

Original Paper

# Machine Learning Model for Predicting Mortality Risk in Patients With Complex Chronic Conditions: Retrospective Analysis

Guillem Hernández Guillamet<sup>1,2</sup>, BSc, MSc; Ariadna Ning Morancho Pallaruelo<sup>2</sup>, BSc; Laura Miró Mezquita<sup>1,2</sup>, BSc, MSc; Ramón Miralles<sup>3,4</sup>, MD, PhD; Miquel Àngel Mas<sup>3,4</sup>, MD, PhD; María José Ulldemolins Papaseit<sup>5,6</sup>, MD; Oriol Estrada Cuxart<sup>1,2</sup>, MD, PhD; Francesc López Seguí<sup>2,7</sup>, MSc, PhD

<sup>1</sup>Research Group on Innovation, Health Economics and Digital Transformation, Institut Germans Trias i Pujol, Badalona, Spain

<sup>2</sup>Hospital Germans Trias i Pujol, Institut Català de la Salut, Badalona, Spain

<sup>3</sup>Direcció Clínica Territorial de Cronicitat Metropolitana Nord, Institut Català de la Salut, Badalona, Spain

<sup>4</sup>Department of Geriatrics, Hospital Germans Trias i Pujol, Badalona, Spain

<sup>5</sup>Direcció d'Atenció Primària Metropolitana Nord, Institut Català de la Salut, Badalona, Spain

<sup>6</sup>Servei d'Atenció Primària Barcelonès Nord, Institut Català de la Salut, Barcelona, Spain

<sup>7</sup>Chair in Palliative Care, Centre for Health and Social Care Research, University of Vic - Central University of Catalonia, Vic, Spain

**Corresponding Author:**

Francesc López Seguí, MSc, PhD

Chair in Palliative Care, Centre for Health and Social Care Research

University of Vic - Central University of Catalonia

Carrer Miquel Martí i Pol, 1

Vic, 08500

Spain

Phone: 1 938863342

Email: [francesc.lopez.segui@gmail.com](mailto:francesc.lopez.segui@gmail.com)

## Abstract

**Background:** The health care system is undergoing a shift toward a more patient-centered approach for individuals with chronic and complex conditions, which presents a series of challenges, such as predicting hospital needs and optimizing resources. At the same time, the exponential increase in health data availability has made it possible to apply advanced statistics and artificial intelligence techniques to develop decision-support systems and improve resource planning, diagnosis, and patient screening. These methods are key to automating the analysis of large volumes of medical data and reducing professional workloads.

**Objective:** This article aims to present a machine learning model and a case study in a cohort of patients with highly complex conditions. The object was to predict mortality within the following 4 years and early mortality over 6 months following diagnosis. The method used easily accessible variables and health care resource utilization information.

**Methods:** A classification algorithm was selected among 6 models implemented and evaluated using a stratified cross-validation strategy with k=10 and a 70/30 train-test split. The evaluation metrics used included accuracy, recall, precision,  $F_1$ -score, and area under the receiver operating characteristic (AUROC) curve.

**Results:** The model predicted patient death with an 87% accuracy, recall of 87%, precision of 82%,  $F_1$ -score of 84%, and area under the curve (AUC) of 0.88 using the best model, the Extreme Gradient Boosting (XGBoost) classifier. The results were worse when predicting premature deaths (following 6 months) with an 83% accuracy (recall=55%, precision=64%  $F_1$ -score=57%, and AUC=0.88) using the Gradient Boosting (GRBoost) classifier.

**Conclusions:** This study showcases encouraging outcomes in forecasting mortality among patients with intricate and persistent health conditions. The employed variables are conveniently accessible, and the incorporation of health care resource utilization information of the patient, which has not been employed by current state-of-the-art approaches, displays promising predictive power. The proposed prediction model is designed to efficiently identify cases that need customized care and proactively anticipate the demand for critical resources by health care providers.

(*Online J Public Health Inform* 2023;15:e52782) doi: [10.2196/52782](https://doi.org/10.2196/52782)

**KEYWORDS**

machine learning; mortality prediction; chronicity; chronic; complex; artificial intelligence; complexity; health data; predict; prediction; predictive; mortality; death; classification; algorithm; algorithms; mortality risk; risk prediction

**Introduction**

Current health care systems are evolving toward a more patient-centered approach, ensuring continuity of care between primary and hospital care. Their goal is to work toward implementing strategies that favor integrated health and social care of excellent quality for individuals with chronic conditions, such as frailty, multimorbidity, or advanced illness, by prioritizing community care led by primary care teams to improve quality of life across the different stage of disease pathways—especially at the end of life [1-3]. This change in health care provision brings out a series of challenges such as predicting populations' needs, selecting high-risk cases, and tracking illness trajectories, as well as tools to analyze which factors determine high patients' needs in different stages, with the aim of optimizing resources, [4,5]. In Institut Català de la Salut (ICS), Catalonia's main public health provider, this model has been designed and is being implemented by the Chronic Care Management Team in Barcelona's North Metropolitan Area under the name "Community-Based Integrated Care Program for People With Complex Chronic Conditions" (ProPCC) [6] with promising results in terms of the decrease in emergency department attendance and hospitalizations [7]. In this scenario, the research team recognizes the necessity of developing supportive tools for care activities and data exploration. These tools aim to predict resource utilization for individual patients, their progression, and the variables influencing it. This initiative is geared toward enhancing resource planning within the program.

The exponential increase in the availability of health data is enabling the application of advanced statistics and artificial intelligence (AI) techniques for the health care system to evolve toward data-driven policy-making and develop decision support systems to improve efficiency in resource planning, prevent diagnoses evolutions, predict patient progression, and make screenings. However, most data currently being collected are not analyzed, either due to a lack of resources or professionals. AI and machine learning (ML) techniques are key to solving these challenges and automating the analysis of large volumes of medical data to reduce professionals' workload and save resources [8,9]. Moreover, there is a severe shortage of professionals in the health care system, which these algorithms could help alleviate. They could greatly aid health care professionals by optimizing resource use and decreasing systemic gaps, with a particular focus on improving structure for the care of high-needs populations. The developed AI models must be robust, ensure privacy, and be easily generalized to other data and contexts in the health care field [10].

In the context of chronic care planning, predicting patient mortality in a short period of time could help the system better plan critical resources, most of which are spent in the last years of life, and provide closer accompaniment during the last phase of a patient's life [11]. In the medical field, predicting mortality has always been carried out in controlled contexts or in groups

of patients with a specific disease. Most of this research has focused on predicting intensive care unit (ICU) mortality. Various studies have advocated the use of ML techniques over the use of logistic regression methods for prediction [12-14], achieving promising results, with an area under the receiver operating characteristic (AUROC) curve of 0.9 [13]. In a hemodialysis cohort, ML techniques were applied to predict death in patients who were critically ill (AUROC=0.86) [15].

In terms of predicting mortality focusing on certain diseases, there are multiple mortality risk scores for different diseases that have been used for a while in the Catalan Health system [16-18]. Regarding ML techniques, many of the models focused on diseases with a high prevalence in the population, such as cardiovascular diseases. ML algorithms such as the support vector machine-radial basis function (SVM-RBF) achieve good results in predicting the mortality onset of cardiovascular diseases [19-21]. Multiple studies use Cox proportional hazards regression models to predict the risk of death due to progressive chronic heart failure [22] and breast cancer [23]. In the case of patients with diabetes, some models have achieved good results in predicting the disease using SVM or NB algorithms, reaching an accuracy of 82% [11,24,25]. Research has prioritized identifying predictors of mortality related to diabetes, including hypercholesterolemia or serum albumin, which have been found to be significant predictors of death in patients undergoing hemodialysis [26,27].

With the COVID-19 pandemic, the need to generate prediction models at a population level has been brought to the forefront. The research paradigm has shifted from predicting mortality in a highly controlled population cohort with a specific clinical profile to the general population, with much more diverse clinical, socioeconomic, and demographic realities. The pandemic has provided the research field with a much higher volume of resources and data to develop models with acceptable accuracy. Recent studies have used deep learning (DL) models to predict the probability of entering the ICU or dying for patients diagnosed with COVID-19 (AUROC=0.84) [28]. Others have tried to predict the mortality rate of critically ill patients diagnosed with COVID-19, achieving promising results (AUROC=0.87) [29]. Other studies sought to predict mortality at the national level [30-32]. Regarding the prediction of mortality in patients with complex chronic conditions, AI and ML algorithms have not yet been used. Currently, only frailty indexes have been used to predict mortality risk in older populations [16].

In this context, this study aims to present an ML model applied to a cohort of patients with highly complex diagnoses, categorized as either a patient with complex chronic disease (CCP) or a patient with advanced chronic disease (ACP). These terms serve as indicators of health and social complexity as defined by the Catalan public health system [17]. The aim is to predict mortality, which is indicative of the severity of their condition by utilizing easily accessible retrospective data. The

model aims to predict a future event, death, by incorporating different population variables to ascertain how their progression would be in each case.

## Methods

### Data Sources and Inclusion Criteria

The data set was obtained from the administrative database of Barcelonès Nord Primary Care Centers linked to the program and the Germans Trias i Pujol Hospital, a reference hospital for the region that covers a population of 1,448,812 people belonging to 70 municipalities. It includes the period between May 2018 and December 2021 (with a slowdown of inclusion due to the COVID-19 pandemic in 2020). The observation study was focused on individuals from the ProPCC cohort, including high-needs high-cost patients cataloged as CCP or ACP based on the Catalonia Department of Health's risk stratification strategy [32,33]. This is a subanalysis from a retrospective observational study to estimate the impact of the introduction of ProPCC.

This study contains the main population characteristics of the patients, including age, sex, main diagnoses, complexity profiles (CCP or ACP), and adjusted mobility groups (ie, grupos de morbilidad ajustados, GMA), and place of habitual residence (home or nursing home). To enter the ProPCC program, patients had to meet one of the inclusion criteria for the program: (1) easily decompensated chronic conditions, (2) frequent visits to hospital emergency rooms for the same reason, (3) dementia with cognitive/behavioral handling challenges, (4) functional dependence with difficult handling, (5) difficulty accepting the

loss of health, (6) polypharmacy with handling difficulty, and (7) advanced complex disease process or caregiver burden.

The database also contains information on the duration of time spent at home and the utilization of health care resources, encompassing primary care and hospital visits within the last 12 and 6 months, specifically, primary care visits and hospital services such as emergency department visits and acute hospitalizations.

### Ethical Considerations

This study was approved by the Medicines Research Ethics Committee of the Institut Universitari d'Investigació en Atenció Primària (IDIAP Jordi Gol; 22/084-P).

### Baseline Characteristics

This study included 264 patients identified as CCP (n=116, 43.9%) or ACP (n=148, 56.1%). The average age at inclusion in the study was 83.6 (SD 10.3) years, and 50% (n=132) of the patients were female. The mean time spent in the program since inclusion was 16.73 (SD 12.8) months. Most patients belonged to the highest adjusted morbidity group (ie, GMA; group 4, n=219, 83%), and 9.1% (n=24) lived in a nursing home. Most were included in the program because they had easily decompensated chronic conditions (n=145, 54.9%). The mean number of active health problems was 24.3 (SD 8.4). [Table 1](#) summarizes all the variables used as predictors in the study. The variable death represents the variable to predict, while the rest act as predictors. Within the first 6 months after inclusion in the ProPCC program, 43.9% (n=116) of the patients in the study died. Subsequently, there were 171 (n=174, 65.9%) additional deaths in the remaining period.

**Table 1.** Clinical characteristics of the implementation cohort (N=264).

Main patient variables	Values
Age (years), mean (SD)	83.6 (10.3)
Female sex, n (%)	132 (50)
<b>Main clinical diagnoses, n (%)</b>	
Respiratory	92 (34.8)
Cardiological	112 (42.4)
Cognitive disorder	86 (32.6)
Diabetes mellitus 2	83 (31.4)
Hypertension	158 (59.8)
Musculoskeletal	156 (59.1)
Renal failure	75 (28.4)
<b>Complexity profiles, n (%)</b>	
CCP <sup>a</sup>	117 (44.3)
ACP <sup>b</sup>	147 (55.7)
<b>Adjusted morbidity groups, n (%)</b>	
Group 4	156 (83)
Group 3	30 (15.9)
Group 1-2	2 (1.3)
<b>Program inclusion criteria, n (%)</b>	
Chronic conditions easily decompensated	145 (54.9)
Frequent visits to the emergency room for the same reason	32 (12.1)
Dementia with cognitive/behavioral management difficulties	23 (8.7)
Hard-to-handle functional dependency	53 (20.1)
Difficulty accepting the loss of health	31 (11.7)
Polypharmacy with complex handling	48 (18.2)
Advanced complex disease process	51 (19.3)
Caregiver burden	31 (11.7)
<b>Consumption of previous care resources, mean (SD)</b>	
Family doctor visit 12 months prior	12.4 (17)
Family doctor visit 6 months prior	8.2 (8.3)
Nursing visit 12 months prior	13.1 (20.7)
Nursing visit 6 months prior	9.3 (11.9)
Social worker visit 12 months prior	1.5 (3.7)
Social worker visit 6 months prior	1 (2.7)
Continued care visit 12 months prior	1.2 (6.6)
Continued care visit 6 months prior	0.7 (3.7)
<b>Consumption of previous hospital care resources, mean (SD)</b>	
Home hospitalization 12 months prior	0.2 (0.6)
Home hospitalization 6 months prior	0.1 (0.4)
Acute hospitalization 12 months prior	1.1 (1.5)
Acute hospitalization 6 months prior	0.7 (1.1)
Visit to the emergency room 12 months prior	2.3 (2.6)

Main patient variables	Values
Visit to the emergency room 6 months prior	1.4 (1.7)
<b>Death (outcome variable), n (%)</b>	
Within the first 6 months of program entry	114 (43.4)
Throughout the rest of the study period	171 (64.7)

<sup>a</sup>CCP: patient with complex chronic disease.

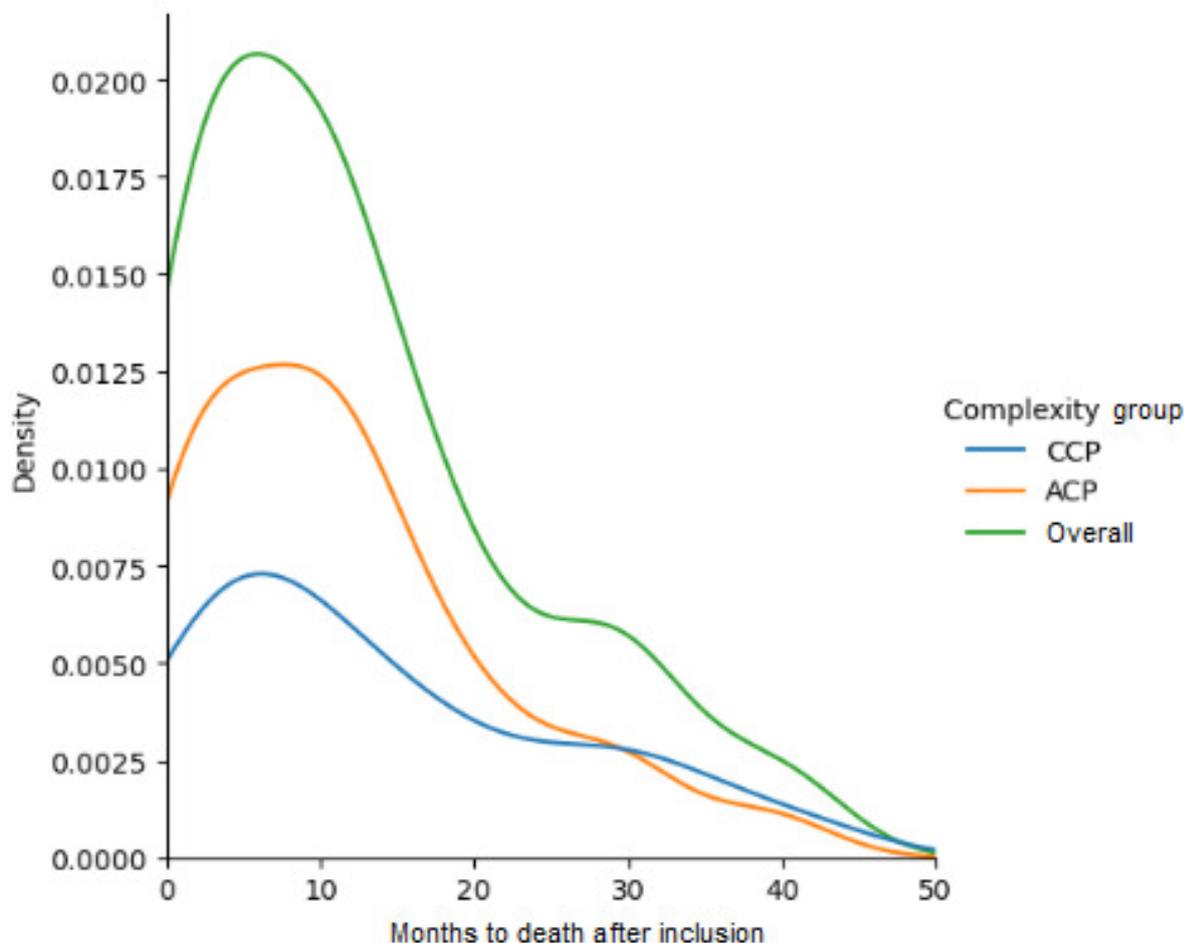
<sup>b</sup>ACP: patient with advanced chronic disease.

### Mortality in High-Risk CCP and ACP

Patients entering the ProPCC program had highly complex and chronic clinical profiles. This caused the death rate to be much higher than in any other health program, with a high probability of having died in the 2 years following the start of the program. Among the 264 patients who were monitored in the ProPCC program during the 48 months of study, 171 (64.7%) died. The case fatality ratio over the first 6 months in the program was 43.2% (n=114). The average patient died after 12.61 (SD 10.8) months. The death distribution over the study time period

representation is shown in Figure 1. Table 2 contains a description of the mortality variable, overall and by complexity subgroups. The study only considers the first 4 years to predict patient mortality. The objective of this analysis was to predict which patients would die within the first 48 months and who would experience premature death within the first 6 months. By definition, a patient with complexity transitions from CCP to ACP. Therefore, it was highly probable that an ACP would die within the first year since the complex diagnosis but less likely to be the case for a CCP.

**Figure 1.** Death distribution over the study period (overall and by complexity groups). ACP: patient with advanced chronic disease; CCP: patient with complex chronic disease.



**Table 2.** Description of ProPCC<sup>a</sup> program prediction variables (mortality up to 4 years).

Prediction variables	Value, n (%)	Mean	SD	Min	Q1	Median	Q3	Max
Months until death overall	171 (64.8)	12.6	10.8	0	4	10	18	43
Months until death for CCP <sup>a</sup>	68 (58% CCP, 39% deaths, 25.8% global)	14.2	11.9	0	4.8	10	22.5	43
Months until death for ACP <sup>b</sup>	103 (70% ACP, 60% deaths, 39% global)	11.6	10	0	3.5	10	16.5	40

<sup>a</sup>CCP: patient with complex chronic disease.

<sup>b</sup>ACP: patient with advanced chronic disease.

## Data Processing

There were no missing values in the database. The numerical variables (ie, age, resource consumption, program inclusion criteria) were normalized and scaled using a min-max scaler. The rest of the categorical variables underwent a one-hot encoding transformation. The classification variable was binarized from the “months until death” variable based on the different experimentation use cases defined in the “Experimental Setup” section. Only retrospective information (up to 1 year prior to inclusion in the ProPCC program) was used to predict death in the following months (48 months or 4 years). No data collected during the response time were used for prediction.

## Model Construction and Evaluation

The developed method aimed to find classification algorithms capable of predicting the mortality of patients with high complexity using retrospective variables available within the health provider's information systems. This approach aimed to use these variables as predictors without the need to conduct surveys or use external data repositories. The ML algorithms we used included SVM with a linear kernel (SVM-Lin) with a one-versus-one decision function, SVM with RBF as a kernel (SVM-RBF, regularization parameter  $c=20$ ), a k-nearest-neighbors classifier (KNN,  $k=25$ ), decision tree classification (Tree), an Extreme Gradient Boosting (XGBoost) classifier, and a Gradient Boosting (GRBoost) classifier (learning\_rate=0.25). All parameters were optimized to identify the best parameterization strategy. All models were implemented using the open-source Python ML library *Scikit-learn* and *dmlc-XGBoost*. A stratified cross-validation (CV) strategy with  $k=10$  and a 70/30 train-test split was used to evaluate and validate the models. Evaluation metrics are averaged over all k-fold CVs.

The metrics used to evaluate the models were accuracy, which measured how many observations, both positive and negative, were correctly classified; recall and precision, extracting the ratio of death patients correctly classified as dead;  $F_1$ -score, which measures the harmonic mean between precision and recall; and AUROC curve, which provides an aggregate measure of performance across all possible classification thresholds.

To obtain the algorithm that yielded the best results in the different experiments, predicting such a critical label as death, we focused on maximizing the recall, and subsequently, the  $F_1$ -score (reducing the number of false negatives). The aim of the model was to save resources by identifying patients with a higher risk of death, and hence more severe health conditions, to participate in the ProPCC study. Thus, it was crucial to reduce

the number of false negatives, as this could imply less care for the patient than necessary.

However, emphasizing precision was not as vital since it could result in offering more services to individuals with lower mortality risk. Participation in the program did not harm the patient; thus, it was not a critical prediction for health. Nonetheless, the primary objective was to minimize the number of patients who are flagged as likely to die but do not; hence, our emphasis was more on  $F_1$ -score than overall accuracy. Employing this methodology and deriving the mean with a 10-fold CV yielded more robust results, balancing the classes and generating results less impacted by the minor imbalance in the data set. The Cochran Q test was performed to assess the robustness of the results.

Once we identified the algorithm with the best results in terms of metrics, we used the open-source Python library *Streamlit* to develop a user-friendly web app with a good user interface and user experience for health care professionals.

## Experimental Setup

This study was divided into 3 experiments based on patients' period of death. All the experiments used the same models, parametrizations, and variables described in Table 1 (ie, age and sex, clinical diagnostics, complexity profile, morbidity profile, program inclusion criteria, and consumption of primary care and hospital care resources in the 6 and 12 months before inclusion). The objective was to detect which model more accurately predicted mortality in each experiment, using the same accessible variables. These experiments were structured as binary classifications (1=death, 0=survival). The first experiment aimed to evaluate the algorithms' ability to predict patient mortality throughout the study period (48 months, 4 years). In this scenario, the binary outcomes were as follows: 1=patient death ( $n=171$ , 64.8%) and 0=patient survival ( $n=93$ , 35.2%). The second scenario was designed to evaluate the algorithms' predictive ability to detect premature mortality (within 6 months). The binary results were as follows: 1=patient death within the first 6 months of the study ( $n=58$ , 22%) and 0=patient survival after 6 months ( $n=206$ , 78%). The third scenario aims to evaluate the algorithm's ability to classify which patients will likely die prematurely (within 6 months) and which will likely die later to check for differences. The binary results were as follows: 1=patient death within the first 6 months of the study ( $n=58$ , 34%) and 0=patient death after 6 months ( $n=113$ , 66%).

## Results

### Overview

The averaged results of all models for the stratified k-fold CV, including all metrics for each algorithm in the different use case scenarios, are presented in Table 3 and Figure 2 to visualize the comparison between the different models.

**Table 3.** Performance results of the algorithms in the different use case scenarios.

Model	Accuracy	Recall	Precision	$F_1$ -score	AUC <sup>a</sup>
<b>Mortality prediction (48 months), mean (SD)</b>					
SVM <sup>b</sup> -Lin <sup>c</sup>	75 (7)	81 (6)	81 (10)	80 (5)	83 (8)
SVM-RBF <sup>d</sup>	74 (10)	80 (9)	81 (11)	80 (7)	82 (9)
KNN <sup>e</sup>	74 (8)	87 (6)	76 (8)	81 (5)	78 (11)
Tree <sup>f</sup>	72 (7)	79 (12)	78 (10)	78 (8)	72 (13)
XGBoost <sup>g</sup>	79 (7)	87 (7)	82 (9)	84 (6)	88 (5)
GRBoost <sup>h</sup>	71 (6)	81 (6)	75 (8)	78 (6)	75 (5)
<b>Early mortality prediction (within 6 months), mean (SD)</b>					
SVM-Lin	82 (9)	61 (31)	63 (22)	57 (22)	83 (14)
SVM-RBF	81 (7)	54 (25)	60 (20)	52 (17)	89 (5)
KNN	79 (7)	31 (18)	48 (35)	34 (23)	84 (8)
Tree	80 (9)	59 (21)	57 (17)	58 (15)	73 (11)
XGBoost	83 (5)	55 (12)	64 (19)	57 (12)	88 (5)
GRBoost	86 (8)	58 (26)	75 (20)	63 (21)	87 (8)
<b>Early mortality prediction vs late mortality (within and after 6 months), mean (SD)</b>					
SVM-Lin	68 (9)	46 (19)	57 (26)	48 (19)	76 (11)
SVM-RBF	66 (13)	45 (21)	62 (26)	47 (14)	72 (14)
KNN	56 (22)	19 (16)	46 (41)	23 (20)	63 (15)
Tree	65 (10)	53 (17)	55 (29)	50 (20)	57 (12)
XGBoost	75 (11)	49 (24)	65 (26)	54 (25)	81 (15)
GRBoost	71 (14)	52 (21)	61 (31)	55 (24)	73 (16)

<sup>a</sup>AUC: area under the curve.

<sup>b</sup>SVM: support vector machine.

<sup>c</sup>Lin: linear kernel.

<sup>d</sup>RBF: radial basis function.

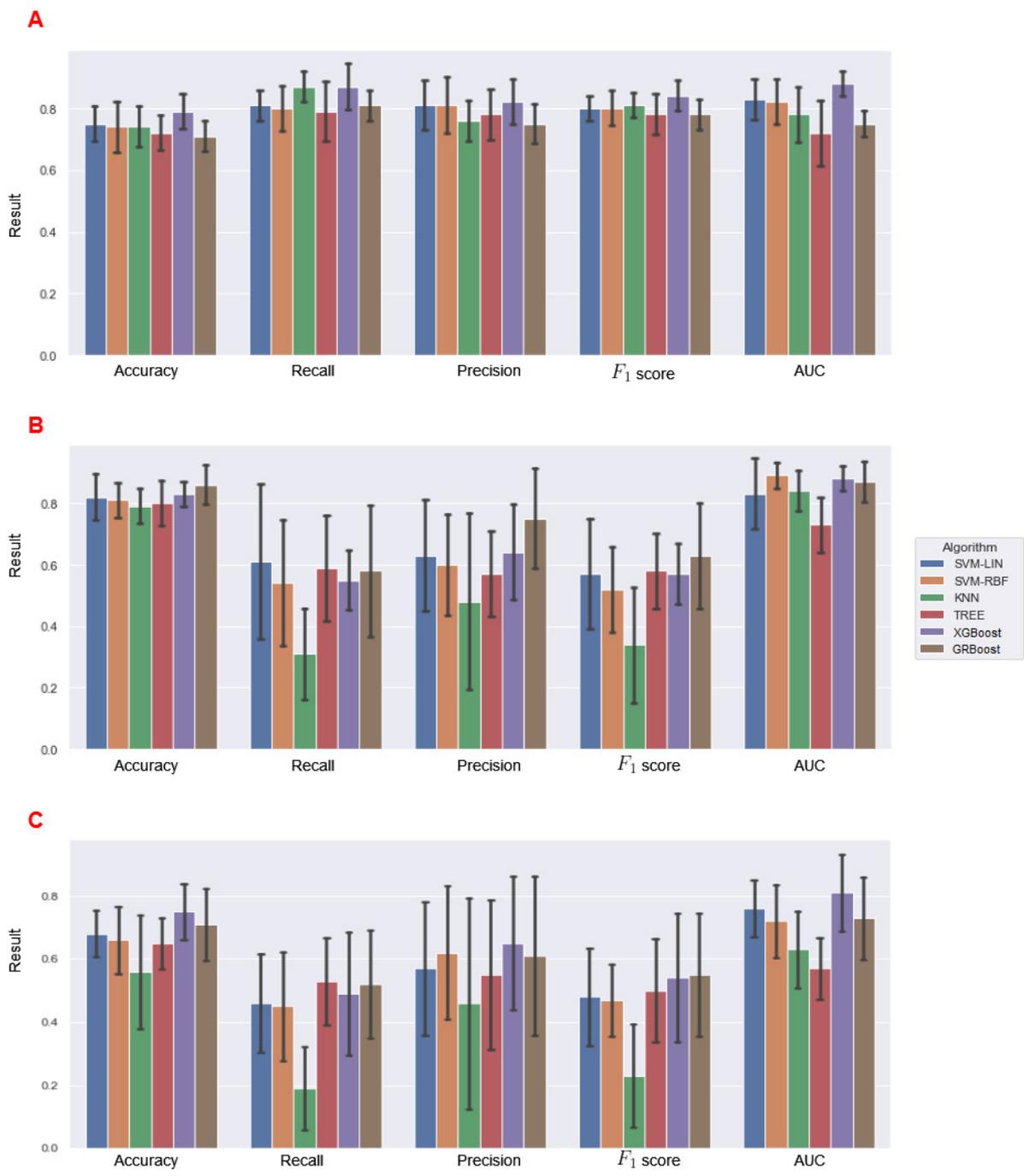
<sup>e</sup>KNN: k-nearest neighbor.

<sup>f</sup>Tree: decision tree.

<sup>g</sup>XGBoost: Extreme Gradient Boosting.

<sup>h</sup>GRBoost: Gradient Boosting.

**Figure 2.** Performance metrics comparison of the different models in the different use case scenarios. (A: mortality in the entire study period, 48 months; B: premature death, 6 months; C: premature death versus dead). AUC: area under curve; GRBoost: Gradient Boosting; KNN: k-nearest neighbor; Lin: linear kernel; RBF: radial basis function; SVM: support vector machine; Tree: decision tree; XGBoost: Extreme Gradient Boosting.



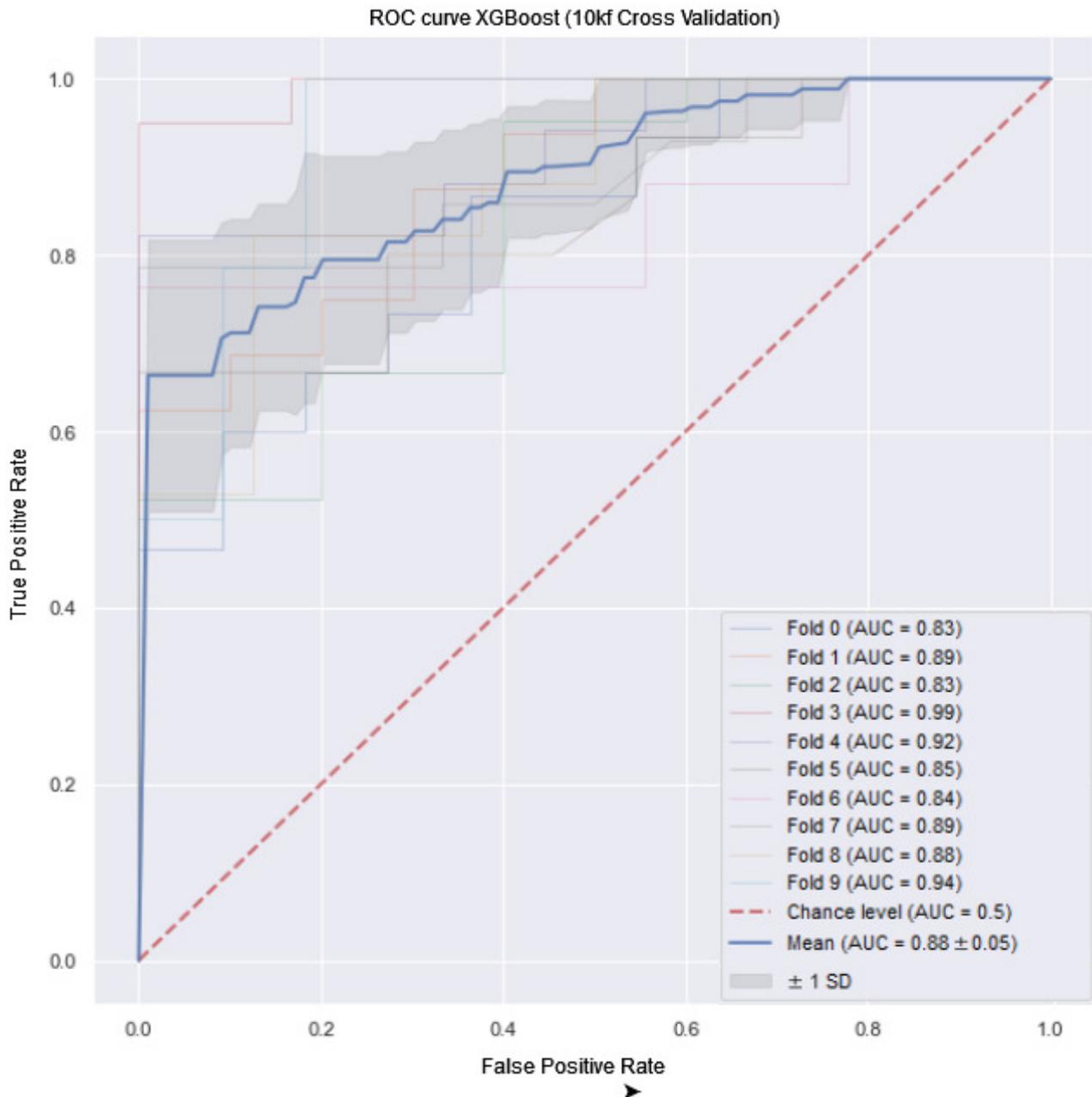
**Model Performance**

The performance of the models varied significantly depending on the use case. The best results were obtained when identifying patient mortality during the entire study period.

In the first experiment, the XBoost algorithm obtained the best results in all metrics. The results were obtained by averaging partial results from a 10 k-fold CV. The SD was small for all metrics, demonstrating robustness across different partial results.

XGBoost achieved the best result, especially in the most critical metrics (recall=87%, F<sub>1</sub>-score=84%, and AUROC=88%). The average Cochran Q test results obtained in different folds were significant in training (Q=137.7, P<.001) but not significant in testing (Q=5.3, P=.43). Regarding testing, there was no significant difference in the accuracy of the different models. However, in terms of the other metrics, the XGBoost model was best suited for predicting patient mortality throughout the study period (48 months, 4 years). Figure 3 shows the AUROC curve of this experiment,

**Figure 3.** Area under the receiver operating characteristic (AUROC) curve for the algorithm with the best performance. Extreme Gradient Boosting (XGBoost) averaged over the k-fold cross-validation (CV, k=10). First experiment (mortality prediction over the entire period: 10 k-fold area under the curve (AUC)=0.88, SD 0.05).



In the second and third experiments, the results are worse. In both cases, the SDs obtained were higher than in the first experiment, indicating that the models were not as robust and were highly sensitive to the cases found in each fold. GRBoost with a 25% learning rate obtained good results in predicting premature deaths (6 months after data collection), especially in terms of accuracy (86%) and AUC (87%), with low SDs. However, in terms of recall (58%), precision (75%), and  $F_1$ -score (63%), the results were less promising and highly sensitive to the fold used (high SDs). SVM-RBF and XGBoost also obtained good results in AUC and accuracy. The average results of the Cochran Q test obtained in the different folds were significant in the training set ( $Q=113.9$ ,  $P<.001$ ) but not significant in the testing set ( $Q=7.1$ ,  $P=.36$ ). Although the results were not as robust and promising as in the first use case, the

overall accuracy and AUC were very high in different models, demonstrating a correlation between the predictors and premature death.

In use case scenario 3, where we aimed to create a model capable of distinguishing between participants who would die prematurely and those who would die during the rest of the study period (without considering survivors), we obtained the poorest results. XGBoost achieved the highest accuracy (75%) and AUC (81%). However, in the other metrics, there was considerable variability across all models, demonstrating that the results were highly sensitive to the fold used. The top  $F_1$ -score (55%) was obtained by GRBoost, although it was much lower than the results obtained in the previous use cases.

As depicted in [Figure 2](#), XGBoost usually obtained the best results in all 3 scenarios. GRBoost and SVM-RBF also performed well in different scenarios.

## Discussion

### Principal Findings

Our research shows that ML and AI models can predict mortality in CCP and ACP with a high degree of accuracy (up to 89%) using easy-to-obtain real-life predictor variables not typically used in this type of experiment. This study illustrates that incorporating these variables has the potential to enhance analysis and accurately predict mortality among patients with significant health and social chronic needs. Previous studies showed similar results [12,14,26,27,29] or better ones [17,22,28,30] in terms of metrics. However, no literature has been found that specifically refers to predicting mortality within the high-need high-cost patient population. Most models for predicting mortality among chronic patients focused on cohorts with a specific diagnosis or establishing basic associations among variables linked to mortality. We have not found a similar use case in the literature that seeks to predict mortality in an identifiable yet heterogeneous group. Additionally, no identified model has been capable of predicting mortality using the consumption of primary care and hospital care resources.

The comparison was carried out with cases of prediction for specific diseases [17,22] or controlled contexts such as ICUs [12,14] or the COVID-19 pandemic outbreak [26-30].

Observing the outcomes, studies predicting mortality in ICUs obtain metrics akin to this study's findings, whereas those addressing specific pathologies present higher metrics. Unlike other death prediction models, the proposed model not only uses patient clinical data such as diagnosis or morbidity but also proposes a framework utilizing medical care resource consumption data during the previous year (visits to different clinical specialties) and sociodemographic health-related variables. These variables include difficulty accepting health loss, dementia, caregiver burden, functional dependence, polypharmacy, or frequent consultations for the same reason. Incorporating these variables as predictors alongside previously established predictive variables resulted in robust outcomes, accurately predicting cases where patients may face mortality. This model can be used to screen patients who will enter the ProPCC program when scaled to all health care areas, aiding in optimizing critical health care resources and supporting patients at a higher risk of mortality.

### Limitations

This study is subject to several limitations that warrant careful consideration. First, it is important to acknowledge the relatively small size of the cohort. The limited number of patients included in our analysis suggests that there is ample room for improvement to bolster the robustness of the findings, particularly through validation in a larger and more diverse cohort. It is worth noting that the incorporation of a 10-fold CV methodology was employed to enhance the reliability of the results and demonstrate the model's predictive capability. However, the inclusion of a larger sample size would provide

more compelling evidence and strengthen the generalizability of our study's outcomes.

Second, an inherent limitation of the study lies in the potential unbalance present within the data sets. This may introduce errors, particularly when predicting mortality, especially in scenarios 2 and 3 where the occurrence of mortality within the sample was relatively lower. Consequently, the model may not adequately capture the intricate patterns and factors associated with mortality in these specific scenarios. It is advisable to address this issue by employing appropriate techniques such as data set balancing if the data set is large enough, thereby potentially improving the model's performance and accuracy in predicting outcomes.

Third, it is essential to consider the quality of data derived from electronic health systems, particularly within noncontrolled environments. While the data set utilized herein exhibited a high standard of coding quality, which was meticulously ensured by a team of proficient professionals, it is crucial to recognize applying these models to uncured data sources within the health care system may yield less favorable results. The inherent variability and potential inconsistencies in data coding across diverse sources may substantially impact the model's performance and its ability to generalize findings to real-world scenarios.

In summary, this study has acknowledged several notable limitations, including the modest cohort size, potential errors arising from data set unbalance, and potential challenges associated with applying the models to real-world data sources characterized by lower data quality. These limitations should be duly considered when interpreting our study findings, and further endeavors should be undertaken to enhance the robustness and applicability of the models.

### Conclusion and Future Work

This study was developed under the hypothesis that mortality among high-need high-cost patients could be predicted in advance using easily accessible real-life patient variables, including clinical, economic resource use, and demographic variables. With this premise, we used an agnostic approach to develop and test a series of ML algorithms to evaluate predictive capacity in a highly controlled cohort of CCP and ACP. The models achieved robust performance in terms of AUC (88%), accuracy (87%), and  $F_1$ -score (84%), indicators that are particularly important in this prediction context.

Our model not only offers a high degree of prediction accuracy but also provides graphs with the importance of each variable in predicting the risk of mortality. The algorithm is encapsulated in a web app developed with the open-source Python library *Streamlit* to improve usability, transportability, and interpretability for medical professionals. This facilitates more efficient utilization of technology resources and enables the system to predict clinical resource usage, thereby offering better support to patients in the end-of-life stage. This type of tool enables process automation and analysis of large volumes of data, while also allowing the seamless transfer of the same infrastructure to different contexts or units.

Regarding future work, there are relevant challenges such as increasing the number of patients and including new clinical and analytical variables that can act as predictors. Our study cohort exhibits high-quality data, meticulously reviewed by multiple professionals to ensure that there are no errors or missing values. In comparison, primary databases, despite

containing a significantly larger patient pool, tend to have lower-quality data. Future objectives include testing the model using this data set and a much larger number of patients, not just from the high-need profile. Deep-learning models could also further refine the results.

## Acknowledgments

This study was carried out with the support of the Generalitat de Catalunya 2021 (SGR 01125) and funded by the Industrial Doctorate Plan 2021 (DI 106), provided by the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR).

## Data Availability

The data supporting this study's findings belong to the Gerència Territorial Metropolitana Nord of the Catalan Health Institute (ICS). Restrictions apply to the availability of these data, which were used after signing a data processing agreement and an ethical agreement that complies with the requirements of the current legal framework.

## Conflicts of Interest

None declared.

## References

1. Busse R, Blümel M, Scheller-Kreinsen D, Zentner A. Tackling Chronic Disease in Europe: Strategies, Interventions and Challenges. Copenhagen, Denmark. WHO Regional Office Europe; 2010.
2. Kodner DL, Spreeuwenberg C. Integrated care: meaning, logic, applications, and implications--a discussion paper. *Int J Integr Care*. Nov 14, 2002;2(4):e12. [FREE Full text] [doi: [10.5334/ijic.67](https://doi.org/10.5334/ijic.67)] [Medline: [16896389](https://pubmed.ncbi.nlm.nih.gov/16896389/)]
3. Gómez-Batiste X, Martínez-Muñoz M, Blay C, Amblàs J, Vila L, Costa X, et al. Prevalence and characteristics of patients with advanced chronic conditions in need of palliative care in the general population: a cross-sectional study. *Palliat Med*. Apr 08, 2014;28(4):302-311. [doi: [10.1177/0269216313518266](https://doi.org/10.1177/0269216313518266)] [Medline: [24403380](https://pubmed.ncbi.nlm.nih.gov/24403380/)]
4. Blumenthal D, Abrams MK. Tailoring complex care management for high-need, high-cost patients. *JAMA*. Oct 25, 2016;316(16):1657-1658. [doi: [10.1001/jama.2016.12388](https://doi.org/10.1001/jama.2016.12388)] [Medline: [27669168](https://pubmed.ncbi.nlm.nih.gov/27669168/)]
5. Bleich SN, Sherrod C, Chiang A, Boyd C, Wolff J, DuGoff E, et al. Systematic review of programs treating high-need and high-cost people with multiple chronic diseases or disabilities in the United States, 2008-2014. *Prev Chronic Dis*. Nov 12, 2015;12:E197. [FREE Full text] [doi: [10.5888/pcd12.150275](https://doi.org/10.5888/pcd12.150275)] [Medline: [26564013](https://pubmed.ncbi.nlm.nih.gov/26564013/)]
6. Mas MA, Miralles R, Heras C, Ulldemolins MJ, Bonet JM, Prat N, et al. ProPCC-Badalona Group. Designing a person-centred integrated care programme for people with complex chronic conditions: a case study from Catalonia. *Int J Integr Care*. Nov 25, 2021;21(4):22. [FREE Full text] [doi: [10.5334/ijic.5653](https://doi.org/10.5334/ijic.5653)] [Medline: [34899101](https://pubmed.ncbi.nlm.nih.gov/34899101/)]
7. Mas Bergas MAA. Metropolitana Nord Community Based Integrated Care Programme to People with Complex Chronic Conditions (Programa ProPCC): description of an ongoing implementation process. *Int J Integr Care*. Sep 01, 2021;21(S1):239. [doi: [10.5334/ijic.icic20395](https://doi.org/10.5334/ijic.icic20395)]
8. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. Jan 7, 2019;25(1):44-56. [doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7)] [Medline: [30617339](https://pubmed.ncbi.nlm.nih.gov/30617339/)]
9. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. May 2019;20(5):e262-e273. [doi: [10.1016/s1470-2045\(19\)30149-4](https://doi.org/10.1016/s1470-2045(19)30149-4)]
10. Johnson AEW, Ghassemi MM, Nemati S, Niehaus KE, Clifton D, Clifford GD. Machine learning and decision support in critical care. *Proc IEEE*. Feb 2016;104(2):444-466. [doi: [10.1109/jproc.2015.2501978](https://doi.org/10.1109/jproc.2015.2501978)]
11. Vela E, Clèries M, Vella VA, Adroher C, García-Altés A. Population-based analysis of the Healthcare expenditure in Catalonia (Spain): what and who consumes more resources? *Gac Sanit*. Jan 2019;33(1):24-31. [FREE Full text] [doi: [10.1016/j.gaceta.2017.05.017](https://doi.org/10.1016/j.gaceta.2017.05.017)] [Medline: [29129491](https://pubmed.ncbi.nlm.nih.gov/29129491/)]
12. Clermont G, Angus DC, DiRusso SM, Griffin M, Linde-Zwirble WT. Predicting hospital mortality for patients in the intensive care unit: a comparison of artificial neural networks with logistic regression models. *Crit Care Med*. Feb 2001;29(2):291-296. [doi: [10.1097/00003246-200102000-00012](https://doi.org/10.1097/00003246-200102000-00012)] [Medline: [11246308](https://pubmed.ncbi.nlm.nih.gov/11246308/)]
13. Awad A, Bader-El-Den M, McNicholas J, Briggs J. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *Int J Med Inform*. Dec 2017;108:185-195. [doi: [10.1016/j.ijmedinf.2017.10.002](https://doi.org/10.1016/j.ijmedinf.2017.10.002)] [Medline: [29132626](https://pubmed.ncbi.nlm.nih.gov/29132626/)]
14. Dybowski R, Weller P, Chang R, Gant V. Prediction of outcome in critically ill patients using artificial neural network synthesised by genetic algorithm. *Lancet*. Apr 27, 1996;347(9009):1146-1150. [doi: [10.1016/s0140-6736\(96\)90609-1](https://doi.org/10.1016/s0140-6736(96)90609-1)] [Medline: [8609749](https://pubmed.ncbi.nlm.nih.gov/8609749/)]

15. Floege J, Gillespie IA, Kronenberg F, Anker SD, Gioni I, Richards S, et al. Development and validation of a predictive mortality risk score from a European hemodialysis cohort. *Kidney Int.* May 2015;87(5):996-1008. [FREE Full text] [doi: [10.1038/ki.2014.419](https://doi.org/10.1038/ki.2014.419)] [Medline: [25651366](https://pubmed.ncbi.nlm.nih.gov/25651366/)]
16. Kojima G, Iliffe S, Walters K. Frailty index as a predictor of mortality: a systematic review and meta-analysis. *Age Ageing.* Mar 01, 2018;47(2):193-200. [doi: [10.1093/ageing/afx162](https://doi.org/10.1093/ageing/afx162)] [Medline: [29040347](https://pubmed.ncbi.nlm.nih.gov/29040347/)]
17. Catalan model of care for people with frailty, complex chronic (CCP) and advanced chronic (ACP) conditions. Department of Health of Catalonia. 2020. URL: [https://scientiasalut.gencat.cat/bitstream/handle/11351/7007/bases\\_conceptuals\\_model\\_atencio\\_persones\\_fragils\\_cronicitat\\_complexa\\_avancada\\_2020\\_ang.pdf?sequence=8&isAllowed=y](https://scientiasalut.gencat.cat/bitstream/handle/11351/7007/bases_conceptuals_model_atencio_persones_fragils_cronicitat_complexa_avancada_2020_ang.pdf?sequence=8&isAllowed=y) [accessed 2023-05-19]
18. Monterde D, Vela E, Clèries M, grupo colaborativo GMA. Adjusted morbidity groups: A new multiple morbidity measurement of use in Primary Care. *Aten Primaria.* Dec 2016;48(10):674-682. [FREE Full text] [doi: [10.1016/j.aprim.2016.06.003](https://doi.org/10.1016/j.aprim.2016.06.003)] [Medline: [27495004](https://pubmed.ncbi.nlm.nih.gov/27495004/)]
19. Anker SD, Gillespie IA, Eckardt K, Kronenberg F, Richards S, Druke TB, et al. On behalf the ARO Steering Committee (collaborators). Development and validation of cardiovascular risk scores for haemodialysis patients. *Int J Cardiol.* Aug 01, 2016;216:68-77. [doi: [10.1016/j.ijcard.2016.04.151](https://doi.org/10.1016/j.ijcard.2016.04.151)] [Medline: [27140339](https://pubmed.ncbi.nlm.nih.gov/27140339/)]
20. Mezzatesta S, Torino C, Meo PD, Fiumara G, Vilasi A. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. *Comput Methods Programs Biomed.* Aug 2019;177:9-15. [doi: [10.1016/j.cmpb.2019.05.005](https://doi.org/10.1016/j.cmpb.2019.05.005)] [Medline: [31319965](https://pubmed.ncbi.nlm.nih.gov/31319965/)]
21. Kearney MT, Fox KA, Lee AJ, Brooksby WP, Shah AM, Flapan A, et al. Predicting sudden death in patients with mild to moderate chronic heart failure. *Heart.* Oct 01, 2004;90(10):1137-1143. [FREE Full text] [doi: [10.1136/hrt.2003.021733](https://doi.org/10.1136/hrt.2003.021733)] [Medline: [15367507](https://pubmed.ncbi.nlm.nih.gov/15367507/)]
22. Kearney MT, Fox KA, Lee AJ, Prescott RJ, Shah AM, Batin PD, et al. Predicting death due to progressive heart failure in patients with mild-to-moderate chronic heart failure. *J Am Coll Cardiol.* Nov 20, 2002;40(10):1801-1808. [FREE Full text] [doi: [10.1016/s0735-1097\(02\)02490-7](https://doi.org/10.1016/s0735-1097(02)02490-7)] [Medline: [12446064](https://pubmed.ncbi.nlm.nih.gov/12446064/)]
23. Candido Dos Reis FJ, Wishart GC, Dicks EM, Greenberg D, Rashbass J, Schmidt MK, et al. An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. *Breast Cancer Res.* May 22, 2017;19(1):58. [FREE Full text] [doi: [10.1186/s13058-017-0852-3](https://doi.org/10.1186/s13058-017-0852-3)] [Medline: [28532503](https://pubmed.ncbi.nlm.nih.gov/28532503/)]
24. Raj R, Sanjay S, Kusuma M, Sampath S. Comparison of support vector machine and naïve Bayes classifiers for predicting diabetes. Presented at: 2019 1st International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE); March 19-20, 2019;41-45; Bangalore, India. [doi: [10.1109/ICATIECE45860.2019.9063792](https://doi.org/10.1109/ICATIECE45860.2019.9063792)]
25. Barakat NH, Bradley AP, Barakat MNH. Intelligible support vector machines for diagnosis of diabetes mellitus. *IEEE Trans Inf Technol Biomed.* Jul 01, 2010;14(4):1114-1120. [doi: [10.1109/TITB.2009.2039485](https://doi.org/10.1109/TITB.2009.2039485)] [Medline: [20071261](https://pubmed.ncbi.nlm.nih.gov/20071261/)]
26. Iseki K, Kawazoe N, Fukiyama K. Serum albumin is a strong predictor of death in chronic dialysis patients. *Kidney Int.* Jul 1993;44(1):115-119. [FREE Full text] [doi: [10.1038/ki.1993.220](https://doi.org/10.1038/ki.1993.220)] [Medline: [8355451](https://pubmed.ncbi.nlm.nih.gov/8355451/)]
27. Li X, Ge P, Zhu J, Li H, Graham J, Singer A, et al. Deep learning prediction of likelihood of ICU admission and mortality in COVID-19 patients using clinical variables. *PeerJ.* 2020;8:e10337. [FREE Full text] [doi: [10.7717/peerj.10337](https://doi.org/10.7717/peerj.10337)] [Medline: [33194455](https://pubmed.ncbi.nlm.nih.gov/33194455/)]
28. Reina Reina A, Barrera JM, Valdivieso B, Gas M, Maté A, Trujillo JC. Machine learning model from a Spanish cohort for prediction of SARS-COV-2 mortality risk and critical patients. *Sci Rep.* Apr 06, 2022;12(1) [doi: [10.1038/s41598-022-09613-y](https://doi.org/10.1038/s41598-022-09613-y)]
29. An C, Lim H, Kim D, Chang JH, Choi YJ, Kim SW. Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Sci Rep.* Oct 30, 2020;10(1) [FREE Full text] [doi: [10.1038/s41598-020-75767-2](https://doi.org/10.1038/s41598-020-75767-2)] [Medline: [33127965](https://pubmed.ncbi.nlm.nih.gov/33127965/)]
30. Zhao Z, Chen A, Hou W, Graham JM, Li H, Richman PS, et al. Prediction model and risk scores of ICU admission and mortality in COVID-19. *PLoS One.* 2020;15(7):e0236618. [FREE Full text] [doi: [10.1371/journal.pone.0236618](https://doi.org/10.1371/journal.pone.0236618)] [Medline: [32730358](https://pubmed.ncbi.nlm.nih.gov/32730358/)]
31. Yan L, Zhang H, Goncalves J, Xiao Y, Wang M, Guo Y, et al. An interpretable mortality prediction model for COVID-19 patients. *Nat Mach Intell.* May 14, 2020;2(5):283-288. [doi: [10.1038/s42256-020-0180-7](https://doi.org/10.1038/s42256-020-0180-7)]
32. Iseki K, Yamazato M, Tozawa M, Takishita S. Hypocholesterolemia is a significant predictor of death in a cohort of chronic hemodialysis patients. *Kidney Int.* May 2002;61(5):1887-1893. [FREE Full text] [doi: [10.1046/j.1523-1755.2002.00324.x](https://doi.org/10.1046/j.1523-1755.2002.00324.x)] [Medline: [11967041](https://pubmed.ncbi.nlm.nih.gov/11967041/)]
33. Generalitat de Catalunya. URL: <https://salutweb.gencat.cat/web/content/ambits-actuacio/Linies-dactuacio/Estrategies-de-salut/Cronicitat/Documentacio-cronicitat/arxiu/catalan-model-care-en.pdfhttps://pubmed.ncbi.nlm.nih.gov/27495004/> [accessed 2023-12-17]

## Abbreviations

**ACP:** patient with advanced chronic disease

**AI:** artificial intelligence  
**AUC:** area under the curve  
**AUROC:** area under the receiver operating characteristic  
**CCP:** patient with complex chronic disease  
**CV:** cross-validation  
**DL:** deep learning  
**GMA:** grupos de morbilidad ajustados  
**GRBoost:** Gradient Boosting  
**ICS:** Institut Català de la Salut  
**ICU:** intensive care unit  
**IDIAP:** Institut Universitari d'Investigació en Atenció Primària  
**KNN:** k-nearest-neighbor  
**ML:** machine learning  
**ProPCC:** Community-Based Integrated Care Program for People With Complex Chronic Conditions  
**RBF:** radial basis function  
**SVM:** support vector machine  
**XGBoost:** Extreme Gradient Boosting

*Edited by E Mensah; submitted 15.09.23; peer-reviewed by N Akpakpan, A Santos Mejías, JJ Beunza, L Carrasco-Ribelles; accepted 18.11.23; published 28.12.23*

*Please cite as:*

*Hernández Guillamet G, Morancho Pallaruelo AN, Miró Mezquita L, Miralles R, Mas MÀ, Ulldemolins Papaseit MJ, Estrada Cuxart O, López Seguí F*

*Machine Learning Model for Predicting Mortality Risk in Patients With Complex Chronic Conditions: Retrospective Analysis  
Online J Public Health Inform 2023;15:e52782*

*URL: <https://ojphi.jmir.org/2023/1/e52782>*

*doi: [10.2196/52782](https://doi.org/10.2196/52782)*

*PMID: [38223690](https://pubmed.ncbi.nlm.nih.gov/38223690/)*

©Guillem Hernández Guillamet, Ariadna Ning Morancho Pallaruelo, Laura Miró Mezquita, Ramón Miralles, Miquel Àngel Mas, María José Ulldemolins Papaseit, Oriol Estrada Cuxart, Francesc López Seguí. Originally published in the Online Journal of Public Health Informatics (<https://ojphi.jmir.org/>), 28.12.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Online Journal of Public Health Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://ojphi.jmir.org/>, as well as this copyright and license information must be included.