# 6

# Validation of the instrumented chair part one: a comparison of four approaches to evaluate the STS movement

An article based on the results given in this chapter has been accepted for publication by IEEE Transactions on Neural Systems and Rehabilitation Engineering. The version presented here is similar to that submitted for publication with some additional information related to the thesis. The citation for this work is presented below:

*Shukla, B.K, Jain, H., Vijay, V., Yadav, S.K., Mathur, A., Hewson, D.J. A Comparison of Four Approaches to Evaluate the Sit-To-Stand Movement. Accepted by IEEE Trans Neural Syst Rehab Eng.*

## 6.1 INTRODUCTION

The importance of the STS test has been highlighted in previous chapters that have used the assessment as a means of screening older adults at risk of falling. Two main variations of the test have been described in which the person either performs 5 STS as quickly as possible or the person performs as many STS as possible within 30 seconds.

Performance in the 5STS is typically measured using a stopwatch. Recently, instrumented versions of both tests have been developed as a means of improving the accuracy of measurement, but also to extract additional information about STS performance that could improve accuracy of predictions. Such tests have used a range of techniques including body-worn accelerometers [Kang *et al.*Weiss *et al.* [2010]], pressure sensors [Arcelus *et al.* [2011]], and visual sensors, often using multiple cameras [Banerjee *et al.* [2013],Allin and Mihailidis [2009]]. In addition to the possibility of such automatic detection, in one study a Kinect was used to evaluate the 5STS [Ejupi *et al.* [2016]], with parameters extracted from the STS more closely related to strength than overall STS time. Such a finding indicates that extracting data on the way in which the STS is performed, rather than simply the time taken to perform the 5STS, could be beneficial.

Here, we analyse the performances of two novel approaches to evaluate the STS and compare them to two previously used instrumented systems to evaluate the STS, the Kinect and a force plate.

The contributions are two-fold:

(1) We propose a design of a novel device in which four force sensors are built into a chair to measure individual STS cycles, which removes the requirement for participants to wear body sensors throughout the experiment.

(2) We propose a low-cost video framework to measure STS time using only a single inexpensive RGB camera. The human skeleton from the frames captured with the RGB camera is extracted using a deep learning network, with frame sequences then segmented into STS cycles using the change in the location of the head.

This framework provides following advantages:

(1) A single RGB camera is a low-cost setup that can be easily extended to android phones [Ejupi *et al.* [2016], Pirsiavash *et al.* [2014], Jain and Harit [2016]].

(2) The method does not involve background subtraction to extract the human silhouette, which although used previously with an RGB-based camera setup [Banerjee *et al.* [2013]], fails in a cluttered environment when silhouette extraction becomes difficult. In contrast, the proposed method uses a deep pose library to extract body position.

(3) The use of visual sensors allows monitoring of both the time taken to perform the STS and the way it is performed, which is not possible in sensor-based approaches alone.

(4) While both STS performance and STS time can be analyzed using an RGB camera, the instrumented chair provides additional information related to the transfer of mass, which could be useful in predicting muscle power as discussed in Chapter 3.

In the next section, we briefly set the context to highlight the need of such comparison. Further, we present the complete chair design and pose estimation using RGB camera. Next, we describe the methodology used to determine STS time and STS velocity using both the visual sensors (RGB and Kinect) and the force-based sensors (chair and force plate). We then present our experimental results, compare the performance of the methods for the four systems, and conclude the work.

## 6.2 EVALUATION OF STS

Previous techniques to evaluate the STS have included the use of wearable and visual sensors. For instance, a waist-mounted triaxial accelerometer was able to classify activities such as sitting, lying, standing, running, as well as transitional activities such as STS and falling [Kang *et al.* [2010]]. Accelerometers have also been used to distinguish between subjects with Parkinson's disease and normal subjects with respect to their STS performance as part of the TUG test [Weiss *et al.* [2010]]. Although, sensor-based tests can be effective, the user is required to wear sensors when the test is being performed, which can be inconvenient. The preferred location of wearable sensors has been reported as the wrist, on glasses, or the arm [Cho [2019]]. In such cases, sensors are not good at detecting the movement of the entire body, such as that performed in the STS [Matsuyama *et al.* [2019]].

Other studies have used visual sensors to evaluate the STS movement. For instance, Allin et al. [Allin and Mihailidis [2009]] used three cameras to extract 3-D features such as the distance between the feet and head, to construct body centroids. Ellipsoid tracking was then used to follow the individual positions of the head, torso, and feet using the Weka Machine Toolkit for classification [Witten *et al.* [2016]], with an excellent correlation observed between measured rise time of the STS and the Berg Balance Score. However, the body parts for each subject had to be manually labeled for at least one image in order for the system to learn the color information for the individual. Moreover, three carefully positioned cameras were required to measure the STS time, such a system is difficult to use in an ecological setting. In another study, pose-based descriptors from volumetric image data were used to identify the STS movement, with features including the number of circles and the area of outer circles from each layer [Pehlivan and Duygulu [2011]]. The nearest neighbor method was used for classification of activity identification, including the STS. More recently, STS time was estimated using 3-D modeling of a human body in voxel space with an ellipse-fitting algorithm and image features used to capture the orientation of the body [Banerjee *et al.* [2013]].

The voxel height, in conjunction with the ellipse fit, gave the best segmentation accuracy for this method. Although, this framework could be used as part of a continuous video monitoring system in the homes of older adults and thus provide valuable information to help detect fall risk, it required two cameras to calculate human voxels. Furthermore, the accuracies of background subtraction are

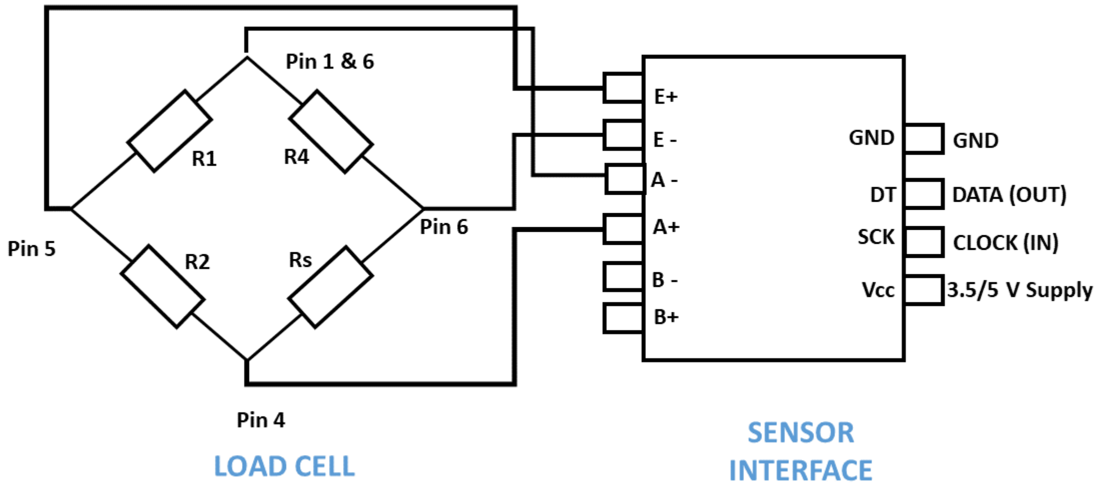


**Figure 6.1 :** Interface with load cell

highly dependent on the type of background, with a cluttered background leading to false silhouette extractions and thus a non-robust solution [Aggarwal and Cai [1999]].

Our goal in the proposed approach is to validate a framework that is suitable for continuous monitoring in an unstructured home setting, without requiring human intervention. This approach is described in the following section.

## 6.3 THE FRAMEWORK

In this section, we propose two new methods to estimate STS time and STS velocity during the STS movement. Firstly, an instrumented chair is designed using four load cells that eliminates the need of subjects to wear body sensors while performing the STS test. Next, we introduce a single RGB camera-based system to capture the STS movement and propose a technique to estimate STS time. A detailed description of both modules follows.

### 6.3.1 Instrumented Chair Design

An Instrumented wooden chair has been designed and comprehensively discussed in previous chapter. Four load cells weighted 40 kg with a precision of 8 g were fixed to the seat of the chair and covered by an piece of wood. Each pair of load cells on one side of the chair was connected to digital converter to convert analog data in to digital data for the analysis.ADC were connected to a microcontroller board (Arduino Mega 2560), with data acquired at 80Hz using a custom-built software (python) program(Figure 6.1). Vertical ground reaction force Fz was taken as the sum of the four ground reaction forces measured by the individual load cells.

Calibration of the chair was carried out for this study separately using a series of known masses, which were placed at different locations on the seat of the chair. This was used to verify the CoP and Fz

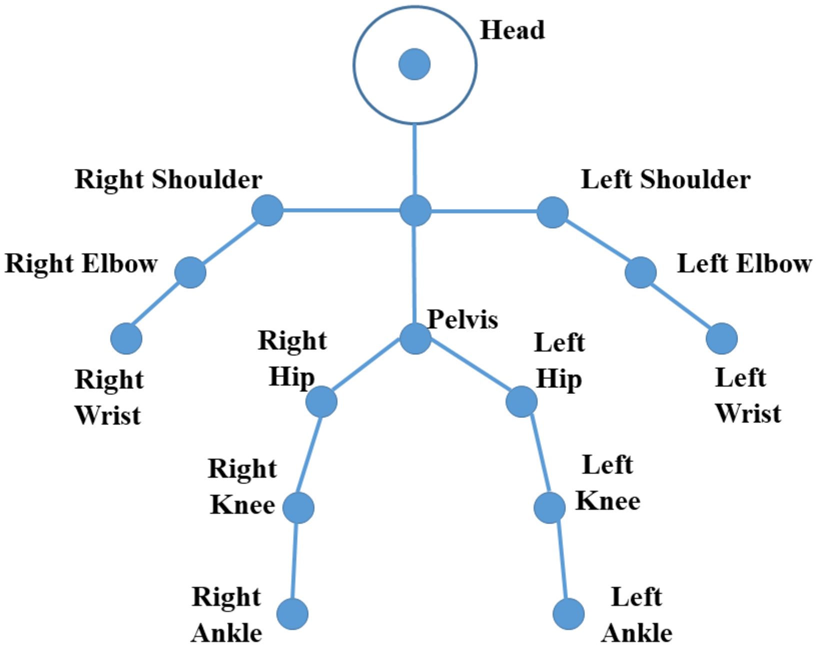


**Figure 6.2 :** The 15-segment model of a pose used to estimate the STS

data, with all values accurate to within the load cell manufacturer's specifications of ± 32 g for the mass and ± 1mm for the CoP.

### 6.3.2 Single Camera-based Posture Analysis

Cameras are readily available in the form of android devices or installed surveillance cameras. These visual sensors can be a useful resource in health care monitoring. Typically, multiple cameras are used in order to extract human silhouettes from video recordings [Banerjee *et al.* [2013],Allin and Mihailidis [2009]]. In the method developed for this study, only a novel single camera solution is used to calculate STS time.

A key step toward identifying people in the frames of a video is an accurate pose estimation. Given a single RGB image, we wish to determine the location of a human body. One way of accomplishing this is by background subtraction and extraction of the human silhouette. Although this technique is relatively simple, it gives false boundaries when the background is cluttered, while the silhouettes do not define body joints distinctively. In contrast, the precise pixel location of important key-points of the body, also referred to joint points, are required for an accurate clinical test [Chen *et al.* [2018]].

A well-established problem in the computer vision community is pose estimation, which has a variety of challenges to researchers. A good pose estimation method must be robust to occlusion, view angles, lighting conditions, clothing and background. With the advent of deep-learning techniques, many solutions to human pose estimation have been introduced, such as the recently-introduced Stacked Hourglass Network method [Newell *et al.* [2016]]. Poses estimated using this library have been shown to be accurate at assessing human movement [Jain and Harit [2017]].

For each frame the network estimates a pose with 15 joint locations (right ankle, right knee, right hip, left hip, left knee, left ankle, pelvis, neck, head, right wrist, right elbow, right shoulder, left

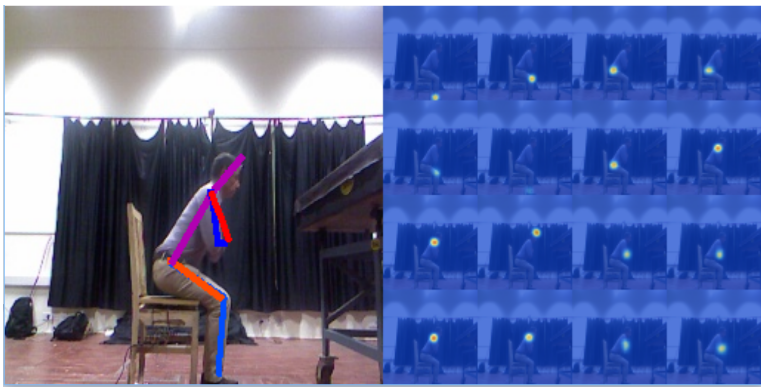

**Figure 6.3 :** Example of pose estimation during the STS movement

shoulder, left elbow, left wrist) as shown in(Figure 6.2). A sample estimation for a subject performing the STS is shown in Fig. 3, with the skeleton on the left and heat maps of joint estimation probability on the right.

Calibration of the camera was performed using the chair as a reference, with the back of the chair measuring 0.5m. This was used to ensure that the pixels within the image that covered the chair corresponded to 0.5m when the other measurements were taken. For all recordings, the camera was placed at a distance of 2.3 m on a line perpendicular to the front of the chair.

### 6.3.3 STS Parameter Calculation

The total time taken for each 5STS was estimated for each of the four recording systems. The method used to estimate STS time for both the RGB and Kinect systems was adapted from that of Ejupi et al. [Ejupi *et al.* [2016]]. This consists of an estimation of the head position obtained from the camera for the duration of the recording, with position data low-pass filtered with a 4th order Butterworth filter with a 2Hz cut-off frequency. The peaks identified were taken to be the middle of the standing positions while the troughs were taken to be the middle of the sitting positions. If the head position was within 5cm of the nearest peak the subject was considered to be standing, while a position within 5cm of the nearest valley was taken to be sitting. An example of head position signals during the 5STS for the RGB and Kinect systems is shown in (Figure 6.4)(a-b).

The mean duration of the 5STS was calculated for the force plate and the chair, as shown in(Figure 6.4) (c-d). For the force plate, the start of each sit-to-stand phase was taken to be $10\%$ of the peak force obtained during the transition to a standing position, which corresponds to the same ratio as the 5cm value used for the two camera-based systems when compared to the mean standing height of 50 cm. A subject was considered to be standing when their force reached $90\%$ of peak force for the individual STS. The standing phase of the STS was considered to have finished when vertical force decreased below $90\%$ of peak force, with subjects considered to have returned to a sitting position when vertical force reached $10\%$ of the previous peak.

In addition to total STS time, a worthwhile parameter that can be obtained from an instrumented STS is sit-to stand velocity, which has been shown to be able to distinguish better between fallers and non-fallers, than does total STS time [Ejupi *et al.* [2016]]. STS velocity was calculated for the two camera-based systems using the method proposed by Ejupi et al. [Ejupi *et al.* [2016]] for the period between the end of the sitting phase and the standing phase of each STS movement. The height change between these two points were divided by the time taken to obtain STS velocity. With respect to the force plate and the chair, velocity was derived using Newton's second law of motion, with STS

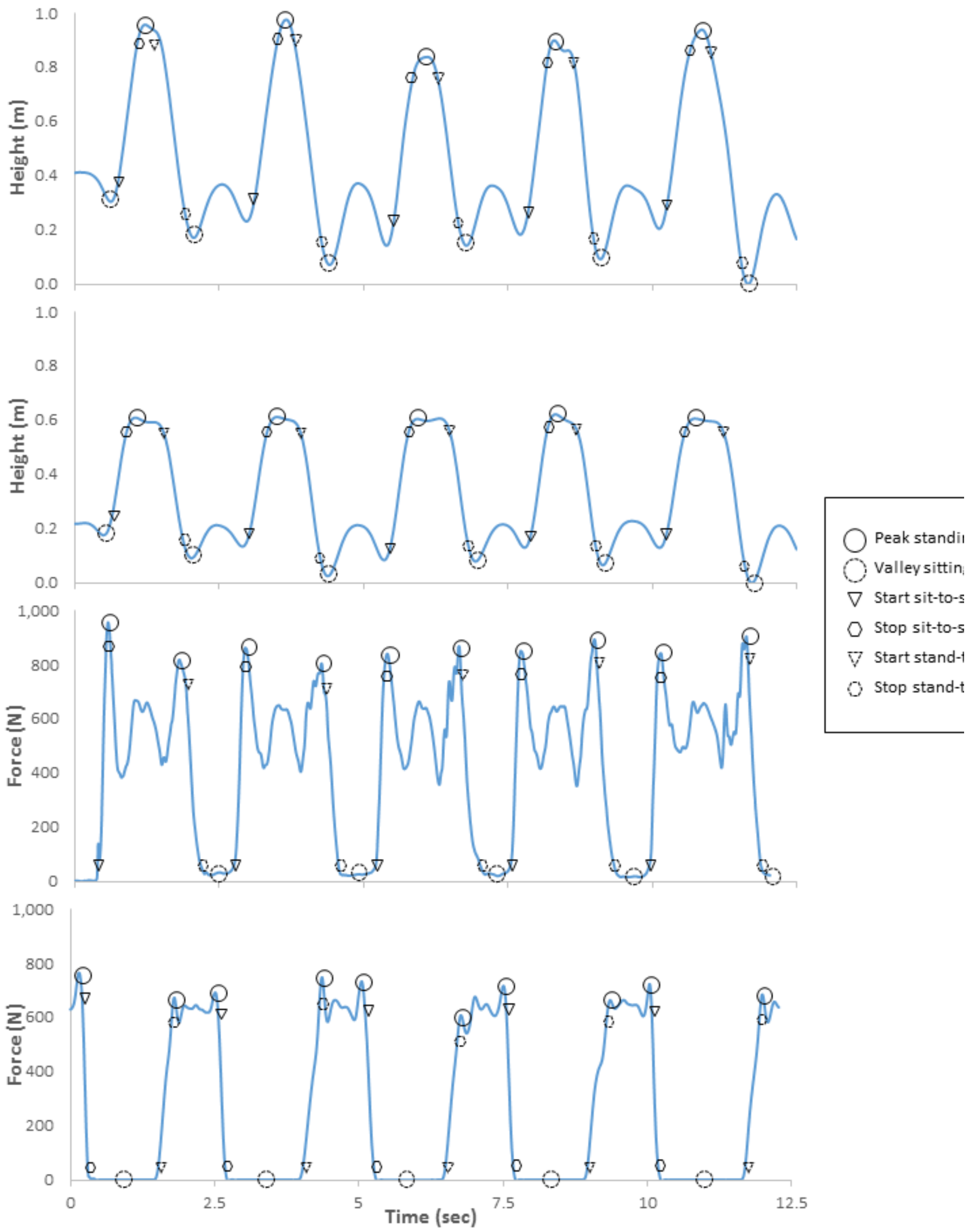


**Figure 6.4 :** Calculation of STS time and STS phases for RGB camera (a), Kinect (b), force plate (c) and chair (d).

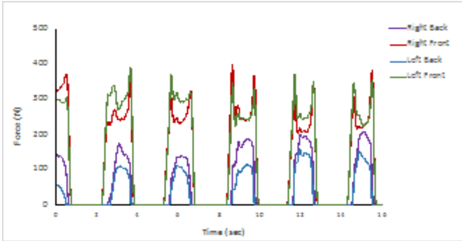


**Figure 6.5 :** Example recording from the instrumented chair during the 5STS test.

velocity calculated as between the time when force was between 10% and 90% of maximal force during the sit-to-stand movement. The average of STS velocity for the five STS movements was used in all subsequent analyses.

### 6.3.4 Comparison of STS Parameters

Comparative performances of the four methods of obtaining STS time and STS velocity were undertaken using correlation analysis and limits of agreement, using Bland-Altman plots [Bland and Altman [1999]]. Overall STS time was compared to a reference time that was obtained from analysis of a frame-by-frame record of each STS from the RGB camera [Banerjee *et al.* [2013]]. The expert manually identified the beginning and end of each STS, with the beginning taken to be when the subject began to move their torso forward in the first STS, while the end of the STS was estimated as the moment when the subject's torso returned to vertical after completing the 5th STS movement. These start and end points were chosen based on the four phases of the STS movement described previously [Millington *et al.* [1992]]. All four methods were compared with that of the expert for total 5STS time using Bland-Altman plots. With respect to STS velocity, no expert velocity was available, therefore Bland-Altman plots were not used.

All data processing was performed using custom-built software developed using LabVIEW (Version 2018, National Instruments Corporation, Austin, Texas, USA). Statistical analysis was performed using SPSS (version 25, IBM Corporation, Armonk, New York, USA).

## 6.4 RESULTS

The performance of the four systems was compared using data collected from a sample of 21 healthy younger subjects and a sample of 16 older fallers. The young subjects were aged 28.3 ± 6.8 years, weighed 67.2 ± 9.6 kg, of height 1.70 ± 0.04 m, with BMI 23.2 ± 3.0 kg/m2, while the fallers were aged 67.2 ± 6.7 years, weighed 64.3 ± 12.0 kg, of height 1.58 ± 0.07 m, with BMI 25.9 ± 4.2 kg/m2. The younger participants performed two trials, the first of which was at a self-selected slow speed, while subjects were instructed to perform the second trial as fast as possible. The older fallers performed a single trial at a self-selected speed. The ethics committee of the Asian Centre For Medical Education, Research & Innovation approved the study (ACMERI/18/001), with all subjects giving informed consent. A typical example recording from the instrumented chair is shown in (Figure 6.5).

### 6.4.1 Total STS Time

The performances of the four systems for young subjects for 5STS time against the expert time of 11.7 ± 2.1 s are shown in Table 6.1. The performance for 5STS time for the older fallers compared to the expert time of 18.0 ± 3.4 s is shown in Table 6.2. Bland Altman plots of the limits of agreement for the four methods for both groups of subjects combined when compared to the expert values are shown in

**Table 6.1 :** Performance of the testing systems for 5 STS time for young

| | Kinect | RGB | Force plate | Chair |
|---|---|---|---|---|
| Time (S) | 10.8±1.9 | 10.6±1.9 | 11.7±2.2 | 11.8±2.2 |
| Correlation | 0.990 | 0.997 | 0.979 | 0.995 |
| 95% CI | 0.975-0.996 | 0.993-0.998 | 0.948-0.991 | 0.988-0.998 |
| Error (S) | -0.84±0.35 | -1.01±0.23 | -0.05.8±0.33 | -0.16±0.17 |
| LOA (S) | 1.38 | 0.91 | 1.31 | 0.67 |
| LOA (%) | 90.9% | 95.5% | 97.7% | 97.7% |

[1]Times and mean errors are means ± SD; limits of agreement are range and percentage of points within this range. LOA: Limits of Agreement

**Table 6.2 :** Performance of the testing systems for 5 STS time for fallers

| | Kinect | RGB | Force plate | Chair |
|---|---|---|---|---|
| Time (S) | 17.6±3.3 | 17.7±3.6 | 17.8±3.5 | 17.7±3.1 |
| Correlation | 0.979 | 0.983 | 0.948 | 0.988 |
| 95% CI | 0.939-0.992 | 0.951-0.994 | 0.854-0.982 | 0.965-0.996 |
| Error (S) | -0.38±0.50 | -0.30±0.51 | -0.19±0.79 | -0.18±0.17 |
| LOA (S) | 1.97 | 1.98 | 3.08 | 0.67 |
| LOA (%) | 93.8% | 93.8% | 100 % | 100% |

[2]Times and mean errors are means ± SD; limits of agreement are range and percentage of points within this range. LOA: Limits of Agreement

Figure 6.6).

All systems performed satisfactorily, with the best performance obtained for the chair, which had the highest correlation, narrowest error SD, the narrowest range for the limits of agreement, and the highest percentage of points within this range.

### 6.4.2 STS velocity

Comparisons for STS velocity are shown in Table 6.3 for younger participants and in Table 6.4for the older fallers. The two camera-based systems obtained higher velocities than the two sensor-based systems. When the younger and older faller results were compared, greater discrepancies for a given system were observed for the two camera-based systems than for the two sensor-based systems, with lower correlations and higher mean differences, especially for the fallers. A comparison of the STS velocity measures from the four devices were made with gait velocity for the group of older fallers. The highest correlation with gait velocity was obtained for chair STS velocity (r=0.76), followed by the force plate (r=0.49), RGB camera (r=0.12), and the Kinect (r=0.07).

**Table 6.3 :** Performance of the testing systems for STS Velocity for young

| | Kinect | RGB | Force plate | Chair |
|---|---|---|---|---|
| Velocity (m/s) | 0.94±0.16 | 1.21±0.22 | 0.89±0.20 | 0.74±0.20 |
| Correlation | 0.811 | | 0.905 | |
| 95% CI | 0.678-0.892 | | 0.832-0.947 | |
| Mean diff. (s) | -0.28±0.13 | | -0.16±0.09 | |

[3]Times and mean differences are mean±SD

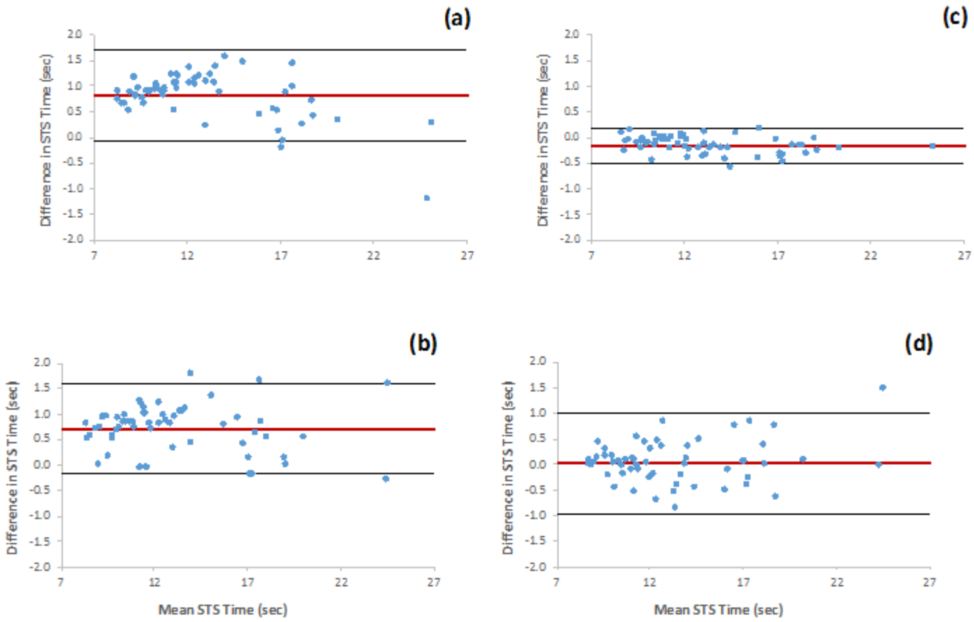


**Figure 6.6 :** Limits of agreement between expert STS time and the four systems; (a) RGB Camera; (b) Kinect; (c) Force plate; (d) Chair.

**Table 6.4 :** Performance of the testing systems for STS Velocity for fallers

| | Kinect | RGB | Force plate | Chair |
|---|---|---|---|---|
| Velocity (m/s) | 0.59±0.16 | 1.06±0.28 | 0.65±0.18 | 0.53±0.15 |
| Correlation | 0.574 | | 0.796 | |
| 95% CI | 0.109-0.833 | | 0.496-0.926 | |
| Mean diff. (s) | -0.47±0.23 | | -0.12±0.11 | |

[4]Times and mean differences are mean±SD

## 6.5 DISCUSSION

In this study, two new methods were proposed to evaluate the STS movement, which is an important functional screening tool in older people. One method used pose estimation from a single RGB camera, while the other method used an instrumented chair. The findings showed that both methods performed as well or better than previously reported methods in which a Kinect and a force plate were used to calculate total STS time and STS velocity, both of which can differentiate between fallers and non-fallers.

All four of the methods had excellent agreement with an expert estimation of STS in terms of the number of data points that fell within 2SD of the mean difference. However, both camera methods underestimated the total STS time compared to the expert by around one second. In contrast, the force-based methods had an average error within 0.2 sec of the expert time. This suggests that the newly developed chair can accurately detect STS time and could offer an alternative to a manual method. The reason for the differences is most likely due to the difficulty in using head height to detect the start of the STS, which begins with a forward movement of the trunk, before the second phase of vertical movement occurs [Millington *et al.* [1992]]. In order to improve the accuracy of the camera-based techniques, it might be necessary to include detection of forward movement, rather than the vertical movement described previously [Ejupi *et al.* [2016]]

The four methods were also compared with respect to the STS velocity parameter. In this case, it wasn't possible to use an expert for comparative purposes as it isn't possible to calculate velocity from a non-instrumented STS test. When the four methods were compared, differences in STS velocity were observed between the four systems. The RGB camera obtained higher values for the younger subjects than the other systems, followed by the Kinect, then the force plate, and finally the chair. When the older fallers were considered, the relative order of velocities differed slightly, with the Kinect values being lower than those of the force plate. The differences between the RGB and the Kinect would have been due to the differences in head height between sitting and standing for the two systems, which differed by around one third due to the scale used for each system.

For all four systems, lower STS velocities were observed for the older fallers, as would be expected. When the estimates of STS velocity for the four devices were compared with measures of physical function, the two force-based measures performed far better than those based on visual analysis. A high correlation was found with gait velocity for the chair, with a moderate correlation for the force plate, suggesting that the force-based measures might be superior, although a larger study is needed to confirm this finding. The differences between the two camera-based systems and the sensor-based systems could be further investigated using an opto-electronic system to record human movement, such as the Vicon, which would confirm the accuracy of the four systems presented here and act as a gold standard.

Results of this study show that the chair could be used to evaluate the STS in clinical settings, providing a potentially cheaper alternative than a force plate. The total cost of the components in the chair was approximately 100 $, which although not the commercial cost of a final product, would be substantially cheaper than a standard force plate, which typically cost thousands of dollars. The use of a force plate would also require the chair to be placed in a specific position in front of the force plate, which would make the protocol more complex than when using a chair with built-in sensors. In addition, when detecting STS time it was also possible to estimate STS velocity, which has been shown to distinguish between older controls and those with a history of hip fracture [Houck *et al.* [2011]]. It would also be possible to estimate the power produced during the STS using the method proposed by Lindemann et al., in which the difference between seated height and standing height is combined with rate of force development to estimate power [Lindemann *et al.* [2007]]. Power during the STS has been shown to be a strong predictor of overall muscle power and even cross-sectional area of the quadriceps [Smith *et al.* [2010],Takai *et al.* [2009]], which means the instrumented chair might be able to estimate muscle mass.

This study also has some limitations. Firstly, the system has thus far only been tested on young healthy subjects and one cohort of older fallers, so needs to be validated on a wider range of older participants to determine the predictive ability of the system in terms of other conditions associated with ageing, such as frailty and sarcopenia. In addition, the present chair design does not enable all of individual phases of the STS to be detected, particularly when the user is no longer in contact with the chair. The absence of a gold standard against which to compare STS velocity obtained from the four systems is also a limitation. Finally, the analysis performed was not automated, which would make the tests more widely applicable in clinical settings.

The limitations in terms of STS phase detection will be addressed in future work using infrared sensors to detect body position and joint angles, such as hip flexion. Future work could also examine whether a fusion of both chair and RGB systems would be of benefit. Finally, it would be worth evaluating whether the system could predict muscle power and/or muscle mass, as has been demonstrated by previous work with STS test [Smith *et al.* [2010],Takai *et al.* [2009]].

## 6.6 CONCLUSION

This chapter presented the development of two novel systems to evaluate the STS movement. The instrumented chair performed the best at detecting the STS when compared to an expert, with encouraging results also obtained when STS velocity was compared to physical function. Future work will use additional sensors to estimate muscle power during the STS.

...