

Analysis of the Correspondence of the Degree of Fragility with the Way to Exercise the Force of the Hand

E.P. Guindal¹, X. Parra¹, M. Musté¹, C. Pérez², O. Macho², A. Català¹

1. Technical Research Centre for Dependency Care and Autonomous Living, Universitat Politècnica de Catalunya, 08800 Vilanova i la Geltrú, Catalonia, Spain; 2. Department of Investigation, Consorci Sanitari Alt Penedès-Garraf, 08800 Vilanova i la Geltrú, Catalonia, Spain

Corresponding Author: Elsa Pérez Guindal, Technical Research Centre for Dependency Care and Autonomous Living, Universitat Politècnica de Catalunya, 08800 Vilanova i la Geltrú, Catalonia, Spain, Email: elsa.perez@upc.edu, <https://orcid.org/0000-0001-7221-054X>

Abstract

BACKGROUND: Frailty is a geriatric syndrome characterized by increased individual vulnerability with an increase in both dependence and mortality when exposed to external stressors. The use of Frailty Indices in routine clinical practice is limited by several factors, such as the cognitive status of the patient, times of consultation, or lack of prior information from the patient.

OBJECTIVES: In this study, we propose the generation of an objective measure of frailty, based on the signal from hand grip strength (HGS).

DESIGN AND MEASUREMENTS: This signal was recorded with a modified Deyard dynamometer and processed using machine learning strategies based on supervised learning methods to train classifiers. A database was generated from a cohort of 138 older adults in a transverse pilot study that combined classical geriatric questionnaires with physiological data.

PARTICIPANTS: Participants were patients selected by geriatricians of medical services provided by collaborating entities.

SETTING AND RESULTS: To process the generated information 20 selected significant features of the HGS dataset were filtered, cleaned, and extracted. A technique based on a combination of the Synthetic Minority Oversampling Technique (SMOTE) to generate new samples from the smallest group and ENN (technique based on K-nearest neighbors) to remove noisy samples provided the best results as a well-balanced distribution of data.

CONCLUSION: A Random Forest Classifier was trained to predict the frailty label with 92.9% of accuracy, achieving sensitivities higher than 90%.

Key words: Comprehensive Geriatric Assessment, frailty identification, hand grip strength, random forests, supervised learning.

Introduction

Frailty is a geriatric syndrome characterized by decreased ability of the body to respond to external stressors. The consequences include, among others, an individual's functional decline, increased risk of falls and dependency or ultimately institutionalization, and even death. Frailty is highly prevalent in older adults, occurring in up to 30% of people over 75 years of age, and its appearance can be anticipated, delayed, or avoided (1). Rehabilitation was more successful when it was applied early. This means that the development of a method capable of objectively identifying frailty for use in clinical practice, would greatly increase the diagnostic capacity and scope, and directly increase the success

of these interventions. The research hypothesis was that the degree of frailty of a patient could be automatically predicted from the signal obtained during a strength test performed using an adapted hand dynamometer, namely, provide support to frailty evaluation through an objective measure processed by Artificial Intelligence.

Frailty measurements

The most important frailty measurement scales are the following. Fried et al. (2) which defined the frailty phenotype by the presence of three or more of the following symptoms: unintended weight loss, weak-ness, low resistance, slow movement, and low activity. The Frail-VIG (3) based on the Comprehensive Geriatric Assessment (CGA) is a global diagnostic tool and methodology designed to identify and quantify biomedical and pharmacological data, functional, psychological, and social problems that older adults may present with. The Barthel scale (4), allows the determination of the degree of dependency of a person or the need for help to carry out ten basic daily actions: eating, washing, going to the bathroom, moving, dressing, etc. The Lawton-Brody scale (5) measures the ability to perform activities of daily living. Each area was scored according to the description that best corresponded to the participant.

Fried and other indices or scales for frailty measurements exclusively use the maximum value of the grip strength of the hand as one of the symptoms, but in clinical practice, the geriatrician uses their experience to make quick diagnoses based on the way or the "form" the patient takes on to perform a handshake, that is, how long it lasts in time and how intense it is. We intend to demonstrate the correlation between Hand Grip Strength form and the frailty level of patients. If we can objectively measure frailty, it would mean savings in medical consultation time and even remote evaluation of patients wearing sensors.

HGS related work

Hand grip strength (HGS) has been widely used by investigators and therapists to diagnose sarcopenia and frailty, as it is a reliable indicator of overall muscle strength that decreases with age. The results obtained from these tests were used to verify whether HGS can function as a predictor of

disability in older men (6). Grip strength appears to be an explanatory or predictive biomarker of specific outcomes such as generalized strength and function, bone mineral density, fractures, falls, comorbidity load, cognition, sleep, hospital-related variables, and mortality. In fact, the force-time curve was originally an assessment meant to identify maximal or submaximal effort and is a graphic representation of the force applied over a period of time during a single HGS trial. The function that was obtained had the form of a step.

There are different studies aimed to investigate the force-time characteristics. One of them focuses on the sustained maximal grip effort according to age and clinical condition (7), using a modified Martin vigorimeter. Another shows a comparative study of the protocols for diagnosing sarcopenia and frailty according to HGS (8), which is an important focus of the present research, as the protocol proposed here will be contrasted with those proposed in previous studies for the purposes of comparison and innovation.

In the preliminary research presented at IWANN 2021 Congress (9), the starting hypothesis for the present work was already stated as follows: “The analysis of the HGS signal, measured by a hand dynamometer, can objectively diagnose frailty in geriatric patients”. In this previous work, we evaluated the behavior of the HGS signal for short periods of time, obtaining the force-time curve, thus expanding the information provided by the maximum value. Signals were recorded with a modified Deyard dynamometer and processed using machine learning strategies to identify brittleness levels.

In the present study, we extended the number of patients from 83 to 138 older adults and use a different machine learning approach to increase the sensitivity and specificity of the previous study. The research hypothesis was that the degree of frailty of a patient could be automatically predicted from the signal obtained during a strength test performed using an adapted hand dynamometer.

Materials and methods

The instrument

The instrument used to measure the HGS in the present study was a modified constant electronic hand dynamometer 14192-709E-EH101–90 kg capacity range, in which the entire electric circuit was modified and replaced by one specific, designed, and constructed at the Technical Research Centre for Dependency Care and Autonomous Living (CETpD) of the Universitat Politècnica de Catalunya (UPC). Because of this modification, the dynamometer continuously measured the HGS over time, and includes the ability to store information, Bluetooth connectivity, and a micro-SD card for data storage using the Inertial Measurement Unit (IMU) developed by CETpD–UPC (10) for long-term monitoring of human pathological movement. The calibration of the system is fully described in at previous study (9).

Pilot protocol and data acquisition

The proposed pilot protocol was approved by the Ethics Committee for Drug Research of the Community of Madrid (Comité de Ética de la Investigación con Medicamentos de la Comunidad de Madrid – Ref 47/916546.9/19).

Participants, who signed informed consents, were selected via convenience sampling (non-probabilistic) from patients who received medical services provided by collaborating entities (“Hospital Central de la Cruz Roja San José” and “Santa Adela de Madrid”, “and “Casa d’Ampara de Vilanova i la Geltrú”), from November 2019 to July 2021. Only researchers from the Health Consortium “Alt Penedès-Garraf” had access to data base during and after data collection.

In the first part of the data acquisition, geriatricians obtained physiological information from each patient. The patient’s labelling between robust and frail of the database was made by the medical staff participating in the clinical trial based on all the information proportioned by the different questionnaires and scales (Fried, Barthel, Lawton-Brody, CGA Index). The second part of the protocol was the HGS test, which consisted of three trials carried out in a sitting position on a chair with the forearm placed on top of the leg in a neutral position (holding the dynamometer perpendicular to the leg), feet firmly on the floor at a shoulder-width distance, shoulders adducted, body neutrally rotated, and dominant hand used (Figure 1). The tests had a duration of six seconds each, with a rest interval of one minute between the tests.

Figure 1. Protocol position example



Inclusion criteria

The sample was stratified by the degree of frailty (robust and frail participants). Participants meeting the following criteria were included in the study: age above 70 years, sufficient reading and writing ability to answer questionnaires, and willingness to participate in the study, and acceptance of the standards of performance and procedures established by researchers. Participants meeting the following criteria were

Table 1. Age, force peak, and geriatric criteria values for ROBUST and FRAIL groups

		Age	Max. Straight	Fried	Barthel	Lawton Brody	Fragil VIG	Women	Men	Total
ROBUST	Max	99,00	37,73	0,00	100,00	8,00	0,16	27	16	43
	Min	71,00	3,02	0,00	75,00	1,00	0,00			
	Mean	83,26	19,34	0,00	96,74	6,26	0,07			
	SD	6,91	8,03	0,57	4,80	2,13	0,05			
FRAIL	Max	103,00	26,93	5,00	100,00	8,00	0,84	54	41	95
	Min	76,00	2,09	1,00	10,00	0,00	0,00			
	Mean	90,44	12,97	3,29	63,37	2,05	0,44			
	SD	5,93	5,53	0,94	22,13	1,84	0,17			
PARTICIPANTS	Max	103,00	37,73	5,00	100,00	8,00	0,84	81	57	138
	Min	71,00	2,09	0,00	10,00	0,00	0,00			
	Media	88,36	14,60	2,78	74,06	3,30	0,32			
	SD	6,96	6,83	1,14	24,44	2,68	0,22			

Table 2. Sociodemographic data for Frail and Robust groups

	MARITAL STATUS				COHABITATION						Total
	Single	Married	Separated	Widower	Living alone	Couple	Family	Caregiver	Community	Nursing Home	
FRAIL	6	28	1	60	15	21	22	11	0	26	95
% Total frail	6,3	29,5	1,1	63,2	15,8	22,1	23,2	11,6	0,0	27,4	
ROBUST	9	21	1	12	7	13	4	0	3	16	43
% Total robust	20,9	48,8	2,3	27,9	16,3	30,2	9,3	0,0	7,0	37,2	
Participants	15	49	2	72	22	34	26	11	3	42	138
% Total Participants	10,9	35,5	1,4	52,2	15,9	24,6	18,8	8,0	2,2	30,4	

excluded: alcohol and/or drug abuse, any type of neurological or osteoarticular affectation (Parkinson's disease, osteoarthritis, or stroke), inability to provide informed consent and to cooperate with study procedures.

The patient's labelling between robust and frail of the database was made by the medical staff participating in the clinical trial based on all the information proportioned by the different questionnaires and scales (2-5). Using the original database, the CRF provides different possible labels. If we choose the Fried Criteria (involuntary loss of weight, low energy or exhaustion, slow mobility, muscle weakness, low physical activity), then the classification is binary: frail and not-frail. Using the Fragile CGA criteria, we can distribute the patients between robust (CGA index < 0.2), prefrail (between 0.2 and 0.5), and frail (CGA index > 0.5). As a first step to achieve some practical results for classification purposes, we decided to combine frail and prefrail into a single class, and to have only two classes for the final classification action. The aim of this simplification was to distinguish between the robustness of frailty tendencies.

The sociodemographic characteristics of the participants, as well as their levels of fragility, are shown in Tables 1 and 2. The present cohort shows an unbalanced data set, so there are more than double frail participants than robust. Balancing techniques are required for obtaining coherent results. An overview of the data shows that there are no direct relationships between the different variables and guides us to search for more complex information, such as temporal variation in exerting the force of the hand (HGS).

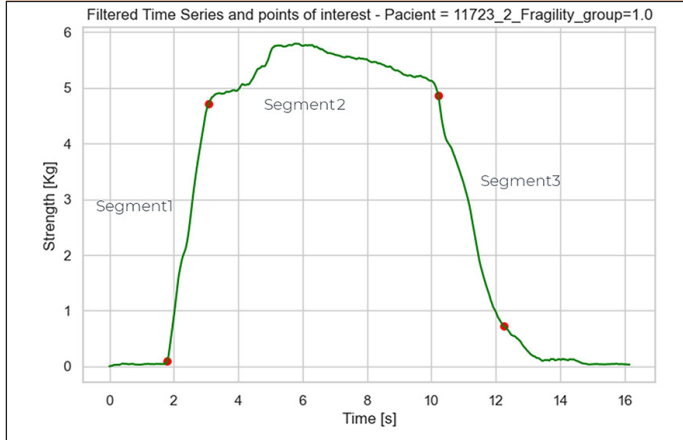
Data base creation

First, geriatricians performed obtained physiological information from each patient and evaluated the following scales in clinical trials: Fried criteria (2), CGA Index (3), Barthel scale (4), and Lawton–Brody scale (5). The second part of the protocol was the HGS test, which consisted of three trials carried out in a sitting position on a chair with the forearm placed on top of the leg in a neutral position. The tests had a duration of six seconds each, with a rest interval of one minute between the tests.

In total, 223 valid HGS temporal series or force-time curves, were recorded for 138 patients. The validity of the HGS signal is based on simple geometric criteria of the recorded shape. Each patient was assigned a set of 19 geriatrician variables obtained from different scales and included in the Case Report Form (CRF).

Two Butterworth filters were used in this study. To obtain the geometric features, a 1 Hz low-pass filter was applied to determine the HGS shape. Three phases can be distinguished in the HGS time series after the first filtering process: the force generation phase (segment 1 in Figure 2), force maintenance phase (segment 2 in Figure 2), and force decay phase (segment 3 in the same Figure). A second 5 Hz filter was applied to obtain the dynamic and frequency features.

Figure 2. Strength (Kg) vs. time (s) signal and the four characteristic points related to the step shape of the time-series signal



A simple algorithm based on the first and second derivatives provided the HGS time series. These phases were used to extract specific characteristics that could be used to determine whether a patient was prone to developing frailty.

Results

The Cohort

The Tables 1 and 2 describe the main characteristics of the cohort studied for each group of interest, “frail” and “robust”, and also for the global number of participants. Table 1 describes the maximum, minimum, average, and standard deviation values of age, strength peak, and Fried, Barthel, Lawton-Brody, and Fragile-VIG criteria. Table 2 describes the two significant sociodemographic variables, Marital Status and Cohabitation, with corresponding percentages for each group.

Feature selection

As a first part of the individual HGS signal processing we extract 177 statistical features such as: areas, the number of peak values inside a signal, the number of cycles calculated

using the rain flow counting method, etc. In order to simplify the classification process, we use the Boruta methodology (11) for feature selection. The results are the 20 most significant features as shown in Table 3.

Table 3. Set of the 20 selected significant features

1 Area (segment 3)	2 Density (segment 2)
3 Maximum (segment 3)	4 Quartile 1 (segment 2)
5 Quartile 2 (segment 2)	6 Peak amplitude (0.004)
7 Peak amplitude (0.01)	8 Number cycles RF (0.0005)
9 Amplitude cycles RF (1.0)	10 Peak amplitude normalized (0.005)
11 Peak amplitude normalized (0.04)	12 Percentile 20 (segment 2)
13 Percentile 25 (segment 2)	14 Percentile 30 (segment 2)
15 Percentile 35 (segment 2)	16 Percentile 45 (segment 2)
17 Percentile 50 (segment 2)	18 Peak amplitude (0.001)
19 Peak amplitude (0.003)	20 Peak Amplitude normalized (0.05)

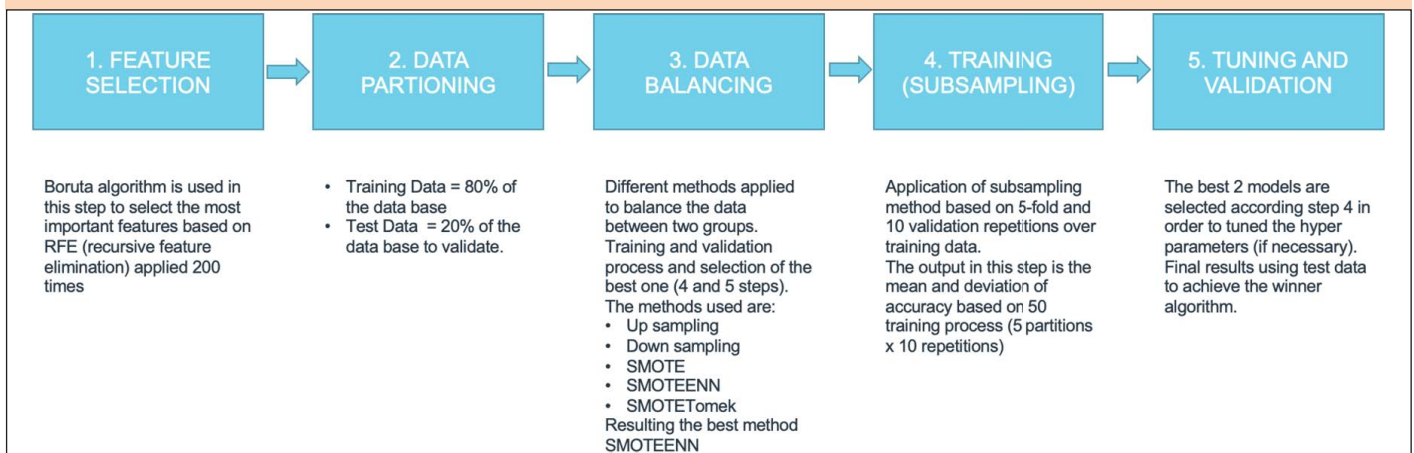
Data balancing techniques

To avoid a model that was biased toward majority class prediction, owing to the larger number of frail participants than robust ones, different balancing techniques were applied: down-sampling, up-sampling, and hybrid methodologies. Finally, a technique based on a combination of the Synthetic Minority Oversampling Technique (SMOTE) to generate new samples from the smallest group and ENN, technique based on K-nearest neighbours (12) to remove noisy samples provided the best results as a well-balanced distribution of data.

Data training and validation

Finally, the database for training consists of the 20 characteristics for each HGS signal as an input vector, and the robust or fragile label coming from the geriatric indices as an output vector. The database was randomly divided into training and validation sets (80% and 20%, respectively). A k-fold validation with k=10 training strategy was performed (Figure 3). The following algorithms were used, considering standardized and non-standardized data: logistic regression,

Figure 3. Data training preparation (k=10), partitioning, subsampling (balancing with down-sampling, up-sampling and SMOTE-hybrid) and resampling steps



linear discriminant analysis, K-nearest neighbor classifier, decision tree classifier, Gaussian naïve Bayes, and support vector machines (Figure 4). The following ensemble algorithms were used for the standardized data: AdaBoost classifier, gradient boosting classifier, random forest classifier (10 estimators), and extra tree classifier (10 estimators) (Figure 5).

Figure 4. Basic algorithms used as a first round of training and classification: Logistic Regression (LR), Linear Discriminant Analysis (LDA), K-Nearest Neighbors Classifier (KNN), Classification and Regression Tree (CART), Gaussian Naive-Bayes (NB), and Support Vector Machines (SVM)

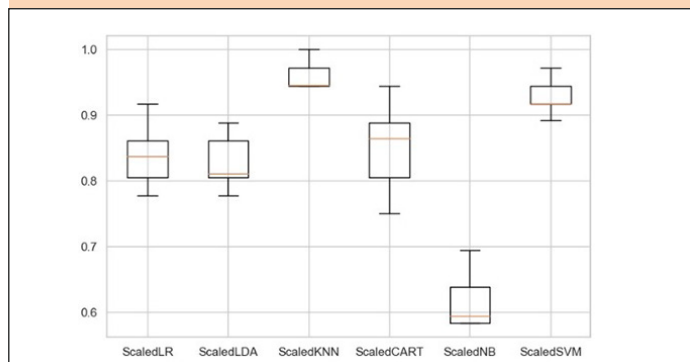
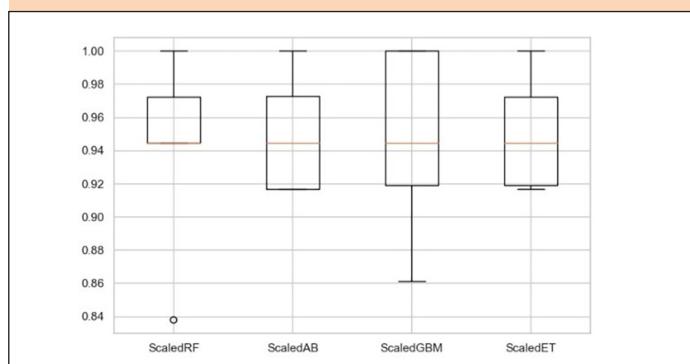


Figure 5. Advanced algorithms achieve the best results in the classification: Random Forest Classifier (RF), Ada Boost Classifier (AB), Gradient Boosting Classifier (GBM), and Extra Trees Classifier (ET)



The balancing techniques that reported the best results were as follows: (1) over-sampling using SMOTE and cleaning using edited nearest neighbors, and (2) a combination of over- and under-sampling using SMOTE and edited nearest neighbors. The scaled random forest classifier achieved the best results, and there was a positive increase in the algorithm optimization.

The confusion matrices obtained for the best algorithm are shown in Figure 6. We can observe the high accuracy of the presented methodology, in the validation phase (data never seen before by the algorithm) 39 patients out of 42 were correctly classified as a fragile or robust.

The percentages shown in the colored cells represent the ratio of cases in each cell compared to the total cases (181 for training and 42 for validation). Precision values are derived by dividing the cases in each row (the color of the percentage denotes the corresponding row) by the total cases in the column. This calculation involves $TP/(TP+FP)$ and $FP/(TP+FP)$. The

recall column (equivalent to sensitivity) mirrors precision but focuses on rows, expressed as $TP/(TP+FN)$ and $TN/(TN+FP)$. The F1-score column is a measure that combines precision and sensitivity, and it is calculated as twice the product of these measures divided by their sum. It reflects the balance between these metrics. The Support column aggregates the number of cases for each class per row in the dataset, providing insight into the balance of pattern distribution.

Figure 6. Confusion matrices for the best solution: Scaled Random Forest

		1.0	2.0	Recall	F1-score	Support
Output class	1.0	106 58,6%	0 0%	100% 0%	100%	106 58,6%
	2.0	0 0%	75 41,4%	100% 0%		100%
Precision		100% 0%	0% 100%	100% 0%		181
		Target class				
		1.0	2.0	Recall	F1-score	Support
Output class	1.0	9 21,4%	1 2,4%	90% 10%	85,7%	10 23,8%
	2.0	2 4,8%	30 71,4%	93,8% 6,3%	95,2%	32 76,2%
Precision		81,8% 18,2%	3,2% 96,8%	92,9% 7,1%		42
		Target class				

Discussion

The protocol designed and proposed in this study was successfully implemented via different tests conducted in the study population, in which the modified Deyard dynamometer was calibrated effectively and yielded satisfactory results. The resulting force–overtime signals were similar to those expected after consulting previous studies that used them (1, 6, 8).

A dataset of 138 patients with valid registration of 223 samples of force-time curves was used to train the different classifiers. The selected Random Forest Classifier was able to predict the frailty label with 92.9% accuracy, achieving sensitivities for robust participants (class 1.0) and fragile participants (class 2.0) of 90% and 93.8%, respectively, when the signal information of hand strength and physiological data of the patient were used.

This study included many layers of development, and we were able to obtain a predictor that could relate the hand force signal to the frailty level. All decisions regarding selection of the best classification algorithm were automated. Data pre-processing methods and their separation can be applied to other signals with similar step-like behaviors.

To achieve a support tool for correctly classifying fragile conditions using a simple test, it is crucial for predictions to have a high level of precision, so that True Positives (fragile) among all subjects can be classified correctly. A value of approximately 93% corroborates this assertion. However, of the same level of importance is the fact that we need to minimize the number of False Negatives, that is, being fragile but predicted to be robust. The sensitivity index, which measures this concept, achieved a high score of approximately 93%. When the combination of precision and sensitivity (recall) was considered (F1-score), a high level of performance was observed for both classes (85.7% and 93.8% for robust and fragile participants, respectively).

Future directions

The database presented here is sufficient for corroborating the possibility of extracting significant knowledge from simple measurements. However, a larger number of patients is necessary to develop a possible instrument for daily practice in medical settings.

An interesting study using the general data set obtained involved a search for correlations between the values of the different criteria and sociodemographic data. Augmenting the granularity of the studied groups, including different groups of women and men, could also be relevant.

Once the best predictive models are obtained, they should be complemented by medical criteria. The present methodology could be a first step for the development of a valuable support tool for general practitioners coping with possible frailty patients.

Acknowledgments: This work was possible because of the collaboration in the selection of participants and database construction by Francisco S. del Corral Guijarro, Aris Somoano, Cristina Gianella, and Luis Ramírez. Thanks also go to the support of the UPC members, Ingrid Magnusson, Francisco J. Ruíz, Pau Martí, Cristóbal Raya, and Pere Ponsa.

Ethics approval: The study was conducted according to the guidelines of the Declaration of Helsinki, and the protocol was approved by the Institutional Review Board (or Ethics Committee) of the “Comité de Ética de la Investigación con Medicamentos de la Comunidad de Madrid” (Ref 47/916546.9/19).

Consent: The study was carried out in accordance with relevant legislation and ethics guidelines. Informed consent was obtained from all subjects involved in the study.

Conflict of Interests: The authors declare that they have no competing interests.

Data availability: due to ethical restrictions imposed by Spanish Agency for Drugs and Medical Devices (Spain), Consorsis Sanitari Alt Penedès-Garraf and el Hospital la Princesa (Madrid’s Community), related to approved consent procedure and protecting privacy (related to a clinical study), all relevant data are available upon request Elsa Pérez, email: elsa.perez@upc.edu

Funding: 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”. Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Open Access: This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

References

1. Alahmari KA, Silvian SP, Reddy RS, Kakarparthi VN, Ahmad I, Alam MM. 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* 2027;45:540–548. <https://doi.org/10.1177/0300060516688976>
2. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: Evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001;56(3):146–56. <https://doi.org/10.1093/gerona/56.3.m146>.
3. Frail-VIG index: Design and evaluation of a new frailty index based on the Comprehensive Geriatric Assessment. *Revista Española de Geriatria y Gerontología*. Elsevier. 2017 DOI:10.1016/j.regg.2016.09.003
4. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. *Md State Med J* 1965;14:61–65
5. Carla G. The Lawton Instrumental Activities of Daily Living Scale. *Am J Nurs* 2008;17(5):343–4. <https://doi.org/10.1097/01.NAJ.0000314810.46029.74>
6. Giampaoli S, Ferrucci L, Cecchi F, et al. Hand-grip strength predicts incident disability in non-disabled older men. *Age Ageing* 1999;28:283–288. <https://doi.org/10.1093/ageing/28.3.283>
7. 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* 2017;98:192–198. <https://doi.org/10.1016/j.exger.2017.08.033>
8. Sousa-Santos AR, Amaral TF. Differences in handgrip strength protocols to identify sarcopenia and frailty - A systematic review. *BMC Geriatr* 2017;17(1):238. <https://doi.org/10.1186/s12877-017-0625-y>
9. Pérez Guindal E, Torres Rangel JE, Musté Rodríguez M, et al. Frailty level prediction in older age using hand grip strength functions over time. In *Advances in Computational Intelligence, Proceedings of the 16th International Work-Conference on Artificial Neural Networks*. Springer, Cham (Switzerland). 2021. https://doi.org/https://doi.org/10.1007/978-3-030-85099-9_29
10. 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* 2017;17(4):827. <https://doi.org/10.3390/s17040827>
11. Mazzanti S. Boruta Explained Exactly How You Wished Someone Explained to You. IOP Publishing. 2020. <https://towardsdatascience.com/boruta-explained-the-way-i-wish-someone-explained-it-to-me-4489d70e154a>. Accessed 17 March 2020
12. Gustavo EAPA Batista, Ana LC Bazzan, Maria Carolina Monard. Balancing training data for automated annotation of keywords: a case study. 2003. In *WOB*, 10–18.

The Author(s) 2024

How to cite this article: E.P. Guindal, X. Parra, M. Musté, et al. Analysis of the Correspondence of the Degree of Fragility with the Way to Exercise the Force of the Hand. *J Frailty Aging* 2024; <http://dx.doi.org/10.14283/jfa.2024.46>