



Design of 1-year mortality forecast at hospital admission: A machine learning approach

Vicent Blanes-Selva 

Universitat Politècnica de València, Spain

Vicente Ruiz-García

Instituto de Investigación Sanitaria La Fe, Spain

Salvador Tortajada

Instituto de Investigación Sanitaria La Fe, Spain

Consejo Superior de Investigaciones Científicas, Spain

José-Miguel Benedí 

Universitat Politècnica de València, Spain

Bernardo Valdivieso

Instituto de Investigación Sanitaria La Fe, Spain

Juan M García-Gómez

Universitat Politècnica de València, Spain

Abstract

Palliative care is referred to a set of programs for patients that suffer life-limiting illnesses. These programs aim to maximize the quality of life (QoL) for the last stage of life. They are currently based on clinical evaluation of the risk of 1-year mortality. The main aim of this work is to develop and validate machine-learning-based models to predict the exitus of a patient within the next year using data gathered at hospital

Corresponding author:

Vicent Blanes-Selva, BDSLab, Instituto de Tecnologías de la Información y Comunicaciones (ITACA), Universitat Politècnica de València, Camino de Vera s/n, 8G Building, Itaca Institute, Valencia 46022, Spain.

Email: viblasel@upv.es



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which

permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

admission. Five machine-learning techniques were applied using a retrospective dataset. The evaluation was performed with five metrics computed by a resampling strategy: Accuracy, the area under the ROC curve, Specificity, Sensitivity, and the Balanced Error Rate. All models reported an AUC ROC from 0.857 to 0.91. Specifically, Gradient Boosting Classifier was the best model, producing an AUC ROC of 0.91, a sensitivity of 0.858, a specificity of 0.808, and a BER of 0.1687. Information from standard procedures at hospital admission combined with machine learning techniques produced models with competitive discriminative power. Our models reach the best results reported in the state of the art. These results demonstrate that they can be used as an accurate data-driven palliative care criteria inclusion.

Keywords

machine learning, palliative care, hospital admission data, mortality forecast

Background and significance

An increasing number of people have multiple morbidities and conditions in the final moments of their lives, current medicine tries to maintain a quality of life of these people, including their needs in the final moments. In this situation, palliative care tries to facilitate the life of people in these conditions from a patient perspective.

Palliative care is a multidisciplinary care that aims to grant comfort to the patient, avoid painful and/or aggressive treatments, alleviate pain, other symptoms, psychological and spiritual distress.¹ In addition, there are some studies which prove that patients receiving early palliative care present a better quality of life, mood, satisfaction with the treatment²⁻⁴ and even a longer survival when compared to patients whose palliative care was delayed.⁵

A criterion for the palliative care inclusion is desirable as early as possible. An adverse event such as a hospital admission could be considered a convenient episode to check this criterion. Nowadays, the main indicator to include a patient in palliative care is the clinical criterion of a potential exitus within the next 12 months. An example of that is the surprise question described in Moss et al.⁶

Mortality forecast has been previously studied by other groups. Buurman et al.⁷ proposed a method for predicting 90-day mortality risk using few clinical features: Barthel test index, Charlson score, and Malignancy and Urea nitrogen (mmol/L). The authors of this study calculated how modifications on the features affect the outcome. The study reported AUC ROC=0.77 (CI 95%, 0.72–0.82). Bernabeu-Wittel et al.⁸ proposed a method for detecting 1-year mortality for polypathological patients. That model computed the PROFUND score, based on some features to assign a mortality risk to the patient: Age, Hemoglobin, and Barthel index, No caregiver or caregiver other than the spouse, hospital admissions ≥ 4 in last 12 months and positive for few diseases. The PROFUND score is mapped into mortality (in less than a year) probability. The reported validation result was AUC ROC=0.7 (CI 95%, 0.67–0.74). Van Walraven et al.⁹ reported a 1-year mortality forecast model based on patient demographics, health burden, and severity of acute illness. The model uses a binomial logistic regression. The AUC ROC ranged from 0.89 (CI 95%, 0.87–0.91) to 0.92 (CI 95%, 0.91–0.92). Recently, Avati et al.¹⁰ presented a deep neural network for 1-year mortality prediction by using 13,654 features, corresponding to the different ICD9 codes in different time windows through the year. They reported and 0.93 of AUC ROC for all validation patients but only 0.87 for admitted patients. A summary of previous studies is provided in Table 1.

Based on the promising results in the literature we have addressed the design of a high-performance predictive model of 1-year mortality exclusively based on observations at hospital admission. The overall aim of our study was to provide quantitative methods to healthcare caregivers to

Table 1. Summary of previous mortality studies.

Approach	No. of features	AUC ROC	Reference
Buurman	4	0.77 (CI 95%, 0.72–0.82)	Buurman et al. ⁷
PROFUND	9	0.70 (CI 95%, 0.67–0.74)	Bernabeu-Wittel et al. ⁸
HOMR	11	0.89 (CI 95%, 0.87–0.91) to 0.92 (CI 95%, 0.91–0.92)	Van Walraven et al. ⁹
Deep learning	13,654	0.93 (all); 0.87 (admitted patients)	Avati et al. ¹⁰

decide the inclusion of patients in the palliative care program during the hospital admission. To this aim, we have designed and evaluated five predictive models from the state-of-the-art machine learning discipline.

These models are meant to be more complex, in terms of algorithm, parameters and amount of data needed than the first studies we presented but with less requirements than Avati's deep learning approximation, being the most adequate option for our dataset size.

The models presented in this work are continued in the InAdvance project (<http://www.inadvanceproject.eu/>) along with other kinds of models such as frailty and resource consumption models to create a complete Computerized Decision Support System (CDSS)¹¹ for the palliative care inclusion decision. This CDSS is meant to join other information systems created to improve the palliative care process, such as O'Connor et al.¹² and Dy et al.¹³

Materials

The data of the study was extracted from the Electronic Health Records from Hospital La Fe. We gathered all the hospitalization episodes of adult patients (≥ 18 years old), excluding those related to mental health, gynecology and obstetrics, from January 2014 to December 2017 (a total number of 114,393 cases) that have been discharged from the hospital. All the patients received standard care, so no effects like a prolonged survival from palliative care⁵ affect the data. To guarantee independent observations, we selected a random single episode for each patient, reaching a total of 65,279 episodes.

The dataset contains information about the previous and current admission (seven features), laboratory test results (seven features) and a list of 28 selected diseases for which the patient is positive or negative. Sex, age, Charlson index, and Barthel tests result are also available. This adds up a total of 36 features which can be obtained straightforwardly in the first hours of admission. Some of these features were used, with positive results, in previous studies.

Target variable was exitus after 1 year from the admission date. The number of patients that have died in less than a year (positive cases) was 8113 (~12.43%), the number of negative cases is 57,166 (~87.57%). The whole variable description can be seen in Table 2. The distributions for Admission Destination, Service, and Admission Cause are represented in Figures 1–3 respectively.

Methods

Development of the models

Five machine learning techniques were employed for developing our predictive models: Gradient Boosting Classifier,¹⁴ Random Forest,¹⁵ K-Nearest Neighbors,¹⁶ Multilayer Perceptron (MLP),¹⁷ and Support Vector Machine.¹⁸ The implementation of the scikit-learn toolkit¹⁹ was employed in

Table 2. Features information.

Variable	Types	Missings	Distribution
Sex	CAT	0	Males: 51.86%
Age	INT	0	61.327 ± 18.375
Urgent admission	BOOL	0	Pos: 56.97%
Admission destination	CAT	0	—
Service	CAT	0	—
Admission cause	CAT	0	—
Prev. stays	INT	0	6.119 ± 9.502
Barthel test	INT	56,214	67.268 ± 37.919
Prev. admissions	INT	0	0.300 ± 0.789
Prev. emergency room	INT	0	0.935 ± 1.691
Charlson score	INT	0	4.233 ± 3.238
Albumin (g/dL)	REAL	46,857	2.955 ± 0.677
Creatinine (mg/dL)	REAL	16,920	0.505 ± 1.063
Hemoglobin (g/dL)	REAL	14,434	11.703 ± 2.228
Leucocytes (Cel/mL)	REAL	14,434	9.457 ± 7.389
PCR (mg/L)	REAL	30,285	63.083 ± 84.481
Sodium (mEq/L)	REAL	17,183	139.672 ± 4.354
Urea (mg/dL)	REAL	18,459	46.255 ± 34.628
Acute myocardial infarction	BOOL	0	Pos: 3.09%
Congestive heart failure	BOOL	0	Pos: 6.14%
Peripheral vascular disease	BOOL	0	Pos: 4.88%
Cerebrovascular disease	BOOL	0	Pos: 6.76%
Dementia	BOOL	0	Pos: 1.5%
Chronic pulmonary disease	BOOL	0	Pos: 10.03%
Rheumatic disease	BOOL	0	Pos: 1.6%
Peptic ulcer disease	BOOL	0	Pos: 1.57%
Mild liver disease	BOOL	0	Pos: 5.77%
Diabetes without complications	BOOL	0	Pos: 13.19%
Diabetes with complications	BOOL	0	Pos: 1.27%
Hemiplegia paraplegia	BOOL	0	Pos: 1.28%
Renal disease	BOOL	0	Pos: 7.46%
Malignancy	BOOL	0	Pos: 18.2%
Moderate severe liver disease	BOOL	0	Pos: 1.49%
Metastasis	BOOL	0	Pos: 3.27%
AIDS	BOOL	0	Pos: 0.57%
Delirium	BOOL	0	Pos: 0.12%

all of them except in the MLP which uses Keras and TensorFlow.²⁰ Moreover, the optimization tool TPOT²¹ was used in order to find a good model to fit the data.

Feature importance

We studied the relevance of each feature for the final prediction by calculating the GINI importance provided by the Gradient Boosting Classifier. The GINI importance measures the average gain of purity by splits of a given variable. If the variable is discriminant for the problem, it tends to split mixed labeled nodes into pure single class nodes.²²

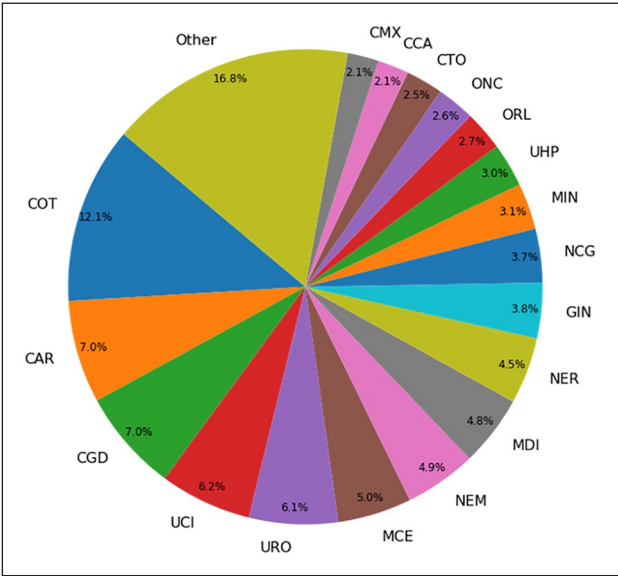


Figure 1. Distribution for admission destination.

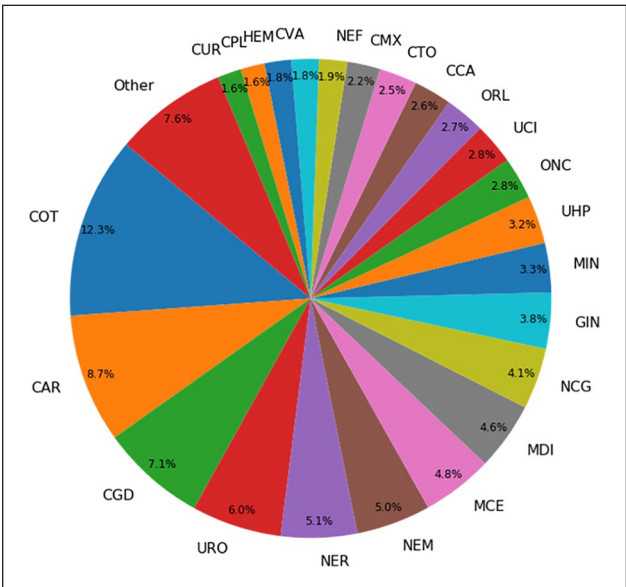


Figure 2. Distribution for service.

Validation of the state-of-the-art models

As a first step, we have compared our model with the PROFUND and Buurman’s model using the same evaluation method. For the Buurman’s model, a clinical committee led by Vicente Ruiz-García at Hospital La Fe adapted the Buurman’s proposal as 1-year mortality index, using a linear regression with the 1-year mortality target variable. Besides, we evaluated the original PROFUND

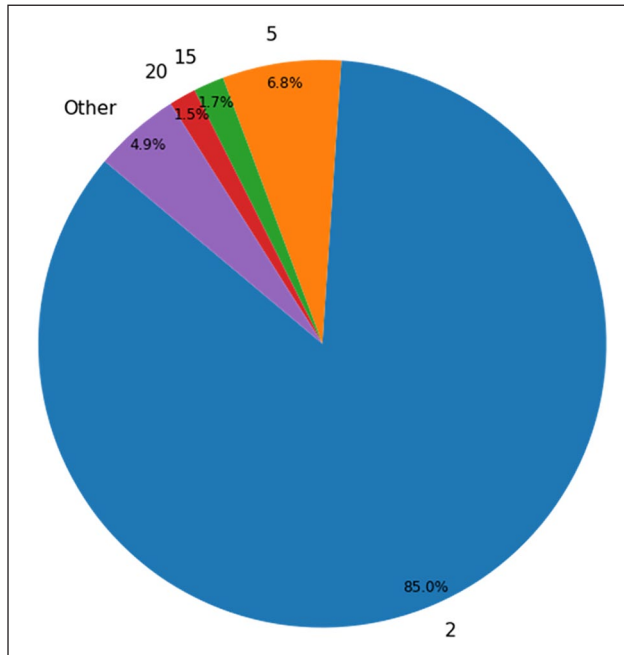


Figure 3. Distribution for admission cause.

model proposed in Bernabeu-Wittel et al.⁸ The validation of the other models in state of the art was not possible due to the lack of part of their features in our data system.

Evaluation of the models

First, we have computed the ROC Curve²³ for each model and calculated the optimum probability threshold (minimum probability to assign the positive class to a sample) running our models using a random split of the data from separating train and test. We iterated over all the different values that could change the specificity and the sensitivity of the model and kept the threshold that minimizes the balanced error rate (BER).²³

Once the threshold is established for each model, we internally validated them using a 100-repetition stratified hold-out (80% of the data in order to train the model and 20% for test it). The missing values have been imputed using an iterative approach that models each feature with missing values as a function of other features, available in scikit-learn as *IterativeImputer*.²⁴ Five metrics have been stored for each experiment (accuracy, AUC ROC, specificity, sensitivity, and balanced error rate).^{22,23} For each metric, the mean and the 95% confidence interval have been computed.

In addition to this process, we performed a second round of experimentations, where we trained the models using a balanced set of data by oversampling the positive class using the SMOTE²⁵ technique. This will help us to determine which are the effect of class imbalance in the ML algorithms for this problem.

Results

Following the above methodology, this section analyses the results of the proposed machine learning models, including a validation from Buurman's modified model and PROFUND from the

literature. In terms of AUC ROC, the model based on Gradient Boosting Classifier achieves the higher score with 0.91, being followed by the Random Forest and the Multi-Layer Perceptron. Table 3 contains means and confidence interval at 95% for all proposed models and selected metrics. Table 4 provides the same experiments as staged before but using SMOTE as an oversampling technique for the positive class. Once again, the Gradient Boosting Classifier is the model with higher AUC ROC (0.908), followed by the Random Forest. Results using SMOTE are slightly worse on the selected metrics. As for the variable importance, Service, Urea and Leucocytes obtained the higher scores (10.6%, 10.32%, 8.65%), followed by PCR (7.65%) and Age (6.99%). Table 5 lists every variable and its Gini Importance.

Discussion

The prediction of death before 1 year could be a relevant criterion to admit the patients into palliative care programs.⁶ Also, the prediction of the death at admission of the patient would help the hospital management to better manage its resources in a more accurate way.

We used the area under the ROC curve as the comparison metric because is the common metric to all other works in the SoA. We also chose the threshold for considering a sample into the positive class taking the value that minimizes the balanced error rate.

The Buurman's modified model and the PROFUND index have been validated, the models described in our work outperform them in terms of AUC ROC, sensitivity and specificity. Whereas our models presented a bigger number of features (36) than the mentioned articles (four for Buurman's modified model and nine for PROFUND index).

Comparing with the most recent work, Avati et al.¹⁰ that presented a neural network with 18 hidden layers of 512 neurons each was trained with 177,011 patients. The models in our approach are trained with 52,223 patients. The network used 13,654 features as input, our model uses only 36. Finally, they achieved an AUC ROC of 0.93 for all their patients but it only achieved 0.87 when only admitted patients are considered. We achieved better results using a significative smaller amount features, this led to a more compact model that also is more interpretable since the best performing model is based on decision trees.

The results obtained using the oversampling technique in the training set are very similar to the ones obtained in the original experiment. The main differences are in the optimal thresholds, which increased in all cases due to the change in proportion in the classes.

The best results in our models achieved the interval reported by van Walraven et al.⁹: 0.89–0.92 AUC ROC. Despite the number of final features is smaller in HOMR (10) two of their features are composed: “charlson comorbidity index score” (15 items) and “diagnostic risk score” (70 items), so at the end, HOMR requires more information about the patients than our models. The performance comparison with Avati et al.¹⁰ and HOMR have been made using their reported results which implies the use of different evaluations and datasets.

We obtained consistent results compared to other studies. In HOMR the features that are capable to add more points to the index are the admitting service (up to 28) and the “age × comorbidity” (other 28 points). We agree with the most important variable (real service code) and the fifth one in importance order (age). Our second most important variable, nitrogen in urea, is included among the Buurman's model. Moreover, creatinine in blood is related to BUN and is a variable also associated with mortality is our results.

It is known from the scientific literature cited above that morbidities, together with age and functionality (in our case measured by the Charlson's index) are the main predictors of mortality. However, our study on the importance of the variables shows that morbidities such as COPD score low, which is consistent with the van Walraven et al.⁹ rule, which gives it only 2 points out of 35. Something similar happens with dementia, which scores only 3 points.

Table 3. Complete results for the machine learning proposed models.

Model	Threshold	Accuracy	AUC ROC	Specificity	Sensitivity	BER
Gradient boosting classifier	0.1	0.814 [0.812, 0.816]	0.91 [0.909, 0.912]	0.808 [0.806, 0.81]	0.858 [0.855, 0.861]	0.167 [0.165, 0.169]
Random forest	0.145	0.802 [0.801, 0.804]	0.9 [0.899, 0.902]	0.795 [0.793, 0.797]	0.851 [0.848, 0.855]	0.177 [0.175, 0.178]
K-nearest neighbors	0.08	0.755 [0.752, 0.757]	0.87 [0.868, 0.872]	0.742 [0.739, 0.745]	0.844 [0.839, 0.849]	0.207 [0.205, 0.209]
Support vector machine	0.105	0.686 [0.613, 0.76]	0.857 [0.855, 0.859]	0.664 [0.576, 0.753]	0.842 [0.811, 0.872]	0.247 [0.217, 0.277]
Multi-layer perceptron	0.15	0.797 [0.788, 0.806]	0.9 [0.897, 0.903]	0.789 [0.778, 0.8]	0.854 [0.845, 0.862]	0.179 [0.175, 0.182]
Baseline: Buurman's modification	0.101	0.794 [0.766, 0.821]	0.681 [0.680, 0.683]	0.838 [0.804, 0.873]	0.483 [0.460, 0.505]	0.340 [0.333, 0.346]
Baseline: PROFUND	0.058	0.569 [0.534, 0.603]	0.739 [0.738, 0.740]	0.534 [0.493, 0.575]	0.814 [0.800, 0.828]	0.326 [0.313, 0.339]

Table 4. Complete results using oversampling strategy for positive class.

Model	Threshold	Accuracy	AUC ROC	Specificity	Sensitivity	BER
Gradient boosting classifier	0.142	0.813 [0.811, 0.815]	0.908 [0.906, 0.909]	0.808 [0.806, 0.81]	0.849 [0.845, 0.853]	0.171 [0.169, 0.174]
Random forest	0.225	0.812 [0.811, 0.814]	0.895 [0.894, 0.896]	0.81 [0.808, 0.812]	0.832 [0.828, 0.836]	0.179 [0.177, 0.181]
K-nearest neighbors	0.39	0.765 [0.763, 0.767]	0.87 [0.869, 0.872]	0.756 [0.753, 0.759]	0.828 [0.824, 0.833]	0.208 [0.206, 0.21]
Support vector machine	0.462	0.772 [0.77, 0.774]	0.869 [0.867, 0.871]	0.765 [0.763, 0.767]	0.819 [0.815, 0.824]	0.208 [0.205, 0.21]
Multi-layer perceptron	0.42	0.776 [0.747, 0.805]	0.867 [0.855, 0.879]	0.772 [0.737, 0.807]	0.805 [0.785, 0.825]	0.212 [0.199, 0.224]
Baseline: Buurman's modification	0.429	0.486 [0.351, 0.621]	0.688 [0.686, 0.691]	0.45 [0.28, 0.62]	0.737 [0.626, 0.847]	0.407 [0.377, 0.436]
Baseline: PROFUND	0.337	0.556 [0.5, 0.611]	0.765 [0.762, 0.769]	0.516 [0.449, 0.583]	0.835 [0.81, 0.86]	0.324 [0.303, 0.346]

Table 5. Variable importance provided by GBC sorted by decreasing importance.

Variable	Importance (%)	Variable	Importance (%)
Service	10.60	Malignancy	0.90
Urea	10.32	Sex	0.68
Leucocytes	8.65	Congestive heart failure	0.51
PCR	7.88	Renal disease	0.42
Age	7.65	Dementia	0.40
Creatinine	6.99	Chronic pulmonary disease	0.35
Albumin	6.66	Diabetes without complications	0.34
Prev. stays	5.83	Acute myocardial infarction	0.30
Hemoglobin	5.67	Moderate severe liver disease	0.26
Sodium	5.07	Cerebrovascular disease	0.25
Charlson score	4.55	Mild liver disease	0.24
Admission destination	3.65	Peripheral vascular disease	0.22
Barthel test	3.20	Rheumatic disease	0.21
Prev. Emergency room	1.98	Hemiplegia paraplegia	0.20
Cause of admission	1.71	Peptic ulcer disease	0.20
Prev. admissions	1.70	Diabetes with complications	0.17
Urgent admission	1.10	Delirium	0.16
Metastasis	0.95	AIDS	0.05

The value of Urea and Creatinine indicate renal failure measured in two different ways, and we also know that these are predictors in other mortality prediction rules such as Buurman's et al.⁷ rule, where they appear as predictive with once again the functional situation.

Our models confirm that the disease variables weight the decision process, but they are less important when no other disease is present. The models also guide us to consider that administrative variables such as previous urgent admissions, previous stays and destination of admission may have an additional weight not considered until now that contribute more to mortality than the fact of suffering a specific disease. Having this information updated permanently will allow for the adjustment of additional risks for patients who are acutely admitted to hospital through the Emergency Services.

We compared the method we have used to assigning importance to the different variables in the dataset, the GINI importance provided by the Gradient Boosting Classifier, to other methods based on features ablation. We used a technique called "Feature ranking with recursive feature elimination" (RFE) from scikit-learn.¹⁹ The RFE starts training a model and obtains the variable importance. The least important variables are removed from the dataset, and the process continues recursively until a minimal set of variables is obtained. The results obtained were almost identical to the ones presented in this work, so we conclude that, in addition to its reputation in other areas such as machine learning interpretability, GINI importance is a robust indicator to feature's significance in the model.

The clinical features included in our work have clinical relevance and appear in other clinical prediction rules. They appears in the records of our hospital databases in Spain and allow the creation of alerts for the clinicians to address patients, to palliative care programs not only for advanced oncology patients but for other chronic pathologies as dementia (a critical literature review exploring the challenges of delivering effective palliative care to older people with dementia, cardiac failure, or chronic obstructive pulmonary disease (COPD) or end-stage renal disease (Comparing the Palliative Care Needs of Those With Cancer to Those With Common Non-Cancer Serious Illness)).²⁶⁻²⁸

This study has caused a direct impact on Hospital La Fe since the model based on the Gradient Boosting Classifier has been implemented in the pre-production information systems and it is on a test stage. Once in the day, a program gathers all the admitted patients' data and extracts the features, this information is passed to the model who gives a posteriori probability and a label prediction, this information is stored on a separated table of the same database including the timestamp.

The main limitation of the study was the use of data from only one hospital, we can't ensure that the models learned with the study population are effective with patients of another country/region, or another type of hospital, Hospital La Fe is a tertiary Hospital a referral in the Valencia region, with different patients and severity.

In addition, the models had only an internal derivation, so we need to refine and validate this model to reproduce the findings with different settings (smaller hospital and with less severity illness) perhaps outside the same city or Valencian community, where we can have a population with different habits such diet or lifestyle. It is necessary to work on additional criteria for palliative care admission besides mortality, for example, introducing the available resources in the decision-making process. Also, an inclusion criterion for chronic patients is needed since their illness trajectories are different from other patients.²⁹

The work presented in this article is being continued in the InAdvance project, a European project about palliative care which aims to research and standardize some palliative care procedures as well as other aspects such as admission criteria. Our work in this project is to develop predictive models for mortality, frailty and resource consumption which integrated into Clinical Decision Support System (CDSS) could be proven as reliable and objective inclusion criteria for palliative care. This palliative care criterion is intended to be unbiased and independent from the place the care takes place, so other institutions such as hospices or nursing homes can take advantage of them with the only restriction of having access to the variables the model needs to run.

Conclusion

This work proposes machine-learning forecast of 1-year exitus using data from hospital admission. Our forecast achieved an area under ROC curve of 0.9 and a BER of 0.17, being the Gradient Boosting Classifier the best model. The features used in the models correspond to basic demographic and administrative information, some laboratory results and a list of positives or negatives for certain diseases. The presented models could have an instant impact on every hospital, only the feature extraction module and the table for results need to be adapted to the particular information system of every hospital, the rest of the components are ready to set in production. Our results have reached the best results in the state-of-the-art, corresponding to the HOMR index which validation in few Canadian hospitals produces AUC ROC from 0.89 to 0.92.

Acknowledgements

Special thanks to the BDSLab researchers for their support.

Clinical relevance statement

The research showed that is possible to predict which patient will have a high risk of death before 1 year after hospital admission. These predictions were for a wide range of population, not only for the suspected patients with a short-life expectancy patients, palliative care must ensure the best quality of life at the end of life's patients.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the InAdvance project (H2020-SC1-BHC-2018-2020 No. 825750).

Human subjects' protections

This work did not imply any risk nor alteration in the habitual health care provided for the subjects. Only the authorized personnel could access the electronic records that were properly anonymized using dissociative and untraceable codes. This investigation followed the International Guideline for Ethical Review of Epidemiological Studies³⁰ and the ethic committee from IIS LA FE that approved the study's protocol.

Summary table

What was already known on the topic

- Different models for 1-year mortality prediction, mostly based on scores
- Patients enrolled in early palliative care improve their quality of life compared with the ones receiving late palliative care or standard care.

What this study added to our knowledge

- One-year forecast at hospital admission using Gradient Boosting Classifier and Random Forest reported the best results in state of the art, achieving performance feasible for clinical use.
- Only features gathered at the first hours of hospital admission may be enough as criteria for palliative care inclusion.
- Hospital service, laboratory analysis, and age are the essential features for 1-year forecast, confirming previous studies and allowing us to create compact predictive models

ORCID iDs

Vicent Blanes-Selva  <https://orcid.org/0000-0002-0056-0329>

José-Miguel Benedí  <https://orcid.org/0000-0001-6516-2746>

References

1. Kelley AS and Morrison RS. Palliative care for the seriously ill. *N Engl J Med* 2015; 373(8): 747–755.
2. Bakitas M, Lyons KD, Hegel MT, et al. Effects of a palliative care intervention on clinical outcomes in patients with advanced cancer: the Project ENABLE II randomized controlled trial. *JAMA* 2009; 302(7): 741–749.
3. Zimmermann C, Swami N, Krzyzanowska M, et al. Early palliative care for patients with advanced cancer: a cluster-randomised controlled trial. *Lancet* 2014; 383(9930): 1721–1730.
4. Temel JS, Greer JA, Muzikansky A, et al. Early palliative care for patients with metastatic non-small-cell lung cancer. *N Engl J Med* 2010; 363(8): 733–742.
5. Bakitas MA, Tosteson TD, Li Z, et al. Early versus delayed initiation of concurrent palliative oncology care: patient outcomes in the ENABLE III randomized controlled trial. *J Clin Oncol* 2015; 33(13): 1438–1445.

6. Moss AH, Ganjoo J, Sharma S, et al. Utility of the “surprise” question to identify dialysis patients with high mortality. *Clin J Am Soc Nephrol* 2008; 3(5): 1379–1384.
7. Buurman BM, Van Munster BC, Korevaar JC, et al. Prognostication in acutely admitted older patients by nurses and physicians. *J Gen Intern Med* 2008; 23(11): 1883–1889.
8. Bernabeu-Wittel M, Ollero-Baturone M, Moreno-Gaviño L, et al. Development of a new predictive model for polypathological patients. The PROFUND index. *Eur J Intern Med* 2011; 22(3): 311–317.
9. van Walraven C, McAlister FA, Bakal JA, et al. External validation of the Hospital-patient One-year Mortality Risk (HOMR) model for predicting death within 1 year after hospital admission. *CMAJ* 2015; 187(10): 725–733.
10. Avati A, Jung K, Harman S, et al. Improving palliative care with deep learning. *BMC Med Inform Decis Mak* 2018; 18(4): 122.
11. Mazzaglia G, Piccinni C, Filippi A, et al. Effects of a computerized decision support system in improving pharmacological management in high-risk cardiovascular patients: a cluster-randomized open-label controlled trial. *Health Inform J* 2016; 22(2): 232–247.
12. O’Connor M, Erwin T and Dawson L. A means to an end: a web-based client management system in palliative care. *Health Inform J* 2009; 15(1): 41–54.
13. Dy SM, Roy J, Ott GE, et al. Tell Us™: a Web-based tool for improving communication among patients, families, and providers in hospice and palliative care through systematic data specification, collection, and use. *J Pain Symptom Manage* 2011; 42(4): 526–534.
14. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001; 29(5): 1189–1232.
15. Breiman L. Random forests. *Mach Learn* 2001; 45(1): 5–32.
16. Cover T and Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory* 1967; 13(1): 21–27.
17. Hinton GE. Connectionist learning procedures. In: *Machine learning, volume III*. San Mateo, CA: Morgan Kaufmann, 1990, pp. 555–610.
18. Guyon I, Boser B and Vapnik V. Automatic capacity tuning of very large VC-dimension classifiers. In: *Advances in neural information processing systems*, Denver, Colorado, 1993, pp. 147–155.
19. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011; 12: 2825–2830.
20. Abadi M, Barham P, Chen J, et al. TensorFlow: a system for large-scale machine learning. In: *OSDI*, Savannah, GA, November 2016, vol. 16, pp. 265–283.
21. Olson RS, Urbanowicz RJ, Andrews PC, et al. Automating biomedical data science through tree-based pipeline optimization. In: *European conference on the applications of evolutionary computation*, Porto, Portugal, 30 March–1 April 2016, pp. 123–137. Cham: Springer.
22. Breiman L. *Classification and regression trees*. United States: CRC Press, 2017.
23. Guyon I, Gunn S, Nikravesh M, et al. (eds). *Feature extraction: foundations and applications*. vol. 207. Germany: Springer, 2008.
24. Buuren SV and Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw* 2010; 45(3): 1–68.
25. Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002; 16: 321–357.
26. Klindedinst R, Kornfield ZN and Hadler RA. Palliative care for patients with advanced heart disease. *J Cardiothorac Vasc Anesth* 2018; 33(3): 833–843.
27. Bostwick D, Wolf S, Samsa G, et al. Comparing the palliative care needs of those with cancer to those with common non-cancer serious illness. *J Pain Symptom Manage* 2017; 53(6): 1079–1084.
28. Birch D and Draper J. A critical literature review exploring the challenges of delivering effective palliative care to older people with dementia. *J Clin Nurs* 2008; 17(9): 1144–1163.
29. Murray SA, Kendall M, Boyd K, et al. Illness trajectories and palliative care. *BMJ* 2005; 330(7498): 1007–1011.
30. Rose S. International ethical guidelines for epidemiological studies: by the Council for International Organizations of Medical Sciences (CIOMS). *Am J Epidemiol* 2009; 170(11): 1451–1452.