Published in final edited form as:

Gerontology. 2018 ; 64(4): 389–400. doi:10.1159/000484241.

Toward Using Smart Watch to Monitor Frailty in Hospital Setting —Using a Single Wrist Wearable Sensor to Assess Frailty in Bedbound Inpatients

Hyoki Lee¹, Bellal Joseph², Ana Enriquez¹, and Bijan Najafi^{1,1*}

¹Interdisciplinary Consortium on Ambulatory Motion Performance (iCAMP), Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, Texas, United States

²Department of Surgery, Division of Trauma, Critical Care, Burns, and Acute Care Surgery, University of Arizona, Tucson, Arizona, United States

Abstract

Background—While various objective tools have been validated for assessing physical frailty in the geriatric population, these are often unsuitable for busy clinics and mobility-impaired patients. Recently we have developed a Frailty Meter (FM) using two wearable sensors, which allows capturing key frailty phenotypes (weakness, slowness, and exhaustion), by testing 20-second rapid elbow flexion-extension test.

Objective—In this study, we proposed an enhanced automated algorithm to identify frailty using a single wrist-worn sensor.

Methods—The data collected from 100 geriatric inpatients (age: 78.9±9.1 years, 49% frail) were re-analyzed to validate the new algorithm. The frailty status of participant was determined using a validated modified frailty index. Different FM phenotypes (31 features) including velocity of elbow rotation, decline in velocity of elbow rotation over 20 second, range of motion, etc. were extracted. A regression model, bootstrap with 2000 iterations, and recursive feature elimination technique were used for optimizing the FM parameters and identifying frailty using a single wrist-worn sensor.

Results—A strong agreement was observed between two-sensor and wrist-worn configuration ($r=0.87$, $p<0.001$). Results suggest that wrist-worn FM with no demographic information still yield high accuracy of 80.0% (95%CI: 79.7-80.3%) and area-under-curve of 87.7% (95%CI: 87.4-87.9%) to identify frailty status. Results are comparable with two sensors configuration, where the observed accuracy and area-under-curve were 80.6% (95%CI: 80.4-80.9%) and 87.4% (95%CI: 87.1-87.6%), respectively.

*Corresponding Author: Bijan Najafi, PhD, Professor, Interdisciplinary Consortium on Advanced Motion Performance (iCAMP), Michael E. DeBakey Department of Surgery, Baylor College of Medicine, 1 Baylor Plz, MS: BCM390, Houston, TX 77030, Phone: (713) 798-7536, Fax: (713) 798-1158, najafi.bijan@gmail.com (BN).

Conflict of Interest: Upper Extremity Frailty meter is protected by a patent pending (US Patent 20,150,332,004). The patent is owned by University of Arizona, and Bijan Najafi is listed as co-inventor on this patent pending. No other conflict of interest was reported by other authors.

Conclusion—The simplicity of FM may open new avenues to integrate wearable technology and mobile health to capture frailty status in busy hospital setting. Furthermore, reduction of needed sensors to a single wrist-worn sensor allows deployment of the proposed algorithm in form of a smart-watch application. From application standpoint, the proposed FM is superior to traditional physical frailty screening tools in which walking test is a key frailty phenotype, and thus has limitation to be used for bedbound patients or in busy clinics where administration of gait test as a part of routine assessment is impractical.

Keywords

Frailty Phenotype; Wearable technology; Geriatric population; Frailty; Frailty meter; mobile health

Introduction

Management and prediction of post-operative complications among older adults is becoming more important given the fact that population aging is rapidly changing the landscape of health care practice for inpatient care and hospital industries. Approximately 50% of surgeries in the United States (US) are performed in patients older than 65 years of age [1]. This group represents the fastest growing segment of the population and is expected to increase to nearly 89 million people by 2050 [2]. Older individuals have significantly higher rates of inpatient (~35.3% of all procedures) and outpatient (~32.1% of all procedures) surgical procedures compared with other age groups [3].

Frailty is a geriatric syndrome of decreased physiologic reserve and resistance to stressors, which leaves patients more vulnerable to poor health outcomes following surgical intervention or post hospital discharge, including adverse events (e.g. dead, cardiac arrest, stroke, etc), unfavorable discharge (e.g. discharge to skill nursing home), increase risk of readmission, worsening mobility post discharge, loss of independency, and prolonged hospitalizations [4-7]. Recognizing the necessity for quality improvements in geriatric surgical care, the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) and the American Geriatrics Society (AGS) recommends determining a baseline frailty score for optimal perioperative decision-making and management of the geriatric surgical patient [3]. Despite this recommendation, frailty as a part of routine pre-operative assessment is not widely used in hospital setting, mainly because of limitations of current screening tools.

The current frailty assessments could be divided into two major groups: 1) physical frailty or frailty phenotype (e.g. Fried's frailty phenotypes [7]), which often include assessing one or more phenotypes including physical activities, gait, motor performance (e.g. timed-up-and-go), exhaustion, weakness, and weight loss [7,8]; 2) deficit accumulation or frailty index (e.g. Rockwood's frailty index [9]): which is based on counting number of health deficits present (determined via survey, chart review, or patient electronic record) divided by the total number of potential health deficits that were considered [9,10]. Although these tools were beneficial to identify frailty, they are clinically cumbersome and time consuming for busy clinical settings. While frailty index screening tools are often considered more practical

for hospital setting compared to frailty phenotype tools due to ease of use and feasibility (e.g. time, training, and costs) [11], they have poor sensitivity to change over time [11,12]. Thus, they are not sensitive to screen benefit of pre-rehabilitation to reduce risk of adverse health outcomes pre-surgery and/or unable to determine modifiable risk factors prior surgical interventions [12-15]. On the other hand, the current frailty phenotype instruments are not suitable for mobility-impaired patients, time consuming, costly, needs trained staff to administer the test, required age and gender adjustment, and could be easily biased by methods of administration of gait tests such as type of footwear, presence of plantar ulcers or lower extremities amputation, or significant pain that limits their walking ability [7, 16-19].

Recent advances in wearable technologies, mobile health, telecommunications, and internet of thing (IoT) opened new opportunities to design a practical and automated tool, which may be able to facilitate clinical screening of frailty irrespective of setting. In particular, wide availability of mobile health platforms such as smart phone and smartwatch, which include inertial sensors (e.g. accelerometer, gyroscopes, etc.) and powerful processor, with ease of sharing digital information and/or exchanging data with electronic health records, open new horizons in development of digital health applications to provide clinicians with individualized objective phenotypic data on function and overall health [20]. The applications of these mobile health platforms have been already explored for management of chronic conditions such as screening tremor in patients with Parkinson's disease, detection of falls, or suspicious inactivity in elderly [21-24]. However, there have been no study done on a feasibility of using mobile health for screening frailty phenotype in geriatric population.

Recently, we have developed a Frailty Meter (FM) using two-sensor configuration, which allows capturing key physical frailty phenotypes (weakness, slowness, and exhaustion) by testing 20-second rapid elbow flexion-extension test [25-28]. We previously demonstrated that the FM is simple, quick, convenient and objective for assessing frailty status in community dwelling population as well as geriatric inpatients [25-28]. The FM was applicable to mobility-impaired patients unlike conventional frailty assessment tools which require mobility tests such as gait assessment [25,26]. However, the initial platform has still significant technical limitations for deployment as a low cost and stand-alone application in a mobile health platform. The previous version of the FM required supervised offline assessment to identify initiation of valid movement, manual controlling of signal quality and manual threshold adjustment based on a magnitude of arm flexion-extension speed which can be hard to be used by non-professional examiners. Besides, the analysis of FM test relies on patient information including age, gender and body mass index (BMI). Moreover, the FM platform needed two independent wearable devices attached on wrist and upper-arm. This in turn, increases technical complexity (e.g. synchronization between two modules) and cost (duplicate electronic elements such as processor, battery, etc.) as well as adds unnecessarily burden for the purpose of sensor attachment.

The aim of this study was to provide a physical frailty phenotype assessment tool using a single wrist-sensor without any additional information from demographic data. These efforts are in direction of facilitating the deployment of the algorithm in a smart phone platform or in widely available wrist-worn platforms such as smartwatch, Fitbit, Jawbone, etc. Our hypotheses are 1) the performance for identifying frailty using rapid arm flexion-extension

test without any patient information and with an automatic threshold setting will be comparable to that with the patient information and the manual threshold setting and 2) the performance using a single-sensor attached on wrist will be comparable to that using two-sensor configuration attached on both wrist and upper-arm like the previous version of FM.

Methods

Subjects and Protocol

A dataset of 100 geriatric inpatients over age of 65 (mean: 78.9 and standard deviation: 9.1, 49% frail) recruited from the Division of Trauma, Critical Care, and Emergency Surgery service at the University of Arizona were re-analyzed. The University of Arizona Institutional Review Board approved the study and all patients provided written consent.

Frailty status was measured by a validated modified frailty index (a shorten version of Rockwood frailty questionnaire [9]) named Trauma-Specific Frailty Index (TSFI) [29], suggested as a new practical tool to identify frailty for predicting discharge disposition in geriatric patients. The TSFI has been validated to be a stronger predictor of frailty than other trauma risk indicators such as Injury Severity Score, head Abbreviated Injury Scale and Glasgow Coma Scale scores [29]. Based on the TSFI score, the patients were divided in to nonfrail (score ≤ 0.27) and frail (score > 0.27) groups. Trained researchers performed all frailty assessments.

We analyzed three-dimensional angular velocity signals recorded by one wearable gyroscope device (LEGSys™, Biosensics LLC, Watertown, MA) attached on interior side of wrist. At the same time, we analyzed angular velocity signals recorded by another same-type wearable device attached on upper-arm near the triceps for demonstrating that use of one device is enough. Two wearable devices were fixed by a flexible band attached with the Velcro (Fig. 1). All signals were simultaneously recorded. Participants performed repetitive full elbow flexion and extension task as fast as possible in supine position (bed-bound position) for 20-second to quantify motor performance (Fig. 1). Participants were instructed that the upper-arm does not move during arm flexion and extension. The device data were recorded at a sampling rate of 100 Hz. Matlab software (The Mathworks, Inc., Natick, MA) was used for processed of the sensor data, design of a model, and statistical analysis of the model performance.

Preprocessing

Of three-dimensional signals, the signal for an axis of arm flexion/extension direction was considered for analyses. Since the aim of this study was to provide the algorithm to identify frailty using a single wearable device, data from the two devices were utilized for comparing performance of the algorithm with data acquired from the single device. When two sensors were used, angular velocity of elbow in respect of upper arm was estimated by calculating the difference between two signals.

Feature Extraction and Preselection

We extracted frailty phenotypes features from gyroscope sensors, which allow measuring angular velocity of wrist and upper arm. Figure 2a shows a typical angular velocity of wrist sensor measuring during 20-second rapid elbow flexion-extension. To computerize the phenotypes extraction (unlike our previous method, which required manual selection of onset of movement), we developed an algorithm that allows identifying the onset of 20-second arm flexion-extension movements using zero-crossing (ZC) technique and criteria described in the following. (Fig. 2):

Step #1: To avoid suppressing information of interest in both high frequency (e.g. jerkiness of movement) and low frequency (e.g. actual range of motion) domains, we didn't use any filtering. Instead, a series of rules were applied to distinguish a real flexion-extension signal from a noisy signal. This was done by estimation of three parameters including 1) the elapsed time between two consecutive detected ZC points as an indicator of flexion duration, 2) the range of angular velocity estimated between three consecutive ZC points, and 3) the magnitude of the maximum value as an indicator of maximum speed of rotation during flexion time. A valid ZC point is determined if each of the parameters described above exceeded a pre-defined threshold.

Step #2: Post processing: After detecting ZC points using the step #1, all maximum values (indicator of maximum angular velocity during flexion) were recalculated. Then if any estimated maximum value was less than 20% of the median value of all detected maximum values, the ZC points pre- and post the maximum value were removed. This allows to personalize the detection algorithm based on the slowness of movement.

Step #3: Detection of the onset and end of the test: After completing the step #2, the very first detected ZC point was considered as the onset of movement and the very last ZC point before 20-second interval in respect to the onset of movement was determined as the end of the test.

After detection all valid ZC points described above, 31 features from angular velocity pattern were extracted, unlike the previous study that introduced eight frailty features [26]. From the sensor, 10 key kinematic parameters (Table 1) were extracted and defined as follows: Angular velocity range (velocity of elbow flexion/extension, Fig 2b-1), angle range (angle was estimated by integrating signal between two consecutive ZC points and applying a drift removal algorithm, assuming that angle at each ZC point should return to zero value), power range (estimated by multiplication of angular velocity and angular acceleration. Angular acceleration was estimated with a single derivation of angular velocity), rising time (Fig 2b-2), falling time (Fig 2b-3), rising+falling time, flexion time (Fig 2b-4), extension time (Fig 2b-5), flexion+extension time, and flexion/extension rate defined as number of elbow flexion-extension per minute. The mean (M), coefficient of variance (CV: standard deviation divided by the mean), and percentage of decline (PD) were calculated from the 10 parameters described above resulting total of 30 features. The PD was defined as the percentage of decline of a parameter between the last and the first 10 seconds of elbow

flexion and extension task. The number of elbow flexion/extension during 20-second test was counted as the 31st feature.

To determine the significant features, between groups comparison (frail v. non-frail) was done using one-way ANOVA. Those features showing significant between group difference ($p < 0.050$) were selected for designing a multivariable model and determination of independent variables.

Optimal Feature Selection Based on Recursive Feature Elimination

Algorithm for optimal selection of parameters was based on linear regression modeling including bootstrap with recursive feature elimination technique. Linear regression model was employed to estimate TSFI as dependent variable from the selected features as independent variables. The recursive feature elimination technique was also utilized to obtain the optimal number of features for identifying between frail and nonfrail groups. The recursive feature elimination allows us to rank the most effective features, which then can be used to select the least number of needed features to results an optimum performance (e.g. accuracy, sensitivity, specificity, etc.) [30]. To be able to generalize the model, we used bootstrapping technique. This technique relies on random sampling with replacement. It allows us to resample available data to form a random set of training dataset and validation dataset. By randomly assign subjects to training and validation datasets, we could test any possible combination of subjects with different sample size [31,32]. In this study and considering our sample size ($n = 100$), we have used 2000 bootstrap iterations to optimize robustness of the model as recommended by the literature [32].

Figure 3 summarizes the schematic diagram of the approach used to select the optimal set of features. The bootstrap technique using 2000 iterations splits the subjects into 2000 pairs of training and validation datasets. This technique enables calculating 95% confidence interval for the algorithm performance (e.g. accuracy, sensitivity, specificity, etc.). The main steps for the recursive feature elimination are as follow. 1) First, the linear regression model is designed using all features except for the first one. 2) Then, performance of the model is calculated based on ability of the model to discriminate between frail and non-frail subjects. 3) The steps #1 and #2 are repeated by including all features expect for 2nd, 3rd, ..., nth (the last feature) from the model, at each iteration, respectively. 4) All estimated performances are compared and the lowest one is used to decide which feature should be eliminated from the training dataset. 5) The steps #1-4 are repeated till only one feature is remained in the model. This process allows ranking the top important features from 1st to nth. 6) The least optimal number of features are decided based on the resulted performance curve as a function of number of considered features. We did this decision, based on minimum number of features needed to reach almost steady state for all targeted performance of interest (e.g. accuracy, sensitivity, specificity, etc.). The method for estimation of these performances has been described in the following.

Quantification of the algorithm performance

The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy were calculated according to the following formulas:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100(\%) \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \times 100(\%) \quad (2)$$

$$\text{Positive predictive value} = \frac{TP}{TP+FP} \times 100(\%) \quad (3)$$

$$\text{Negative predictive value} = \frac{TN}{TN+FN} \times 100(\%) \quad (4)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \times 100(\%) \quad (5)$$

where true positive (TP) represents the number of correctly identified frail, false negative (FN) indicates the number of misdiagnosed as frail, true negative (TN) indicates the number of correctly identified nonfrail, and false positive (FP) stands for the number incorrectly identified as nonfrail. The area under the curve (AUC) was calculated as an additional evaluation of the performance.

Evaluation of the model performance

After determining the top features of interest (e.g. top 5) using the training dataset and recursive feature elimination method described above, we evaluated the model performance (e.g. accuracy, sensitivity, specificity, etc.) using the validation set only. The performance was assessed using wrist-sensor only as well as two-sensor configuration. Mean and 95% confidence interval value were calculated from the validation datasets (2,000 bootstrap iterations) for all performance parameters of interest.

Based on the linear model generated from the training dataset, we estimated the frailty index using each sensor configuration. To demonstrate the agreement between one-sensor (wrist-sensor) and two-sensor configuration, we examined the Pearson's correlation coefficient using the entire data set.

Results

Feature Extraction and Preselection

Table 2 shows the values of all 31 defined features in each frailty group extracted from wrist sensor and two-sensor configurations. Nine features were not shown to significantly differentiate between groups ($p < 0.05$). The other 22 features from each sensor configuration, marked with an asterisk in the Table 2, were chosen as independent variables in the linear regression model with recursive feature elimination technique.

Optimal Parameter Selection Based on Recursive Feature Elimination

Figure 4 illustrates the change in the algorithm performances (e.g. accuracy, sensitivity, etc.) as a function of number of features ranked using the recursive feature elimination technique. According to the figure 4, we set optimum target performance as 80% for sensitivity, specificity, accuracy, PPV, and NPV and greater than 90% for the AUC. Based on these criteria, the minimum number of needed features was 5 for either sensor configurations. The final equation for estimation of frailty index based on the optimized linear regression model is as follows:

$$FI_{est} = b + a_1Ph_1 + a_2Ph_2 + a_3Ph_3 + a_4Ph_4 + a_5Ph_5 \quad (6)$$

where FI_{est} is estimated frailty index, ' b ' is an intercept (0.24495 for the wrist-sensor, 0.39770 for two-sensor configuration), ' a ' denotes constant coefficients ($a_1 = -1.7357 \times 10^{-3}$, $a_2 = -1.2026 \times 10^{-3}$, $a_3 = 0.36848 \times 10^{-3}$, $a_4 = -0.49396$ and $a_5 = 0.48974$ for wrist-sensor and $a_1 = -0.43126 \times 10^{-3}$, $a_2 = -0.57419 \times 10^{-6}$, $a_3 = -0.50113 \times 10^{-3}$, $a_4 = 0.18680 \times 10^{-3}$ and $a_5 = 0.39429$ for two-sensor configuration) and ' Ph ' represents the selected features (frailty phenotypes). For the wrist-sensor, the selected phenotypes were Ph_1 : mean of angle range, indicator of rigidity, Ph_2 : percentage of decline of power range, indicator of exhaustion, Ph_3 : mean of elbow flexion time, indicator of slowness, Ph_4 : CV of elbow flexion time, indicator of lack of flexion steadiness, and Ph_5 : CV of elbow extension time, indicator of lack of extension steadiness. For the two-sensor configuration model, the selected phenotypes were Ph_1 : mean of angular velocity range, indicator of slowness, Ph_2 : mean of power range, indicator of weakness, Ph_3 : PD of power range, indicator of exhaustion, Ph_4 : mean of flexion time, indicator of slowness and Ph_5 : CV of elbow flexion/extension rate, indicator of lack of flexion-extension steadiness.

Evaluation of algorithm performance

Table 3 presents the performance of the model to identify frailty. The performance parameters were calculated when each model (i.e. wrist-sensor and two-sensor configuration) was tested against the validated dataset. A relatively high performance (in range of 79-83%) was obtained for all performance parameters including sensitivity, specificity, PPV, NPV, and accuracy irrespective of sensor configuration (Table 3). In addition, a strong agreement was observed between frailty index estimated using wrist-worn sensor and two-sensor configuration ($r = 0.870$, $p < 0.001$, Fig. 5).

Discussion

This study demonstrated that physical frailty phenotypes could be objectively assessed using a 20-seconds rapid elbow flexion-extension and a single wrist-sensor. This was done based on pre-identification of 31 features (physical frailty phenotypes) measurable by the wrist-sensor and an optimization approach based on recursive feature elimination technique. This model yielded five dominant features including 1) mean of elbow flexion time, indicator of slowness, 2) mean of elbow range of motion, indicator of rigidity, 3) inter-cycle variability of elbow extension time, indicator of lack of extension steadiness, 4) inter-cycle variability of elbow flexion time, indicator of lack of flexion steadiness, and 5) magnitude of decline in elbow rotation power after 20-second rapid elbow flexion-extension, indicator of exhaustion. When, the results of this model was compared to the gold standard (i.e. validated modified frailty index), a relatively high sensitivity, specificity, and AUC were observed, which were comparable with the results extracted from two-sensor configuration model. In addition, a strong agreement between wrist-worn and two-sensor configuration was observed when the frailty index (as a continuous scale between 0 and 1) was estimated from each sensor configuration.

The new developments could facilitate deployment of the physical frailty phenotype algorithm in widely available wearable platforms such as smart phone, smartwatch, and other wrist-worn wearable technologies (e.g., Jawbone, Fitbit, etc.). The model needs a single inertial sensor (i.e. gyroscope) attached to subject's wrist. From implementation stand point and calculation cost, only linear combination of five features was needed, which could reduce memory allocation and power cost. This in turn, could reduce battery consumption as well as ease of real-time implementation. In addition, the method is not dependent on subject's clinical and demographic information, which could make the system simpler without the need of an interface to enter any additional data for frailty status identification. The use of gyroscope instead of accelerometer could facilitate the ease of use by inexperienced users. For instance, unlike accelerometer based tools, which sensor's output is highly dependent on sensor location and method of sensor fixation, gyroscope is less dependent on sensor attachment since it measures speed of rotation, which by theory is remained unchanged along a rigid segment and less noisy than acceleration.

Results suggested that using optimal FM feature extraction from kinematic data measured by a wrist-sensor during elbow flexion and extension yield almost similar accuracy in comparison with using two-sensor configuration attached on both wrist and upper-arm suggested in a previous study [26]. This is a significant advance for the purpose of development and commercialization. Using two independent sensor, a module requires two independent processors, memories, and batteries. In addition, the synchronization between two wireless units is always challenging and drained significantly battery power reducing overall autonomy of the platform. Thus from commercialization standard point and technical design, a single sensor module to screen frailty is considered as a significant advantage compared to a two independent modules platform.

Interestingly, our results suggest that increasing the number of body segments/sensors (e.g. after including upper arm) does not necessarily increase the accuracy of physical frailty

assessment using rapid 20-second elbow flexion/extension test. Specifically, we noticed slightly lower specificity for two-sensor configuration compared to wrist-sensor. This could be explained by potentially higher noise/artifacts generated by skin movement at the upper arm location compared to wrist location. This is similar to observation in Najafi et al. study [33], in which it was demonstrated that addition of wearable sensors to body segments with significant soft tissue mass (e.g. thigh segment) is not necessary increase the accuracy of human body model to estimate kinematic parameters of interest.

Automated and quick identification of frailty may empower doctors and institutions to better inform patients and their family about potential risk of hospitalization and/or major surgical interventions [34]. In particular, quick screening of physical frailty phenotypes irrespective of age could assist in identifying potential modifiable risk factors to improve health outcomes and/or management of post discharge care [16,35]. Although age is strongly associated with postoperative adverse outcomes, the overall response of an older individual to surgical intervention is highly variable and can be affected by their level of fitness [36]. This variability makes age-based clinical decisions difficult. For instance, our recent findings showed that while age was not significantly associated with readmission, physical frailty phenotype screening has the potential to predict 30-day and 60-day readmission among geriatric trauma patients with a relatively high odds ratio [28]. Thus, integrating a rapid and practical frailty phenotype screening tool in hospital setting, may have potential to identify high risk individuals and implement personalized intervention to reduce risk of readmissions or poor health outcomes post discharge.

This study suffers from few limitations. First, this is a cross-sectional cohort study with aim to examine whether a single sensor derived parameters would be sufficient to identify frailty status in inpatient geriatric population. Another prospective study should be addressed to examine whether such algorithm is also sensitive to predict adverse health outcomes (e.g. readmission, dead, etc.) and discharge disposition (e.g. discharge to skill nursing home). Second, while the method proposed in this study is based on determining physical frailty phenotypes to identify frailty status, a validated frailty index was used as a gold standard instead of a validated frailty phenotype assessment tool. This selection was because of limitation in conventional frailty phenotype assessment tools for our population, who had walking limitation. However, in our previous study, in which frailty status was assessed in ambulatory community dwelling older adults using two sensor configuration, a relatively good agreement with Fried's frailty criteria, which is considered as one of the most-established frailty phenotype screening tool, was observed [25]. Another study, needs to be addressed whether a single sensor derived parameters has also similar agreement with Fried's criteria in community dwelling older adults. Third, this study was limited to assess frailty status in inpatient geriatric population admitted to hospital because of ground level falls, which could be too homogeneous. Thus, the results may not be generalized to all inpatient geriatric population including those who are admitted to hospital for the purpose of surgical intervention. Lastly, the sample size ($n = 100$) may be insufficient for the purpose of model validation. However, since this was a cohort study design and half of the recruited subjects were frail, and five independent predictors were selected using recursive feature eliminations and bootstrap method with 2000 iterations, we believe the selected sample should be sufficient for the purpose of this accuracy assessment study. While approximately

80% accuracy to detect frailty status in comparison to the frailty index (i.e. cumulative deficit frailty) seems to be reasonable [29], further improvement may be possible and should be addressed in future studies. This could be done by adding demographic and clinical information (e.g. age, BMI, history of falls, gender, etc.), additional biomechanical measurements (e.g. muscle activities), adding other sensors (e.g. surface EMG), and/or using a different or additional physical tests (e.g. fist making/fist release).

Despite, the above limitations, we believe the results are promising and support the idea of using a simple 20-second physical frailty phenotype assessment tool based on a single wrist-worn sensor and elbow rapid arm test to quickly assess frailty status in geriatric inpatients. In our previous studies [25,26], we have demonstrated that 20-second rapid arm flexion-extension using two wearable sensor configuration yield a relatively high agreement with two well-established frailty assessment tools including physical frailty (Fried's frailty phenotypes) [25] and deficit accumulation (i.e. frailty index) [26]. In addition, we have demonstrated that such short test is sufficiently accurate to predict adverse health outcomes post hospital discharge in geriatric population [28]. Thus, considering that this study demonstrate a high agreement between wrist-sensor and two-sensor configuration, we may conclude that wrist-worn sensor configuration may be sufficient to predict adverse health outcomes in geriatric population. This simple tool may address some of the limitations of both frailty index and frailty phenotype tools from point of view of time-efficiency, ease of use and practicality for both ambulatory and non-ambulatory patients, lack of requirement for trained staff or dedicated infrastructure, objectivity, and real-time assessment of frailty status.

In conclusion, we developed an automatic frailty screening algorithm which allows identifying physical frailty phenotypes using a single wrist-worn sensor without any subject's demographics including gender, age and BMI. This will facilitate development of frailty assessment in low cost or widely available platforms such as smartwatch and smart phones. It would be a more practical tool for busy clinics without the need of specific training or additional measurement. The use of a single sensor and its low calculation cost (i.e. using only five parameters) could facilitate deployment of the proposed algorithm in widely available wearable platforms such as smartwatch and smartphone. From application standpoint, the proposed FM is superior to traditional frailty phenotype screening tools in which walking test is a key physical frailty phenotype, and thus has limitation for screening of bedbound patients or in busy clinics where administration of gait test as a part of routine assessment is impractical.

Acknowledgments

We thank Bardiya Zangbar, Nima Toosizadeh, Michelle R Heusser and Tahereh O Jokar for data collection and patients' recruitment, and clinical coordination. Partial support was provided by the National Institutes of Health/ National Institute of Aging (Award number: 1R44AG050338-0). The content is solely the responsibility of the authors and does not necessarily represent the official views of the sponsor.

References

1. Etzioni DA, Liu JH, O'Connell JB, Maggard MA, Ko CY. Elderly patients in surgical workloads: A population-based analysis. *The American surgeon*. 2003; 69:961–965. [PubMed: 14627256]

2. Center for Disease Control and Prevention. Health related quality of life: nationwide trends. Centers for Disease Control and Prevention; WhangHTaSC:
3. Chow WB, Rosenthal RA, Merkow RP, Ko CY, Esnaola NF. American College of Surgeons National Surgical Quality Improvement P, American Geriatrics S. Optimal preoperative assessment of the geriatric surgical patient: A best practices guideline from the american college of surgeons national surgical quality improvement program and the american geriatrics society. *Journal of the American College of Surgeons*. 2012; 215:453–466. [PubMed: 22917646]
4. Makary MA, Segev DL, Pronovost PJ, Syin D, Bandeen-Roche K, Patel P, Takenaga R, Devgan L, Holzmueller CG, Tian J. Frailty as a predictor of surgical outcomes in older patients. *Journal of the American College of Surgeons*. 2010; 210:901–908. [PubMed: 20510798]
5. Robinson TN, Wu DS, Pointer L, Dunn CL, Cleveland JC Jr, Moss M. Simple frailty score predicts postoperative complications across surgical specialties. *American journal of surgery*. 2013; 206:544–550. [PubMed: 23880071]
6. Lasithiotakis K, Petrakis J, Venianaki M, Georgiades G, Koutsomanolis D, Andreou A, Zoras O, Chalkiadakis G. Frailty predicts outcome of elective laparoscopic cholecystectomy in geriatric patients. *Surgical endoscopy*. 2013; 27:1144–1150. [PubMed: 23052539]
7. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, McBurnie MA. Cardiovascular Health Study Collaborative Research G. Frailty in older adults: Evidence for a phenotype. *The journals of gerontology Series A, Biological sciences and medical sciences*. 2001; 56:M146–156.
8. Schwenk M, Howe C, Saleh A, Mohler J, Grewal G, Armstrong D, Najafi B. Frailty and technology: A systematic review of gait analysis in those with frailty. *Gerontology*. 2014; 60:79–89. [PubMed: 23949441]
9. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A. A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005; 173:489–495. [PubMed: 16129869]
10. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. *BMC geriatrics*. 2008; 8:24. [PubMed: 18826625]
11. Cesari M, Gambassi G, van Kan GA, Vellas B. The frailty phenotype and the frailty index: Different instruments for different purposes. *Age and ageing*. 2014; 43:10–12. [PubMed: 24132852]
12. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. *Lancet*. 2013; 381:752–762. [PubMed: 23395245]
13. Yang F, Gu D. Predictability of frailty index and its components on mortality in older adults in china. *BMC geriatrics*. 2016; 16:145. [PubMed: 27455962]
14. Rockwood K, Mitnitski A. Frailty defined by deficit accumulation and geriatric medicine defined by frailty. *Clinics in geriatric medicine*. 2011; 27:17–26. [PubMed: 21093719]
15. Robinson TN, Walston JD, Brummel NE, Deiner S, Brown CHt, Kennedy M, Hurria A. Frailty for surgeons: Review of a national institute on aging conference on frailty for specialists. *Journal of the American College of Surgeons*. 2015; 221:1083–1092. [PubMed: 26422746]
16. Rockwood K, Andrew M, Mitnitski A. A comparison of two approaches to measuring frailty in elderly people. *The journals of gerontology Series A, Biological sciences and medical sciences*. 2007; 62:738–743.
17. Carriere I, Colvez A, Favier F, Jeandel C, Blain H. group Es. Hierarchical components of physical frailty predicted incidence of dependency in a cohort of elderly women. *Journal of clinical epidemiology*. 2005; 58:1180–1187. [PubMed: 16223662]
18. Ravaglia G, Forti P, Lucicesare A, Pisacane N, Rietti E, Patterson C. Development of an easy prognostic score for frailty outcomes in the aged. *Age and ageing*. 2008; 37:161–166. [PubMed: 18238805]
19. Studenski S, Hayes RP, Leibowitz RQ, Bode R, Lavery L, Walston J, Duncan P, Perera S. Clinical global impression of change in physical frailty: Development of a measure based on clinical judgment. *Journal of the American Geriatrics Society*. 2004; 52:1560–1566. [PubMed: 15341562]
20. Mohler J, Najafi B, Fain M, Ramos KS. Precision medicine: A wider definition. *J Am Geriatr Soc*. 2015; 63:1971–1972. [PubMed: 26390003]

21. Wile DJ, Ranawaya R, Kiss ZH. Smart watch accelerometry for analysis and diagnosis of tremor. *J Neurosci Methods*. 2014; 230:1–4. [PubMed: 24769376]
22. Deutsch M, Burgsteiner H. A smartwatch-based assistance system for the elderly performing fall detection, unusual inactivity recognition and medication reminding. *Stud Health Technol Inform*. 2016; 223:259–266. [PubMed: 27139412]
23. Casilari E, Oviedo-Jimenez MA. Automatic fall detection system based on the combined use of a smartphone and a smartwatch. *PLoS One*. 2015; 10:e0140929. [PubMed: 26560737]
24. Mortazavi B, Nemati E, VanderWall K, Flores-Rodriguez HG, Cai JY, Lucier J, Naeim A, Sarrafzadeh M. Can smartwatches replace smartphones for posture tracking? *Sensors (Basel)*. 2015; 15:26783–26800. [PubMed: 26506354]
25. Toosizadeh N, Mohler J, Najafi B. Assessing upper extremity motion: An innovative method to identify frailty. *Journal of the American Geriatrics Society*. 2015; 63:1181–1186. [PubMed: 26096391]
26. Toosizadeh N, Joseph B, Heusser MR, Orouji Jokar T, Mohler J, Phelan HA, Najafi B. Assessing upper-extremity motion: An innovative, objective method to identify frailty in older bed-bound trauma patients. *Journal of the American College of Surgeons*. 2016; 223:240–248. [PubMed: 27155751]
27. Toosizadeh N, Najafi B, Reiman EM, Mager RM, Veldhuizen JK, O'Connor K, Zamrini E, Mohler J. Upper-extremity dual-task function: An innovative method to assess cognitive impairment in older adults. *Frontiers in aging neuroscience*. 2016; 8:167. [PubMed: 27458374]
28. Joseph B, Toosizadeh N, Orouji Jokar T, Heusser MR, Mohler J, Najafi B. Upper-extremity function predicts adverse health outcomes among older adults hospitalized for ground-level falls. *Gerontology*. 2017; 63:299–307. [PubMed: 27941328]
29. Joseph B, Pandit V, Zangbar B, Kulvatunyou N, Tang A, O'Keeffe T, Green DJ, Vercruysse G, Fain MJ, Friese RS, Rhee P. Validating trauma-specific frailty index for geriatric trauma patients: A prospective analysis. *Journal of the American College of Surgeons*. 2014; 219:10–17 e11. [PubMed: 24952434]
30. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Machine Learning*. 2002; 46:389–422.
31. Zhu W. Making bootstrap statistical inferences: A tutorial. *Res Q Exerc Sport*. 1997; 68:44–55. [PubMed: 9094762]
32. Efron, B., Tibshirani, RJ. *An introduction to the bootstrap*. CRC press; 1994.
33. Najafi B, Lee-Eng J, Wrobel JS, Goebel R. Estimation of center of mass trajectory using wearable sensors during golf swing. *J Sports Sci Med*. 2015; 14:354–363. [PubMed: 25983585]
34. Lin HS, Watts JN, Peel NM, Hubbard RE. Frailty and post-operative outcomes in older surgical patients: A systematic review. *BMC geriatrics*. 2016; 16:157. [PubMed: 27580947]
35. Van Epps P, Oswald D, Higgins PA, Hornick TR, Aung H, Banks RE, Wilson BM, Burant C, Graventstein S, Canaday DH. Frailty has a stronger association with inflammation than age in older veterans. *Immunity & ageing : I & A*. 2016; 13:27. [PubMed: 27777599]
36. Partridge JS, Harari D, Dhesei JK. Frailty in the older surgical patient: A review. *Age and ageing*. 2012; 41:142–147. [PubMed: 22345294]
37. Coley B, Jolles BM, Farron A, Bourgeois A, Nussbaumer F, Pichonnaz C, Aminian K. Outcome evaluation in shoulder surgery using 3d kinematics sensors. *Gait Posture*. 2007; 25:523–532. [PubMed: 16934979]

Figure 1. Elbow flexion/extension test associated with the frailty meter (FM). Using a 20-second rapid arm flexion-extension different physical frailty phenotypes including slowness, exhaustion, weakness, and flexibility were extracted using a wrist-worn sensor.

Figure 2. Sensor output of a 20-second rapid arm flexion/extension test. (a) Example of onset and end of the test detected automatically by the proposed algorithm. (b) Definition of angular velocity range (1), rising time (2), falling time (3), elbow flexion time (4), and elbow extension time (5).

Figure 3. Flowchart for optimal feature selection. PPV stands for positive predictive value, NPV refers to negative predictive value and AUC indicates area under curve.

Figure 4.

Selection of the optimal number of features based recursive feature elimination. Five features (FM phenotypes) were selected for each sensor configuration as the least number of features to yield steady state and relatively highest performance among all performance parameters of interest. The selected number of features yields an AUC of over 90% and over 80% performance among the rest.

Figure 5.
Association between estimated frailty index using wrist-sensor and two-sensor configuration.

Table 1

Definition of kinematic parameters extracted from the wearable sensors. Mean, coefficient of variance (CV) and percentage of decline (PD) of all measures except for the number of flexion/extension were calculated as features (total of 30 features). The number of flexion/extension was utilized as the 31st feature.

Extracted measures	Definition
Angular Velocity Range	Range of angular velocity estimated by difference between maximum and minimum angular velocity peaks (Fig. 2b-1)
Angle Range	Range of flexion-extension angle
Power Range	Product of the angular velocity range and angular acceleration range [37]
Rising time	Elapsed time to reach the maximum angular velocity during flexion (Fig. 2b-2)
Falling time	Elapsed time to reach the minimum angular velocity during extension (Fig. 2b-3)
Rising+Falling time	Sum of rising and falling times
Elbow Flexion time	Duration of elbow flexion (Fig. 2b-4)
Elbow Extension time	Duration of elbow extension (Fig. 2b-5)
Elbow Flexion+Extension time	Sum of elbow flexion and extension times
Elbow Flexion/Extension rate	Number of elbow flexion/extension per minute
Number of flexion/extension	Number of flexion/extension during 20 seconds test

Table 2

Statistical results for comparing features between nonfrail and frail groups.

Features	Wrist-sensor (1-sensor)						Two-sensor					
	Nonfrail			Frail			Nonfrail			Frail		
	Mean	S.D.	p-value	Mean	S.D.	p-value	Mean	S.D.	p-value	Mean	S.D.	p-value
Angular Velocity Magnitude (deg/s)	Mean	834.03	256.45	479.85	228.10	<.001*	879.48	272.55	<.001*	469.25	219.50	<.001*
	CV	0.10	0.04	0.14	0.08	.002*	0.10	0.04	.001*	0.15	0.08	.001*
	PD (%)	-7.40	8.83	-11.40	12.09	.064*	-7.41	8.92	.071*	-11.35	12.21	.071*
Angle Magnitude (deg)	Mean	106.67	25.89	81.35	31.00	<.001*	106.18	31.15	<.001*	78.51	30.69	<.001*
	CV	0.09	0.05	0.13	0.10	.013*	0.09	0.05	.002*	0.14	0.10	.002*
	PD (%)	-6.38	7.66	-9.82	12.29	.099*	-6.96	7.02	.084*	-10.52	12.43	.084*
Power Magnitude (deg ² /s ³)	Mean ($\times 10^4$)	130.23	104.68	37.30	53.89	<.001*	135.25	115.40	<.001*	36.45	53.26	<.001*
	CV	0.24	0.10	0.32	0.13	.003*	0.25	0.10	.001*	0.33	0.13	.001*
	PD (%)	-9.03	26.95	-19.58	24.01	.043*	-8.87	27.91	.049*	-19.27	23.33	.049*
Rising Time (ms)	Mean	267.67	78.34	362.38	132.00	<.001*	260.16	84.79	<.001*	358.39	120.93	<.001*
	CV	0.19	0.11	0.27	0.24	.038*	0.21	0.13	.050*	0.30	0.25	.050*
	PD (%)	1.20	17.91	12.50	40.78	.079*	2.37	19.09	.108*	13.78	45.36	.108*
Falling Time (ms)	Mean	261.56	91.34	338.98	135.73	.001*	257.58	90.67	.001*	341.33	149.48	.001*
	CV	0.19	0.09	0.26	0.20	.023*	0.18	0.10	.007*	0.26	0.19	.007*
	PD (%)	-0.93	14.18	0.90	23.52	.641*	-1.57	13.84	.732*	-0.24	23.31	.732*
Rising+Falling Time (ms)	Mean	264.61	77.80	350.68	122.16	<.001*	258.87	77.92	<.001*	349.86	122.12	<.001*
	CV	0.23	0.10	0.31	0.22	.020*	0.25	0.10	.034*	0.32	0.23	.034*
	PD (%)	0.37	14.53	3.84	22.13	.360*	0.02	15.13	.227*	4.76	22.69	.227*
Flexion Time (ms)	Mean	419.98	129.98	644.18	357.60	<.001*	419.54	124.79	<.001*	641.17	332.14	<.001*
	CV	0.09	0.05	0.15	0.15	.005*	0.09	0.05	.003*	0.16	0.16	.003*
	PD (%)	0.69	10.11	5.63	19.19	.114*	0.77	10.96	.128*	6.04	21.41	.128*
Extension Time (ms)	Mean	427.36	132.86	669.61	343.72	<.001*	427.71	137.60	<.001*	672.65	369.50	<.001*
	CV	0.09	0.05	0.17	0.23	.014*	0.09	0.04	.019*	0.17	0.23	.019*

Features	Wrist-sensor (1-sensor)						Two-sensor					
	Nonfrail Mean	S.D.	Frail Mean	S.D.	p-value		Nonfrail Mean	S.D.	Frail Mean	S.D.	p-value	
Flexion+Extension Time (ms)	PD (%)	1.27	11.75	-0.90	16.34	.452 *	1.02	10.59	-1.06	15.88	.448 *	
	Mean	423.67	127.51	656.89	344.40	<.001 *	423.63	127.46	656.91	344.41	<.001 *	
	CV	0.11	0.05	0.19	0.18	.004 *	0.11	0.05	0.20	0.19	.003 *	
Flexion/Extension Rate (per min.)	PD (%)	0.34	9.51	3.61	15.53	.213 *	0.36	9.50	3.69	15.50	.203 *	
	Mean	77.28	22.02	54.73	19.86	<.001 *	77.22	22.01	54.73	19.87	<.001 *	
	CV	0.06	0.04	0.09	0.07	.007 *	0.06	0.04	0.09	0.07	.007 *	
Number of Flexion Extension	PD (%)	-0.49	7.54	-0.67	7.42	.905 *	-0.49	7.55	-0.65	7.41	.918 *	
		25.08	7.24	17.29	6.61	<.001 *	25.08	7.24	17.29	6.61	<.001 *	

* Asterisk indicates significantly difference between nonfrail and frail groups; CV: coefficient of variation; PD: percentage of decline; S.D.: standard deviation

Table 3

Performance of each model (sensor configuration) for identifying nonfrail and frail groups.

	Wrist-sensor <i>Mean [95% CI]</i>	Two-sensor <i>Mean [95% CI]</i>
Optimal number of features	5	5
Sensitivity (%)	79.60 [79.08 to 80.13]	82.94 [82.54 to 83.35]
Specificity (%)	81.23 [80.72 to 81.73]	78.84 [78.38 to 79.30]
Positive predictive value (%)	80.95 [80.49 to 81.42]	79.52 [79.12 to 79.93]
Negative predictive value (%)	80.94 [80.48 to 81.40]	82.98 [82.59 to 83.37]
Accuracy (%)	79.99 [79.69 to 80.29]	80.62 [80.37 to 80.87]
Area under the curve (%)	87.66 [87.41 to 87.91]	87.37 [87.14 to 87.60]