3

Prediction of muscle power using functional movement tests

We follow the steps of EWGSOP2 algorithm given in Chapter 1 for diagnosis of sarcopenia from SARC-F to Severity. Present chapter focuses on the Confirm (step 3) of the algorithm. As shown in Figure 1.1, this step requires muscle quantity or quality. This muscle power for the STS is either obtained by using the standard clinical tests, such as, DXA, CT, MRI etc. or estimated, using physical model, such as, Takai [Takai et al, 2009] or regression model, such as, Smith [Smith *et al.* [2010]] or using some other alternative techniques that are easy to use, inexpensive, and requires expertise and equipment which is routinely used in a clinical setting. One method to examine muscle power, to detect functional decline, is to use a stair-climbing test such as a stair climb power test [Bean *et al.* [2007]]. Such a test is inexpensive, although it does require the testing site to have stairs. Other functional screening tests that might estimate power include the Timed-Up and Go (TUG) and the One Leg Stance (OLS), which are very simple functional tests require no equipment other than a chair and a stopwatch. This chapter focuses on prediction of muscle power using BMI, Grip Strength, Gate Velocity, TUG and OLS. Although, we are using the supervised set-up, we do not have the supervisor, that is, muscle power data. We rely on the estimated muscle power provided by the earlier mentioned models, due to Takai and Smith. Results of these models, to predict muscle power using the STS and functional screening tests, are compared in part one of this chapter, followed up by an evaluation in which lower-limb power is measured using the gold-standard isokinetic dynamometer in part two. A predictive equation is developed by linear regression using the STS and basic subject characteristics as inputs. We then use, in the next chapter, machine learning techniques to classify older people as sarcopenic or non-sarcopenic according to the EWGSOP algorithm for physical performance.

3.1 INTRODUCTION

As mentioned earlier, muscle mass is best measured using body imaging techniques such as a magnetic resonance imaging (MRI) scan or CT Scan or dual energy X-ray absorptiometry (DEXA) etc. However, all of these techniques are expensive and can only be used under the guidance of an expert. The EWGSOP also recommended grip strength as a good measure of muscle strength [Cruz-Jentoft *et al.* [2010]], however, a measure of knee flexion and extension power could be more beneficial in the diagnosis of sarcopenia as older people lose power quicker than strength [Cooper *et al.* [2013]]. This was confirmed in the updated version of the EWGSOP in which the STS was added as a measure of lower-limb muscle strength [Cruz-Jentoft *et al.* [2018]].

Other than measuring strength and functional performance, the STS could also be used to estimate muscle power and muscle mass. In an early work in this field, the STS was identified as a useful measure of lower-limb power [Lindemann et al, 2003]. In a subsequent study, an accurate estimation of lower-limb muscle power was obtained using a regression equation in which only body weight and the number of sit-to-stands performed in 20 seconds were used, see, equation 1.2.An earlier study had also reported a good estimation of muscle mass obtained from an MRI, again using only a linear regression equation in which only three basic variables, leg length, body mass, and the time taken for a single sit-to-stand movement, were used, see equation 1.1. If estimates of all three variables were obtained from the 5STS, it may become a single screening tool for sarcopaenia, which could then be validated by clinical tests using MRI and an isokinetic dynamometer.

This study presents a regression-based approach to design a model that can take inputs from these tests and give us an estimate of muscle power. In order to estimate muscle power, we need to establish a relationship between the inputs from these screening tests and the muscle power. We use both the Takai [Takai *et al.* [2009]T] and Smith [Smith *et al.* [2010]] models to estimate the muscle power for the STS. This model can potentially help in detecting problems like sarcopenia in older people at an early stage. In addition, we also aim to verify the relationships due Takai and Smith by measuring lower-limb power using the gold-standard isokinetic dynamometer for lower-limb functional assessment.

A set of experiments is carried out and results obtained are compared in to two parts. The first part provides all experimental setup and results for predicting muscle power using STS and functional screening tests, while the second part presents experimental setup and result for measuring lower-limb power using a gold standard isokinetic dynamometer.

3.2 PART 1

3.2.1 Methodology

Participants

The dataset used for this study was from an initial study carried out in Jodhpur in 2015 for the Culturally Appropriate Geriatric Screening study, in partnership with the University of Bedfordshire (UK). A total of 98 participants (24 females, 74 males) aged above 65 years with mean age (70.3 ± 5.4) years were included in the study, with 11 subjects (11%) having fallen in previous 12 months. Ethical approval for the study was obtained from both the Institutions for Health Research Ethics Committee of the University of Bedfordshire (IHREC574), and the Ethics Committee (Human studies) of the Asian Centre for Medical Education, Research & Innovation (C/18-10-15) Jodhpur. All participants were medically screened before participating in this study. Inclusion criteria included no known musculoskeletal or neurological disorder for all subjects, being community dwelling, and aged over 65 years. Participants were provided with an information sheet pertaining to the study and gave their informed consent and were told of their rights to withdraw at any time from the study.

Testing Protocol

The above participants were asked to participate in five tests, the OLS, STS, TUG, gait velocity and grip strength. They were also asked to answer a questionnaire that included questions about their past medical records and health. As part of the TUG test, the subjects were asked to stand up from a chair, walk for 3 meters (9.84 feet) and return to the chair and sit. To check for the gait velocity, the subjects were asked to walk for a distance of 15 feet at their normal pace without over reaching themselves. Both tests were performed twice, with the best time taken as their performance measure. Participants were further instructed to perform the OLS on their preferred leg with their eyes open. Following this initial familiarisation test, participants repeated the OLS three times. Subjects were asked to remain standing on one leg for 30 seconds, if possible, with the best time taken as their OLS performance. For the grip strength test, a Jamar dynamometer was used (Lafayette Instrument Company, Lafayette, Indian, USA). Participants were required to hold the dynamometer in their dominant hand, with the arm at right angles and the elbow by the side of the body. The dynamometer was squeezed with maximum effort, with force maintained for three seconds, with their best effort of three attempts taken as their grip strength measurement. Finally, participants performed a 5STS test that required them to stand up and sit down as quickly as possible five times, without stopping, and with their arms folded across their chest. The fastest of the two attempts was taken as their STS performance.

Muscle Power

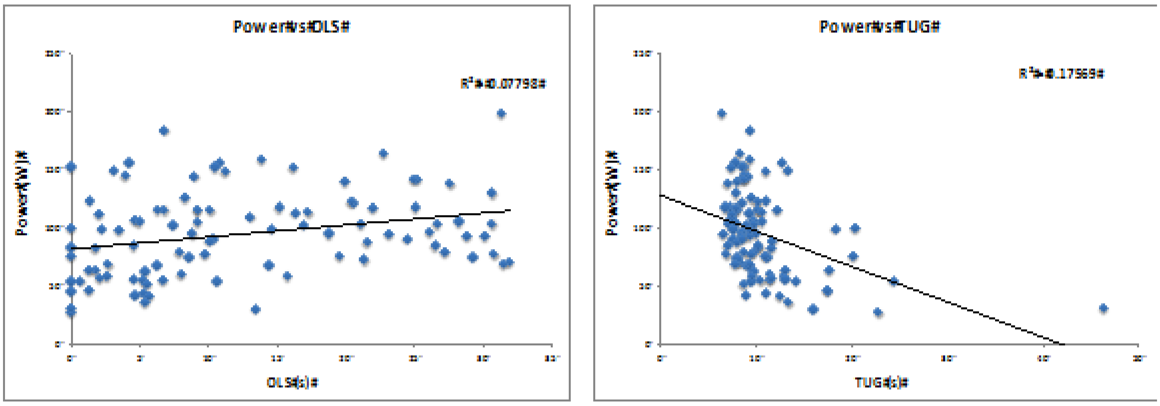

Figure 3.1 : Power vs OLS (left) and TUG (right)

We now estimate the reference muscle power against which the dependency would be evaluated. We use Takai [Takai *et al.* [2009]] and Smith [Smith *et al.* [2010]] models to derive estimates of muscle power from the 5STS.

The Takai model uses the STS time, body mass, and leg length as per the following equation to estimate power generated during the STS:

$$P(T) = \frac{(L-0.4)*Mass*10*g)}{STS_{10}} \tag{3.1}$$

where, L is leg length, which was taken as 41% of the subject's height, Mass is participant's body mass in N, g is acceleration due to gravity of $9.81m/s^2$ and STS_{10} is the time taken to perform the 10STS used in the Takai study. In the present study the 5STS time was doubled as an approximation of 10STS.

The Smith model is a linear regression model and much simpler than that of Takai, and only uses STS performance and body mass :

$$P(S) = -715.218 + 13.915*Mass + 33.425*STS_{20} \tag{3.2}$$

where Mass is the participant's body mass in N and STS_{20} is the number of STS the subject is able to be performed in 20 seconds, which was extrapolated from 5STS time.

We use linear regression to predict muscle power using five screening tests, BMI, TUG, gait velocity, grip strength and OLS. The plots of OLS and TUG are shown in Figure 3.1 . It can be seen that the performance of both relationships was weak, with $(R^2) = 0.18$ and $(R^2) = 0.08$ for TUG and OLS, respectively. In contrast, stronger relationships are evident for both gait velocity $(R^2 = 0.33)$ and grip strength $(R^2 = 0.39)$.

We now build a linear regression equation to predict muscle power, that is,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon \tag{3.3}$$

where X_1=BMI, X_2=Gait Velocity, X_3=TUG, X_4=Grip strength, X_5=OLS

To assess how these results would generalize to the entire dataset, we performed k-fold cross validation, where in every iteration on a different set of known data is given on which training is run (training dataset). Cross validation limits problems like over fitting and shows how the model

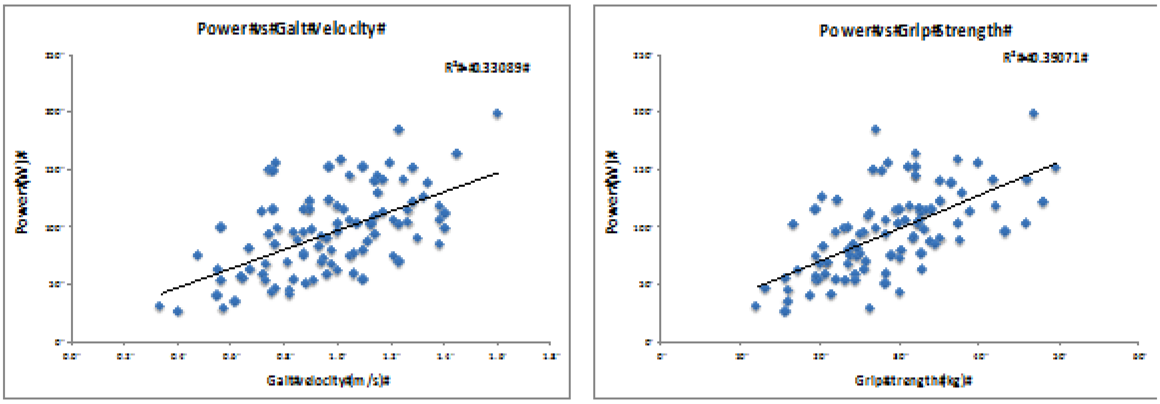

Figure 3.2 : Power vs gait velocity (left) and grip strength (right)

will generalize to an independent dataset. To reduce variability, multiple rounds of cross-validation are performed using different partitions, and the validation results are averaged to estimate a final predictive model.

3.2.2 Result

We now fit the parameters of linear regression model in both the cases, when $P(T)$ is considered as estimated power then we obtain

$$Y = -73.7748 + 2.2839X_1 + 62.5587X_2 - 0.2696X_3 + 2.0252X_4 - 0.5308X_5 \tag{3.4}$$

This model in equation 3.4 is significant at 95% confidence level (R^2)=0.61, $F(Statistic)$ = 28.94, P=0.000). When stepwise regression is used, rather than entering all variables simultaneously, we use various combinations of inputs and find that the variability explained by grip strength, BMI, and gait velocity is equally good. Thus, a 3-factor model in equation 3.5 performed as good as the 5-factor version (R^2)=0.60, F Statistic = 23.39, p=0.000), with the equation given below:

$$Y = -79.730 + 2.411X_1 + 56.408X_2 + 1.918X_4 \tag{3.5}$$

Next, we fit the model 3.6 considering $P(S)$ as estimated power:

$$Y = -72.936 + 3.328X_1 + 50.769X_2 + 56.408X_3 + 1.386X_4 - 0.171X_5 \tag{3.6}$$

This equation 3.6 is also seen significant at the 95% confidence level $(R^2 = 0.70, F(Statistic) = 42.39, p = 0.000)$. When stepwise regression was used, observation was very similar to that of the earlier one. A 3-factor model in equation 3.7 is found to perform as good as 5-factor version $(R^2 = 0.69, F(Statistic) = 69.49, p = 0.000)$.

$$Y = -91.133 + 3.343X_1 + 58.308X_2 + 1.372X_4 \tag{3.7}$$

It is clear from both the above predicted models that BMI, grip strength and gate velocity can be used to estimate muscle power while other functional parameters are less relevant.

3.2.3 Discussion

Both the above models perform well, with the $P(S)$ based model achieving slightly better results. The $R^2 = 0.70$ equates to a correlation of 0.83, which can be considered to be a very large

Table 3.1 : Subject characteristics

Group	Age (years)	Height (m)	Weight (kg)	BMI (kg/m2)
Young	20.8 ± 2.0	1.69 ± 0.10	72.6 ± 9.4	25.5 ± 3.4
Middle aged	53.7 ± 5.4	1.66 ± 0.09	72.5 ± 15.0	26.4 ± 5.3
Old	72.7 ± 8.2	1.64 ± 0.09	72.6 ± 15.6	26.9 ± 4.9

effect. The $R^2 = 0.61$ for the $P(T)$ based model equates to a correlation of 0.78, which is also a very large effect [Hopkins *et al.* [2009]]. The current study suggests that it might be possible to construct a factor model to estimate lower limb muscle power without using any expensive equipment. Such a model could help to identify functional decline, loss of independence and physical frailty in older people. It could also help detect patients suffering from Sarcopenia.

The present study did have some limitations. Firstly, the data was collected from an Indian population, with Indians reported to have lower than normal values for gait velocity and grip strength [Gunasekaran *et al.* [2016]]. Indeed, the stepwise 3-factor versions, $P(T)$ and $P(S)$ based models, incorporated grip strength and gait velocity, as well as BMI. A second limitation is that we use estimated value of power. However, similar results have been seen for both the models. An additional study is therefore needed in which power, measured using isokinetic dynamometry, is used instead of the estimated power. The above discussion motivates us for the following study.

3.3 PART TWO

3.3.1 Methodology

Participants

The data for the following study was collected at the U.K.. Three groups of participants took part in this pilot study, with ten people in each group. The young adult groups consisted of people aged between 18-30 years (6 females, 4 males). The second group, middle-aged adults, consisted of people aged between 40-60 years (6 females, 4 males). The third group of older adults consisted of people aged 65 years and over (6 females, 4 males). Subjects characteristics are presented in table 3.1. The study was granted ethical approval by the Institute for Health Research Ethics Committee of the University of Bedfordshire (IHREC798). Participants were provided with an information sheet pertaining to the study and a health screening questionnaire and gave their informed consent.

Testing protocol

All testing was completed in the Sports Therapy laboratories at the University of Bedfordshire. Anthropometric measures of the participants were taken, including height (m), weight (kg), with each participant's body mass index (BMI) (kg/m^2) calculated. Leg length was measured using the 'true' leg length difference method [Sabharwal and Kumar [2008]],by measuring the distance between the anterior superior iliac spine (ASIS) and the medial malleolus. Participants completed a 5-minute warm-up on a cycle ergometer (Monark Ergomedic 818E, Monark Exercise, Vansbro, Sweden) to reduce the possibility of any injury occurring during the testing. The warm-up workload was set at 50W, with a self-selected cadence between 60-80 RPM. After completing the warm-up, a demonstration of the STS was performed, with participants permitted a practice trial of the test if they wished.

The participant performed two trials of the 5STS and the 30STS, with order of the 5STS and 30STS randomised. Participants were given a two-minute rest period between each trial and each test to reduce the effect of fatigue. The time taken to complete each test was recorded using a

Table 3.2 : STS performance

Group	5STS (s)	30STS (reps)
Young	8.5 ± 1.9	20.6 ± 4.7
Middle aged	10.7 ± 1.9	16.5 ± 1.4
Old	13.5 ± 3.1	13.2 ± 3.0

stopwatch. After completion of the two STS tests, participants performed two sets of five repetitions of knee flexion and extension of both lower extremities on Isokinetic Dynamometer (IKD) (Biodex 3, Biodex Medical Systems Inc., Shirley, New York, USA). Participants were given a two-minute rest period between each trial. The angular velocity of the IKD was set at 60° per second.

The best performance of each STS was used in all subsequent analysis. With respect to the IKD data, values of both peak torque and total work were calculated [Kannus [1994]]. Peak torque values for knee flexion and knee extension were calculated for each repetition of each of the two sets, with the maximal value of Moment Angle Position (MAP) curve considered to be peak torque [Baltzopoulos and Brodie [1989]]. The maximal values of peak torque for both knee flexion and knee extension from the two sets of five repetitions were used in all subsequent analyses. Total work was calculated as the Area Under the Curve (AUC) for the torque curve for each repetition for the period when velocity was equal to 60°per second [Neder *et al.* [1999]]. The maximal values of work for both knee flexion and knee extension from the two sets of five repetitions were used in all subsequent analyses. All isokinetic data analysis was performed using custom-made LabVIEW software (National Instruments Ltd, Austin, Texas, USA).

Data analysis

All analyses were carried out using SPSS version 25 (IBM Corp, Armonk, New York, USA). Data were assessed for normality using the Shapiro-Wilk owing to a sample size less than 50. All variables were normally distributed, therefore parametric statistics were used. The effect of age group on STS performance was determined using a one-way analysis of variance, with Bonferroni adjustments made for post-hoc pairwise comparisons. Stepwise linear regression analysis, discussed in previous section, was used to construct a model to predict the measures of muscular performance from the isokinetic dynamometer (peak torque and maximal total work for a single repetition) using data from the STS tests and anthropometric measures.

3.3.2 Result

Sit-to-stand performances of the three groups are shown in Table 3.2. There was a significant difference between the three groups of participants for both 5STS $(F = 10.834, p = 0.000)$ and 30STS $(F = 12.431, p = 0.000)$. When pairwise comparisons were made, both 5STS and 30STS performance was significantly better between each successive comparison of age-groups, with effects sizes (Cohen's d) ranging from (1.1-1.9).

Performance of the 5STS was negatively correlated with peak torque for both knee flexion and knee extension Figure 3.3 (left). r=0.601, p<0.05 and Figure 3.3.(right) r=0.599, p<0.05, respectively). Performance of the 30STS was also negatively correlated with peak torque for both knee flexion and knee extensionFigure 3.4 (left) r=0.568, p<0.05 andFigure 3.4 (right). r=0.466, p>0.05, respectively).

When total work was considered, the 5STS was negatively correlated with work for both Figure 3.5 (left). r=0.587, p<0.05 andFigure 3.5 (right). r=0.576, p>0.05, respectively). Performance of the 30STS was also negatively correlated with peak torque for both knee flexion and knee extension

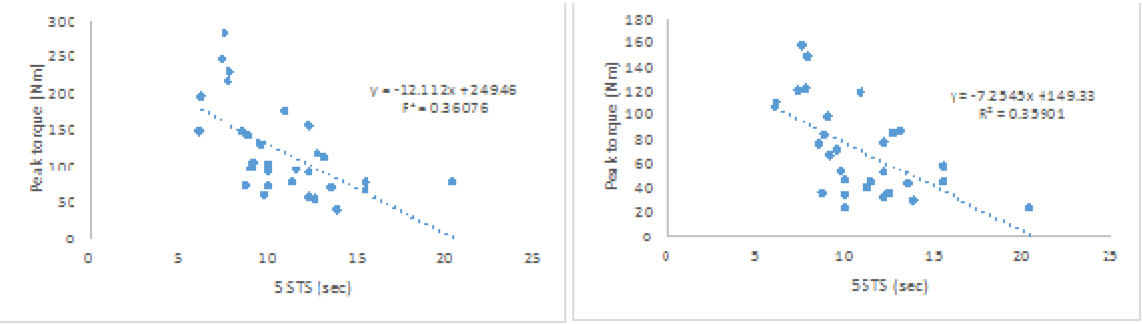

Figure 3.3 : Peak torque versus 5STS performance for knee extension (left) and knee flexion (right)

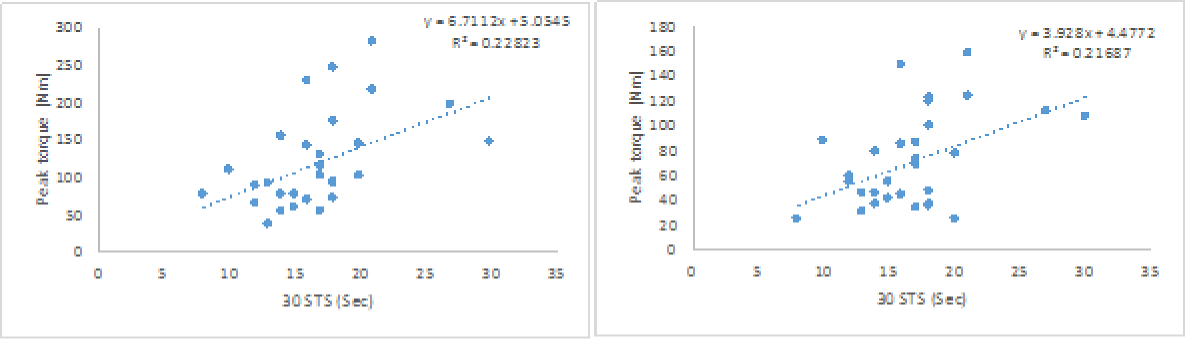

Figure 3.4 : Peak torque versus 30STS performance for knee extension (left) and knee flexion (right)

Figure 3.6 (left). r=0.478, p<0.05 and Figure 3.6 (right) r=0.466, p>0.05, respectively).

The results of the stepwise linear regression for the 5STS combined with anthropometric parameters are shown in Table 3.3. Models were developed for peak torque and total work for knee extension, knee flexion, and the sum of knee extension and knee flexion. In all cases, the stepwise regression retained only one anthropometric variable in the analysis, which was always the height of the subject. The variance explained by these models ranged from 62-73%, which equates to correlations from 0.76-86, which are all very strong using the Hopkins' scale [Hopkins et al, 2009].

3.3.3 Discussion

The aim of the first part of this study was to identify the relationship between the 5STS and other components of the original EWGSOP algorithm as well as other commonly used tests of physical function. The results showed that the 5STS test can be used to construct a factor model to estimate lower limb muscle power without requiring expensive equipment. However, in the first part of this chapter it was not possible to verify this relationship due to the lack of an isokinetic dynamometer in the Jodhpur laboratory. It should be noted that the methods used in this study to develop predictive models used simple linear regression, rather than more advanced methods. The role of more complex methods from machine learning will be introduced in the following Chapter of the thesis.

Following on from the first part of the study, a second study was carried out using data collected in the Luton laboratory of the Institute for Health Research. In this part of the study, the aim was to verify the relationship between the STS performance and power identified in previous work [Smith et

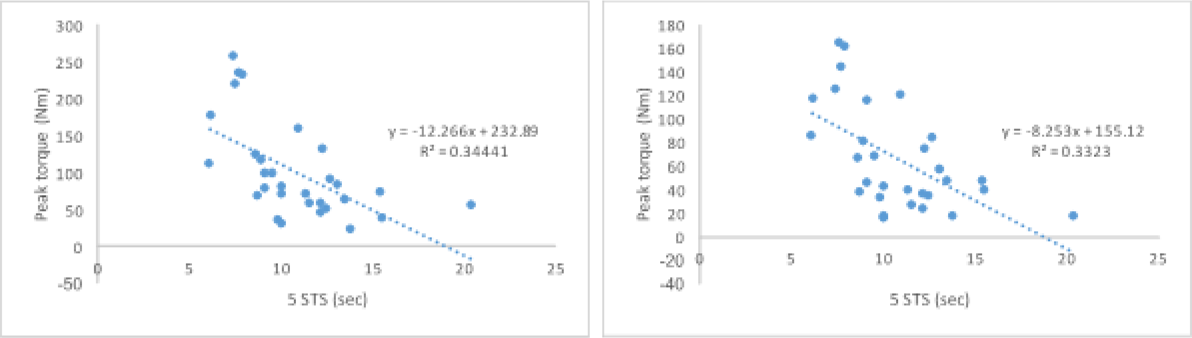

Figure 3.5 : Total work versus 5STS performance for knee extension (left) and knee flexion (right)

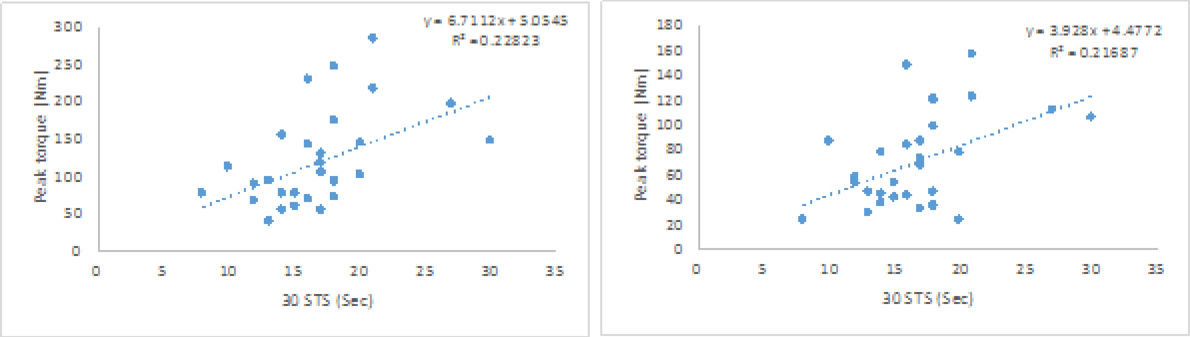

Figure 3.6 : Total work versus 305STS performance for knee flexion (left) and knee extension (right)

Table 3.3 : Sit to Stand performance

	Knee extension	Knee flexion	Sum of extension and flexion
Peak torque	r=0.760 $R^2 = 0.578$ F=18.513 p=0.000	r=0.845 $R^2 = 0.713$ F=33.573 p=0.000	r=0.808 $R^2 = 0.652$ F=25.320 p=0.000
Total work	r=0.789 $R^2 = 0.622$ F=22.254 p=0.000	r=0.855 $R^2 = 0.730$ F=36.590 p=0.000	r=0.830 $R^2 = 0.690$ F=30.005 p=0.000

al, 2010, Takai et al, 2009].

The results obtained in this protocol confirm the findings of the two previous studies. In the study of Takai [Takai et al, 2009], a correlation of r=0.73 was reported between their STS power equation using leg length, body mass, and the fastest STS time with power from an isometric knee extension test. This corresponds well with the results of the current study of r=0.76 for knee extension, while even better results were obtained for knee flexion. Likewise, Smith [Smith et al, 2010] reported a correlation of r=0.90 between their STS equation using body mass and time for 20 STS and peak power during the STS, which was measured using motion capture and force plate data. The best results in the present study were obtained for total work, rather than peak power, although these differences were small for all comparisons made (knee flexion, knee extension, combined flexion and extension).

Of note in the present study is the stepwise regression equations that systematically included only 5STS time and height. Firstly, this could suggest that the 5STS is a better test of power than the 30STS, for which a fatigue component could be worth assessing. This could be carried out using an iSTS device that was able to compare the time for the earlier STS in comparison to the later STS during the 30-second test. In addition, the use of height in the prediction equation underlines the importance of having a chair of an appropriate height for people of different heights, with taller people having less work to do to stand up for a given height of chair than shorter people. This would suggest that if an iSTS is developed using a chair, this would need to height-adjustable to maximise the precision of any predicted models of sarcopenia developed.

3.4 CONCLUSION

These findings confirm that the STS is a suitable test of muscle power in subjects of all ages. Further work will need to determine whether better results will be obtained using machine learning methods when the requirement of the models includes classification. Future work in this area should evaluate the effect of using iSTS parameters such as those identified in Chapter 2 (iSTS Systematic Review), while a validation of power developed during the 5STS as a predictor of muscle mass will also be needed to confirm the second finding of the Takai study [Takai et al, 2009].

…