A novel approach for predicting acute hospitalizations among elderly recipients of home care? A model development study

Udsen, Flemming Witt; Stausholm, Mads Nibe; Hejlesen, Ole; Cichosz, Simon Lebech

Published in:
International Journal of Medical Informatics

DOI (link to publication from Publisher):
10.1016/j.ijmedinf.2022.104715

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Udsen, F. W., Stausholm, M. N., Hejlesen, O., & Cichosz, S. L. (2022). A novel approach for predicting acute hospitalizations among elderly recipients of home care? A model development study. *International Journal of Medical Informatics*, *160*, Article 104715. https://doi.org/10.1016/j.ijmedinf.2022.104715

International Journal of Medical Informatics 160 (2022) 104715

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

A novel approach for predicting acute hospitalizations among elderly recipients of home care? A model development study

Udsen Flemming Witt *, Stausholm Mads Nibe, Hejlesen Ole, Cichosz Simon Lebech

Department of Health Science and Technology, Aalborg University, Fredrik Bajers Vej 7E, DK-9220 Aalborg, Denmark

A R T I C L E I N F O

Keywords:
Aged
Geriatric Assessment
Machine Learning
Home Health Nursing
Hospitalization

A B S T R A C T

Background: Frail elderly are at high risk of hospitalizations and have a complex pattern of risk factors that makes it hard to foresee potential needs for additional treatment and care. Machine learning algorithms are potentially well-suited to discover hidden patterns in registrations that are routinely made across sectors.

Objective: To investigate predictors and performance of machine learning algorithms designed to predict acute hospitalizations in elderly recipients of home care services.

Materials and methods: A development study based on a retrospective social sector cohort with 1,282 participants was designed. Included subjects were aged 65 or older and received home care services in Aalborg Municipality at least once a week from 1/1–2016 to 31/12–2017. Data were collected from a newly developed triage tool in combination with administrative and clinical data routinely collected in the Danish healthcare and social care sector. 857 predictors were tested and evaluated based on the area under the precision-recall curve (PR-AUC). The data was divided into a 70/30 training and test split with 5-fold cross-validation. A sliding window approach combining random under-sampling with a boosting algorithm (RUSBoost) was applied with a standard logistic regression included for comparison.

Results: The logistic regression achieved a PR-AUC of 0.01 (CI 0.00; 0.01) while the PR-AUC was 0.71 (CI 0.56; 0.76) for the RUSBoost algorithm. Four of the five most important citizen-level features used to accurately predict an acute hospitalization was the total number of services provided by the municipality to the citizen, the number of personal care registrations as well as number of medication handlings and nutritional registrations. A final important predictor was the number of physical complaints derived from the triage tool.

Conclusions: Municipalities routinely collect valuable administrative and clinical data that can be used for early prediction of acute hospitalizations. However, future studies are needed to validate the results.

1. Introduction

Around one in six citizens in the world will be over 65 years old within the next 30 years, rising from a proportion of 9% in 2019 to a projected 16% of all people in 2050 [1]. In Denmark, population forecasts has estimated the percentage of persons over 65 years to increase from 20% in 2019 to 22.6% in 2030, which is higher than the expected average for Northern Europe [1]. In the latest population projection, Statistics Denmark also estimate that the proportion of citizens aged 80 or over will double from 5% in 2021 to 10% in 2060 [2].

Older adults have relatively frequent contact with emergency departments, which often converts into hospitalizations [3,4]. 12% of Danish citizens over 65 years has at least one hospitalization per year [5]. Moreover, there is a correlation between hospital contact and home care use for the elderly. A US analysis showed that 30% of home healthcare patients are hospitalized or visit an emergency department [6]. In Denmark, 50% of those over 65 years who were hospitalized also received home care or lived in a nursing home (for patients hospitalized aged 80 or above the proportion was 80%) [5] resulting in expenses that are four times higher for this group than for the elderly in general [5].

If possible, hospitalizations should be avoided for dependent and frail elderly, since they can exacerbate challenges with mobility and daily activities [7] and lead to adverse events during hospitalizations, e. g. acquired infections [8], or in transitioning from the hospital to homes [9]. An overview of systematic reviews has highlighted common characteristics of effective interventions to prevent hospitalizations for frail older individuals [10]. Results show that the most effective interventions are those that integrate community care (e.g. at-home

* Corresponding author.
E-mail addresses: fwu@hst.aau.dk (U.F. Witt), okh@hst.aau.dk (H. Ole), simcich@hst.aau.dk (C.S. Lebech).

https://doi.org/10.1016/j.ijmedinf.2022.104715
Received 18 November 2021; Received in revised form 25 January 2022; Accepted 7 February 2022
Available online 10 February 2022

preventive geriatric screening) with hospital care [10].

To facilitate an integrated approach, and to help municipalities responsible for community care in Denmark, the Danish Health Authority have previously published guidelines to assist community health personnel in assessing the everyday health of the elderly receiving home care and to act upon any abnormalities detected in order to avoid preventable hospitalizations [11]. Based on these recommendations, Aalborg Municipality implemented a digital version of a triaged changing table suggested in the guidelines [11]. The tool consisted of a checklist used to detect changes in the citizen's habitual state by answering questions within five subgroups of health status: Mental and social state, behavior in the home, activities of daily living, consumption of food and beverages as well as physical complaints. The current state was given a color code (green, yellow or red) depending on the number of changes registered from the citizen's normal habitus within the five subgroups.

A high risk of hospitalizations and a complex pattern of risk factors of the elderly adults was also found in the overview of systematic reviews [10]. This makes machine learning (ML) a suitable way to help primary care and community health care personnel with decisions surrounding potential hospitalizations, because ML was developed specifically for prediction purposes with an intention to identify hidden patterns and interactions among various predictor variables [12–14]. However, not much is known about which variables to include in prediction models aimed at forecasting hospitalizations for the elderly in a community care context and expected predictive performance. A systematic review from 2014 concluded that, in general, models based on administrative and/or routine clinical data from community-dwelling adults performs well [15] and that the best performing models included historical activity and concrete codes for diagnoses and medicine [15]. However, a recent study has found good performance for prediction models based only on triage of functional assessments, care needs and geriatric syndromes from the elderly [16].

The objective of this study was therefore to investigate whether it was possible to predict acute hospitalizations in elderly recipients of home care services based on a newly developed triage tool in Aalborg Municipality in combination with administrative and clinical data routinely collected in the Danish healthcare and social care sector.

2. Materials and methods

2.1. Study design and data

A development study based on a retrospective cohort was designed [17]. Eligible participants were initially extracted from the electronic home care record (EHCR) in Aalborg Municipality and included subjects aged 65 or older, who had received home care services at least once a week from 1/1–2016 to 31/12–2017. Terminal subjects and subjects with severe mental disorders were excluded from the study.

The data material included 1,282 subjects. By applying a civil registration number that all Danish citizens receive at birth (a so-called CPR number) [18], precise linkage between registers across sectors can be made and in this case using data from three sources:

(1) The electronic home care record (EHCR) from Aalborg Municipality, which holds both citizen-level socio-demographic data as well as activity-based information primarily used for costing or compensation purposes. Municipality personnel routinely record the type, timing and duration for almost all social care activities such as personal care, practical help, home nursing care and rehabilitation activities for each contact.
(2) The triaged changing table database (TCT) from Aalborg Municipality [11], where all registrations and triage state changes are were stored. In Aalborg Municipality, the community personnel routinely utilize the triage tool for assessing the individual citizen's current health state.
(3) The Danish National Patient Register (DNPR), which contains patient-level data of all inpatient, outpatient and emergency contacts to all hospitals in Denmark, and the register forms the basis for hospital statistics in Denmark [19].

The principles of the International Committee of Medical Journal Editors (ICMJE) checklist for assessment of medical AI were adhered to [20].

2.2. Model development

The models were implemented using Matlab R2020b (The Mathworks Inc., Natick, Massachusetts). The outcome of interest was whether acute hospitalizations has taken place, which in the DNPR is defined by two variables (a hospitalization date and an indicator of acute or planned hospitalization). Therefore, the problem was a binary classification problem [12–14]. A sliding window approach was implemented [21], which is illustrated in Fig. 1.

Periods in which there is a hospitalization (an event, marked by a red line) was called event periods and periods with no hospitalizations was called control periods. There was a fixed lead time of 3 days for each period to make it possible for community health care personnel to act and potentially avoid the predicted hospitalizations from the algorithm. Features from a fixed period of two weeks prior to the lead time period was used to predict acute hospitalizations. 90,375 periods were extracted with 345 of those resulting in an acute hospitalization.

Before pre-processing and model development, the data material was randomly divided at patient-level such that 70% was used for training and validating the models and 30% was reserved for testing the model. Hence, the 30% reserved for testing was not used in the development of the models. This procedure ensured that the results were not prone to overfit allowing for better transferability to a similar cohort [22]. This approach is illustrated in Fig. 2.

2.2.1. Pre-processing

Periods with and without acute hospitalizations were derived based on the availability of data from the EHCR, TCT, and DNPR. Some subjects did not have registered data in a given period within the time frame, this could be because of e.g. death of the subject, moving to another municipality, or unknown registration failure. This study utilized a complete-case-approach such that only periods with available data were included for analysis.

2.2.2. Feature extraction and reduction

857 features were extracted for each period from the data sources presented in Table 1. Extracted features were calculated from the history of hospitalizations, socio-demographics, municipal information, usage of past and current home care services, and absolute and relative changes from the triaged changing table. The features were extracted from the observation in the periods for each variable by calculating mean, standard deviation, maximum value, minimum value, number of observations, slope of linear regression, and most frequent values. Furthermore, the difference between the first 7 days and last 7 days of the period were calculated to account for relative changes, which could indicate change in health status. Current home care service features also included the relative between visitation time and service time delivered. This large number of features were extracted because we did not know which information from the data sources would have relevant information about potential exacerbations which leads to hospitalization.

The initial 857 extracted features were before model training reduced to a subset of 40 features. This procedure was conducted using the area under the receiver operating characteristics curve (ROC AUC) from each feature to rank their ability to discriminate between control and event periods. The reduction procedure was utilized to reduce computation time. Moreover, many of the potential features were assumed to hold redundant information. We included a feature set from

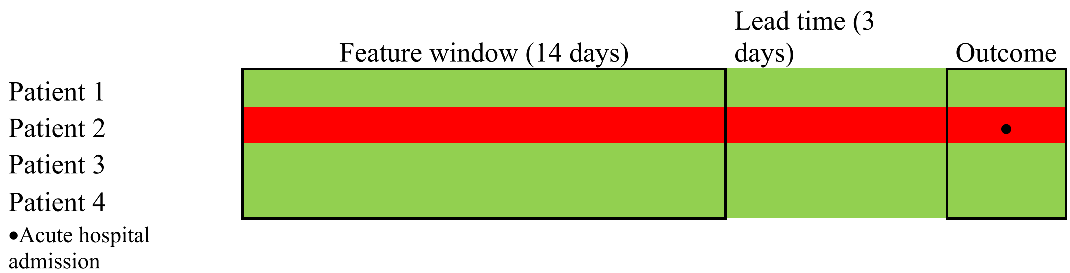

Fig. 1. Sliding window approach. Each day the window will slide one step forward while feature window and lead time remains the same.

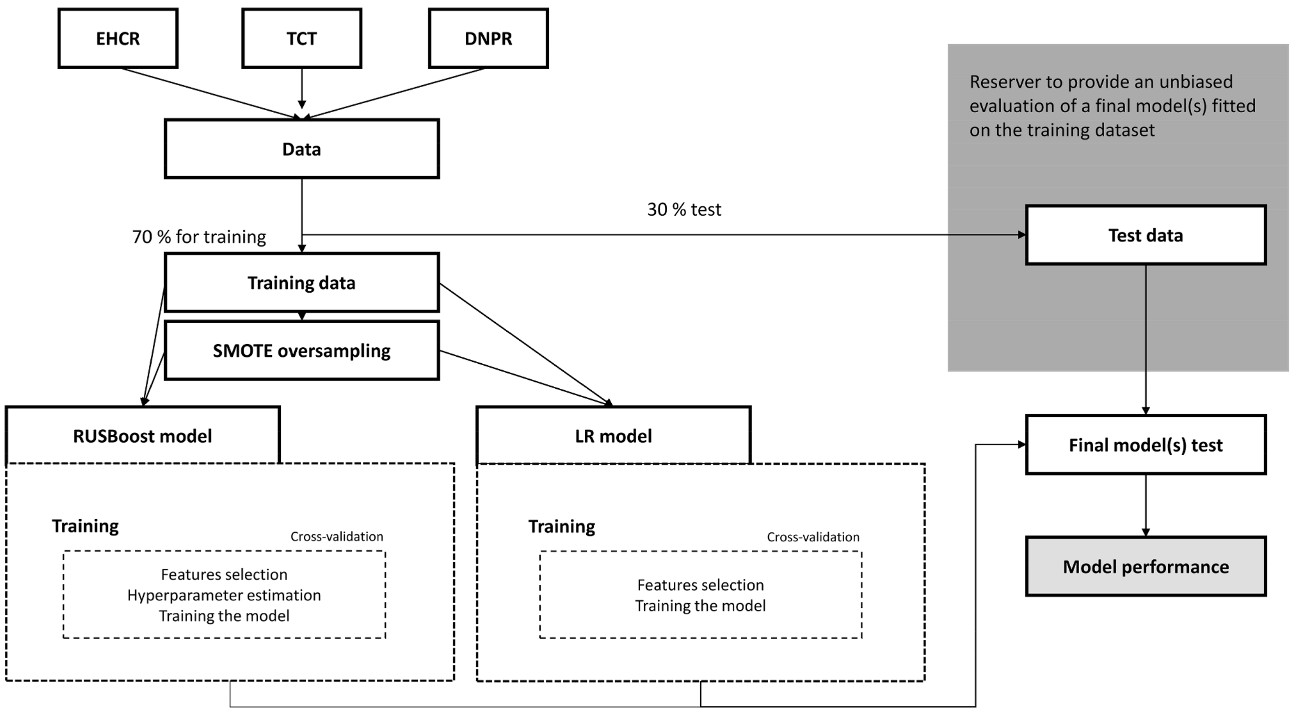

Fig. 2. Illustration of the model development. The data sources (EHCR, TCT database, and DNPR) were combined into one dataset before splitting the dataset into a training dataset (70%) and test dataset (30%). The test dataset was reserved for testing the final model(s). The training dataset was used to develop models. Cross-validation was used for features selection, hyperparameter estimation and final training of the model(s). Moreover, a training dataset with oversampling (using SMOTE) of the minority class was also used to train the model(s).

Table 1
Overview of the extracted information from the data sources.

Data source	Variables
Electronic home care record (EHCR)	Age, gender, marital status, comorbidities, medical prescriptions, past and current home care services including start and stop dates, and time spent
Triaged changing table database (TCT)	Habitual health status, registrations from each home care visit, and information concerning deviations from the habitual health status
Danish National Patient Register (DNPR)	Hospitalizations including acute/planned contact, primary diagnosis, procedural codes, and length of stay.

the subset using forward selection and 5-fold cross-validation to reduce overfitting of the model.

2.2.3. *Choice of machine learning algorithm*

We chose to compare a simple linear model (logistic regression) and a more complex nonlinear model (RUSBoost). Logistic regression was used due to the simplicity of the model with linear combination of features, which gives an easy interpretable model.

Because of the imbalance in the training data between periods without hospitalization and periods with hospitalization RUSBoost ensemble method was also used to model data [23]. RUSBoost have been reported to be a fast and robust classifier for dataset with imbalanced data [24]. RUSBoost has been reported to be computationally less extensive with slightly better performance than more complex methods for dealing with class imbalance [25]. This combination of simplicity, speed and performance makes RUSBoost an interesting technique for learning from imbalanced data [25].

We also compared training the models with and without over-sampling the training dataset 1:1 between the two classes. The training dataset was balanced using Synthetic Minority Over-sampling Technique (SMOTE, k = 5) with five nearest neighbors [26].

Hyper-parameters for the RUSBoost model (i.e. learning cycles [range, 10:10:1000], learn rate [range, 0:0.1:10]) were determined using 5-fold cross-validation to minimize overfitting. The final model was then tested on the test dataset. Our study did not recalibrate the model post the training procedure, which is considered best practices to get an unbiased estimate of performance. To investigate the effect of the search space of hyper-parameters on generalization, a 'restricted' model was trained with a smaller search space (i.e. learning cycles [range, 10:1:100] and learn rate [range, 0:0.01:1]).

2.2.4. Evaluation metrics

To evaluate the performance of the model, the area under the precision-recall curve (PR-AUC) was chosen with standard ROC-AUC used for comparison. Precision-recall is a better measure for evaluating imbalanced classification problems [27]. Confidence intervals (CI) for PR-AUC and ROC-AUC were estimated using bootstrap replicas (n = 1000).

3. Results

An overview of the characteristics is presented in Table 2. The data material included 1,282 subjects with a total of 90,375 periods. Median age was 84 years (interquartile range (IQR): 11.3) and 63.7% were women. 41.7% had no hospitalizations, 31.8% had one, and 26.5% had two or more hospitalizations. Median duration for personal care were 257 min per week (IQR: 364.2), 12 min per week for medication (IQR: 83.1), 45 min per week for cleaning (IQR: 22), 82.6 min for nutrition services (IQR: 345) and 0 min per week (IQR: 0) were spent on shopping. Furthermore, the median Triage level was 1 (IQR: 0.4). Finally, the training dataset included 242 event periods with acute hospitalizations (among 198 patients) and 63,021 control periods whereas the test dataset included 103 event periods with acute hospitalizations (among 85 patients) and 27,009 control episodes.

3.1. Performance

3.1.1. Logistic regression model

The PR-AUC for the training dataset was 0.01 (CI 0.01; 0.01) and 0.01 (CI 0.00; 0.01) for the test dataset. The ROC-AUC from the training and test dataset were 0.70 (CI 0.67; 0.73) and 0.73 (CI 0.71; 0.75), respectively. Using SMOTE oversampling increased the PR-AUC on the training data to 0.42 (CI 0.40; 0.46), but the results from the test data was not improved (0.01 [CI 0.00; 0.01]).

3.1.2. Rusboost model

The PR-AUC for the training dataset was 0.99 (CI 0.99; 0.99) and 0.71 (CI 0.56; 0.76) for the test dataset. The ROC-AUC from the training and test dataset were 0.99 (CI 0.99; 0.99) and 0.99 (CI 0.98; 0.99), respectively. The ROC and PR curves for the test data are plotted in Fig. 3. The PR-curve in Fig. 3 (right) illustrates a tradeoff between the sensitivity and positive predictive value of the algorithm that arises due to falsely predicting a hospitalization when none actually occurs. E.g. if a sensitivity of 74% is required (i.e. accurately finding roughly ¾ of all actual hospitalizations) that would correspond to a positive predictive value of 47% (i.e. the algorithm is then able to correctly predict hospitalizations in roughly half of all predictions that flag a risk of a coming hospitalization). If lower sensitivity thresholds for sensitivity are acceptable (e.g. 59%) then a higher positive prediction value can be achieved (87%).

Table 3 show prediction metrics for several threshold values. Using SMOTE oversampling in the training set did not increase the PR-AUC on the test data, which was 0.60 (CI 0.48; 0.64). The PR-AUC results from the search 'restricted' model were 0.78 (CI 0.77; 0.78) for the training set and 0.70 (CI 0.54; 0.75) for the test dataset. The ROC-AUC from the training and test dataset was 0.99 (CI 0.99; 0.99) and 0.99 (CI 0.98; 0.99), respectively.

Table 2
Overview of the cohort characteristics; presented as proportion (%) or median [25 percentile; 75 percentile].

Characteristic	
Subjects (n)	1,282
Gender, women %	63.7
Age, years	84 [77.7; 89.0]
Charlson Comorbidity Index (CCI)	
CCI 0, %	26.2
CCI 1, %	25.8
CCI 2, %	19.5
CCI 3+, %	28.5
Admissions last 12-months	
0 admissions, %	41.7
1 admission, %	31.8
2+ admissions, %	26.5
Triage level	1 [1.0; 1.4]
Duration of personal care, minutes per week	257 [85.8; 450.0]
Duration for medication, minutes per week	12 [0.0; 83.1]
Duration for shopping, minutes per week	0 [0.0; 0.0]
Duration for nutrition, minutes per week	82.6 [0.0; 345.0]
Duration for cleaning, minutes per week	45 [37.5; 59.5]

3.2. Most important features

3.2.1. Rusboost model

Five features from the feature set were included in the final model; number of personal care registrations (median: 11, range 0–15), number of medication handling registrations (median: 15, range 0–29), number of nutrition services registered (median: 12, range 0–15), number of services provided by the municipality (median: 45, range 15–85), number of physical complaints registrations from the home care visits (median: 3 range 0–8), Charlson Comorbidity Index (CCI) (CCI0: 26.2%, CCI1: 25.8%, CCI2: 19.5%, CCI > 2: 28.5%). The contribution from each feature is illustrated Fig. 4 (using forward selection of features and the corresponding PR-AUC on the training dataset).

4. Discussion

The result in this study demonstrates that the triage tool, in combination with administrative and clinical data routinely collected in the Danish healthcare and social care sector, have a potential for being utilized to predict elderly people at high risk of impending hospitalization. To our knowledge, this study is the first to investigate the potential of using these types of data for prediction of acute hospitalizations.

The performance of the logistic regression was low, especially the PR-AUC of 0.01 makes it unusable in practice. Furthermore, the implementation of SMOTE oversampling in the training of the model did not improve performance on the test dataset. An explanation of the low performance for the logistic regression model could be due to the linear capability of the model, which does not capture non-linear association in the data.

However, the performance of the RUSBoost model was acceptable with a PR-AUC of 0.71 (ROC-AUC was 0.99). The most important predictor of acute hospitalizations was the number of personal care registrations the last 14 days, followed by the number of medication handling registrations and number of nutrition services registered.

Four of the five most important features used to accurately predict an acute hospitalization came from routine municipality registrations in the feature period (i.e., the total number of services provided by municipalities as well as personal care, medication handling and nutritional registrations). The final important predictor stemmed from the developed triage tool (i.e., the number of physical complaints).

Using a 'restricted' search space for the training of the RUSBoost model did not improve the performance of the model. However, the gap (in PR-AUC) between the training and test dataset was narrowed.

4.1. Implications

The model and results are of relevance to both municipalities and clinicians conducting geriatric assessments within a community or home health care setting. The included patient socio-demographics (age, gender, marital status), the patients' historical activity in hospitals, outpatient clinics, comorbidities and clinical diagnoses from previous hospitalizations does not seem to be important in predicting future acute hospitalizations. What matters most are routine administrative and clinical registrations in municipalities that are well suited to predict

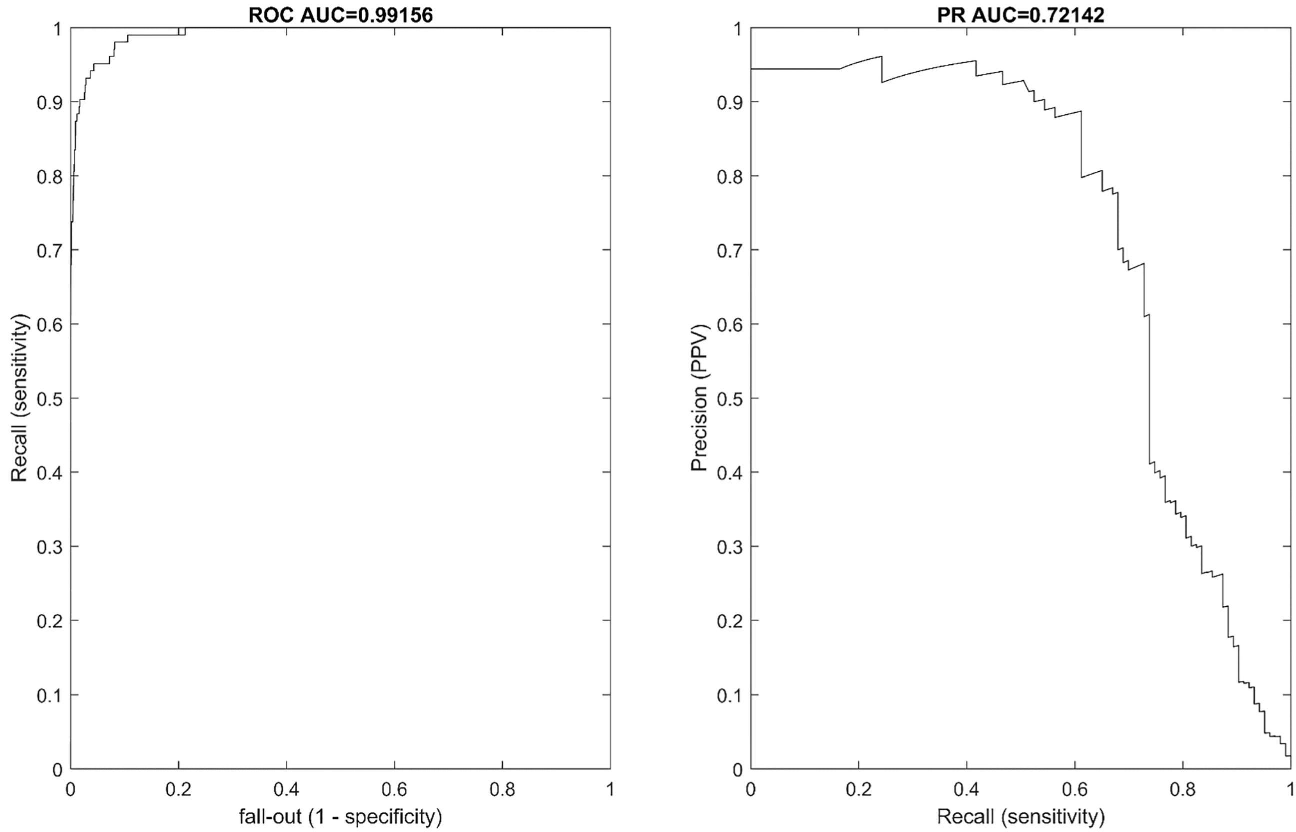

Fig. 3. Receiver operating curve (ROC) for the RUSBoost classifier (left). Precision-Recall (PR) curve for the RUSBoost classifier (right). The performance is based on the test dataset.

Table 3
Specific metrics for several thresholds of the RUSBoost model (based on the test dataset). TPR: true positive rate; TNR: true negative rate; PPV: positive predictive value; NPV: negative predictive value; TP: true positive; TN: true negative; FP: false positive; FN: false negative.

TPR	TNR	PPV	NPV	TP	TN	FP	FN
0.74	0.99	0.47	0.99	76	26,924	85	27
0.66	0.99	0.73	0.99	68	26,985	24	35
0.59	0.99	0.87	0.99	61	27,000	9	42

future hospitalizations for the elderly with one item from the triage tool implemented in Aalborg Municipality was also highly relevant. We can only speculate as to why this is the case. Routine registrations might be more rigorously registered if they are used for compensation or costing purposes. The triage tool is also a first version seeking to translate potential predictors into everyday measurements and it might therefore need further revision if it should be used as a prediction tool for acute hospitalizations.

5. Strengths and limitations

This study is to our knowledge the first to apply administrative and clinical registrations routinely documented by Danish regions and municipalities and to combine them with a community-based triage tool to predict acute hospitalizations. The data was derived from a heterogenic sample of citizens above age 65, who received home care services. However, several limitations are present. Firstly, the prediction relied heavily on assumptions about the causal relation between municipality registrations, triage and the hospitalization, and this relationship might not be fully understood (e.g. hospitalizations might also influence home care and future triage registrations). Secondly, despite including a

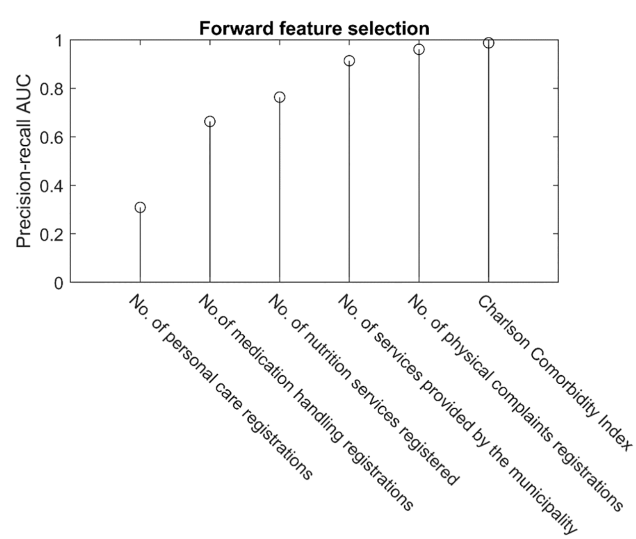

Fig. 4. Forward selection of predictors in relation to precision-recall-AUC. The predictors are selected from the training data using 5-fold cross-validation.

sufficiently sized sample for prediction purposes, the number of outcome events (345 hospitalizations) were rather limited and the results could be subject to bias [28]. This is also seen from the test sample results, where the PR-AUC is significant lower that the PR-AUC from the training data. While we cannot exclude overfitting, even though cross-validation was applied, the difference in results might also be because of a large heterogeneity among participants. This mean that the expected performance on new data is not yet solid, even though the results clearly point toward that the data holds valuable information in relation to disease exacerbation. Thirdly, the impact of the prediction model is not validated. From a technological readiness perspective, this would require a proof of principle setup [29] with e.g. a large randomized controlled trial, which could investigate if the proposed predictions could lead to earlier treatment, reduced hospitalizations or length of stay. Fourthly, multiple sources of undesirable biases can arise in the development of ML algorithms, that can lead to unintended discrimination of gender and minority groups. Investigation of model fairness was not performed. However, models were trained and tested on a dataset that is fairly representative of the gender distribution of the elderly receiving home health care. Also, gender was not included in the final model due to lack of explanatory power. Assessing ethnicity bias based on register data was impossible, since this information, should it exist, is highly restricted in Denmark. So, representation bias is expected to be low across gender and unknown for ethnic minorities. Finally, access to home health care is universal and free of charge in Denmark leading to a low risk of algorithmic bias.

5.1. Comparison with other research

Few studies have been published that seeks to predict acute hospitalizations in a home care or community context. Topaz and colleagues [6] achieved a similar performance in predicting a combined outcome of emergency department visits or hospitalizations in the US with a random forest algorithm (PR-AUC 0.76). The study relied on text mining of clinical notes and a much younger population (mean age of the elderly were 70.8 compared to 82.9 in this study). In our study did not rely on text mining of patient journals (or claims data), but rather on structured registrations used nationally by Danish regions and common language standards applied for elderly in municipalities [30,31], which should allow for greater generalizability at least in a Danish context. Veyron and colleagues [32] found a lower performance seeking to predict emergency department visits (ROC-AUC 0.7 with a random forest) using a home care screening application with four domains. The population was slightly older than in our study (mean age 88 vs 82.9). Finally, using a gradient boosting algorithm, Jones and colleagues [33] sought to predict emergency department utilization using a standardized clinical assessment instrument (the Resident Assessment Instrument: Home Care (RAI-HC)) applied routinely in Ontario for recipients of home and community care in Ontario, Canada. With a similar population age (mean age around 82 years), the algorithm achieved a lower performance (ROC-AUC 0.689).

5.2. Future research

There exist plenty of alternative machine learning algorithms for binary classification, including but not limited to naïve bayes algorithms, K-nearest neighbors, support vector machines, decision trees, neural networks, and other ensemble methods than the RUSBoost applied in this study. Although the RUSBoost was chosen based on experiences with performance and computational requirements for imbalanced problems, we acknowledge that other methods could potentially lead to better performance, especially if a larger dataset with more events was available for investigation. Recently, there have been developments in algorithms specifically targeted the clustered nature of repeated measures in longitudinal data [34]. Future research could compare the performance of more methods and tuning strategies simultaneously. Machine learning algorithms could also be used to predict other relevant outcomes, such as repeated hospitalizations, length of stay or adverse events associated with hospitalization. Moreover, future research should have focus on maturing the algorithm and validating the results in other cohorts for geriatric assessment in a home health care setting. It is not yet clear if these results could be transferred to another country due to structural differences between registration methods. It could be interesting to explore deeper if especially routine administrative registration is a valuable unexplored data resource for predicting outcome for patients.

6. Conclusion

The novel approach utilized in this study, point towards that especially administrative and clinical data routinely collected in municipalities could be a valuable resource for early prediction of disease exacerbation leading to acute hospitalizations. Future studies are needed to validate the results further and explore the relation between the administrative data and exacerbations.

Author contributions

All authors have made substantial contributions to the following: (1) the conception and design of the study, acquisition of data, and analysis and interpretation of data, (2) drafting the article and revising it critically for important intellectual content, (3) final approval of the submitted version.

Funding

This project is partially supported by the Division for Elderly and Disabled and the Centre for Practice and Healthcare Research Collaboration, Aalborg Municipality, Denmark. Funders had no role in the study design, analyses, and interpretation of the results.

Data sharing

No additional data are available. Anonymized data cannot be shared because of Danish General Data Protection Regulations (GDPR).

Summary table

What was already known on the topic:

- Hospitalizations should be avoided for dependent and frail elderly due to a risk of adverse events and potential exacerbation of challenges with mobility and daily activities.
- The elderly have a complex pattern of risk factors associated with hospitalization making machine learning a suitable way to assist health professionals with decisions surrounding potential hospitalizations.
- Few studies have been published that seeks to predict acute hospitalizations in a home care or community context and little is known about the performance and the most suitable data sources and concrete predictors.
- Studies that have applied machine learning in a community context have used text mining of clinical notes and more or less standardized screening tools to obtain possible predictions of acute hospitalizations.

What this study added to our knowledge:

- By combining a newly developed triage screening tool with administrative and clinical data routinely collected across the Danish healthcare and social sector, we were able to achieve comparable or higher performance in predicting acute hospitalizations than found in existing studies.
- Administrative data registered in a community care context is a very important information source when predicting acute hospitalizations among the elderly. 4 of the 5 most important variables were found amongst routine administrative registrations entered in municipalities and only one item from the triage tool implemented in Aalborg Municipality was also highly relevant.
- Included patient socio-demographics, historical activity in hospitals and outpatient clinics, comorbidities and clinical diagnoses from

previous hospitalizations were of much less importance in predicting future acute hospitalizations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Aalborg Municipality for facilitating the implementation of the study.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2022.104715.

References

[1] United Nations: Department of Economic and Social Affairs, World Population Ageing 2019: Highlights, 2019.
[2] Statistics Denmark, Befolkningstilvæksten højest blandt de ældste (Danish), 2021.
[3] D.C. Roberts, M.P. McKay, A. Shaffer, Increasing Rates of Emergency Department Visits for Elderly Patients in the United States, Ann. Emerg. Med. 51 (6) (2008) 769–774.
[4] J.M. Pines, P.M. Mullins, J.K. Cooper, L.B. Feng, K.E. Roth, National Trends in Emergency Department Use, Care Patterns, and Quality of Care of Older Adults in the United States, J. Am. Geriatr. Soc. 61 (1) (2013) 12–17.
[5] Sundhedsdatastyrelsen, Ældre medicinske patienters kontakt med det regionale sundhedsvæsen og den kommunale pleje (in Danish), 2016.
[6] M. Topaz, K. Woo, M. Ryvicker, M. Zolnoori, K. Cato, Home Healthcare Clinical Notes Predict Patient Hospitalization and Emergency Department Visits, Nurs. Res. 69 (6) (2020) 448–454.
[7] W. Duan-porter, T.N. Vo, K. Ullman, L. Langsetmo, E.S. Strotmeyer, B.C. Taylor, et al., Hospitalization-Associated Change in Gait Speed and Risk of Functional Limitations for Older Adults, J. Gerontol. A Biol. Sci. 74 (10) (2019) 1657–1663.
[8] M. Avci, O. Ozgenc, S.A. Coskuner, A.I. Olut, Hospital acquired infections (HAI) in the elderly : Comparison with the younger patients, Arch. Gerontol. Geriatr. 54 (1) (2012) 247–250.
[9] M. Mesteig, J.L. Helbostad, O. Sletvold, T. Røsstad, I. Saltvedt, Unwanted incidents during transition of geriatric patients from hospital to home : a prospective observational study, (2010) 1–9.
[10] I. Philp, F. Lond, C. Medical, E. Yorkshire, N.H.S. Hospitals, E. Riding, Reducing hospital bed use by frail older people: results from a systematic review of the literature, Int. J. Integr. Care 13 (December) (2013) 1–19.
[11] The Danish Health Authority, Tidlig opsporing af sygdomstegn hos ældre medicinske patienter: Værktøjer til hverdagsobservationer, 2013.
[12] B. Lantz, Machine Learning with R, Packt Publishing, Birmingham, 2013.
[13] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second ed., John Wiley & Sons, Inc., New York, 2001.
[14] J. Brownlee, Machine Learning Mastery with R. Vol. 1.1, 2016.
[15] E. Wallace, E. Stuart, N. Vaughan, K. Bennett, T. Fahey, S.M. Smith, Risk Prediction models to predict emergency hospital admission in community-dwelling adults: A systematic review, Med. Care 52 (8) (2014) 751–765.
[16] F. Mowbray, M. Zargoush, A. Jones, K. de Wit, A. Costa, Predicting hospital admission for older emergency department patients: Insights from machine learning, Int. J. Med. Inform. 140 (2020) 104163, https://doi.org/10.1016/j.ijmedinf.2020.104163.
[17] M.N. Stausholm, P.H. Secher, S.L. Cichosz, O.K. Hejlesen, Predicting Preventable Hospitalizations among Elderly Recipients of Home Care: A Study Protocol, in: Proc from 16th Scand Conf Heal Informatics 2018, Aalborg, Denmark August 28–29, 2018, pp. 75–79.
[18] C.B. Pedersen, H. Gøtzsche, J.O. Møller, P.B. Mortensen, The Danish Civil Registration System. A cohort of eight million persons, Dan. Med. Bull. 53 (4) (2006) 441–449.
[19] The State Serum Institute, The Danish National Patient Register, 2014.
[20] F. Cabitza, A. Campagner, The need to separate the wheat from the chaff in medical informatics: Introducing a comprehensive checklist for the (self)-assessment of medical AI studies, Int. J. Med. Inform. 153 (2021) 104510, https://doi.org/10.1016/j.ijmedinf.2021.104510.
[21] T.G. Dietterich, Machine learning for sequential data: A review, in: T. Caelli, A. Amin, R.P.W. Duin, D. de Ridder, M. Kamel (Eds.), Structural, Syntactic, and Statistical Pattern Recognition Lecture Notes in Computer Science, Berlin Heidelberg, Springer, 2002, pp. 15–30.
[22] R. Kohavi, A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection [Internet], 1995 [cited 2021 Mar 9]. Available from: http://robotics.stanford.edu/~ronnyk.
[23] C. Seiffert, T.M. Khoshgoftaar, J. Van Hulse, A. Napolitano, RUSBoost: A hybrid approach to alleviating class imbalance, IEEE Trans. Syst. Man. Cybern. Part A Syst. Hum. 40 (1) (2010) 185–197.
[24] C. Seiffert, T.M. Khoshgoftaar, J. Van Hulse, A. Napolitano, RUSBoost: Improving classification performance when training data is skewed. Proceedings - International Conference on Pattern Recognition, Institute of Electrical and Electronics Engineers Inc, 2008.
[25] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, F. Herrera, A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches, IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 42 (4) (2012) 463–484.
[26] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: Synthetic Minority Over-sampling Technique, J. Artif. Intell. Res. 16 (2002) 321–357. https://www.jair.org/index.php/jair/article/view/10302.
[27] T. Saito, M. Rehmsmeier, G. Brock, The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets, PLoS One 10 (3) (2015) e0118432, https://doi.org/10.1371/journal.pone.0118432.
[28] A. Rajkomar, M. Hardt, M.D. Howell, G. Corrado, M.H. Chin, Ensuring Fairness in Machine Learning to Advance Health Equity, Ann. Intern. Med. 169 (12) (2018) 866. www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M18-1990, https://doi.org/10.7326/M18-1990.
[29] A. Lavin, C.M. Gilligan-Lee, A. Visnjic, S. Ganju, D. Newman, A.G. Baydin, et al., Technol. Readiness Levels Mach. Learn. Syst. (2021) 1–26.
[30] Local Government Denmark. Common Language 3.
[31] M. Hansen, Evaluation of the use of common language in the Danish municipalities - survey with directors of social services and head of elderly, 2004.
[32] J.H. Veyron, P. Friocourt, O. Jeanjean, L. Luquel, N. Bonifas, F. Denis, et al., Home care aides' observations and machine learning algorithms for the prediction of visits to emergency departments by older community-dwelling individuals receiving home care assistance: A proof of concept study, PLoS One 14 (8) (2019) 1–13.
[33] A. Jones, A.P. Costa, A. Pesevski, P.D. McNicholas, Predicting hospital and emergency department utilization among communitydwelling older adults: Statistical and machine learning approaches, PLoS One 13 (11) (2018) 1–15.
[34] J.L. Speiser, A random forest method with feature selection for developing medical prediction models with clustered and longitudinal data, J. Biomed. Inform. 117 (2021) 103763, https://doi.org/10.1016/j.jbi.2021.103763.