4 Project N°: IST-2002-507424
Acronym: ALLADIN

Task 4.3.4
Results of the Standardised Outcome Measures (SOM)

List of Partners: Arteveldehogeschool (B)
        Language and Computing NV (B)
        Budapest University of Technology and Economics (HU)
        Univerza v Ljubljani, Fakulteta za Elektrotehniko (SL)
        Zenon SA, Robotics and Informatics (EL)
        University Campus Bio-Medico (I)
        Multitel ASBL (B)
        Trinity College Dublin (IRL)
        Országos Orvosi Rehabilitációs Intézet (HU)
        Scuola superiore di studi universitari e di perfezionamento
        Sant'Anna (I)

Document identifier:	Internal D.3.4
Version:	1
Date:	30/04/2007
Organisation:	University of Dublin, Trinity College
Deliverable:	Internal Deliverable 4.3.4
Milestone:	M39
Workpackage:	4
Sub Task:	4.3.4
Nature	Report
Dissemination:	Public
Authors:	Emma Stokes, AC Varghese, Caroline O'Connell, Eleanor O'Grady.

Table of Contents

Index of Tables

Index of Figures

1. Executive Summary

The objective of this document is to report the standardised outcome measure results of participants who completed at least eight weeks of the ALLADIN trial at TCD and NIMR.

The standardised outcome measures used in the course of the study were the Fugl-Meyer motor assessment (FM), the Motor Assessment Scale (MAS) and the Stroke Impact Scale (SIS). These outcome measures are commonly used in clinical research trials in the area of stroke and measure at the level of impairment, activity and limitation in participation of activities of daily living.

For the purpose of this report, 65 participant data sets were analysed at entry to the ALLADIN study, eight weeks after entry and at the end of the study (24 weeks after entry). For each of the three scales, results are reported not only as a total score, but also as an upper limb and a mobility score, as pertaining to the relevant subsections of the scales.

The overall results show strong correlations between scores on all three scales at the three time-points analysed. This finding is not surprising, as it has been reported previously that there is a correlation between measures that focus on impairment and activity (Hazard *et al.* 1994; Sullivan *et al.* 2000). The differences in scores from entry to end of study are statistically significant for all three scales, indicating that all participants demonstrated clinically meaningful improvements throughout the course of the study. The greatest difference in scores occurred within the first eight weeks. Participants also displayed improvements on three scales when outcomes were compared at eight and 24 weeks, however, these findings did not reach statistical significance. These findings are similar to other research findings (Kwakkel *et al.* 1999) and are positive in terms of the pattern of stroke recovery observed in the first six months following the stroke.

2. Introduction

The ALLADIN project focuses on the development of a user-friendly natural language based decision support software for neuro-rehabilitation, specifically in stroke. The project monitored and measured the recovery process of post stroke individuals through 24 weeks following stroke. Three different methods of measurement were employed for this purpose. Quantitative measure of isometric force/torque while the subject performed selected activities of daily living, audio recording of clinical description of the subject and three standardised outcome measures (SOMs). This reports the results and the analysis of the three measures used in this study.

2.1 Standardised Outcome Measures

The SOMs used in this study were – the Fugl-Meyer motor assessment (FM), the Motor Assessment Scale (MAS) and the Stroke Impact Scale (SIS).

2.1.1 Fugl-Meyer

The Lindmark adaptation of the Fugl-Meyer scale (Lindmark and Hamrin 1988) was chosen for this research project as it combines the examination of functional limitations with underlying impairments. This adaptation uses a similar scale as the original measure and its convergent validity has been considered through comparison with scores on the FM and significant relationship noted. Galdstone *et al.* (2002) suggested that expanding the grading system might have actually been beneficial in maximising the ability to detect change- which is the main purpose of the scale. The Lindmark adaptation (Lindmark and Hamrin 1988) has expanded the grading of motor function and balance to a four point scale and walking to a seven point scale. This scale thus has the advantage of the benefits of FM scale with additional benefit of sensitivity to change.

Objective scoring with a clear description for each point of the scale and well-established levels of inter and intra rater reliability makes this a very suitable choice of outcome assessment for specific conditions. The Lindmark adaptation of the Fugl-Meyer scale provides separate scored sections to test both impairment of upper and lower limbs and limitations of activities, specifically mobility and balance. By assessing both the affected and the non-affected sides, the instrument measures

functionality as well as disability. The scale gives a total description of the motor capacity of the individual and not just of the hemiparetic side.

The maximum total score on the scale is 440. For the purpose of this study, the upper limb score was taken from the section relating to ability to perform active movements in the upper extremity, for which the maximum available score is 54. Similarly, the mobility score was taken as the combined score of sections relating to mobility and balance, for which the maximum available score is 48.

2.1.2 Motor Assessment Scale

The Motor Assessment Scale (Carr *et al.* 1985) is a stroke specific performance-based measure. It comprises seven items. Each item is scored from 0-6 with a higher score indicating optimal motor function. Clear guidelines are provided to optimise reliability (Loewen and Anderson 1988). Predictive validity has been established. Arm function scores at one week and one month after stroke were very good predictors of functional arm movement at discharge (Dean and Mackey 1992).

The maximum total score on the MAS is 48. For the purpose of this study, the upper limb score was taken as the combined score of "upper arm function", "hand movements" and "advanced hand activities", for which the maximum available score is 18. Similarly, the mobility score was taken as the combined score of "supine to side lying onto intact side", "supine to sitting over side of bed", "balanced sitting", "sitting to standing" and "walking", for which the maximum available score is 30.

2.1.3 Stroke Impact Scale

The Stroke Impact Scale (Duncan *et al.* 1999) is a stroke specific outcome measure that is a comprehensive measure of health outcomes. It is a 59-item scale assessing eight domains: strength, hand function, activities of daily living (ADLs and IADLs), mobility, communication, emotion, memory/thinking and participation. The first four domains can be combined to produce a physical domain score, but the other four domains must be scored separately. Performance on this scale is self-reported according to the difficulty experience by the respondent. The items in each domain are ordered hierarchically from the least to most difficult based on clinical perception of difficulty and Rasch analysis (Duncan *et al.* 2003). The final question uses the

visual analogue scale to assess the person's global perception of the amount of recovery since the onset of stroke.

The total score on the scale is in the range 59-295. For the purpose of this study, the upper limb score was taken as the combined scores for items regarding the participants' ability to use the hand that was most affected by their stroke, for which the available range is 5-25. Similarly, the mobility score was taken as the combined score of items regarding the participants' ability to be mobile at home and in the community, for which the available range is 9-45.

2.2 Inter-rater Reliability of the MAS, FM and SIS

Confidence in the reliability of the assessment instrument and the raters must exist for researchers to draw valid conclusions from clinical studies (Loewen and Anderson 1988). Reliability refers to the consistency, reproducibility and repeatability of the instrument or measurement procedure. It is also a measure of the degree to which the outcome measure is free of random or variable errors. Inter-rater reliability is determined by the same group of subjects being measured at the same time by a number of raters. Alladin project had three clinical sites located in different countries in Europe and the assessors/raters were located at these locations. It was therefore important to test the inter rater reliability among these assessors.

2.2.1 Aim

The purpose of this reliability study was to examine the inter-rater reliability of the four individual assessors when performing the MAS and Fugl-Meyer (Lindmark adaptation) assessments. The study explored the extent to which results or scores obtained by different raters, on the same occasion, using the same measurement method correlate.

2.2.2 Methodology

Participants presenting with diagnosis of stroke, as defined by the WHO definition, were invited to participate in this study. The sample group comprised of two inpatients in the age related health care unit, and three outpatients who were receiving treatment in the *stroke service* in St. James's Hospital, Dublin (SJH). Following a

detailed explanation of the process, the participant information leaflet was given to the participant and informed consent was obtained.

An assessment area was created, with the required facilities as follows: a plinth, chair, table, along with the props required for the MAS and FM specifically. Four chairs were then assembled each with a clear view of the assessment area. An experienced clinician (physiotherapist with more than 25 years experience) administered the scale and four physiotherapists involved in ALLADIN research and who were administering the scales during the ALLADIN clinical trial independently scored on the scales.

2.2.3 Statistical Analysis

The authors examined the percentage level of agreement between the raters and used the weighted kappa coefficients to quantify level of agreement between the raters. Kappa is a measure of agreement between two different tests or raters. It relates to the amount of observed agreement beyond chance as a proportion of the potential agreement beyond chance. The score ranges from 0 for no agreement beyond chance to 1 for perfect agreement (Daly and Bourke 2000). Criteria for each section on both scales are provided to assist the examiners in reliably grading the performance of each item together with general rules for administering the MAS and FM.

2.2.4 Previous Reliability Studies on the scales

The inter-rater reliability of the MAS was assessed in two separate studies. The authors carried out the first study in 1985. Five participants at various stages of recovery were selected for inter-rater reliability testing. The assessments were videotaped, and a time was visible on the tape. Twenty physical therapists and twenty physical therapy students rated the scores of the participants. The Pearson Product Moment Correlation Coefficient, *r*, ranged from 0.89 – 0.99 (excluding general tonus)(Carr *et al.* 1985). The greater the numerics value of *r*, the stronger the relationship between the two variables (Daly and Bourke 2000). The average percentage agreement between the raters was 87% (range 78 – 95). In a second study by Poole and Whitney (1988), 24 individuals with stroke were assessed. The mean length of time post stroke was 12 months (range 0.5 – 96 months). Two examiners observed and scored each subject individually. Spearman Rank Correlation

Coefficient for the total MAS score was 0.99 and ranged from 0.92 to 1.0 for the individual items, excluding general tonus, which was 0.29.

Initial studies involving the Fugl-Meyer concentrated mainly on the construct validity of the scale. Since then, there have been two main studies in the area of inter-rater reliability. In the paper by Duncan *et al.* (1983), a multiple range test was used to establish where inter-rater means differed. They concluded that the upper and lower limb components of the FM had reliability in the range of *r=0.86-0.99.* The greater the numeric value of *r*, the stronger the relationship between the two variables (Daly and Bourke 2000). Sanford *et al.* (1993) used 'interclass correlation' as a measure of inter-rater reliability and deemed their results in a 95% confidence interval range of 0.91-0.99 to be comparable with the earlier study. These scores are both for the total FM scores generated following a reliability study. Sanford *et al.* (1993) fail to prove the reliability of the pain subscale, attaining reliability scores of only 0.61. All other ICC scores were between 0.85 and 0.97.

2.2.5 Results of the study

Table 2.1 and table 2.2 outline the average percentage agreement of the different subsections of the MAS and FM scores obtained from the study.

Table 2.1 MAS

Sections	Average % Agreement	Range % Agreement	Average weighted Kappa	Range weighted Kappa
Supine to lying on intact side	74	49 - 93	0.44	0.063 – 0.833
Supine to sitting over side of bed	96	92 - 100	0.97	0.933 - 1
Balanced Sitting	100	100	1	1
Sitting to Standing	100	100	1	1
Walking	100	100	1	1
Upper Arm Function	87	80 - 100	0.94	0.909 - 1
Hand Movements	83	65 - 100	0.95	0.917 - 1
Advanced Hand Activities	100	100	1	1
General Tonus	91	83 - 100	0.516	0 - 1
TOTAL MAS SCORES	96	94 - 99	0.91	0.84 – 0.957

Table 2.2 FM

NON PARETIC

	Avg % agr.	Range % agr.	Avg Kw	Range Kw
Active Mvt	97.2	95.5-99.2	0.804	0.740-0.934
Rapid Mvt Changes	91.4	84.1-96.0	0.501	0.046-0.902
Mobility	90.0	85.1-90.1	0.681	0.600-0.818
Balance	92.3	89.4-98.9	0.876	0.731-0.937
Sensation	91.1	86.1-95.8	0.271	0.076-0.777
Joint pain	96.0	94.3-98.9	0.321	0.118-0.839
Joint Motion	96.9	94.9-99.1	0.831	0.716-0.924
Totals	97.2	94.8-99.8	0.853	0.770-0.940

PARETIC

	Avg % agr.	Range % agr.	Avg Kw	Range Kw
Active Mvt	87.8	81.3-94.0	0.883	0.802-0.939
Rapid Mvt Changes	75.8	63.5-83.2	0.625	0.357-0.792
Mobility				
Balance				
Sensation	93.2	87.0-96.8	0.898	0.822-0.934
Joint pain	90.8	89.1-92.6	-0.085	-0.296-0.259
Joint Motion	90.4	86.7-95.8	0.571	0.377-0.874
Totals	92.1	87.5-94.8	0.798	0.693-0.891

2.2.6 Discussion

Four sections of the MAS demonstrated very good reliability, namely the balanced sitting, sitting to standing, walking and advanced hand activity sections (see table 1.1). Although, the four raters demonstrated excellent reliability in the advanced hand activity section, questions arose regarding the scoring hierarchy. In this section, to score on items 3 and 4, the participant was required to perform timed writing tasks. To obtain a score of five, the participant is required to bring a dessertspoon of liquid to the mouth without spilling. To obtain a maximum score of six, the participant must hold a comb and comb the hair at the back of the head. Although all participants were scored by MAS criteria, those who could perform the writing tasks in this section were also asked to perform the feeding and hair combing tasks. Generally, individuals can perform simpler tasks such as feeding and hair combing before they can perform more complex skills such as writing (Poole and Whitney 1988). One participant was unable to perform items 3 and 4 but performed the higher scoring hair combing task.

This finding is in agreement with that of Poole and Whitney (1988) suggesting that further evidence is needed to validate the order of advanced hand activities.

The first section – supine to lying on intact side, demonstrated the lowest level of agreement. In the case of a particular participant, who pulled himself into side lying in 3 seconds, one rater recorded a score of one for this section, while three raters recorded a score of six. A score of one indicates that the participant 'pulls self into side lying' whereas a score of six indicates that the participant can 'roll to side in 3 seconds'. A possible reason for this discrepancy is the lack of clarity in the scoring criteria. It is unclear whether the quality of the movement should be taken into account if the participant can perform the movement within 3 seconds.

Unlike previous studies, the average reliability of the general tonus section was moderate with an average percentage agreement of 91% (range 83 – 100). The average weighted Kappa coefficient was 0.516 (range 0 – 1). In the study by Carr *et al.* (1985), the category of general tonus was excluded as it was deemed too difficult to assess from videotape. Poole and Whitney (1988) reported the inter-rater reliability of the general tone section to be 0.29. This section was included in the study because the authors felt that one could determine tone from direct observation of a participant's performance. However, they concluded that the low correlation scores indicated that the evaluation of tone through observation cannot be replicated and is not accurate. They further suggested that some physical contact with a participant might be necessary for accurate assessment of tone. Therefore, in this study, the raters individually assessed tone in the upper limb in each participant by passively moving the elbow through range. Reliability may have been improved if specific instructions were provided to the raters for the general tone section.

In terms of the total FM scores obtained, percentage agreement scores show that there is good reliability between the raters. For the non-paretic limb the range of agreement was 94.8%-99.8%, average 97.2%. Percentage agreement indicates the percentage of the time that scores are repeatable by different raters. 97.2% therefore is in the range of good to very good agreement. This came with a weighted Kappa coefficient of 0.853. It has been suggested that a weighted Kappa coefficient of greater than 0.75 indicated excellent agreement beyond chance (Carr *et al.* 1985).

The reliability scores obtained varied from section to section, with the greatest percentage agreement being obtained in the sections 'active movement', 'joint motion' and total score. This correlates with the scores obtained by the other authors (Duncan *et al.* 1983; Sanford *et al.* 1993). It can be suggested that this reliability is due to the ease of observation of range of motion, even without hands on contact.

However, Sanford *et al.* (1993) failed to confirm the reliability of the 'Joint Pain' subsection, scoring r=0.61, in comparison with our score of 96% for the non-paretic side and 96.9 for the 90.8% for the paretic side. However the Weighted Kappa Co-efficients obtained were 0.321 and –0.085 respectively. This suggests that the reliability cannot be confirmed for this section, in fact the score of –0.085 indicates that the reliability obtained is less than what could be expected by chance. As mentioned earlier, this may be due to the choice of methodology. However, the Lindmark instruction for this section relies on patient reporting of pain, so perhaps more discussion into the interpretation of patients' pain report is needed.

Lower reliability scores were also obtained in the 'rapid movement changes' section. Percentage agreement was 91.4% (non-paretic side) and 75.8% (paretic side). 75.8% was the lowest percentage agreement in the study. Discussion with the raters suggests that there was a considerable amount of confusion over the definition of rapid movement changes – how fast was required, should quality of the movement be taken into consideration? The participants also tended to move quickly within the mid range, avoiding taking the movements to the outer ranges in favour of achieving a rapid movement.

An encouraging aspect of the study was that reliability can be ascertained for the total scores on both the paretic and non-paretic side, with weighted Kappa scores of 0.798 and 0.853 respectively, both of which are in the range that suggests very good agreement. This is an important conclusion, as previous studies in this area did not differentiate between the two sides.

2.2.7 Conclusion

The authors conclude that MAS and FM both provide reliable and quantifiable means of documenting motor abilities in terms of function. However, as the reliability of

each individual rater is considered an essential factor in the use of the Motor Assessment Scale, Carr *et al.* (1985) recommend physiotherapists intending to use the scale should become familiar with the criteria for scoring by testing at least six subjects before formally using it in clinical practice. The authors suggest this recommendation be extended to the Fugl-Meyer scale so as to improve efficacy of assessment.

2.3 Time-points of measurements and analysis

Participants in the ALLADIN trial were recruited within eight weeks following stroke and their progress followed for 24 weeks. The outcome measures were administered as illustrated in table 2.3.

The analysis includes commentary on the results at entry to the study, at eight weeks after entry and at the end of the study (24 weeks). At eight weeks, each participant will have completed four FMs, four MASs and two SISs.

Table 2.3 Sequence of outcome measurement

Week number	SOM
1	FM 1
2	MAS 1 & SIS 1
3	FM 2
4	MAS 2
5	FM 3
6	MAS 3
7	FM 4
8	MAS 4 & SIS 2
10	FM 5
12	MAS 5
14	FM 6
16	MAS 6
18	FM 7
20	MAS 7
22	FM 8
24	MAS 8 & SIS 3

2.4 Statistical analysis

The statistical analysis to follow is largely descriptive in nature. The analysis was completed using Microsoft Excel 2003 and Minitab statistical software (release 13.1). The tests included in this report are the Wilcoxon signed-rank test and Spearman Rank correlation co-efficient.

2.4.1 Wilcoxon Signed-Rank Test

The Wilcoxon signed-rank test is a non-parametric test used to compare paired observations. It is considered appropriate for use with non-parametric (ordinal) data from two related groups, for example, when the same individuals are measured in two different circumstances, such as before and after a given intervention (Petrie and Sabin 2000). In this report, the results of the Wilcoxon signed-rank test will be reported with the point estimate for the difference and p-value (at all times alpha=5%)

2.4.2 Spearman Rank Correlation Coefficient

The Spearman Rank correlation coefficient is a non-parametric measure of correlation. It is designed for use with ordinal data. The Spearman Rank determines independence between two variables. It is similar to the Pearson product, but data is ranked prior to analysis. The results of the Spearman rank correlation will be reported as the correlation coefficient (between 0 and 1, 1 indicating strong correlation) and the associated p-value.

Studies have suggested that the correlation between physical impairments and activity measures typically varies between 0.2 and 0.5 (Hazard *et al.* 1994; Sullivan *et al.* 2000).

2.4.3 Comparisons of SOMs

The standardised outcome measures in this report have very different score ranges. For this reason, the scores were normalised to the percentages of maximum available score wherever it was necessary, to compare across the scales. This technique was utilised only in the preparation of graphical representations of the data and in the commentary.

The actual scores were used for the statistical tests, the Wilcoxon signed-rank test and Spearman rank as described above.

2.4.4 Box plots

The box plots (also known as a 'box and whisker plot') used throughout the report were generated using Minitab statistical software (release 13.1). Data is presented in a single column consisting of a rectangular box, a horizontal line within the box, vertical lines and asterisks in some cases. The rectangular box represents the interquartile range, i.e. the middle 50% of the data. The upper and lower edges of the box indicate the 75th and 25th percentile of the data set respectively. The horizontal line within the box represents the median. The vertical lines or 'whiskers' extending either side of the box indicate the general extent of the data. The extreme points on the line indicate the minimum and maximum data values. In cases where there are outliers in the data, they are represented by asterisks.

Figure 2.1. Box plots
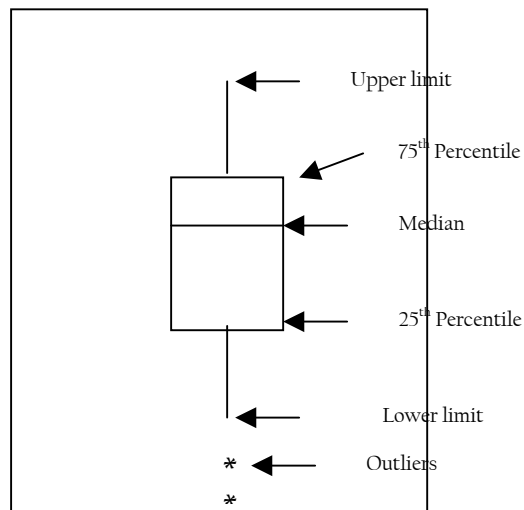
3. Demographic Details

There were 37 men (57%) and 28 women (43%) included in the study. The age ranged from 34-93 years, with a mean age 64 years.

There was an even distribution of left and right hemiparesis, 52% and 48% respectively. Participants presenting with hemiparesis at the dominant side accounted for 49%.

Many (25%) of the participants were working at the time of stroke, and most (45%) were married.

Table 3.1 contains a summary of the demographic details recorded on participants in the ALLADIN trial.

Table 3.1 Demographic Details

Gender	Male 57% (n=37)
	Female 43% (n=28)
Age	Mean 64.25 years, SD 12.03 years, Range 34-93 years
Hemiparesis	Right 47.69% (n=31)
	Left 52.31% (n=34)
Hemiparesis at Dominant Side	49.23% (n=32)
Time Since Stroke	Mean 16.125 days, SD 14.36 days, Range 0-57 days
Marital Status	Currently Married 44.62% (n=29)
	Never Married 9.23% (n=6)
	Separated 7.69% (n=5)
	Divorced 10.77% (n=7)
	Widowed 10.77% (n=7)
	Unspecified 16.92% (n=11)
Occupation	Working 24.62% (n=16)
	Retired 58.46% (n=38)
	Homemaker 1.54% (n=1)
	Unemployed (Health Reasons) 4.62% (n=3)
	Unspecified 10.77% (n=7)

## 4. Standardised Outcome Measure: results at entry

4.1 Scores at entry- total

Analysis of the total scores on the three scales at entry is reported in table 4.1 and the boxplot (figure 4.1) is a graphical illustration of this results and the relationship between the three scales.

Table 4.1. Scores at entry- total

| SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	65	377.43	388	55.35	201	440	330	426.5
MAS	65	33.65	36	12.34	12	48	21	45.5
SIS	56	198.96	199.5	41.84	116	289	159.5	230

Figure 4.1 Scores at entry- total
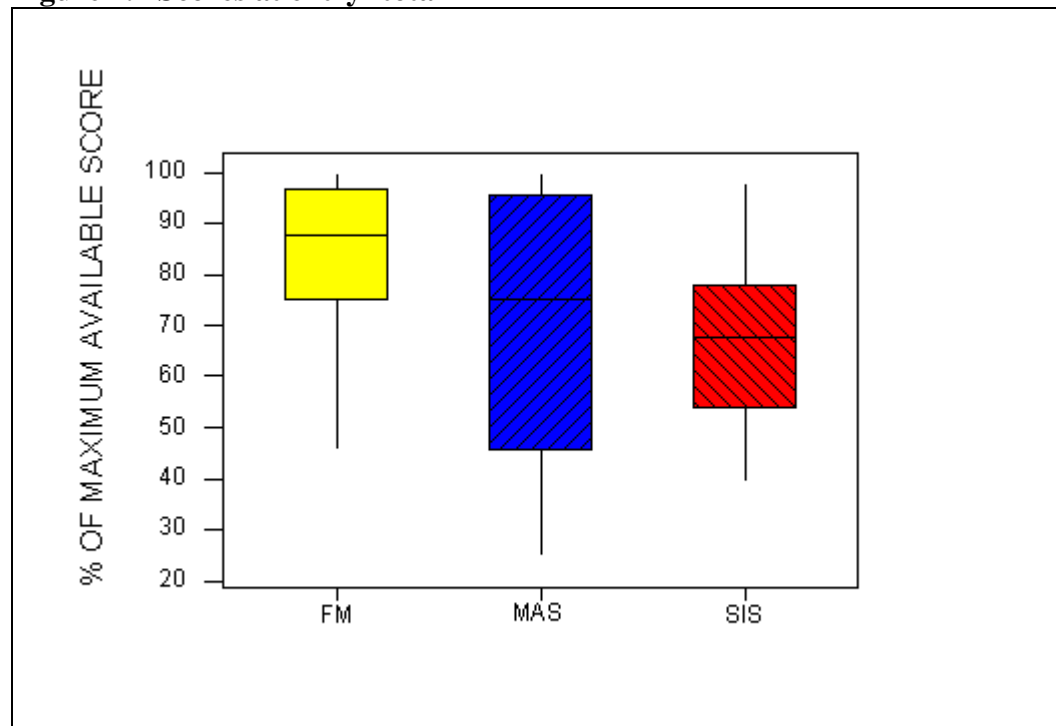
4.2 Scores at entry- upper limb

Analysis of the upper limb scores on the three scales at entry is reported in table 4.2, and the boxplot (figure 4.2) is a graphical illustration of this results and the relationship between the three scales.

Table 4.2 Scores at entry- upper limb

| SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	65	31.42	42	21.88	0	54	5.5	52.5
MAS	65	10.14	12	6.84	0	18	2.5	17
SIS	56	12.21	11.5	7.06	5	25	5	17

Figure 4.2 Scores at entry- upper limb
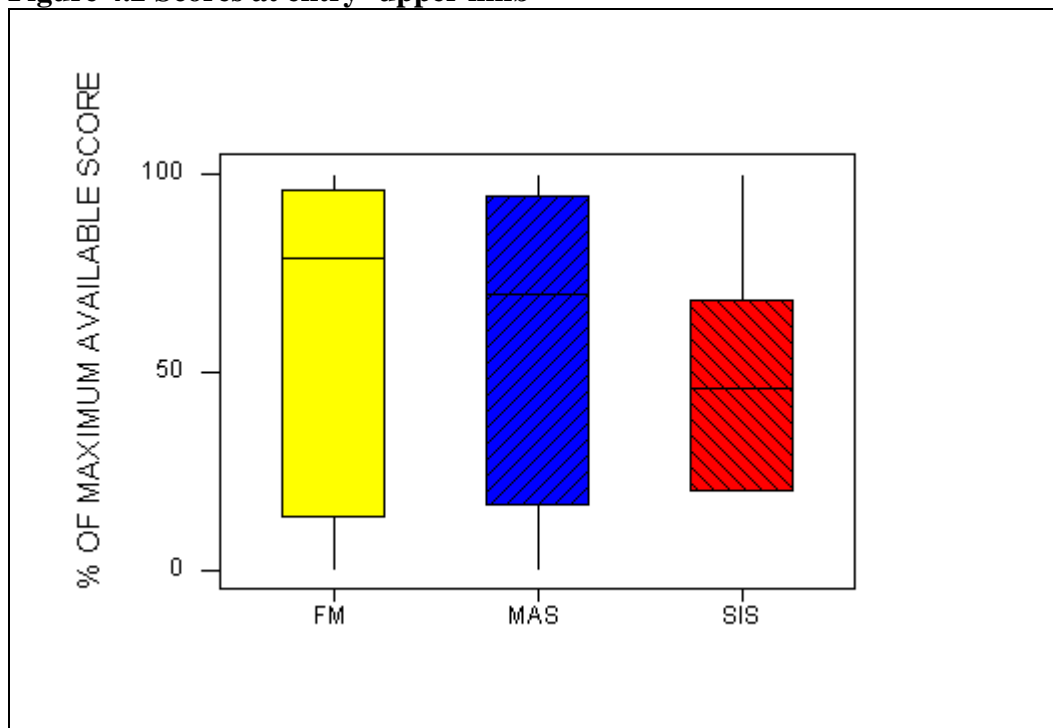
4.3 Scores at entry- mobility

Analysis of the mobility scores on the three scales at entry is reported in table 4.3 and the boxplot (figure 4.3) is a graphical illustration of this results and the relationship between the three scales.

Table 4.3 Scores at entry- mobility

| SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	65	35.23	37	12.43	8	48	25	47
MAS	65	23.51	27	7.135	3	30	18	30
SIS	56	28.82	28	11.83	10	45	18	39

Figure 4.3 Scores at entry- mobility
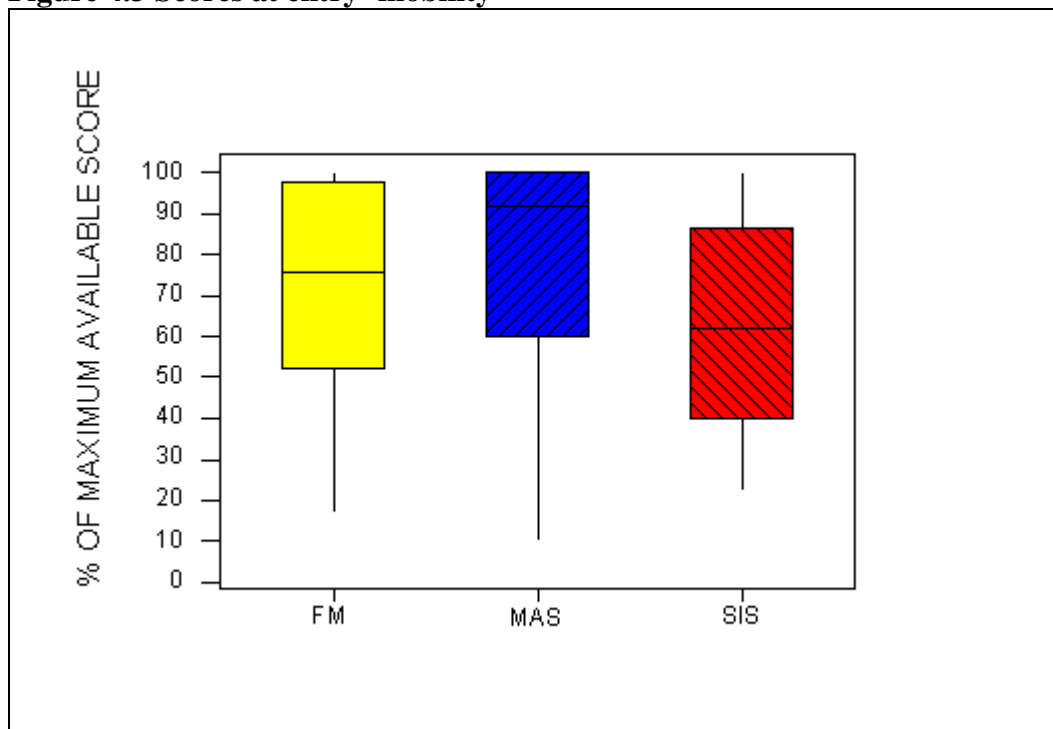
4.4 Summary of findings at entry

Overall, there was a strong correlation between scores on the three scales (see table 4.4). Figure 4.4 below illustrates the scores obtained in the three scales for total score, upper limb and mobility.

Figure 4.4 FM, MAS and SIS scores at entry
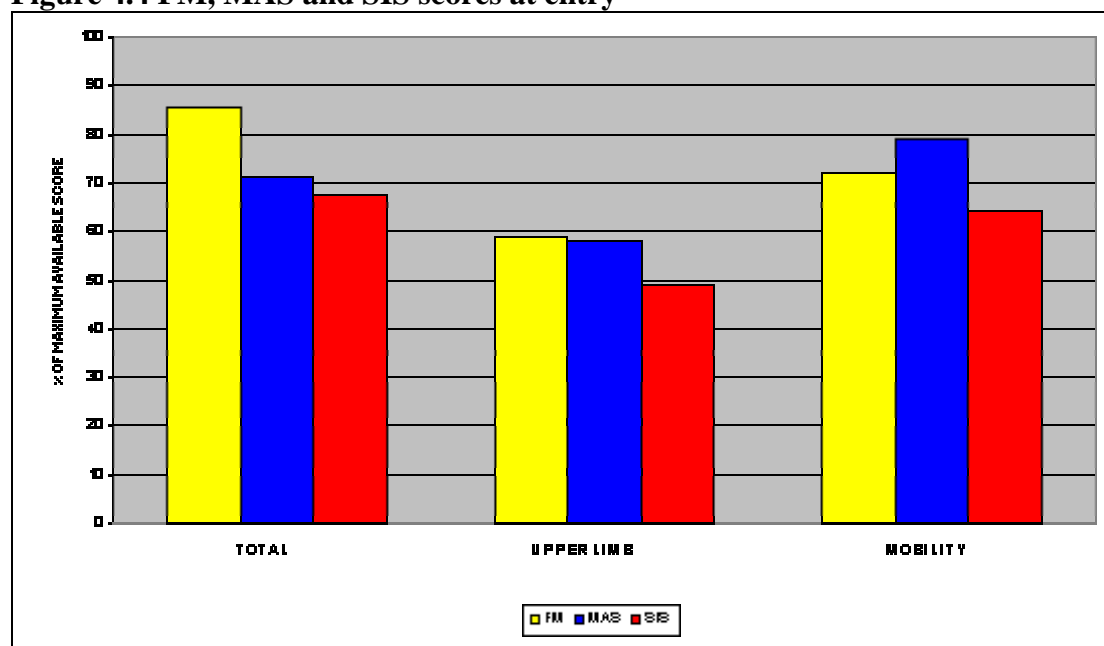
The results recorded at admission indicate that the participants in the trial had moderate disability. The upper limb was more severely affected in many cases, as indicated by the lower percentage scores on all scales. Two of the scales (FM and MAS) have a similar focus, physical ability. However, the FM measures at the level of impairment and the MAS measures at the level of limitation in activities. Despite this difference, both scales recorded similar levels of physical functioning. A correlation was also seen between FM, MAS and SIS, which measure at the level of limitation in activities and participation. The later is a self-reported measure, and it is interesting to note that despite the correlation between the three scales, the SIS scores are significantly lower for all dimensions analysed. This suggests that the perceived impact of the stroke may be greater than what the objective measure of function reveals. SIS takes into consideration the other effects of the stroke such as depression, fatigue and communication difficulties, whereas the FM and the MAS do not.

Table 4.4 Correlation of scores at entry

SOM	Spearman Rank Correlation		
	Total	Upper Limb	Mobility
FM & MAS	0.89	0.90	0.78
FM & SIS	0.82	0.78	0.71
MAS & SIS	0.77	0.79	0.58

5. Standardised Outcome Measure: results at eight weeks

A minimum requirement of eight weeks participation was sought from all participants. This represented two months of recovery post stroke.
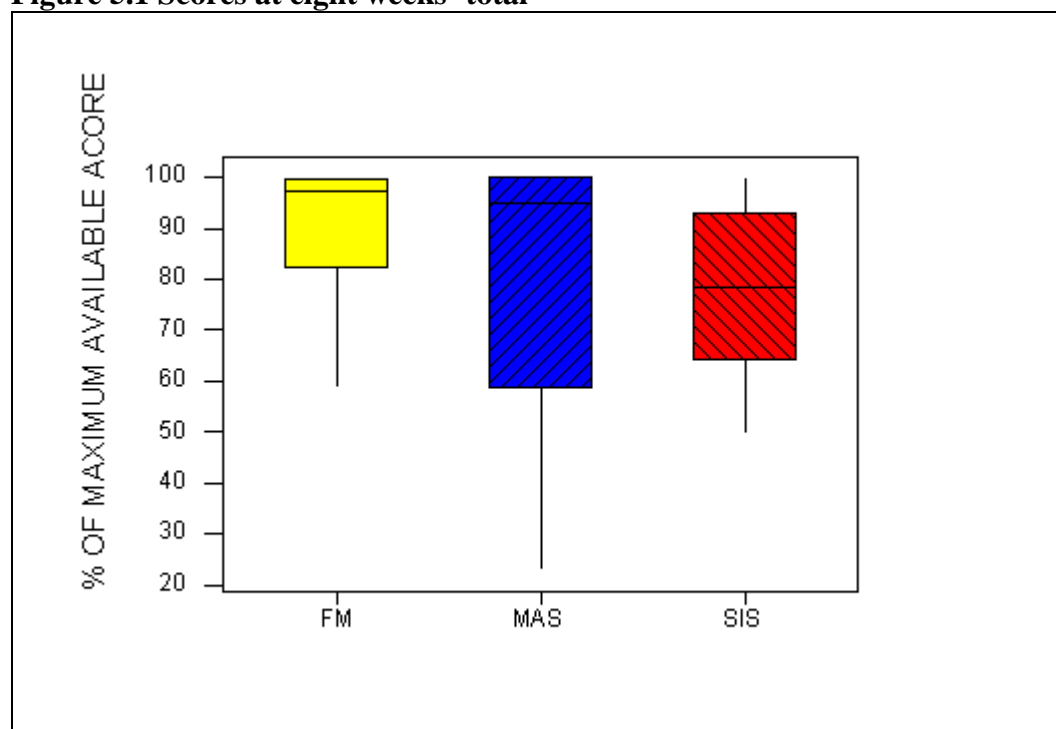
5.1 Scores at eight weeks- total

Analysis of the total scores on the three scales at eight weeks is reported in table 5.1 and the boxplot (figure 5.1) is a graphical illustration of this results and the relationship between the three scales.

Table 5.1 Scores at eight weeks- total

SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	65	402.78	428	43.99	258	440	377.5	437.5
MAS	65	40.08	46	9.93	11	48	32.5	48
SIS	52	229.62	231.5	43.16	147	295	189.75	274

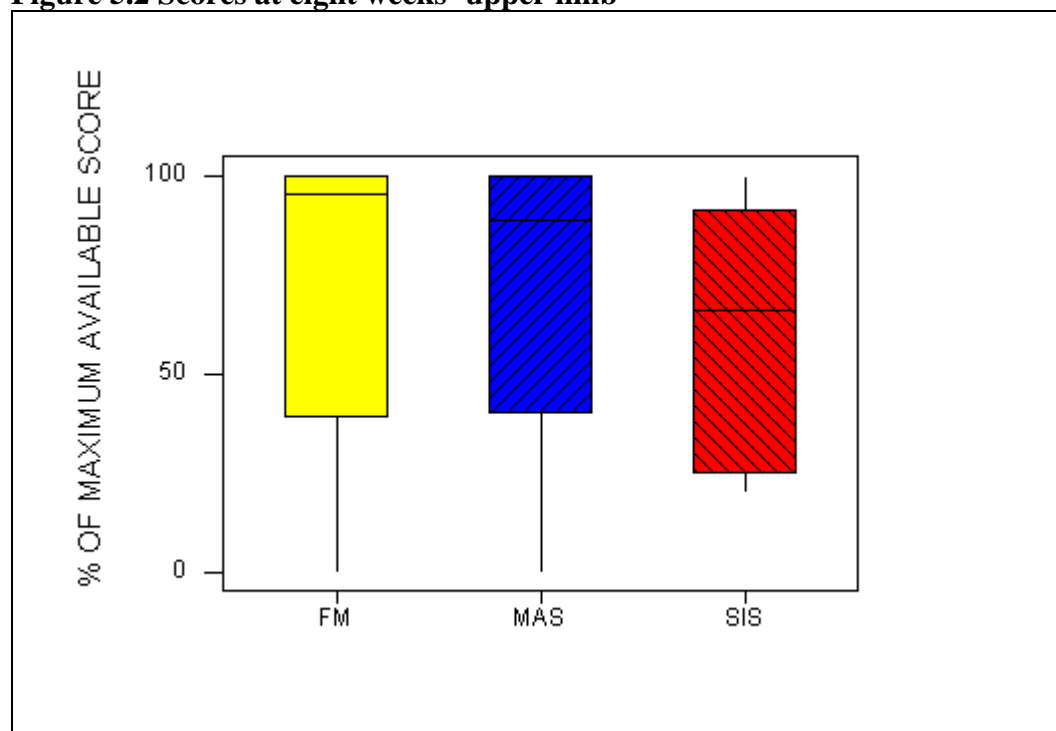**Figure 5.1 Scores at eight weeks- total**

5.2 Upper limb scores at eight weeks

Analysis of the upper limb scores on the three scales at eight weeks is reported in table 5.2 and the boxplot (figure 5.2) is a graphical illustration of this results and the relationship between the three scales.

Table 5.2 Scores at eight weeks- upper limb

SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	65	39	52	19.82	0	54	23	54
MAS	65	13.03	16	6.35	0	18	8.5	18
SIS	52	15.23	16.5	7.99	5	25	6.25	22.75

**Figure 5.2 Scores at eight weeks- upper limb**

5.3 Scores at eight weeks- mobility

Analysis of the mobility scores on the three scales at eight weeks is reported in table 5.3 and the boxplot (figure 5.3) is a graphical illustration of this results and the relationship between the three scales. Please note outliers.

Table 5.3 Scores at eight weeks- mobility

SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	65	41.67	47	9.56	7	48	38.5	48
MAS	65	27.05	30	5.32	4	30	26	30
SIS	52	36.9	40.5	8.87	10	45	32	44.75

Figure 5.3 Scores at eight weeks- mobility
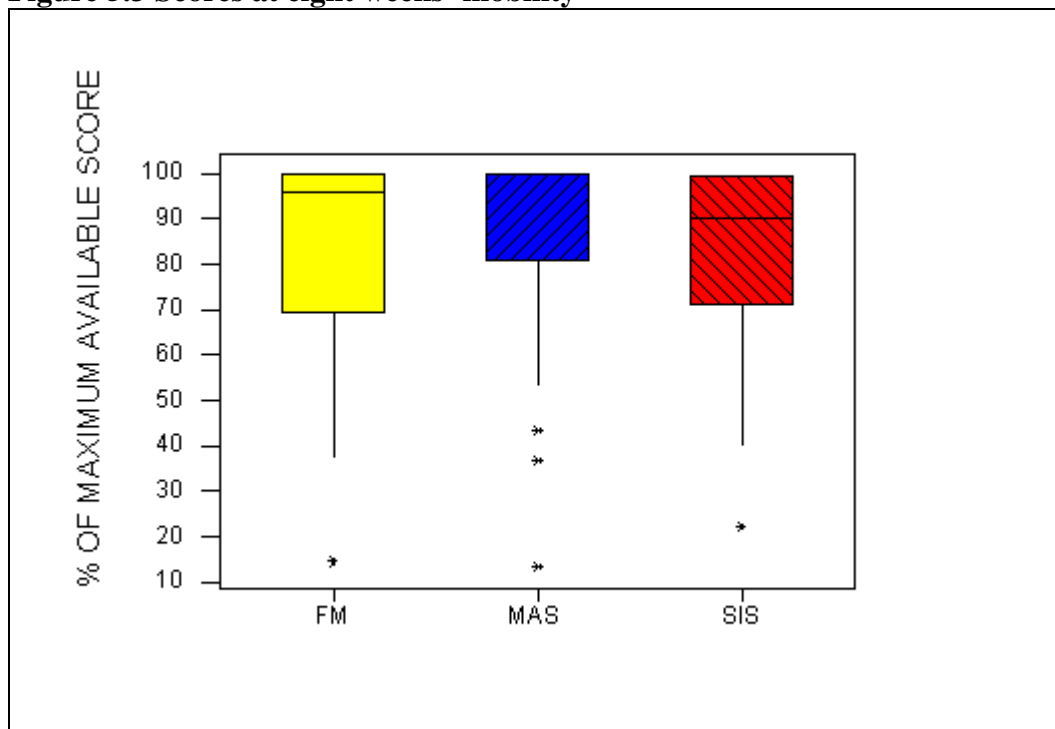
5.4 Summary of findings at eight weeks

After eight weeks, participants showed definite improvement on the scales as illustrated in figure 5.4 below. The correlation between scores on all three scales during this period was shown to be very strong (table 5.4).

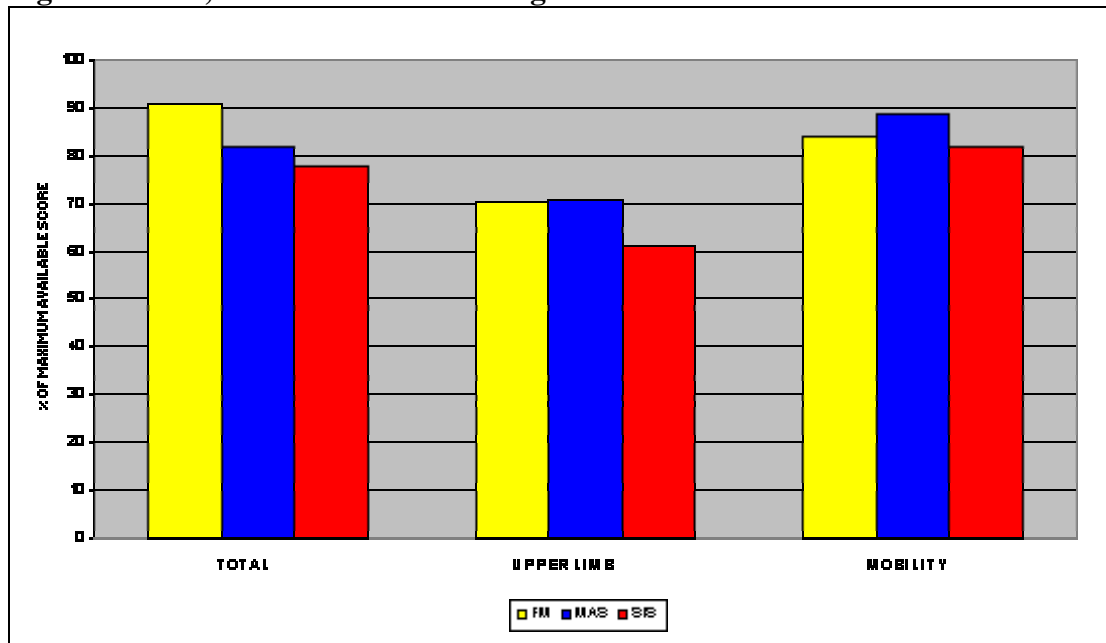**Figure 5.4 FM, MAS and SIS scores eight weeks**

Table 5.4 Correlation of scores at eight weeks

SOM	Spearman Rank Correlation		
	Total	Upper Limb	Mobility
FM & MAS	0.84	0.90	0.87
FM & SIS	0.69	0.80	0.64
MAS & SIS	0.70	0.81	0.69

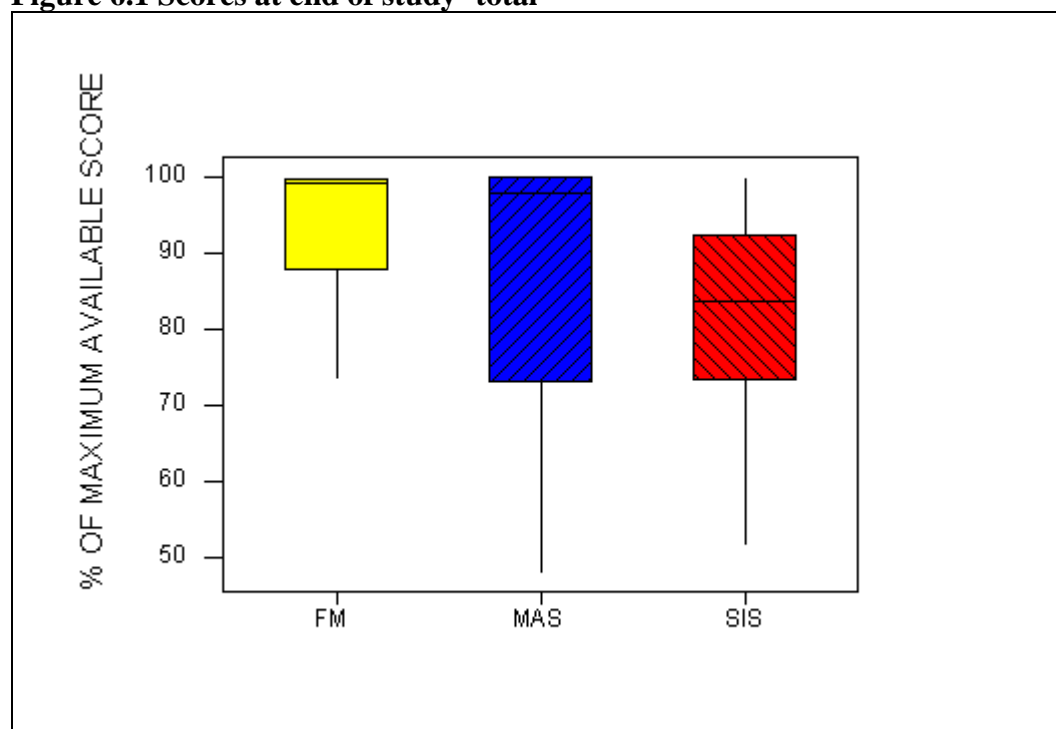**6. Standardised Outcome Measure: results at end of study**

6.1 Scores at end of study- total

Analysis of the total scores on the three scales at end of study is reported in table 6.1 and the boxplot (figure 6.1) is a graphical illustration of this results and the relationship between the three scales.

Table 6.1 Scores at end of study- total

| SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	57	414.72	434	35.36	323	440	396.5	438
MAS	58	42.17	46.5	7.95	23	48	37.25	48
SIS	44	239.93	247	36.7	152	295	216	272.5

Figure 6.1 Scores at end of study- total

6.2 Scores at end of study- upper limb

Analysis of the upper limb scores on the three scales at end of study is reported in table 6.2 and the boxplot (figure 6.2) is a graphical illustration of this results and the relationship between the three scales.

Table 6.2 Scores at end of study- upper limb

SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	57	42.23	54	17.77	0	54	31.5	54
MAS	58	12.85	16.5	6.39	0	18	6	18
SIS	44	16.91	17.5	7.32	5	25	11.25	25

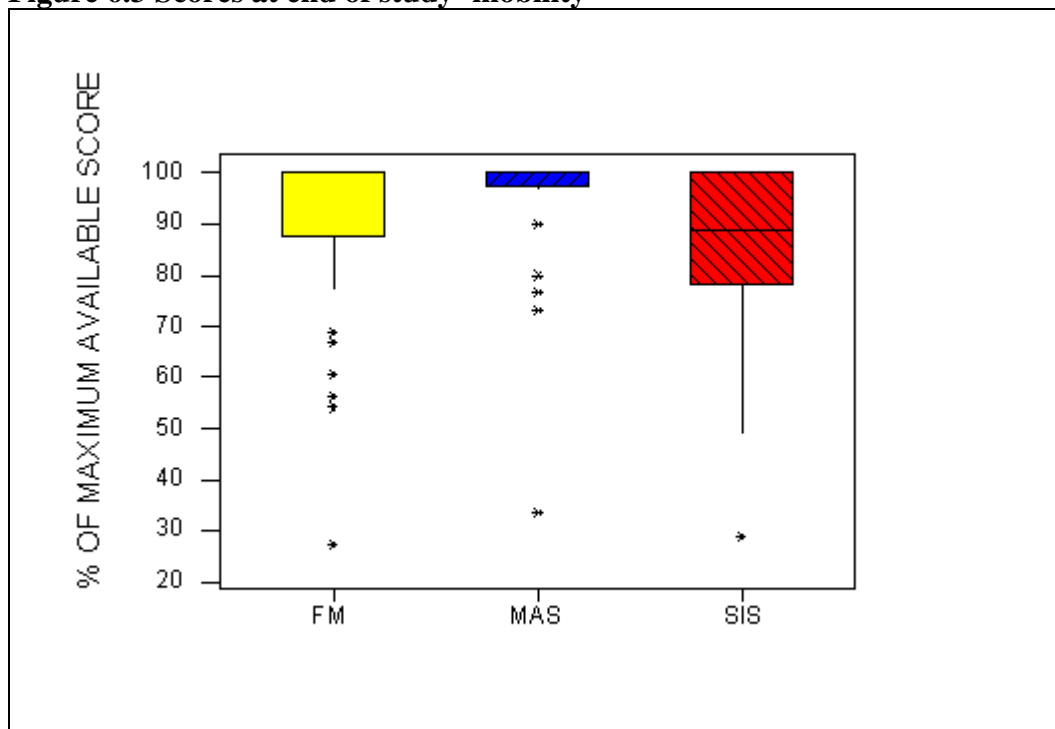**Figure 6.2 Scores at end of study- upper limb**
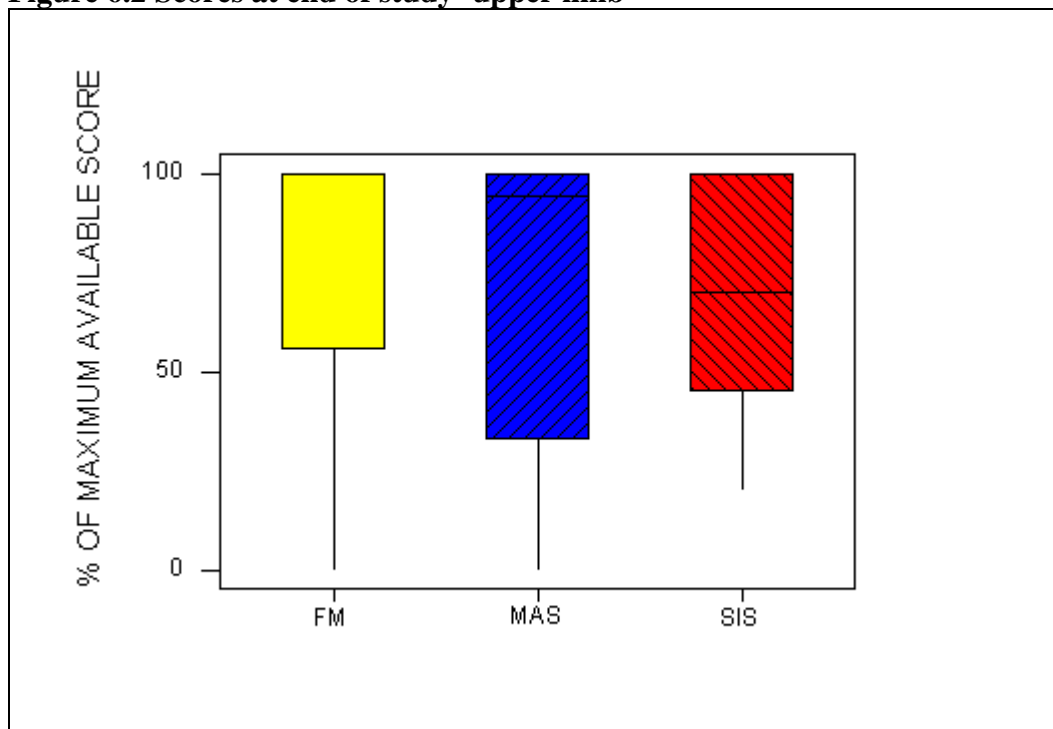
6.3 Scores at end of study- mobility

Analysis of the mobility scores on the three scales at end of study is reported in table 6.3 and the boxplot (figure 6.3) is a graphical illustration of this results and the relationship between the three scales. There are a number of outliers in this case.

Table 6.3 Scores at and of study- mobility

| SOM	n	Mean	Median	SD	Min	Max	Q1	Q3
FM	57	44.54	48	7.32	13	48	45	48
MAS	58	28.81	30	3.14	10	30	29.75	30
SIS	44	38.2	40	7.29	13	45	35.25	45

Figure 6.3 Scores at end of study- mobility

6.4 Summary of findings at end of study

The relationship of the three scales for total score, upper limb score and mobility score at end of study is illustrated in figure 6.4.

Figure 6.4 FM, MAS and SIS scores at end of study
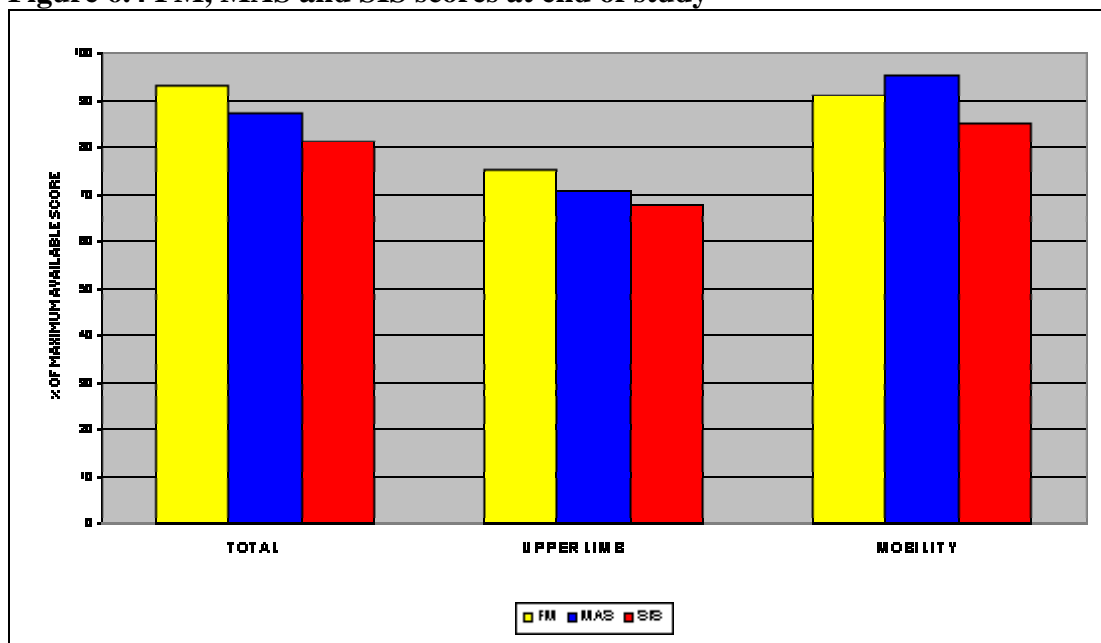
Over the three sections (total, upper limb and mobility), the SOMs show strong correlation. The total scores and mobility section scores are nearing maximum of available score.

Table 6.4 Correlation of scores at end of study

| SOM	Spearman Rank Correlation		
	Total	Upper Limb	Mobility
FM & MAS	0.88	0.84	0.84
FM & SIS	0.84	0.84	0.71
MAS & SIS	0.88	0.86	0.56

7. Change in Standardised Outcome Measure scores from entry to end of study

7.1 *Change in scores- total*
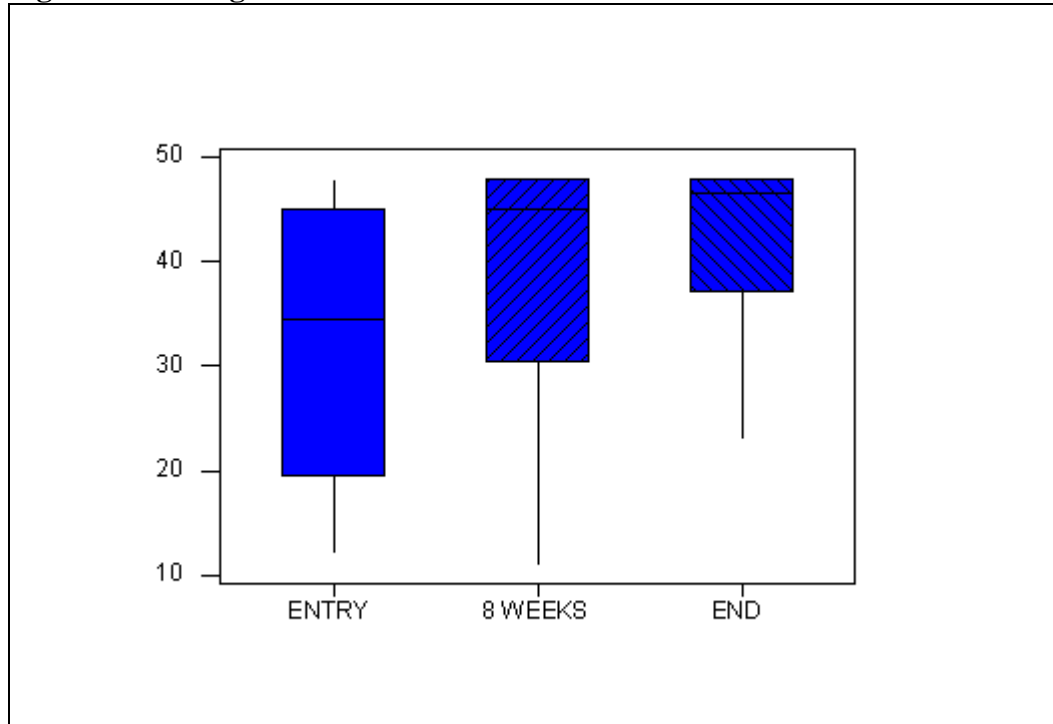
7.1.1 Change in FM scores- total

FM scores of participants changed from a mean score at entry of 377/440 to a mean score at discharge of 415/440. This is a change of 38/440 points, and a percentage change of 8.64%, from 85.68% to 94.32% of the maximum score. This change is statistically significant ($p < 0.001$). Figures 7.1 illustrate the change in FM scores over the course of the study.

Figure 7.1 Change in FM scores- total
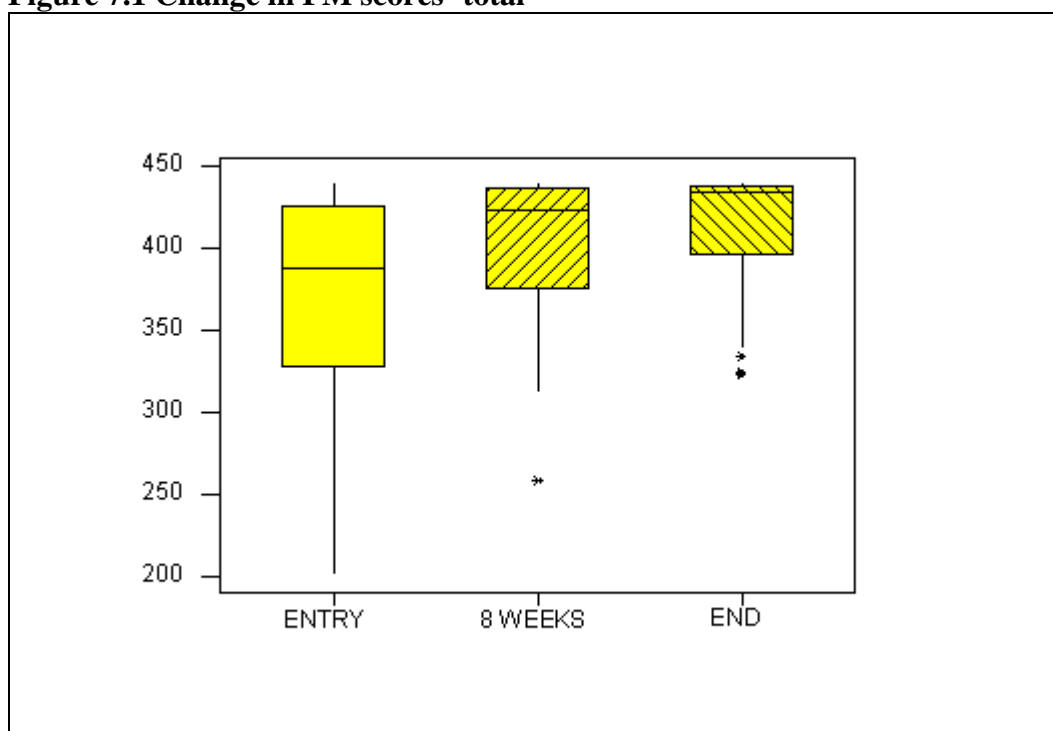
*7.1.2 Change in MAS scores- total*

MAS scores of participants changed from a mean score at entry of 34/48 to a mean score at discharge of 42/48. This is a change of 8/48 points, and a percentage improvement of 16.67%, from 70.83% to 87.5% of the maximum score. This change is statistically significant ($p < 0.001$). Figures 7.2 illustrate the change in MAS scores over the course of the study.

Figure 7.2 Change in MAS scores- total

7.1.3 Change in SIS scores- total

SIS scores of participants changed from a mean score at entry of 199/295 to a mean score at discharge of 240/295. This is a change of 41/295 points, and a percentage improvement of 13.9%, from 67.46% to 81.36% of the maximum score. This is a statistically significant change ($p < 0.001$). Figures 7.3 illustrate the change in SIS scores over the course of the study.
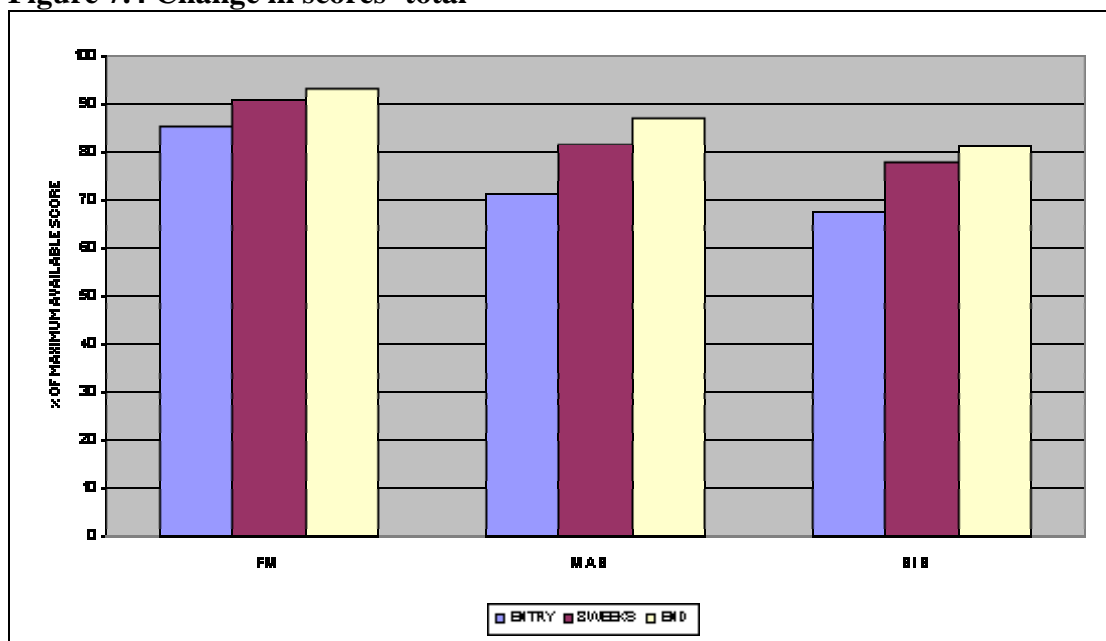
Figure 7.3 Change in SIS scores- total
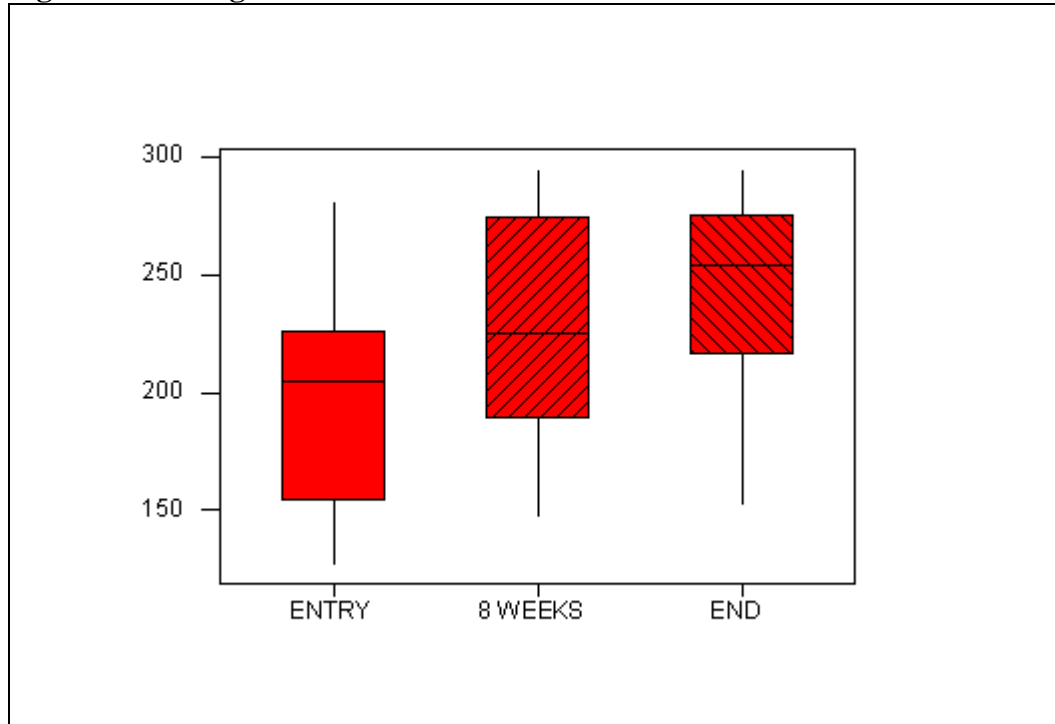
7.1.4 Overview of change in scores- total

If all scales are taken together, the improvement follows a similar pattern. Figure 7.4 illustrates the relationship of the total scores on the three scales at admission, midpoint and end of study.

Figure 7.4 Change in scores- total

7.2 Change in scores- upper limb

The change in the upper limb sections is similar to the change in the total scores for the FM, MAS and SIS. An improvement was demonstrated on all three scales, as illustrated in figures 7.5-7.8.
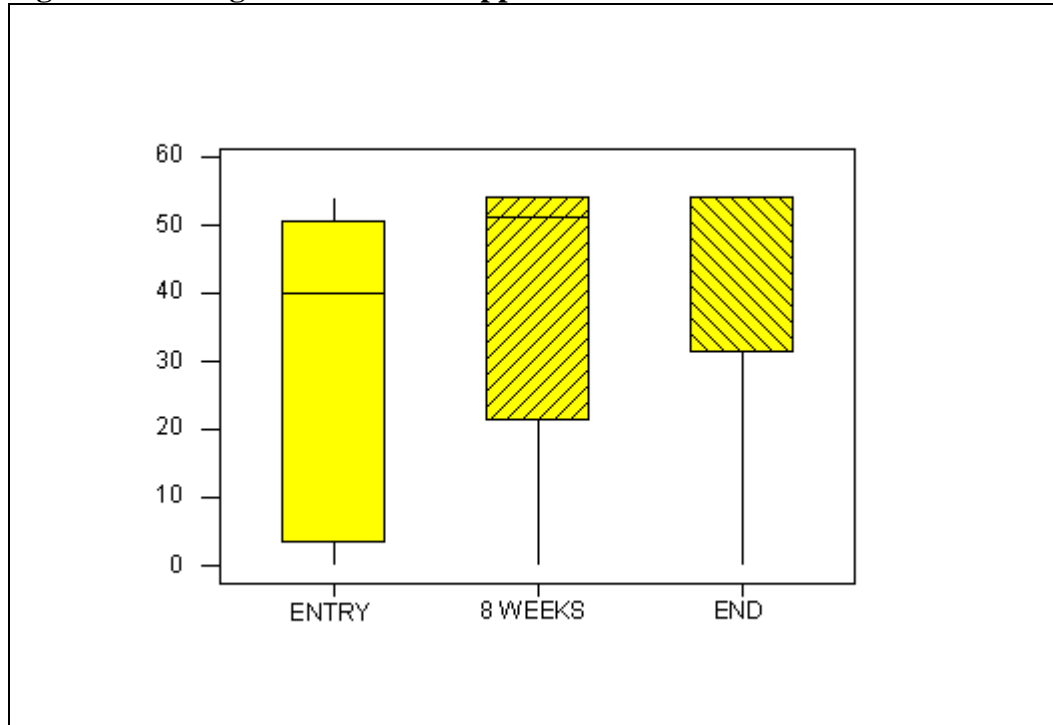
Figure 7.5 Change in FM scores- upper limb

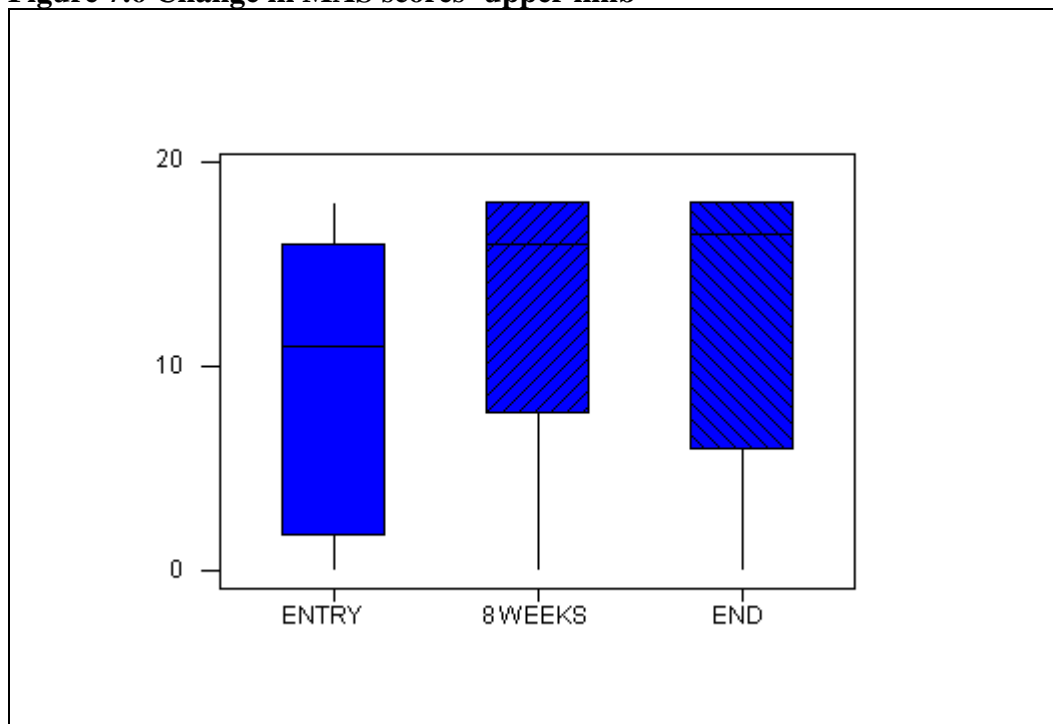**Figure 7.6 Change in MAS scores- upper limb**
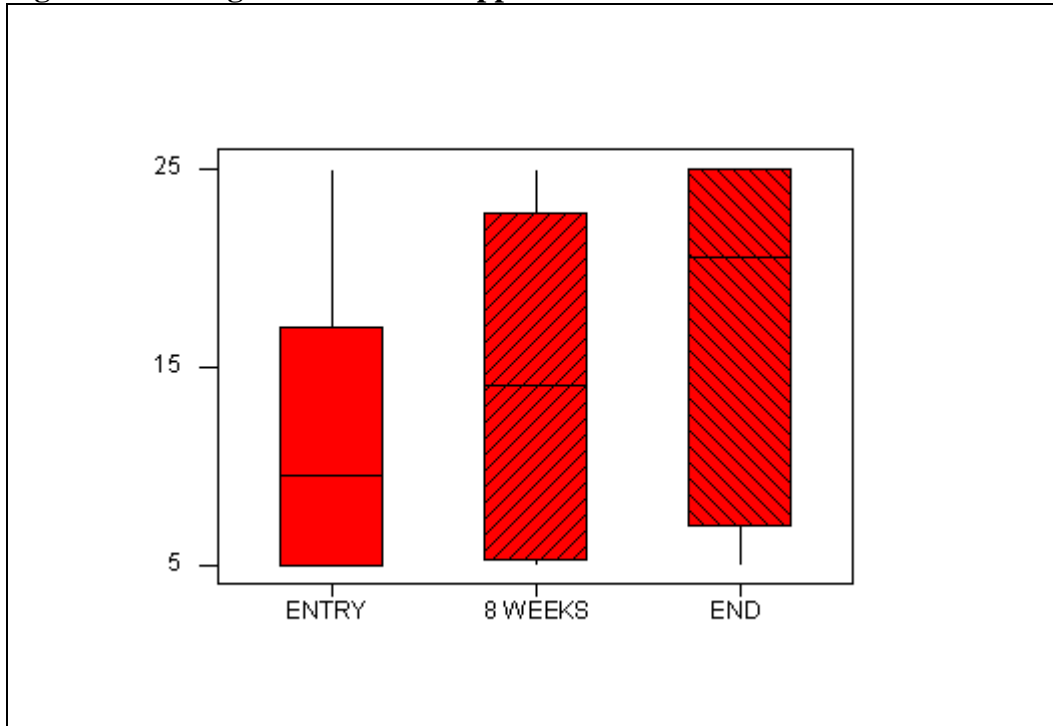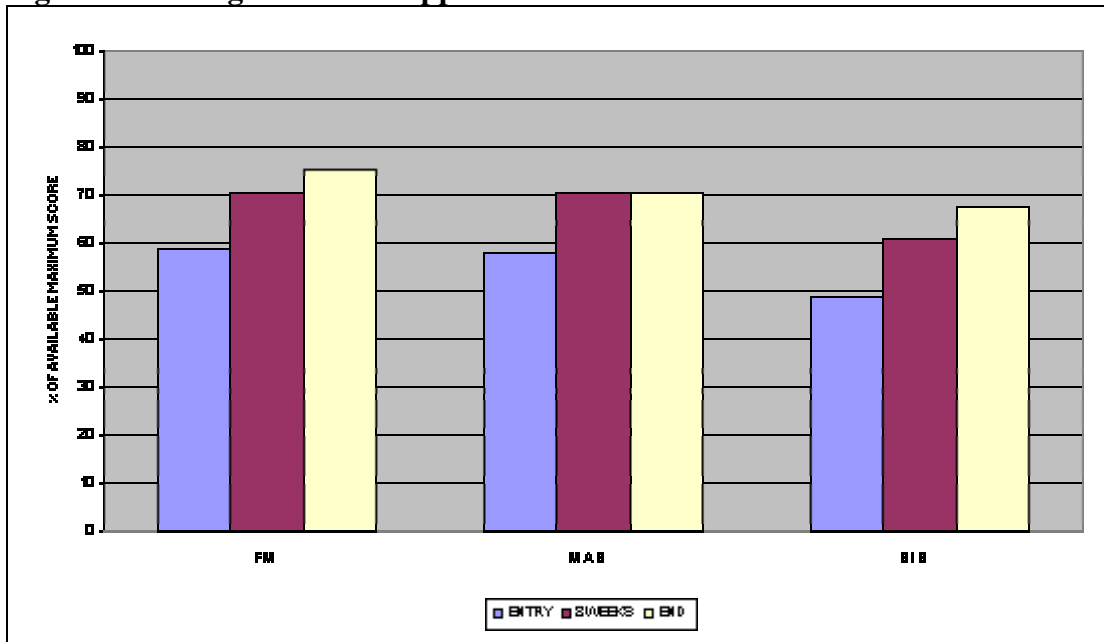
Figure 7.7 Change in SIS scores- upper limb

Figure 7.8 Change in scores- upper limb

7.3 Change in scores- mobility

Similar to the total and upper limb scores, improvement occurred in the mobility scores of the three scales. This is illustrated in figures 7.9-7.12.

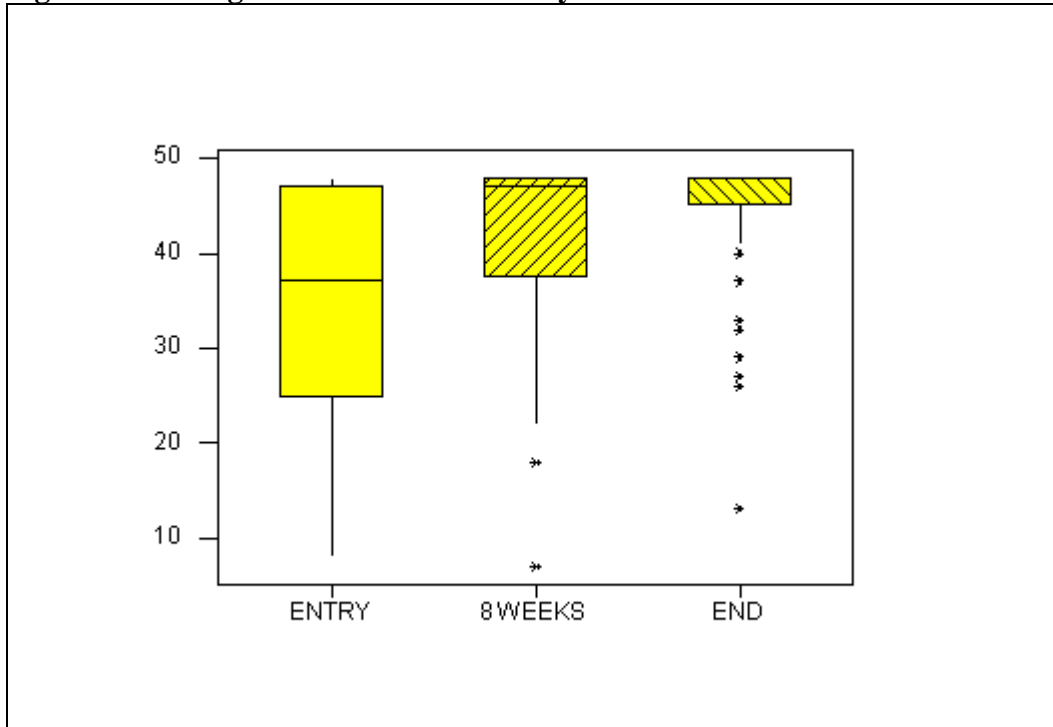**Figure 7.9 Change in FM scores- mobility**
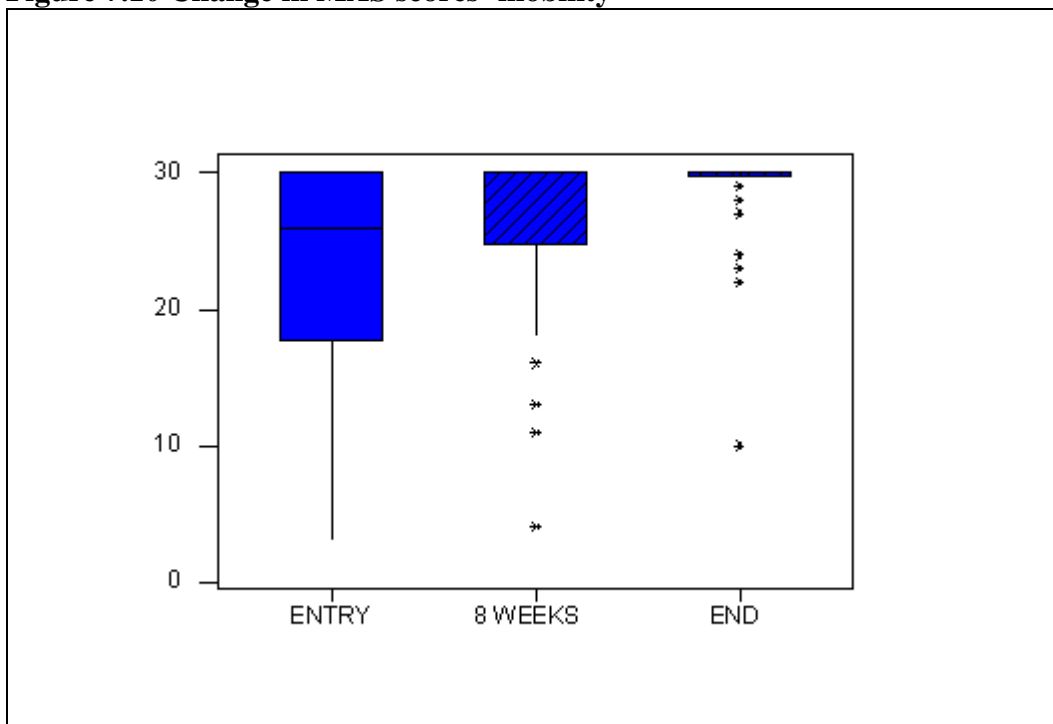

Figure 7.10 Change in MAS scores- mobility
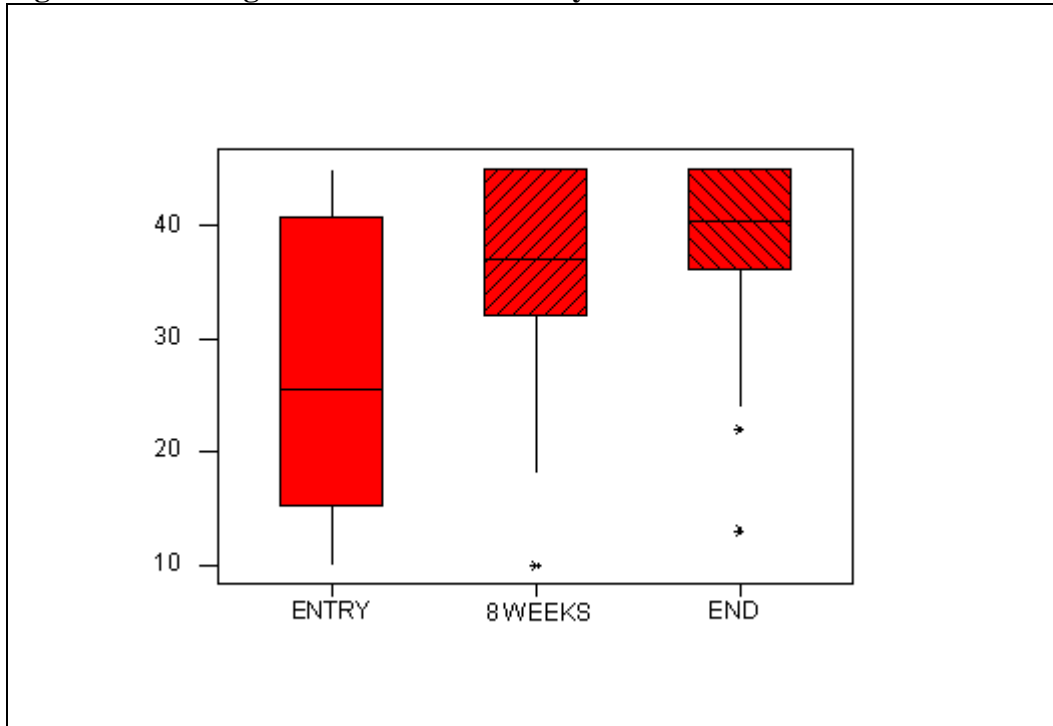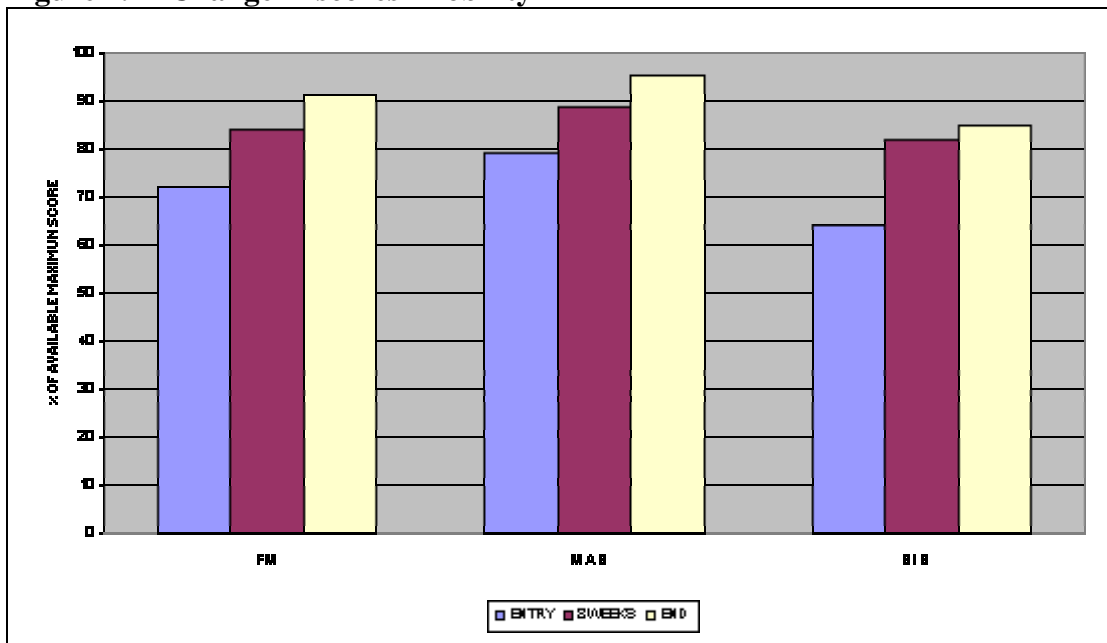
Figure 7.11 Change in SIS scores- mobility

Figure 7.12 Change in scores- mobility

7.4 Summary of change in scores

The graphs show increased scores with increasing time post stroke, thus indicating that the effects of a stroke on function and participation of the individual are less after six months. Table 7.1 below contains the statistical analysis of this difference. Scores at entry and end are analysed using the Wilcoxon signed-rank test. All subscales reach statistical significance.

Table 7.1 Summary of changes in scores

SOM	Point Estimate for difference from entry to end of study	p-value
Total Score		
FM	33	< 0.0001
MAS	7	< 0.0001
SIS	46	< 0.0001
Upper Limb Score		
FM	9	< 0.0001
MAS	2	0.0004
SIS	5	0.0002
Mobility Score		
FM	12	< 0.0001
MAS	3	0.0001
SIS	11	< 0.0001

8. Conclusion

The results therefore suggest that people make significant functional improvements in the first six months post stroke. This functional recovery has an effect on the perceived impact of the stroke. The correlation between functional scores and SIS scores confirms this. In addition, it is encouraging to see that recovery is continuing up to six months post stroke. There are studies reporting recovery in chronic stroke patients (Sullivan and Hedman 2007; Fischer *et al.* 2007). Further follow-up in this group may show that recovery can continue to take place for a longer period.

The results of the ALLADIN clinical trial are positive in terms of stroke recovery. Overall recovery on the objective measures was 88% and 94% on the MAS and FM respectively. Furthermore, an overall recovery on the subjective measure (SIS) of 81% was reported.

It is natural to assume that recovery in motor impairment should reflect in the improvement in function. Melchior *et al.* (2007) reports a relationship between sensory impairment and hand function. But not all studies agree with this finding (Chae *et al.* 2007). However the strong correlation between all three measures found in this study is indicative of the relationship between impairment, function and participation. This also confirms that the objective measures (FM & MAS) correlates with the self-reported subjective measure (SIS).

It can be seen in the results that mobility improvements are greater than those in the upper limb. The participants upper limb scores are 78.2%, 71.4% and 67.6% expressed as a percentage of maximum available scores on the FM, MAS and SIS respectively, as opposed to those of 92.8%, 96% and 84.9% for mobility, suggest that even six months post stroke, upper limb function is 14.6%, 24.6 % and 17.3% less than mobility as measured on the FM, MAS and SIS respectively. This indicates a higher rate of recovery in mobility than in the upper limb. Most of the participants in the study had middle cerebral artery (MCA) infarct, affecting the MCA territory.

The complex nature of upper limb function as compared to the lower limb function of support and locomotion will require a higher degree of motor recovery in the upper limb to translate to a meaningful functional recovery. It is also possible that more

importance is given to early mobilisation of the individual, by the therapist, the rehabilitation team, patient and the family.

The largest change occurred within the first eight weeks, up to FM 4. There is no statistically significant difference between scores obtained at eight weeks and those at end of study showing a slowing down of the rate of recovery. The results of the MAS mirrored those of the FM, with no statistically significant difference between results obtained at two months post stroke and six months post stroke (end of study).

Interestingly, there was no statistically significant difference between SIS scores also at eight weeks and those at end of study. This also suggests that the subjective self reported measure is comparable to administered measurement scales.

References

Carr, J.H., Shepherd, R.B., Nordholm, L., Lynne, D. (1985) 'Investigation of a new motor assessment scale for stroke patients', *Physical Therapy* **65**(2): 175-180.

Chae, J., Mascarenhas, D., Yu, D.T., Kirsteins, A., Elovic, E.P., Flanagan, S.R., Harvey, L., Zorowitz, R.D., Fang, Z.P. (2007) 'Poststroke shoulder pain: its relationship to motor impairment, activity limitation, and quality of life', *Archives of Physical Medicine and Rehabilitation 88(3):298-301*.

Daly, L.E., Bourke, G.J. (2000) *Interpretation and uses of medical statistics*, 5th Edition, Blackwell Science, Oxford.

Dean, C.M. and Mackey, F.H. (1992) 'Motor assessment scale scores as a measure of rehabilitation outcome following stroke', *The Australian Journal of Physiotherapy* **38**(1): 31-35.

Duncan, P.W., Bode, R.K., Min Lai, S., Perera, S., Glycine Antagonist in Neuroprotection Americas Investigators. (2003) 'Rasch analysis of a new stroke specific outcome scale: the stroke impact scale', *Archives of Physical Medicine and Rehabilitation* **84**(**7**): 950-63.

Duncan, P.W., Wallace, D., Min Lai, S., Johnson, D., Embretson, S., Laster, L.J. (1999) 'The Stroke Impact Scale Version 2.0. Evaluation of reliability, validity, and sensitivity to change', *Stroke* **30**(10): 2131-40.

Duncan, P.W., Propst, M. and Nelson, S.G. (1983) 'Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident', *Physical Therapy* **63**(10): 1606-10.

Finch, E., Brooks, D., Stratford, P.W., Mayo, N.E. (2002) *Physical rehabilitation outcome measures: a guide to enhanced clinical decision making.* 2nd Edition. Canadian Physiotherapy Association. Lippincott, Williams & Williams.

Fischer, H.C., Stubblefield, K., Kline, T., Luo, X., Kenyon, R.V., Kamper, D.G. (2007) 'Hand rehabilitation following stroke: a pilot study of assisted finger extension training in a virtual environment', *Stroke Rehabilitation* **14**(1):1-12.

Fugl-Meyer, A.R., Jaasko, L., Leyman, I., Olsson, S., Steglind, S. (1975) 'The post stroke hemiparetic patient. 1. A method for evaluation of physical performance', *Scandinavian Journal of Rehabilitation Medicine* **7**(1): 13-31.

Geyh, S., Cieza, A., Schouten, J., Dickson, H., Frommelt, P., Omar, Z., Kostanjsek, N., Ring, H., Stucki, G. (2004) 'ICF Core Sets for stroke', *Journal of Rehabilitation Medicine* suppl **44**:135-41.

Gladstone, D.J., Danells, C.J., Black, S.E. (2002) 'The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties', *Neurorehabilitation and Neural Repair* **16**(3): 232-40.

Hazard, R.G., Haugh, L.D., Green, P.A., Jones, P.L. (1994) 'Chronic low back pain: the relationship between patient satisfaction and pain, impairment, and disability outcomes', *Spine* **19**(8): 881-7.

Kwakkel, G., Wagenaar, R.G., Twisk, J.W.R., Lankhorst, G.J., Koetsier, J.C. (1999) 'Intensity of leg and arm training after middle-cerebral-artery stroke: a randomized trial', *The Lancet* **354**(9174):189-94.

Lindmark, B. and Hamrin, E. (1988) 'Evaluation of functional capacity after stroke as a basis for active intervention. Validation of a modified chart for motor capacity assessment', *Scandinavian Journal of Rehabilitation Medicine* **20**(3): 111-5.

Loewen, S.C. and Anderson, B.A. (1988) 'Reliability of the modified motor assessment scale and the Barthel Index', *Physical Therapy* **68**(7):1077-1081.

Melchoir, H., Vatine, J.J., Weiss, P.L. (2007) 'Is there a relationship between light touch-pressure sensation and functional hand ability?', *Disability Rehabilitation* **29**(7):567-75.

Petrie, A. and Sabin, C. (2000) *Medical Statistics at a Glance,* Blackwell Science, London.

Poole, J.L. and Whitney, S.L. (1988) 'Motor assessment scale for stroke patients: concurrent validity and interrater reliability', *Archives of Physical Medicine and Rehabilitation* **69**(3):195-7.

Sanford, J., Moreland, J., Swanson, L.R., Stratford, P.W., Gowland, C. (1993) 'Reliability of the Fugl-Meyer assessment for testing motor performance in patients following stroke', *Physical Therapy* **73**(7):447-54.

Sullivan, J.E. and Hedman, L.D. (2007) 'Effects of home-based sensory and motor amplitude electrical stimulation on arm dysfunction in chronic stroke', *Clinical Rehabilitation* **21**(2):142-50.

Sullivan, M.S., Shoaf, L.D., Riddle, D.L. (2000) 'The relationship of lumbar flexion to disability in patients with low back pain', *Physical Therapy* **80**(3): 240-50.

World Health Organization. (2001) *International Classification of Functioning, Disability and Health*. Geneva.

World Health Organization. (1992-19994) *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision, Vols. 1-3. Geneva.