Informatics in Medicine Unlocked 24 (2021) 100609

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Skeleton avatar technology as a way to measure physical activity in healthy older adults

Alisa Lincke [a,*], Cecilia Fagerström [b,d], Mirjam Ekstedt [b,c], Welf Löwe [a], Sofia Backåberg [b]

[a] Department of Computer Science and Media Technology, Faculty of Technology, Linnaeus University, Sweden
[b] Department of Health and Caring Sciences, Faculty of Health and Life Sciences, Linnaeus University, Sweden
[c] Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden
[d] The Research Section, Region Kalmar County, 39185, Kalmar, Sweden

ARTICLE INFO

Keywords:
Accelerometer
Balance
Machine learning
Older adults
Physical activity
Self-reported assessments

ABSTRACT

Background: Nowadays, self-reported assessments (SA) and accelerometer-based assessments (AC) are commonly used methods to measure daily life physical activity (PA) in older adults. SA is simple, cost-effective, and can be used in large epidemiological studies, but its reliability and validity have been questioned. Accelerometer measurement has proven valid to provide accurate and reliable measurement of everyday life physical activities regarding frequency, duration, and intensity in older populations, but is expensive and requires a long-time measurement. Here is, furthermore, a lack of well-defined and reliable accelerometer cut-off points to measure PA among older adults. Therefore, there is a need to develop a simple and reliable method to complement/replace self-assessment methods of daily life physical activity and facilitate the future development of cut-off points to measure daily life physical activities among older adults. In this study, we explore how skeleton avatar technology (SAT) can be used to measure PA among older adults.
Objectives: 1. To explore the association between accelerometer data and self-reported assessment data of daily life physical activities in older adults, and 2. To explore how the SAT of a standardized functional (balance) test can be used to measure daily life physical activity among older adults.
Method: The correlation analysis was used to explore the association between response variables, and deep neural networks were used to predict the response variables (AC and SA outcomes).
Results: The results indicate that there is a moderate ($r = 0.31$) significant ($p = 0.029$) correlation between AC of PA and SA of PA. The functional balance test assessed with SAT was able to predict AC with *3.89%* Mean Absolute Error (MAE), and SA with *11.07%* MAE.
Conclusion: Overall, these results indicate that one functional balance test measured with SAT can be used to predict PA outcomes measured with accelerometer devices. SAT can predict PA outcomes better than SA outcomes within the same population. More research is needed to explore the ability of SAT predicting PA among older adults with various functional abilities, and how SAT can be developed using 2D recordings, such as mobile phone recordings, to predict PA efficiently.

1. Introduction

Maintaining daily life physical activity (PA) among older adults has a significant impact on the quality of life, independent living, risk of falls, and cardiovascular [1] and metabolic health [2–4]. Levels of PA decrease with age [5,6], which has severe implications for the burden of chronic disease and mortality [7]. Thus, older people are recommended to have regular and various physical activities, 150–300 min per week at a moderate to vigorous intensity. In addition to that, older adults are recommended to perform physical activities to enhance balance three or more days per week and muscle-strengthening activities two or more days a week to prevent falls and prevent physical function (WHO 2020).

PA is a modifiable behavior that contributes substantially to maintaining functional capacity and health [7]. Thus, measures to prevent physical impairment and fall-related injuries for older adults are particularly important. Through regular assessment of PA, interventions could be initiated early, which might prevent mobility loss, improve quality of life, and prolong independent living among older adults. In

* Corresponding author.
E-mail address: alisa.lincke@lnu.se (A. Lincke).

https://doi.org/10.1016/j.imu.2021.100609
Received 6 April 2021; Received in revised form 17 May 2021; Accepted 17 May 2021
Available online 21 May 2021

community care, however, everyday rehabilitation and preventive work are not always prioritized alongside domestic care. The reason is that current assessment methods are insufficient, i.e., too costly, complicated, inaccessible, and/or inaccurate. The consequences may be that PA and function of older adults become less visible, and that early interventions are not being performed.

Measuring PA in older adults is therefore important to identify older adults at risk and prevent prolonged disease and injuries. Today, however, there is a lack of routinely repeated, reliable, valid, and inexpensive methods to measure PA among older adults and new methods need to be developed [8].

2. Related work

Current approaches to measuring PA are *self-reported assessments* (SA), and wearables capturing motion such as *accelerometers*. Several SA questionnaires exist measuring PA in older adults such as International Physical Activity Questionnaire [9], the Standard 7-day Physical Activity Recall Questionnaire [10], Zutphen Physical Activity Questionnaire [11]. Self-assessed physical activity may be completed through online questionnaires or during an appointment with, e.g., a physiotherapist or clinician and is commonly focused on PA performance the past days, weeks or months. The advantages of using SA are that it is cost-effective and that it can be used in large epidemiological studies to generate large data sets. However, some studies show limitations of SA methods for older adults, such as memory retention (cognition) [12,13], over-reporting [14], tendencies to overestimate/underestimate time spent on physical activity [12], influences by mood, depression etc. [15].

The limitations of accelerometers are manifold. They are costly in large-scale studies and they lack information about upper body movement due to attachment to the individual's hip or leg. Metabolic equivalent tasks (METs) are not 100% correct as they cannot detect whether a person is carrying any weight [16]. Established and reliable accelerometer cut-off values for older adults are missing [17]. For moderate to vigorous PA, the suggested cut-off value of 1.041 counts/-min [18] varies depending on the type of accelerometer and placement on the body [19]. A recent literature review [17] shows that the validity and reliability increase when combining accelerometry with an inclinometer for measuring sedentary behavior. Hence, they are currently mainly used in research studies [15]. Other studies have combined accelerometer with GPS sensors in order to measure PA in outdoor and indoor environments [20,21]. The results show a positive association and better health outcomes for those who are physically active both indoor and outdoor [21]. However, not all older adults feel safe in the outdoor activities, due to risk of falling [21]. In addition, bad quality of the GPS signals led to a reduced amount of useable data [20]. Thus, using extra sensors and devices does not necessarily reduce the complexity (data collection, data synchronization, and analysis) and the time effort to measure PA in older adults. In our study, we are interested in a low-barrier tool to measure PA with minimal time and effort.

Skeleton Avatar Technology (SAT) records human movements in 2D or 3D, estimates the position of joints in the movement recording, and analyzes the resulting stick figure sequences with artificial intelligence (AI). 3D SAT enables the efficient and sufficient assessment of human movements, e.g., for detecting the risk of falls among older adults [22] and the assessment of balance and postural control [23]. It can also be used for an accurate, objective, and automated assessment of movement quality [24–26].

In our previous study [27], we have measured PA in 54 healthy older adults using SA, and *functional abilities* (mobility and balance) using functional tests (FT) and SAT. We explored the association between FT, SAT, and SA. It appears that *functional balance* assessed with SAT was a good predictor to predict functional tests assessments, such as 30-s Chair Stand Test (30sCST), and the 4-Stage Balance Test (4SBT). While the previous study has focused on the assessing *functional abilities* (mobility and balance), in this study we seek a new solution to routinely measure *physical activity* in older adults, with minimal efforts in time and complexity. Thus, the objectives of this follow-up study are to (1) explore the association between *accelerometer data* (AC) and self-reported assessment data (SA) of daily life physical activities in older adults, and (2) how SAT of a standardized functional (balance) test can be used to measure daily life physical activity (PA) in older adults.

In the literature, PA is defined as body movement produced by the construction of skeletal muscle that results in energy expenditure (EE) [28]. The total amount of EE depends on the type of activity (such as walking, running, swimming, etc.), its frequency, intensity, and duration [29]. PA can be measured in terms of EE and/or METs with the use of accelerometer devices. Many studies have shown the validity and reliability of the use of accelerometer/motion sensors as a measurement tool for sedentary behavior (SB) and PA in older adults [17,18,30–35]. The reliability of the accelerometers were tested in terms of intra-class correlation coefficient (ICC) for continuous data, and show an acceptable level of reliability (ICC = 0.80) in measuring SB in older adults [17,36]. The validity of accelerometers refers to accurately assessing PA and/or EE [17]. Previous studies have shown the moderate correlation ($0.48 > r > 0.63$) between EE measured by the doubly labeled water (DLW) method and accelerometers [35], and high accuracy (>80%) in measuring/classifying daily life activities (walking, standing, sitting, etc.) in older adults [17,31,37]. However, some other studies reported inaccurate estimation of EE in older adults by using accelerometers [35,38,39]. Most commonly used accelerometers that have been validated on older adults are ActiGraph (GT3X, GTM1 models), ActivPAL, and Triaxial research tracker (RT3) [16,17,19,40,41].

Much of the previous research has explored the relationship between self-reported PA levels, and everyday life activities (such as walking, running, sitting, etc.), PA in METs recorded by accelerometer in older adults [11,42–45]. They have found a moderate correlation between SA of PA and accelerometer-based PA with Spearman's correlation coefficient $r = 0.33$ [43], $r = 0.34$ [11], and $r = 0.28$ [45], respectively. Males tend to report 25% higher PA than females [44], whereas no significant difference in accelerometer-based PA was found. Also, people have reported less sedentary behavior and higher levels of vigorous-intensity PA compared with the accelerometer data [44]. Several studies showed that the correlation between total PA measured by the self-reported questionnaires and accelerometer data varied widely but to a lower degree than the acceptable level [11,44].

Only a few studies have investigated the association between functional balance measures and daily life physical activity among older adults [46–48]. One study provides evidence that PA positively affects balance outcomes in older healthy adults [49]. Another study [46] has found a strong association between functional balance measured by accelerometer and PA also measured by accelerometer. Thus, more research is needed to investigate the associations between balance and PA in the long term [49] and to find more factors that may influence physical activity [46].

A number of studies tackle the problem of defining MET cut points and accelerometer cut points to objectively assess physical activity in older adults [9,36,50–52]. The cut points are meaningful for clinicians to defining the patient's symptom/health status. Therefore, we support MET and accelerometer cut points research, as it is crucial for objectively measuring PA in older adults. However, it is challenging to create a general method to identify METs and accelerometer cut points for older adults that would be invariant to e.g., age, health, and chronicle disease. Therefore, in this study, we have used the continuation scores instead of a specific threshold/cut-point, where the higher values mean a person has a high PA and lower values mean the person has a low PA. There is a need to develop a simple and reliable method to complement/replace self-assessment methods of daily life physical activity and facilitate the future development of cut-off points to measure daily life physical activities among older adults. The development of SAT that might objectively measure physical activity and based on a short balance test should be beneficial.

3. Materials and methods

3.1. Study design and participants

The present study applied a cross-sectional design and was a secondary analysis of data collected from 54 older adults in our previous study [27]. Community-dwelling older adults (65+ years) were recruited via four pensioners' associations through emails/phone calls. In total, 54 older adults (38 females and 16 males) signed up for this study. All participants signed informed consent, and the study was approved by the Swedish Ethical Review Authority (Dnr: 2019–02553).

3.2. Data collection

Participants first completed a self-assessment questionnaire about PA (see Table 1), and a questionnaire with demographic information about their gender, age, weight, height, diagnosis, and symptoms (see Table 2).

The answers from the questionnaires were saved with Excel in comma-separated value (csv) format and further condensed to a self-assessed PA score, cf. Section 2.3, processed with Matlab version R2020a [53].

The ActivePAL (PAL Technologies Ltd, Glasgow, UK) device was attached to the participant's left or right thigh and used to collect daily life activities within 7 days. The device is small (35 × 53 × 7 mm), light (20 g), has a 10 Hz sampling frequency, and data with 15-s epochs. The ActivePAL device allows collecting the accelerometer raw data as well as different activities such as *stepping*, *sitting to stand transition*, *sitting*, *standing*, *walking*, *biking*, and *ActivPAL's activity score* in MET. However,

Table 1
Self-assessment questions of PA.

N	Question	Scores	Scores Description
1	How much do you sit and lie in total during a normal day if you count off sleeping time?	1 Never 2 1–3 h 3 4–6 h 4 7–9 h 5 10–12 h 6 13–15 h 7 Most all-day	score 0–6, the higher score the more sedentary lifestyle
2	How much time do you spend a regular week on everyday physical activity?	1 >300 min 2 150–299 min 3 90–149 min 4 60–89 min 5 30–59 min 6 <30 min 7 No time	score 0–6, higher score means more physical activity
3	How much time do you spend on a regular week for physical exercise that makes you feel short of breath?	1 >300 min 2 150–299 min 3 90–149 min 4 60–89 min 5 30–59 min 6 <30 min 7 No time	score 0–6, higher score means more physical activity
4	Currently, do you engage in regular physical activity?	Yes/no	0-no,1-yes
5	Have you been engaged in regular physical activity for the past 6 months?	Yes/no	0-no,1-yes
6	Do you have a training program? If yes: - How often do you do your exercises? - How long do you do your exercises? - How strenuous are you experiencing your exercise program?	Yes/no 1 No training program 2 Few times per week 3 Sometimes per week 4 Every day 1 No training program 2 Less than 30 min 3 More than 30 min 4 No training program 1 Not strenuous 2 Little strenuous 3 Moderate strenuous 4 Very strenuous	0-no,1-yes scores 0–3, higher score means more exercises/ longer exercise duration/ less strenuous exercise program

Table 2
Demographic, PA measured by self-assessment, and accelerometer data of participants (n = 54, missing 0–1.9%).

Variable	Men (n=16)	Women (n=38)	Total (n=54)
Age in years, Mean (SD)	75.6 (3.2)	73.7(4.7)	74.3(4.4)
BMI, Mean (SD)	26.2 (4.5)	25.9(5.4)	26.0(5.1)
PA measured by SA			
Sitting and lying, hours/day, n(%)			
Never	0	0	0
1-3 hours	4 (25.0)	3 (8.1)	7 (13.2)
4-6 hours	9 (56.3)	22 (59.5)	31 (58.5)
7-9 hours	1 (6.3)	8 (21.6)	9 (17.0)
10-12 hours	1 (6.3)	4 (10.8)	5 (9.4)
13-15 hours	1 (6.3)	0	1 (1.9)
Most all-day	0	0	0
Physical activity, minutes/week, n (%)			
>300 min	8 (50)	20 (54.1)	28 (52.8)
150-299 min	5 (31.3)	7 (18.9)	12 (22.6)
90-149 min	1 (6.3)	5 (13.5)	6 (11.3)
60-89 min	0	1 (2.7)	1 (1.9)
30-59 min	0	3 (8.1)	3 (5.7)
<30 min	2 (12.5)	1 (2.7)	3 (5.7)
No time	0	0	0
Exercise (strenuous activity), minutes/ week, n (%)			
>300 min	0	2 (5.4)	2 (3.8)
150-299 min	4 (25.0)	6 (16.2)	10 (18.9)
90-149 min	4 (25.0)	9 (24.3)	13 (24.5)
60-89 min	2 (12.5)	8 (21.6)	10 (18.9)
30 - 59 min	2 (12.5)	6 (16.2)	8 (15.1)
<30 min	3 (18.8)	3 (8.1)	6 (11.3)
No time	1 (6.3)	3 (8.1)	4 (7.5)
How often do you do your exercises? n (%)			
No training program	8 (50)	14 (41.7)	22 (45.0)
Few times per week	4 (25)	10 (29.4)	14 (29.0)
Sometimes per week	1 (6.6)	8 (23.5)	9 (18.3)
Every day	2 (13.3)	2 (5.88)	4 (8.2)
How long do you do your exercises? n (%)			
No training program	8 (50)	14 (50)	22 (45.0)
Less than 30 min	5 (33.3)	9 (26.4)	14 (29.0)
More than 30 min	2 (13.3)	11 (32.4)	11 (22.0)
How strenuous are you experiencing your exercise program?			
No training program	8 (50)	14 (50)	22 (45.0)
Not strenuous	0 (0.0)	1 (3)	1 (2.0)
Little strenuous	3 (20)	6 (17)	9 (18.4)
Moderate strenuous	3 (20)	13 (38)	16 (33.0)
Very strenuous	1 (6.2)	0 (0.0)	1 (2.0)
PA measured by ActivPAL			
Steps per week, in average	7956	7803	7917
Sit to Stand transitions per week, in average	43	44	44
Sitting hours per week, in average	8.7	8.7	8.7
Standing hours per week, in average	4	4	4
Walking hours per week, in average	1.6	1.6	1.6
Biking hours per week, in average	0.1	0.06	0.07

the ActivPAL's activity score values (MET) are significantly different from the criterion of oxygen uptake at various speeds (2–4 mph) [54] and have low accuracy in estimating EE in METs [55]. In addition, ActivPAL does not provide a transparent way of calculating METs from its accelerometer raw data. Therefore, the decision was taken to not use the activity score provided by ActivPAL in the machine learning analysis. As for SA and using the same approach, the daily life activity data collected by the accelerometers was exported into comma-separated value (csv) files with the help of the ActivPAL4[†] software and further condensed to an accelerometer PA score, cf. Section 2.3, processed with Matlab version R2020a [53].

3.3. Data preprocessing for response variables

In the data preprocessing step, we applied the following filters and transformations to the collected SA and accelerometer (raw csv) data: (a) removing columns that had zero variance; (b) normalizing the data using the (complementary) cumulative distribution functions (C)CDF; (c) aggregating the data to a common *self-assessment activity score* (SA), and *accelerometer activity score* (AC) using the joint CDF of the individual variable scores.

The normalized scores of the individual variables in SA and AC computed in step (b) are values between 0 for the lowest activity and 1 for the highest activity. Note that for variables such as walking and standing where lower observed values (in minutes/week or hours/day) indicate a lower activity and higher values indicate higher activity. For other variables such as sitting and lying, higher values (in hours/day) indicate low activity. To achieve a normalization of scores between 0 and 1 from the lowest to the highest contribution to activity, we use CDF for the normalization of variables positively correlated with activity (high variable value also means high activity) and CCDF for variables negatively correlated with activity (high variable value actually means low activity). Consequently, for SA, physical activity, exercises, training exercises, duration of exercises, and strenuous exercises are normalized with CDF, whereas sitting and lying are normalized using CCDF. For AC, stepping, sit to stand transitions, walking, biking, standing are normalized with CDF, and sitting, lying variables with CCDF. Details of this normalization are described and motivated in our previous study [27].

3.4. Data preprocessing for predictors

As mentioned earlier, in our previous study [27], the stage four (one-leg-stand) exercise of the 4SBT test was recorded with a Kinect camera and used as SAT data in this study. Each recorded exercise movement is a sequence of posture frames. Each posture frame encodes the subject's joint positions in 3D at a point in time of the recording. While there are 25 joint positions recorded, because of the low reliability of some joints, we only used the following 13: head, left/right shoulder, left/right elbow, left/right wrist, left/right hip, left/right knee, and left/right ankle.

Each frame is a record of *features*. A feature is called *direct* if it is directly measured by the 3D camera, and *indirect* if it is computed form direct or other indirect features. The direct features include the x, y, and z coordinates of 13 skeleton joints. Indirect features include the *angles* between different limbs and the axes of the 3D coordinate system.

Because number of recorded subjects was low, we applied well-known data augmentation technique [56] to artificially increase the number of training and test sequences. Therefore, we *stretched* each frame in the x and y directions by the same constant factors around 1, we *rotated* each frame around the y axis by the same constant angle around 0°, and we *mirrored* the frames. Cascading these transformations led to an increase in the number of individual movement sequences for machine learning by a factor of about 1000.

3.5. Statistical analysis

Descriptive statistics and Pearson's correlation analysis was conducted using Matlab version R2020a [53]. Significance was set at $p < 0.05$. Pearson's correlation coefficients r were used to determine the dependencies between the PA measured with SA and accelerometers. The correlation results were interpreted as *low* ($r < 0.30$), *moderate* ($0.30 = r < 0.60$), or *high* ($r \geq 0.60$). The correlation analysis between predictors (said 3D SAT data of 4SBT movements) and response variables (SA and AC scores) was not conducted due to the predictor's data type (time series over a multi-dimensional feature vector).

3.6. Machine learning

We applied the same *deep learning* approaches as in our previous study [27], here to map 3D SAT data to the SA and AC scores. The input data is the described sequence of records of *direct* features (x, y, and z coordinates of skeleton joints) and, optionally, of *indirect* features (angles between limbs and the axes of the 3D coordinate system).

We conducted each machine learning experiment twice, once with and once without indirect features included. Orthogonally to that, we conducted the experiments with the cut sequences (where frame sequence starts with lifting the leg and ends with setting it down again) and the uncut sequences (including some frames where subjects get into the position before starting the movement or left the scene, respectively). We cut the sequences automatedly using the winning dynamic time warping (DTW) approach reported in the comparison [57]. Each of these variants was tested twice again, once with the standardized, and once with the raw (direct and indirect) features. This results in altogether $2^3 = 8$ setups regarding the input data for each machine learning experiment.

The response/output variables are the normalized—normalization as described in subsection 3.3 (b)—daily life physical activity scores SA and AC, respectively.

Our machine learning experiments applied standard neural network technology [58] implemented in Python 3 using the Tensorflow framework [59]. We tested two principally different neural network architectures with roughly the same number of parameters to learn:

- *Model 1.* A *convolutional neural networ*k (CNN) with three 1D convolutional layers with a depth of 128, 64, and 32 neurons, respectively, and each followed by a 1D maximum pooling layer of size two and activated with a ReLU, followed by an output layer with a single output (the score) activated with a *sigmoid* activation function.
- *Model 2.* A *recurrent neural network* (RNN) with three long short-term memory layers (LSTM) of 32 neurons each, followed by an output layer, as in Model 1.

In all architectures, we used either dropout, with a rate of 0.5 in the first layer, or kernel and activation regularization (L2 norm, penalty of 0.001) of the first two layers in order to avoid overfitting, or both or none of the regularizations. This results in $2^2 = 4$ setups for each of the two the machine learning networks and, with the two input variants, in altogether $2 \times 4 \times 8 = 64$ experimental setups. All systematically tested variants and their parameter settings were experimentally selected based on the model performance in the previous study [27]. Beyond the simple grid-search described above, we did not manually fine-tune nor automatically optimize the variants and their parameters. Since, the sample size is rather small, we were rather interested in the principle predictive power of SAT on AC. For developing and optimizing a production-quality model, which is future work, a bigger dataset would be needed.

For each experiment (setup) we conducted 10-fold cross-validation. We randomly split the input sequences into 10 folds each about 10% of the data. We did not mix the augmented sequences, i.e., all transformed sequences remained in the same fold as its original. Each of the

48 experiments was conducted 10 times, once with each fold as test data and the remaining 9 folds as training data. We did not put aside additional validation data.

We selected the mean absolute error (*MAE*) for assessing model accuracy on the training and test data, respectively. As the SA and AC scores are between 0 and one, so is the theoretic $MAE \in [0,100\%]$. The machine learning results were interpreted as *good* ($MAE < 10\%$), *moderate* ($10\% = MAE < 20\%$), or *bad* ($MAE \geq 20\%$). The results reported are the average *MAE*s of the predictor functions applied to the test data using cross-validation.

4. Results

4.1. Descriptive statistics and correlation

Table 2 provides an overview of the collected data of the PA measured by the self-assessment questionnaire, and the ActivPAL accelerometer. Twenty-eight persons (52.8%) carried out leisure-time physical activities for more than 5 h/week. All but four persons (7.5%) stated that they performed moderate to high levels of physical activities every week. Six persons (11.3%) reported that they spent 10–15 h sitting or lying down every day. About half of the sample (52.8%) had an exercise program that they followed. How many times a week they performed the program varied, as did the length of the programs. Of those who followed an exercise program, 65.5% had a program that was 30 min or longer. Participants were also asked if they considered themselves physically active in general (71.7% did).

Table 3 presents descriptive statistics of the 7 days activities collected by the ActivPAL device. Outcome variables for daily life activities were measured in hours per day spent on sitting/lying, standing, stepping (counts/d), sit to stand transitions (counts/d), walking, and biking. On average 44 sit to stand transitions per day were recorded, 4 h per day standing time (16.6%), and sitting/lying 36% (8.7 h/day). Only a few minutes (0.07 h/day) were spent on biking, and around 1.6 h on walking per day.

In the correlation analysis, we look at the relationship between *self-assessed PA*, and accelerometer *daily activities* characteristics (see Table 4). There is moderate (significant) correlation between AC stepping and SA reported exercise ($r = 0.30$, $p = 0.03$), and SA reported exercise duration ($r = 0.31$, $p = 0.02$). Also, moderate (significant) correlation was found between AC sitting (i.e., avoiding sitting; recall the normalization direction) and SA exercise duration ($r = 0.33$, $p = 0.01$); AC walking and SA exercise duration ($r = 0.34$, $p = 0.01$). AC walking also significantly correlates with SA strenuous exercise program ($r = 0.30$, $p = 0.03$). No significant correlation was found between sitting time measured by accelerometer (AC Sitting) and self-reported sitting and lying time (SA Sitting and lying). As shown in Table 4, self-reported physical activity (SA physical activity) does not significantly correlate with accelerometer daily life activities.

However, there is moderate (significant) correlation between common AC and SA scores ($r = 31$, $p = 0.029$) (see Table 5). Moreover, a moderate (significant) correlation was found between our AC and ActivPAL's activity score ($r = 0.48$, $p = 0.000$). No correlation could be supported between SA and ActivPAL's activity score.

4.2. Prediction of accelerometer score using the SAT

As shown in Table 6, the recurrent neural network (RNN) and the convolutional neural network (CNN) lead to *good* predictors of PA as assessed by AC with $MAE = 3{,}8\%$ and $MAE = 5{,}09\%$, resp. using all normalized features and the uncut video sequences. This means all features and the leading and trailing frames before and after, resp., the exercise contribute to the overall accuracy.

4.3. Prediction of self-assessment score of physical activity using the SAT

As shown in Table 7, both models (RNN and CNN) are *moderate* predictors for self-reported PA—$MAE = 11.07$ and $MAE = 14.94\%$, resp. —using all normalized features and the cut video sequences. This means all features contribute to the overall accuracy, but not the leading and trailing frames before and after, resp., the exercise. In both models (Tables 6 and 7) RNN performed better than CNN.

5. Discussion

The present study was designed to explore the correlation between self-reported physical activity and daily life activities measured by an accelerometer; and how the SAT can be used to predict PA among older adults using the corresponding SA and AC scores. The correlation results indicate that there is a moderate ($r = 0.31$, see Table 5) significant ($p = 0.029$) correlation between self-reported physical activity and daily life activities measured by an accelerometer. It can thus be concluded that there is a difference between PA measured by self-assessment and using an accelerometer.

The machine learning results of this study indicate that SAT based on a functional balance test (stage 4 of 4SBT) can be used to predict PA as

Table 3
Descriptive statistics of daily life activities measured by ActivPAL.

Activity	Mean	Standard Deviation (SD)	Min-Max
Steps (per day)	7917	4168	2540–18162
Sit to Stand transitions (per day)	44.00	17.00	2–106
Sitting (in hours per day)	8.70	2.43	1–19
Standing (in hours per day)	4.00	1.46	0–12
Walking (in hours per day)	1.61	0.75	0–4.39
Biking (in hours per day)	0.07	0.20	0–1.49

Table 5
Correlation coefficients of aggregated overall SA, AC, and normalized ActivPAL activity scores.

r (p-value)	Accelerometer Score (AC)	Self-Assessment Score (SA)
Self-Assessment Score (SA)	**0.31 (*0.029*)**	
ActivPAL's Activity Score (PAL)	**0.48 (*0.000*)**	0.06 (0.68)

Table 4
Correlation coefficients of SA, and AC daily life activities.

Activity/*r (p-value)*	AC Stepping	AC SitToStand	AC Sitting	AC Standing	AC Walking	AC Biking
SA Sitting and lying	0.09 (*0.53*)	0.02 (*0.85*)	0.09 (*0.53*)	0.14 (*0.32*)	0.08 (*0.56*)	0.26 (*0.06*)
SA Physical activity	0.20 (*0.46*)	0.10 (*0.46*)	0.03 (*0.82*)	−0.02 (*0.87*)	0.17 (*0.22*)	0.13 (*0.33*)
SA Exercise	**0.30 (*0.03*)**	0.08 (*0.56*)	0.24 (*0.09*)	0.24 (*0.09*)	0.22 (*0.11*)	0.27 (*0.05*)
SA Often exercising	0.06 (*0.63*)	−0.05 (*0.68*)	0.31 (*0.02*)	0.10 (*0.45*)	0.08 (*0.56*)	0.08 (*0.56*)
SA Exercise duration	**0.31 (*0.02*)**	0.10 (*0.46*)	**0.33 (*0.01*)**	0.26 (*0.06*)	**0.34 (*0.01*)**	−0.03 (*0.80*)
SA Strenuous exercise program	0.29 (*0.04*)	0.04 (*0.75*)	0.22 (*0.11*)	0.14 (*0.32*)	**0.30 (*0.03*)**	0.05 (*0.72*)

Table 6
MAE in predicting AC score of daily life PA using SAT balance data.

	Recurrent Neural Network (RNN)	Convolutional Neural Network (CNN)
MAE in %	3,89%	5,09%
data transformation/ features	*uncut, normalized, all features*	*uncut, normalized, all features*

Table 7
MAE in predicting SA score of PA using SAT balance data.

	Recurrent Neural Network (RNN)	Convolutional Neural Network (CNN)
MAE in %	11,07%	14,94%
data transformation/ features	*cut, normalized, all features*	*cut, normalized, all features*

assessed with an accelerometer with high accuracy ($MAE = 3,89\%$, see Table 6). Predicting self-reported PA using the same SAT balance data provides only moderate accuracy ($MAE = 11.07\%$, see Table 7). This is no contradiction to having a moderate (significant) correlation between accelerometer score AC and self-reported activity score SA ($r = 0.31$).

Based on the deep learning results, it is likely that some association (perhaps a non-linear relationship) exists between balance and daily life activities. Previous studies have observed association between daily life PA in older adults and balance, gait, and that persistent physical activity may improve the balance in general.

Our own accelerometer score AC aggregating on ActivePAL raw data is somehow competing with ActivPAL's score aggregating the same data but differently. We do not suggest that our AC is superior to ActivPAL's score. There are, however, several reasons in favour of an overall accelerometer activity score and replacing/adding to the activity score provided by ActivPAL: (a) We wanted to make sure to use all relevant data collected from accelerometer devices (walking, stepping, etc.). (b) We could not repeat, hence, not rely on the way of calculating and representing the physical activity score of the ActivPAL software based on the raw accelerometer data.[‡] (c) We only found a moderate (but significant) correlation ($r = 0.48$) between our activity score and ActivPAL's score. In the future, we could repeat our study in a new sample using ActivPAL's activity score to see if we will get similar results. (d) ActivPAL's score has a very low (insignificant) correlation with the self-reported score (SA) in comparison to our activity score AC (see Table 5). Note that we avoided the use of existing METs cut points for healthy adults and did not introduce ours since, in this study, we did not have a representative sample size for defining general METs/accelerometer cut points for older adults.

The knowledge from this study shows that SAT may be supportive for clinicians to retrieve accurate information about a person's PA behaviuor based on the single movement (stage four of the 4SBT) scanned and analyzed by SAT. One selected movement performance (such as a balance test) would reduce both the time (it required only 10 s) and the effort (no need to wear accelerometer and other sensors for several days/weeks) to measure patients' PA. It would therefore be easier for the clinicians to apply on routinely basis. On the other hand, for the older adults that have large balance difficulties (cannot stand at all on one leg), this test is not applicable. For a more fine-grained PA assessment of these persons, we probably need to add other movements performances (e.g., stages 1–3 of 4SBT) in order to make our approach a more applicable tool. SAT can facilitate clinicians to both assess balance [27], and physical activity in general. The latter means a new possibility to measure PA in clinic that is more reliable than self-assessment. At the same time, it opens up for up-scale studies to gain valuable knowledge about i. e. PA patterns over time, cut-off points among older adults and fall risk assessment.

There are advantages of deep learning models and SAT for predicting PA: (a) its faster as it is based on only 10 s of the stage 4 of a 4SBT test instead of wearing an accelerometer for 7 days; (b) with the same SAT data (stage 4 of a 4SBT) other functional tests (FT) outcomes and even expert's movement quality assessments can be predicted [27]. Hence, the overall effort over several relevant factors for the well-being of older adults (PA, FT, movement quality) get reduced.

However, the disadvantage of this approach is that a deep learning model based on SAT cannot be interpreted by humans, cf. the machine learning black box problem [60]. While we observe the predictive power of the model predicting AC/SA outcomes based on SAT data with low/moderate errors, it is hard to transform this back to human knowledge.

In order to assess the validity of SAT-measured PA among older adults, we discuss the possible threats to validity [61,62] of this study and results in the following subsection.

5.1. Validity threats

Construct validity. In this study, we do neither rely on expert assessments of PA nor on METs/accelerometer cut-points. The participants were consulted to behave as usual in their everyday life during the collection of data with the accelerometer. However, there could be a risk that some participants were more active during this study then in general.

Conclusion validity. Objective and reliable methods were used to measure daily PA in older adults. A possible threat could be that no correlation analysis was not conducted between predictors (SAT data of 4SBT movements) and outcomes (SA and AC scores). Thus, the statistical significance of the relationship between predictors (SAT data of balance movements) and the outcome (AC, SA scores) is unknown. Also, the black box deep learning model does not provide details and explanations of the relationship between predictors and outcome variables. In addition, the small and maybe not representative sample of the population is a threat to validity.

Internal validity. Standard algorithms/techniques were used to analyze the data. For deep-learning cross-validation approach was used, and splitting the dataset on training and testing. However, in data-preprocessing the original dataset was artificially increased with well-known data augmentation technique, because of the otherwise too small sample size for deep leaning technique. A second possible threat is the use of uncut sequences (frames that do not belong to the fourth stage of 4SBT) to predict AC and SA outcomes. Maybe these leading and trailing frames were significant for the prediction, not the balance test. A third threat is that researchers were involved in the supervision of the data collection process (SAT data was collected in community care settings under instructions and technical supervision), which might have impacted on the collected data.

External validity. This study was conducted with randomly selected 54 participants (age 65+) who have multiple health conditions and diagnosis (such as diabetes, cardiomyopathy, stroke, asthma), and with different levels of PA reported in self-assessed questionnaire (see Table 2). However, the selection was driven by the availability of the subjects; we did not put effort in finding a representative sample of the (65+) population. Thus, we cannot generalize the study results to that population. The SAT approach was not clinically tested before and after physical therapy/or rehabilitation conditions and we cannot claim predictive power in these conditions.

Finally, we can neither claim any predictive power of SAT at home nor when based on other everyday movements of older adults on the daily life PA.

Dependability. Mathematically well-defined standard techniques and algorithms were used to analyze the data implemented in libraries with a wide user base, which makes the dependable. Our own implementations by-and-large used the libraries and were carefully tested. Cross-validation creates different random splits of training and testing data, which may cause small deviations in the results with other random splits. However, we are confident that repeating this study under the same conditions will lead to the same correlation and prediction accuracy results.

6. Conclusion

There is a need to develop a simple and reliable method to complement/replace self-assessment methods, and objectively measure physical activities among older adults. Therefore, the first objective of this study was to explore the association between self-assessment methods and accelerometer. This study has shown that self-assessment data have a moderate (significant) correlation with accelerometer data. The second objective was to facilitate the assessment of daily life physical activity among older adults using SAT of standardized functional (balance) tests. The obtained results indicate that one functional balance test measured with SAT can be used to predict PA outcomes measured with accelerometer devices. However, the obtained prediction results indicate that there is some non-linear association between functional balance tests recorded and assessed by SAT and physical activity as assessed by an accelerometer. The SAT can predict PA outcomes better than SA outcomes within the same population. Further research should be undertaken to investigate (a) the use of SAT in everyday movement/person transfers such as getting up from a chair or a bed to measure the daily life PA in older adults; (b) the use of SAT to predict PA among older adults with various functional abilities; and (c) how SAT can be developed using 2D information, such as mobile phone recordings, to measure PA. Positive result in (a) would allow for installed cameras—both in the homes of the older adults or at the clinic—and continuous assessments. Improvements in (b) are needed to generalize and effectively utilize the suggested approach in larger, more diverse populations. Regarding (c), there is ongoing and promising development to replace the 3D Kinect camera with a mobile phone camera, which would enable measurements of physical activity more conveniently, independent of technical experts and practically useable, e.g., in the home settings. This may furthermore facilitate and add value to routinely measurement of functional ability and fall risk among older adults, e.g. in primary care, which has been asked for [8].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to all the participants in the study, to the Linnaeus University Center for Data Intensive Sciences and Application and its High-Performance Computing Center for the financial support and the provision of the computing infrastructure, and to AIMO for developing a research version of their SAT software used in this study.

Abbreviations

4SBT	4-Stage Balance Test
AC	Accelerometer-based assessments
CNN	Convolution Neural Network
DTW	Dynamic Time Warping
EE	Energy Expenditure
FT	Functional Test
ICC	Intra-class Correlation Coefficient
METs	Metabolic Equivalent Tasks
MAE	Mean Absolute Error
RNN	Recurrent Neural Network
PA	Physical Activity
SA	Self-Reported Assessments
SAT	Skeleton Avatar Technology
SB	Sedentary Behavior

References

[1] Ejlersen H, Andersen ZJ, von Euler-Chelpin MC, Johansen PP, Schnohr P, Prescott E. Prognostic impact of physical activity prior to myocardial infarction: case fatality and subsequent risk of heart failure and death. Eur. J. Prev. Cardiol. 2017;24(10):1112–9.
[2] Talarska D, Tobis S, Kotkowiak M, Strugała M, Stanisławska J, Wieczorowska-Tobis K. Determinants of quality of life and the need for support for the elderly with good physical and mental functioning. Med. Sci. Monit. Int. Med. J. Exp. Clin. Res. 2018;24:1604.
[3] Aunger JA, Doody P, Greig CA. Interventions targeting sedentary behavior in non-working older adults: a systematic review. Maturitas 2018;116:89–99.
[4] Rejeski WJ, et al. The MAT-sf: identifying risk for major mobility disability. J. Gerontol. Ser. A Biomed. Sci. Med. Sci. 2015;70(5):641–6.
[5] Stierlin AS, et al. A systematic review of determinants of sedentary behaviour in youth: a DEDIPAC-study. Int J Behav Nutr Phys Activ 2015;12(1):1–19.
[6] Lara J, et al. A proposed panel of biomarkers of healthy ageing. BMC Med 2015;13 (1):1–8.
[7] Lee I-M, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012;380 (9838):219–29.
[8] Meekes WMA, Leemrijse CJ, Korevaar JC, Henquet J, Nieuwenhuis M, van de Goor LAM. Implementation and evaluation of a fall risk screening strategy among frail older adults for the primary care setting: a study protocol. Clin Interv Aging 2020;15:1625.
[9] Ryan DJ, Wullems JA, Stebbings GK, Morse CI, Stewart CE, Onambele-Pearson GL. Reliability and validity of the international physical activity questionnaire compared to calibrated accelerometer cut-off points in the quantification of sedentary behaviour and physical activity in older adults. PloS One 2018;13(4): e0195712.
[10] Jørstad-Stein EC, et al. Suitability of physical activity questionnaires for older adults in fall-prevention trials: a systematic review. J Aging Phys Activ 2005;13(4): 461–81.
[11] Harris TJ, Owen CG, Victor CR, Adams R, Ekelund ULF, Cook DG. A comparison of questionnaire, accelerometer, and pedometer: measures in older people. Med Sci Sports Exerc 2009;41(7):1392–402.
[12] Chan CS, Slaughter SE, Jones CA, Wagg AS. Measuring activity performance of continuing care residents using the ActivPAL: an exploratory study. J Frailty Aging 2016;5(3):158–61.
[13] Washburn RA, Jette AM, Janney CA. Using age-neutral physical activity questionnaires in research with the elderly. J Aging Health 1990;2(3):341–56.
[14] Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. Res Q Exerc Sport 2000;71:1–14. sup.2.
[15] Rikli RE. Reliability, validity, and methodological issues in assessing physical activity in older adults. Res Q Exerc Sport 2000;71:89–96. sup.2.
[16] Lee I-M, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. Br J Sports Med 2014;48(3): 197–201.
[17] Heesch KC, Hill RL, Aguilar-Farias N, Van Uffelen JGZ, Pavey T. Validity of objective methods for measuring sedentary behaviour in older adults: a systematic review. Int J Behav Nutr Phys Activ 2018;15(1):1–17.
[18] Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. J Aging Phys Activ 2009;17(1):17–30.
[19] Migueles JH, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports Med 2017;47(9):1821–45.
[20] Almanza E, Jerrett M, Dunton G, Seto E, Pentz MA. A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. Health Place 2012;18(1):46–54.
[21] Kerr J, et al. The relationship between outdoor activity and health in older adults using GPS. Int J Environ Res Publ Health 2012;9(12):4615–25.
[22] Ejupi A, Brodie M, Gschwind YJ, Lord SR, Zagler WL, Delbaere K. Kinect-based five-times-sit-to-stand test for clinical and in-home assessment of fall risk in older people. Gerontology 2016;62(1):118–24.
[23] Clark RA, et al. Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control. Gait Posture 2015; 42(2):210–3.
[24] Dressler D, Liapota P, Löwe W. "Towards an automated assessment of musculoskeletal insufficiencies," in *Intelligent Decision Technologies 2019*. Springer; 2020. p. 251–61.
[25] Dressler D, Liapota P, Löwe W. Data-driven human movement assessment. In: Intelligent decision Technologies 2019. Springer; 2019. p. 317–27.

[26] Hagelbäck J, Liapota P, Lincke A, Löwe W. The performance OF some machine learning approaches IN human movement assessment. Proc Int Conf e-Health 2019: 35–42. 2019.
[27] Backåberg S, et al. Evaluation of the skeleton avatar technique for assessment of mobility and balance among older adults. Front Comput Sci 2020;2:1–13.
[28] Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Publ Health Rep 1985;100(2):126.
[29] Allison MJ, Keller C, Hutchinson PL. Selection of an instrument to measure the physical activity of elderly people in rural areas. Rehabil Nurs 1998;23(6):309–14.
[30] Hansen BH, Kolle E, Dyrstad SM, Holme I, Anderssen SA. Accelerometer-determined physical activity in adults and older people. Med Sci Sports Exerc 2012; 44(2):266–72.
[31] Garatachea N, Luque GT, Gallego JG. Physical activity and energy expenditure measurements using accelerometers in older adults. Nutr Hosp 2010;25(2):224–30.
[32] Denkinger MD, et al. Accelerometer-based physical activity in a large observational cohort-study protocol and design of the activity and function of the elderly in Ulm (ActiFE Ulm) study. BMC Geriatr 2010;10(1):1–14.
[33] Sumukadas D, Laidlaw S, Witham MD. Using the RT3 accelerometer to measure everyday activity in functionally impaired older people. Aging Clin Exp Res 2008; 20(1):15–8.
[34] Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults: a systematic review and data analysis. Eur. Rev. Aging Phys. Act. 2014;11(1): 35–49.
[35] Colbert LH, Matthews CE, Havighurst TC, Kim K, Schoeller DA. Comparative validity of physical activity measures in older adults. Med Sci Sports Exerc 2011;43 (5):867.
[36] Matthews CE, Ainsworth BE, Thompson RW, Bassett Jr DR. Sources of variance in daily physical activity levels as measured by an accelerometer. Med Sci Sports Exerc 2002;34(8):1376–81.
[37] Phillips LJ, Petroski GF, Markis NE. A comparison of accelerometer accuracy in older adults. Res Gerontol Nurs 2015;8(5):213–9.
[38] Aguilar-Farias N, Peeters G, Brychta RJ, Chen KY, Brown WJ. Comparing ActiGraph equations for estimating energy expenditure in older adults. J Sports Sci 2019;37(2):188–95.
[39] Hall KS, Howe CA, Rana SR, Martin CL, Morey MC. METs and accelerometry of walking in older adults: standard versus measured energy cost. Med Sci Sports Exerc 2013;45(3):574.
[40] Klenk J, et al. Concurrent validity of activPAL and activPAL3 accelerometers in older adults. J Aging Phys Activ 2016;24(3):444–50.
[41] Wullems JA, Verschueren SMP, Degens H, Morse CI, Onambélé GL. Performance of thigh-mounted triaxial accelerometer algorithms in objective quantification of sedentary behaviour and physical activity in older adults. PloS One 2017;12(11): e0188215.
[42] Van Dyck D, et al. Environmental and psychosocial correlates of accelerometer-assessed and self-reported physical activity in Belgian adults. Int J Behav Med 2011;18(3):235–45.
[43] Sabia S, et al. Association between questionnaire-and accelerometer-assessed physical activity: the role of sociodemographic factors. Am J Epidemiol 2014;179 (6):781–90.
[44] Dyrstad SM, Hansen BH, Holme IM, Anderssen SA. Comparison of self-reported versus accelerometer-measured physical activity. Med Sci Sports Exerc 2014;46(1): 99–106.
[45] Kochersberger G, McConnell E, Kuchibhatla MN, Pieper C. The reliability, validity, and stability of a measure of physical activity in the elderly. Arch Phys Med Rehabil 1996;77(8):793–5.
[46] Dawe RJ, et al. Association between quantitative gait and balance measures and total daily physical activity in community-dwelling older adults. J Gerontol Ser A 2018;73(5):636–42.
[47] Galperin I, et al. "Associations between daily-living physical activity and laboratory-based assessments of motor severity in patients with falls and Parkinson's disease. Park Relat Disord 2019;62:85–90.
[48] Pau M, Leban B, Collu G, Migliaccio GM. Effect of light and vigorous physical activity on balance and gait of older adults. Arch Gerontol Geriatr 2014;59(3): 568–73.
[49] McMullan II, McDonough SM, Tully MA, Cupples M, Casson K, Bunting BP. The association between balance and free-living physical activity in an older community-dwelling adult population: a systematic review and meta-analysis. BMC Publ Health 2018;18(1):1–21.
[50] Barnett A, van den Hoek D, Barnett D, Cerin E. Measuring moderate-intensity walking in older adults using the ActiGraph accelerometer. BMC Geriatr 2016;16 (1):1–9.
[51] Hooker SP, et al. Validation of the actical activity monitor in middle-aged and older adults. J Phys Activ Health 2011;8(3):372–81.
[52] Aguilar-Far\'\ias N, Brown WJ, Peeters GG. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. J Sci Med Sport 2014;17(3):293–9.
[53] Matlab. Version 9.8.0 (R2020a). Natick, Massachusetts: The MathWorks Inc.; 2020.
[54] Hart TL, McClain JJ, Tudor-Locke C. Controlled and free-living evaluation of objective measures of sedentary and active behaviors. J Phys Activ Health 2011;8 (6):848–57.
[55] Florez-Pregonero A, Meckes N, Buman M, Ainsworth BE. Wearable monitors criterion validity for energy expenditure in sedentary and light activities. J Sport Heal Sci 2016;6(1):103–10.
[56] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019;6(1):1–48.
[57] Hagelbäck J, Liapota P, Lincke A, Löwe W. Variants of dynamic time warping and their performance in human movement assessment. In: Proceedings on the international Conference on artificial intelligence (ICAI). Las Vegas: July, CSREA Press; 2019. p. 9–15.
[58] Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning, vol. 1. MIT press Cambridge; 2016. 2.
[59] Abadi M, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems. *arXiv Prepr. arXiv1603.04467*; 2016.
[60] Papernot N, McDaniel P, Goodfellow I, Jha S, Celik ZB, Swami A. Practical black-box attacks against machine learning. In *Proceedings of the 2017 ACM on Asia conference on computer and communications security*. 2017. p. 506–19.
[61] Feldt R, Magazinius A. Validity threats in empirical software engineering research-an initial survey. Seke 2010:374–9.
[62] Mustafa N, Labiche Y, Towey D. "Mitigating threats to validity in empirical software engineering: a traceability case study," in *2019*. IEEE 43rd Ann Comput Softw Appl Conf (COMPSAC) 2019;2:324–9.