EVALUATION OF PREDICTIVE ANALYTIC TECHNIQUES IN HEALTHCARE RESEARCH: A PRSIMA STYLE REVIEW

Hayden Wimmer, Georgia Southern University, hwimmer@georgiasouthern.edu
Carl M. Rebman, Jr., University of San Diego, carlr@sandiego.edu
Queen E. Booker, Minnesota State University, Mankato, queen.booker@mnsu.edu

ABSTRACT

Preventing hospital admissions and readmission has the potential to reduce healthcare costs nationwide. Disease and readmission prevention can be assisted by applying data science and predictive analytics to healthcare data. This paper presents a PRISMA style literature review of pneumonia readmissions based on the Medline, Healthsource Academic, and CINHAL databases. These databases were searched for articles that contained 'pneumonia' and 'readmission' in the titles. While disease and readmission predication are well represented in the literature, the application of more advanced data science techniques is under represented. Regression appeared to be the most dominant technique applied and future research studies should study more data science and predictive analytic approaches. Adopting analytical techniques can help make more robust and precise analysis and can aid in the classroom instruction of data analytics and health informatics.

Keywords: Predictive Analytics, Data Science, PRISMA, Readmission, Pneumonia, Health Informatics

INTRODUCTION

Hospital readmission is a critical issue facing healthcare providers, payers, and patients. Pay for performance measures have been shown to reduce costs and improve care (Hsieh et al, 2015). Specifically, readmission within 30 days for some conditions results in reduced reimbursement from payers such as Medicare. The readmissions reduction program by the Center for Medicare and Medicaid services, or CMS, defines payment adjustments and measures and defines reports that are required and patients who are covered (Medicare.gov, 2017). In fact, Medicare fined 2610 hospitals in the third round of readmission penalties (Rau, 2014). One such condition, pneumonia, has growing interest from the academic community; however, studies and methods for readmission detection in pneumonia are underrepresented in the current state of the literature.

Disease prediction using machine learning and data mining techniques has been studied using techniques such as support vector machine (Stiglic et al, 2012) decision trees and neural networks (Shadmi et al, 2015). Predictive approaches have garnered interest with the advent of the pay-for-performance incentive design model. Specifically, 30-day readmissions criteria are targets for research; however, the literature lacks studies using predictive analytics approaches. Standard statistical approaches such as t-tests (Tang et al, 2014), univariate analysis (Hebert et al, 2014), logistic regression (Mather et al, 2013; Shorr et al, 2013), and analysis of variance (Ranasignhe et al, 2014) are represented. However, machine learning and data science approaches are lacking.

PRISMA REVIEW PROCESS

This review follows the PRISMA method (Moher et al, 2010) in identifying literature and systematically conducting the review. The PRISMA review process differs from meta-analysis approach in that it does not review the statistical results to determine a quantitate summary, effect size, or other data mining techniques to investigate the analytical methods used. It is also commonly utilized and accepted in the medical profession. In order to focus the results, specific terms were employed in the search and databases searched restricted to medical databases. Medline, CINHAL, and Health Source Academic were searched for pneumonia and readmission in the title which resulted in 31 total articles. Of the 31, 11 were duplicated resulting in 20 unique articles. Overlap occurred by having 3 articles occurring

in all 3 databases, 5 articles shared between Healthsource and Medline, 0 shared between Healthsource and CINHAL, and 1 shared between Medline and CINHAL. One article was removed and labeled as irrelevant as neither of the search terms were included in the title and therefore, the article should not have been returned. The article by Goetz et al (1988), was titled "Surgical Complications Related to Insertion of Penile Prostheses with Emphasis on Infection and Cost". After an abstract and full text analysis of the article, it was deemed irrelevant to the current literature search. Figure 1 shows the PRISMA flow of included studies and Table 1 lists a summary of included studies.
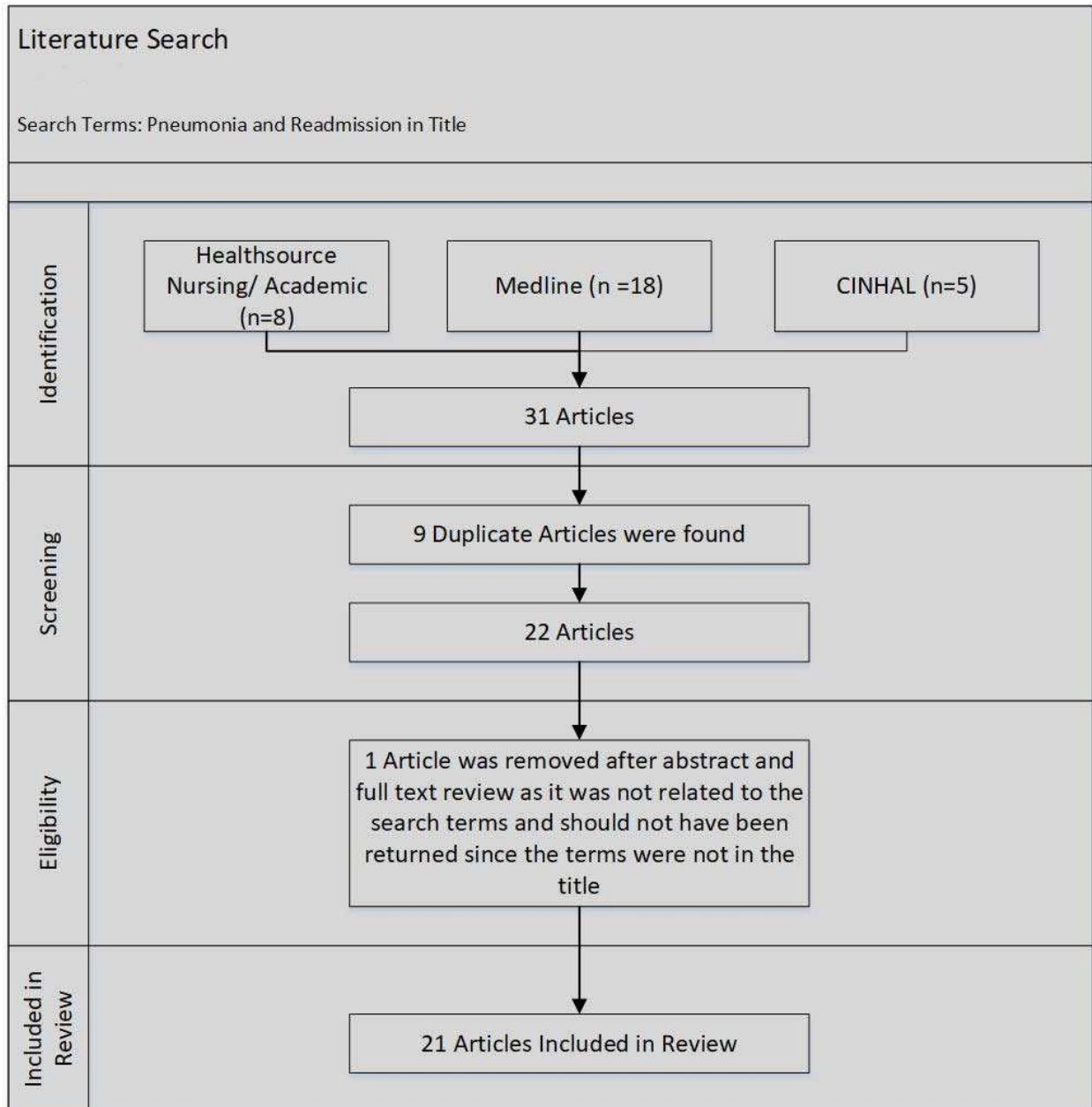
Figure 1. PRISMA Flow of Included Studies

LITERATURE REVIEW

The literature in this section is organized by the database and order from which it was returned during the database search process. Duplications are only described in a single section, Medline, as there was overlap between Healthsource and Medline, CINHAL and Medline, but no overlap between CINHAL and Healthsource. The overlap is indicated accompanying table for summarizing each article review.

Medline

Sjoding et al. (2015) identified the potential for gaming readmission rates via medical coding. Specifically, pneumonia readmission, in lieu of coding as pneumonia, may be coded as respiratory failure thereby evading the reimbursement penalty for readmission. The authors sampled 100 hospitals with pneumonia mortality rates greater than 50% and recoded eligible patients to sepsis or respiratory failure thereby reducing the mortality and readmission rates of both readmission and mortality. Recommendations were made to the CMS to change the calculations for hospital readmission and mortality to prevent hospitals from gaming the system.

Hilbert et al. (2014) identified the lack of readily available tools for determining readmission criteria for acute myocardial infraction, heart failure, and pneumonia. The authors employed R, a freely available statistical package, to create decision trees based on structured discharge data. To demonstrate the method, decision trees are tested on California's State Inpatient databases. Results showed that decision trees may not perform as well as other models; however, the authors state the added transparency of decision trees are well worth the tradeoff in performance.

Andruska et al. (2014) evaluated 9624 patients discharged from pneumonia admissions and found that 46.1% were classified with a culture-negative pneumonia, 20.2% with antibiotic resistant pneumonia, 31.1% with pneumonia caused by antibiotic resistant bacteria, and 2.7 having a viral form of pneumonia. Multiple logistic regression was conducted and results indicate that pneumonia attributed to antibiotic resistant strains increased risk for readmission.

Shorr et al (2013). followed a cohort of 977 admitted subjects with evidence of bacterial pneumonia infection where 78.9 survive to discharge. Healthcare associated pneumonia (HCAP) was 7.5% more likely to suffer readmissions. Readmission criteria is coupled with pneumonia type and variables with which the hospital has no control. The authors state that using rules to guide payment that fail to account for HCAP are inappropriate.

Steer et al. (2012) collected clinical and demographic data on 924 patients hospitalized with acute exacerbations of chronic obstructive pulmonary disease (AECOPD). The population experienced 10.4% mortality rate while admitted and a 19.1% readmission rate of survivors. Patients with evidence of pneumonia had significantly higher mortality rates of 20.1% compared to 5.8% of non-pneumonic admissions.

Lindenauer et al. (2010) investigated the development, validation and results of a risk-standardized measure of 30-day hospital readmission rates of pneumonia by studying a cohort of Medicare beneficiaries over 65 years of age in 4675 hospitals across the United States which included 226,545 total hospitalizations. A hierarchical logistic regression was employed and model performance measured by area under the receiver operating curve (ROC). While risk-standardized readmission rates varied among hospitals, readmission rates derived from administrative claims are similar to those based on medical records review.

Lindenauer et al. (2011) studied patterns of hospital and regional performance of 4813 nonfederal acute hospitals in the United States based on 30-day readmission and mortality rates of Medicare beneficiaries age 65 or older. In the 3 year analysis of 1.1 million subjects, 11.6% of patients in the study die within 30 days of hospital admission and 18.3% were readmitted within 30 days of discharge. The study was descriptive in nature suggesting variation among hospitals and regions and the authors claim the outcome of pneumonia patients is influenced by both hospital and region.

Eurich et al (2010). investigated whether the presence of dysglycaemia at admission is correlated with adverse effects in 90 days and 1 year in patients with community-acquired pneumonia by collecting data from 6 hospitals on 2366 adults. Contrary to increasing evidence, no significant relationship existed between the presence of dysglycaemia and morbidity or mortality at both 90 days and 1 year suggesting that dysglycaemia is a maker of pneumonia as opposed to a target for reducing mortality or readmission.

Schroeder (2009) described CMS reporting the 30-day readmission measures for acute myocardial infarction, heart failure, and pneumonia. The aforementioned conditions are listed as the most common among Medicare beneficiaries with pneumonia being the second most common hospital discharge diagnosis and has a high prevalence of readmission. The relationship between readmission and quality of care is defined according to the relevant literature. Finally, the author discussed how 30-day readmission will be used to evaluate hospital performance and assign hospitals to groups based on performance.

Bohannon et al. (2004) followed 153 patients following successful discharge for pneumonia and 62.6% had died or were readmitted within 1 year of discharge. Patients were at least 65 years of age, admitted directly to the pulmonary unit for diagnosed pneumonia, and survived the hospitalization to discharge. 83% of the patients had at least 1 comorbidity including CPOD, diabetes mellitus, congestive heart failure, cerebrovascular disease, cancer, and renal disease. Authors suggested the outcomes may be based on the comorbidities and physical therapy and counseling may prove successful to lower the mortality and readmission rates.

Herzog et al. (2003) studied the effect of influenza vaccine on mortality and readmission rates of elderly patients admitted for pneumonia via randomly reviewing the medical records of 12,566 Medicare beneficiaries hospitalized for pneumonia. Subjects were grouped into either a) vaccinated prior to hospitalization, b) during hospitalization, and c) unknown vaccination. The study demonstrated that patients vaccinated prior to hospitalization had significantly lower mortality and readmission rates than those with an unknown vaccination status but no significant difference between vaccination during hospitalization and unknown vaccination. The results suggested that influenza vaccine before hospital admission may lower readmission and mortality rates in the elderly.

Health Source Academic/ Nursing

Neuman et al. (2014) studied readmission of children following pneumonia hospitalization in 43 hospitals with 82,566 subjects. Results show that 30-day readmission rates reach 3.1%. Younger children (< 1 year of age) experienced higher readmission rates as did children with chronic comorbidities. Additional factors affecting readmission include complicated pneumonia or treatment at hospitals with a low volume of pneumonia admissions.

Dharmarajan et al. (2013) analyzed 2007-2009 Medicare fee-for-service claims to identify patterns for 30-day readmission following heart failure, acute myocardial infarction, and pneumonia. The authors showed that the majority of readmission occur within the first 15 days of discharge. A relationship was not found between readmission and demographics, namely age, sex, and race. The authors suggested that additional attention within the first 15 days may lower the 30-day readmission rate.

CINHAL

Fisher and Swartz (2014) explained that COPD and acute ischemic stroke were included in the CMS readmission program as of 2012 and implemented a program to reduce CPOD and pneumonia readmissions by 10%. The study revealed inconsistent patterns due to a lack of consistent evidence-based care; however, the program addresses this concern. Following implementation, CPOD readmissions were reduced to 8.9% from 14.4% and pneumonia readmission reduced to 11.03% from 14.47%. Coordinating a system wide effort including consistent evidence-based care strengthened outpatient follow-up thereby reducing readmissions.

Medline, Health Source, and CINHAL

Cabré et al. (2013) followed 2359 patients for an average of 24 months to determine if oropharyngeal dysphagia is a risk factor for readmission for pneumonia in the elderly. Of those readmitted (7.9%), 24.2% had aspirational pneumonia. A multivariate Cox regression showed that oropharyngeal dysphagia is a prevalent risk factor for readmission in the elderly for both aspirational and non-aspirational pneumonia.

Krumholz et al. (2013) studied the relationship between 30-day readmission and mortality rates for patients with acute myocardial infarction (AMI), heart failure (HF), and pneumonia. Correlations between readmission and mortality was 0.03 for AMI, -0.17 for HF, and 0.002 for pneumonia indicated a weak relationship for heart failure but no relationship between AMI or pneumonia.

Medline and Health Source

Suter et al (2014). performed a cross-sectional analysis of hospital readmission for acute myocardial infarction (AMI), heart failure (HF), and pneumonia to identify national patterns of readmission and mortality. Authors demonstrated that between 2009 and 2012, improvement in all annual risk-standardized outcome rates. The results indicate that the CMS reimbursement criteria may be having a positive effect on reducing readmission and mortality in AMI, HF, and pneumonia.

Calvillo–King et al (2013). performed a systematic review of social factors on risk of readmission or mortality in community acquired pneumonia (CAP) and heart failure (HF). OVID, PubMed, and PSYCHINFO was searched from 1980 through 2012 returning 630 records of which 360 were excluded. Further refinement of the results led to 72 articles included in the review. Factors that appeared frequently were age, gender and race. Education, income, and unemployment were less represented but shown to impact readmission following CAP. Other factors for readmission and mortality included socioeconomic status, living situation, lack of social support, and risk behaviors.

Palepu et al. (2003) extended the literature by straying from studying the elderly and study readmission among the HIV-infected patients with bacterial pneumonia. Multivariate logistic regression was employed to identify factors associated with readmission. Results indicated that the most prevalent factors for readmission in the HIV-infected population are leaving the hospital against medical advice, living in a poor urban environment, and being hospitalized in the 6 months prior to and during the summer months and suggest social work as an effective mediator of readmission in this population.

Grant et al. (1999) also studied readmission factors among the AID population hospitalized with bacterial pneumonia. Results from this study included increased risk for readmission if patients left the hospital unaccompanied by family or a friend, used crack cocaine, and one of more coincident AIDS diagnoses, or had been admitted in the preceding 6 months. Other demographic factors such as alcoholism or intravenous drug use were not successful readmission predictors.

Medline and CINHAL

Schmeida and Savrin (2012) employed multivariate regression with 50 state level data samples to determine factors associated with higher than national average readmission rates. Results revealed that incorrect initial antibiotic and lack of discharge information given to patients is associated with a 30-day readmission rate greater than the national average. Additionally, factors that were previously reported by the CMS to lower readmission, were shown to increase the chances a hospital would have worse than average readmission rates. These factors included higher percentage of white Medicare enrollees, higher median income, more total days of care, and more Medicare enrollees with prescription drug coverage. Recommendations included proper discharge patient preparedness and correct initial antibiotic administration and that hospitals should not overlook non-minority and higher income patients.

Table 1. Summary of Studies

Study	Source	Synopsis	Primary Method(s)
Sjoding, et al, 2015	Medline	Test possibility of recoding pneumonia mortality and readmission and demonstrates the current reimbursement calculations can be gamed via recoding.	Descriptive/ Review
Hilbert et al, 2014	Medline	Applies decision trees to California's Inpatient Database and demonstrates, while model performance is not as well as more sophisticated tools, added transparency is worth the reduced performance.	Decision Trees in R

Study	Source	Synopsis	Primary Method(s)
Andruska et al, 2014	Medline	Studies 9624 patients discharged from pneumonia admissions and determined that antibiotic resistant strains of bacteria as the initial cause of pneumonia increased readmission risk.	T-Test, U-Test, Logistic Regression
Shorr et al, 2013	Medline	Follows 977 subjects on determines that healthcare associated pneumonia (HCAP) is 7.5% more likely to suffer readmission and suggests that formulae that guides payment that does not account for HCAP is inappropriate.	T-Test, U-Test, Logistic Regression
Steer et al, 2012	Medline	Pneumonia was a strong predictor of readmission for patients admitted with acute exacerbations of chronic obstructive pulmonary disease.	X^2, T-Test, U-Test
Lindenauer et al, 2011	Medline	Studies 30 day readmission criteria where Medicare beneficiary age > 65 and determines that risk-standardized readmission rates varied among hospitals,	Linear Regression
Lindenauer et al, 2010	Medline	Describes patterns of hospital and regional performance on 30-day readmission and mortality rates of Medicare beneficiaries age 65 or older.	GLIMMIX by SAS (General Linear Mixed Model)
Eurich et al, 2010	Medline	Studies the relationship between presence of dysglycaemia at admissions and 90 day and 1 year readmission and mortality rates and determined a relationship does not exist suggesting dysglycaemia is likely a cause of community-acquired pneumonia.	Cox, Multivariate
Bohannon et al, 2004	Medline	Studies patients 65 or older admitted for pneumonia and determined comorbidities were common among this population and suggests that physical therapy and counseling may serve as appropriate interventions to reduce 1 year mortality and readmission following pneumonia.	Pearson Correlation, Logistic Regression
Herzog et al, 2003	Medline	Studies influenza vaccination on readmission and mortality of pneumonia and determines that vaccination prior to hospitalization leads to a lower rate of mortality and readmission.	CHI Squared, Cox, ANOVA
Neuman et al, 2014	Healthsource Academic	Studies children 30 day pneumonia readmission and determines that children less than 1 year in age or children with comorbidities are most likely to experience readmission following discharge for pneumonia.	Logistic Regression
Dharmarajan et al, 2013	Healthsource Academic	Studies Medicare fee-for-service claims data for AMI, HF, and pneumonia then determines the majority of 30 day readmissions occurs within the first 15 days; therefore, additional attention following discharge may lower 30 day readmission rates.	Cox, Logistic Regression

Study	Source	Synopsis	Primary Method(s)
Fisher et al, 2014	CINHAL	Reports the results of implementing a comprehensive program including consistent evidence-based care which reduced readmissions in both COPD and pneumonia.	Descriptive/ Review
Cabre et al, 2013	Medline, Healthsource Academic, CINHAL	Follows 2359 elderly patients for 2 years and determines that oropharyngeal dysphagia is a relevant risk factor in readmission of both aspirational and non-aspirational pneumonia.	Cox, Regression
Krumholz et al, 2013	Medline, Healthsource Academic, CHINAL	Studied correlations between risk-standardized mortality rates and risk-standardized readmission rates and found no correlation for acute myocardial infarction and pneumonia and a weak relationship with heart failure.	Hierarchical Logistic Regression
Suter et al, 2014	Medline, Healthsource Academic	Study readmission and mortality from 2009-2012 and demonstrate the rates are on the decline suggesting the payment incentive programs by the CMS are having the desired effect of reducing mortality and readmission of heart failure, acute myocardial infarction, and pneumonia.	Hierarchical Logistic Regression
Calvillo-King et al, 2013	Medline, Healthsource Academic	Systematic review from 1980-2012 that identifies factors for mortality and readmission including demographics as well as education, income, living situation, and socioeconomic factors to name a few.	Descriptive/ Review
Palepu et al, 2003	Medline, Health Source Academic	Studies readmission following pneumonia of HIV-infected patents and determines that leaving the hospital against medical advice, living in a poor urban environment, and time frame for hospitalization as the primary contributing factors for readmission and suggests social work as a possible intervention.	Wilcox, Logistic Regression
Grant et al, 1999	Medline, Healthsource Academic	Studies readmission from hospitalization of pneumonia among the AIDS population and determined that leaving the hospital alone, use of crack cocaine, coincident AIDS diagnoses, and prior admission within 6 months were related to readmission; however, other demographic factors such as IV drug use and alcoholism were not determining factors.	X^2,Wilcox, Logistic Regression
Schmeida et al, 2012	Medline, CINHAL	Determined that correct initial antibiotic administration and properly preparing patients for discharge are key variables in readmission prevention and that hospitals should not overlook non-minority and higher income populations as, against CMS reports, show high readmission potential.	Multivariate Regression

RESEARCH DISCUSSION

Incentive programs have been demonstrated effective (Suter et al, 2014); on the other hand, recoding can be employed to game the system (Sjoding et al, 2015). Medicare fee-for-service beneficiaries where the primary targets for 30-day pneumonia readmission studies. This is likely due to the financial incentive involved with preventing readmission. Authors did study other populations besides the elderly or patients 65 and older with examples as children (Neuman et al, 2014), HIV and AIDs infected patients (Palepu et al, 2003; Grant et al, 1999), and adult patients over 18 years of age (Shorr et al, 2013).

The literature indicates there is a mix of demographic (i.e. income) and cultural factors (i.e. living arrangements) with clinical factors (i.e. initial medication, presence of comorbidities). In order to address the demographic, cultural, and clinical factors, implementing comprehensive programs that address aftercare as well as utilize evidence-based medicine have proven effective (Fisher et al, 2014). The literature likewise finds that analyzing administrative claims produced similar results to medical record review highlighting the need to examine additional dimensions to reduce hospital readmission for pneumonia.

The literature also indicates that most analysis and classification is performed with standard statistical methodologies such as logistic or multivariate regression. Only one study, (Hilbert et al, 2014), employed data science techniques outside of statistical analysis methodologies, namely classification and regression trees (CART) inducted using the R statistical package. There was did not seem to be any attempt to study other decision tree induction algorithms such as ID3 (Quinlan, 1986; Cheng et al, 1988) or C4.5 (Quinlan, 1993). Other data science and machine learning techniques were also omitted from the literature on 30-day readmission of pneumonia. Besides decision trees, neural networks, support vector machines, Bayesian classifiers, k-means clustering, association rule mining, and text analytics may be employed to learn knowledge from data in order to make predictions about 30-day readmission factors. Similarly, no studies reviewed seek to study big data processing likely due to the sizes of the datasets; nevertheless, medical data is growing exponentially creating a demand for big data algorithms and processing on medical datasets.

The lack of machine learning and data science techniques demonstrates the need for future research to integrate data science and predictive analytics into healthcare, specifically 30-day pneumonia readmission rates. In addition to data science and analytics approaches to reducing readmissions, clinicians and patients may not have the time nor the resources to keep abreast of the current state of the art in readmission prevention; therefore, combining outreach in the form of knowledge dissemination to clinical workers and patients with data science efforts has the ability to address both the clinical and aftercare dimensions of reducing 30-day pneumonia readmissions.

LIMITATIONS AND CONCLUSION

Hospital readmission of various conditions such as myocardial infarction, COPD, and more recently pneumonia is of concern. Payers, such as Medicare, withholds reimbursement in 30-day readmission cases. In this PRISMA style review, we examined techniques used to predict pneumonia readmission by searching for "Pneumonia" and "Readmission" in the title of articles published in Medline, Healthsource academic, and CINHAL. We discovered the dominant method employed for prediction appears to be regression analysis with little support for data science and predictive analytics approaches. Perhaps this because the techniques utilized by researchers were the most familiar or because they were the best bit or best known.

There are some limitations of this paper which indicate the need for future research. First, there might be other articles that did not appear in our literature review. To address these concerns, additional databases may be searched for literature not covered by this work. Second, there are limitations of the PRISMA approach as it does not employ data mining or other analytic techniques. This may be addressed by using text mining, sentiment analysis, or topic modeling techniques. Finally, there might have been some articles not included in the database or not published when this research was conducted. Regardless of limitations, this work can serve as a step in identifying the need for more advanced techniques in clinical research.

The intent of this study was to analyze what analytical techniques had been used in previous research. It would be interesting if the data from previous studies were reinvestigated using new data science and predictive analytics methods on predicting hospital readmission. Again, based on our review close to over 50% utilized logistical regression. This suggests that current and future researchers could adopt newer analytical techniques could help provide a more robust analysis towards a very pressing problem. Building on past research and data can also assist future students in the realm of data analytics and health informatics as the data would be real, relevant, interesting, and provide a meaningful impact on society.

REFERENCES

Andruska, A, Micek, S.T., Shindo, Y., Hampton, N., Colona, B., McCormick, S., & Kollef, M. H. (2014). Pneumona pathogen characterization is an independent determinant of Hospital Readmission. *CHEST Journal, 148*(1), 103-111.

Bohannon, R. W., Maljanian, R., & Ferullo, J. (2004). Mortality and readmission of the elderly one year after hospitalization for pneumonia. *Aging Clinical and Experimental Research, 16*(1), 22-25.

Cabré, M., Serra-Prat, M., Force, L., Almirall, J., Palomera, E., & Clavé, P. (2013). Oropharyngeal dysphagia is a risk factor for readmission for pneumonia in the very elderly persons: observational prospective study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 5 July, glt099

Calvillo–King, L., Arnold, D., Eubank, K. J., Lo, M., Yunyongying, P., Stieglitz, H., & Halm, E. A. (2013). Impact of social factors on risk of readmission or mortality in pneumonia and heart failure: systematic review. *Journal General Internal Medicine, 28*(2), 269-282.

Cheng, J., Fayyad, U. M., Irani, K. B., & Qian, Z. (1988). Improved decision trees: a generalized version of id3. In *Proceedings of the Fifth International. Conference on Machine Learning* (pp. 100-107).

Dharmarajan, K., Hsieh, A.F., Lin, Z., Bueno, H., Ross, J. S., Horwitz, L. I., Barreto-Filho, J. A., Kim, N., Bernheim, S. M., & Suter, L. G. (2013). Diagnoses and timing of 30-day readmissions after hospitalization for heart failure, acute myocardial infarction, or pneumonia. *JAMA,* 309(4), 355-363.

Eurich, D., Gamble, J., Marrie, T., & Majumdar, S. (2010). Dysglycaemia and 90 day and 1 year risks of death or readmission in patients hospitalised for community-acquired pneumonia. *Diabetologia, 53*(3), 497-503.

Fisher, J., & Swartz, J. (2014). The Impact of a Multi-Disciplinary COPD and Pneumonia Readmission Reduction Program Across the Healthcare Continuum. *Respiratory Care, 59*(10).

Goetz, A., Yu, V., & O'Donnell, W. (1988). Surgical Complications Related to Insertion of Penile Prostheses with Emphasis on Infection and Cost. *Infection and Control Hospital Epidemiology, 9*(6), 250-254.

Grant, R.W., Charlebois, E. D., & Wachter, R. M. (1999). Risk factors for early hospital readmission in patients with AIDS and pneumonia. *Journal of General Internal Medicine, 14*(9), 531-536.

Hebert, C., Shivade, C., Foraker, R., Wasserman, J., Roth, C., Mekhjian, H., Lemeshow, S., & Embi, P. (2014). Diagnosis specific readmission risk prediction using electronic health data: a retrospective cohort study. *BMC Medical Informatics and Decision Making, 14*(1), 65.

Herzog, N. S., Bratzler, D. W., Houck, P. M., Jiang, H., Nsa, W., Shook, C., & Weingarten, S. R. (2003). Effects of previous influenza vaccination on subsequent readmission and mortality in elderly patients hospitalized with pneumonia. *The American Journal of Medicine, 115*(6), 454-461.

Hilbert, J.P., Zasadil, S., Keyser, D. J., & Peele, P. B. (2014). Using Decision Trees to Manage Hospital Readmission Risk for Acute Myocardial Infarction, Heart Failure, and Pneumonia. *Applied Health*

Economics and Health Policy, 12(6), 573-585.

Hsieh, H.-M., Tsai, S.-L., Shin, S.-J., Mau, L.-W., & Chiu, H.-C. (2015). Cost-Effectiveness of Diabetes Pay-for-Performance Incentive Designs. *Medical Care, 53*(2), 106-115.

Lindenauer, P. K., Bernheim, S. M., Grady, J. N., Lin, Z., Wang, Y., Wang, Y., Merrill, A. R., Han, L. F., Rapp, M. T., & Drye, E. E. (2010). The performance of US hospitals as reflected in risk-standardized 30-day mortality and readmission rates for medicare beneficiaries with pneumonia. *Journal of Hospital Medicine, 5*(6), E12-E18

Lindenauer, P. K., Normand, S. L.T., Drye, E. E., Lin, Z., Goodrich, K., Desai, M. M., Bratzler, D. W., O'Donnell, W. J., Metersky, M. L., & Krumholz, H. M. (2011). Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia. *Journal of Hospital Medicine, 6*(3), 142-150.

Krumholz, H. M., Lin, Z., Keenan, P. S., Chen, J., Ross, J. S., Drye, E. E., Bernheim, S. M., Wang, Y., Bradley, E. H., & Han, L. F. (2013). Relationship between hospital readmission and mortality rates for patients hospitalized with acute myocardial infarction, heart failure, or pneumonia. *JAMA*, 309(6), 587-593.

Mather, J. F., Fortunato, G. J., Ash, J. L., Davis, M. J., & Kumar, A. (2013). Prediction of pneumonia 30-day readmissions: a single-center attempt to increase model performance. *Respiratory Care*, respcare-02563.

Medicare.gov https://www.medicare.gov/hospitalcompare/Data/30-day-measures.html Accessed 8 May 2017.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery, 8*(5), 336-341.

Neuman, M. I., Hall, M., Gay, J. C., Blaschke, A. J., Williams, D. J., Parikh, K., Hersh, A. L., Brogan, T. V., Gerber, J. S., & Grijalva, C. G. (2014). Readmissions among children previously hospitalized with pneumonia. *Pediatrics*, peds-2014.

Palepu, A., Sun, H., Kuyper, L., Schechter, M. T., O'Shaughnessy, M. V., & Anis, A. H. (2003). Predictors of Early Hospital Readmission in HIV-infected Patients with Pneumonia. *Journal of General Internal Medicine, 18*(4), 242-247.

Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning, 1*(1), 81-106.

Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.

Ranasinghe, I., Wang, Y., Dharmarajan, K., Hsieh, A. F., Bernheim, S. M., & Krumholz, H. M. (2014). Readmissions after Hospitalization of Heart Failure, Acute Myocardial Infarction, or Penumonia among Young and Middle-Aged Adults: A Retrospective Observational Cohort Study. *PLoS Medicine*, *11*(9), e1001737.

Rau, J. (2014). *A Guide to Medicare's Readmissions Penalties and Data. Kaiser Health News.* http://kaiserhealthnews.org/news/a-guide-to-medicare-readmissions-penalties-and-data /. Accessed May 8, 2017

Schmeida, M., & Savrin, R. A. (2012). Pneumonia rehospitalization of the Medicare fee-for-service patient: a state-level analysis: exploring 30-day readmission factors. *Professional case management, 17*(3), 126-131.

Schroeder, S. (2009). Public reporting of 30-day risk-standardized readmission measures for acute myocardial infarction, heart failure and pneumonia. *South Dakota Medicine: The Journal of the South Dakota State Medical Association, 62*(12), 488.

Shadmi, E., Flaks-Manov, N., Hoshen, M., Goldman, O., Bitterman, H., & Balicer, R. D. (2015). Predicting 30-day readmissions with preadmission electronic health record data. *Medical Care, 53*(3), 283-289.

Shams, I., Ajorlou, S., & Yang, K. (2015). A predictive analytics approach to reducing 30-day avoidable readmissions among patients with heart failure, acute myocardial infarction, pneumonia, or COPD. *Health Care Management Science, 18*(1), 19-34.

Shorr, A. F., Zilberberg, M. D., Reichley, R., Kan, J., Hoban, A., Hoffman, J., Micek, S. T., & Kollef, M. H. (2013). Readmission following hospitalization for pneumonia: the impact of pneumonia type and its implication for hospitals. *Clinical Infectious Diseases,* cit254.

Sjoding, M. W., Iwashyna, T. J., Dimick, J. B., & Cooke, C. R. (2015). Gaming Hospital-Level Pneumonia 30-Day Mortality and Readmission Measures by Legitimate Changes to Diagnostic Coding. *Critical Care Medicine, 43*(5), 989-995.

Steer, J., Norman, E., Afolabi, O., Gibson, G., & Bourke, S. (2012). Dyspnoea severity and pneumonia as predictors of in-hospital mortality and early readmission in acute exacerbations of COPD. *Thorax, 67*(2), 117-121.

Stiglic, G., Pernek, I., Kokol, P., & Obradovic, Z. (2012). Disease prediction based on prior knowledge. In: *Proceedings of the ACM SIGKDD Workshop on Health Informatics, in Conjunction with 18th SIGKDD Conference on Knowledge Discovery and Data Mining*

Suter, L. G., Li, S.-X., Grady, J. N., Lin, Z., Wang, Y., Bhat, K. R., Turkmani, D., Spivack, S. B., Lindenauer, P. K., & Merrill, A. R. (2014). National patterns of risk-standardized mortality and readmission after hospitalization for acute myocardial infarction, heart failure, and pneumonia: update on publicly reported outcomes measures based on the 2013 release. *Journal of General Internal Medicine, 29*(10), 1333-1340.

Tang, V. L., Halm, E. A., Fine, M. J., Johnson, C. S., Anzueto, A., & Mortensen, E. M. (2014). Predictors of rehospitalization after admission for pneumonia in the veterans affairs healthcare system. *Journal of Hospital Medicine, 9*(6), 379-383.