAP, BBT, and FAMm Change Scores**

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Wilcoxon P</th>
</tr>
</thead>
<tbody>
<tr>
<td>mEFAP total</td>
<td>61.29</td>
<td>37.35</td>
<td>59.11</td>
<td>7.79</td>
<td>121.65</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Floor</td>
<td>6.89</td>
<td>5.86</td>
<td>5.125</td>
<td>-0.97</td>
<td>23.78</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Carpet</td>
<td>6.66</td>
<td>4.89</td>
<td>6.81</td>
<td>-0.93</td>
<td>16.57</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Up and go</td>
<td>14.23</td>
<td>11.09</td>
<td>11.06</td>
<td>-1.3</td>
<td>45.72</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Obstacles</td>
<td>21.57</td>
<td>16.81</td>
<td>17.83</td>
<td>3.38</td>
<td>62.4</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Stairs</td>
<td>11.96</td>
<td>10.38</td>
<td>11.1</td>
<td>-10.35</td>
<td>36.26</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>BBT</td>
<td>-9.65</td>
<td>7.43</td>
<td>-9.5</td>
<td>-28</td>
<td>1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FAMm</td>
<td>-9.65</td>
<td>4.03</td>
<td>-10</td>
<td>-18</td>
<td>0</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
1 level and 7 (27%) improved by 2 or more levels. Spearman correlational analysis revealed a correlation between change in the level of MA and change in the mEFAP floor ($r=0.528$, $P<0.006$) and carpet ($r=0.533$, $P<0.005$) subtasks but not with the other subtasks or with the total mEFAP change ($r=0.343$, $P<0.086$). A significant ($r=-0.553$, $P<0.003$) correlation was detected between change in MA and change in FAMm but not between change in MA and change in the BBT ($r=-0.348$, $P<0.08$). Subjects who demonstrated the most dramatic decreases in required MA (ie, improvement of 2 or more levels) had higher final BBT ($P<0.020$) and FAMm ($P<0.015$) scores and a greater amount of positive change on the both the BBT ($P<0.019$) and the FAMm ($P<0.008$, Wilcoxon rank sum). An improvement of 2 or more levels in MA was not associated with higher mEFAP change scores (Wilcoxon, $P<0.908$).

### Subject Variables

Gender did not significantly influence performance on the mEFAP, BBT, or FAM (Wilcoxon 2-sample test, $P>0.05$). Subject age was correlated with final scores on the BBT ($r=-0.580$, $P<0.002$), FAMm change scores ($r=0.576$, $P<0.002$), FAMm final scores ($r=-0.494$, $P<0.010$), and change in MA required ($r=-0.565$, $P<0.003$) by nonparametric analysis. No correlation was found between subject age and mEFAP scores ($P>0.05$). Visual analysis of the data (scattergrams) revealed a tendency toward better functional performance on these measures in younger subjects. Differences in the number of days between stroke onset and initial testing were not associated with differential performance on any of the initial, final, or change measures by Spearman correlations ($P>0.05$).

### Number of Sessions

There was no significant relationship between the number of sessions attended and change in performance on the FAMm ($r=-0.174$, $P=0.395$), the BBT ($r=-0.159$, $P=0.437$), or the mEFAP ($r=0.048$, $P=0.815$). Grouping of the mEFAP scores according to the number of completed sessions (ie, Wilcoxon signed ranking of $<14$ and $\geq14$ sessions and $<18$ and $\geq18$ sessions) did not reveal any statistically significant effect of session number on mEFAP change scores ($P>0.813$ and $P>0.306$).

### Discussion

The mEFAP was evaluated for its interrater reliability, test-retest reliability, sensitivity to change, and concurrent validity in the assessment of a heterogeneous group of subjects undergoing outpatient rehabilitation for poststroke gait dysfunction.

### Normality

In spite of the continuous nature of the timed mEFAP data, a lack of normality was found in the distribution of many of the mEFAP scores. Skewness of the data appeared to account for the lack of normality in all instances. Given the range in the severity of impairments and gait dysfunction in this subject pool, it is not surprising that we found striking differences in the distribution of performance on the 5 subtasks and differences in the magnitude of change in gait speed among the subjects.

### Reliability

Interrater and test-retest reliabilities for the timed portion of the mEFAP were high and were consistent with previous findings for similarly constructed scales.\(^4,^5,^7\) The relatively poor test-retest reliability of the BBT seen in the present study has not been reported previously. Berg et al\(^14\) found a considerably higher test-retest reliability (ICC 0.97) for 6 poststroke subjects when measured by the same observer twice, 1 week apart. In the present study, a practice effect may have been responsible for some of the improvements that were seen on intrasession repeat testing. Indeed, all 4 of the subjects who were tested by the BBT twice during the first session had higher scores on the second administration (within that session) of the test. It is difficult to draw firm conclusions regarding the test-retest reliability of the BBT in this population because of the small sample size.

### Sensitivity to Change

A few subjects demonstrated no change, or a decrement in performance, on 1 of the 3 final measures. A poorer performance on only 1 of the measures likely reflects both the differences in content of the 3 tools and the intersubject differences in impairments, disabilities, and comorbidities. Differences in the magnitude of an individual subject’s change scores between the mEFAP and the BBT, or the mEFAP and the FAMm, support the task-specific nature of the mEFAP. Additionally, subjects demonstrated substantial variability in the magnitude of change on the mEFAP subtasks compared with change on the summed mEFAP, which indicates that the individual subtasks may each offer unique information and argues against the inclination to remove any of the tasks because they may appear redundant.

### Concurrent Validity: Comparison of Measures

The strong agreement in initial-to-initial and final-to-final comparisons between the mEFAP and BBT and between the mEFAP and FAMm scores supports the concurrent validity...
of the mEFAP. Gait-related impairments and disability, when present, often improve during the first 3 months after stroke. The overall consistency of improvement, both within and between the tests, suggests that these changes are related to spontaneous or therapy-enhanced functional recovery and do not represent an artifact of repeated testing.

The significant correlation between initial, final, and change scores on the mEFAP and the BBT may be explained by an overlap in the fundamental abilities tested by these measures (particularly balance and strength). Poststroke subject performance on the BBT has shown similarly strong correlation with performance on other measures of functional ambulation (the EFAP and the Barthel Mobility Subscale), balance (the Fugl-Meyer Balance Subscale), and motor function (the Fugl-Meyer Arm and Leg Subscales). Individual differences in baseline impairments or comorbidities may contribute to differences in the magnitude of change observed between the BBT and the mEFAP scores.

The correlation between the mEFAP and FAMm initial-to-final and final-to-final scores suggests that improvement in subject performance can be detected with either scale. Lack of correlation between mEFAP and FAMm change scores may reflect a difference in the tasks evaluated (content), the difficulty in comparing the FAMm (an ordinal scale) with the mEFAP (a continuous scale), or the heterogeneity of impairments displayed by these poststroke subjects.

**ADs and MA**
The inclusion of subjects who required more MA reflects a shift in our patient population toward a higher severity of impairments and disabilities since the original scale was developed and supports the decision to adapt the mEFAP to account for greater MA required by these subacute poststroke subjects. In the present study, 12 subjects needed minimal assistance (ie, the subject required manual support for up to 25% of the physical work) on at least 1 of the subtasks at the time of initial testing. No subject needed greater external support (eg, AD or MA) at the final testing session, and most required less. Although no statistical relationship was observed between the absolute amount of mEFAP change and changes in required support, the possibility still exists that subjects would have been faster or slower at final testing if they had been tested under the same support conditions on both the initial and final measurements. Post hoc analysis comparing subjects who improved by ≥2 ranks of AD use or ≥2 levels of MA with those with improvement of a lesser degree revealed no significant differences in the absolute amount of change on the mEFAP. This negative finding suggests that the mEFAP change scores were of sufficient magnitude not to be significantly altered by changes in AD or MA over time. The apparent independence of the AD and MA changes from the time changes on mEFAP supports the separation of these entities in the reporting system of the mEFAP.

**Subject Variables**
Improved functional outcome has been reported in younger stroke patients. An association between younger age and improved final performance on the BBT and FAMm, greater absolute change on the FAMm, and greater reduction in the amount of required manual support was observed in the present study. Why age was not an influential factor on mEFAP performance remains unclear. Successful completion of the BBT, the FAMm, and the mEFAP requires significant balance, strength, and complex motor planning capabilities. Age-related physiological changes, or the accumulation of more significant comorbidity with advancing age, may have greater impact on the tasks encountered on the BBT and FAMm relative to those encountered with the mEFAP. Alternatively, age in and of itself may not be a factor that influences performance on the specific tasks that make up the mEFAP.

One potentially important issue is the relationship between location of lesion and subject performance on the mEFAP. Parenthetically, we observed a trend toward poorer performance on the initial BBT in the small number of subjects with bilateral strokes, as well as lower initial FAMm scores in subjects with cortical and subcortical lesions. These data suggest that future evaluation of lesion type with a larger sample size is warranted.

**Number of Sessions**
In the present study, we were unable to detect an association between the number of physical therapy sessions attended and performance on any of the 3 outcome measures. This finding may be due in part to nonrandom events, such as the criteria for determining discharge and/or fiscal limitations set by external case reviewers.

**Ceiling Effects**
At present, discharge criteria from the comprehensive outpatient program are essentially set in accordance with the functional levels that are defined by the FIM+FAM. Walking ability, as measured on the FIM+FAM, reflects the distance achieved and the external support requirements. There is no mechanism for tracking objective changes specific to gait parameters, such as velocity. Ceiling effects are of particular concern when a tool that was designed for an acutely impaired group is subsequently adapted for use with patients who are at a more advanced stage of recovery. The developers of the FIM+FAM have elaborated on the problems of ceiling effects in their review of outcome data for their subjects with chronic traumatic brain injury. Although the maximal FAMm target score of 49 was never reached in the present subject pool, the discharge criteria for the RDP actually appear to be centered around the attainment of the level of modified independence (ie, an item score of 6), with an effective maximal attainable FAMm score of 42. On further analysis, 6 of the subjects had an FAMm score ≥42 at discharge, again suggesting a potential ceiling effect.

As anticipated with a continuous measure, the mEFAP did not demonstrate evidence of a ceiling effect. In contrast, 5 of the subjects in this study obtained the maximal BBT score of 56 at final testing, suggesting the presence of a ceiling effect with this measure. This finding is in contrast to the recent work of Wood-Dauphinee et al, who did not encounter evidence of a ceiling effect with the BBT when administered.
serially to poststroke individuals within the first 12 weeks after the event.

**Limitations and Future Directions**

The BBT and the FAMm may not have been the most suitable criterion measures for this study. Direct comparison of the mEFAP to a well-validated test of gait velocity (such as the timed 10-meter walk) may have strengthened the establishment of concurrent (as well as construct) validity. Previous comparison of the EFAP, the predecessor of the mEFAP, to the timed 10-meter walk revealed a strong association between the measures in the performance characteristics of both poststroke subjects and healthy controls.  

The mEFAP is sensitive to changes in time taken to complete challenging ambulation tasks. Comparison of subject performance on the mEFAP and on measures of community and household ambulation is now needed to investigate the relationship between improved speed on the mEFAP and functional ambulation in a real-world setting. Age-matched normal values, as well as the minimal mEFAP values associated with successful household and community ambulation, can be established to focus treatment efforts, predict caregiver burden, and plan for appropriate discharge disposition.

The mEFAP provides clear and specific functional information, has no apparent ceiling effect, and looks at an activity (walking) that is often a high priority for the patient. Clinical scales that are currently in use, such as the FIM, Barthel, and Rankin scales, may not provide appropriately detailed task-specific information and may be limited by ceiling effects in the assessment of functional ambulation for individuals after stroke. Further research will be required to determine whether the mEFAP will be sufficiently sensitive and specific as an outcome measure for defined therapeutic interventions.

Ambulation status after intensive rehabilitation may greatly affect a stroke survivor’s sense of self, ease of community reentry, and vocational prospects. Future trials should look at the relationship between mEFAP score changes and measures of stroke-specific quality of life, such as the Stroke Impact Scale.

Future investigations should also address the possible added benefit of serial measurements with this tool during rehabilitation. Performing repeated measurements might allow treatment planning to be altered when the patient does not follow an expected course of functional recovery.

With the economic pressures that are exerted on provision of care in acute and postacute stroke rehabilitation settings, functionally based assessments (such as the FIM and the FIM+FAM) have gained widespread support for setting rehabilitation goals, documenting change, and justifying the use of resources. These measures, however, may be better at describing the patient’s success in mastering adaptive strategies and functional substitutions than they are at reflecting more subtle changes in the performance of specific tasks. The mEFAP can provide task-specific feedback regarding the speed of ambulation across common environmental barriers and the need for external support over the course of post-stroke rehabilitation. The mEFAP is inexpensive, easily understood, and can be administered in approximately 20 minutes.

**Acknowledgments**

The authors thank the Center for Rehabilitation Medicine at Emory University for permitting access to patients from their Rehabilitation Day Program. We also thank Deborah Desamours, PT, and Lisa Gillig, PT, for their contributions to data collection and George Cotsonis for providing excellent biostatistical assistance.

**References**


Modified Emory Functional Ambulation Profile: An Outcome Measure for the Rehabilitation of Poststroke Gait Dysfunction
Heather R. Baer and Steven L. Wolf

*Stroke.* 2001;32:973-979
doi: 10.1161/01.STR.32.4.973
*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2001 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/32/4/973

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/