Frailty status can be accurately assessed using inertial sensors and the TUG test

BARRY R. GREENE1, EMER P. DOHENY1, AISLING O’HALLORAN1, ROSE ANNE KENNY2,3

1TRIL Centre, University College Dublin, Dublin, Ireland
2Dept Medical Gerontology, Trinity College Dublin, Ireland
3Falls and Blackout Unit at St James’s Hospital, Dublin, Ireland

Address correspondence to: Barry R. Greene. Tel: +353-1-716-5402. Email: barry.greene@ucd.ie

Abstract

Background: frailty is an important geriatric syndrome linked to increased mortality, morbidity and falls risk.
Methods: a total of 399 community-dwelling older adults were assessed using Fried’s frailty phenotype and the timed up and go (TUG) test. Tests were quantified using shank-mounted inertial sensors. We report a regression-based method for assessment of frailty using inertial sensor data obtained during TUG. For comparison, frailty was also assessed using the same method based on grip strength and manual TUG time.
Results: using inertial sensor data, participants were classified as frail or non-frail with mean accuracy of 75.20% (stratified by gender). Using TUG time alone, frailty status was classified correctly with mean classification accuracy of 71.82%. Similarly, using grip strength alone, the frailty status was classified correctly with mean classification accuracy of 77.65%. Stratifying sensor data by gender yielded significantly (p<0.05) increased accuracy in classifying frailty when compared with equivalent manual TUG time-based models.
Conclusion: results suggest that a simple protocol involving assessment using a well-known mobility test (Timed Up and Go (TUG)) and inertial sensors can be a fast and effective means of automatic, non-expert assessment of frailty.

Keywords: community dwelling older adults, frailty, inertial sensor, mobility, older people, TUG

Introduction

In recent years, the concept of frailty in older adults has emerged as an important geriatric syndrome [1–3]. Although frailty is a recognisable and common phenomenon in ageing, it is also a rather nebulous concept, making it difficult to accurately define and diagnose. It is a multi-factorial condition, influenced by the combination of a person’s physical, psychological and social health. Fried et al. [3] showed that the ‘frailty phenotype’ has statistically significant predictive association with five important health outcomes which are: death, first hospitalisation, first fall, worsening activities of daily living disability and worsening mobility. These findings have been confirmed in other populations [1, 4]. However, manual application of Frailty scales can be time consuming and may require clinical expertise for interpretation.

Prompt and accurate identification of a person’s frailty state could allow effective multi-factorial intervention which has been shown to improve health outcomes [5].

The Timed Up and Go (TUG) test [6, 7] is a standard mobility assessment, the time taken to complete the test has been shown to be a strong predictor of frailty [8] and is commonly used for assessing risk of falls in older adults. Recent research has investigated using inertial sensors for quantitative evaluation of movement. A number of recent studies have employed inertial sensors for quantitative gait analysis [9, 10] and to quantify movement in the TUG test [11, 12]. Higashi et al. [13] employed body-worn gyroscopes to evaluate movement in hemiplegic patients with pathological gait while performing the TUG test. Salarian et al. [12] showed that an instrumented TUG was both a reliable and sensitive method for quantifying gait and mobility in Parkinson’s disease patients. Similarly, Weiss et al. [11] used body-worn accelerometers to quantify the gait of Parkinson’s disease patients during the TUG test. Martinez-Ramirez [14] used accelerometers during a standing balance tasks to examine the utility of parameters derived using a wavelet based algorithm to discriminate between frail, pre-frail and robust participants. To date, automated classification of frailty using inertial sensors has not been investigated.

The present study investigates a fast method for automatic, quantitative assessment of the frailty state of a patient based on a simple protocol employing body-worn inertial
sensors. The value of frailty assessment is in early intervention (based on targeted risk factors) and prevention of decline. Methodologies which can be applied at population level rather than requiring highly skilled application would be of particular utility. A system that would allow non-expert, objective assessment of frailty, coupled with an assessment of mobility, could be of significant clinical benefit.

Data set

Participants

A convenience sample of 479 (138 M, 341 F) community-dwelling older adults were recruited for a wider study on ageing, conducted in the TRIL clinic, St James hospital, Dublin, Ireland. The inclusion criteria were persons aged 60 and over, who were able to walk independently with or without walking aid, cognitively intact (Mini-Mental State Exam >18 [15, 16]) and able to provide informed consent. Ethical approval was received from the local ethics committee.

Clinical assessment

Each participant received a comprehensive geriatric assessment [17], which included tests of visual acuity (Binocular LogMar) and visual contrast sensitivity (Pelli-Robson test), maximum grip strength and blood pressure, to check for orthostatic hypotension (defined as orthostatic SBP drop >20 mmHg). Each participants balance and mobility was evaluated using the TUG test and also the Berg balance scale (BBS) [18]. The results of these assessments along with demographic characteristics are tabulated in Table 1.

Maximum grip strength (lbs) (taken as the maximum of the left and right hand grip strength) was used as a reference measure of frailty [19]. Values were measured using a handheld dynamometer (Baseline® Hydraulic Hand Dynamometers, NexGen Ergonomics Inc., Quebec, Canada).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Robust</th>
<th>Pre-frail</th>
<th>Frail</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (M/F)</td>
<td>61/123</td>
<td>47/138</td>
<td>7/23</td>
</tr>
<tr>
<td>Age (years)</td>
<td>71.38 ± 6.69</td>
<td>74.95 ± 7.41</td>
<td>78.17 ± 6.18</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>166.77 ± 8.76</td>
<td>162.41 ± 9.20</td>
<td>160.00 ± 6.86</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>74.04 ± 13.40</td>
<td>72.43 ± 14.82</td>
<td>68.79 ± 13.18</td>
</tr>
<tr>
<td>Binocular LogMar</td>
<td>0.11 ± 0.15</td>
<td>0.16 ± 0.17</td>
<td>0.23 ± 0.23</td>
</tr>
<tr>
<td>Contrast Sens</td>
<td>1.63 ± 0.19</td>
<td>1.56 ± 0.21</td>
<td>1.55 ± 0.16</td>
</tr>
<tr>
<td>Max grip (lbs)</td>
<td>64.81 ± 20.58</td>
<td>45.98 ± 17.29</td>
<td>36.91 ± 10.55</td>
</tr>
<tr>
<td>BBS</td>
<td>54.41 ± 2.63</td>
<td>51.34 ± 5.14</td>
<td>43.29 ± 5.25</td>
</tr>
</tbody>
</table>

Bonferroni post hoc analysis is also reported. Data are reported as mean ± SD. BBS refers to Berg balance scale [18], Contrast Sen. refers to the Pelli-Robson contrast sensitivity test.

Frailty assessed using inertial sensors and TUG

Method

Frailty model

The Fried frailty formulation defined frailty as a syndrome in which three or more of the following criteria are present:

- unintentional weight loss
- self-reported exhaustion
- weakness (as measured by grip strength)
- slow walking speed
- low physical activity

Participants were then categorised into three classes: robust, pre-frail and frail based on the five frailty criteria outlined above. A participant was classed as frail if they met three or more of the criteria listed above. Participants meeting one or two criteria were classed as pre-frail according to the modified Fried criteria used by NiMhaoilain et al. [20].

Sensor data acquisition

The participant was asked to perform the TUG test [6], as fast as was safely possible, by getting up from a standard chair (46 cm high seat, 65 cm arm rests), walking three metres, turning at a designated spot, returning to the seat and sitting down. The time taken to complete the task was also recorded by the clinician using a stopwatch. The time was measured from the moment the clinician says ‘go’ to the moment the participant sits back on the chair (referred to hereafter as the manual TUG time). The task was demonstrated to each participant and participants were given time to familiarise themselves with the test. Participants completed the TUG once but were allowed to repeat the test if they did not complete the first one correctly. In order to quantify movement, kinematic data for each participant were acquired using two wireless body-worn inertial sensors (SHIMMER, Dublin, Ireland), which were attached by a research nurse, using elasticised bandages, to the mid-point of each anterior shank (shin) [21]. Each sensor contained a tri-axial accelerometer and a tri-axial gyroscope and sampled at 102.4 Hz. Sensors were calibrated using a standard method [22]. The raw gyroscope signal was low pass filtered with zero-phase 2nd order Butterworth filter with a 20 Hz corner frequency. All analysis was performed offline in Matlab version 7.11 (Natick, VA, USA).

Sensor data analysis

The mobility of each participant performing the TUG test was evaluated using a previously reported method for quantitative assessment of movement during the TUG test [21, 23]. Features were calculated from the angular velocity signals obtained from the tri-axial gyroscope sensors mounted on each shank. The 44 sensor-derived features can be grouped into four categories: temporal gait parameters, spatial gait parameters, tri-axial angular velocity parameters and turn parameters. Clinical parameters were included in analysis (see Table 1). Coefficient of variation features were transformed using a log-transform to ensure a more normal distribution.
All features were then normalised to have zero mean and unity standard deviation.

**Statistical analysis**

A one-way analysis of variance (ANOVA) was used to test for differences in each clinical parameter between robust, pre-frail and frail classes (Table 1). Bonferroni post hoc analysis was performed to examine differences between each of the classes. A logistic regression model with interaction terms included was used to evaluate frailty, using a data set containing the sensor-derived features detailed above, combined with gender, age, height and weight. The features included in each model were selected using a cross-validated sequential forward feature selection procedure. Models were then evaluated using a separate repeated cross-fold validation, with 10-fold and 10 repetitions. Frailty was considered as a binary classification problem; grouping participants listed as frail and pre-frail together into one frail class and comparing this to a non-frail (robust) class. The output of this model was an estimate of the frailty category (frail/non-frail). This estimated frailty category was then compared with the true frailty category (as defined using modified Fried criteria) to yield an estimate of the accuracy in classifying each participant according to frailty category. A number of different classifier configurations and feature sets were investigated. Data were stratified by gender as well as considered in a single model. Regression models based only on the maximum grip strength or the manual TUG time for each participant were also generated to provide a reference method for quantitative assessment of frailty. As with the inertial sensor method, maximum grip strength and TUG time data were stratified by gender as well as considered in a single model.

The classification accuracy (Acc) is defined as the percentage of participants correctly identified by the system as frail or non-frail. The sensitivity (Sens) is defined as the percentage of each class (frail or non-frail) correctly classified as such by the algorithm. The predictivity value (Pred) is defined as the proportion of participants, classified as a given class by the algorithm, who were correctly classified. The classifier performance measures were then taken as the mean of each measure across all folds and repetitions, providing an unbiased, low variance estimate of classification performance. 95% confidence intervals (95% CI) were also calculated for each classifier performance metric. A two-sided t-test was used to determine if the classification results obtained were significantly (P < 0.05) more accurate than those obtained using only the TUG time or maximum grip strength.

**Results**

Problems with data acquisition during the initial assessment, including sensor/software failure and human error led to sensor data for a number of participants being excluded. It should be noted that the observed errors were in no way related to the outcome of interest and arose due to operational issues at the start of the study. Sensor data for 399 of 479 (83.33%) participants were then included in the final analysis (115 males, 284 females see Tables 2 and 3).

Thirty participants were categorised as frail, 185 as pre-frail and 184 as robust using the Fried frailty criteria. Grouping the participants as frail or non-frail gave 184 non-frail and 215 participants categorised as frail. The mean age of the cohort at the time of initial evaluation was 73.6 ± 7.3 years, while the mean height and weight were 164.2 ± 9.2 cm and 72.9 ± 14.1 kg, respectively. Clinical

<table>
<thead>
<tr>
<th>Inertial sensor</th>
<th>Max grip</th>
<th>TUG time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acc (%)</strong></td>
<td><strong>Sens (%)</strong></td>
<td><strong>Pred (%)</strong></td>
</tr>
<tr>
<td>Non-frail Frail</td>
<td>Non-frail Frail</td>
<td>Non-frail Frail</td>
</tr>
<tr>
<td>All</td>
<td>72.88</td>
<td>74.33</td>
</tr>
<tr>
<td>Male</td>
<td>78.09</td>
<td>83.00</td>
</tr>
<tr>
<td>Female</td>
<td>72.30</td>
<td>79.07</td>
</tr>
<tr>
<td>Mean</td>
<td>75.20</td>
<td>77.05</td>
</tr>
</tbody>
</table>

Participants were classified using a binary definition of frailty. Classification accuracy (Acc) is defined as the percentage of participants correctly identified by the system as frail or non-frail. Sensitivity (Sens) is defined as the percentage of each class (frail or non-frail) correctly classified as such by the algorithm. The predictivity value (Pred) is defined as the proportion of participants, classified as a given class by the algorithm, who were correctly classified.

**Table 2. Cross-validated results for logistic regression models developed using: inertial sensor parameters obtained during the TUG test (left panel); maximum grip strength (centre panel); TUG time**

<table>
<thead>
<tr>
<th>Inertial sensor (%)</th>
<th>TUG time (%)</th>
<th>Max grip (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (M/F)</td>
<td>All (M/F)</td>
<td>All (M/F)</td>
</tr>
<tr>
<td>72.88 (71.57–74.18)</td>
<td>72.09 (70.64–73.55)</td>
<td>66.93 (65.60–68.26)</td>
</tr>
<tr>
<td>72.09 (70.54–74.07)</td>
<td>73.97 (71.65–76.29)</td>
<td>76.83 (74.52–79.15)</td>
</tr>
<tr>
<td>72.30 (70.54–74.07)</td>
<td>69.76 (68.12–71.41)</td>
<td>78.47 (77.03–79.91)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inertial sensor (%)</th>
<th>TUG time (%)</th>
<th>Max grip (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (M/F)</td>
<td>All (M/F)</td>
<td>All (M/F)</td>
</tr>
<tr>
<td>72.88 (71.57–74.18)</td>
<td>72.09 (70.64–73.55)</td>
<td>66.93 (65.60–68.26)</td>
</tr>
<tr>
<td>72.09 (70.54–74.07)</td>
<td>73.97 (71.65–76.29)</td>
<td>76.83 (74.52–79.15)</td>
</tr>
<tr>
<td>72.30 (70.54–74.07)</td>
<td>69.76 (68.12–71.41)</td>
<td>78.47 (77.03–79.91)</td>
</tr>
</tbody>
</table>

**Table 3. Mean classification accuracies (95% confidence intervals) for the inertial sensor, TUG time and max grip frailty classifier models**
information for the cohort from the clinical assessment is detailed in Table 1. Statistically significant differences across the three frailty categories were observed in a number of the clinical parameters (age, height, binocular logMAR, contrast sensitivity, max grip, TUG and BBS).

Combining the frail and pre-frail classes into a single Frail class reduces the problem to a binary classification problem. Using inertial sensor data applied to a single regression model the mean cross-validated classification accuracy was 72.88% (95% CI: 71.57–74.18%) while models stratified by gender (male: 78.09% (95% CI: 75.67–80.52%), female: 72.30%, (95% CI: 70.54–74.07%) produced a mean classification accuracy of 75.20%). Inclusion of eye gaze parameters into the inertial sensor models did not improve results significantly.

Models using only maximum grip strength or the TUG time produced classification accuracies of 66.93% (95% CI: 65.60–68.26%) and 72.09% (95% CI: 70.64–73.55%), respectively. When stratified by gender, maximum grip strength and manual TUG time produced mean classification accuracies of 77.65% (male: 76.83% (95% CI: 74.52–79.15%), female: 78.47% (95% CI: 77.03–79.91%)) and 71.82% (male: 73.97% (95% CI: 71.65–76.29%), female: 69.76% (95% CI: 68.12–71.41%)), respectively.

When stratified by gender, logistic regression models based on the sensor data were found to be significantly more accurate ($P < 0.05$) than the manual TUG time. Maximum grip strength was not found to be significantly more accurate than the sensor based model for females but was found to be significantly more accurate for males. Table 2 reports classification results for sensor data-based models, along with results for models obtained using only the maximum grip strength or the TUG time for each participant.

The combined male and female regression model contained max X-axis angular velocity (deg/s) and the TUG time (s). The male data model contained the maximum Y-axis angular velocity (deg/s) and the TUG time (s). The female data model contained single support (%), minimum Z-axis angular velocity (deg/s) and the TUG time. A definition of the included variables can be found elsewhere [21].

**Discussion**

We report a method for assessing the clinical phenotype of frailty using body-worn inertial sensors and the TUG test. The TUG is a simple, widely accepted test of mobility. The inertial sensors used are small, lightweight and can be easily applied by non-experts in a supervised clinical setting. Due to the necessity for correct orientation of the sensors on the shank and the need to ensure the sensors are charged between sessions, unsupervised use by the patient may not be feasible.

We found that a model classifying participants according to two frailty categories (frail and non-frail) yielded a mean cross-validated classification accuracy of 72.88%. Stratifying the inertial sensor data by gender to produce separate male and female regression models yielded improved results (75.20%).

Maximum grip strength and TUG time were found to be very strong independent predictors of frailty which is in broad agreement with the literature [8, 19, 24]. Grip strength is used by the Fried frailty formulation as a surrogate for ‘weakness’ so this result is not surprising. Similarly, slow walking speed (gait velocity) as used in the Fried definition of frailty can be related to TUG time which have both been linked with falls, frailty and cognitive decline [25–27]. Regression models with maximum grip strength and TUG time as the only predictors were less accurate in assessing frailty than the inertial sensor based method. However, when maximum grip strength data were stratified by gender (as per the Fried frailty formulation), the resulting regression models were more accurate in classifying frailty category. Stratifying sensor data by gender yielded significantly increased accuracy in classifying frailty when compared with manual TUG time-based models. Results suggest that inertial sensor methods may have utility in automatic quantitative assessment of frailty in a clinical environment.

The main advantage of the present method over that reported by Fried et al. [3] is that the present method is fast and automated. In the Fried et al. formulation, continuous variables (i.e. grip strength, walking speed and physical activity) need to be retrospectively dichotomised using a variety of stratifications. This requires considerable statistical expertise and also a reference sample, both of which are not always available in a primary care setting. A number of the subcomponents of the Fried scale are derived from longer questionnaires (e.g. Center for Epidemiologic Studies Depression Scale [28] and the Minnesota Leisure Time activity questionnaire [29]) which can be time consuming. Additionally, the quantitative mobility assessment provided by the present method may provide insight into specific mobility impairments that may be associated with frailty.

A limitation of the present study is the small number of participants ($N = 30$) categorised as ‘frail’. To increase the statistical power, the pre-frail and frail categories were combined to produce two classes: frail and non-frail. Given a larger cohort, it may be possible to create robust multi-class statistical models that can reliably classify participants into each of the three frailty classes. A large proportion of the present cohort were self-referred which could indicate differences when compared with cohorts of hospital in-patients or nursing home residents. Every effort was made to ensure the statistical models used in the present study were generalised across the study population, differences may exist when compared with the general population given the relatively robust nature of the study population.

Martinez-Ramirez et al. [14] found that measures derived from accelerometers during a standing balance task could discriminate between the three frailty classes as defined by Fried et al. [3] but did not provide a means to automatically classify participants according to frailty class. Authors [14] reported that inertial sensors methods may be able to distinguish between robust and frail, and robust and pre-frail subjects. However, they suggest that such methods did not seem to be useful for distinguishing between pre-frail and frail
A cohort of 399 community-dwelling older adults were assessed using Fried’s frailty phenotype and the TUG test, movement was quantified using shank-mounted inertial sensors. A regression-based method classified participants as frail or non-frail with mean accuracy of 75.20% (stratified by gender). Using TUG time alone, frailty status was classified correctly with mean classification accuracy of 71.82%.

A simple protocol using the TUG and inertial sensors can be a fast and effective means of automatic, non-expert assessment of frailty.

Key points

- Frailty is an important geriatric syndrome linked to increased mortality, morbidity and falls risk.
- A cohort of 399 community-dwelling older adults were assessed using Fried’s frailty phenotype and the TUG test.
- A regression-based method classified participants as frail or non-frail with mean accuracy of 75.20% (stratified by gender).
- Using TUG time alone, frailty status was classified correctly with mean classification accuracy of 71.82%.
- A simple protocol using the TUG and inertial sensors can be a fast and effective means of automatic, non-expert assessment of frailty.

Supplementary data

Supplementary data mentioned in the text are available to subscribers in Age and Ageing online.

Acknowledgements

This research was completed as part of a wider programme of research within the TRIL Centre (Technology Research for Independent Living). The TRIL Centre is a multi-disciplinary research centre, bringing together researchers from UCD, TCD, NUIG & Intel, funded by Intel, GE and IDA Ireland. http://www.trilcentre.org. We also acknowledge the help and support of the staff of the TRIL Clinic, St James, hospital, Dublin and the participants involved in this study.

Conflicts of interest

B.R.G. is a director of a company that has a license to develop and market the technology reported in this study.

References

23. Greene BR, Doheny EP, Walsh C, Cunningham C, Crosby I, Kenny RA. Evaluation of falls risk in community-dwelling...
Male sex and vascular risk factors affect cystatin C-derived renal function in older people without diabetes or overt vascular disease

KARIN BIRGITTA WERNER1,2, SÖLVE ELMSTÅHL1,2, ANDERS CHRISTENSSON3,4, MATS PIHLSGÅRD1

1Department of Health Sciences, Lund University, Geriatriksa kliniken Jan Waldenströmgata 35, Malmö 205 02, Sweden
2Department of Geriatrics, Skåne University Hospital, Jan Waldenströmsga 35, Malmö 205 02, Sweden
3Department of Clinical Sciences, Lund University, Malmö, Sweden
4Department of Nephrology and Transplantation, Skåne University Hospital, Malmö, Sweden

Address correspondence to: K. B. Werner. Email: karin.werner@med.lu.se

Abstract

Background/objectives: to explore the effect of ageing on renal function with cystatin C as the marker of glomerular filtration rate (GFR) in the general population without vascular disease or diabetes.

Design: a cross-sectional analysis of a healthy subset from the Good Aging in Skåne-cohort study representative of the Swedish general population.

Subjects: 1252 participants without vascular disease and diabetes (43.9% men) of whom 203 were over 80 years old were included from the original cohort of 2931.

Methods: plasma cystatin C and plasma creatinine were used as markers for GFR. Estimated GFR (eGFR) was calculated with three chronic kidney disease epidemiology collaboration (CKD-EPI) formulas involving cystatin C, creatinine or both.

Results: the median for plasma cystatin C was 0.93 mg/l (60–69 years old), 1.04 (70–79 years old) and 1.24 (80+ years old). The difference in mg/l between the 5th and 95th percentile was 0.46, 0.62 and 0.90 for these age groups. Male sex increased the age effect on plasma cystatin C levels with 0.004 mg/l/year (P = 0.03), adjusted for vascular risk factors. Smoking, lower HDL and higher diastolic blood pressure were associated with higher cystatin C levels. 54.7% (CKD-EPI creatinine) to 73.9% (CKD-EPI cystatin C) of the 80+ had an eGFR < 60 ml/min/1.73 m2.

Conclusion: non-diabetics without overt vascular disease exhibit an age related but heterogeneous decline in renal function. The ageing effect is more pronounced in men. At least half of healthy 80+ years old could be expected to have at least CKD Stage 3 with eGFR < 60 ml/min/1.73 m2.