Predicting Length of Stay in Geriatric Patients Using an Ensemble Learning Method

Mariachiara Savino[1], Chiara Dachena[2](✉), Carlotta Masciocchi[2], Stefania Orini[3], Roberto Gatta[3], Stefano Patarnello[2], Riccardo Rinaldi[2], Nicola Acampora[4], Eleonora Meloni[4], Giovanni Arcuri[4], Francesco Landi[4], Graziano Onder[4], and Christian Barillaro[4]

[1] Diagnostica per immagini, Radioterapia Oncologica ed Ematologia, Università Cattolica del Sacro Cuore, Rome, Italy
[2] Real World Data Facility, Gemelli Generator, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
chiara.dachena@policlinicogemelli.it
[3] Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy
[4] Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

Abstract. Prolonged hospital stays that exceed the expected duration pose a challenge to patient care and utilization of hospital resources. Leveraging routine clinical data and knowledge about the intensity of medical care, a comprehensive predictive model was developed and validated to assess the risk of prolonged hospitalization. The model, adjusted for patients aged 50 years and older, integrates classification and regression models, applied from the fourth to the fourteenth day of hospitalization to account for updated patient data. Separate predictive models were trained for homogeneous subgroups of patients based on emergency department diagnosis. Moreover, a dashboard was developed to support physicians in clinical practice, facilitating the prompt identification of patients at high risk of delayed discharge. This approach addresses the pressing need to optimize resource allocation and mitigate the risks associated with prolonged hospitalization, particularly among elderly patients.

Keywords: Bed Blocking · Ensemble Predictive Model · Intensity of Care · Length of Stay · Machine Learning

1 Introduction

Prolonged hospital stays in older patients are linked to higher risks of complications, such as infections, cognitive decline, and loss of autonomy. Ackroyd-

M. Savino and C. Dachena—Contributed equally to this work and share first authorship.

R. Bellazzi et al. (Eds.): AIME 2025, LNAI 15735, pp. 356–360, 2025.
https://doi.org/10.1007/978-3-031-95841-0_66

Stolarz et al. [1] found that longer stays in the emergency room (ER) significantly increase the likelihood of adverse events during hospitalization. Prolonged stays often result from treatment complications or delayed discharge, a situation known as bed blocking [4], which strains hospital resources and increases wait times. Identifying patients at risk of prolonged stays early in their admission is essential for improving care and resource use. Prior studies, including one by Curiati et al. [2] showed that machine learning models using basic ER data can help predict extended stays. These studies agree on two points: early detection supports better resource planning, and initial visit data can aid in identifying high-risk patients. However, further research is needed to pinpoint the most predictive variables and to develop tools for clinical decision-making. In our study, we developed an ensemble model to estimate prolonged hospitalization risk for patients aged 50 years or older admitted through the ER at Fondazione Policlinico Universitario A. Gemelli IRCCS. Each model was tailored to a specific hospitalization day starting from day 4, using updated clinical and care-intensity data, such as exam and consultation frequency, to improve accuracy. By including patients 50 and older and validating on a large, independent cohort within a short recruitment window, we ensured model robustness and minimized bias from procedural shifts. In this work we don't propose a direct comparison with other models given the difference in the proposed ones and the population analyzed.

2 Methods

2.1 Study Design, Data Collection and Data Preprocessing

This study aimed to develop a predictive model for early identification of patients aged 50 or longer at risk of prolonged hospital stay after ER admission to Fondazione Policlinico Universitario A. Gemelli IRCCS (Jan 2022-Jun 2023). Patients were excluded if they were under 50, had COVID-19, were admitted to ICU/surgical units, or had a hospital stay under 4 d. The 2022 cohort was used for model training, and the first half of 2023 as the independent test set. Clinical and administrative data were extracted using the Gemelli Generator Real World Data (G2 RWD) facility [3]. A retrospective datamart included demographics, comorbidities, nursing scales, tracheostomy, bedsores, and care intensity, measured as the frequency of radiological exams, lab tests, and medical consultations.

Patients were grouped into 9 diagnostic categories based on ER diagnosis, and prolonged stay was defined as a length of stay exceeding the 4th quartile within each group. Missing nursing scale values were imputed by forward-filling the last recorded measurement. Care intensity data were aggregated by exam type (e.g., MRI, CT), medical specialty, and clinical focus (e.g., renal, infectious). In this case, intensity of care is used as an aggregative feature to reduce the number of possible model variables. A preprocessing pipeline generated daily datasets from day 4 to day 14 of hospitalization, incorporating newly available patient data. Separate models were trained per diagnosis group and day to improve predictive granularity.

day	Q1Q2/Q1	degenza	rTree.KM	rRF	Q1	cTree.KM	cRF	cLR	cEL
4	13/20	27	11	20	1	FN	1	1	1
5	13/20	27	12	21	1	FN	1	1	1
6	13/21	27	12	21	1	FN	1	1	1
7	14/21	27	14	22	1	FN	1	1	1
8	14/22	27	13	22	1	FN	1	1	1
9	15/22.25	27	13	22	1	FN	1	FN	1
10	16/23	27	15	21	1	FN	FN	FN	FN
11	17/24	27	15	22	1	FN	FN	FN	FN
12	18/25	27	16	23	1	FN	FN	FN	FN
13	19/26	27	21	24	1	FN	FN	FN	FN
14	21/28	27	17	24	0	0	0	0	0

Fig. 1. Monitoring dashboard presenting details about models predictions on the training set. *rTtree.Km* is a decision tree regressor and *rRF* is a random forest regressor. *cTree.KM*, *cRF*, *cLR*, *cEL* are the fourth quartile classifiers based on, respectively: decision tree, random forest, logistic regression and ensemble learning.

2.2 Ensemble Prediction Model

An ensemble model was built to predict prolonged hospital stays, combining two regression models (Decision Tree, Random Forest) and three classification models (Random Forest, Logistic Regression, Survival Decision Tree). Regression models estimate the exact length of stay, while classification models assess whether it exceeds the fourth quartile of the reference group. A weighted voting method assigns patients to high, medium, or low risk. High risk is assigned when both regression models predict a stay in the fourth quartile and all classifiers show a PPV > 0.4. Medium risk applies if at least one regression model and one classifier meet this criterion, or if both regressors predict a high stay regardless of classification. Low risk includes all other cases. Separate ensemble models were trained for each day from day 4 to 14, incorporating newly available data. Models were also customized by ER diagnosis group to reduce patient heterogeneity. Performance was tested on an independent set, focusing on PPV to better capture true high-risk cases.

2.3 Monitoring Dashboards

Two dashboards were developed post-model training to: a) visualize model performance on training and testing sets, and b) simulate clinical use for decision support. The first dashboard enables the technical team to monitor and analyze predictions. The second simulates a clinical scenario by randomly selecting 80 patients from the testing set, converting them into inpatients on random admission days. The ensemble model is applied to this simulated cohort, generating a real-time list of potential bed blockers (patients at medium or high risk of prolonged stay). It highlights key variables and their trends impacting each prediction.

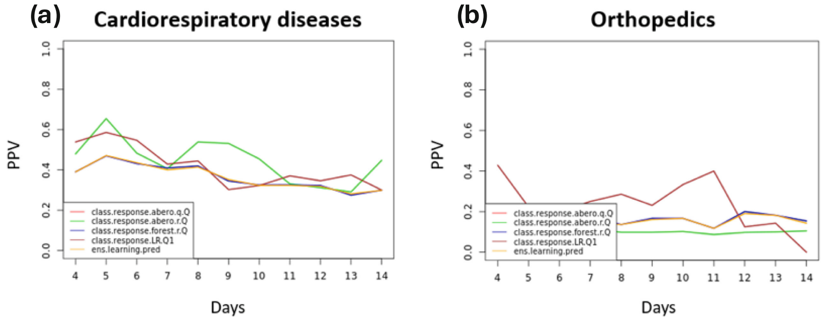

Fig. 2. Performances of each predictive models and the final ensemble model on two different diagnosis groups, in predicting the fourth quartile.

3 Results

A total of 9,915 patients were included and split into a training set (6,290; 63.44%) from 2022 admissions and a testing set (3,625; 36.56%) from the first half of 2023. Patients were categorized into 9 homogeneous groups based on ER admission diagnoses. Cardiorespiratory diseases were the most common in both sets (28.66% training, 28.88% testing), while orthopedic diagnoses were the least frequent (1.03% training, 1.43% testing). Median hospital length of stay (LOS) across most groups was 13 days, except for orthopedics (18 days in training, 17 in testing). The two sets showed no significant differences across diagnosis groups (p-values < 0.05), but the incidence of frailty unit consultations was significantly higher in 2023. To align with the Graphical User Interface, quartiles were inverted: Q1 represents the highest quartile, and Q4 the lowest. Figure 1 presents the monitoring dashboard used to track model performance. The table details predictions on training set, emphasizing the proportion of correctly identified high-risk patients. Figure 2 summarizes model performance across two diagnosis groups. In the "cardiorespiratory diseases" showed declining PPV, and "orthopedics" consistently underperformed across all models and time point.

The second dashboard, showed in Fig. 3, simulates a list of patients in charge by randomly sampling 80 patients from the testing set and transforming them from discharged patients to inpatients. The column labeled "LOS.risk" in the table provides the risk for extended hospitalization, i.e. length of stay above the fourth quartile of the reference subpopulation, as estimated by the ensemble model. Moreover, the summary table also includes data on model performance, specifically the PPV, to evaluate prediction reliability.

day	repartoIngresso	diagnosiPS	cluster	median/Q1	LOS.risk	PPV > mediana	PPV > Q1
11	PNEUMOLOGIA	POLMONITE BATTERICA, NON SPECIFICATA	infettivi	17/25	2 - high	0.7	0.41
5	PNEUMOLOGIA	BRONCHITE CRONICA OSTRUTTIVA, CON RIACUTIZZAZIONE	infettivi	13/21	2 - high	0.7	0.41
7	PNEUMOLOGIA	POLMONITE IN ALTRE MALATTIE INFETTIVE CLASSIFICATE ALTROVE	infettivi	14/21	2 - high	0.7	0.41
12	MALATT. INF.	MALATTIA DA SARS-COV-2 (COVID-19) CONCLAMATA, VIRUS IDENTIFICATO	infettivi	19/26	2 - high	0.7	0.41
7	GERIATRIA	BRONCOPOLMONITE, NON SPECIFICATA	infettivi	14/21	1 - medium	0.6	0.34
7	NEUROLOGIA	POLMONITE BATTERICA, NON SPECIFICATA	infettivi	14/21	1 - medium	0.6	0.34
7	NEUROLOGIA	POLMONITE BATTERICA, NON SPECIFICATA	infettivi	14/21	1 - medium	0.6	0.34
9	MED. INTERNA GER. COLUMBUS	CISTITE ACUTA	infettivi	15/22	1 - medium	0.6	0.34
14	UNITA COGN.-FUNZ.	POLMONITE BATTERICA, NON SPECIFICATA	infettivi	21/29	1 - medium	0.6	0.34

Fig. 3. Clinical dashboard simulating a list of in charge patients with estimated risk of prolonged hospitalization.

4 Discussion and Conclusions

This study presents an ensemble model combining classification and regression approaches to predict prolonged hospital stay in patients aged 50+. Models are dynamically updated from day 4 to 14 of hospitalization using real-time clinical data. A dashboard simulates clinical integration and supports daily decision-making. By incorporating care intensity, the model enhances risk stratification and early identification of bed blockers. This is the first work to integrate routine data with care intensity for this purpose. Limitations include missing nursing scale data and manual entry inconsistencies. Future improvements include automating risk stratification and validating the model across multiple settings or external validation data. Moreover, we could expand our work with a sensitivity analysis or varying the weights of regression and classification models for each diagnosis group to improve the performance.

References

1. Ackroyd-Stolarz, S., Read Guernsey, J., Mackinnon, N.J., Kovacs, G.: The association between a prolonged stay in the emergency department and adverse events in older patients admitted to hospital: a retrospective cohort study. BMJ Qual. Saf. **20**(7), 564–569 (2011)
2. Curiati, P.K., Gil-Junior, L.A., Morinaga, C.V., Ganem, F., Curiati, J., Avelino-Silva, T.J.: Predicting hospital admission and prolonged length of stay in older adults in the emergency department: The PRO-AGE scoring system. Ann. Emerg. Med. **76**(3), 255–265 (2020)
3. Damiani, A., et al.: Building an artificial intelligence laboratory based on real world data: The experience of gemelli generator. Front. Comput. Sci. **3** (2021)
4. Warren, M.B., et al.: Prolonged length of stay in ED psychiatric patients: a multivariable predictive model. Am. J. Emerg. Med. **34**(2), 133–139 (Feb 2016)