Reliability of Modified Iowa Level of Assistance Scale by patients with total knee joint replacement

Background.
High quality measurements are needed to judge the effectiveness of physical therapy. This pilot study determines the reliability of the Modified Iowa Level of Assistance Scale (MILAS) by patients with total knee joint replacement.

Aim of the study.
Aim of this study is to access reliability of the Modified Iowa Level of Assistance Scale. We hypothesized that this test will be close or equal to the reliability (0.86) of tests of total hip joint replacement.

Subjects and methods.
32 patients were observed at three different times during 5 different activities by 1 therapist and two observers. (5 therapists and 1 observing student) to establish the reliability.

Results. The ICC statistic supported moderate interested reliability (K = 0.804). The total score was responsive to 4 days of therapy postoperative.

Conclusion.
We can conclude that the five activities graded with the modified Iowa level of assistance scale appeared to have moderate to good inter-reliability. Reliability be higher with guidelines for every physical therapist and observer.

Keywords: Modified Iowa Level of Assistance Scale; MILAS; tests and measurements, total knee replacement.
Introduction

In hospital care we often talk about independency of patients with a total knee replacement. Unfortunately objective measurements are often not recorded. Before patients can be discharged from the hospital they have to be examined if they are able to function independently. The Barthel-index can look at the independency of patients, but this measurement list is not good enough for physical care, because it also looks at issues that are not treated by the physiotherapist. Therefore we have looked for another lists and we found the MILAS.

MILAS stands for Modified Iowa Levels of Assistance Scale and looks at the independency of patients in the hospital. The MILAS has five assessments: from lying to sit, from sit to lying, from sit to stand, walking en walking on stairs. The five functionally related activities assessed are: ability to get out of bed, ability to get back in to bed, ability to stand from the bed, ability to ambulate 4/ 6 m, ability to climb, up and down 10 steps. Each activity is graded on an ordinal scale from 0 to 6 for the level of assistance required. A five-point ordinal scale provides a weighting factor for the assistive device used during the assessment of individual activities.

Independence, which is counted as 0, indicates that the therapist could leave the room and the patient could safely perform the activity being assessed. Standby assistance, which is counted as 1, indicates that the therapist would not feel comfortable leaving the patient; the therapist, however, provides no physical assistance. Minimal assistance, counted as 2, indicates that the therapist provides one point of contact with the patient. Moderate assistance, counted as 3, indicates that the therapist applies two points of contact. Maximal assistance, counted as 4, indicates that the therapist or therapists are applying a total of three or more points of contact. An activity that is attempted but is not completed with maximal assistance indicates, counted as 5, that the patient has failed maximal assistance.
A patient who is not tested for reasons of safety, as determined by the therapist's judgment, is graded as not tested (6).

The MILAS is not used often by physiotherapists in the hospital. We found one article about the reliability of the MILAS on patients with a Total Hip Replacement. It is therefore not clear whether the MILAS is reliable on patients with a Total Knee Replacement. In the article we found that the MILAS has a high score for reliability, validity and responsiveness. All viewed variables scored high on where the reliability had a score 0.8 at the Kappa scale. In this pilot study we focus on reliability in people with a Total Knee replacement.

**Patients and Methods**

**Inclusion and exclusion criteria**

Inclusion criteria were between the age of 35 and 95 with a total knee joint replacement. And the patients had to be operated on Monday or Tuesday so the patients could be followed for the rest of the week. Exclusion criteria were patients multiple disorders, or operated on other days of the week then Monday and Tuesday.

**Patients**

32 patients participated in this study the group ranged in age from 42 to 88. 17 patients were female and 15 were male. The female subject ranged from 54 to 88. The male subject ranged from 42 to 85. This procedure served as a mechanism to assess the reliability during the acute phase of therapy following total knee joint replacement.

**Procedure**

All patients were 3 times assessed during their hospitalization. Eight physical therapists with a range of 1 to 15 years of experience and one physical therapist student where testers in this study. There was an one hour introduction before starting the study. Each assessment was done by 2 physical therapists and one physical therapist student. The initial assessment was done after the first day of surgery in the afternoon (T0), the second assessment was done after the second day of surgery in the morning (T1) and the third assessment was done after the third day of surgery in the afternoon(T2). Each subject was assessed for the assistance required from lying in bed to sit, from sit back to lying in bed, from sit to stands, 4/6m walking and climb the stairs.
Each assessment required a total time of 20 to 30 minutes. After each assessment the 2 therapists and the physical therapist student administrated their results individually on paper and turned upside down in a tray. Each result was then immediately administrated in SPSS.

Statistics

The nominal data are progressed by the Intra Correlation Coefficient (ICC). ICC describes how strongly units in the same group resemble each other. The ICC will be high when there is little variation between the scores given to each item by the raters, compared to, if all raters give the same, or similar scores to each of the items. The model we used was two-way random and for type absolute agreement. The reliability was measured three ways. One, the overall inter-reliability was measured. Two, each reliability was measured for every item separately. Three, the reliability was measured for T0, T1, T2 separately.

Results

The ICC outcome can be seen in table 1 and in table 2. In table 1 shows the outcome of the different parts of the MILAS in total.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The reliability of the different parts of the MILAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lying-sit</td>
<td>0.440</td>
</tr>
</tbody>
</table>

Table 2 shows the outcome of T0, T1 and T2. The total reliability of the MILAS with patients with a TKR is 0.804. Table 1 shows that from lying to sit (0.440) and from sit-lying (0.634) has a poor reliability. Walking (0.876) and walking on stairs (0.952) have an excellent reliability. From sit to standing (0.753) has a position between these two extremes. Table 2 shows that T0 (0.835) and T1 (0.799) have a high reliability while T2 has a surprising outcome of 0.548.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>The reliability of T0, T1 and T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0</td>
<td>0.835</td>
</tr>
<tr>
<td>T1</td>
<td>0.799</td>
</tr>
<tr>
<td>T2</td>
<td>0.548</td>
</tr>
</tbody>
</table>
Discussion

Our study showed that reliability was 0.804. This is not the reliability we expected at the start. There are some methodological reasons why our result was not as high as we expected. In literature we could find almost no research towards the reliability of the MILAS with a Total Knee replacement. Due to the lack of research, there are no guidelines among the physical therapists for evaluation of patients with a total knee replacement. This makes, that there is absolutely no consensus about how to fill in the MILAS.

T0, T1 and T2 had a big difference in reliability. Reliability at T0 was high because a lot of Items were by all observers scored as a 6. At T1 and T2 real scores were given. Summarizing, our study showed, that inter-observer variability was high. Scoring instruction of observers before the study would probably have improved variability and reliability. The same holds true for scoring individual items.

For T0, T1 and T2 not all the 32 patients couldn’t be measured. T0 wasn’t measured in 5 of the 32 patients because of pain or a femoral block. Therefore these patients started at T1. Associated with the short time for this study and the number of patients on the department we decided not to exclude these patients from our study. One of the patients had a bleeding of her surgical wound during measure moment T2, as a result the knee was bandaged again to stop the bleeding, this meant she couldn’t perform as well compared to her fellow patients.

Another reason was, that during the study there were three holidays. Associated with understaffing on these days and the number of patients that moment (1 or 2) we decided to skip these days of research. This means that there are 2 different patients that missed T1 and 2 different patients that missed T2 in this research. These patients stayed included in this study. During another holiday all patients were seen in the morning. Consequently T1 and T2 fused to getter on this day. This could not be handled differently as a result of the work hours of the physical therapists that day.

In the inclusion and exclusion criteria we haven’t described what we would do with patients that had 2 prostheses in 2 weeks time. These patients have been included in the study but there results weren’t as good as their fellow patients.

During week 7 of this study one of the physical therapists accidentally treated one of the patients without including the researchers. Consequently there has been decided to leave out this measurement moment. The fact that some patients got behind on their schedule had no influence on the inter-reliability
During this study we came to items we missed or though could be handled differently. The score 6 is not a proper description for every patient. There are patients that don’t need to do certain items on the list of the MILAS, like walking the stairs because of their home environment. If you would give these patients a 6 as a score. It would look if the patient performed less than he/she actually can. It would be better if items that don’t need attention would be described as a nt (not tested). Than the final score is not affected.

During this study it became clear that the MILAS missed one item. This item was from stand back to sit.

Finally, it is not even clear what to do if some activities are not done. Hereby the therapist could make the decision to use the score of the previous day, or the therapist could make the decision to enter a nt (not tested).

**Conclusion and summary**

This study included 32 patients in 8 weeks. In these 8 weeks the reliability of the modified Iowa level of assistance scale was examined. From here we can conclude that the five activities graded with the modified Iowa level of assistance scale appeared to have moderate to good inter-reliability. But the reliability could have scored higher if we made sharp cut guidelines for every physical therapist and observer.

Considering 32 patients were seen during this study and the results look positive, we recommend further investigation.
1. Reliability, Validity, and Responsiveness of Functional Tests in Patients with Total Joint Replacement, 1995


