Establishing the Reliability of Test Procedures in Order to Investigate the Effect of Penetrating Trunk Trauma and Prolonged Mechanical Ventilation on the Recovery of Adult Survivors

ABSTRACT: Gunshot wounds and/or stab wounds to the trunk are injuries commonly seen in South African hospitals. Such injuries often necessitate exploratory surgical intervention to identify and treat injuries to the internal organs. These patients are managed in the intensive care unit and often undergo prolonged mechanical ventilation with immobilization that often results in some degree of muscle dysfunction. In order to monitor the recovery of adult survivors over a 6-month period after hospital discharge, validation of the six-minute walk test and hand-held dynamometry between data collectors was necessary.

Methods: Data was collected for the six-minute walk test and hand-held dynamometry by three data collectors. Subsequently a correlation coefficient (Pearson product moment correlate) was calculated and one-way analysis of variance (ANOVA) performed to assess the degree of association between the data.

Results and Discussion: A strong positive correlation existed for data collected on the 6-minute walk test between data collectors 1, 2 and 3 with r ranging from 0.71 to 1 (p < 0.025 to p < 0.0005). The intra-class correlation (ICC) scores between data collectors ranged from 0.69 to 0.90 confirming this finding. A weak to fair correlation existed between data collectors for dynamometry measurements of the triceps and quadriceps muscles. Intra-class correlation scores ranged from 0.11 to 0.71.

Conclusion: A strongly positive inter-observer correlation existed for data sets for the six-minute walk test. A weak intra-observer correlation existed for data sets collected by data collectors for dynamometry however this improved with intervention. The degree of association between data sets for dynamometry should be assessed again during the main study.

KEY WORDS: RELIABILITY, 6-MINUTE WALK TEST, HAND-HELD DYNAMOMETRY.

INTRODUCTION
South Africa has a high incidence of violence and death due to unnatural causes. Unnatural causes include assault, motor vehicle accidents, suicide and hanging. Unspecified unnatural deaths accounted for a mortality of 15.6% amongst South African males over the period 1997 – 2001 and males were three times more likely to die of unspecified unnatural causes than females for the same time period (Statistics South Africa, 2002).

Gunshot wounds and/or stab wounds to the trunk are injuries commonly seen in South African hospitals. These injuries often necessitate exploratory surgical intervention to identify and treat injuries to the internal organs. These patients often require re-look laparotomies and return to theatre frequently for abdominal lavage and eventual wound closure (Wilson and Bender, 1996). During this period patients are managed in the intensive care unit and frequently undergo prolonged mechanical ventilation with immobilization.

The 2002 Brussels Roundtable Discussions amongst intensive care practitioners concluded that survivors of intensive care frequently (1) had poor functional capabilities, (2) had decreased quality of life, (3) did not return to work and (4) posed an increased burden and stress on families and informal caregivers. These factors led to increased economic costs for the patient, families...

We propose to monitor the recovery of adult survivors of penetrating trunk trauma and prolonged mechanical ventilation for 6-months after hospital discharge. Previous studies have shown that the six-minute walk test (6MWT) is a reliable, valid and reproducible tool to determine the functional abilities of patients with chronic cardiac failure (Demers et al 2001; O’Keefe et al 1998; Opasich et al 1998) and chronic respiratory disease (AACVPR, 2004; Steele, 1996; Butland et al 1982). We propose to utilise this test in this study to monitor the functional capacity of the above-mentioned adult patients in the follow up period.

Similarly, the hand-held dynamometer that can be used to measure strength of a variety of muscles over a prolonged period of time would be a suitable tool to monitor the functional capacity of these patients. Functional capacity will invariably be influenced by muscle weakness. Results from previous studies have shown that hand-held dynamometry has high intra-observer and inter-observer reliability for knee extensor strength (Bohannon, 2001; Roy, and Doherty, 2004), elbow flexor and extensor strength (Visser et al 2003; Burns et al 2005) and shoulder abductors and knee flexors (Visser et al 2003) when tested in various patient populations.

Prior to the execution of the study the inter-observer and intra-observer reliability must be established between the novice data collectors involved in the study in order to test the reliability of measurements taken during the main study with the 6MWT and hand-held dynamometer. The aim of this reliability study is therefore to establish the degree of association (a positive correlation or a negative correlation) between the data collected by the respective novice data collectors in a group of healthy subjects for the above outcome measurements.

**METHODS**

A study design was used in which the one-tailed hypotheses were:
- High inter-observer correlation exists for data collected during the 6-minute walk test among data collectors.
- High intra-observer correlation exists for data collected during hand-held dynamometry among data collectors.

Fifteen people of good health were approached to participate in the study and nine volunteered. The study was conducted on the University of the Witwatersrand Education Campus, Johannesburg, South Africa on two successive Friday afternoons in April 2005 (testing sessions were separated by one week). Inter-observer correlation was determined for the 6MWT as two data collectors will be used to obtain these measurements during the main study. Intra-observer correlation was determined for dynamometry measurements as one data collector will be used to obtain these measurements in the main study due to space limitations. This study was approved by the Human Research Ethics Committee (Medical) of the University of the Witwatersrand.

**Six-minute walk test**

Each volunteer was assessed on the six-minute walk test (6MWT). A standardized protocol was used to administer this test to the volunteers over a distance of 30 meters in the corridor of the Cardiac Rehabilitation Unit. The protocol consisted of the following instructions:

- Volunteers were instructed to walk the measured distance at their own pace while trying to cover as much distance as possible in the 6 minutes.
- The test was supervised and the remaining time was called out every 2 minutes.
- Volunteers were encouraged at standard intervals of 30 seconds in the form of “You are doing well”.
- At the end of the 6 minutes, the volunteer was told to stop and the distance covered was recorded.
- The number of rest stops and the duration of each rest stop were also recorded.

Each volunteer was given 30 minutes to rest before the next 6MWT procedure was performed. All data collectors had an opportunity to familiarize themselves with the procedure on one occasion prior to the testing sessions. Data collectors 1 and 3 tested each volunteer at the same time during the first testing session. Thereafter data collector 2 tested each volunteer individually. At the second testing session data collectors 2 and 3 tested each volunteer at the same time and thereafter data collector 1 tested each volunteer individually. To establish the inter-observer reliability between the data collectors the following measurements were documented on an outcome measure sheet:

- Heart rate was measured, prior to, directly after and two-minutes after completion of the test. This was done manually over the radial artery.
- Respiratory rate was measured, prior to, directly after and two-minutes after completion of the test, by observation of chest wall movement.
- These measurements were made with the volunteer sitting in an upright position.
- The distance covered by each volunteer was also recorded by multiplying the number of completed laps at the end of the test with 30. The additional meters covered during the final lap were added on and the total distance expressed in meters.

The data collected between all data collectors was then compared for correlation.

**Dynamometry**

Dynamometry measurements were made of the left triceps and right quadriceps muscle on each volunteer with the hand-held dynamometer [Wagner Instruments (SDK 4)]. These tests were performed with each volunteer in the following positions:

- Left triceps muscle strength was assessed with the volunteer in a prone position. The shoulder was placed at 90° abduction and the elbow hung over the edge of the plinth at 90° flexion to the horizontal plane with the forearm in a neutral position. The dynamometer was placed just above the radial and ulnar styloid processes perpendicular to the shaft of the radius.
- Right quadriceps muscle strength was assessed with the volunteer seated in an upright position on a plinth with the legs perpendicular to the floor and the knees flexed at 90°. The left foot was stabilized on a footstool.
The dynamometer was placed just above the malleoli perpendicular to the tibial crest. The “break” test technique was used. This consists of the application of adequate force through the dynamometer to overcome the volunteer’s muscular effort. The test is essentially a concentric muscle strength test. All data collectors familiarized themselves with the handheld dynamometer on one occasion prior to the testing sessions. Volunteers were assessed separately (apart from the other data collectors), three measurements were recorded and the mean calculated for each muscle group. Each volunteer had a 10-minute rest period between testing sessions. Verbal encouragement was provided to each volunteer during testing.

The data collected between all data collectors was then compared for correlation.

Statistical analysis
The data collected was of an interval/ratio nature. The association between the sets of data was calculated by using the Pearson product moment correlation coefficient. The Excel 2000 programme for Microsoft Windows XP was used for the abovementioned calculations. The one-way analysis of variance (ANOVA) was used to determine the intra-class correlation (ICC) between data collectors (inter-observer correlation) for both the 6-minute walk test and the dynamometry. The ANOVA calculations were performed with the STATA 8 statistical software package.

RESULTS
A total of nine volunteers participated in this study. Three were female and six male and ages ranged from 21 – 26 years. Data was collected twice for the nine volunteers on two separate occasions and the data analysed as described above. Pearson product moment correlation coefficient results for the 6-minute walk test for heart rate (HR), respiratory rate (RR) and distance walked are presented in Tables 1 and 2. The subjective nature of manual measurements of HR and RR could have contributed to the weaker correlation found between data collectors as a difference of 1 beat/minute counted over 15 seconds and 1 breath/minute counted over 30 seconds might have influenced data to some extent.

The ANOVA between data collectors 1 and 3 continued during the second test session with r ranging from 0.71 to 0.98 for 6 of 7 measurements taken (p < 0.025 to 0.0005, Table 2). The first test session revealed a strong positive correlation between data collectors 2 and 3 with r ranging from 0.60 to 0.92 for 5 of 7 measurements taken (p < 0.05 to 0.0005). Table 2 showed a stronger positive correlation between data collectors 2 and 3 with r ranging from 0.81 to 0.98 for 6 of 7 measurements taken (p < 0.005 to 0.0005). A high degree of association existed between the data collected for all data collectors on the 6MWT. The ICC scores in tables 3 and 4 showed a high inter-observer correlation between data collectors with ICC scores ranging from 0.69 to 0.90.

Pearson product moment correlation coefficient results for dynamometry measurements of the right quadriceps muscles and left triceps muscles of the nine volunteers are presented in Table 5. A strong positive correlation existed for right quadriceps and left triceps measurements between data collectors at the

<table>
<thead>
<tr>
<th>Data Collectors (Raters)</th>
<th>Pre HR</th>
<th>Post HR</th>
<th>Post HR</th>
<th>Pre RR</th>
<th>Post RR</th>
<th>Post RR</th>
<th>Distance walked</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2 mins)</td>
<td></td>
<td></td>
<td>(2 mins)</td>
<td></td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td>r = 0.85</td>
<td>r = 0.59</td>
<td>r = 0.80</td>
<td>r = 0.50</td>
<td>r = 0.52</td>
<td>r = 0.75</td>
<td>r = 0.94</td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.005</td>
<td>p &lt; 0.05</td>
<td>p &lt; 0.005</td>
<td>p = NS</td>
<td>p = NS</td>
<td>p &lt; 0.01</td>
<td>p &lt; 0.0005</td>
</tr>
<tr>
<td>1 &amp; 3</td>
<td>r = 0.91</td>
<td>r = 0.76</td>
<td>r = 0.84</td>
<td>r = 0.72</td>
<td>r = 0.92</td>
<td>r = 0.95</td>
<td>r = 0.98</td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.0005</td>
<td>p &lt; 0.01</td>
<td>p &lt; 0.005</td>
<td>p = 0.025</td>
<td>p &lt; 0.0005</td>
<td>p &lt; 0.0005</td>
<td>p &lt; 0.0005</td>
</tr>
<tr>
<td>2 &amp; 3</td>
<td>r = 0.85</td>
<td>r = 0.54</td>
<td>r = 0.85</td>
<td>r = 0.60</td>
<td>r = 0.55</td>
<td>r = 0.79</td>
<td>r = 0.92</td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.005</td>
<td>p = NS</td>
<td>p &lt; 0.005</td>
<td>p &lt; 0.05</td>
<td>p = NS</td>
<td>p &lt; 0.01</td>
<td>p &lt; 0.0005</td>
</tr>
</tbody>
</table>

[Pre HR = pre-test heart rate; Post HR = heart rate immediately after test; Post HR (2 mins) = heart rate after 2 minute rest period; Pre RR = respiratory rate prior to test; Post RR = respiratory rate immediately after test; Post RR (2 mins) = respiratory rate after 2 minute rest period; NS = not significant]
second testing session as $r$ ranged from 0.62 to 0.85 with the p-value < 0.05 to 0.005. However a weak correlation existed between data collectors for these measurements during the first testing session as $r$ ranged from 0.07 to 0.54 and p-values were mostly insignificant.

The ANOVA results for intra-observer reliability for dynamometry measurements are presented in Table 6.

ICC scores ranging from 0.59 to 0.71 demonstrated a moderate to high intra-observer reliability for left triceps measurements. ICC scores ranging from 0.11 to 0.43 demonstrated weak intra-observer reliability for right quadriceps measurements.

**DISCUSSION**

The Pearson product moment correlation coefficient is a parametric test and can be used for correlational study designs as long as the data is of an interval/ratio nature (Hicks, 2004). This design looks for the degree of association between sets of data (Hicks, 2004). A correlation coefficient close to +1 indicates a strong positive relationship, a value closer to −1 indicates a strong negative relationship. A value closer to 0 indicates a weak relationship between data sets (Hicks, 2004).

**Six minute walk test**

Each volunteer in this study was given a 30 minute rest period in-between the six minute walk test (6MWT) measurements to allow recovery time prior to the execution of the next test. This decision was supported by the research of Opasich et al in 1998 in which these researchers investigated the reproducibility of 6MWT test results in 233 patients by performing 2 tests on the same day and on 2 consecutive days. They showed that the data collected on the same day (provided that there was a 30-minute rest between tests) was equivalent to that collected over 2 consecutive days.

Demers and colleagues (2001) performed the 6MWT test on 60 patients.

**Table 2: Pearson product moment correlation coefficient results for 6-minute walk test – week 2 (df = 7).**

<table>
<thead>
<tr>
<th>Data Collectors (Raters)</th>
<th>Pre HR</th>
<th>Post HR (2 mins)</th>
<th>Pre RR</th>
<th>Post RR (2 mins)</th>
<th>Distance walked</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 2</td>
<td>$r = 0.21$</td>
<td>$r = 0.54$</td>
<td>$r = 0.93$</td>
<td>$r = 0.83$</td>
<td>$r = 0.98$</td>
</tr>
<tr>
<td></td>
<td>$p = NS$</td>
<td>$p = NS$</td>
<td>$p &lt; 0.0005$</td>
<td>$p &lt; 0.005$</td>
<td>$p &lt; 0.0005$</td>
</tr>
<tr>
<td>1 &amp; 3</td>
<td>$r = 0.35$</td>
<td>$r = 0.73$</td>
<td>$r = 0.85$</td>
<td>$r = 0.74$</td>
<td>$r = 0.98$</td>
</tr>
<tr>
<td></td>
<td>$p = NS$</td>
<td>$p &lt; 0.025$</td>
<td>$p &lt; 0.005$</td>
<td>$p &lt; 0.025$</td>
<td>$p &lt; 0.0005$</td>
</tr>
<tr>
<td>2 &amp; 3</td>
<td>$r = 0.80$</td>
<td>$r = 0.83$</td>
<td>$r = 0.83$</td>
<td>$r = 0.81$</td>
<td>$r = 1$</td>
</tr>
<tr>
<td></td>
<td>$p &lt; 0.005$</td>
<td>$p &lt; 0.005$</td>
<td>$p &lt; 0.005$</td>
<td>$p &lt; 0.005$</td>
<td>$p &lt; 0.0005$</td>
</tr>
</tbody>
</table>

[Pre HR = pre-test heart rate; Post HR = heart rate immediately after test; Post HR (2 mins) = heart rate after 2 minute rest period; Pre RR = respiratory rate prior to test; Post RR = respiratory rate immediately after test; Post RR (2 mins) = respiratory rate after 2 minute rest period; NS = not significant]

**Table 3: ANOVA results for inter-observer correlation for 6-minute walk test (n = 18).**

<table>
<thead>
<tr>
<th>Raters 1 &amp; 3</th>
<th>HRpre6</th>
<th>RRpre6</th>
<th>HRpost6</th>
<th>RRpost6</th>
<th>HR2mins</th>
<th>RR2mins</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC</td>
<td>0.85</td>
<td>0.70</td>
<td>0.74</td>
<td>0.90</td>
<td>0.84</td>
<td>0.90</td>
</tr>
<tr>
<td>CI</td>
<td>0.68 – 1.03</td>
<td>0.36 – 1.04</td>
<td>0.43 – 1.04</td>
<td>0.77 – 1.02</td>
<td>0.65 – 1.03</td>
<td>0.78 – 1.02</td>
</tr>
</tbody>
</table>

[HRpre6 = heart rate prior to 6-minute walk test (6MWT); RRpre6 = respiratory rate prior to 6MWT; HRpost6 = heart rate immediately after 6MWT; RRpost6 = respiratory rate immediately after 6MWT; HR2mins = heart rate after 2-minute rest period; RR2mins = respiratory rate after 2-minute rest period; ICC = intra-class correlation; CI = 95% confidence interval]

**Table 4: ANOVA results for inter-observer correlation for 6-minute walk test (n = 18).**

<table>
<thead>
<tr>
<th>Raters 2 &amp; 3</th>
<th>HRpre6</th>
<th>RRpre6</th>
<th>HRpost6</th>
<th>RRpost6</th>
<th>HR2mins</th>
<th>RR2mins</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC</td>
<td>0.75</td>
<td>0.79</td>
<td>0.84</td>
<td>0.69</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>CI</td>
<td>0.46 – 1.04</td>
<td>0.54 – 1.04</td>
<td>0.65 – 1.03</td>
<td>0.34 – 1.04</td>
<td>0.67 – 1.03</td>
<td>0.68 – 1.03</td>
</tr>
</tbody>
</table>

[HRpre6 = heart rate prior to 6-minute walk test (6MWT); RRpre6 = respiratory rate prior to 6MWT; HRpost6 = heart rate immediately after 6MWT; RRpost6 = respiratory rate immediately after 6MWT; HR2mins = heart rate after 2-minute rest period; RR2mins = respiratory rate after 2-minute rest period; ICC = intra-class correlation; CI = 95% confidence interval]
with stable heart failure 3 times during the course of their study and produced ICC scores ranging from 0.88 to 0.91. O’Keefe and colleagues (1998) tested 24 elderly heart failure patients in whom the ICC score for reproducibility was > 0.75. Pankoff et al (2000) tested the reliability of 6MWT in 26 people with fibromyalgia in whom ICC scores were 0.91 to 0.98. Although the current study was performed on healthy subjects, the intra-class correlation coefficient results were similar to those found in the above-mentioned studies.

Positive inter-observer correlations existed for data collected by all three data collectors during the performance of the 6MWT. The hypothesis that the 6MWT is reproducible is confirmed and that the high inter-observer reliability found between the data collectors for measurement of heart rate, respiratory rate and distance walked is likely to be similar for measurements of blood pressure, oxygen saturation and level of perceived exertion during the execution of the 6MWT in the subsequent study.

Dynamometry
Burns et al (2005) compared the make and break techniques for hand-held dynamometry on elbow flexor and extensor muscles of 19 people with upper limb weakness due to tetraplegia. They used two novice testers for their study and the testing sessions were separated by 10-minutes. The inter-observer and intra-observer reliability was strong with ICC scores exceeding 0.9 for both make and break tests. Roy et al (2004) investigated the reliability of hand-held dynamometry in the assessment of knee extensor strength after hip fracture in 16 subjects. These researchers showed high test-retest ICC scores of 0.91 for the fractured leg and 0.90 for the unfractured leg.

In our study the Pearson product moment correlation coefficient showed a weak correlation between data sets collected for left triceps and right quadriceps muscles at the first testing session in this study. The Pearson correlation coefficient showed a stronger positive correlation between data collected at the second test session. The ICC scores obtained for left triceps muscles ranged from 0.59 to 0.71, which constitute a fair degree of intra-observer correlation. However, the ICC scores for right quadriceps muscles ranged from 0.11 to 0.43, a low intra-observer correlation. On these grounds, the one-way hypothesis for dynamometry would have to be rejected. Possible explanations for poor reproducibility of this test within data collectors could be the difference in strength between the data collector and the volunteer. The volunteers were young, healthy and strong. It is possible that as a consequence some data collectors might not have been able to apply sufficient vertical resistance to the muscle contractions of some of the volunteers. Mulroy found that female examiners’ maximal push force with a hand-held dynamometer was not significantly different when measuring quadriceps strength in male and female patients but was only 60% and 40% of quadriceps strength force generated in a group of normal women and men (Mulroy et al 1997). The planned study will involve muscle testing of patients who suffered penetrating trunk trauma, prolonged mechanical ventilation and prolonged bed rest. This group of patients frequently suffer from muscle weakness for a prolonged time period after hospital discharge (Winkelman, 2004; Lewis, 2003; Bruton et al 2002; Convertino et al 1997). If the strength of the volunteers contributed to the discrepancies seen in the current study, these may not be present in the main study.

The time frame between the two testing sessions was a limitation to the study as this might have influenced the performance of the volunteers. If a day or two separated the two testing sessions, less variation in data might have been seen especially for the dynamometry measurements. Another limitation was the level of inexperience in using the hand-held dynamometer amongst the data collectors. In accordance with this it is interesting that a learning effect was seen between the results of the first and second testing sessions for dynamometry. It is possible that further experience with this device may dramatically

<table>
<thead>
<tr>
<th>Data Collectors</th>
<th>Week 1</th>
<th>Week 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quadriceps Right</td>
<td>Triceps Left</td>
<td>Quadriceps Right</td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td>r = 0.07</td>
<td>r = 0.48</td>
</tr>
<tr>
<td>p = NS</td>
<td>p = NS</td>
<td></td>
</tr>
<tr>
<td>r = 0.62</td>
<td>r = 0.64</td>
<td></td>
</tr>
<tr>
<td>p &lt; 0.05</td>
<td>p &lt; 0.05</td>
<td></td>
</tr>
<tr>
<td>1 &amp; 3</td>
<td>r = 0.48</td>
<td>r = 0.85</td>
</tr>
<tr>
<td>p = NS</td>
<td>p &lt; 0.005</td>
<td></td>
</tr>
<tr>
<td>r = 0.80</td>
<td>r = 0.85</td>
<td></td>
</tr>
<tr>
<td>p &lt; 0.005</td>
<td>p &lt; 0.005</td>
<td></td>
</tr>
<tr>
<td>2 &amp; 3</td>
<td>r = 0.30</td>
<td>r = 0.54</td>
</tr>
<tr>
<td>p = NS</td>
<td>p = NS</td>
<td></td>
</tr>
<tr>
<td>r = 0.29</td>
<td>r = 0.67</td>
<td></td>
</tr>
<tr>
<td>p = NS</td>
<td>p &lt; 0.025</td>
<td></td>
</tr>
</tbody>
</table>

[NS = not significant]

| ICC scores obtained for left triceps muscles ranged from 0.59 to 0.71, which constitute a fair degree of intra-observer correlation. However, the ICC scores for right quadriceps muscles ranged from 0.11 to 0.43, a low intra-observer correlation. On these grounds, the one-way hypothesis for dynamometry would have to be rejected. Possible explanations for poor reproducibility of this test within data collectors could be the difference in strength between the data collector and the volunteer. The volunteers were young, healthy and strong. It is possible that as a consequence some data collectors might not have been able to apply sufficient vertical resistance to the muscle contractions of some of the volunteers. Mulroy found that female examiners’ maximal push force with a hand-held dynamometer was not significantly different when measuring quadriceps strength in male and female patients but was only 60% and 40% of quadriceps strength force generated in a group of normal women and men (Mulroy et al 1997). The planned study will involve muscle testing of patients who suffered penetrating trunk trauma, prolonged mechanical ventilation and prolonged bed rest. This group of patients frequently suffer from muscle weakness for a prolonged time period after hospital discharge (Winkelman, 2004; Lewis, 2003; Bruton et al 2002; Convertino et al 1997). If the strength of the volunteers contributed to the discrepancies seen in the current study, these may not be present in the main study.

The time frame between the two testing sessions was a limitation to the study as this might have influenced the performance of the volunteers. If a day or two separated the two testing sessions, less variation in data might have been seen especially for the dynamometry measurements. Another limitation was the level of inexperience in using the hand-held dynamometer amongst the data collectors. In accordance with this it is interesting that a learning effect was seen between the results of the first and second testing sessions for dynamometry. It is possible that further experience with this device may dramatically
improve intra-observer reliability. This, however, would have to be investigated during the execution of the main study. Based on the above results, data collector 2 will be used for dynamometry measurements and data collectors 1 and 3 will be used for the 6MWT measurements during the main study.

CONCLUSION
A strong inter-observer reliability was established between data collectors for the 6MWT. A low to fair intra-observer correlation was established for the dynamometry measurements in this study which may improve over time. A possibility exists that the patients in the main study might find the prone position uncomfortable for triceps strength measurement postoperatively. If this is found, triceps muscle strength will be assessed in supine.

ACKNOWLEDGEMENTS
The researchers would like to thank the following people/organizations:
- Prof P Becker from the Medical Research Foundation, Pretoria, South Africa for the statistical analysis of the data.
- National Research Foundation of South Africa for funding this study through the Thuthuka Women in Research Programme.
- Medical Faculty Research Endowment Fund of the University of the Witwatersrand, Johannesburg, South Africa for funding this study.

DISCLAIMER
Any opinion, findings and conclusions or recommendations expressed in this article are those of the authors and therefore the NRF does not accept any liability in regard thereto.

REFERENCES
Winkelman C 2004 Inactivity and inflammation: Selected cytokines as biologic mediators in muscle dysfunction during critical illness. AACN Clinical Issues 15: 74 – 82