Frailty level prediction in older age using hand grip strength functions over time

Elsa Pérez[1], Jose E. Torres Rangel[1], Marta Musté[1], Carlos Pérez[2], Oscar Macho[2], Francisco S. del Corral Guijarro[3], Aris Somoano[4], Cristina Gianella[4], Luis Ramírez[4], Andreu Català[1]

[1] CETpD-UPC, Technical University of Catalonia, Vilanova i la Geltrú, Spain
[2] Consorci Sanitari de l'Alt Penedès i Garraf
[3]Universidad Politécnica de Madrid, Spain
[4] Hospital Central de la Cruz Roja San José y Santa Adela de Madrid, Spain
elsa.perez@upc.edu

Abstract. Frailty syndrome can be defined as a clinical state in which there is a rise in individual vulnerability, developing an increase in both the dependence of the person and mortality. Frailty is completely related to age. A fundamental factor to apply rehabilitative interventions successfully resides in having a simple and reliable method capable of identifying frailty syndrome.

Frailty indexes (FI) have several sources of uncertainty trough the opinion of the patients, white coat effect and external factors. Moreover, in the clinical practice, the experience of the geriatricians led them to determine an approximation of the frailty level only with a simple handshake. Hand grip strength (HGS) has been widely used in tests by investigators and therapists to be able to diagnose sarcopenia and frailty, as it is a reliable indicator of the overall muscle strength, which decreases with age. Most researches focused mainly on peak HGS, which will not give insight on how the patient's strength was distributed over time. In the present work it is proposed to evaluate HGS behavior over a period of time, and to develop a system based on Machine Learning for the identification of frailty levels using physiological features, FI and the classical signal processing based on statistics of the HGS signals.

The starting hypothesis is that it can be identified the "way" of performing HGS correlated with the level of frailty. To achieve this goal a clinical study was designed and carried out with a cohort of 70 elderly persons, in two Hospitals.

Keywords: Frailty identification, Hand grip strength, Machine Learning

1 Introduction

Frailty syndrome can be defined as a clinical state in which there is a rise in individual vulnerability, developing an increase in both the dependence of the person and mortality when exposed to a stressor. Frailty is completely related to age, being highly prevalent in the elderly, reaching up to 30% in people over 75 years of age.

There are many questionnaires that define various frailty factors, but there is no accepted standardization on this. In addition, these methods require the opinion of the patient, whose criteria may vary depending on the patient, thus generating an imprecise diagnosis. Geriatricians consider that this clinical condition considerably increases health risks [1] and even death [2].

On the other hand, there is evidence through several studies that show that the appearance of frailty can be anticipated, delayed or even avoided [9]. Therefore, a fundamental factor to apply rehabilitative interventions successfully resides in having a simple, effective and reliable method capable of identifying people with frailty syndrome.

At present there is no common criterion to quantify frailty. Most questionnaires are based on asking the patient about various symptoms and noting which ones he or she manifests or perceives [3], [4], [5]. One of these methods is that of Fried et al. [17], which involved 5210 people over 65 years of age, who proposed that the frailty phenotype is defined by the presence of three or more of the following symptoms: unintended weight loss, weakness, low resistance, slowness of movement, low activity.

In the 90s it was demonstrated the usefulness of VIG (comprehensive geriatric assessment) [18] to evaluate frailty in the elderly. The VIG is a global diagnostic tool or methodology at all levels of care, it is designed to identify and quantify biomedical and pharmacological data, physical, functional, psychological and social problems that the elderly may present.

Both Fried and other indices or scales exclusively use the maximum value of the grip strength of the hand as one of the symptoms, but in clinical practice the geriatrician uses his experience to make quick diagnoses based on the "form" in which the patient performs a handshake.

1.1 HGS Related work

Hand grip strength (HGS) has been widely used in tests by investigators and therapists to be able to diagnose sarcopenia and frailty, as it is a reliable indicator of the overall muscle strength, which decreases with age. Results obtained from these tests have been used to verify if HGS can indeed work as a predictor of disability in older men [6]. These tests are tied to recent protocols, such as Southampton protocol or the one proposed by the ASHT (American Society of Hand Therapists), made to try and establish a common ground for different studies.

However, even after updating these protocols recently (as recent as 2015), there is still a lack of consistency when it comes to evaluate HGS over a period of time. There are studies which aim to gather data from other studies that measured HGS to diagnose sarcopenia and frailty and identify the differences in the protocols used [7], which is an important focus for the present research, as the protocol that is to be proposed will use others as means for comparison and innovation.

When a protocol is taken for a specific study, there are a few main elements that can be appreciated and need to be highlighted from the beginning, such as the dynamometer used to measure HGS, which hand was used, the subject’s posture, arm position, handle position, how long did the measurement take or how long were the intervals between the measurements. the recommended protocol to follow is the most recent ASHT protocol, as it is the most detailed one, and if a modification should be made, it is to be mentioned [8].

Another parameter which has also shown to have correlation to grip strength is the body mass index (BMI). Again, data was collected with a Jamar dynamometer and using healthy males as subjects aged 20 – 74 years [9].

Even though a wide range of instruments was used among the majority of studies (Smedley, Martin, Tekdyne, among others), the predominant dynamometer used was de Jamar dynamometer [10].

Regarding sincerity of effort, that is, whether a genuinely maximal effort is being given during clinical strength testing, there are several studies that have examined the force-time curve produced by maximal and submaximal effort [11]. By using a specialized dynamometer with a force transducer (Biopac Instruments), with a test time of 5 seconds, a 30-second rest interval between trials, the function obtained had the form of a step.

Respecting the time for test, a 6-second test was found to have a higher reliability despite gender or hand dominance, in contrast to a 10-second test, which did not have results as reliable as the first [12].

If HGS was to be measured over time, and plot a strength curve for the same period of the procedure, more valuable information could be obtained regarding how HGS really determines patients' muscle strength, or perhaps, even go as far as being able to diagnose more efficiently frailty or sarcopenia by extracting determined features from it. Some studies studied the slope of the force-time curve related to sincerity of effort with a Jamar dynamometer and following the protocol recommended by the ASHT [13, 14].

Another study that aimed to investigate the force-time characteristics during a sustained maximal grip effort, according to age and clinical condition [15] was consulted, in which a sustained maximal grip was continuously recorded by using a modified Martin vigorimeter. The investigators concluded that the force-time characteristics during a sustained maximal handgrip effort are significantly different according to age and clinical condition. Old patients were characterized by a rather fast decline in muscle work during the first part of sustained grip.

1.2 HGS and the level of frailty

In the present work, a system based on Machine Learning for identifying the levels of frailty is developed using the features of the grip force signal in a determined period of time. The objectives of this study are two: to perform the main frailty indexes VIG, Fried, Frail, with a cohort of elderly persons, done by geriatricians, in a transversal pilot together with a designed test of HGS in a period of time, and with the created database to develop different Machine Learning strategies to extract significant information and knowledge useful for the detection of frailty tendencies using the results of only one simple test. The starting hypothesis is that we can identify the "way" of performing HGS correlated with the level of frailty.

2 Material and Methods

The following lines are dedicated to the description of the instrument and the cohort .

2.1 Instrumentation

The instrument used for the present study was a modified Deyard dynamometer, being the modified part the whole electronic circuit, which was replaced by one designed and made by the CETpD, the rest of the model, meaning the mechanical design, remained the same as the original. Thanks to this modification, the dynamometer measures the HGS continuously in time.

Fig. 1. Modified Deyard dynamometer

The modification includes the ability to store information and Bluetooth connectivity with the IMU (Inertial Measurement Unit) developed by the CETpD [16] 15 for long-term monitoring of human pathological movement.

2.2 Calibration

Calibration was necessary to set the accuracy of the modified version to an acceptable level. For this, weights that ranged from 5 to 40 Kg were used (which is more than the max force expected for the tested population), with an increasing rate of 5 Kg per trial. The dynamometer was held by two metallic bars, which were placed in the space between the handle and the screen, where no disruption should be presented for the test. A belt was tied to the base for the weights (extra 731.8 grams) and to the handle, as centered as possible, note that the belt is made of a non-stretchable material, as to not influence the result of the calibration tests.

Fig. 2. Calibration process

Its calibration curve resulted as follows:

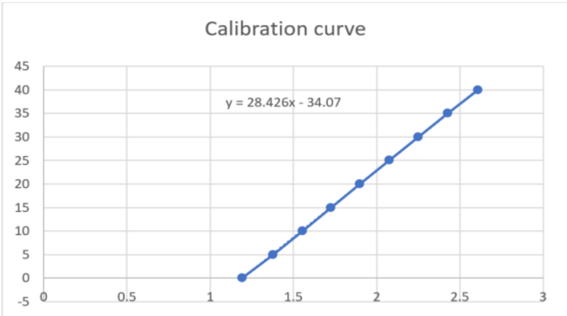


Fig. 3. Calibration curve Force (Kg) Vs Voltage (V)

For which the slope equation for the curve is as shown:

$$y = 28.426 * x - 34.07$$

Considering the weight of the balance used for the calibration trials (731.8 g) it is then:

$$y = 28.426 * x - (34.07 - 0.7318)$$

Where y is the force total (Kg), and x represents the voltage (V) measured.

2.3 Pilot Protocol

The protocol designed includes the cohort, the clinical study, which includes the registration of several frailty indexes, and the performance of 3 HGS test recorded with a modified Deyard dynamometer to store and transmit the produced signals. This protocol was carried out by geriatricians, with a cohort of 70 elderly persons, in two Hospitals, (*Hospital Central de la Cruz Roja San José* and *Santa Adela de Madrid, Consorci Sanitari de l'Alt Penedès i Garraf*). All data were captured in a single visit, where the general inclusion criteria were applied and each participant was assigned their corresponding level of frailty. Geriatricians performed the HGS test and evaluated the following scales in clinical trials: Fried criteria [17], Fragile Vig Index (VIG) [18], Barthel scale [19] and Lawton – Brody scale [20]. The protocol was approved by the "Comité de Ética de la Investigación con Medicamentos de la Comunidad de Madrid" (Ref 47/916546.9/19).

The HGS test consists of 3 trials carried out in a sitting stance take on a chair, forearm placed on top of the leg in neutral position (holding the dynamometer perpendicular to the leg), feet firm on the floor at shoulder-width distance, shoulder adducted, neutrally rotated and using the dominant hand. Encouragement was also one aspect to keep into consideration, as it was used as well. Tests had a duration of 6 seconds each, and a rest interval of 1 minute between tests.

Fig. 4. Protocol position example

2.4 Data acquisition and processing

To plot the force-time curves, the data was stored in the memory card inside the IMU and inserted in the PC to run the Matlab script, which acquired the signal, filter it, establish the desired range for treatment, and then segment the resulting signal in three phases: the Force-generation phase (FGP), force maintenance (FM) and the Force-decay phase (FDP) (Fig. 5). These phases will be used to extract different proposed features with a Matlab script, specially designed for this function. These specific characteristics can be identified to truly be able to detect whether a patient is prone to developing frailty in the future or not.

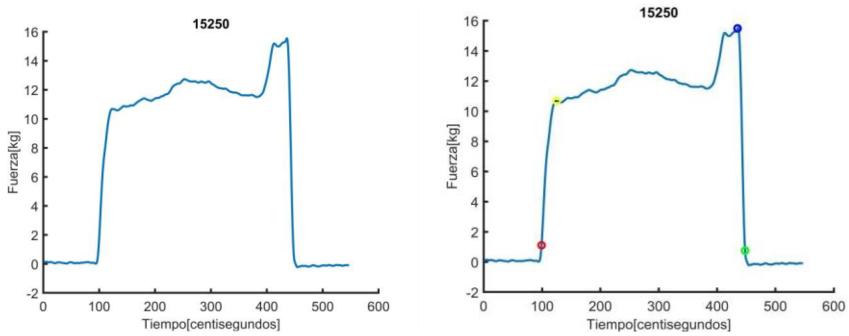


Fig. 5. Filtered signal sample from a patient and the 3 segments of the resulting signal.

2.5 Features

The features to consider for each patient are divided in 2 large groups: Physiological information compiled during the VIG & FRIED Tests performed to each patient and a group of between 1 to 3 HGS signals for each patient.

Physiological information compiled during the VIG & FRIED Tests performed to each patient.

In total we gather around 92 features from the VIG and Fried Tests from those we only took interest in 5: age, gender, weight, height, ICM. This was because the need was to select as few features as possible that were related with the muscular strength of the patients and that were easy and simple to get.

HGS signals for each patient.

First is important to remember that the signal is a non-structured data because it does not have a fixed number of samples and behaves as a time dependent variable. So, the first step was to convert the signal into a structured group of data. To do so we took inspiration in the work of industrial control signal featurization since the HGS could be interpreted as a response to a step input function and the features selected to represent each of the 3 segments (Generation, Maintenance and Decay phases) of the HGS signal were: Initial time, Initial Strength, Final time, Final Strength, Area, Density, Minimum Strength, Time of occurrence of Minimum Strength, Maximum Strength, Time of occurrence of Maximum Strength, Mean Strength, Median Strength, Strength Scope (Last-First),Mean Slope, Median Slope, Maximum Slope, Minimum Slope and Over-Peak. Meaning that the HGS Signals were converted into 54 structured features.

Summarizing, each sample of each patient have 59 features. It was decided to work with the samples because we could upgrade our number of observations from 83 patients to 235 samples.

2.6 Targets

The original DDBB possessed 3 possible outputs or diagnosis the ones from VIG, FRIED and Estratos. Each of them had 3 possible output classifications: Frail, prefrail and Sturdy. The ideal distribution of classifications should be 33.33/33.33/33.33 so the more the actual distributions approach to it the better.

The Output of the VIG Diagnosis was selected to use as target, since: FRIED diagnosis had only 1 patient classified as sturdy, Estratos diagnosis depends on both the FRIED and VIG Test and the VIG diagnosis has a good distribution between the frail, prefrail and sturdy patients.

Also, to simplify the number of classifications, a one vs all focus was applied. Meaning that the classifications were changed from prefrail, frail and sturdy (3-class) to frail or not-frail (2-class).

2.7 Structured Data Base Creation

With the features and the targets selected we created a new Data Base that can be use as input for a predictor with the following size: 235 rows (samples) x 60 Columns (59 features and 1 target).

2.8 Predictor Structure

The predictor proposed was a SNN (Shallow Neural Network) available in Matlab as patterned a NN (Neural Network) with 1 hidden layer. The final amount of hidden size or internal neurons of the hidden layer were determined during the training.

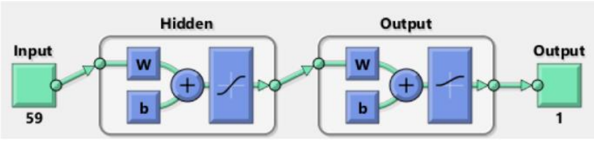


Fig 6. SNN without Training

2.9 Predictor Training

To create the final predictor (Fig. 7), the following steps were followed: a) Separate the Structured DDBB in 3 subsets called train, check and test b) Perform the basic training of an SNN c) Train 10000 SNN for the same hidden size d) train 120 different hidden size. And then the best SNN is picked from 1.200.000 trained SNN. This Final SNN was called FragilNET.

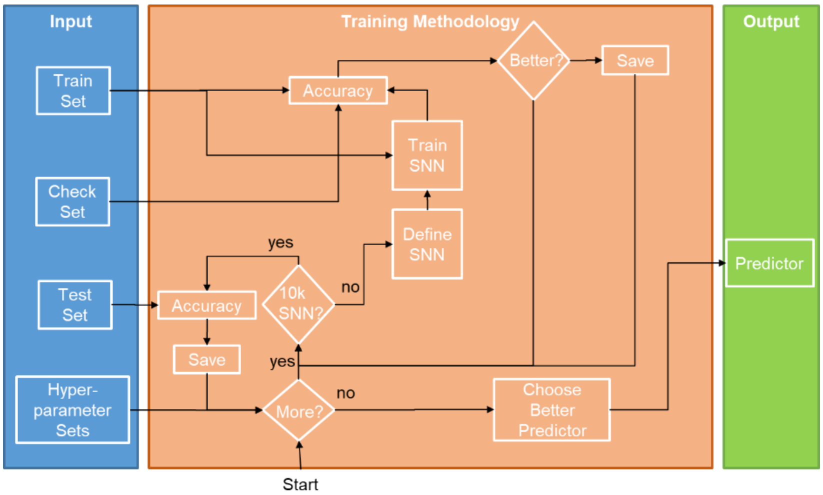
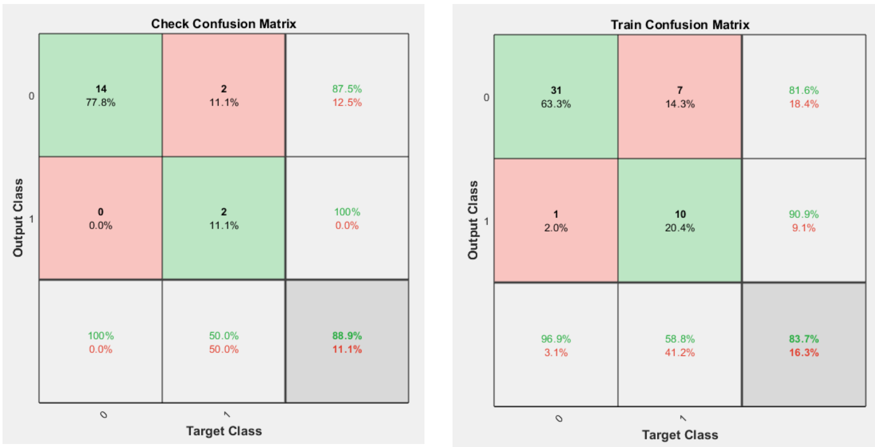


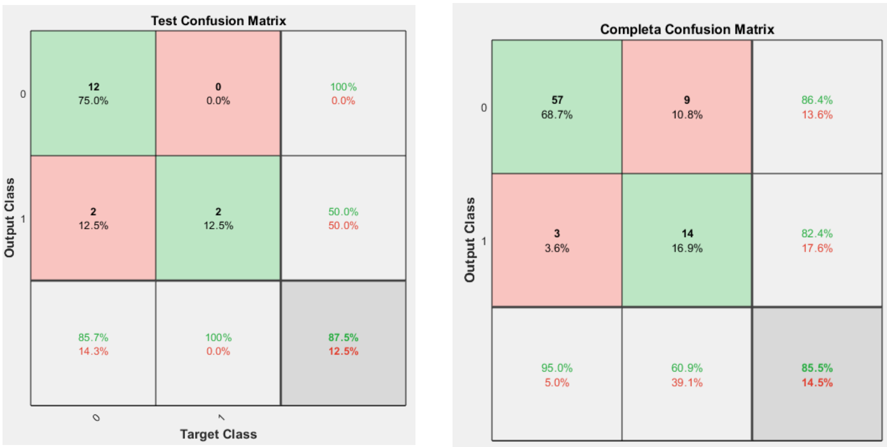
Fig 7. Training Methodology

3 Tests and Results

The confusion matrix of each set of the FragilNET are presented in figure 8.



Train Check



Test Complete

Fig 8. Confusion Matrix associated with each data set.

A summary of the data comparations is found in table 1.

Table 1. Confusion Metrics comparation.

	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
Train	83.7	90.9	58.8	95.9
Check	88.9	100	50	100
Test	87.5	50	100	85.7
Complete	85.5	82.4	60.9	95
Mean±std	86.4±2.3	80.8±21.8	67.4±22.2	94.2±6.0

4 Conclusions

The protocol design proposed was successfully implemented during the different tests conducted in the population, and the modified Deyard dynamometer was calibrated effectively and yielded satisfactory results for these tests and the resulting force overtime signals were similar to what was expected after consulting past studies that used them as well.

The Data of 83 patients (235 samples) was used to build a SNN called FragilNET that can predict the frailty label with 85,5% of accuracy and a sensibility of 67% using the signal information of the hand strength and the physiological data of the patient.

This research has many layers of development and finally we got the predictor that can relate the signal of hand force to the frailty level.

Even with the good news, we found it convenient to increase SNN training with more iterations, as it is slightly lower than check performance meaning that we hadn't reach the best training performance yet.

All decision to pick the best SNN were automatized. Data preprocessing Methods and how to separate it could work on other signals with similar step-like behaviors.

In order to achieve a support tool to correctly classify with a simple test the fragile condition it is crucial to have a high level of precision in our predictions, that means the correct classification of True Positives (Fragile) among all the subject. A value around 86% corroborates that assertation. On the other hand, and in the same level of importance, it is the fact that we need to minimize the number of False Negatives i.e., being fragile and predict robust. In this case the Sensitivity index that measures this concept has still a low value, around 67%. This is the weakness part of the whole prediction system and further work is needed to improve this sensitivity.

4.1 Recommendations for future work

Apply a data augmentation algorithm or get more data to balance the labels. We need more data from frail patients or perform data augmentation strategies like SMOTE to achieve a better balance between frail and robust patients.

Follow the classification strategy of the VIG, FRIED and Estratos index, built a three-class predictor including the "prefrailty level".

Combine all 3 predictors to enhance the final frailty level predictor with a 3-level definition.

We can work also in the redesign of the data base, for example doing all de previous steps to a data set with one observation per patient (only last sample, the average of the samples or the weighted average of the samples).

Acknowledgements This work was partially supported by the Spanish Ministry of Ciencia, Innovación y Universidades under project RTI2018-096701-B-C22, and by the Catalonia FEDER program, resolution GAH/815/2018 under the project, PECT Garraf : Envelliment actiu i saludable i dependència.

5 References

1. Campbell, A. J.; Buchner, D. M. Unstable disability and the fluctuations of frailty. *Age Ageing* **1997**, *26*, 315–318.
2. Morley, J. E.; Vellas, B.; Abellan van Kan, G.; Anker, S. D.; Bauer, J. M.; Bernabei, R.; Cesari, M.; Chumlea, W. C.; Doehner, W.; Evans, J.; Fried, L. P.; Guralnik, J. M.; Katz, P. R.; Malmstrom, T. K.; McCarter, R. J.; Gutierrez Robledo, L. M.; Rockwood, K.; von Haehling, S.; Vandewoude, M. F.; Walston, J. Frailty Consensus: A Call to Action. *J Am Med Dir Assoc* **2013**, *14*, 392–397
3. Collard, R. M.; Boter, H.; Schoevers, R. A.; Oude Voshaar, R. C. Prevalence of frailty in community-dwelling older persons: A systematic review. *J Am Geriatr Soc* **2012**, *60*, 1487–1492.

4. Clegg, A.; Young, J.; Iliffe, S.; Olde Rikkert, M. G. M.; Rockwood, K. Frailty in older people. *Lancet* **2013**, *381*, 752–762.
5. Rockwood, K.; Song, X.; MacKnight, C.; Bergman, H.; Hogan, D. B.; McDowell, I.; Mitniski, A. A global clinical measure of fitness and frailty in elderly people. *Can Med Assoc J* **2005**, *173*, 489–495.
6. Giampaoli, S., Ferrucci, L., Cecchi, F. et al.: Hand-grip strength predicts incident disability in non-disabled older men. Age Ageing 28, 283–288 (1999).
7. Sousa-Santos, A.R., Amaral, T.F.: Differences in handgrip strength protocols to identify sarcopenia and frailty - A systematic review. BMC Geriatr 17 (2017).
8. Richards, L.G., Olson, B., Palmiter-Thomas, P.: How Forearm Position Affects Grip Strength. Am J Occup Ther 50, 133–138 (1996).
9. Alahmari, K.A., Silvian, S.P., Reddy, R.S. et al.: Hand grip strength determination for healthy males in Saudi Arabia: A study of the relationship with age, body mass index, hand length and forearm circumference using a hand-held dynamometer. J Int Med Res 45, 540–548 (2017).
10. Desrosiers, J., Bravo, G., Hébert, R., Dutil, E. Normative data for grip strength of elderly men and women. Am J Occup Ther Off Publ Am Occup Ther Assoc 49, 637–644 (1995).
11. Innes, E.: Handgrip strength testing: A review of the literature. In Australian Occupational Therapy Journal 46(3), 120–140 (1999).
12. Kamimura, T., Ikuta, Y.: Evaluation of grip strength with a sustained maximal isometric contraction for 6 and 10 seconds. J Rehabil Med 33 (2001).
13. Sindhu, B.S., Shechtman, O. Using the force-time curve to determine sincerity of effort in people with upper extremity injuries. J Hand Ther 24, 22–30 (2011).
14. Bhuanantanondh, P., Nanta, P., Mekhora, K.: Determining Sincerity of Effort Based on Grip Strength Test in Three Wrist Positions. Saf Health Work 9, 59–62 (2018).
15. De Dobbeleer, L., Beyer, I., Njemini, R. et al.: Force-time characteristics during sustained maximal handgrip effort according to age and clinical condition. Exp Gerontol 98, 192–198 (2017).
16. Rodríguez-Martín, D., Pérez-López, C., Samà, A. et al.: A waist-worn inertial measurement unit for long-term monitoring of Parkinson's disease patients. Sensors 17(4), (2017).
17. Fried, L.P., Tangen, C.M., Walston, J. et al.: Frailty in older adults: Evidence for a phenotype. Journals Gerontol - Ser A Biol Sci Med Sci 56(3), (2001).
18. Redin, J.M.: Valoración geriátrica integral (I). Evaluación del paciente geriátrico y concepto de fragilidad. ANALES del Sistema Sanitario de Navarra 22 (1), 41-50 (1999).
19. Mahoney, F.I., Barthel, D.W.: Functional evaluation: the barthel index. Md State Med J 14, 61-65 (1965).
20. Carla, G.: The Lawton Instrumental Activities of Daily Living Scale. The American Journal of Nursing 17(5), 343-4 (2008)