2
Systematic Review of the Instrumented Sit-to-Stand Test

Results of this chapter published in Reviews in Bioengineering. The version presented here is similar to that submitted for publication with suitable addition of comments related to the thesis. The citation for this work is presented below:

Shukla, B.K, Bassement, J.N.C., Vijay, V., Yadav, S.K., Hewson, D.J. Instrumented analysis of the sit-to-stand movement for geriatric screening: A systematic review. Bioengineering.

2.1 INTRODUCTION

As stated in the previous chapter, STS test plays a vital role in the diagnosis of sarcopenia. Different researchers have used different versions of the test, such as, 30 seconds STS, 20 seconds STS, 5STS etc. This chapter presents a systematic review in search of recent articles on STS test for sarcopenia, especially for older people.

In addition, we also expand our search to technological devices which are being used for testing. This is due to the reason that a number of investigators have recently begun to add technological devices to test in order to provide more precise information about the physical capacity of the person being tested rather than limit the information collected from functional screening tests to a simple measure of the time taken to perform the task. For instance, an instrumented version of the TUG is now widely used, with the qTUG recommended for use by the National Institute for Health and Care Excellence in the UK [Smith *et al.* [2016]]. Several studies have recently been published in which the standard STS has been augmented with the use of technology such as cameras or body-worn sensors [Van Lummel *et al.* [2013]]. There has been one previous review of instrumented STS, however this was performed in 2014, with many new studies published recently [Millor *et al.* [2014b]]. In addition, this review focuses only on motion sensor devices, with no studies of video technology included in the review, while the effectiveness of the iSTS were not evaluated with respect to any diagnostic accuracy. Therefore, the aim of this systematic review, is to identify whether an instrumented version of the STS offers a better alternative to a standard test to detect older people at risk of falling, frailty, and sarcopenia.

2.2 METHODS

2.2.1 Search strategy

The search for articles was carried out based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [Liberati *et al.* [2009]]. The electronic databases searched were MEDLINE, CINAHL, Web of Science, and IEEE Xplore. Hand searches of reference lists of the selected articles was used to identify other relevant studies. Publication date was limited to dates between 1994, when the STS was first published by Guralnik and colleagues [Guralnik *et al.* [1994]], to July 2019.

Searches were limited to title and abstract of articles, with the following key words used

1. Old people: Old OR geriatric OR senior (people OR adult OR person)

2. Sit to Stand: Sit-to-stand OR stand-to-sit OR chair stand OR STS OR 5STS OR 30STS

3. Technology: sensor OR instrument OR accelerometer OR gyroscope OR magnetom OR ICT OR device OR smartphone OR motion capture OR video OR Kinect OR camera.

Types of studies

All types of quantitative study designs were included in the review.

Types of participants

Studies were limited to community-dwelling people aged 60 years and older. Any studies in which participants with a specific disease or condition, other than frailty, sarcopenia, or being fallers, such as Parkinson's Disease or dementia, were excluded. Articles in which instrumented STS were developed, but not evaluated on older people were also excluded.

Primary outcomes

The primary outcome of the studies selected needed to include an evaluation of iSTS performance to discriminate between older people with and without one of three chosen health factors or conditions, fall risk, frailty, or sarcopenia.

Inclusion and exclusion criteria

Studies with subjects aged under 60 years or with a specific medical condition were excluded. The STS test performed needed to be a recognised variation of the STS, such as the 5STS or 30STS, with studies in which only the sit-stand transition was reported being excluded.

Data extraction

Articles were identified by the combined keyword searches for each database separately, with duplicates removed. The titles and abstracts of these articles were screened to identify relevant studies to retain for full-text screening. Any articles for which a full-text version could not be obtained was excluded from the review. Data was extracted from the full text versions of the studies separately by both reviewers, with information subsequently pooled. Data extracted included study design and characteristics, the variants of the STS, the technology used, and ability of the instrumented device to discrminate between participants with frailty, sarcopenia, or fallers from non-fallers.

Quality appraisal

The criteria adopted to appraise articles for this review was the method of Loney and colleagues [Loney *et al.* [1998]] and Sanderson and colleagues [Sanderson *et al.* [2007]], and modified by Payette and colleagues [Payette *et al.* [2016]], which is suitable for observational studies. This method uses ten criteria, with each one scored as zero or one, and the total score taken as an index of methodological quality. The questions used for this appraisal were:

(1) Are the recruitment sources described?

(2) Are the criteria for exclusion or inclusion well defined?

(3) Are required sample size calculations presented?

(4) Is the method of calculating the iSTS parameters clearly described?

(5) Is the evaluation procedure clearly described?

(6) Has the outcome measure for the health condition been adequately described?

(7) Does the results section present a minimum of descriptive information about the participants, such as age (mean, range or standard deviation) and gender?

(8) Is the statistical analysis for evaluation of the association between iSTS parameters and the healthcare condition described?

(9) Are effect sizes reported with measures of precision?

(10) Are the study's limitations adequately presented?

Studies that scored at least 5 out of 10 were considered to be satisfactory and were included in the review [Loney *et al.* [1998]].

2.2.2 Results

Article selection A PRISMA flowchart of the search is shown in (Figure 2.1). A total of 740 articles were retrieved from the databases searched, with a further six articles identified from other sources. After duplicates were removed, 679 articles remained for title and abstract screening, resulting in 624 articles being removed for not meeting the inclusion criteria. The remaining 55 articles underwent full-text appraisal, with 45 articles rejected. The remaining 10 articles were retained for the systematic review, with characteristics of the articles presented in. The details of the quality appraisal for each article are presented in Appendix 2.12.

Study characteristics

The 10 studies selected used six different approaches to obtain an instrumented STS. Eight studies used sensors that were attached to the body. Four studies from two research groups used Inertial Measurement Units (IMU), with three of the studies from the same research group, which used a single IMU placed on the third lumbar vertebra [Millor *et al.* [2017],Millor *et al.* [2013],Millor *et al.* [2014a]]. The remaining studies used five IMU, three of which were placed on the lower limbs, with the remaining two placed on the fifth lumbar vertebra and the sternum [Greene *et al.* [2014]]. Two studies from the same research group used two triaxial accelerometers placed on the thigh and sternum [Doheny *et al.* [2013], Doheny *et al.* [2011]], while one group used a hybrid device consisting of a triaxial accelerometer and a pressure sensor worn in a pendant around the neck [Zhang *et al.* [2015]]. The study in which a sensor was attached to the body used a linear position transducer that was attached to the belt by a cable [Vincenzo *et al.* [2018]]. The remaining two studies used a Kinect sensor placed perpendicular to the chair [Ejupi *et al.* [2016]], and four force plates integrated into the chair [Houck *et al.* [2011]].

In total, 1357 participants were included in the 10 studies, none of which evaluated people with sarcopenia. Five studies compared fallers and non-fallers, one of which used fallers with hip fracture, four studies compared frail, pre-frail and robust-participants classified using the Fried frailty phenotype [Batista *et al.* [2014]], while the remaining study used both fall and frailty classification in the same participant group [Greene *et al.* [2014]]. With respect to the STS test, three different versions were used in the 10 selected studies. Six studies focussed on the 5STS test, three studies used the 30STS, while one study used a less-common version of the STS in which only three repetitions were performed [Houck *et al.* [2011]]. All studies in which the 30STS was used contained participants with frailty, rather than falls as the health condition of interest.

Evaluation of fallers and non-fallers

<u>Temporal parameters</u>

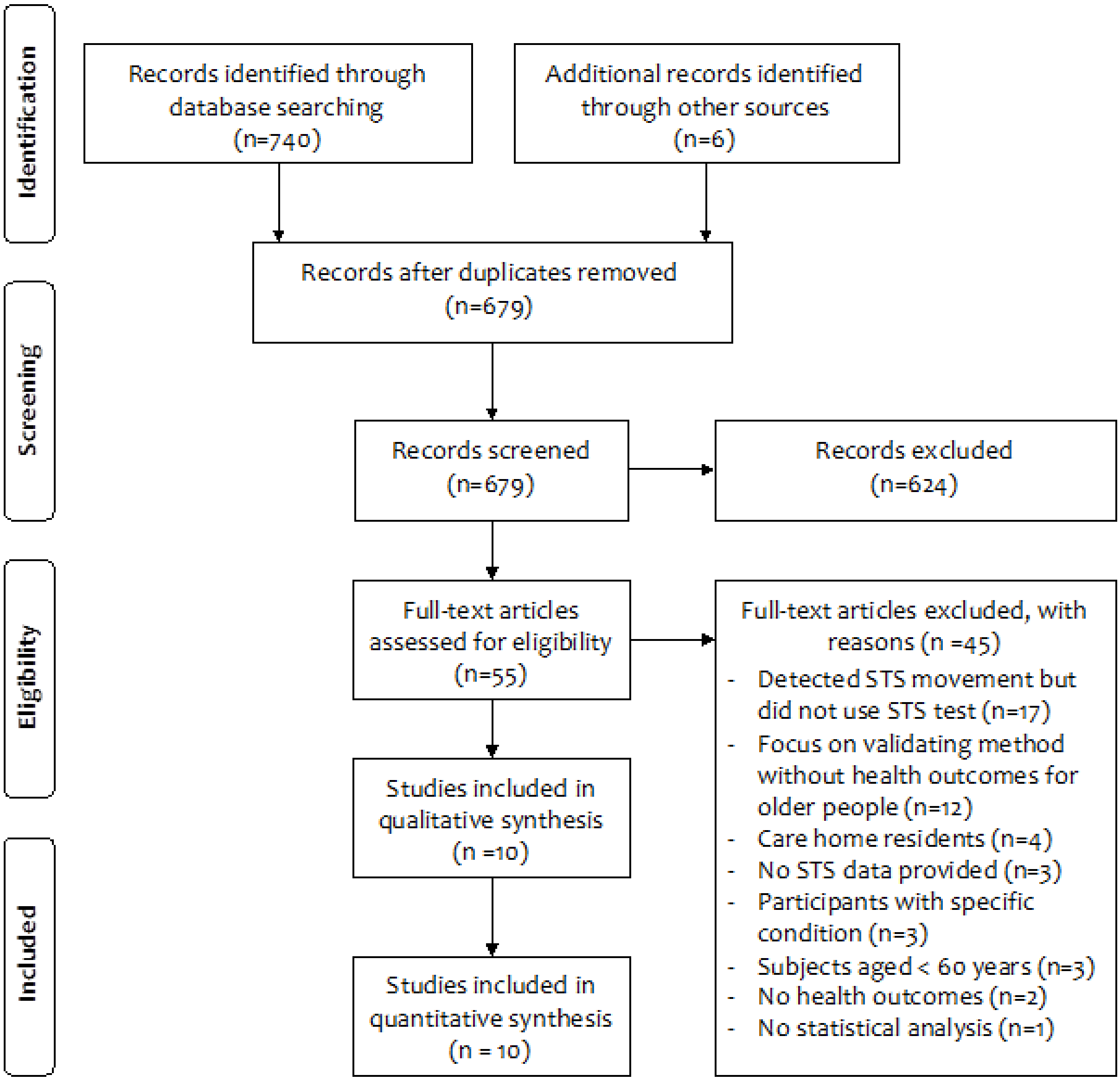

Figure 2.1 : Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) flow chart of study selection [17]

Table 2.1 : Characteristics of the articles accepted after full-text screening

Authors	Technology	Version	Subject	Age (years)	Condition	Quality
Doheney et. al (2011)	Triaxial accelerometers	5STS	40 (60%)	71.4 ± 7.3	Fallers	6
Doheny et al. (2013)	Triaxial accelerometers	5STS	39 (59%)	Fallers: 74.9 ± 7.0 Non-fallers: 68.4 ± 6.2	Fallers	7
Ejupi et al. (2016)	Kinect sensor	5STS	94 (70.2%)	79.7 ± 6.4	Fallers	9
Greene et al. (2014)	IMU	5STS	124 (73.4%)	Non-frail: 73.7 ± 6.0 Frail: 77.8 ± 6.4 Fallers: 76.0 ± 6.2 Non-fallers: 75.8 ± 6.8	Frailty, Fallers	8
Houck et al. (2011)	Force plates in a chair	3STS	28 (71.4%)	Control: 69.4 ± 10.9 Hip fracture: 76.4 ± 7.1	Fallers with injury	9
Millor et al. (2013)	IMU	30STS	47 (44.7%)	Frail: 85 ± 5 Prefrail: 78 ± 3 Robust: 54 ± 6	Frailty	6
Millor et al. (2014)	IMU	30STS	431 (N/S)	Frail: 79 ± 6 Pre-frail: 73 ± 5 Robust: 74 ± 5	Frailty	5
Millor et al. (2017)	IMU	30STS	431 (N/S)	Frail: 79 ± 6 Pre-frail: 73 ± 5 Robust: 74 ± 5	Frailty	6
Vincenzo et al. (2018)	Linear position transducer	5STS	98 (62.2%)	77.5 ± 7.3	Fallers	9
Zhang et al. (2017)	Triaxial accelerometer, air pressure sensor	5STS	25 (80.0%)	79.7 ± 5.7	Frailty	7

The results for fallers and non-fallers with respect to performance in the standard parameters obtained from the 5STS are shown in Table 2.2. Only one study reported a significant difference between fallers and non-fallers [Ejupi *et al.* [2016]], with this difference corresponding to a small effect using the scale of [Cohen [2013]]. When the time to complete different phases of the STS was compared, the only significant difference observed was for sit to stand time in one of the three studies [Doheny *et al.* [2011]], with this difference corresponding to a small effect Table 2.3.

<u>Other parameters</u>

A range of different parameters was calculated, which has been classified as force/power, frequency, and velocity. The results of these parameters in terms of differences between fallers and non-fallers is shown in Table 2.4. Thirteen of these 19 comparisons showed significant differences between the two groups, with three large effects found from one study for rate of force development in the injured leg of fallers with hip fracture, peak Ground-Reaction Force (GRF) and area under the force curve [Houck *et al.* [2011]]. In addition, eight moderate effects including RFD bilaterally [Houck *et al.* [2011]], average velocity [29] and sit-to-stand velocity [Ejupi *et al.* [2016]].

Several studies evaluated the association between STS parameters and functional capacity using tests of strength, balance and mobility. In total, five large correlations were reported, along with 17 moderate correlations Table 2.5. The largest correlations were found for force/power variables [31]

Table 2.2 : Comparison of fallers and non-fallers using 5STS time

Authors	Fallers	Non-Fallers	Cohen's d	p
Doheny et al. (2011)	17.18 ± 4.70	15.96 ± 4.10	0.28	0.160
Doheny et al. (2013)	17.56 ± 4.70	15.78 ± 3.84	0.42	0.200
Ejupi et al. (2016)	16.8 ± 5.68	14.33 ± 4.53	0.50	0.028

Table 2.3 : Comparison of fallers and non-fallers for different phases

Authors	Phase	Fallers	Non-Fallers	Cohen's d	p
Doheny et al. (2011)	Sit to stand to sit	2.34 ± 0.68	2.21 ± 0.64	0.20	0.330
Doheny et al. (2011)	Stand to sit	0.54 ± 0.19	0.55 ± 0.19	0.05	0.740
Doheny et al. (2013)	Stand to sit	0.36 ± 0.20	0.34 ± 0.17	0.11	0.710
Doheny et al. (2011)	Sit to stand	0.49 ± 0.20	0.41 ± 0.20	0.42	0.040
Ejupi et al. (2016)	Sitting time	1.75 ± 0.88	1.46 ± 0.59	0.42	0.071
Ejupi et al. (2016)	Standing time	1.02 ± 0.38	0.85 ± 0.31	0.51	0.063

and for sit-to-stand velocity [Ejupi *et al.* [2016]]. Finally, three studies reported models with respect to discriminating or classifying between fallers and non-fallers [Greene *et al.* [2014], Doheny *et al.* [2013], Vincenzo *et al.* [2018]]. The results of these models that were developed using different methods are shown in Table 2.6 , with the best results obtained for peak velocity [29].

Evaluation of frailty sub-groups

Temporal parameters

In one study, the robust participants were not aged 60 years and above, therefore the results for this study have not been included for the frailty comparisons [Millor *et al.* [2017]]. Two of the studies used participants from the same study with different parameters for the iSTS, but the same results for the number of cycles performed in the STS [Millor *et al.* [2017], Millor *et al.* [2014a]]. Accordingly, the results of only one of these studies for this parameter are reported here. There was a significant difference between frail and pre-frail, with frail participants performing less STS cycles (Frail 6.24 ± 2.53, pre-frail 8.16 ± 2.42, d=0.79; $p<0.05$) [Millor *et al.* [2017]]. There was also a significantly greater number of STS performed by robust than pre-frail participants (Pre-frail 8.16 ± 2.42, Robust 9.86 ± 3.00, d=0.63; $p<0.05$) [22]. Both of these effects can be considered as moderate.

When the time taken for the specific phases of the STS were compared, all comparisons were significantly different between frail and pre-frail participants for all phases of the STS, with three large and one moderate effect observed Table 2.7. Likewise, all differences between phase durations for the comparison between pre-frail and robust participants were also significantly different, with small effects observed for all phases except for a single STS cycle, which had a moderate effect Table 2.8.

Other parameters

Results for the parameters calculated from the STS are shown in Table 2.9for the comparison between frail and pre-frail. In total four large effects were observed for the AUC and velocity, while six moderate and four small effects were also reported. The comparison between pre-frail and robust participants for the same parameters are shown in Table 2.10. Two large effects were observed, for acceleration and RMS, while six moderate and six small effects were also observed. Only one study examined the correlation between frailty scores and the iSTS, with a non-significant correlation of -0.16

Table 2.4 : Comparison of fallers and non-fallers using other parameters

Authors	Catogry	Variable	Fallers	Non-Fallers	Cohen's d	p
Houck et al. (2011)	Force / Power	RFD injured leg	22.8 ± 8.5	35.7 ± 12.4	1.21	0.008
Houck et al. (2011)	Force / Power	RFD Non injured leg	33.4 ± 11.1	32.8 ± 9.3	0.06	0.349
Houck et al. (2011)	Force / Power	RFD bilateral	53.2 ± 17.8	64.9 ± 17.2	0.67	0.089
Houck et al. (2011)	Force / Power	Power	4.4 ± 1.3	5.0 ± 1.8	0.38	0.256
Houck et al. (2011)	Force / Power	Area	1.16 ± 0.84	0.59 ± 0.21	0.93	0.021
Houck et al. (2011)	Force / Power	Peak GRF injured	4.50 ± 0.82	5.10 ± 0.54	0.86	0.004
Houck et al. (2011)	Force / Power	Peak GRF non-injured	5.30 ± 0.86	5.10 ± 0.41	0.30	0.371
Houck et al. (2011)	Force / Power	Arm impulse	1.20 ± 0.58	0.86 ± 0.58	0.59	0.040
Vincenzo et al. (2018)	Force / Power	Average power	4.06 ± 1.27	4.83 ± 1.54	0.54	0.011
Vincenzo et al. (2018)	Force / Power	Peak power	6.22 ± 1.92	7.32 ± 2.24	0.52	0.014
Doheny et al. (2011)	Frequency	Total SEF	1.87 ± 0.20	1.77 ± 0.21	0.44	0.020
Doheny et al. (2011)	Frequency	Mean sit-stand-sit SEF	3.73 ± 1.20	3.20 ± 1.04	0.48	0.020
Doheny et al. (2011)	Frequency	Mean stand -sit SEF	15.13 ± 3.50	13.30 ± 3.81	0.50	0.010
Doheny et al. (2011)	Frequency	Mean sit-stand SEF	13.12 ± 4.80	11.29 ± 3.11	0.45	0.020
Doheny et al. (2011)	Frequency	Mean SEF AP sit-stand-sit	3.91 ± 1.80	3.38 ± 1.02	0.36	0.270
Vincenzo et al. (2018)	Velocity	Average velocity	0.41 ± 0.13	0.50 ± 0.16	0.60	0.007
Vincenzo et al. (2018)	Velocity	Peak velocity	0.64 ± 0.20	0.74 ± 0.23	0.46	0.017
Ejupi et al. (2016)	Velocity	Sit-stand	0.78 ± 0.20	0.94 ± 0.24	0.70	0.019
Ejupi et al. (2016)	Velocity	Stand-sit	0.65 ± 0.20	0.76 ± 0.22	0.51	0.151

Table 2.5 : Correlation between STS parameters and physical function

Authors	Category	Variable	Functional test	r	p
Houck et al. (2011)	Force/Power	RFD injured leg	LEF	0.593	0.001
Houck et al. (2011)	Force/Power	RFD injured leg	Gait speed	0.579	0.002
Houck et al. (2011)	Force/Power	Peak GRF injured	LEF	0.517	0.004
Houck et al. (2011)	Force/Power	Peak GRF injured	Gait speed	0.499	0.008
Houck et al. (2011)	Force/Power	RFD bilateral	LEF	0.487	0.010
Houck et al. (2011)	Force/Power	RFD injured leg	Balance	0.433	0.061
Houck et al. (2011)	Force/Power	RFD bilateral	Gait speed	0.417	0.030
Houck et al. (2011)	Force/Power	Area	TUG	0.334	0.089
Houck et al. (2011)	Force/Power	RFD bilateral	TUG	-0.301	0.127
Houck et al. (2011)	Force/Power	Arm impulse	LEF	-0.307	0.286
Houck et al. (2011)	Force/Power	Peak GRF injured	Balance	-0.31	0.116
Houck et al. (2011)	Force/Power	Power	TUG	-0.323	0.101
Houck et al. (2011)	Force/Power	Peak GRF injured	TUG	-0.349	0.057
Houck et al. (2011)	Force/Power	RFD injured leg	TUG	-0.357	0.068
Houck et al. (2011)	Force/Power	Area	Gait speed	-0.443	0.021
Houck et al. (2011)	Force/Power	Area	LEF	-0.551	0.003
Ejupi et al. (2016)	Temporal	Sitting time	Knee extension	-0.304	<0.01
Ejupi et al. (2016)	Temporal	Total time	Knee extension	-0.316	<0.01
Ejupi et al. (2016)	Temporal	Standing time	Knee extension	-0.326	<0.01
Ejupi et al. (2016)	Velocity	Sit-stand	Knee extension	0.533	<0.01
Ejupi et al. (2016)	Velocity	Stand-sit	Knee extension	0.432	<0.01
Ejupi et al. (2016)	Velocity	Sit-stand	Reaction time	-0.321	<0.01

Table 2.6 : Results of classification models for fallers and non-fallers

Authors	Comparison	Model type	Variables	Accuracy
Doheny et al. (2013)	Faller vs non-faller	Logistic regression	STS time, SEF, RMS, stand-sit time	74.4%
Greene et al. (2014)	≥2 falls or injury vs ≤2 falls	SVM	STS time, SEF, RMS, stand-sit time	72.6%
Vincenzo et al. (2018)	Faller vs non-faller	Logistic regression	Peak power	OR 0.69 (0.51, 0.93)
Vincenzo et al. (2018)	Faller vs non-faller	Logistic regression	Average power	OR 0.78 (0.63, 0.96)
Vincenzo et al. (2018)	Faller vs non-faller	Logistic regression	Peak velocity	OR 0.02 (0.001, 0.39)
Vincenzo et al. (2018)	Faller vs non-faller	Logistic regression	Average velocity	OR 0.09 (0.01, 0.69)

Table 2.7 : Comparison of frail and pre-frail participants for temporal STS parameters

Authors	Phase	Fallers	Non-Fallers	Cohen's d	p
Millor et al.(2014)	Impulse	1.90 ± 1.00	1.31 ± 0.57	0.92	<0.05
Millor et al.(2014)	Stand-up	1.38 ± 0.50	1.13 ± 0.33	0.70	<0.05
Millor et al.(2014)	Sit-down	1.67 ± 0.52	1.30 ± 0.42	0.85	<0.05
Millor et al.(2014)	Single cycle	4.94 ± 1.77	3.76 ± 1.16	0.94	<0.05

Table 2.8 : Comparison of pre-frail and robust participants for temporal STS parameters

Authors	Phase	Pre-frail	Robust	Cohen's d	p
Millor et al.(2014)	Impulse	1.31 ± 0.57	1.07 ± 0.45	0.47	<0.05
Millor et al.(2014)	Stand-up	1.13 ± 0.33	0.96 ± 0.28	0.55	<0.05
Millor et al.(2014)	Sit-down	1.30 ± 0.42	1.10 ± 0.37	0.50	<0.05
Millor et al.(2014)	Single cycle	3.76 ± 1.16	3.12 ± 0.97	0.60	<0.05

Table 2.9 : Comparison of frail and pre-frail participants for the STS parameters

Authors	Catogry	Phase	Frial	Pre-frail	Cohen's d	p
Millor et al. (2014)	Acceleration	Max Peak Stand-up	1.91 ± 0.93	2.62 ± 0.92	0.77	<0.05
Millor et al. (2014)	Acceleration	Max Peak sit-down	2.47 ± 1.77	3.62 ± 2.18	0.54	<0.05
Millor et al. (2014)	Acceleration	Min Peak Stand-up	-1.74 ± 0.97	-2.47 ± 1.08	0.68	<0.05
Millor et al. (2014)	Acceleration	Min Peak sit-down	-1.25 ± 0.75	-1.90 ± 0.88	0.75	<0.05
Millor et al. (2014)	AUC	AUC Stand-up	1.01 ± 0.36	1.29 ± 0.34	0.82	<0.05
Millor et al. (2014)	AUC	AUC-Sit-down	0.34 ± 0.13	0.49 ± 0.18	0.86	<0.05
Millor et al. (2014)	Velocity	Peak Stand-up	0.48 ± 0.16	0.61 ± 0.16	0.81	<0.05
Millor et al. (2014)	Velocity	Peak Sit-down	-0.37 ± 0.13	-0.51 ± 0.17	0.85	<0.05
Millor et al. (2017)	Displacement	AP Impulse Phase	18.81 ± 9.60	22.01 ± 9.73	0.33	<0.05
Millor et al. (2017)	Acceleration	Stand-up	1.36 ± 0.37	1.28 ± 0.35	0.23	<0.05
Millor et al. (2017)	Acceleration	Sit-down	1.21 ± 0.37	1.10 ± 0.39	0.28	<0.05
Millor et al. (2017)	Power	Stand-up	88.4 ± 50.8	65.4 ± 40.2	0.55	<0.05
Millor et al. (2017)	RMS	AP Impulse Phase	1.53 ± 0.70	1.18 ± 0.52	0.64	<0.05
Millor et al. (2017)	Power regularity	Up and down	0.68 ± 0.14	0.63 ± 0.16	0.32	<0.05

reported for the association between chair rise peak power and the Groningen Frailty Index [Zhang *et al.* [2015]]. Finally, two studies classified participants into frailty categories using STS parameters [Millor *et al.* [2017], Greene *et al.* [2014]], with the results shown in (Table 2.11). These results include separate classifications for each frailty level and also for the number of STS cycles, gait velocity, and a model containing four parameters [Millor *et al.* [2017]]. When these results were combined for all frailty levels, the least accurate was the number of cycles (52.9%), followed by gait velocity (67.2%), with the best classification obtained for the model using iSTS parameters (89.9%).

2.3 DISCUSSION

Overview

The aim of this systematic review was to determine whether an instrumented version of the STS offers a better alternative than a standard STS with respect to detection of older people at risk of falling, frailty, and sarcopenia. A total of 10 articles were identified, all of adequate quality, with six evaluations of fallers and five evaluations of frailty. None of the articles selected used an iSTS to compare differences between older people with and without sarcopenia. In one respect, this could be expected given that the STS is one component of the SPPB, which is one of the tests of physical function used in sarcopenia screening such as the updated version of the EWGSOP sarcopenia algorithm, discussed in Chapter 1. On the other hand, given that the STS is already used for sarcopenia screening, the use of iSTS parameters related to power and velocity might offer an alternative to some of the other tests already used, thus decreasing the time needed for screening.

iSTS and Fallers

Table 2.10 : Comparison of pre-frail and robust participants for the STS parameters

Authors	Category	Phase	Frail	Pre-frail	Cohen's d	p
Millor et al. (2014)	Acceleration	Max Peak Stand-up	2.62 ± 0.92	3.10 ± 1.03	0.49	<0.05
Millor et al. (2014)	Acceleration	Max Peak sit-down	3.62 ± 2.18	4.43 ± 2.67	0.33	<0.05
Millor et al. (2014)	Acceleration	Min Peak Stand-up	-2.47 ± 1.08	-3.23 ± 1.39	0.61	<0.05
Millor et al. (2014)	Acceleration	Min Peak sit-down	-1.9 ± 0.88	-2.44 ± 1.11	0.54	<0.05
Millor et al. (2014)	AUC	AUC Stand-up	1.29 ± 0.34	1.44 ±0.36	0.43	<0.05
Millor et al. (2014)	AUC	AUC-Sit-down	0.49 ± 0.18	0.58 ± 0.17	0.51	<0.05
Millor et al. (2014)	Velocity	Peak Stand-up	0.61 ± 0.16	0.68 ± 0.16	0.44	<0.05
Millor et al. (2014)	Velocity	Peak Sit-down	-0.51 ± 0.17	-0.59 ± 0.16	0.48	<0.05
Millor et al. (2017)	Displacement	AP Impulse Phase	22.01 ± 9.73	25.8 ± 12.0	0.34	<0.05
Millor et al. (2017)	Acceleration	Stand-up	1.28 ± 0.35	1.01 ± 0.37	0.75	<0.05
Millor et al. (2017)	Acceleration	Sit-down	1.10 ± 0.39	0.79 ± 0.30	0.89	<0.05
Millor et al. (2017)	Power	Stand-up	65.4 ± 40.2	38.1 ± 34.8	0.72	<0.05
Millor et al. (2017)	RMS	AP Impulse Phase	1.18 ± 0.52	0.78 ± 0.46	0.81	<0.05
Millor et al. (2017)	Power regularity	Up and down	0.63 ± 0.16	0.51 ± 0.21	0.65	<0.05

Table 2.11 : Results of classification models for frailty categories

Authors	**Test**	**Comparison**	**Model type**	**Variables**	**Accuracy**	**AUC**
Greene et al. (2014)	5STS	Frail/Pre-Frail vs Robust	SVM	STS time, SEF, RMS, stand-sit time	76.7%	-
Millor et al. (2017)	30STS	Robust	Decision tree	Number of cycles	55.6%	0.650
Millor et al. (2017)	30STS	Pre-frail	Decision tree	Number of cycles	47.2%	0.531
Millor et al. (2017)	30STS	Frail	Decision tree	Number of cycles	75.0%	0.657
Millor et al. (2017)	30STS	Robust	Decision tree	Gait velocity	69.4%	0.763
Millor et al. (2017)	30STS	Pre-frail	Decision tree	Gait velocity	62.5%	0.484
Millor et al. (2017)	30STS	Frail	Decision tree	Gait velocity	84.7%	0.850
Millor et al. (2017)	30STS	Robust	Decision tree	Imp phase, acceleration and power in Sit-Stand phase, impulse Stand-Sit phase	100.0%	1.000
Millor et al. (2017)	30STS	Pre-frail	Decision tree	Imp phase, acceleration and power in Sit-Stand phase, impulse Stand-Sit phase	83.3%	0.938
Millor et al. (2017)	30STS	Frail	Decision tree	Imp phase, acceleration and power in Sit-Stand phase,impulse Stand-Sit phase	70.8%	0.936

The iSTS evaluations produced parameters of different types, of which the most common were temporal, power and force, velocity, and acceleration. With respect to fallers, the results of temporal parameters tended to be similar to those of the time taken for the entire STS. This could indicate that, rather than there being a particular phase of the STS such as the sit-to-stand transition that is difficult for fallers to perform, it might be equally difficult to perform all phases. However, another reason for the lack of differences observed could be that many different methods of dividing the STS into phases were used, while some studies did not segment the STS at all [Zhang *et al.* [2015], Vincenzo *et al.* [2018]]. The most phases identified in the STS was four [Ejupi *et al.* [2016]], which included sit-to-stand transition, stand-to-sit transition, and standing and sitting phases. The smallest number of phases used was two although one of these methods only used the sit-to-stand and stand-to-sit transitions, discarding the time spent in between [Greene *et al.* [2014],Doheny *et al.* [2011],Doheny *et al.* [2013]], while the other method contained only the preparation for the sit-to-stand transition and the actual transition, without analysing the stand-to sit [Houck *et al.* [2011]]. It seems clear that, in order to compare studies more readily, it would be worthwhile standardising the phases used in order to make it easier to compare results between studies.

One potential benefit of using an iSTS is the possibility of identifying differences related to force, power, velocity and acceleration, which cannot be obtained from the STS. This was shown by the results of this review in which consistently good results were obtained for power and velocity parameters. This was particularly the case for the sit-to-stand transition, when power needs to be developed to ensure the person can stand up. Indeed, previous work has already shown that lower-limb power developed can be predicted using a logistic regression equation that contained STS performance body mass [Smith *et al.* [2010]]. Other work has shown that an equation using STS time, leg length and body mass can predict knee extensor strength and even the cross-sectional area of the quadriceps [Takai *et al.* [2009]]. Given that these models predict muscle power based only on the time taken to perform an STS test or the time taken to complete a number of cycles, it would be interesting to investigate whether the power directly calculated from an iSTS would give better results than the predictive equations of these studies.

With respect to the technologies used to instrument the STS in fallers, the majority of studies used body-worn sensors that were either IMU or triaxial accelerometers. Other methods used included force plates, a linear transducer, and several studies by the same research group using a Kinect sensor. When different techniques are compared, the best results in terms of differences between fallers and non-fallers were found for the chair equipped with force plates [Houck *et al.* [2011]], although it should be noted that these results were for fallers with hip fracture, so further work is needed for this technology with different subject groups. Classification of fallers and non-fallers was only reported by two studies, one on which did not include accuracy, making it difficult to compare. Although, the accuracy from an IMU-based system was only moderate [Greene *et al.* [2014]], an excellent odds ratio was obtained with the linear transducer [Vincenzo *et al.* [2018]] to distinguish fallers from non-fallers.

iSTS and Frailty

The majority of the frailty results were from the same research group, with the other two studies only providing comparisons between frailty categories and a single correlation between the iSTS and the overall frailty score. This means that the results for frailty were almost exclusively from one group using the Kinect sensor and the 30STS [Millor *et al.* [2017],Millor *et al.* [2014b]Millor *et al.* [2014a]], rather than the 5STS, which was used by the studies on fallers. In the case of the 30STS, rather than the time taken to perform a test, the number of STS cycles performed in 30 seconds is used to define STS. Frailty comparisons were then made between the three categories of frail, pre-frail and robust, with moderate effect sizes overserved for the number of STS cycles between frail and pre-frail and between pre-frail and robust groups. However, when the time, spent in the three phases of the STS used in this analysis, was considered, larger differences were observed between frail and pre-frail categories than between pre-frail and robust. The other iSTS parameters had similar results, with a greater number of large effects observed between frail and pre-frail than between pre-frail and robust.

When the iSTS was used in two studies to classify people into frailty categories. The best classification was obtained in a study using the Kinect sensor, which had 90% accuracy when using a decision tree model with iSTS parameters, which was much better than the 53% classification for the number of STS cycles alone. Obviously, a model in which multiple parameters are used is more likely to perform well than a single-factor model, but nevertheless, the result demonstrates the utility of the iSTS approach.

However, there are some advantages on sensors that are not body-worn in terms of acceptability and convenience. For instance, the requirement to place the sensors on the body means that issues of acceptability and ease of placement need to be considered, as well as the type of clothes worn by the person being evaluated. There has been little research into user preferences in this area, although in one report, the preferred location for wearable sensors was reported to be the wrist [Cho [2019]], which could be less effective at detecting whole body movement of the STS [Matsuyama *et al.* [2019]]. It is also possible that results would differ if sensors were incorrectly placed on the body, which would be particularly relevant for comparison of results between individuals and across sessions. In contrast, sensors that do not need to be worn such as the Kinect, might be quicker and easier for testing, including standardisation.

In addition to the articles chosen for this systematic review, there were many articles rejected at full-text screening as they did not include a specific version of the STS, and instead used a standalone STS movement meaning that the iSTS and STS could not be compared. Despite this, it would be worth examining the parameters developed in such studies to see whether some of the methods used could be incorporated into 5STS or 30STS tests for clinical use.

There were also many articles that were not selected as their focus was validating new methodologies. Although some of these articles were earlier work by authors whose articles were included in this review, many were from other research teams that could potentially provide valuable information on the STS when tested in populations of older people with falls history, frailty, and sarcopenia. Finally, one area that was not evaluated in any of the studies was changes between the different STS cycles within the STS test. For instance, in the 30STS, it might be worth comparing parameters from the first STS cycles with those from later cycles to determine whether any fatigue effect is present.

2.4 CONCLUSION

The results of this systematic review have identified the emergence of the instrumented STS as an area with great potential to improve the detection of strength-related conditions such as physical frailty, and to assess the risk of falling in older people. A range of parameters extracted from an iSTS appear to provide better differentiation between groups that the time taken for the STS alone. Future work in this area should focus on a greater standardisation of segmenting the STs into phases so that results can be compared more easily between studies. In addition, the development of more devices integrated into the chair could be beneficial to decrease any problems of standardisation of protocols and sensor placement for body-worn devices. This actually provides further motivation to design an integrated device for the diagnosis of sarcopenia.

Table 2.12 : Appendix 1

iSTS Criteria	Doheny et al. (2011)	Doheny et al. (2013)	Ejupi et al. (2016)	Greene et al .(2014)	Houck et al. (2011)	Millor et al. (2013)	Millor et al. (2014)	Millor et al. (2017)	Vincenzo et al. (2018)	Zhang et al. (2017)
Are the recruitment sources described?	1	0	1	1	1	1	0	1	1	0
Are the criteria for exclusion or inclusion well defined?	0	0	1	0	1	0	0	0	1	1
Are required sample size calculations presented?	0	0	0	0	0	0	0	0	0	0
Is the method of calculating the iSTS parameters clearly described?	1	1	1	1	1	1	1	1	1	1
Is the evaluation procedure clearly described?	1	1	1	1	1	1	1	1	1	1
Has the outcome measure for the health condition been adequately described?	1	1	1	1	1	1	1	1	1	1
Does the methods or results section present a minimum of descriptive information about the participants, such as age (mean, range or standard deviation) and gender?	1	1	1	1	1	1	1	1	1	1
Are the statistical analyses for evaluation of the association between iSTS parameters and the healthcare condition described?	1	1	1	1	1	1	1	1	1	1
Are effect sizes reported with measures of precision?	0	1	1	1	1	0	0	0	1	0
Are the study's limitations adequately presented?	0	1	1	1	1	0	0	0	1	1
TOTAL	6	7	9	8	9	6	5	6	9	7

…