International Journal of COPD

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Validation of an administrative claims-based diagnostic code for pneumonia in a US-based commercially insured COPD population

This article was published in the following Dove Press journal:
International Journal of COPD
23 July 2015
Number of times this article has been viewed

David M Kern[1]
Jill Davis[2]
Setareh A Williams[3]
Ozgur Tunceli[1]
Bingcao Wu[1]
Sally Hollis[4]
Charlie Strange[5]
Frank Trudo[2]

[1]HealthCore, Inc., Wilmington, DE, [2]AstraZeneca Pharmaceuticals, Wilmington, DE, [3]AstraZeneca Pharmaceuticals, Gaithersburg, MD, USA; [4]AstraZeneca Pharmaceuticals, Cheshire, UK; [5]Department of Medicine, Division of Pulmonary and Critical Care Medicine, Medical University of South Carolina, Charleston, SC, USA

Objective: To estimate the accuracy of claims-based pneumonia diagnoses in COPD patients using clinical information in medical records as the reference standard.

Methods: Selecting from a repository containing members' data from 14 regional United States health plans, this validation study identified pneumonia diagnoses within a group of patients initiating treatment for COPD between March 1, 2009 and March 31, 2012. Patients with ≥1 claim for pneumonia (International Classification of Diseases Version 9-CM code 480.xx–486.xx) were identified during the 12 months following treatment initiation. A subset of 800 patients was randomly selected to abstract medical record data (paper based and electronic) for a target sample of 400 patients, to estimate validity within 5% margin of error. Positive predictive value (PPV) was calculated for the claims diagnosis of pneumonia relative to the reference standard, defined as a documented diagnosis in the medical record.

Results: A total of 388 records were reviewed; 311 included a documented pneumonia diagnosis, indicating 80.2% (95% confidence interval [CI]: 75.8% to 84.0%) of claims-identified pneumonia diagnoses were validated by the medical charts. Claims-based diagnoses in inpatient or emergency departments (n=185) had greater PPV versus outpatient settings (n=203), 87.6% (95% CI: 81.9%–92.0%) versus 73.4% (95% CI: 66.8%–79.3%), respectively. Claims-diagnoses verified with paper-based charts had similar PPV as the overall study sample, 80.2% (95% CI: 71.1%–87.5%), and higher PPV than those linked to electronic medical records, 73.3% (95% CI: 65.5%–80.2%). Combined paper-based and electronic records had a higher PPV, 87.6% (95% CI: 80.9%–92.6%).

Conclusion: Administrative claims data indicating a diagnosis of pneumonia in COPD patients are supported by medical records. The accuracy of a medical record diagnosis of pneumonia remains unknown. With increased use of claims data in medical research, COPD researchers can study pneumonia with confidence that claims data are a valid tool when studying the safety of COPD therapies that could potentially lead to increased pneumonia susceptibility or severity.

Keywords: positive predictive value, pneumonia, validation, claims data, medical record review

Introduction

Claims data are generated primarily from administrative transactions and are essential for the accounting and reimbursement purposes associated with health care delivery. Increasingly, however, administrative claims data are repurposed for real-world observational studies in the United States and other developed countries. Typically maintained in regularly updated repositories, claims data provide convenient and easy access to researchable details on demographics, clinical characteristics, health care utilization – frequency and type of prescription medication fills, inpatient

Correspondence: David M Kern
HealthCore, Inc., 800 Delaware Avenue, Wilmington, DE 19801, USA
Tel +1 302 230 2102
Fax +1 302 230 2020
Email dkern@healthcore.com

hospitalization, and the use of outpatient, emergency, and physician offices – and the cost of care.[1,2] Claims databases efficiently and comprehensively capture structured data such as diagnostic codes from the International Classification of Diseases Version 9 (ICD-9-CM) or generic product identifier medication codes. In general, these databases do not enable access to data on patients' race or ethnicity, laboratory test values and other clinical measures or to unstructured data, including clinicians' progress notes, all of which could be valuable in identifying disease and establishing patient status. It is essential, therefore, when assessing outcomes related to disease identified with claims data, to understand the degree of agreement between claims-based observations and actual physician diagnoses, and such validations are not available for several diseases.

Pneumonia was associated with more than 1.1 million inpatient hospitalizations and 50,000 deaths in 2010,[3,4] and is more common and more fatal in COPD patients than in those without COPD.[5] In the first year after a COPD diagnosis, individuals are at 16 times the risk for pneumonia compared to those without COPD.[6] In a recent study the incidence rate of community acquired pneumonia was 22.4 events per 1,000 person years in the 10 years following the diagnosis of COPD, and more than 50% higher in those categorized as having severe COPD.[7] Furthermore, the economic impact of pneumonia is greater for those with COPD, illustrated by a doubling of direct medical costs following an inpatient hospitalization for pneumonia compared to those without COPD in a study of older individuals.[8] Thus, pneumonia is especially of interest in individuals with COPD, and administrative claims databases can be a useful tool for studying this disease if the data are determined to be sufficiently valid.

In the past few years alone, pneumonia has been widely researched using claims data in patients with and without COPD, with studies focusing on the economic and clinical impact of the disease, the development of algorithms to predict inpatient readmission, and measuring the safety and efficacy of treatments.[8–22] Although the accuracy of claims-based diagnoses is critical to the sensitivity and validity of observational research, this remains largely unexamined for pneumonia. To our knowledge, no study has validated claims-based pneumonia diagnoses across different points of service such as inpatient hospitalizations, emergency departments (EDs), and physicians' offices.

The current study aimed to determine the validity of pneumonia diagnosis codes in a large US-based administrative claims database by comparing patients' claims-based records and their medical records at the time of their observed pneumonia diagnosis. The objective was to confirm the accuracy of claims-based diagnoses of pneumonia by using the physicians' diagnoses (medical record-based) as the reference standard.

Materials and methods

Data source

This retrospective cohort study utilized claims data for commercially insured individuals from the HealthCore Integrated Research Environment. The HealthCore Integrated Research Environment contains a diverse spectrum of longitudinal claims data for more than 32 million individuals, and has the capability of linking claims data to providers, which facilitates the selection of medical records.

All study materials were handled in compliance with the Health Insurance Portability and Accountability Act, and a limited dataset was used for all analyses, as defined by the Privacy Rule. A national, institutional review board (IRB), the New England IRB, reviewed the protocol and approved this study prior to the researchers obtaining patient medical records. HealthCore was granted a waiver of authorization to obtain the medical records without patient authorization of release after the IRB determined the use or disclosure of protected health information in this research project involved no more than minimal risk to the individuals.

Study population

This validation study was nested within a larger, retrospective cohort study (NCT01921127)[23,24] comparing the effectiveness of two common inhaled corticosteroid/long-acting β2-agonist combination (ICS/LABA) medications in COPD patients (ICD-9-CM code 491.xx, 492.xx, or 496.xx). Pneumonia, a complication in many COPD patients, was analyzed as a secondary outcome, and the validation study was performed to assess the accuracy of the pneumonia diagnosis code (ICD-9-CM code 480.xx–486.xx) to reflect a documented diagnosis of pneumonia reported in the medical records. Patients initiating ICS/LABA between March 1, 2009 and March 31, 2012 were followed for 1 year following treatment initiation, during which patients diagnosed with pneumonia were identified. Only the first event was assessed for patients with more than one claim with a diagnosis of pneumonia.

Medical record abstraction

A professional medical record abstraction agency, employing trained and experienced medical data reviewers, obtained and abstracted the patient medical records using a study-specific instrument developed by the research team. The medical record

abstractors had advanced education in nursing (registered nurse [RN], licensed practical nurse [LPN], licensed vocational nurse [LVN]), pharmacy (PharmD, registered pharmacist [RPh]) and/or medicine (medical doctor, doctor of osteopathic medicine [DO]). Prior to the start of the study, all data abstractors participated in a project-specific data collection training session utilizing the data collection manual and materials specifically developed for the study. Each individual medical record was abstracted by a single abstractor.

The validity of the diagnostic claims was determined by the presence of a diagnosis of pneumonia documented in the medical records. The confirmation of a pneumonia diagnosis in the patient record was obtained via either a diagnosis code for pneumonia in the record or text from a physician note, assessment, or diagnostic summary stating that pneumonia was present. Additional information abstracted from the records included demographic characteristics (sex, race, ethnicity, and smoking status), symptoms pertaining to pneumonia (cough, fever, chest pain, chills, dyspnea, rales, rhonchi, wheezing, distant breath sounds, and temperature >100°F), diagnostic tests and the presence of abnormal results related to pneumonia (respiratory rate, heart rate, chest X-ray, chest computerized axial tomography [CAT] scan, sputum Gram stain, sputum culture, and blood culture), and the administration or prescribing of antibiotics.

Statistical analysis

A target final sample size of 400 patients was chosen a priori to allow for the estimation of validity with a maximum margin of error <5%. A threshold of 5% margin of error was selected to achieve a confidence interval (CI) no more than 10% wide. Additionally, robust sample sizes, ideally of ≥100 records, were wanted for each of the subgroup analyses by place of service (inpatient/ED versus [vs] outpatient) and chart type (paper, electronic, or a combination of both paper and electronic records [hybrid]), with the assumption that no subgroup would account for less than one quarter of all records based on the claims results and prior experience.

The primary outcome of the validation analysis was calculated using the positive predictive value (PPV) of the claims diagnosis for pneumonia relative to the medical record review, the reference standard. PPV was calculated as the number of pneumonia cases with a documented pneumonia diagnosis in the medical records ("diagnosed pneumonia") (true positives) divided by the total number of cases identified from the claims data (true positives plus false positives). Results were reported overall and stratified by place of service (inpatient or ED vs outpatient) and medical record type (paper, electronic, or hybrid). A 95% CI for the estimated PPV was constructed using Clopper-Pearson (exact) confidence limits. Additional descriptive statistics are reported for each of the demographic characteristics, symptoms, diagnostic tests, and antibiotic use. All analyses were performed using SAS version 9.2 (SAS Institute Inc., Cary, NC, USA).

Results

Medical record collection

A total of 1,345 patients had ≥1 claim with a pneumonia diagnosis during an inpatient hospitalization, ED visit, or outpatient visit over the 12-month follow-up period; 1,022 of the diagnoses were from claims containing provider contact information, and were deemed eligible for medical record abstraction. A random subset of 800 patients was selected with the goal of abstracting information on 400 patients, allowing for the possibility that up to 50% of targeted records may be unobtainable. There were 730 record requests to providers because an initial review showed 70 patients with insufficient contact information. Medical records were obtained for 407 patients; however 19 were unusable due to missing information on birth dates or sex, or had unmatched names, or incorrect time periods, which resulted in 388 abstracted records (Figure 1).

Patient characteristics

The average age of patients with abstracted records was 67.8 (±11.6) years old, and 53.9% were female (Table 1). Of the 47.9% of patients with information on race noted in the record, 92.5% were white. Patients' smoking status was reported as active (23.2%), former (36.3%), never (10.3%), and status unknown for the remaining 30.2%. Patients whose records were abstracted (n=388) and pneumonia patients identified with claims data (n=1,345) had similar mean age (67.8 and 67.4 years), sex (53.9% and 52.9% female), and health plan type (60.3% and 60.6%, enrolled in preferred provider organizations), respectively. A slightly greater proportion of patients were from the Midwest in the record-reviewed vs the claims group (51.8% vs 47.1%) and a slightly smaller proportion from the Northeast (12.1% vs 17.5%).

Primary analysis

A documented pneumonia diagnosis was found in 311 out of the total 388 available medical records (Table 2), indicating that 80.2% (PPV 95% CI: 75.8% to 84.0%) of pneumonia positive patients identified with the claim code algorithm were diagnosed with pneumonia according to the medical records. Stratification by health care setting showed that

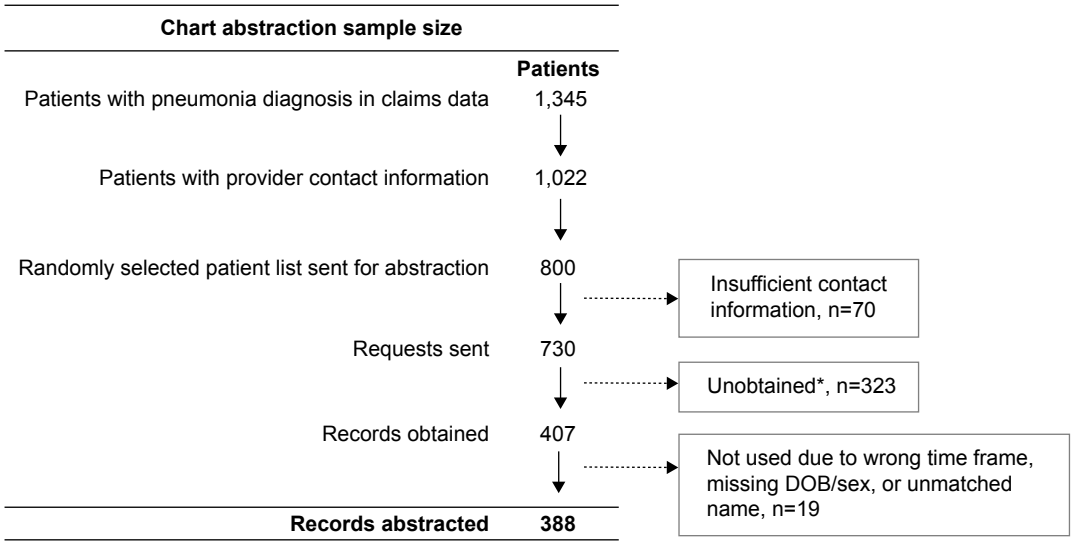

Figure 1 Patient selection.
Notes: *Could not locate provider (n=36), no record of patient (n=44), no record of patient in timeframe requested (n=32), provider refuses to participate (n=52), provider requires patient consent or IRB approval (n=41), chart lost/destroyed/archived (n=11), no response from office (n=63), payment denied/record not needed (n=3), requests suspended (met quota) (n=41).
Abbreviations: DOB, date of birth; IRB, institutional review board.

Table 1 Patient demographics

	All pneumonia patients identified in claims (N=1,345)		Medical record status					
			All medical records (N=388)		Documented pneumonia diagnosis (N=311)		No documented pneumonia diagnosis (N=77)	
Female (n, %)	711	52.9%	209	53.9%	160	51.4%	49	63.6%
Age (n, %)								
40 to 49 years	108	8.0%	29	7.5%	22	7.1%	7	9.1%
50 to 59 years	285	21.2%	66	17.0%	57	18.3%	9	11.7%
60 to 64 years	195	14.5%	62	16.0%	51	16.4%	11	14.3%
65+ years	757	56.3%	231	59.5%	181	58.2%	50	64.9%
Mean ± SD	67.4	12.5	67.8	11.6	67.5	11.4	68.9	12.3
Geographic region (n, %)								
Northeast	236	17.5%	47	12.1%	41	13.2%	6	7.8%
Midwest	634	47.1%	201	51.8%	157	50.5%	44	57.1%
South	303	22.5%	87	22.4%	70	22.5%	17	22.1%
West	172	12.8%	53	13.7%	43	13.8%	10	13.0%
Healthplan type (n, %)								
HMO	434	32.3%	122	31.4%	99	31.8%	23	29.9%
PPO	815	60.6%	234	60.3%	185	59.5%	49	63.6%
CDHP	48	3.6%	11	2.8%	10	3.2%	1	1.3%
Other commercial	48	3.6%	21	5.4%	17	5.5%	4	5.2%
Race (n, %)	n/a*							
Caucasian/White			172	44.3%	147	47.3%	25	32.5%
African American/Black			9	2.3%	8	2.6%	1	1.3%
Asian/Pacific Islander			5	1.3%	4	1.3%	1	1.3%
Not specified			202	52.1%	152	48.9%	50	64.9%
Smoking status (n, %)	n/a*							
Active smoker			90	23.2%	81	26.0%	9	11.7%
Former smoker			141	36.3%	122	39.2%	19	24.7%
Never smoked			40	10.3%	34	10.9%	6	7.8%
Unknown			117	30.2%	74	23.8%	43	55.8%

Notes: Age, geographic region, and healthplan type were captured via claims data at the index date, while race and smoking status were identified from the medical records.
*Not applicable, data are only available in medical records, not the claims data.
Abbreviations: HMO, health maintenance organization; PPO, preferred provider organization; CDHP, consumer-driven health plan; SD, standard deviation.

Table 2 Primary outcome – positive predictive value for pneumonia in claims

	Number abstracted	Number of documented pneumonia diagnosis	PPV (%)	(95% CI)
All pneumonia patients	388	311	80.2	(75.8–84.0)
By place of service				
Inpatient hospitalization/emergency department visit	185	162	87.6	(81.9–92.0)
Outpatient/office visit	203	149	73.4	(66.8–79.3)
By medical record type				
Paper	101	81	80.2	(71.1–87.5)
Electronic medical records (EMR)	150	110	73.3	(65.5–80.2)
Hybrid of paper and EMR	137	120	87.6	(80.9–92.6)

Abbreviation: PPV, positive predictive value; CI, confidence interval.

claims originating from an inpatient or ED (n=185) had a higher PPV than those from outpatient settings (n=203) (87.6% vs 73.4%, respectively). Claims that were linked to paper charts had the same PPV as the overall study population, 80.2% (95% CI: 71.1%–87.5%), while those linked to electronic medical records (EMR) had a lower PPV, 73.3% (95% CI: 65.5%–80.2%), and those with data from both paper charts and EMR had a higher PPV, 87.6% (95% CI: 80.9%–92.6%).

Diagnosed pneumonia cases (n=311) vs those without a documented diagnosis in the medical records (n=77) included fewer females (51.4% vs 63.6%) and a higher proportion of smokers (26.0% active and 39.2% former smokers vs 11.7% and 24.7%, respectively).

Pneumonia symptoms

Within all examined medical records, coughing (61.3%), was the most prevalent symptom reported, followed by dyspnea (53.9%), wheezing (42.5%), distant breath sounds (31.2%), rhonchi (28.9%), fever (26.0%), and rales (25.5%) (Table 3). There was a clear discrepancy in the frequency of reported symptoms between records with and without a pneumonia diagnosis; just 15.8% of patients with a pneumonia diagnosis documented in the medical records had none of the symptoms listed, while more than half (53.2%) of those without a diagnosis had no symptoms mentioned. Furthermore, individual symptoms were 2–5 times more prevalent in diagnosed cases compared with those without a diagnosis documented in the medical records.

Table 3 Pneumonia symptoms captured from medical charts

	All medical records (N=388)		Medical record status			
			Documented pneumonia diagnosis (N=311)		No documented pneumonia diagnosis (N=77)	
Symptoms (n, %)						
Cough	238	61.3%	215	69.1%	23	29.9%
Fever	101	26.0%	91	29.3%	10	13.0%
Chest pain	70	18.0%	66	21.2%	4	5.2%
Chills	64	16.5%	59	19.0%	5	6.5%
Dyspnea	209	53.9%	188	60.5%	21	27.3%
Rales	99	25.5%	88	28.3%	11	14.3%
Rhonchi	112	28.9%	107	34.4%	5	6.5%
Wheezing	165	42.5%	151	48.6%	14	18.2%
Distant breath sounds	121	31.2%	111	35.7%	10	13.0%
Temperature >100°F (37.8°C)	51	13.1%	45	14.5%	6	7.8%
Number of symptoms (n, %)						
No symptoms reported	90	23.2%	49	15.8%	41	53.2%
1 symptom	39	10.1%	29	9.3%	10	13.0%
2 symptoms	39	10.1%	29	9.3%	10	13.0%
3 symptoms	40	10.3%	37	11.9%	3	3.9%
4+ symptoms	180	46.4%	167	53.7%	13	16.9%

Diagnostic testing and antibiotic use

Diagnostic testing was more common in diagnosed pneumonia cases vs those without a documented diagnosis of pneumonia in the medical record (Table 4). Respiratory rate testing was performed in 74.3% of diagnosed pneumonia cases compared with 42.9% of records that had no pneumonia diagnosis. Within those that had a test, abnormal respiratory rate results were reported for 55.0% of those with a documented diagnosis of pneumonia and 45.5% of those without. Similarly, chest X-ray and CAT scans were performed more often in those with vs without a documented diagnosis (77.8% and 24.8% vs 51.9% and 11.7%, respectively), and abnormal results in those with diagnosed pneumonia were found for 89.6% of X-rays and 90.5% of CAT scans compared to 61.5% and 77.8% of X-rays and CAT scans within patients without a diagnosis documented in the charts. Results of other diagnostic tests showed similar differences.

Medical record reviews indicated that antibiotics were administered or prescribed to 79.7% of the diagnosed pneumonia cases compared with 35.1% that had no documented diagnosis (Table 5). Levofloxacin (25.8%), azithromycin (25.0%), and ceftriaxone (24.2%) were the three most commonly used antibiotics in this patient sample; all three are commonly used to treat pneumonia. The lack of evidence of antibiotic use in 20% of confirmed cases may suggest that many of those cases were suspected to be viral in origin.

Discussion

This study tested the validity of pneumonia diagnoses among COPD patients identified in a large US administrative claims database using medical records as a reference standard. This study differs from prior investigations on this subject in a number of important ways. Prior studies have examined the validity of claims data in diagnosing pneumonia.[25–29] However, those studies were either limited to a single hospital or ED,[25,26] to children in pediatric hospitals,[28] or to only hospitalizations due to pneumonia.[29]

We used ICD-9-CM codes to identify pneumonia diagnoses in COPD patients from outpatient, ED, and inpatient settings using a claims database that covers more than 30 million managed care patients across the US. Although there are major differences in the design of our study compared to others, the PPV calculated for claims-based pneumonia diagnoses in the current study was similar to those reported by Aronsky et al (79.4%–80.8% PPV based on the algorithm that used different combinations of ICD-9-CM codes)[25] and Whittle et al (89%).[26] Both of these PPVs were similar to the 87.6% PPV found in inpatient or ED cases in our study.

Table 4 Patients with available results for diagnostic tests recorded in medical charts and the proportion with abnormal results

	All medical records (N=388)		Medical record status			
			Documented pneumonia diagnosis (N=311)		No documented pneumonia diagnosis (N=77)	
Respiratory rate test	264	68.0%	231	74.3%	33	42.9%
Results available in chart	264		231		33	
Abnormal (<12 or >20 breaths per minute)	142	53.8%	127	55.0%	15	45.5%
Heart rate test	295	76.0%	256	82.3%	39	50.6%
Results available in chart	294		255		39	
Abnormal (<60 or >100 bpm)	131	44.6%	120	47.1%	11	28.2%
Chest X-ray test	282	72.7%	242	77.8%	40	51.9%
Results available in chart	270		231		39	
Abnormal*	231	85.6%	207	89.6%	24	61.5%
Chest CAT scan test	86	22.2%	77	24.8%	9	11.7%
Results available in chart	83		74		9	
Abnormal*	74	89.2%	67	90.5%	7	77.8%
Sputum Gram stain test	71	18.3%	67	21.5%	4	5.2%
Results available in chart	58		55		3	
Abnormal (positive)	40	69.0%	37	67.3%	3	100.0%
Sputum culture test	78	20.1%	74	23.8%	4	5.2%
Results available in chart	62		58		4	
Abnormal (positive, isolated organism)	27	43.5%	25	43.1%	2	50.0%
Blood culture test	144	37.1%	133	42.8%	11	14.3%
Results available in chart	133		123		10	
Abnormal (positive, isolated organism)	12	9.0%	10	8.1%	2	20.0%

Notes: The proportion of charts with abnormal results are based on the denominator of charts with available results. *Abnormal included any of the following notes: "consolidation", "pleural effusion", "fluid", "infiltrate", "haze", or "consistent with pneumonia".
Abbreviations: bpm, beats per minute; CAT, computerized axial tomography.

Table 5 Antibiotics use identified in medical charts

	All medical records (N=388)		Medical record status			
			Documented pneumonia diagnosis (N=311)		No documented pneumonia diagnosis (N=77)	
Antibiotics (n, %)						
Azithromycin	97	25.0%	90	28.9%	7	9.1%
Clarithromycin	9	2.3%	9	2.9%	0	0.0%
Erythromycin	0	0.0%	0	0.0%	0	0.0%
Doxycycline	11	2.8%	10	3.2%	1	1.3%
Ciprofloxacin	22	5.7%	20	6.4%	2	2.6%
Gemifloxacin	1	0.3%	1	0.3%	0	0.0%
Gatifloxacin	1	0.3%	1	0.3%	0	0.0%
Levofloxacin	100	25.8%	93	29.9%	7	9.1%
Moxifloxacin	60	15.5%	51	16.4%	9	11.7%
Sparfloxacin	0	0.0%	0	0.0%	0	0.0%
Trovafloxacin	0	0.0%	0	0.0%	0	0.0%
Ceftriaxone	94	24.2%	88	28.3%	6	7.8%
Amoxicillin	1	0.3%	1	0.3%	0	0.0%
Amoxicillin with clavulanate	10	2.6%	8	2.6%	2	2.6%
Ampicillin	1	0.3%	1	0.3%	0	0.0%
Patients with ≥1 antibiotics prescribed	275	70.9%	248	79.7%	27	35.1%

Compared to outpatient claims, pneumonia diagnoses in inpatient or ED claims were found to be more accurate (87.6% vs 73.4% PPV). Furthermore, the presence of a chest X-ray or chest CAT scan was more common in inpatient and ED cases, of which, all but one individual (99.4%) with documented pneumonia had evidence of either procedure; conversely, chest imaging was present in just 61.1% of cases originating in an outpatient setting with documented pneumonia. Further research is needed to investigate differences in validity by place of service since there is more detailed information in inpatient medical records compared to those in physicians' offices. Our study suggests that differences in pneumonia outcomes may be more difficult to study in the outpatient setting; since neither medical records nor claims data strongly support the accuracy of the medical diagnosis of pneumonia as evidenced by the lack of imaging in significant numbers of individuals.

One goal of this research was to confirm that a patient was diagnosed with pneumonia during the same health care encounter in which we observed a diagnosis of pneumonia in the claims data. The goal was not to determine whether the diagnosis given by the physician was accurate; however, there was evidence that pneumonia diagnoses that were simultaneously documented in the medical records were usually accurately diagnosed by the physician. Many claims-only diagnoses may not have truly been pneumonia. For example, patients with a pneumonia diagnosis documented in the medical records had abnormal diagnostic test results and symptoms of pneumonia at least twice as often compared with those who did not have a diagnosis in the medical records. Additionally, the higher smoking rates in those with a documented diagnosis is consistent with research that has shown smoking to be the strongest independent risk factor of pneumonia in immunocompetent non-elderly adults.[30,31] Conversely, although nearly 20% of patients in the claims data did not have a pneumonia diagnosis documented in the medical records, there was evidence that some of those patients may have had pneumonia. The prevalence of individual symptoms being less common in unconfirmed cases may be a function of the pneumonia being truly absent, but also may be a result of additional omissions in medical records (ie, physicians who omitted a note of pneumonia in the record may also be more likely to leave out notes of specific symptoms). Furthermore, more than 20% of patients with an unconfirmed diagnosis had three or more symptoms that are associated with pneumonia; of the 40 patients who had a chest X-ray, 61.5% had an abnormal result, and 35.1% unconfirmed cases were given antibiotics during their visit. Thus, although the medical record was considered the reference standard in this analysis, they might not be 100% accurate, and the results could be an underestimate of the true PPV of the claims diagnosis.

Strengths

This study was able to directly link patients' records from an administrative claims database to their medical records for the exact time they had an insurance claim for a diagnosis of pneumonia in outpatient, inpatient, and EDs. This allowed the confirmation that pneumonia diagnoses observed in the claims database were present in the medical records, where

the medical record was considered the reference standard in the validation analysis. Along with capturing the presence of a pneumonia diagnosis, we were also able to obtain information on the presence of pneumonia symptoms, diagnostic lab testing, and administration of antibiotics during the visit of interest. The high rates of symptoms and abnormal test results in the medical record documented cases provided further evidence that most were true pneumonia patients.

Limitations

Medical records, though used as the reference standard here, are not 100% accurate. For example, it was possible for records to be at multiple locations for a single patient during the same time period, due to different providers having separate paper charts and/or EMR systems; from which a record from one provider may have had a pneumonia diagnosis while another did not. Each event had a medical record pulled from a single location and, thus, it was possible to miss some records with a pneumonia diagnosis. Further, human error plays a role in accurate recording and abstracting of diagnoses. It is possible, that a patient was diagnosed with pneumonia but it was not noted in the record, or that part of the medical record was missing, or that a diagnosis for pneumonia was in the record but missed by the abstractor. However, all medical record abstractors were trained and experienced in the health care field and participated in project-specific data collection training that utilized the data collection manual and materials specifically developed for this study. Less than 80% of patients with pneumonia documented in the medical record had a chest X-ray noted, a procedure that should be performed prior to diagnosing pneumonia.[32] It may be that imaging was performed but not included in the chart, performed at an alternative medical site that was not captured, or that the provider felt there was enough evidence to make a diagnosis of pneumonia without an imaging test. While a clinical diagnosis of pneumonia without a chest X-ray may not be clinically sound, the goal of this study was not to validate whether or not the clinician made a correct diagnosis, instead it was to determine whether or not the administrative claims accurately reflected the diagnosis made by the clinician.

The reference standard of medical record abstraction has limitations. Some of the comorbidities of COPD that include congestive heart failure, pulmonary emboli, and mucus plugging during a COPD exacerbation may be diagnosed as pneumonia with an abnormal radiograph and pulmonary symptoms, treated with antibiotics, and yet not be pneumonia at all. Biomarkers such as brain natriuretic peptide for congestive heart failure and D-dimer for thromboembolic disease should be captured in future prospective studies. However, medical records were considered the gold standard as they are the repository for the clinical information documented by the medical provider and other health care professionals and utilized for the care of the patient. Furthermore, medical records have been used as the gold standard in a number of other studies to determine the validity of pneumonia diagnoses.[27–29,33]

Because this study examined only patients with a pneumonia diagnosis observed in claims, the negative predictive value, sensitivity, and specificity of pneumonia identified via administrative claims could not be estimated from this study.

Lastly, all patients in the larger study design were required to have 12 months of follow-up and all pneumonia diagnoses were captured during those 12 months. Thus, only pneumonia patients who survived until the end of the follow-up period were included, and so the most severe cases of pneumonia resulting in death were not represented in this validation study. The validity recorded in this study might have been higher if cases of pneumonia that resulted in death were included, since these cases in an inpatient setting would be expected to have higher validity compared to outpatient diagnoses.

Conclusion

This study demonstrated that administrative claims data provide an accurate basis for the diagnosis of pneumonia in COPD patients across multiple service settings. Diagnostic accuracy varies by treatment setting and tends to be better in inpatient and EDs vs outpatient settings. Additional research will help to validate whether the service setting has a bearing on the accuracy of pneumonia diagnosis and whether differences were a consequence of data type or the actual content of the data inherent in claims and medical records. With the increasing amount of research being performed utilizing administrative claims it is important to understand the validity of such data. Confirming the validity of pneumonia diagnoses in claims data allows researchers to confidently study COPD associated pneumonia and investigate the safety of COPD therapies utilizing longitudinal observational data.

Acknowledgment

Bernard Tulsi provided editorial support for this manuscript. AstraZeneca Pharmaceuticals provided funding.

Disclosure

JD, FT, SAW, and SH have disclosed that they are employed by AstraZeneca. DMK, OT, and BW have disclosed that they are employed by HealthCore, Inc., which received funding from AstraZeneca for this study. CS has been a consultant for Astra Zeneca LP, CSL Behring, Grifols, PneumRx, Pulmonx,

and Uptake Medical on the topic of COPD. The authors report no other conflicts of interest in this work.

References

1. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health*. 2014;68(3):283–287.
2. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med*. 1997;127(8 Pt 2):666–674.
3. CDC/NHS. Number and rate of discharges from short-stay hospitals and of days of care, with average length of stay and standard error, by selected first-listed diagnostic categories: United States, 2010. CDC/NCHS National Hospital Discharge Survey; 2010. Available from: http://www.cdc.gov/nchs/data/nhds/2average/2010ave2_firstlist.pdf. Accessed June 22, 2015.
4. Centers for Disease Control and Prevention. Murphy SL, Xu J, Kochanek KD. *Deaths: Final Data for 2010*. National Vital Statistics Reports; 2013;61(4). Available from: http://www.cdc.gov/nchs/data/nvsr/nvsr61/nvsr61_04.pdf. Accessed June 13, 2015.
5. Restrepo MI, Mortensen EM, Pugh JA, Anzueto A. COPD is associated with increased mortality in patients with community-acquired pneumonia. *Eur Respir J*. 2006;28(2):346–351.
6. Soriano JB, Visick GT, Muellerova H, Payvandi N, Hansell AL. Patterns of comorbidities in newly diagnosed COPD and asthma in primary care. *Chest*. 2005;128(4):2099–2107.
7. Mullerova H, Chigbo C, Hagan GW, et al. The natural history of community-acquired pneumonia in COPD patients: a population database analysis. *Respir Med*. 2012;106(8):1124–1133.
8. Ryan M, Suaya JA, Chapman JD, Stason WB, Shepard DS, Thomas CP. Incidence and cost of pneumonia in older adults with COPD in the United States. *PLoS One*. 2013;8(10):e75887.
9. Bratzler DW, Normand SL, Wang Y, et al. An administrative claims model for profiling hospital 30-day mortality rates for pneumonia patients. *PLoS One*. 2011;6(4):e17401.
10. Broulette J, Yu H, Pyenson B, Iwasaki K, Sato R. The Incidence Rate and Economic Burden of Community-Acquired Pneumonia in a Working-Age Population. *Am Health Drug Benefits*. 2013;6(8):494–503.
11. Davydow DS, Hough CL, Zivin K, Langa KM, Katon WJ. Depression and risk of hospitalization for pneumonia in a cohort study of older Americans. *J Psychosom Res*. 2014;77(6):528–534.
12. Janson C, Larsson K, Lisspers KH, et al. Pneumonia and pneumonia related mortality in patients with COPD treated with fixed combinations of inhaled corticosteroid and long acting beta2 agonist: observational matched cohort study (PATHOS). *BMJ*. 2013;346:f3306.
13. Lin J, Li Y, Tian H, et al. Costs and health care resource utilization among chronic obstructive pulmonary disease patients with newly acquired pneumonia. *Clinicoecon Outcomes Res*. 2014;6:349–356.
14. Lindenauer PK, Normand SL, Drye EE, et al. Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia. *J Hosp Med*. 2011;6(3):142–150.
15. Thomas CP, Ryan M, Chapman JD, et al. Incidence and Cost of Pneumonia in Medicare Beneficiaries. *Chest*. 2012;142(4):973–981.
16. Polsky D, Bonafede M, Suaya JA. Comorbidities as a driver of the excess costs of community-acquired pneumonia in U.S. commercially-insured working age adults. *BMC Health Serv Res*. 2012;12:379.
17. Reynolds CA, Finkelstein JA, Ray GT, Moore MR, Huang SS. Attributable healthcare utilization and cost of pneumonia due to drug-resistant streptococcus pneumonia: a cost analysis. *Antimicrob Resist Infect Control*. 2014;3:16.
18. Ruhnke GW, Coca-Perraillon M, Kitch BT, Cutler DM. Marked Improvement in 30-Day Mortality among Elderly Inpatients and Outpatients with Community-Acquired Pneumonia. *Am J Med*. 2012;142(2):171–178.
19. Sato R, Gomez Rey G, Nelson S, Pinsky B. Community-acquired pneumonia episode costs by age and risk in commercially insured US adults aged ≥50 years. *Appl Health Econ Health Policy*. 2013;11(3):251–258.
20. Thornton-Snider J, Luna Y, Wong KS, et al. Inhaled corticosteroids and the risk of pneumonia in Medicare patients with COPD. *Curr Med Res Opin*. 2012;28(12):1959–1967.
21. Tichopad A, Roberts C, Gembula I, et al. Clinical and economic burden of community-acquired pneumonia among adults in the Czech Republic, Hungary, Poland and Slovakia. *PLoS One*. 2013;8(8):e71375.
22. Yawn BP, Li Y, Tian H, Zhang J, Arcona S, Kahler KH. Inhaled corticosteroid use in patients with chronic obstructive pulmonary disease and the risk of pneumonia: a retrospective claims data analysis. *Int J Chronic Obstruct Pulmon Dis*. 2013;8:295–304.
23. AstraZeneca. Comparative Effectiveness of Symbicort vs Advair Among COPD Patients. Available from: http://clinicaltrials.gov/show/NCT01921127. NLM identifier: NCT01921127. Accessed June 13, 2015.
24. Kern DM, Davis J, Williams SA, et al. Comparative effectiveness of budesonide/formoterol combination and fluticasone/salmeterol combination among chronic obstructive pulmonary disease patients new to controller treatment: a US administrative claims database study. *Respir Res*. 2015;16(1):52.
25. Aronsky D, Haug PJ, Lagor C, Dean NC. Accuracy of administrative data for identifying patients with pneumonia. *Am J Med Qual*. 2005;20(6):319–328.
26. Whittle J, FIne MJ, Joyce DZ, et al. Community-acquired pneumonia: can it be defined with claims data? *Am J Med Qual*. 1997;12(4):187–193.
27. Skull SA, Andrews RM, Byrnes GB, et al. ICD-10 codes are a valid tool for identification of pneumonia in hospitalized patients aged > or =65 years. *Epidemiol Infect*. 2008;136(2):232–240.
28. Williams DJ, Shah SS, Myers A, et al. Identifying pediatric community-acquired pneumonia hospitalizations: Accuracy of administrative billing codes. *JAMA Pediatr*. 2013;167(9):851–858.
29. Yu O, Nelson JC, Bounds L, Jackson LA. Classification algorithms to improve the accuracy of identifying patients hospitalized with community-acquired pneumonia using administrative data. *Epidemiol Infect*. 2011;139(9):1296–1306.
30. File TM. Community-acquired pneumonia. *Lancet*. 2003;362(9400):1991–2001.
31. Nuorti JP, Butler JC, Farley MM, et al. Cigarette smoking and invasive pneumococcal disease. Active Bacterial Core Surveillance Team. *N Engl J Med*. 2000;342(10):681–689.
32. Mandell LA, Wunderink RG, Anzueto A, et al. Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-Acquired Pneumonia in Adults. *Clin Infect Dis*. 2007;44 Suppl 2:S27–S72.
33. Drahos J, Vanwormer JJ, Greenlee RT, Landgren O, Koshiol J. Accuracy of ICD-9-CM codes in identifying infections of pneumonia and herpes simplex virus in administrative data. *Ann Epidemiol*. 2013;23(5):291–293.