Predicting hospital admissions and its cost due to respiratory diseases in Brazil using Machine Learning Time Series Forecasting

Rafaela Ferreira
IQVIA Solutions (Brazil)
Vinicius Goularte
IQVIA Solutions (Brazil)
Arthur Pate
IQVIA Solutions (Brazil)
Marion Coting Braga
IQVIA Solutions (Brazil)
Ariane AbreuB
IQVIA Solutions (Brazil)
Ramon Gonçalves Pereira (rpereira2@br.imshealth.com)
IQVIA Solutions (Brazil)

Research Article

Keywords: respiratory disorders; machine learning; artificial intelligence, health management

Posted Date: September 29th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2050214/v1

Abstract

Background: Respiratory diseases (RD) impose an immense health burden and over 1 billion people suffer from acute or chronic RD. Artificial Intelligence (AI) can improve the quality of healthcare, with the potential of assisting in the medical diagnosis of many diseases and reducing unnecessary hospitalizations and costs. This article aims to develop a Machine Learning (ML) model to predict the healthcare resources utilization (HCRU) and costs associated to RD hospitalizations in the Brazilian public health system (SUS).

Methods: Data were extracted from three public databases: Hospital Information System (SIH), "e-saúde" database and Meteorological Database, in the city of Curitiba, between 2017 and 2019. All analyzes considered the number of hospitalizations per day. The outcomes predicted by ML were the cost and the number of hospitalizations in the next seven days after a RD claim. The models were created by data mining process. Different algorithms were tested by the model building process up to five times. The best model for the seven-day cost and utilization forecasts was defined according to mean absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE). The SHAP method was used to analyze the interpretability of the best selected model.

Results: There were, on average, 315.41 hospitalizations and 97,596 primary care services for RD per week in the city of Curitiba between 2017 and 2019, with an average cost of 246,390.30 US dollars (R$ 549,332.87). The Recurrent Neural Network (RNN) methods (LSTM and GRU) presented the best results for forecasting costs and HCRU. LSTM model outperformed all other algorithms in both models with a RMSE of 0.07 and 0.04 respectively. The most impacting variables in the model (SHAP analysis) were the meteorological ones. However, the forward to specialist, type of attendance and medical specialty on the ambulatorial records were also important. High average temperatures support the model to make a prediction of a smaller number of hospitalization days for that period.

Conclusion: The prediction model used was robust enough to predict information about hospitalization and costs related to RD, demonstration its applicability as a tool to optimize resources allocation and health promotion strategies.

1. Introduction

Respiratory diseases (RD) impose an immense worldwide health burden [1]. This group includes a wide spectrum of diseases, ranging from infectious diseases such as pneumonia to chronic non-communicable diseases like asthma and chronic obstructive pulmonary disease (COPD). They account for more than 10% of all disability-adjusted life-years and are second only to cardiovascular diseases, including stroke [1]. RD include five of the thirty most common causes of death in the world: COPD, lower respiratory tract infection, tracheal, bronchial and lung cancer, tuberculosis and asthma [1–3]. Altogether, more than 1 billion people suffer from either acute or chronic respiratory conditions, including infants and young children are particularly susceptible[4,5].

In Latin America, lower respiratory tract infections like pneumonia cause 6% of deaths, while prevalence of asthma symptoms in adolescents exceeded 15% in countries like Paraguay, Peru, and Brazil [6]. In Brazil, asthma and COPD have higher prevalence than other chronic respiratory conditions, similar to other countries [7].

Several companies and organizations have already demonstrated how AI can improve the quality of care and/or decrease it, including new therapies and diseases diagnosis [1,8]. AI can help doctors make better decisions, leading to fewer unnecessary hospitalizations, which reduces costs [9]. Recently, Machine Learning (ML) techniques have been implicated on healthcare data for variable proposes [10]. Supervised ML models are a particular type of AI, that aim to accurately predict future outcomes based on a "learning" process from historical data. Those models are employed for instance in regression tasks, where an unknown continuous variable is predicted based on known information [11]. Considering health care management applications, ML models has been employed to forecast health care events such as inpatient admissions [12–14].

In that sense, understanding primary health care management in relation to RD can be crucial to optimize resources and health promotion strategies. For that AI models can be important methods to support primary health care to rapidly evolving not only in terms of health policies but also technologically [15]. In recent years different studies shown the applicability of AI into primary health care [16–20], however despite all the methodological and computational advances in AI, very few are translated into routine clinical practice [21,22]. Interpretability is a particularly important aspect in health-related ML models, which means to understand the influence and of each feature to the model, helping to understand the reasoning behind the predictions [23].

In this way, this study aimed to build ML models based on climatic factors and characteristics of outpatient visits for RD to predict hospitalization cost and healthcare resource utilization for RD within seven days frame.

2. Material And Methods

2.1. Study Design

Data was extracted from three public databases for the city of Curitiba and includes data between Januarys 1st, 2017 and December 31st, 2019. Data from 2020 and 2021 was excluded due to COVID-19 pandemic. The Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT) code for "Disorder of respiratory system" (320136) was used as primary criterion in the selection of the ICDs for RD, excluding the ICDs regarding neoplasms (*i.e* C00-C97 and D00-D48). ML models for forecasting HCRU and hospitalization costs related to RD were developed and evaluated.

Located in the southern region of Brazil, Curitiba is the 5th biggest city in Brazil, with an estimated population of over 1,9 million habitants and one of the highest (0.823) Human Development Index of

Brazil [24]. Curitiba provides open access to public data of primary attention care. The database of this study includes data extracted from three databases:

- Brazil's Hospital Information System (SIH), which contains hospitalization claims, was based on ICDs of Respiratory System Disorders. Variables include data on sex, age, ICD-10, type of care and type of professional [25].
- The “e-saúde” database, an informatized system which gathers telemedicine and telehealth data in the Unified Health System (SUS), and the extraction of data includes the number of urgent and general care, medical specialties, requests for medical exams, prescription and drug dispensing [26].
- Meteorological Database of the National Institute of Meteorology (INMET), which contains daily data on weather conditions following the technical standards of the World Meteorological Organization. The INMET database reports temperature, humidity, wind measures, and precipitation per hour [27].

2.2. Statistical Analysis

In the present study, data extracted from the three databases were adjusted and merged into one final database. The data were presented in the format of tables and graphs and Python version 2.7.3 was used. All variables are reported in Supplementary Appendix Table 1.

Table 1
Cost and HCRU models algorithms evaluation metrics results for the test set

Model algorithms	Linear Regression	XGBOOST	Neural Network	LSTM	GRU
Cost					
RMSE¹	0.16	0.15	0.15	0.07	0.08
MAE² (USD PPP)	25,149.20	24,849.20	24,669.92	11,821.83	13,776.29
MAE³ (USD PPP)	31,480.51	29,573.34	29,548.68	13,853.00	15,893.89
MAE² (BRL)	57,358.09	56,673.88	56,265.00	26,962.19	31,419.75
MAE³ (BRL)	71,798.00	67,448.28	67,392.05	31,594.70	36,249.39
MAPE (%)	15.26	13.96	13.80	6.05	7.22
HCRU					
RMSE¹	0.12	0.12	0.14	0.04	0.06
MAE² (Hospitalization days)	21.00	23.00	21.00	10.00	13.00
MAE³ (Hospitalization days)	30.99	29.95	32.78	12.35	17.11
MAPE (%)	11.73	11.90	12.48	4.19	6.32
¹Standardized; ²Median Absolute Error; ³Mean Absolute Error.					
BRL: Brazilian real; GRU: Gated Recurrent Units; HCRU: healthcare resources utilization; LSTM: Long Short-Term Memory; MAPE: Mean Absolute Percent Error; PPP: purchasing power parity; RMSE: Root Mean Squared Error; USD: United States dollar; XGBoost: Extreme Gradient Boosting.					

2.3. Machine Learning models built and Evaluation

Data was collected, cleaned, standardized, modeled and integrated for each one of the databases collected. Different algorithms (Linear Regression, Extreme Gradient Boosting [XGBoost], Artificial Neural Network, Long Short-Term Memory [LSTM], and Gated Recurrent Units [GRU]) were tested and had their performances compared according to 2.3.1.

For the cost model, a summary of the cost for each procedure performed during hospitalization was made and the final cost of each hospitalization was calculated per day and per seven consecutive days. This value was used to forecast the costs of all hospitalization for RD seven consecutive days. The HCRU model was developed to predict the frequency of hospitalizations within seven days in Brazil. For both models, training was performed using data between 2017 and 2018 (67%), while data from 2019 (33%) was tested. All algorithms were tested using 5-fold cross-validation evaluating optimal parameters for each algorithm.

For descriptive proposes, the costs in Brazilian Reais (BRL) were converted to US Dollars (USD) considering the official purchasing power parity (PPP) exchange rates for 2017 (1 USD = 2.182 BRL), 2018 (1 USD = 2.226 BRL), 2019 (1 USD = 2.281 BRL), which already considers inflation rate [28].

2.4. Defining the best model and interpretability analysis

The best model for the cost and utilization forecasts in seven days was defined according to performance measured by i) Mean Absolute Percent Error (MAPE): calculated as the sum of individual absolute errors divided by the actual value and presented as the average of all percentage errors; ii) Mean Absolute Error (MAE): used to identify an outlier and compare models; iii) Root Mean Squared Error (RMSE): also used to identify an outlier and, to minimize its sensitivity to outliers, real values were used.

Interpretability analysis was developed for the best performing model by SHapley Additive exPlanations (SHAP). Figure 1 shows the main processes involving the development of the model using ML. See the Supplementary Appendix, available with the full text of this article, for additional details.

3. Results

The database presented, on average, 315.41 hospitalizations and 97,596 primary care services for RD per week in the city of Curitiba from January 2017 to December 2019, with a weakly mean cost of USD PPP 246,390.30 (BRL 549,332.87). The most frequent (73%) group of ICD-10 for RD in the database was acute upper respiratory infections (J00-J06), while chronic lower RD (J40-J47) represent 8% of the database. Supplementary material Table 2 presents the percentage of primary care service in the database according to ICD-10 for RD.

3.1. Costs forecast results

The Recurrent Neural Network (RNN) methods (LSTM and GRU) presented the best results for cost forecasting, with a median absolute error prediction of USD PPP 13,853.00 (6.05%) for LSTM and US$ 15,893.89 (7.22%) for GRU. Table 1 shows the performance of the models for forecasting hospitalization costs and HCRU in the next seven days, while Fig. 2 shows the forecasting results for disorder of respiratory system hospitalization with public data for costs and HCRU in Curitiba during the study period.

Table 1. Cost and HCRU models algorithms evaluation metrics results for the test set

Regarding the variables related to the seasons, there are higher hospitalization costs related to RD in the autumn and winter seasons. A high level of the “mean humidity” and “mean temperature” content had a low and positive impact on the hospitalization costs; therefore, a high humidity and temperature levels present lower hospitalization costs, while a less humidity weather and cold days means higher hospitalization costs. The variables for emergency care and specialized emergency care (attend_UPA, special_urgency_attend_UPA, attend_SIACE) have high and positive impact on the hospitalization costs for this population, meaning that higher numbers of cases in the emergency care units (UPA) implies higher costs (Fig. 2).

3.2. HCRU forecast results

The best model for HCRU forecasting over the next 7 days was also the LSTM model, with median absolute error of 10.00 (4.19%) hospitalization days (Figs. 2 and Table 1). The most impacting variable in the model according to SHAP analysis was “mean_temperature”, with a low and positive impact on the HCRU in this study; therefore, high average temperatures support a model to make a prediction of a lower number of hospitalization days, for that timeframe. imply lower hospitalization days, and lower temperature stands for higher hospitalization days. In this model, the seasons autumn and summer impacts on higher hospitalization days. Also, more drug dispensed and more attendance in an emergency care unit resulted in less HCRU for RD in this study (Fig. 3).

4. Discussion

For both cost and HCRU models, the LSTM model outperformed all other models in forecasting the cost and HCRU over the next 7 days. According with the SHAP results for this study population, climatic factors have a major influence in RD hospitalizations: the low humidity tends to cause more cost in hospitalizations, while lower temperature tends to generate more HCRU and costs. Winter and autumn seasons also presented high influence in hospitalizations cost.

The LSTM and GRU algorithms were able to better understand the relationship between time variables and their influence on the future due to the algorithm characteristics. In addition, linear regression, XGBoost and Neural Network commonly show more difficulty in predicting abrupt changes, although linear regression may be one of the largely used methods, with a variety application in healthcare [29,30].

As expected, SHAP results showed that climatic factors including lower temperature and humidity levels, and winter and autumn seasons may be linked to higher hospitalizations costs and HCRU for RD, which was already observed in diseases such as COPD, viral respiratory infections and others [31–36]. Summer was related to positive and negative contributions to costs and HCRU, with a cost reduction in early January and an increase at the end of the summer, possibly related to a behavior during holidays and school return [37–39]. The prediction results have already shown the model's ability to handle the most complex behavior, and SHAP allows us to see the pattern of nonlinear separability, which public managers may use as a methodology to forecast public spending on RD [40].

It is possible to reproduce this study for other cities and other countries. Other diseases, in addition to RD (such as asthma, bronchitis, flu), are also sensitive to climatic factors, such as diseases with seasonal variation (such as malaria and dengue). A study in the Philippines used ML techniques on meteorological variables to develop models and accurately predict the temporal pattern of dengue incidence or occurrence [41].

Limitation includes a risk of the same patient be included more than once, since database was developed according to hospitalization visits (analyzes are not at patient level); therefore, the database could have patients that were hospitalized more than once. Our findings might not be applicable for other cities realities. Our study is based on the total volume of data available for hospitalizations for Curitiba, those hospitalizations are not directly linked with outpatient visits register. The hospitalizations used for the model could have included cases of patients not followed before in outpatient setting (e.g., emergency cases). Despite the aforementioned limitations, our study demonstrated a novel, generalizable, and high-performing method to predict and support health judgment for resources optimization.

5. Conclusion

The prediction model used was robust enough to predict information about hospitalization and costs related to RD, demonstration its applicability to be used as a tool to optimize resources and for health promotion strategies. We hope that researchers, policy makers, and medical professionals will use this approach for other research questions or scenarios related to judgments of health and decision-making.

Abbreviations

AI: Artificial intelligence; BRL: Brazilian reais; COPD: Chronic obstructive pulmonary disease; GRU: Gated Recurrent Units; HCRU: Healthcare resources utilization; INMET: Meteorological Database of the National Institute of Meteorology; LSTM: Long Short-Term Memory; MAE: Mean absolute error; MAPE: Mean absolute percentage error; ML: Machine learning; PPP: purchasing power parity; RD: Respiratory diseases; RMSE: Root mean squared error; RNN: Recurrent Neural Network; SHAP: SHapley Additive exPlanations; SIH: Hospital Information System; SNOMED-CT: Systematized Nomenclature of Medicine - Clinical Terms; SUS: *Sistema Único de Saúde* (Brazilian public health system); UPA: Emergency care units; USD: United States dollars; XGBoost: Linear Regression, Extreme Gradient Boosting.

Declarations

Acknowledgments: The authors are grateful to Guilherme Julian, Marina Birck, Pedro Antonacio and Renato Watanabe de Oliveira for their support with study design and medical writing services.

Authors' contributions: RF: data processing, research methodology development, statistical analysis, model evaluations, performance assessment; VG: data processing, research methodology development, statistical analysis, article writing; MCB: medical writing and health insights support; AP: medical writing and health insights support; AA: guidance and review of the article, research methodology in health sub-area; RGP: guidance and review of the article, data modeling, data processing, research coordinator and research methodology in computer science and statistics sub-area.

Funding: This study has been developed funded by IQVIA Solutions Brazil. This is part of a special program of thoughtful leadership and research development in real world studies.

Availability of data and materials: All data used in this manuscript are publicly available data. Data from the Information Technology Department of the Brazilian National Health System (DATASUS) database are available at https://www.datasus.saude.gov.br/, data from the primary care database of Curitiba city are available at https://www.curitiba.pr.gov.br/dadosabertos/busca/?grupo=1, and data from the Meteorological Database of the National Institute of Meteorology are available at https://bdmep.inmet.gov.br/.

Ethics approval and consent to participate: Not applicable. As this is a retrospective study involving the analysis of secondary databases made available online (open data), submission to the ethics and research committee involving human beings was not necessary, as there is no possibility of identifying the patients in question.

Consent for publication: Not applicable.

Competing interests: The authors declare that they have no competing interests.

References

1. Wang H, Naghavi M, Allen C, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388(10053):1459–544.
2. Organization WH. Global status report on noncommunicable diseases 2014. Geneva PP - Geneva: World Health Organization.
3. World Health Organisation. Asthma Key facts. World Heal. Organ.:1.
4. Valiulis A, Bousquet J, Veryga A, et al. Vilnius Declaration on chronic respiratory diseases: multisectoral care pathways embedding guided self – management, mHealth and air pollution in chronic respiratory diseases. Clin Transl Allergy [Internet] 2019;1–10. Available from: https://doi.org/10.1186/s13601-019-0242-2.
5. Goals MD. Pneumonia: the leading killer of children. 2006;368:1048–50.
6. Murray CJL, Lopez AD. Mortality by cause for eight regions of the world: Global Burden of Disease Study. 1997;349:1269–76.
7. Beran D, Zar HJ, Perrin C, et al. Burden of asthma and chronic obstructive pulmonary disease and access to essential medicines in low-income and middle-income countries. Lancet Respir [Internet] 3(2):159–70. Available from: http://dx.doi.org/10.1016/S2213-2600(15)00004-1.
8. Meskó B, Hetényi G, Gy Z. Will artificial intelligence solve the human resource crisis in healthcare ? 2018;1–4.
9. Mesko B. The role of artificial intelligence in precision medicine. Expert Rev Precis Med Drug Dev [Internet] 2017;2(5):239–41. Available from: https://doi.org/10.1080/23808993.2017.1380516.
10. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. JAMA. 2018;319(13):1317–8.

11. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018;2(10):719–31.
12. Zhou L, Zhao P, Wu D, Cheng C, Huang H. Time series model for forecasting the number of new admission inpatients. BMC Med Inform Decis Mak. 2018;18(1):1–11.
13. Olsavszky V, Dosius M, Benecke J, Vladescu C. Time Series Analysis and Forecasting with Automated Machine Learning on a National ICD-10 Database. Int J Environ Res Public Health. 2020;17(14):1–17.
14. Tong Y, Messinger AI, Wilcox AB, et al. Forecasting future asthma hospital encounters of patients with asthma in an academic health care system: Predictive model development and secondary analysis study. J Med Internet Res 2021;23(4).
15. The Lancet. Artificial intelligence in health care: within touching distance. Lancet. 2017;390(10114):2739.
16. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. npj Digit Med. 2018;1(1):39.
17. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 2016;316(22):2402.
18. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. npj Digit Med. 2018;1(1):40.
19. Salazar LHA, Leithardt VRQ, Parreira WD, da Rocha Fernandes AM, Barbosa JLV, Correia SD. Application of Machine Learning Techniques to Predict a Patient's No-Show in the Healthcare Sector. Futur Internet. 2021;14(1):3.
20. Nunes BP, Vissoci J, Delpino FM, et al. Machine learning analysis to predict health outcomes among emergency department users in Southern Brazil: a protocol study. Rev Bras Epidemiol 2021;24.
21. Stead WW. Clinical Implications and Challenges of Artificial Intelligence and Deep Learning. JAMA. 2018;320(11):1107.
22. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care — Addressing Ethical Challenges. N Engl J Med. 2018;378(11):981–3.
23. ElShawi R, Sherif Y, Al-Mallah M, Sakr S. Interpretability in healthcare: A comparative study of local machine learning interpretability techniques. Comput Intell. 2021;37(4):1633–50.
24. IBGE. IBGE- Dados demográficos e informações sobre Curitiba- Paraná [Internet]. Available from: https://cidades.ibge.gov.br/brasil/pr/curitiba/panorama.
25. Regulação DDE, Controle AE. Sistema de Informação Hospitalar. 2015;(nível 3).
26. Sistema E-Saude. - Perfil de atendimento de Enfermagem nas Unidades Municipais de Saúde de Curitiba [Internet]. Available from: https://www.curitiba.pr.gov.br/dadosabertos/busca/?termo=e-sa%25c3%25bade&grupo=1.
27. Brasil. Ministério da agricultura pecuária e abastecimento. Instituto Nacional de Meteorologia (INMET) [Internet]. Available from: https://portal.inmet.gov.br/sobre.
28. World Bank. PPP conversion factor. World Bank. 2022;1.

29. Martinelli DD. Generative machine learning for de novo drug discovery: A systematic review. Comput Biol Med. 2022;145:105403.
30. Xiong P, Lee SM-Y, Chan G. Deep Learning for Detecting and Locating Myocardial Infarction by Electrocardiogram: A Literature Review. Front Cardiovasc Med 2022;9.
31. Silva Júnior JLR da, Padilha TF, Rezende JE, Rabelo ECA, Ferreira ACG, Rabahi MF. Efeito da sazonalidade climática na ocorrência de sintomas respiratórios em uma cidade de clima tropical. J Bras Pneumol 2011;37(6):759–67.
32. Li Y, Reeves RM, Wang X, et al. Global patterns in monthly activity of influenza virus, respiratory syncytial virus, parainfluenza virus, and metapneumovirus: a systematic analysis. Lancet Glob Heal. 2019;7(8):e1031–45.
33. Wang X, Li Y, O'Brien KL, et al. Global burden of respiratory infections associated with seasonal influenza in children under 5 years in 2018: a systematic review and modelling study. Lancet Glob Heal. 2020;8(4):e497–510.
34. Goka EA, Vallely PJ, Mutton KJ, Klapper PE. Single and multiple respiratory virus infections and severity of respiratory disease: A systematic review. Paediatr Respir Rev. 2014;15(4):363–70.
35. Santo AH, Fernandes FLA. Chronic Obstructive Pulmonary Disease-Related Mortality in Brazil, 2000–2019: A Multiple-Cause-of-Death Study. COPD J Chronic Obstr Pulm Dis. 2022;19(1):216–25.
36. Tsuchiya LRRV, Costa LMD, Raboni SM, et al. Viral respiratory infection in Curitiba, Southern Brazil. J Infect. 2005;51(5):401–7.
37. Nascimento AP, Santos JM, Mill JG, Souza JB de, Reis Júnior NC, Reisen VA. Association between the concentration of fine particles in the atmosphere and acute respiratory diseases in children. Rev Saude Publica 2017;51.
38. Furlanetto KC, Demeyer H, Sant'anna T, et al. Physical Activity of Patients with COPD from Regions with Different Climatic Variations. COPD J Chronic Obstr Pulm Dis. 2017;14(3):276–83.
39. Johnston NW. The Similarities and Differences of Epidemic Cycles of Chronic Obstructive Pulmonary Disease and Asthma Exacerbations. Proc Am Thorac Soc. 2007;4(8):591–6.
40. Lundberg S, Lee S-I. A Unified Approach to Interpreting Model Predictions. 2017.
41. Carvajal TM, Viacrusis KM, Hernandez LFT, Ho HT, Amalin DM, Watanabe K. Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines. BMC Infect Dis. 2018;18(1):183.

Figures

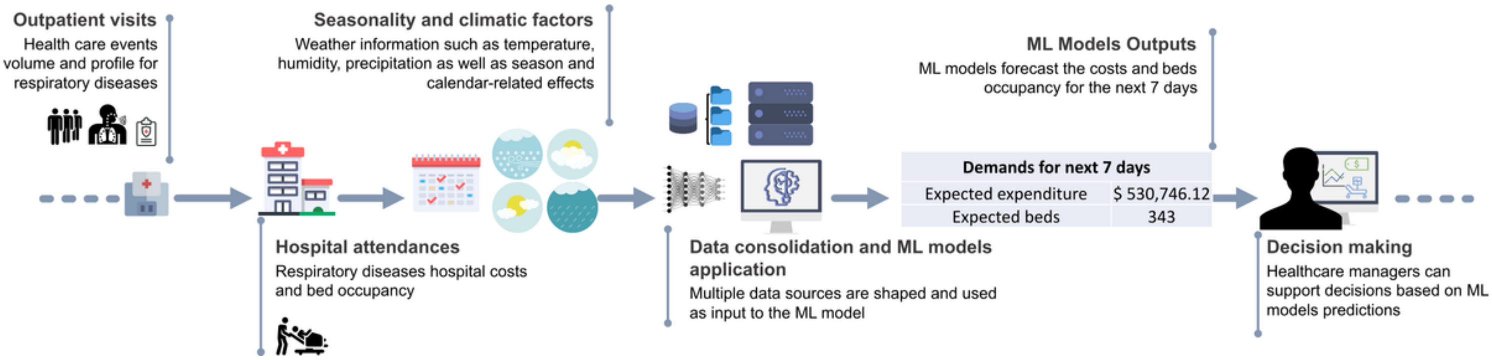

Figure 1

Flowchart of the main processes development for building the model

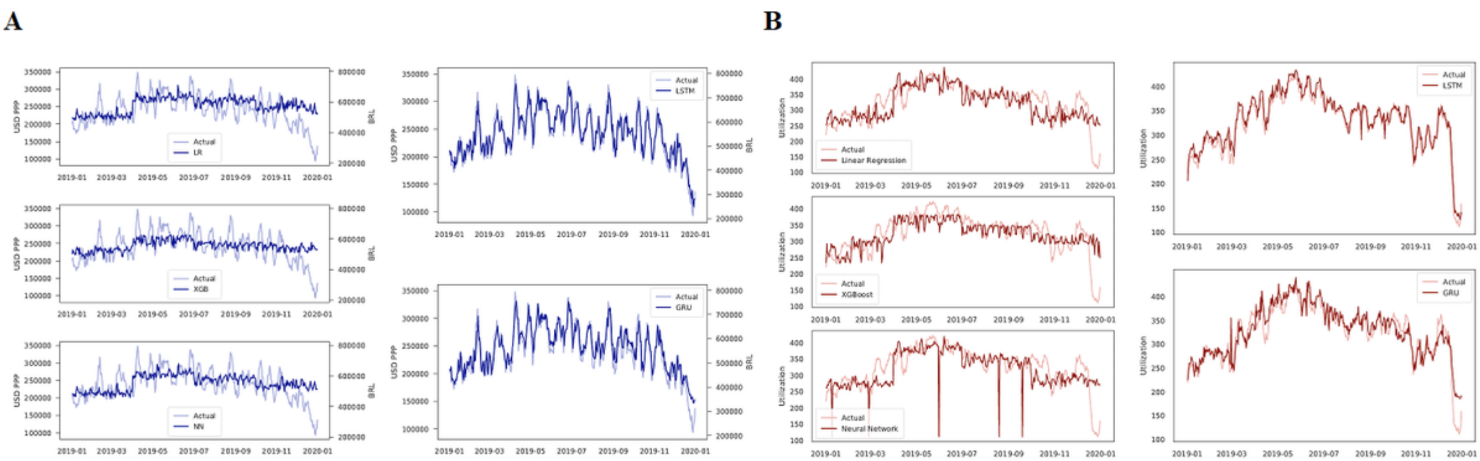

Figure 2

Forecasting for disorder of respiratory system with public data from Brazilian Hospital Information System (SIH) in Curitiba from Januarys 1st, 2017 and December 31st, 2019. A) Hospitalization costs; B) healthcare resources utilization.

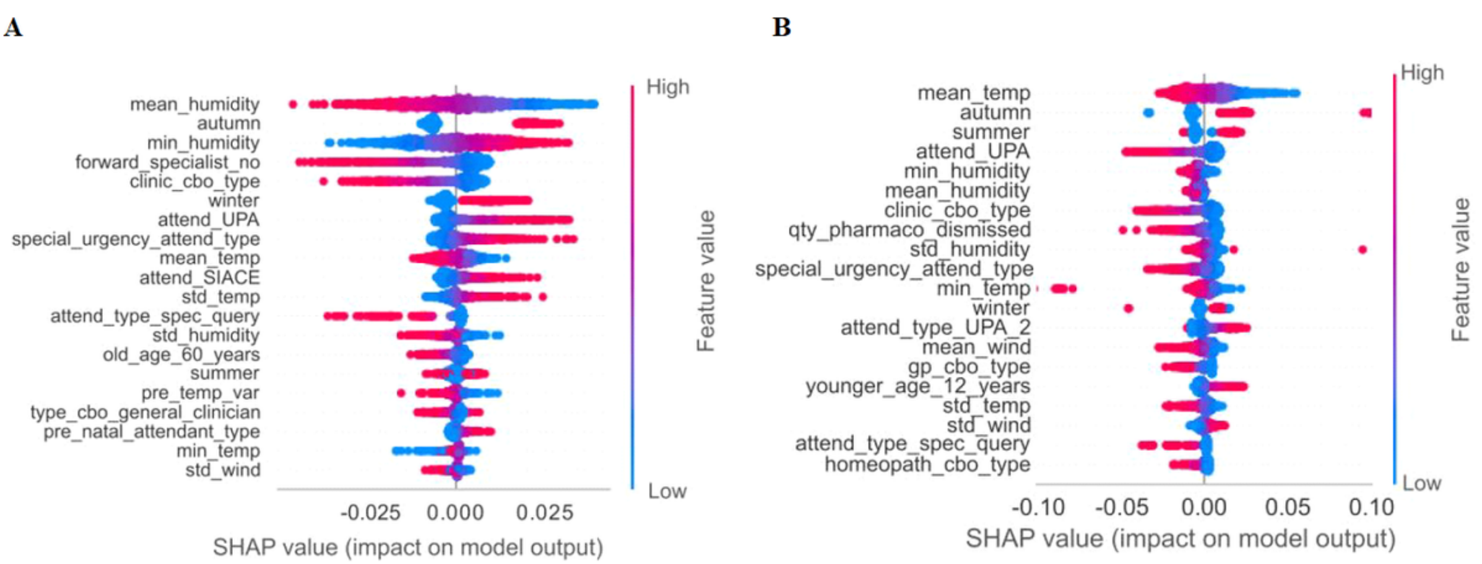

Figure 3

SHapley Additive exPlanations (SHAP) for prediction features using Long Short-Term Memory (LSTM). A) Cost model; B) Healthcare resources utilization model. The SHAP analysis classified variables by importance, where variables are ranked in descending order and with high (red) and low (blue) correlation to cost and HCRU. The horizontal location shows whether the effect of that value is associated with a positive or negative prediction.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryappendix.docx