medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Development of quantitative frailty and mortality prediction models on older patients as a palliative care needs assessment tool

Authors: Vicent Blanes-Selva[1], Ascensión Doñate-Martínez[2], Gordon Linklater[3], Juan M. García-Gómez[1]

1: Biomedical Data Science Lab, Instituto Universitarios de Tecnologías de la Información y Comunicaciones (ITACA), Universitat Politècnica de València, Valencia, Spain.

2: Polibienestar Research Institute, University of Valencia, Spain.

3: Highland Hospice and NHS Highland, Inverness, UK

Keywords: Palliative Care, Machine Learning, Frailty, Mortality, Admission, Prediction

Abstract

Background: Palliative care (PC) has demonstrated benefits for life-limiting illnesses. Cancer patients have mainly accessed these services, but there is growing consensus about the importance of promoting access for patients with non-malignant disease. Bad survival prognosis and patient's frailty are usual dimensions to decide PC inclusion.

Objectives: The main aim of this work is to design and evaluate three quantitative models based on machine learning approaches to predict frailty and mortality on older patients in the context of supporting PC decision making: one-year mortality, survival regression and one-year frailty classification.

Methods: The dataset used in this study is composed of 39,310 hospital admissions for 19,753 older patients (age >= 65) from January 1st, 2011 to December 30th, 2018. All prediction models were based on Gradient Boosting Machines. From the initial pool of variables at hospital admission, 20 were selected by a recursive feature elimination algorithm based on the random forest's GINI importance criterion. Besides, we run an independent grid search to find the best hyperparameters in each model. The evaluation was performed by 10-fold cross-validation and area under the receiver operating characteristic curve and mean absolute error were reported. The Cox proportional-hazards model was used to compare our proposed approach with classical survival methods.

Results: The one-year mortality model achieved an AUC ROC of 0.87 ± 0.01; the mortality regression model achieved an MAE of 329.97 ± 5.24 days. The one-year frailty classification reported an AUC ROC of 0.9 ± 0.01. The Spearman's correlation between the admission frailty index and the survival time was –0.1, while the point-biserial correlation between one-year frailty index and survival time was –0.16.

Conclusions: One-year mortality model performance is at a state-of-the-art level. Frailty Index used in this study behaves coherently with other works in the literature. One-year frailty classifier demonstrated that frailty status within the year could be predicted accurately. To our knowledge, this is the first study predicting one-year frailty status based on a frailty index. We found mortality and frailty as two weakly correlated and complementary PC needs assessment criteria. Predictive models are available online at http://demoiapc.upv.es.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

1. Background and Significance

Palliative Care (PC) is a holistic approach that improves the quality of life of patients with life-limiting disease. It is recommended that it be incorporated early in the disease trajectory, even in conjunction with potentially curative treatments [1]. PC can improve quality of life [2], mood [3], symptom control [4], reduce emergency department visits and hospitalisation [5], and even increase the one-year survival [6]

PC services have traditionally been mainly accessed by cancer patients but there is growing consensus about the importance of promoting access for patients with non-malignant disease at earlier stages [7, 8, 9]. In the same line, Raudonis et *al*. [10] suggest in their study that frail older adults could benefit from involvement in PC programs. Koller et *al*. 11] show how frailty is associated with poor health outcomes and death. They emphasize the importance of quantifying frailty so when the patient becomes frailer the focus of care can change to palliation.

It is estimated that at least 75% of patients would benefit from access to PC during their terminal illness [12]. Nevertheless, uncertainty about prognostication is cited as the most common barrier to PC referral, particularly for patients with non-malignant disease [13]

Different strategies have been used to try to aid prognostication. Clinical intuition was harnessed with the Surprise Question ("Would I be surprised if this patient died in the next year?") which has been promoted as a tool to prompt clinicians to recognise patients with a limited prognosis [14] However, in 2017 Downar et *al*. [15] published a systematic review of the surprise question, concluding that more accurate tools are required given its poor to modest performance as a mortality predictor.

Functional status was used in the Palliative Performance Scale [16]. Risk of death increased with lower performance levels and with falling performance levels, but survival data varied across different healthcare systems [17].

The Supportive and Palliative Care Indicators Tool (SPICT) uses clinical indicators of poor prognosis which were developed through a consensus of expert opinion [18]. It has been shown to have a predictive accuracy of up to 78% [19].

Other studies have used data analysis to propose alternative tools to predict short-term mortality. Bernabeu-Wittel in 2010 developed the PROFUND index [20], a predictive model for patients with multimorbidity with a reported area under the ROC curve (AUC ROC) of 0.7 (95% CI 0.67-0.74) in testing. Van Walraven et *al.* in 2015 proposed HOMR [21], a tool for predicting one-year mortality on adults (>=18 years and >= 20 years for the different cohorts) reporting 0.89 (95% CI 0.87 to 0.91) to 0.92 (95% CI 0.91 to 0.92). In 2018 Avati et *al.* [22] proposed a deep learning approach to identify patients with a survival between 3 and 12 months and reported an AUC ROC of 0.93 (0.87 for admitted patients), in 2019 Wegier et *al*. [23] proposed a version of HOMR but using only variables available at the admission achieving 0.89 of AUC ROC.

In addition to life-expectancy prognostication tools a wide array of frailty indexes (FI) have been proposed to assess the health status in older adults. These are usually based on deficits accumulation [24]. Frailty index has been used as a tool to predict mortality and poor health outcomes [25]. Some studies have tried to predict frailty status: Babič et al. in 2019 [26] use a clustering approach to identify clusters considering the prefrail, non-frail and frail status using 10 numerical variables for adults over 60 years old. Sternberg et al. [27] in 2012 tried to identify frail patients with their methods against the VES frailty score [28] for patients over 65 years old. Bertini et al. [29] in 2018 created two predictive models for patients over 65 years old: one to assess frailty

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

risk using the probability of hospitalization or death within the year and a second one to assess worsening risk to each subject in the lower risk class. However, to our knowledge, no study has tried to predict frailty status within a year without using proxies such as mortality. The authors hypothesize that frailty status can be predicted and used as part of PC inclusion criteria.

Based on these previous results, our overall aim is to develop machine learning tools capable of making predictions about mortality and frailty for older adults so that health professional can benefit from quantitative approaches based on data-driven evidence. First, we choose one year as a horizon to make the mortality prediction, as stated elsewhere [22], longer than 12 months is not desirable due to the difficulty in the predictions and the limited resources of the programs which are better to focus on immediate needs. To fulfil this need, we aim to create one-year mortality model that will work as a binary predictor.

In addition to the one-year mortality model, we propose the creation of a survival regression model, this model aims to obtain a prediction in days from admission to death. Despite being more difficult to predict, we think this information will help to contextualise the results of the one-year mortality model providing the health professional with additional information such as the magnitude of the remaining time until death (days, weeks or months).

Finally, we aim to create a frailty model to estimate the health status, assessed by the Frailty Index, of a patient in the future. Following the work of Searle et al. [30] we created a frailty index from our dataset and stratified it in 4 categories according to the work of Hoover et al. [31] and aggregate together the two less severe frailty conditions (Non-Frail + Vulnerable) and the two more frail status (Frail + Most Frail). To complement the information provided by the one-year mortality model, we set this frailty model to make predictions for 12 months after the admission. The authors think that the combination of both criterions mortality and frailty can have a positive impact on detecting PC needs.

2. Methods

2.1 Ethics

The data used in this study comes from the University and Polytechnic La Fe Hospital of Valencia and was retrospectively collected on the Electronic Health Records (EHR). This procedure was assessed and approved by the Ethical Committee of the University and Polytechnic La Fe Hospital of Valencia (registration number: 2019-88-1).

2.2 Materials

The data used in this study comes from the electronic health records (EHR) from the University and Polytechnic La Fe Hospital of Valencia. The dataset contained hospital admissions records for older patients (age >= 65) from January 1st, 2011 to December 31st, 2018.

The dataset is composed of variables from several sources: socio-demographic information, administrative details about the admissions, laboratory results, Barthel scale variables, cognitive function, and a set of ICD9 diagnosis. A total of 147 candidate variables were selected based on a literature review performed within the framework of the H2020 European Project entitled InAdvance (ref.: 825750) and clinician's expert opinion.

Mortality target variables were extracted from admission administrative data and recorded death date. Patients without a record of death were considered alive at the time of extraction and consequently were excluded from the mortality risk model development. Those patients alive at the

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

extraction date were not eligible for the survival regression problem, which is known as right-censored data (20959 available samples).

As for the frailty target, we have calculated the FI of every episode (admission frailty), sorted them chronologically from oldest to newest and grouped them by patient. The frailty target of one episode is the admission frailty of the next one if that episode has occurred in the time threshold specified (one year). Otherwise, their target frailty is the admission frailty. Most recent episodes and patients with only one episode have been removed because no posterior data was available, so they became censored data. Finally, we transformed the obtained FI into categories following the work from Hoover et al. [31] Variables composing this frailty index are listed in Table 1.

Table 1: List of variables included in the frailty index

Difficulties in dressing	Difficulties in urinating	Difficulties in bathing
Difficulties in grooming	Difficulties in eating	Difficulties in stooling
Difficulties in moving	Hypertension	Chronic renal failure
COPD	Heart Failure	Cancer
Stroke	Parkinson	Atrial fibrillation
Thyroid Disorders	Diabetes Mellitus	Psychiatric disease
Gastrointestinal or liver disease	Musculoskeletal Diseases	Depression
Dementia	Malnutrition	Constipation
Pressure Ulcers	Anaemia	Visual impairment
Hear impairment	Gastrointestinal problems	

2.3 Pre-processing

We mapped the two categorical variables: *Admission Diagnose Code and Real Service Code* to integers using a dictionary, no prior information was provided so in each run categories receive different codification depending on the shuffling of the data. The method used does not support missing values in the model's input, so an imputing was strategy. We calculated the median for each variable in the training set and use it to impute the missing values in both training and test sets.

2.4 Feature selection

The authors used a recursive feature elimination process as a filter method. This process starts with the whole set of features, trains a tree-based model, calculates the relevance of each variable using the GINI importance [32], which measures the average gain of purity in the tree splits. Finally, less relevant variables are eliminated. The process is repeated until the desired number of features is obtained. Alternatively, the same method can be performed with a regression using its coefficients as importance measure. In this case, we used the Random Forests as the model for both classification and regression problems, set to 20 the number of features and set to 2 the number of features to eliminate each iteration.

2.5 Gradient Boosting Machine and Grid Search

Gradient Boosting Machines (GBM) [33] are ensemble models based on regression trees. The training algorithm behind the GBM is an iterative optimization method consisting of adding a tree to the ensemble that best reduces a loss function in the whole ensemble at every step. GBMs have been used in different problems with notable performance [34, 35, 36].

Since GBM is available for both classification and regression, we made use of this model to solve all three proposed problems. We used a traditional Grid Search to estimate some of the most relevant hyper-params: number of trees, maximum depth of each tree, learning rate and loss function. We

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

used logistic regressions as baselines to compare the performance of the models in the classification task: one-year mortality and frailty classification.

2.6 Evaluation

For the evaluation of the models, we have used the classical K-Fold cross-validation technique. Where the dataset is split into K different sets and each iterations K-1 sets are used to train the model, and one is used to evaluate the model, aggregating the various results and metrics at the end. The experiments were set with K=10.

To evaluate the performance of the one-year mortality and the frailty binary classifier we selected following metrics: area under ROC curve (AUC ROC), accuracy, sensitivity (or True Positive Rate) and specificity (or True Negative Rate). For the survival regression model, we selected the mean squared error (MSE) and the mean absolute error (MAE). Results are rounded to two decimals as in other relevant works already introduced [20, 21, 22, 23]. In addition, since the GBM offers the GINI variable importance for the built model, we present the average and standard deviation for each variable in each iteration.

2.7 Survival models

In order to compare our mortality regression model with state of the art, we have performed survival analysis over the data processed with same pipelines described above. For that, we chose the Cox regression model [37], from which we obtained survival estimations for patients by calculating the survival expected time E[T].

2.8 Software

The whole experimentation described in this work has been carried using the python 3 programming language [38], and the following scientific libraries and packages: numpy [39], pandas [40], scikit-learn [41] and lifelines [42].

3. Results

3.1 Descriptive results

Data contains a total of 39,310 cases corresponding to 19,753 unique patients. Cohort was composed by 9780 males and 9973 females with a mean age of 80.75 years, complete description is available in Table 2.

Table 2: Patient demographic information

Sex	N-individuals	Mean Age (years)	STD Age (years)
Female	9,973	80.75	8.67
Male	9,780	77.44	8.24
All	19,753	79.11	8.62

The one-year mortality target variable distribution is: 35905 (55.9%) are negative cases (death > 365 days) and 28368 (44.1%) are positive (death <= 365 days) as shown in Figure 1.A. The survival regression target variable (mean 368.59; std 496.46) presents a right-skewed shape as can be observed in its density plot in Figure 1.B. The admission FI (mean 0.27; std 0.12) and the FI target variable (mean 0.32; std 0.14), both are shaped like a normal distribution (plot in Figure 1.C and Figure 1.D), while the distributions of the different categories are: Non-Frail 986 (2.2%), Vulnerable 10911 (24.34%), Frail 25638 (57.19%), Most Frail 7294 (16.27%). As aggrupation of two categories: Non-Frail + Vulnerable 11897 (26.54%), Frail + Most Frail 32932 (73.46%), data represented in Figure 1.E.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

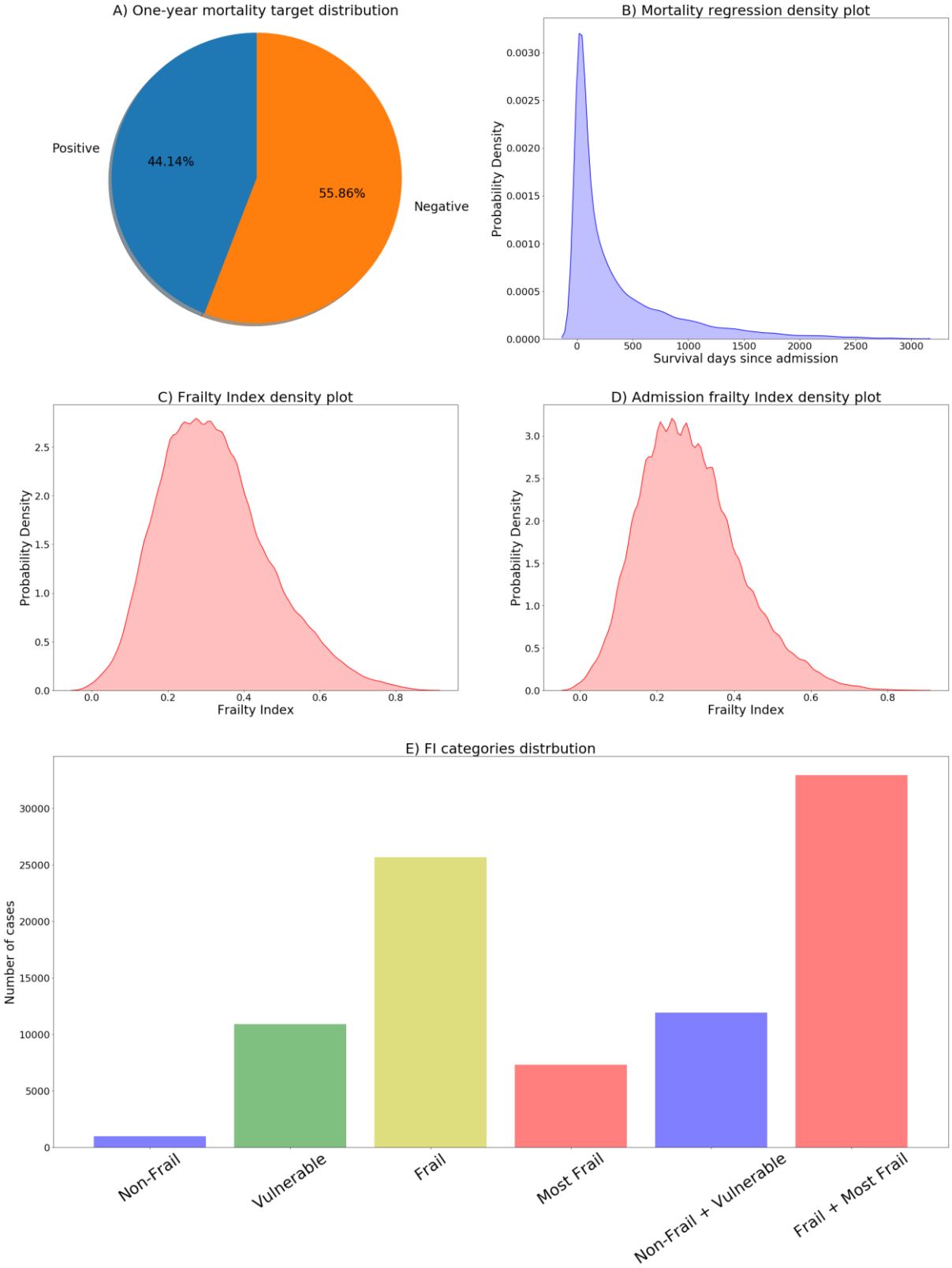

Figure 1: A) One-year mortality target distribution; B) Density plot from survival regression target variable; C) Density plot from the FI target variable; D) Density plot from the admission FI; E) FI categories distribution.

The calculated spearman's correlation coefficients are survival regression target vs admission FI = -0.1 and survival regression target vs FI target = -0.16. The one-year mortality vs stratified (binary) FI

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

was statistically significant in the two-sided fisher's exact test (p=2.96e-86). All correlations are statistically significant with p-values < 0.001. Complete details available in Table 6.

3.2 One-year morality classifier

Following the described methodology, the recursive feature elimination process provided the list of most significant variables, the three more relevant were: Number of active groups, Real Service Code and Charlson. Table 3a contains the list of the selected variables and the mean GINI importance assigned by the GBM model during the k-fold cross validation evaluation. The logistic regression one-year mortality model achieved and AUC of 0.79, while the one based on GBM presented a great discriminative power with an AUC ROC of 0.87, complete results on Table 3b.

Table 3a: Importance assigned at the evaluation of the one-year mortality model

Variable	GINI Importance (%)
Number of active groups	18.90 ± 0.44
Real service code	16.50 ± 2.60
Charlson index	13.37 ± 1.15
Age	8.70 ± 0.40
Barthel index	6.16 ± 0.75
Leukocyte	5.73 ± 0.17
Urea	5.31 ± 0.25
RDW-SD	4.80 ± 0.36
DRG	3.77 ± 0.21
Sodium	3.05 ± 0.13
RDW-CV	2.52 ± 0.36
PCR	2.44 ± 0.10
Creatinine	1.78 ± 0.14
Glucose	1.48 ± 0.10
Number of previous stays	1.35 ± 0.09
Haematocrit	1.23 ± 0.14
Admission diagnose code	1.14 ± 0.09
Potassium	0.75 ± 0.05
Haemoglobin	0.53 ± 0.08
Number Previous ER	0.50 ± 0.04

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Table 3b: One-year mortality classifier evaluation

Model	AUC ROC	Sensitivity (TPR)	Specificity (TNR)	Accuracy
GBM	**0.87 ± 0.01**	0.80 ± 0.03	0.77 ± 0.05	0.80 ± 0.03
Logistic Reg.	0.79 ± 0.01	0.73 ± 0.04	0.69 ± 0.03	0.70 +- 0.01

3.3 Survival regression

For the survival regression model, the most important variables in the model were: real service code, number of active groups and the Charlson index. Table 4a shows the whole list of variables and the importance computed in the model validation. The cox regression produced a MAE of 956.55, while the GBM model achieved a MAE of 329.97, complete performance for survival regression models on Table 4b.

Table 4a: Importance assigned at the evaluation of the survival regression model

Variable	**GINI Importance (%)**
Real Service Code	11.55 ± 0.90
Number Active groups	11.11 ± 0.46
Charlson	8.80 ± 0.56
Admission Diagnose Code	7.39 ± 0.40
Filtered Glomerular CKD	7.18 ± 0.51
Age	5.90 ± 0.37
Leukocyte	5.80 ± 0.25
DRG	5.15 ± 0.29
Urea	5.15 ± 0.52
PCR	4.89 ± 0.52
RDW-SD	4.19 ± 0.37
Number Previous ER	3.76 ± 0.43
Aspartate Aminotransferase	3.00 ± 0.17
Creatinine	2.87 ± 0.26
Glucose	2.76 ± 0.30
RDW-CV	2.75 ± 0.44
Previous Stays	2.63 ± 0.15
Haematocrit	2.18 ± 0.21
Potassium	1.82 ± 0.19
Number Previous ER 365d	1.12 ± 0.14

Table 4b: Mortality regressor evaluation

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

Model	MSE	MAE
Cox	1300783.79 ± 59500.27	956.55 ± 26.89
GBM Survival Reg.	215175.28 ± 6084.36	329.97 ± 5.24

3.4 One-year frailty classifier

Finally, we performed the same evaluation for the frailty classifier. The most important variable according to the GINI index were: Charlson index, emergency room visits and Hypertension (complete details on Table 5a). The classification model based on the logistic regression achieved an AUC ROC of 0.84, while the model based the GBM outperformed it with an AUC ROC of 0.9. Complete metrics for the frailty classification are available in Table 5b.

Table 5a: Importance assigned at the evaluation of the survival regression model

Variable	**GINI Importance (%)**
Charlson index	37.19 ± 0.51
Number Previous ER	9.14 ± 0.21
Hypertension	6.96 ± 0.21
Psychiatric Disease	6.31 ± 0.14
Atrial Fibrillation	6.12 ± 0.16
Gastrointestinal or Liver Disease	5.96 ± 0.12
Barthel index	4.95 ± 0.24
Age	4.91 ± 0.17
Real Service Code	3.52 ± 0.33
Haematocrit	2.26 ± 0.1
RDW-SD	1.94 ± 0.19
RDW-CV	1.85 ± 0.21
Glucose	1.81 ± 0.17
Creatinine	1.38 ± 0.12
DRG	1.27 ± 0.11
Leukocyte	1.11 ± 0.07
Admission Diagnose Code	0.9 ± 0.06
Urea	0.85 ± 0.12
Potassium	0.83 ± 0.07
Previous Stays	0.75 ± 0.07

Table 5b: Frailty classifier evaluation

Model	AUC ROC	Sensitivity (TPR)	Specificity (TNR)	Accuracy
GBM	**0.90** ± 0.01	0.79 ± 0.03	0.84 ± 0.03	0.81 ± 0.01
Logistic Reg.	**0.84** ± 0.01	0.74 ± 0.04	0.78 ± 0.05	0.75 ± 0.02

4. Discussion

As previous studies have shown, mortality and frailty could be relevant criteria to admit patients to PC programs. Therefore, health professionals could benefit from the use of data-driven accurate predictions of these two dimensions. In addition to the benefits experienced by patients and their families, the early identification of these patients' needs can help better manage the available health and social care resources and may even reduce costs overall.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

We set the number of variables on each model to the 20 most relevant because the original number of variables was arguably too high to be used by a human operator, also, not every variable is relevant in all three problems. The authors decided to select the most important variables using the Random Forest's GINI importance criteria with recursive feature elimination as a data-driven method, this is a standard machine learning procedure and applies to all three tasks. As shown in Table B, all three problems share a great number of variables, being only 27 different variables. The selected variables by the recursive feature elimination algorithm are coherent with the different mortality works in the literature [20, 21].

Our one-year mortality model rank among the best general admission models in terms of AUC ROC (0.87 ± 0.01). Outperforming PROFUND (0.77) [20] and scoring slightly below HOMR (0.89-0.92) [21], mHOMR (0.89) [23] and Avati's deep learning approach (0.93, 0.87 for admitted only patients) [22]. However, our model is not fully comparable since it targeted older adults (>= 65 years old) meanwhile all the mentioned studies use inclusion criteria of >= 18, except Avati, which also includes paediatric records. As expected, the GBM model performed significantly better than the Logistic Regression counterpart.

Most of the works in the literature dealing with survival are based in cox regression model [37]. The Cox model offers some advantages with respect to the standard machine learning models, the acceptance of censored data being one of the most important. The Cox model allows the analysis of the simultaneous effect of a set of covariables in the survival expressed by the hazard ratio. The Cox model implemented in lifelines [43] also allows to obtain a prediction of survival by calculating the survival expected time.

In the survival regression problem, our model scores a mean absolute error of 329.97 days, outperforming the 956.55 days scored by the cox model. Despite obtaining better predictions than one of the most used models when dealing with survival time, a mean error of almost a year seems not quite useful for the original purpose of this model. The authors suggest the use of this kind of model only when the one-year mortality model provides a positive prediction. A preliminary result for the regression model trained only with those patients whose survival target was <= 365 provided a mean absolute error of 68.02 days. This result fits much better with the purpose of the survival model. It seems to indicate that the poor performance of our original model is due to the presence of huge values (much bigger than 365) in the target variable, as can be observed in the long right tail of Figure B.

We composed our frailty index using 29 variables, some of them evaluate the ability of the patients for taking care for him/herself in daily activities such as grooming, dressing or bathing, and the others designate the presence or absence of a certain diagnosed diseases (complete list on Table 2). The authors followed the steps for the creation of the frailty index by Searle et al. [30].

The binary frailty model scored a 0.9 of AUC ROC, 0.79 of sensitivity and 0.84 of specificity, outperforming the logistic regression version (0.84 AUC ROC, 0.74 Sensitivity, 0.78 Specificity). These results demonstrate a great predictive power for assessing a patient's frailty index category one year from the admission. As far as we know, this is the first study where a model is used to predict a future frailty status without using proxies such as mortality or disability. These results provide a complementary perspective based on an objective measure of frailty to initiate early PC.

The mean admission FI was 0.27 ± 0.12, and its shape resembles a normal distribution. This is a coherent behaviour with the findings in the Mitnitski et al. study [24], where the most impaired groups have a bigger FI mean, and the distribution is shaped like a normal distribution, as opposed

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

to the less impaired groups which had a smaller mean FI and can be approximated using a gamma distribution. The correlation between our admission FI and MR target in days is –0.10, lower than the one reported in [24] which was –0.234. These means that the frailty index proposed in this work, for this sample, is less associated with mortality.

The relationship between frailty and mortality have been studied previously [25] confirming the association between both. Despite the similarity in the used input variables, the target variable distributions are poorly correlated (Table 6). Both criteria have been highlighted as important for accessing PC in previous studies and are related. However, they reflect two different distributions, and the authors think of them as two complementary criterions. Therefore, we suggest their aggregation together will increase the information to support the decision-making process.

Table 6: Spearman's correlations (continuous variables) and point-biserial correlation (binary vs continuous) coefficient between FI and mortality variables. All values are statistically significant (p-value < 0.001). Target 1YM (binary) vs Target FI (binary) produced statistical significance using the Fisher's exact test.

	Admission FI (cont.)	Target FI (cont.)	Target FI (binary)	Target Surv. (cont.)
Target FI (cont.)	0.69			
Target FI (binary)	0.54	0.67		
Target Surv. (cont.)	-0.10	-0.16	-0.14	
Target 1YM (binary)	0.10	0.19	*	-0.77

Despite the existence of studies on one-year mortality using machine learning methods, the authors believe that the focus of these models on adults over 65 and, most importantly, the incorporation of the criterion of frailty may represent an added value for those health professionals deciding about inclusion in PC services. We developed a web user interface to incorporate the three models and made them publicly available at http://demoiapc.upv.es/.

The main limitation in our study resides in the use of data from only one hospital, an internal validation only assures the performance of the models with similar data. We cannot ensure the reported efficiency in other hospitals and/or with other patient populations [43]. Also, data from the same centres can change over time due a wide variety of reasons such as change in protocols or external agents such a pandemic [44, 45]. Additional external validations are needed as future work.

5. Conclusion

This work proposes the use of three different machine learning models based on hospital admission data to assess PC needs on older adults and help health professionals in the decision-making process. The authors constructed a version of the one-year mortality models achieving results in state of the art for this problem (AUC ROC = 0.87), a regression mortality model has also created to provide more information about the first prediction. A predictive model to assess frail within the year was developed, presenting a great discriminative power (AUC ROC = 0.9). To our knowledge, this is the first study predicting one-year frailty status based on a frailty index. The frailty index used in this study is coherent with previous studies observations but is less correlated with mortality. The

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

authors propose the use of predictions in both mortality and frailty as complementary predictions to help assess PC needs due to its individual relevance but weak correlation and the reliability and great predictive power. The described models have been implemented and publicly available at http://demoiapc.upv.es/.

6. Conflict of Interest

The authors declare that they have no conflicts of interest in the research.

7. Acknowledgements

This work was supported by the InAdvance project (H2020-SC1-BHC-2018-2020 No. 825750.) and the ALBATROSS project (PID2019-104978RB-I00). The authors thanks their contributions to María Soledad Giménez-Campos, María Eugenia Gas-López, María José Caballero Mateos and Bernardo Valdivieso. Special thanks to Ángel Sánchez-García for his contributions to the website.

8. References

[1] Callaway MV, Connor SR, Foley KM. World Health Organization public health model: a roadmap for palliative care development. Journal of pain and symptom management. 2018 Feb 1;55(2):S6-13.

[2] Temel JS, Greer JA, Muzikansky A, Gallagher ER, Admane S, Jackson VA, Dahlin CM, Blinderman CD, Jacobsen J, Pirl WF, Billings JA. Early palliative care for patients with metastatic non–small-cell lung cancer. New England Journal of Medicine. 2010 Aug 19;363(8):733-42.

[3] Bakitas M, Lyons KD, Hegel MT, Balan S, Brokaw FC, Seville J, Hull JG, Li Z, Tosteson TD, Byock IR, Ahles TA. Effects of a palliative care intervention on clinical outcomes in patients with advanced cancer: the Project ENABLE II randomized controlled trial. Jama. 2009 Aug 19;302(7):741-9.

[4] Yennurajalingam S, Urbauer DL, Casper KL, Reyes-Gibby CC, Chacko R, Poulter V, Bruera E. Impact of a palliative care consultation team on cancer-related symptoms in advanced cancer patients referred to an outpatient supportive care clinic. Journal of Pain and Symptom Management. 2011 Jan 1;41(1):49-56.

[5] Quinn KL, Stukel T, Stall NM, Huang A, Isenberg S, Tanuseputro P, Goldman R, Cram P, Kavalieratos D, Detsky AS, Bell CM. Association between palliative care and healthcare outcomes among adults with terminal non-cancer illness: population based matched cohort study. bmj. 2020 Jul 6;370.

[6] Bakitas MA, Tosteson TD, Li Z, Lyons KD, Hull JG, Li Z, Dionne-Odom JN, Frost J, Dragnev KH, Hegel MT, Azuero A. Early versus delayed initiation of concurrent palliative oncology care: patient outcomes in the ENABLE III randomized controlled trial. Journal of clinical oncology. 2015 May 1;33(13):1438.

[7] McIlfatrick S. Assessing palliative care needs: views of patients, informal carers and healthcare professionals. Journal of advanced nursing. 2007 Jan;57(1):77-86.

[8] Addington-Hall JM, Higginson IJ. Palliative care for non-cancer patients. Oxford University Press; 2001 Jul 15.

[9] Kingston AE, Kirkland J, Hadjimichalis A. Palliative care in non-malignant disease. Medicine. 2020 Jan 1;48(1):37-42.

[10] Raudonis BM, Daniel K. Frailty: an indication for palliative care.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

[11] Koller K, Rockwood K. Frailty in older adults: implications for end-of-life care. Cleve Clin J Med. 2013 Mar 1;80(3):168-74.

[12] Etkind SN, Bone AE, Gomes B, Lovell N, Evans CJ, Higginson IJ, Murtagh FE. How many people will need palliative care in 2040? Past trends, future projections and implications for services. BMC medicine. 2017 Dec;15(1):1-0.

[13] Murray SA, Firth A, Schneider N, Van den Eynden B, Gomez-Batiste X, Brogaard T, Villanueva T, Abela J, Eychmuller S, Mitchell G, Downing J. Promoting palliative care in the community: production of the primary palliative care toolkit by the European Association of Palliative Care Taskforce in primary palliative care. Palliative medicine. 2015 Feb;29(2):101-11.

[14] Moss AH, Ganjoo J, Sharma S, Gansor J, Senft S, Weaner B, Dalton C, MacKay K, Pellegrino B, Anantharaman P, Schmidt R. Utility of the "surprise" question to identify dialysis patients with high mortality. Clinical Journal of the American Society of Nephrology. 2008 Sep 1;3(5):1379-84.

[15] Downar J, Goldman R, Pinto R, Englesakis M, Adhikari NK. The "surprise question" for predicting death in seriously ill patients: a systematic review and meta-analysis. CMAJ. 2017 Apr 3;189(13):E484-93.

[16] Lau F, Downing M, Lesperance M, Karlson N, Kuziemsky C, Yang J. Using the Palliative Performance Scale to provide meaningful survival estimates. Journal of pain and symptom management. 2009 Jul 1;38(1):134-44.

[17] Linklater G, Lawton S, Fielding S, Macaulay L, Carroll D, Pang D. Introducing the Palliative Performance Scale to clinicians: the Grampian experience. BMJ Supportive & Palliative Care. 2012 Jun 1;2(2):121-6.

[18] Highet G, Crawford D, Murray SA, Boyd K. Development and evaluation of the Supportive and Palliative Care Indicators Tool (SPICT): a mixed-methods study. BMJ supportive & palliative care. 2014 Sep 1;4(3):285-90.

[19] Woolfield A, Mitchell G, Kondalsamy-Chennakesavan S, Senior H. Predicting Those Who Are at Risk of Dying within Six to Twelve Months in Primary Care: A Retrospective Case–Control General Practice Chart Analysis. Journal of palliative medicine. 2019 Nov 1;22(11):1417-24.

[20] Bernabeu-Wittel M, Ollero-Baturone M, Moreno-Gaviño L, Barón-Franco B, Fuertes A, Murcia-Zaragoza J, Ramos-Cantos C, Alemán A, Fernández-Moyano A. Development of a new predictive model for polypathological patients. The PROFUND index. European journal of internal medicine. 2011 Jun 1;22(3):311-7.

[21] van Walraven C, McAlister FA, Bakal JA, Hawken S, Donzé J. External validation of the Hospital-patient One-year Mortality Risk (HOMR) model for predicting death within 1 year after hospital admission. Cmaj. 2015 Jul 14;187(10):725-33.

[22] Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. BMC medical informatics and decision making. 2018 Dec 1;18(4):122.

[23] Wegier P, Koo E, Ansari S, Kobewka D, O'Connor E, Wu P, Steinberg L, Bell C, Walton T, van Walraven C, Embuldeniya G. mHOMR: a feasibility study of an automated system for identifying inpatients having an elevated risk of 1-year mortality. BMJ Quality & Safety. 2019 Dec 1;28(12):971-9.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

[24] Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. TheScientificWorldJournal. 2001 Aug 8;1.

[25] Shamliyan T, Talley KM, Ramakrishnan R, Kane RL. Association of frailty with survival: a systematic literature review. Ageing research reviews. 2013 Mar 1;12(2):719-36.

[26] Babič F, Majnarić LT, Bekić S, Holzinger A. Machine Learning for Family Doctors: A Case of Cluster Analysis for Studying Aging Associated Comorbidities and Frailty. InInternational Cross-Domain Conference for Machine Learning and Knowledge Extraction 2019 Aug 26 (pp. 178-194). Springer, Cham.

[27] Sternberg SA, Bentur N, Abrams C, Spalter T, Karpati T, Lemberger J, Heymann AD. Identifying frail older people using predictive modeling. The American journal of managed care. 2012 Oct;18(10):e392-7.

[28] Saliba D, Elliott M, Rubenstein LZ, Solomon DH, Young RT, Kamberg CJ, Carol Roth RN, MacLean CH, Shekelle PG, Sloss EM, Wenger NS. The Vulnerable Elders Survey: a tool for identifying vulnerable older people in the community. Journal of the American Geriatrics Society. 2001 Dec;49(12):1691-9.

[29] Bertini F, Bergami G, Montesi D, Veronese G, Marchesini G, Pandolfi P. Predicting frailty condition in elderly using multidimensional socioclinical databases. Proceedings of the IEEE. 2018 Feb 5;106(4):723-37

[30] Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC geriatrics. 2008 Dec;8(1):24.

[31] Hoover M, Rotermann M, Sanmartin C, Bernier J. Validation of an index to estimate the prevalence of frailty among community-dwelling seniors. Health Rep. 2013 Sep 18;24(9):10-7.

[32] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press; 1984.

[33] Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of statistics. 2001 Oct 1:1189-232.

[34] Touzani S, Granderson J, Fernandes S. Gradient boosting machine for modeling the energy consumption of commercial buildings. Energy and Buildings. 2018 Jan 1;158:1533-43.

[35] Chen X, Huang L, Xie D, Zhao Q. EGBMMDA: extreme gradient boosting machine for MiRNA-disease association prediction. Cell death & disease. 2018 Jan 5;9(1):1-6.

[36] Zhou J, Li E, Yang S, Wang M, Shi X, Yao S, Mitri HS. Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories. Safety Science. 2019 Oct 1;118:505-18.

[37] Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological). 1972 Jan;34(2):187-202.

[38] Van Rossum G, Drake FL. The python language reference manual. Network Theory Ltd.; 2011 Mar 1.

[39] Walt SV, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation. Computing in science & engineering. 2011 Mar;13(2):22-30.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.22.21249726; this version posted January 25, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-ND 4.0 International license .

[40] McKinney W. Data structures for statistical computing in python. InProceedings of the 9th Python in Science Conference 2010 Jun 28 (Vol. 445, pp. 51-56).

[41] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. 2011 Nov 1;12:2825-30.

[42] Davidson-Pilon C. lifelines: survival analysis in Python. Journal of Open Source Software. 2019 Aug 4;4(40):1317.

[43] Sáez C, Romero N, Conejero JA, García-Gómez JM. Potential limitations in COVID-19 machine learning due to data source variability: a case study in the nCov2019 dataset. Journal of the American Medical Informatics Association. 2020 Oct 7.

[44] Sáez C, García-Gómez JM. Kinematics of big biomedical data to characterize temporal variability and seasonality of data repositories: functional data analysis of data temporal evolution over non-parametric statistical manifolds. International journal of medical informatics. 2018 Nov 1;119:109-24.

[45] Sáez C, Gutiérrez-Sacristán A, Kohane I, García-Gómez JM, Avillach P. EHRtemporalVariability: delineating temporal dataset shifts in electronic health records. medRxiv. 2020 Jan 1.