

University of Louisville

Journal of Respiratory Infections



ORIGINAL RESEARCH

Predicting 30-Day Mortality in Hospitalized Patients with Community-Acquired Pneumonia Using Statistical and Machine Learning Approaches

Timothy L Wiemken,[†] Stephen P Furmanek, William A Mattingly, Brian E Guinn, Rodrigo Cavallazzi, Rafael Fernandez-Botran, Leslie A Wolf, Connor L English, Julio A Ramirez

Abstract

Background: Predicting if a hospitalized patient with community-acquired pneumonia (CAP) will or will not survive after admission to the hospital is important for research purposes as well as for institution of early patient management interventions. Although population-level mortality prediction scores for these patients have been around for many years, novel patient-level algorithms are needed. The objective of this study was to assess several statistical and machine learning models for their ability to predict 30-day mortality in hospitalized patients with CAP.

Methods: This was a secondary analysis of the University of Louisville (UofL) Pneumonia Study database. Six different statistical and/or machine learning methods were used to develop patient-level prediction models for hospitalized patients with CAP. For each model, nine different statistics were calculated to provide measures of the overall performance of the models.

Results: A total of 3249 unique hospitalized patients with CAP were enrolled in the study, 2743 were included in the model building (training) dataset, while the remaining 686 were included in the testing dataset. From the full population, death at 30-days post discharge was documented in 458 (13.4%) patients. All models resulted in high variation in the ability to predict survivors and non-survivors at 30 days.

Conclusions: In conclusion, this study suggests that accurate patient-level prediction of 30-day mortality in hospitalized patients with CAP is difficult with statistical and machine learning approaches. It will be important to evaluate novel variables and other modeling approaches to better predict poor clinical outcomes in these patients to ensure early and appropriate interventions are instituted.

DOI: 10.18297/jri/vol1/iss3/10

Received Date: April 19, 2017

Accepted Date: May 4, 2017

Website: <https://ir.library.louisville.edu/jri>

Affiliations:

Division of Infectious Diseases, Department of Medicine, University of Louisville School of Medicine, Louisville, KY: (TLW, SPF, WAM, BEG, LW, CLE, JAR)

Division of Pulmonary, Critical Care and Sleep Disorders Medicine, University of Louisville School of Medicine, Louisville, KY: (RC)

Department of Pathology & Laboratory Medicine, University of Louisville School of Medicine, Louisville, KY: (RF-B)

© 2017, The Author(s).



1 Introduction

CAP is the leading cause of infectious diseases-related death worldwide¹. Predicting if a hospitalized patient with CAP will or will not survive after admission to the hospital is important for research purposes as well as for institution of early patient management interventions. Historically, clinical judgment was used to predict mortality, make site-of-care decisions, and to inform clinical management. In 1997, the Pneumonia Severity Index provided clinicians with a more objective method for making

mortality predictions through provision of a simple population-based score using demographic, social, physical, laboratory, and radiographic data^{2,3}.

Over time, this score was shown to not be as ideal for patient-level decision making as it was for population-level research adjustment of disease severity⁴. For example, the Pneumonia Severity Index risk classes outline the risk for 30-day mortality. This can be useful for site-of-care decision making, but does not provide an individual prediction of if a single patient may survive or die, only aggregated estimates for a population of hospitalized patients. In the years since this score was developed, additional laboratory tests with predictive power for mortality have become standard, computing power is substantially stronger, and artificial intelligence computation is readily accessible. Although these patient care advances have occurred, few investigators have evaluated the utility of novel computational approaches to predicting mor-

[†]Correspondence To: Timothy L Wiemken PhD

Assistant Professor of Medicine

University of Louisville Division of Infectious Diseases

501 E Broadway Suite 120

Louisville, KY 40202

Office Phone: 502-852-4627

Email: tim.wiemken@louisville.edu

tality in hospitalized patients with CAP⁵⁻⁷. While traditional approaches in statistical modeling/learning such as logistic regression are widely used, novel machine learning approaches such as random forests, recursive partitioning, and other decision tree analyses provide a more robust approach to patient-level predictive modeling. Moreover, these methods can be combined in multiple ways, allowing for powerful computation and accurate prediction that is only relatively recently possible on personal computers.

The objective of this study was to assess several statistical and machine learning models for their ability to predict 30-day mortality in hospitalized patients with CAP.

2 Methods

This was a secondary analysis of the University of Louisville (UofL) Pneumonia Study database. The UofL Pneumonia Study was a prospective, population-based cohort study of all hospitalized adults with CAP who were residents of Louisville, Kentucky. Although this was a two-year study, only patients enrolled in study year 1 were included in the current study. These patients were enrolled from June 1, 2014 to May 31, 2015. All hospitalized adult patients in Louisville underwent screening for participation in the study.

2.1 Inclusion Criteria

A patient was defined as having CAP when the following three criteria were met: 1) presence of a new pulmonary infiltrate on chest radiograph and/or chest computed tomography scan at the time of hospitalization, defined by a board-certified radiologist's reading; 2) at least one of the following a) new cough or increased cough or sputum production, b) fever $>37.8^{\circ}\text{C}$ (100.0°F) or hypothermia $<35.6^{\circ}\text{C}$ (96.0°F), c) changes in leukocyte count (leukocytosis: $>11,000$ cells/ mm^3 ; left shift: $>10\%$ band forms/ mL ; or leukopenia: $<4,000$ cells/ mm^3); and 3) no alternative diagnosis at the time of hospital discharge that justified the presence of criteria 1 and 2.

2.2 Exclusion Criteria

With the intent to enroll only hospitalized patients with CAP who lived in Louisville, Kentucky and who were counted in the 2010 U.S. Census, patients were excluded from analysis if they: 1) did not have a permanent or valid Louisville address based on U.S. Census Bureau data, 2) did not have a valid Social Security Number (SSN), or 3) were in the correctional system.

2.3 Unique Patients Hospitalized with CAP

A unique patient hospitalized with CAP was counted as the first hospitalization during each study year. A re-hospitalization due to a new episode of CAP was identified by a repeat of the same SSN in the same study year. Only unique patients were included in the current study to limit bias in the study outcome.

2.4 Study Definitions

Predictor Variables: The following variables were included as candidate variables in our models: age, sex, body mass index (kg/m^2), nursing home residence, smoking status, active cancer, history of congestive heart failure, renal disease, liver disease, chronic renal failure, diabetes, cirrhosis, chronic obstructive pulmonary disease, HIV infection, asplenia, coronary artery disease, atrial fibrillation, prior myocardial infarction, hyperlipidemia, arterial hypertension, need for home wound care, need for chronic dialysis, home infusion therapy, intravenous drug use, pleural effusion on chest radiograph or computed tomography scan, suspicion of aspiration, need for intensive care on admission, altered mental status, need for invasive mechanical ventilation on admission, need for blood pressure support on admission, hospitalization in the prior 90 days, use of intravenous antibiotics in the prior 90 days, heart rate respiratory rate, systolic and diastolic blood pressure, oxygen saturation, FiO_2 , hematocrit, hemoglobin, white blood cell count, platelets, serum sodium, serum potassium, blood urea nitrogen, creatinine, serum bicarbonate, serum glucose, albumin, alanine aminotransferase, aspartate aminotransferase, and bilirubin.

Outcome Variable: The outcome variable in this study was all-cause mortality up to 30 days after hospitalization. Mortality was obtained through medical record abstraction and using data from the Kentucky Department for Public Health Office of Vital Statistics.

2.5 Quality Control/Data Management Plan

The UofL Pneumonia Study Coordinating Center provided research support for the University of Louisville Pneumonia Study. Trained study coordinators and/or research associates collected clinical data from the patient's medical record onto a paper case report form. A separate research associate entered these data into a secure, web-based electronic data capture system called REDCap. The data dictionary for the project was designed by clinical and analysis experts in the center to ensure an appropriate mapping from the case report form to the electronic database. The Center's REDCap instance is hosted at the enterprise security datacenter at the University of Louisville, with security procedures and protocols for HIPAA compliant database operation. REDCap supports user access controls, audit trails of all accessed data, and timed automatic logouts to prevent accidental exposure of patient data. Each of these features were actively used throughout the study. Data quality rules based on good clinical practice and standard of care were used to limit out-of-range errors and inappropriate data types. After any data quality issues were resolved using REDCap's query resolution workflow, cases were locked in REDCap for analysis.

Statistical Analysis: For each analysis, the dataset was split into a training and a testing set. A random sample of eighty percent of the subjects were included in the training set, while the residual twenty percent were included in a testing dataset. Several statistical and machine learning models were used in the current study. Since many of the models used are considered to function

poorly in the presence of an unbalanced outcome (e.g. mortality was not 50%), we conducted several of the analyses in the full training dataset, as well as various sampling schemes in order to provide the models with an equal balance of survivors (here, known as the majority class since there are more cases of survivors in our dataset) and non-survivors (here, known as the minority class since there are fewer cases of non-survivors in our dataset). First, we chose a random down-sampled selection of the cases who survived, next, we chose a random up-sampling of cases who did not survive, and finally an alternative up-sampling scheme termed "Synthetic Minority Over-Sampling Technique" or SMOTE. The down-sampling scheme selected a random sample of those who survived to match the number of individuals who did not survive. The up-sampling scheme repeated individuals who did not survive randomly into the dataset until the frequency of survivors and non-survivors was the same. The SMOTE sampling approach up-sampled those who died, but using synthetic cases based on data in the dataset. For a more-in depth overview of the SMOTE sampling process, see Chawla et. al.⁸. Each of these datasets provided an equal proportion of survivors and non-survivors for analysis. The caret package in R was used for up and down sampling of cases, while the package DMwR was used for SMOTE sampling. The testing dataset was not up or down sampled as it allowed us to evaluate each model in the context of what would occur in clinical practice (e.g. 30-day mortality being relatively rare). Each model and the datasets used for analysis are described below.

1. Logistic Regression: The first model used for analysis was a traditional logistic regression model. Logistic regression functions well in the context of an unbalanced outcome therefore only the full training dataset was used to develop the model. The R function glm with a binomial family was used for logistic regression analysis.
2. Least Absolute Shrinkage and Selection Operator (LASSO) Regression with Two-Way Interaction Genetic Learning Variable Selection and Logistic Regression. Each of these functions well in the context of an unbalanced outcome therefore only the full training dataset was used to develop the model. The R packages lars and glmulti were used for LASSO regression and genetic learning, respectively, while the function glm with a binomial family was used for logistic regression analysis. LASSO regression was used to identify a subset of variables that best predict mortality for use in the genetic learning model. The genetic algorithm creates many candidate logistic regression models, using subsets of all variables as well as all possible 2-way interactions between the variables. Using model selection criteria (specifically, the Akaike Information Criterion), after several hundred "generations" of the model building, the best-fit model is identified. This approach allows for identification of a model with interactions between variables that would not necessarily be identified through traditional means, leading to a model that is more robust. Since the genetic algorithm has a limit on the number of variables one can utilize to check for all possible interactions, the LASSO approach allowed us to identify

the best subset. Once the final model was identified, traditional logistic regression was used to estimate mortality predictions.

3. Random Forest: The Random Forest algorithm is an ensemble learning algorithm, which takes random samples of both cases (with replacement) as well as variables from the list of candidate variables (listed previously) and creates many decision trees from them. Each tree makes a prediction as to if the patient would be a survivor or non-survivor and the results of all samples and all trees are aggregated together to make the final prediction. For each model 500 trees were created, while a sample of seven variables were selected for each tree and 63.2% of the cases are used (with replacement) for each tree. Random Forest analysis was used in the full training set, as well as in the down sampled, up sampled, and SMOTE sampled sets. The R package randomForest was used for random forest analysis.
4. Recursive Partitioning Tree: Recursive partitioning trees are decision algorithms that provide a tree-like decision rule to classify patients into the outcome. Splits are made on variables as necessary in order to arrive at the best classification with the least error and ends when the sample size at a split is small or until no improvement in error is made. In our implementation, the splits are based on an information criterion known as the Gini Index. Recursive partitioning analysis was used in the full training set, as well as in the down sampled, up sampled, and SMOTE sampled sets. The R package rpart was used for recursive partitioning analysis.
5. Conditional Inference Tree: Conditional inference trees are similar to the recursive partitioning trees described above with the exception that the binary splits between the independent variables are chosen based on splits being statistically significant (based on Bonferonni-corrected P-values), as opposed to maximization of the information criterion selected (e.g. Gini Index). Conditional inference tree analysis was used in the full training set, as well as in the down sampled, up sampled, and SMOTE sampled sets. The R package party was used for conditional inference tree analysis.
6. Naïve Bayes: The Naïve Bayes classification algorithm uses Bayes' Theorem to compute the conditional probability of the outcome given the complete set of independent variables supplied to the model (described previously). Given the likelihood of the variables occurring with and without the outcome, as well as the prior probabilities of the outcome, a prediction can be derived as to if a particular case would survive or not survive, providing a model for predicting mortality. Naïve Bayes analysis was used in the full training set, as well as in the down sampled, up sampled, and SMOTE sampled sets. The R package e1071 was used for Naïve Bayes analysis.

For each model, several statistics were calculated to provide measures of the overall performance (e.g. predictive ability) of the

Table 1 Performance measures of each model in the testing dataset.

Model	Percent Correct Prediction Alive	Percent Correct Prediction Dead	Accuracy	Balanced Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
Logistic Regression	11.7%	44.4%	13.4%	40.9%	2.7%	79.1%	44.4%	11.7%	83.2%
LASSO/Genetic Regression	59.5%	12.3%	14.9%	43.9%	3.7%	84.2%	59.4%	12.3%	81.3%
Random Forest									
Actual Sample	12.6%	15.4%	12.7%	44.4%	0.0%	88.5%	15.4%	12.6%	85.3%
Up-sampled minority class	11.7%	13.7%	11.7%	40.4%	0.1%	80.2%	13.6%	11.6%	85.0%
Down-sampled minority class	4.4%	69.8%	28.8%	25.0%	30.2%	19.8%	69.8%	4.4%	83.8%
SMOTE-sampled	4.7%	66.4%	24.5%	23.9%	24.8%	22.9%	66.4%	4.7%	83.2%
Recursive Partitioning Tree									
Actual Sample	52.8%	88.1%	86.3%	58.5%	97.1%	19.8%	88.1%	52.8%	66.7%
Up-sampled minority class	26.7%	94.0%	67.9%	70.4%	66.9%	74.0%	94.0%	26.7%	75.2%
Down-sampled minority class	29.4%	92.8%	73.3%	69.6%	74.7%	64.5%	92.8%	29.4%	70.2%
SMOTE-sampled	28.2%	94.8%	69.3%	72.6%	68.1%	77.1%	94.8%	28.2%	74.3%
Conditional Inference Tree									
Actual Sample	56.5%	87.5%	86.4%	55.9%	98.3%	13.5%	87.5%	56.5%	80.5%
Up-sampled minority class	29.1%	91.7%	74.5%	67.3%	77.7%	57.3%	91.7%	29.1%	70.3%
Down-sampled minority class	28.5%	93.9%	70.8%	71.3%	70.6%	71.9%	93.9%	28.5%	77.3%
SMOTE-sampled	24.4%	92.3%	66.7%	66.3%	66.9%	65.6%	92.3%	24.4%	73.8%
Naïve Bayes Classification									
Actual Sample	40.9%	94.4%	81.6%	76.6%	83.5%	69.8%	94.4%	40.9%	79.9%
Up-sampled minority class	32.8%	95.8%	74.2%	76.7%	73.2%	80.2%	95.8%	32.8%	80.0%
Down-sampled minority class	39.5%	94.2%	80.9%	75.8%	82.9%	68.8%	94.2%	39.5%	83.1%
SMOTE-sampled	36.2%	94.4%	78.4%	75.2%	79.6%	70.8%	94.4%	36.2%	81.7%

model in the testing dataset (e.g. clinical practice). The following statistics were calculated: 1) Percent of survivors correctly predicted as survivors, 2) Percent of non-survivors correctly predicted as non-survivors, 3) Accuracy, 4) Balanced Accuracy, 5) Sensitivity, 6) Specificity, 7) Positive Predictive Value, 8) Negative Predictive Value, 9) Area Under the Receiver Operating Characteristic Curve (AUC).

3 Results

A total of 3249 unique hospitalized patients with CAP were enrolled in the study, 2743 were included in the model building (training) dataset, while the remaining 686 were included in the testing dataset for which performance metrics are reported. From the full population, death at 30-days post discharge was documented in 458 (13.4%) patients. This proportion remained consistent in the training and testing datasets (13.3% and 13.8%, respectively).

Performance metrics for each of the models' performance on the testing dataset can be found in **Table 1**. A visual representation of the performance statistics can be found in **Figure 1**. Each model and each sample resulted in different performance with respect to predicting 30-day mortality. Overall, LASSO Regression Genetic Learning Variable Selection and Logistic Regression performed the best in predicting survivors, but was also the worst at predicting non-survivors. Naïve Bayes Classification in an up-sampled dataset had the best prediction of non-survivors, but only predicted one third of survivors correctly. When evaluating sensitivity and specificity together, Naïve Bayes classification algorithms had the best predictive power overall, and appeared the most powerful prediction algorithm across the majority of statistical metrics calculated in the actual sample as well as in each of the up/down and SMOTE sampled sets.

4 Discussion

This study suggests that some machine learning algorithms perform better than traditional statistical modeling approaches when predicting 30-day mortality in hospitalized patients with CAP. However, none of the models or samples assessed offered overall accurate predictions of patient-level mortality and each exhibited a wide variation in performance based on the measure utilized.

The majority of studies evaluating prediction of mortality in hospitalized patients with CAP use the AUC as the sole measure of the performance or accuracy of the model or score. The AUC is a measure of the area under a curve defined by the continuous sensitivity versus false-positive rate of a variable. Although this statistic is widely reported, it can be misinterpreted in several contexts. Since the AUC is only evaluating one overall aspect of model accuracy (Sensitivity vs False Positive Rate), it can be misleading particularly in the context of an unbalanced outcome. When the outcome is relatively rare or unbalanced as was mortality in our dataset, predicting that a particular patient will survive will more often prove to be correct just by chance since there were far more survivors than non-survivors. In this context, the AUC may be high even though the model cannot predict the non-survivor class accurately. Furthermore, since the AUC is not penalized by the number of predictor variables, it will be biased toward models with more variables, regardless of how well those models actually predict the outcome and how well those models actually fit the data. These scenarios make it difficult to properly evaluate a predictive model without several measures of accuracy. Therefore, an important implication of our findings is that over-reliance on the AUC may lead to faulty predictive models. It may be better to evaluate modeling strategies across a wide array of performance statistics.

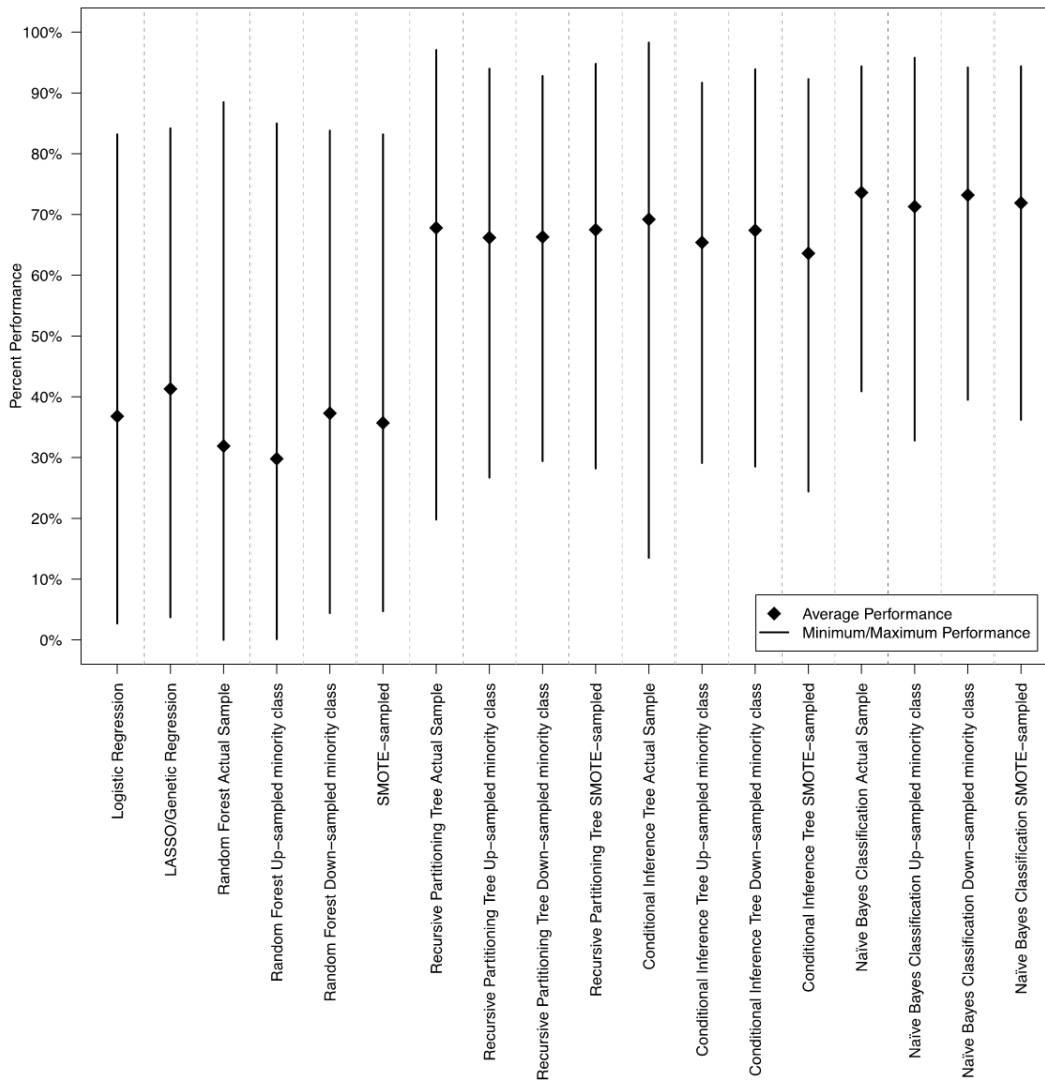


Fig. 1 Model performance for predicting 30-day mortality in hospitalized patients with community-acquired pneumonia.

One potential reason explaining the lack of overall accurate predictions in any of the models is that we may have reached the maximum predictive power of demographic, medical/social and basic laboratory variables. Using recent standard of care laboratory values and biomarkers such as brain natriuretic peptide (BNP), C-reactive protein, and procalcitonin, or novel "research only" laboratory values such as cytokines may enhance the accuracy for predicting mortality in hospitalized patients with CAP. Recently, investigators have begun to evaluate biomarkers as potential predictors of various outcomes in these patients with some success⁹⁻¹³.

Our study reports similar AUC values for predicting mortality in hospitalized patients with CAP compared to the original publication of the Pneumonia Severity Index across almost all models evaluated³. However, other investigators have reported lower AUCs which are in agreement with the lower performing models in our study¹⁴. As previously discussed, reliance on the AUC for model performance, particular in the context of an unbalanced outcome, can lead to substantial error in prediction of patients who do not survive.

Our study has several limitations. First, since there is no gold standard definition of CAP, it is possible our sample has misclassified patients. Second, there are many statistical and machine learning models as well as many other methods for balancing the outcomes of datasets for machine learning models. It is possible that some combination of models and/or samples we did not assess provide accurate predictions of 30-day mortality.

The primary goal of future research in the field of predicting outcomes in hospitalized patients with CAP should be to create models that can predict the clinical outcome at an individual patient level. Machine learning models may provide enhanced ability to do this, particularly in light of using a large number of novel variables such as biomarkers and cytokines.

In conclusion, this study suggests that accurate patient-level prediction of 30-day mortality in hospitalized patients with CAP is difficult with statistical and machine learning approaches. It will be important to evaluate novel variables and other modeling approaches to better predict poor clinical outcomes in these patients to ensure early and appropriate interventions are instituted.

Role of the Funding Source: This study was not funded.

Conflict of Interest: None of the authors have any conflicts of interest to report.

References

- 1 World Health Organization, "The top 10 causes of death," 2015 [cited 2017]. [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs310/en/>
- 2 M. J. Fine, T. E. Auble, D. M. Yealy, B. H. Hanusa, L. A. Weissfeld, D. E. Singer, C. M. Coley, T. J. Marrie, and W. N. Kapoor, "A prediction rule to identify low-risk patients with community-acquired pneumonia," *New England journal of medicine*, vol. 336, no. 4, pp. 243-250, 1997.
- 3 M. J. Fine, L. J. Hough, A. R. Medsger, Y.-H. Li, E. M. Ricci, D. E. Singer, T. J. Marrie, C. M. Coley, M. B. Walsh, M. Karpf *et al.*, "The hospital admission decision for patients with community-acquired pneumonia: results from the pneumonia patient outcomes research team cohort study," *Archives of Internal Medicine*, vol. 157, no. 1, pp. 36-44, 1997.
- 4 T. Wiemken, R. Kelley, and J. Ramirez, "Clinical scoring tools: which is best to predict clinical response and long-term outcomes?" *Infectious disease clinics of North America*, vol. 27, no. 1, pp. 33-48, 2013.
- 5 G. F. Cooper, V. Abraham, C. F. Aliferis, J. M. Aronis, B. G. Buchanan, R. Caruana, M. J. Fine, J. E. Janosky, G. Livingston, T. Mitchell *et al.*, "Predicting dire outcomes of patients with community acquired pneumonia," *Journal of biomedical informatics*, vol. 38, no. 5, pp. 347-366, 2005.
- 6 G. F. Cooper, C. F. Aliferis, R. Ambrosino, J. Aronis, B. G. Buchanan, R. Caruana, M. J. Fine, C. Glymour, G. Gordon, B. H. Hanusa *et al.*, "An evaluation of machine-learning methods for predicting pneumonia mortality," *Artificial intelligence in medicine*, vol. 9, no. 2, pp. 107-138, 1997.
- 7 C. Wu, R. Rosenfeld, and G. Clermont, "Using data-driven rules to predict mortality in severe community acquired pneumonia," *PLoS One*, vol. 9, no. 4, p. e89053, 2014.
- 8 N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321-357, 2002.
- 9 M. Alan, E. Grolimund, A. Kutz, M. Christ-Crain, R. Thomann, C. Falconnier, C. Hoess, C. Henzen, W. Zimmerli, B. Mueller *et al.*, "Clinical risk scores and blood biomarkers as predictors of long-term outcome in patients with community-acquired pneumonia: a 6-year prospective follow-up study," *Journal of internal medicine*, vol. 278, no. 2, pp. 174-184, 2015.
- 10 R. Menéndez, R. Martinez, S. Reyes, J. Mensa, X. Filella, M. A. Marcos, A. Martinez, C. Esquinas, P. Ramirez, and A. Torres, "Biomarkers improve mortality prediction by prognostic scales in community-acquired pneumonia," *Thorax*, vol. 64, no. 7, pp. 587-591, 2009.
- 11 P. Ramírez, M. Ferrer, V. Martí, S. Reyes, R. Martínez, R. Menéndez, S. Ewig, and A. Torres, "Inflammatory biomarkers and prediction for intensive care unit admission in severe community-acquired pneumonia," *Critical care medicine*, vol. 39, no. 10, pp. 2211-2217, 2011.
- 12 A. Torres, P. Ramirez, B. Montull, and R. Menéndez, "Biomarkers and community-acquired pneumonia: tailoring management with biological data," in *Seminars in respiratory and critical care medicine*, vol. 33, no. 03. Thieme Medical Publishers, 2012, pp. 266-271.
- 13 S. Yende, G. D'angelo, J. A. Kellum, L. Weissfeld, J. Fine, R. D. Welch, L. Kong, M. Carter, and D. C. Angus, "Inflammatory markers at hospital discharge predict subsequent mortality after pneumonia and sepsis," *American journal of respiratory and critical care medicine*, vol. 177, no. 11, pp. 1242-1247, 2008.
- 14 M. R. Ananda-Rajah, P. G. Charles, S. Melvani, L. L. Burrell,

P. D. Johnson, and M. Lindsay Grayson, "Comparing the pneumonia severity index with CURB-65 in patients admitted with

community acquired pneumonia," *Scandinavian journal of infectious diseases*, vol. 40, no. 4, pp. 293–300, 2008.