METHODOLOGY

A comparison of methodologies for the real-time identification of hospitalized patients with acute exacerbations of COPD

Parth Shah[1]
Andrew McWilliams[2]
Daniel Howard[3]
Jason Roberge[4]

[1]UNC School of Medicine, Center for Outcomes Research and Evaluation, Atrium Health, Charlotte, NC 28203, USA; [2]Carolinas Hospitalist Group, Center for Outcomes Research and Evaluation, Atrium Health, Charlotte, NC 28203, USA; [3]Medical Group Division, Atrium Health, Charlotte, NC 28232, USA; [4]Center for Outcomes Research and Evaluation, Atrium Health, Charlotte, NC 28203, USA

Background: COPD is a lung disease characterized by chronic, irreversible airway obstruction that can precipitate into acute exacerbations of COPD (AECOPD) often requiring hospitalization. Improving these outcomes will require proactive innovations in care delivery to at-risk populations. Data-driven models to identify patients with AECOPD on admission to the hospital are needed, but do not exist.

Objective: This study aimed to compare the performance of several models designed to identify patients with AECOPD within 24 hours of hospital admission.

Methods: Clinical factors associated with admissions for AECOPD that are available within 24 hours of an encounter were combined into six different models and then tested retrospectively to evaluate each model's performance in predicting AECOPD. The data set incorporated billing and clinical data from patients who were older than 40 years of age with an inpatient or observation encounter in 2016 at one of the nine hospitals within a large integrated healthcare system.

Results: Of the 116,329 encounters, 6,383 had a billing diagnosis for AECOPD. The models showed a wide range of sensitivity (0.473 vs 0.963) and positive predictive value (0.190 vs 0.827).

Conclusion: It is possible to leverage clinical and administrative data to identify patients admitted with AECOPD in real-time for quality improvement or research purposes. Because models relied on clinical data, local variation in care delivery also likely contributed to performance variation across hospitals. These findings emphasize the importance of testing model performance on local data and choosing the model that best aligns with the specific goals of the targeted initiative.

Keywords: quality improvement, outcomes research, AECOPD, model, validity

Introduction

COPD is a disease state of the lungs characterized by progressive airway obstruction with underlying irreversible pathophysiology. Disease progression is often a result of long-term exposure to the inflammation-inducing particles and gases found in cigarette smoke.[1] In the United States alone, COPD affects ~15 million people and has become the third leading cause of death.[2] While considered a chronic disease, COPD patients often experience acute episodes during which symptoms worsen. These exacerbations will usually present with acute onset of dyspnea, cough, and sputum production, frequently precipitated by a respiratory illness or noncompliance with medication.[1] Often, these acute exacerbations of COPD (AECOPD) require hospitalization contributing to a large financial burden on the US healthcare system. AECOPD is estimated to cause 110,000 deaths and over 500,000 hospitalizations per year, resulting in $13.2 billion

Correspondence: Jason Roberge
Center for Outcomes Research and Evaluation, Atrium Health, 1540 Garden Terrace, Charlotte, NC 28203, USA
Tel +1 704 355 0268
Email jason.roberge@atriumhealth.org

in healthcare expenditures in 2012.[3] In-hospital mortality rates range from 2% to 5%, while 30- and 90-day mortality rates are 8.6% and over 15%, respectively.[4] According to the Centers for Medicare and Medicaid Services (CMS), around 20% of COPD patients who suffer an acute exacerbation are readmitted in 30 days. Of these 20%, ~75% were deemed preventable.

Since 2015, CMS began penalizing hospitals with high readmission rates for patients with AECOPD as part of its Hospital Readmissions Reduction Program (HRRP). Along with HRRP, the advent of other new value-based payment models has created an environment in which hospitals are looking at ways to lower readmission rates and length of stay for AECOPD patients.[2]

To successfully implement proactive interventions, health systems must have a reliable method to identify hospitalized AECOPD patients in near real-time. The mainstay of AECOPD identification relies on billing data, which is inherently delayed by weeks if not months from the actual hospitalization.[3] Shah et al have called for the development of accurate methods for the real-time identification of AECOPD.[3] Because no currently published methods exist to guide quality improvement or research, we hypothesized that a model using data available within the first 24 hours of admission could reliably identify patients with AECOPD.

Methods
Study design
This retrospective study used health system clinical and administrative data available within 24 hours of admission to create six different models to predict if a patient was admitted with AECOPD. We analyzed model performance against final hospital billing data and compared performance across models. The Chesapeake Institutional Review Board approved the study protocol with a waiver of consent and a waiver of patient authorization (Pro00021790). The outcomes were based on aggregated data pulls. Patient health information was not gathered for research purposes. These practices complied with recommendations made in the Declaration of Helsinki.

Study population
The study population included patients above the age of 40 years who were admitted in one of the nine hospitals within Atrium Health under inpatient or observation status from January 1, 2016 through December 31, 2016. Participating hospitals are in or near Charlotte (NC, USA). No other exclusions were applied.

Data collection
The hospital system maintains an Enterprise Data Warehouse (EDW) that daily integrates clinical data from the Cerner Electronic Medical Record (EMR) and administrative/ billing data from various systems. Demographics, billing codes, and administrative data were extracted from the EDW. Using convenience sampling of patients admitted to the hospital with AECOPD and those without AECOPD, the research team conducted chart reviews of patients at two facilities to empirically identify additional variables for data extraction and eventual inclusion in the models. The chart review helped to identify an EMR phenotype for patients with suspected AECOPD based on order use patterns and administrative data available at the time of admission. Typically, the following orders were used for patients with AECOPD: "Duonebs," an inhaled combination therapy of 0.5 mg ipratropium bromide and 3.0 mg albuterol sulfate; "oxygen management," an order for oxygen administration for patients as needed; any variation of systemic steroids such as prednisone or methylprednisolone; and an order to alert respiratory therapy of a patient with a COPD exacerbation.

Within Cerner, healthcare organizations can create combinations of order sets called "PowerPlans" which can detail evidence-based care steps for patients with a specific disease and target multiple provider disciplines. The COPD PowerPlan available at Atrium Health includes order options like antibiotics, steroids, bronchodilators, discharge planning, ancillary equipment, COPD education, and smoking cessation. Additional orders can be linked to clinical protocols that launch a series of events like respiratory assessments or medication weaning. Providers at Atrium Health can also use a COPD bronchodilator protocol, which includes standing orders for inhaled bronchodilator administration, respiratory therapist assessment every 4 hours, and subsequent spacing of frequency based on clinical improvement. Individual clinical orders are also available to providers for use outside of the PowerPlans and protocols. Other information available in the EMR at the time of admission included historical billing diagnoses of COPD and the presence of COPD on the problem lists maintained by providers.

We combined these empirically selected variables into six different models. Each model created a dichotomous result of "at risk for having AECOPD" if any criteria were met.

Model 1
1. COPD PowerPlan use
2. Respiratory therapy navigator notification of COPD exacerbation

3. COPD bronchodilator protocol use
4. History of COPD listed in prior billing diagnosis, admitting documentation, or problem list AND an order of systemic steroids.

Model 2

1. COPD PowerPlan use
2. Respiratory therapy navigator notification of COPD exacerbation
3. COPD bronchodilator protocol use
4. History of COPD listed in prior billing diagnosis, admitting documentation, or problem list AND an order of systemic steroids with a dose ≥40 mg.

Model 3

1. COPD PowerPlan use
2. Respiratory therapy navigator notification of COPD exacerbation
3. History of COPD listed in prior billing diagnosis, admitting documentation, or problem list AND an order of systemic steroids with a dose ≥40 mg.

Model 4

1. COPD PowerPlan use
2. Respiratory therapy navigator notification of COPD exacerbation.

Model 5

1. Duoneb order AND oxygen management order
2. COPD PowerPlan use
3. Respiratory therapy navigator notification of COPD exacerbation
4. COPD bronchodilator protocol use
5. History of COPD listed in prior billing diagnosis, admitting documentation, or problem list AND an order of systemic steroids with a dose ≥40 mg.

Model 6

1. Duoneb order AND oxygen management order
2. COPD PowerPlan use
3. Respiratory therapy navigator notification of COPD exacerbation
4. COPD bronchodilator protocol
5. History of COPD listed in prior billing diagnosis, admitting documentation, or problem list AND an order of oxygen management.

Statistical analysis

Each model was applied to all patients discharged and classified as inpatient or observation during 2016 (n=116,329).

For each model, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value were measured. These measurements were calculated based on a comparison with ICD10 billing codes for AECOPD as the gold standard. AECOPD were defined as an ICD10 code of J44.1 as either a primary or a secondary diagnosis. Measures of validity and frequencies were calculated using SAS.[5]

Results

The eligible population included 116,329 encounters, 5.5% (n=6,383) of which had a final billing diagnosis of AECOPD. Demographic characteristics of the population with AECOPD are described in Table 1. The population is on average 67 years of age, mostly comprises non-Hispanic Caucasians, and has a slight female predominance. This population is similar to national AECOPD admission data showing a mean age of 70.6 years with slight majority of females (52.8%) and large majority of non-Hispanic Caucasians.[6]

The model performance results are summarized in Table 2. Models 4 and 6 show the wide range of sensitivity (0.473–0.963, respectively) and PPV (0.827–0.190, respectively). The addition of a dosage criterion of ≥40 mg for steroid administration contributed to an absolute increase in PPV of 10.7% (Model #2) and 11.6% (Model 3) in comparison with Model 1, which did not include a dosage criterion. Consequentially, Models 2 and 3 also experience a decrease in sensitivity of 2.9% and 3.2%, respectively, in comparison with Model 1.

The sensitivity and PPV were also compared across each of the nine hospitals with Models 1–3 show considerable

Table 1 Baseline characteristics among adults ≥40 years of age upon admission with an AECOPD billing code (J44.1) in 2016

	Encounters (n=6,383)
Male	2,800 (43.9%)
Ethnicity	
Non-Hispanic	6,120 (95.9%)
Hispanic	57 (0.9%)
Unknown	206 (3.2%)
Race	
American Indian/Alaska Native	10 (0.15%)
Asian	10 (0.15%)
African American	1,388 (21.80%)
Caucasian	4,885 (76.50%)
Other	69 (1.08%)
Unknown	21 (0.32%)
Age, mean (SD)	67 (11)

Abbreviation: AECOPD, acute exacerbation of COPD.

Table 2 Validity measures by model

	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Model 1	0.880	0.961	0.567	0.993
Model 2	0.851	0.976	0.674	0.991
Model 3	0.848	0.977	0.683	0.991
Model 4	0.473	0.994	0.827	0.970
Model 5	0.900	0.943	0.483	0.994
Model 6	0.963	0.760	0.190	0.997

variability with SD for sensitivity of 0.036, 0.047, and 0.046 and for PPV of 0.072, 0.062, and 0.056 (Table 3). Overall across hospitals, Model 1 shows the least variation in sensitivity, while Model 3 shows the least variation in PPV.

Discussion

This study demonstrates that by leveraging various clinical and administrative data sources, one can reliably identify patients with AECOPD in near real-time. While the performance of these models varies widely based on variable selection, the tradeoffs between sensitivity and PPV highlight the challenge of matching any given model's performance to the specific needs of a health system initiative. Typically, an AECOPD model would be deployed to assist in identifying a specific segment of the population in order to allocate additional resources as part of a quality improvement initiative or research study. This would favor using a model with a high PPV so as not to "waste time and resources" on false positives. Whereas, a health system or study focused on changing an overall population's outcome or metric wishes to reach the maximum number of eligible patients, thus placing emphasis on sensitivity,

to minimize the "false negatives" or number of AECOPD patients who are missed.

The specific effects of variables on sensitivity and PPV from this study will serve as a guide for health systems adapting models to fit the needs of their quality improvement and research initiatives. For example, the increased PPV achieved in Models 2 and 3 resulted from the inclusion of a dosage criterion of ≥40 mg for steroid administration, which follows the current guidelines for AECOPD care. The dosage criteria helped to exclude those patients who did not have AECOPD but were taking lower dose chronic steroids for an independent indication, like severe COPD without exacerbation or rheumatoid arthritis. Conversely, the decreased sensitivity in these models can be explained by "missing" the AECOPD patients who were either not given steroids at all or given a dose less than recommended by the guidelines.

Other variables, like the order for bronchodilator protocol do little to improve sensitivity while decreasing PPV and increasing variability in AECOPD identification accuracy across hospitals. Model 3 showed that excluding the bronchodilator protocol had a very marginal decrease in sensitivity (0.3%) while slightly increasing the PPV (1.1%) and

Table 3 Validity measures for Models 1–3 by hospital

	Beds	Model 1		Model 2		Model 3	
		Sensitivity	PPV	Sensitivity	PPV	Sensitivity	PPV
Hospital 1	241	0.932	0.617	0.923	0.700	0.923	0.704
Hospital 2	67	0.848	0.668	0.811	0.748	0.808	0.752
Hospital 3	101	0.831	0.678	0.790	0.735	0.790	0.737
Hospital 4	457	0.892	0.588	0.848	0.718	0.848	0.72
Hospital 5	235	0.891	0.521	0.873	0.593	0.856	0.634
Hospital 6	175	0.892	0.611	0.864	0.729	0.862	0.733
Hospital 7	100	0.883	0.586	0.866	0.647	0.866	0.649
Hospital 8	1,132	0.827	0.442	0.795	0.578	0.792	0.585
Hospital 9	173	0.913	0.573	0.904	0.680	0.901	0.683
SD		0.036	0.072	0.047	0.062	0.046	0.056

Note: Models 1–3 were used to specifically evaluate the effects of the inclusion of the steroid >40 mg dosage criterion and COPD bronchodilator protocol criterion.
Abbreviation: PPV, positive predictive value.

decreasing variation in PPV among the different hospitals (decrease in SD of 0.6%). This indicates that the bronchodilator protocol is for the most part a redundant criterion for Model 1 and should be excluded, as it could be a source for identifying false-positive patients.

On the other hand, when a variable that is common in AECOPD management is added, we expect improved sensitivity and decreased PPV. Indeed, this was the case when comparing Models 1 and 5, which adds orders of "duonebs and oxygen management" as a criterion. This change likely identified patients who were having exacerbations, but for whom the AECOPD order set was not used, thus improving sensitivity. However, this less specific criterion also resulted in the inclusion of patients with stable respiratory disease, who use duonebs, or their equivalent, daily, thus generating false positives and decreasing PPV.

In addition to looking at sensitivity and PPV, it is also important to consider the generalizability of a model across different hospitals. We found that while Model 1 had the least variation for sensitivity, Model 3 demonstrated the least variation for PPV. The variation across models that incorporated different clinical variables is likely a result of local level differences in clinical care delivery, such as PowerPlan usage and billing patterns. In addition to provider preference and behavior, different utilization of order sets may be based on variation in patient level comorbidities and acuity.[7] Due to the potential effect of local differences in care suggested by these results, similar analyses for variance should be conducted by any large health system prior to selecting a model that relies on clinical and billing data.

While there are no published models to identify AECOPD, the performance of models in this study is similar to published results for a similar model designed to identify admitted patients with acute lung injury (ALI) in real-time. The algorithm for detecting ALI had a sensitivity of 96.3% and a PPV of 46.0%, which are relatively similar to Model 3 in our study (sensitivity 84.8% and PPV 68.3%).[8]

This study has several limitations. Most importantly, the EMR and EDW inherently contain variables with data that are missing or entered in error. If these fields were more accurately and completely populated, the likely effect would be to increase the sensitivity of Models 2, 3, and 5. Second, this was a retrospective analysis, and model performance has not yet been prospectively validated. Prospective model validation should be the subject of future studies. Third, we used the final billing ICD10 diagnosis for AECOPD as the gold standard for comparison of model performance. Final billing codes may over- or under-represent the true diagnosis

of AECOPD; however, a better alternative is currently not available. Finally, we examined model performance within a single health care system. While we did include a wide variety of hospital sizes and types, generalizability to other healthcare systems may be limited due to the dependence of variables on provider billing and order set usage behavior.

Conclusion

This study demonstrates that a healthcare system can successfully leverage clinical and administrative data to reliably identify AECOPD patients in real-time, thus guiding resource utilization for population health interventions. However, such models will inherently have variability in their performance as measured by sensitivity and PPV. Understanding the unique needs of the scenarios and users that will rely on the model can help in selecting the highest performing model. Furthermore, the variable performance of models across hospitals suggests that each health system should test model performance within their local environment prior to deployment. This study provides a starting point for incorporating risk models into quality improvement initiatives and research for patients with AECOPD, while setting the stage for future research in AECOPD predictive models.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Celli BR, MacNee W, Agusti A, ATS/ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J*. 2004;23(6):932–946.
2. Bhatt SP, Wells JM, Iyer AS, et al. Results of a Medicare bundled payments for care improvement initiative for chronic obstructive pulmonary disease readmissions. *Ann Am Thorac Soc*. 2017;14(5):643–648.
3. Shah T, Press VG, Huisingh-Scheetz M, White SR. COPD readmissions: addressing COPD in the era of value-based health care. *Chest*. 2016; 150(4):916–926.
4. Lindenauer PK, Grosso LM, Wang C, et al. Development, validation, and results of a risk-standardized measure of hospital 30-day mortality for patients with exacerbation of chronic obstructive pulmonary disease. *J Hosp Med*. 2013;8(8):428–435.

5. *SAS Enterprise Guide version 6.1 for Windows on platform 9.4.1*, 2013 SAS Institute Inc., Cary, NC.

6. Perera PN, Armstrong EP, Sherrill DL, Skrepnek GH. Acute exacerbations of COPD in the United States: Inpatient burden and predictors of costs and mortality. *COPD*. 2012;9(2):131–141.

7. Goodridge D, Lawson J, Rennie D, Marciniuk D. Rural/urban differences in health care utilization and place of death for persons with respiratory illness in the last year of life. *Rural Remote Health*. 2010;10(2):1349.

8. Herasevich V, Yilmaz M, Khan H, Hubmayr RD, Gajic O. Validation of an electronic surveillance system for acute lung injury. *Intensive Care Med*. 2009;35(6):1018–1023.