Reliability of the GAITRite® walkway system for the quantification of temporo-spatial parameters of gait in young and older people

Hylton B. Menz, Mark D. Latt, Anne Tiedemann, Marcella Mun San Kwan, Stephen R. Lord∗

Prince of Wales Medical Research Institute, University of New South Wales, High St, Randwick, Sydney NSW 2031, Australia

Received 12 January 2003; received in revised form 9 April 2003; accepted 26 May 2003

Abstract

The purpose of this study was to evaluate the test–retest reliability of an instrumented walkway system (the GAITRite® mat) for the measurement of temporal and spatial parameters of gait in young and older people. Thirty young subjects (12 males, 18 females) aged between 22 and 40 years (mean 28.5, S.D. 4.8) and 31 older subjects (13 males, 18 females) aged between 76 and 87 years (mean 80.8, S.D. 3.1) walked at a self-selected comfortable walking speed across the pressure-sensor mat three times and repeated the process approximately 2 weeks later. Intra-class correlation coefficients (ICC), coefficients of variation (CV) and 95% limits of agreement were then determined. For both groups of subjects, the reliability of walking speed, cadence and step length was excellent (ICCs between 0.82 and 0.92 and CVs between 1.4 and 3.5%). Base of support and toe in/out angles, although exhibiting high ICCs, were associated with higher CVs (8.3–17.7% in young subjects and 14.3–33.0% in older subjects). It is concluded that the GAITRite® mat exhibits excellent reliability for most temporo-spatial gait parameters in both young and older subjects, however, base of support and toe in/out angles need to viewed with some caution, particularly in older people.

© 2003 Elsevier B.V. All rights reserved.

Keywords: Locomotion; Ageing; Measurement

1. Introduction

Measurements of temporal and spatial parameters of gait are commonly used for the identification of gait disorders [1] and for the evaluation of therapeutic interventions such as exercise [2,3]. A variety of techniques have been used for this purpose, including simple visual observation [4–6], stopwatches [7,8] and paper walkways [9–11]. More recently, sophisticated gait analysis tools have become commercially available, including the in-shoe Clinical Stride Analyser® [12,13] and the GAITRite® mat [14–17]. Although the reliability of the Clinical Stride Analyser® is well established [12,13], there is limited information available pertaining to the reliability of the GAITRite® system.

The GAITRite® mat is a portable walkway embedded with pressure sensors that detect footfalls as the subject walks the length of the mat. The software enables the documentation of a wide range of temporo-spatial gait parameters, including walking speed, cadence, step length, base of support and foot placement angles. A number of studies have been performed to evaluate the validity of these measurements against existing techniques. A single case study by McDonough et al. [14] evaluated the concurrent validity of the GAITRite® mat against chalk footprints and a hand-held stopwatch. The results revealed that while there was good agreement for spatial variables between the GAITRite® mat and the chalk footprints, the association between the GAITRite® mat temporal parameters and timed measures with the stopwatch was somewhat lower. The authors attributed this poor association to the subjectivity involved in timing gait events with a stopwatch.

A similar investigation by Selby-Silverstein and Besser [15] compared the GAITRite® mat to powdered footprints and an in-shoe pressure measurement tool (the Parotec® system), and reported moderate correlations for temporal variables. Cutlip et al. [16] compared the GAITRite® mat to the Peak Performance Technologies Motus 3.1® system, and reported strong associations between the systems for temporospatial measurements.

∗ Corresponding author. Tel.: +61-2-9382-2721; fax: +61-2-9382-2722.
E-mail address: s.lord@unsw.edu.au (S.R. Lord).

0966-6362/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0966-6362(03)00068-7
all parameters, however, significant differences between the systems were found for the measurement of step length and stride velocity. Recently, Bilney et al. [17] reported very high correlations between the GAITRite® mat and the Clinical Stride Analyzer®, and concluded that the GAITRite® system is a valid tool for gait analysis. The reliability of the GAITRite® system has received only limited attention. Gertz et al. [18] reported good immediate retest and 2-week retest reliability (Intra-class correlation coefficients (ICCs)>0.75) for temporo-spatial measurements of normal adults and adults with Down’s syndrome, but noted that base of support measurements were more variable than other parameters in both groups. Bilney et al. [17] reported high intraclass correlation coefficients for normal and fast-walking measurements obtained from three repeat trials recorded on the same day in 25 healthy subjects aged 21–71 years, however, the measurements were more variable between trials when subjects walked at a slow speed. Although these results indicate that the GAITRite® mat provides both valid and reliable measurements of temporo-spatial gait parameters, the test-retest reliability of the system has not been evaluated in older people [17].

As ageing is associated with increased variability in gait patterns, particularly in elderly people who fall [19,20], it is necessary to establish whether measurements obtained on different days using the GAITRite® mat are repeatable to determine whether the technique can be used to measure the effects of interventions in this population. Therefore, the aim of this study was to evaluate the test–retest reliability of GAITRite® system in healthy young subjects in addition to a sample of community-dwelling older people by obtaining measurements on two separate occasions, approximately 2 weeks apart.

2. Methods

2.1. Subjects

A convenience sample of thirty healthy young subjects (12 males, 18 females) aged between 22 and 40 years (mean 28.5, S.D. 4.8) were recruited from the staff and student population of the Prince of Wales Medical Research Institute. All were free of musculoskeletal or neurological pathology. Thirty-one community-dwelling people (13 males, 18 females) aged between 76 and 87 (mean 80.8, S.D. 3.1) were recruited from an existing database developed for previous research studies by our group. Subject characteristics are shown in Table 1. The Human Studies Ethics Committee at the University of New South Wales gave approval for this study, and informed consent was obtained from all subjects prior to their participation.

2.2. Apparatus

Spatial and temporal parameters of gait were measured with the GAITRite® mat (CIR Systems Inc. Clifton, NJ 07012). The standard GAITRite® mat is a carpet 460 cm long with an active sensor area of 366 cm long and 61 cm wide. The active area contains 13824 pressure sensors arranged in a grid pattern with a spatial resolution of 1.27 cm and a sampling frequency of 80 Hz.

2.3. Procedure

Subjects were fitted with Oxford-style lace-up shoes of the appropriate size with a suede upper and nitrile rubber sole. Each subject was instructed to walk over the mat at his or her usual, comfortable walking speed. In order to exclude the first and last few steps of each trial, subjects started walking from a point 2 m in front of the mat and stopped at a point 2 m behind the mat. Three trials were recorded for each subject and the average used for subsequent analysis. The procedure was then repeated approximately 2 weeks later.

2.4. Statistical analysis

Data for the young and older subjects were examined independently to compare the reliability of the GAITRite® system in each age group. All data were explored for normality using the skewness statistic, and data were considered normally distributed if the skewness of the distribution was <1. Differences in gait parameters between the young and older subjects (using data obtained from the first measurement session) were determined using a series of t-tests. Level of significance was set at $\alpha=0.05$. For each age group, ICCs of the type (3,1) were then used to evaluate the test–retest reliability of each gait parameter between the two measurement sessions. Detailed explanations of the calculation of the ICC are provided elsewhere [21]. To interpret ICC values we used benchmarks suggested by Shrout and Fleiss [21] (>0.75 excellent reliability, 0.40–0.75 fair-to-good reliability and <0.40 poor reliability). Paired t-tests were performed to test for any systematic differences between sessions, and to determine the absolute between trial variability in scores, both coefficients of variation (CV) of method error [22] and 95% limits of agreement were calculated. The coefficient of variation of method error was calculated as follows. First, method error was calculated using the standard deviation of the difference scores ($S_d$) between test and retest:

$$ME = \frac{S_d}{\sqrt{2}}$$

Table 1

<table>
<thead>
<tr>
<th>Subject characteristics (mean±S.D.)</th>
<th>Young</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (cm)</td>
<td>171.8 (10.3)</td>
<td>164.9 (9.79)*</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>68.4 (17.0)</td>
<td>68.2 (14.0)</td>
</tr>
<tr>
<td>Body mass index</td>
<td>22.9 (3.9)</td>
<td>24.8 (3.4)</td>
</tr>
</tbody>
</table>

*P<0.05
This value reflects the amount of variation in the difference scores between the two trials. To convert ME into a percentage, the coefficient of variation is then calculated:

\[ CV_{ME} = \frac{2ME}{X_1 + X_2} \times 100 \]

The resultant value expresses the ratio of the standard deviation to the mean as a percentage [22].

The 95% limits of agreement statistics express the degree of error proportional to the mean in the units the measurement was taken in, and was calculated using the methods described by Bland and Altman [23].

3. Results

3.1. Differences between young and older subjects

The older subjects were significantly shorter than the younger subjects \((t_{50} = -2.65, P < 0.01)\), see Table 1. A comparison of gait parameters between young and older subjects is shown in Table 2. Older subjects walked with a significantly slower velocity \((t_{50} = -5.77, P < 0.003)\), and shorter step lengths for both right \((t_{50} = -7.18, P < 0.001)\) and left limbs \((t_{50} = -7.43, P < 0.001)\). However, there were no differences between young and older subjects for cadence, base of support or toe in/out angles.

3.2. Reliability

Mean values for each of the gait parameters were very similar across the two measurement sessions (Table 2). Paired t-tests revealed that there were no significant systematic differences in any of the gait parameters between the two trials.

Test-retest reliability data for each of the gait parameters are shown in Table 3. In the young subjects, the ICCs for all gait parameters were excellent (ranging from 0.83 to 0.94). CV were small for walking speed, cadence and step length (1.4–3.5%), but relatively large for base of support (14.3–15.2%) and toe in/out angles (24.4–33.0%). This was also reflected in the 95% limits of agreement.

4. Discussion

Reliability is a fundamental requirement for any gait analysis system to ensure that any observed differences in gait performance between testing sessions reflect real changes in locomotor function, rather than random or systematic error in the measurement technique. This study has shown that the GAITRite® mat generally provides highly reliable measurement of temporospatial gait parameters when young and older subjects are tested 2 weeks apart, and that the system can be used with confidence to measure the effects of interventions over different testing sessions in both young and elderly populations.

Although ICCs are generally considered to be more appropriate indicators of reliability than simple correlation coefficients (i.e. Pearson’s r and Spearman’s rho), a high ICC does not necessarily mean that a test has acceptable reliability. In situations where there is a large range of scores in the sample, a high ICC may be obtained, despite there being large within-subject differences in the actual scores between trials. To overcome this, some authors recommend the complementary use of an absolute, as opposed to relative, measure of reliability, such as the coefficient of variation [22] or limits of agreement [23]. The coefficient of variation expresses the percentage variation of subjects’ scores between trials, thereby providing a clinically useful indicator of an intervention that is not affected by the presence of a heterogenous sample. Thus, a clinical measurement that has a high ICC may not have sufficient reliability if the coefficient of variation is unacceptably large. The determination of what constitutes an “acceptable” coefficient of variation depends on what size difference the researcher wants to detect.
Test–retest reliability for each of the gait parameters in young and older subjects

<table>
<thead>
<tr>
<th></th>
<th>Young subjects</th>
<th></th>
<th>Older subjects</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC (95%CI)</td>
<td>CV (%)</td>
<td>Mean (95%LoA)</td>
<td>ICC (95%CI)</td>
</tr>
<tr>
<td>Walking speed (m/s)</td>
<td>0.88 (0.77–0.94)</td>
<td>1.8</td>
<td>1.44 (1.31–1.55)</td>
<td>0.91 (0.83–0.98)</td>
</tr>
<tr>
<td>Cadence (steps/min)</td>
<td>0.83 (0.67–0.91)</td>
<td>1.9</td>
<td>111.48 (102.34–117.77)</td>
<td>0.82 (0.66–0.91)</td>
</tr>
<tr>
<td>Step length, R (cm)</td>
<td>0.92 (0.84–0.96)</td>
<td>1.4</td>
<td>77.30 (73.02–82.81)</td>
<td>0.89 (0.78–0.94)</td>
</tr>
<tr>
<td>Base of support, R (cm)</td>
<td>0.87 (0.74–0.93)</td>
<td>8.3</td>
<td>8.49 (5.25–12.11)</td>
<td>0.56 (0.26–0.76)</td>
</tr>
<tr>
<td>Toe in/out, R (°)</td>
<td>0.86 (0.77–0.94)</td>
<td>17.7</td>
<td>7.03 (1.51–11.31)</td>
<td>0.71 (0.48–0.94)</td>
</tr>
</tbody>
</table>

ICC, intra-class correlation coefficient; CV, coefficient of variation; LoA, limits of agreement. NB: R, right limb; L, left limb.

With this in mind, the measurement of toe in/out angles and base of support appear to be exceptions to the generally excellent reliability of the GAITRite® mat. Although these two variables exhibited excellent ICCs in young subjects (0.85–0.94), the ICCs for the older subjects were somewhat lower (ranging from 0.49 to 0.82) and in both groups of subjects, these variables had the highest CV and broadest 95% limits of agreement. This result is similar to Gretz et al. [18], who reported that while most tempo-spatial parameters of gait had good re-test reliability in adults with and without Down’s syndrome, base of support measurements were somewhat less reliable.

The toe in/out angle is the angle between the line of progression and the line connecting the heel point to the forward point of the footprint. This angle represents the transverse plane orientation of the stance limb [24], and is thought to influence lower extremity motion during gait [25]. The base of support is the distance between the heels when walking, and is thought to be an important parameter in the maintenance of stability. Abnormal bases of support have been observed in older people [26–28], people with Parkinson’s disease [29], pregnant women [11] and children with Down’s syndrome [30]. As such, both these parameters are of considerable interest to researchers and clinicians. However, the reliability reported here for these parameters may not be sufficiently high to enable the confident detection of small but clinically important differences over time in older people. In particular, the coefficient of variation for toe in/out measurement indicates that this parameter can be expected to vary by up to 33% between testing sessions, meaning that only quite large changes in this measurement could be confidently detected with the system.

There are two explanations for the comparatively low reliability of toe in/out measurement. The most likely explanation relates to the spatial resolution of the system (1.27 cm). As the GAITRite® mat calculates toe in/out angle as the angle formed by the line of progression and the line connecting the heel point to the forward point of the footfall, the accuracy of this measure is strongly influenced by the accuracy of identifying three reference points from two different footprints (see Fig. 1). Although all subjects in this study wore standardised footwear with clearly defined, sturdy heels, triggering of peripheral sensors may have been quite variable, leading to relatively large errors.

The spatial resolution of the GAITRite® mat may also explain the variability between trials for the measurement of base of support. This parameter is measured as the perpendicular distance from the centre of the heel of one footfall to the line of progression of the opposite foot (which is itself determined by the centre heel points of two contralateral footfalls) and as such is also very sensitive to variations in...
sensor detection between measurement sessions (see Fig. 1). Given that the width of the heel section of the shoes and the base of support were both approximately 8 cm, it can be easily appreciated that the occasional failure to detect one or two 1.27 cm wide sensors from a footfall could result in quite large differences between trials. Although the ICC for this measurement was excellent in young subjects, the CV and 95% limits of agreement were relatively large, indicating that the system would only be capable of detecting large differences if used in an intervention study or to compare clinical groups. The spatial resolution would appear to be adequate for step length measurements, as the measured distance between the two footfalls is considerably larger (ca. 65 cm) than the base of support measurement. However, care would need to be taken when interpreting step length data in populations who walk with very small steps, such as children and people with a disability.

Although the spatial resolution of the GAITRite® mat may partly explain the lower reliability of toe-in/out angle and base of support measurement, it is also possible that these parameters are inherently variable. Besser et al. [31] reported that while recording six to eight strides was sufficient to obtain representative data in unimpaired adults (defined as 95% confidence intervals within 5% of the mean), base of support measurements were significantly more variable and did not meet the 5% criterion even after ten strides were recorded. This inherent variability is likely to be even greater in older people [32]. Indeed, gait patterns in older people who fall have been shown to exhibit greater variability in cadence [19,20] and step length [33,34] than those who do not fall, suggesting that older people with poor balance exhibit impaired motor control when walking. It is yet to be determined whether older people who fall also exhibit increased variability in toe-in/out angles or base of support. However, it is likely that inherent stride-to-stride variability in the gait patterns of the elderly is responsible for the slightly lower reliability of these parameters in older subjects. The findings reported here need to be viewed in light of the limitations of the study. Firstly, we only recorded three trials from each subject, using the standard length GAITRite® mat. The number of trials was selected to prevent the older subjects from becoming fatigued and to keep the duration of the assessment to a manageable level. However, it is acknowledged that taking an average of a greater number of trials may have improved further the reliability of measurements obtained with the system [31]. Secondly, due to time constraints it was not possible to allow any time for subjects to accommodate to the test footwear. However, we do not feel that this would have altered the results significantly, as all subjects reported that the shoes were very comfortable, and the range of half-sizes and soft suede upper of the shoes allowed them to be appropriately fitted to each subject. Thirdly, it was not possible to ensure that recordings were taken at the same time during the day across the two testing sessions, so diurnal variability may explain the lower reliability of some measures, such as toe-in/out angles and base of support. Finally, it is possible that actual changes in the gait patterns of our elderly subjects may have occurred during the 2-week interval between measurement sessions, thereby resulting in lower retest reliability of base of support and toe-in/out angles. However, as the other variables (velocity, cadence and step length) were very stable over time, it is unlikely that base of support and toe-in/out angles would change without also influencing velocity, cadence and step length.

5. Conclusion

This study has shown that the GAITRite® produces highly reliable measurements of gait parameters, with the possible exception of base of support and toe-in/out angle in older subjects, when measured on two occasions approximately 2 weeks apart. The system can be used with confidence to evaluate the effects of various interventions on walking speed, cadence and step length. The reliability of base of support and toe-in/out angle measurement, though lower than the other parameters, may be sufficient for some applications. However, clinicians or researchers requiring more precise measurement of base of support or toe-in/out angles in older people may need to use ink or powder foot-print techniques, as these methods, although cumbersome and time-consuming are not limited by the spatial resolution of walkway pressure sensors and may be more reliable.

Acknowledgements

This project was funded by the NHMRC Health Research Partnerships Grant: Prevention of Injuries in Older People (Grant No. 209799) and an equipment grant from the Perpetual Foundation of Australia.

References