ELSEVIER

Journal of Clinical Epidemiology 60 (2007) 834−838

Journal of Clinical Epidemiology

International classification of diseases codes showed modest sensitivity for detecting community-acquired pneumonia

Ewoudt M.W. van de Garde[a,b], Jan Jelrik Oosterheert[c], Marc Bonten[d], Robert C. Kaplan[e], Hubert G.M. Leufkens[a,*]

[a]*Division of Pharmacoepidemiology and Pharmacotherapy, Utrecht Institute for Pharmaceutical Sciences, Faculty of Science, Utrecht University, PO Box 80082, 3506 TB Utrecht, The Netherlands*

[b]*Department of Clinical Pharmacy, St. Antonius Hospital, Nieuwegein, The Netherlands*

[c]*Department of Internal Medicine and Dermatology, Division of Medicine, Infectious Diseases & Geriatrics, University Medical Center Utrecht, The Netherlands*

[d]*Eijkman Winker Institute for Microbiology and Infectious Diseases, Julius Center for Health Science and Primary Care, University Medical Center Utrecht, The Netherlands*

[e]*Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA*

Accepted 17 October 2006

Abstract

Objective: To estimate the sensitivity of International Classification of Diseases (ICD-9-CM) coding for detecting hospitalized community-acquired pneumonia and to assess possible determinants for misclassification.

Study Design and Setting: Based on microbiological analysis data, 293 patients with a principal diagnosis of community-acquired pneumonia at seven hospitals in the Netherlands were assigned to three categories (pneumococcal pneumonia, pneumonia with other organism, or pneumonia with no organism specified). For these patients, the assigned principal and secondary ICD-9-CM codes in the hospital discharge record were retrieved and the corresponding sensitivity was calculated. Furthermore, pneumonia-related patient characteristics were compared between correctly and incorrectly coded subjects.

Results: The overall sensitivity was 72.4% for the principal code and 79.5% for combined principal and secondary codes. For pneumococcal pneumonia (ICD-9-CM code 481) and pneumonia with specified organism (ICD-9-CM code 482−483), the sensitivities were 35% and 18.3%, respectively. Patient characteristics were not significantly different between correctly and incorrectly coded subjects except for duration of hospital stay, which correlated negatively with coding sensitivity ($P = 0.01$).

Conclusion: ICD-9-CM codes showed modest sensitivity for detecting community-acquired pneumonia in hospital administrative databases, leaving at least one quarter of pneumonia cases undetected. Sensitivity decreased with longer duration of hospital stay.

Keywords: Pneumonia; International Classification of Diseases; Databases; Sensitivity and specificity; Outcome assessment; Inpatients

1. Introduction

Community-acquired pneumonia (CAP) is a common and potentially fatal infection of lung tissue and is associated with high health care costs. Therefore, CAP is subject to many epidemiological and economical studies. In about 20% of all pneumonia cases, inpatient treatment is required because the clinical situation does not allow outpatient therapy [1,2]. Many of the studies on pneumonia, therefore, identify cases of hospitalized CAP because these are most likely to result in death and resource use. A common approach to identify cases of hospitalized CAP is using hospital discharge records as coded according to the International Classification of Diseases-9th revision-Clinical Modification (ICD-9-CM). The common ICD-9-CM codes used for this purpose are 481, 482.x, 483.x, 485, and 486 [3−8]. Information, however, on the validity of such data is inconclusive or lacking in many cases. Several researchers have assessed the validity of hospital discharge records of various diseases, including pneumonia, by identifying cases through ICD-9-CM codes and subsequently reviewing medical charts to confirm or reject the correct diagnosis [9−12]. This approach, however, leaves cases of CAP without an ICD-9-CM code for pneumonia undetected and provides no information on the sensitivity of

* Corresponding author. Tel.: +31-30-2537324; fax: +31-30-2539166.
E-mail address: h.g.m.leufkens@pharm.uu.nl (H.G.M. Leufkens).

0895-4356/07/$ − see front matter
doi: 10.1016/j.jclinepi.2006.10.018

ICD-9-CM coding for detection of cases of CAP. The aim of this study is to estimate the sensitivity of ICD-9-CM code assignment in a population of patients admitted with a principal diagnosis of CAP. Furthermore, we aim to assess possible determinants for misclassification.

2. Patients and methods

This study used patient data from a randomized open label clinical trial (July 2000−March 2004) on efficacy of an early switch of intravenous antibacterial treatment to oral treatment of CAP [13]. All adult patients hospitalized for CAP in seven hospitals (two university medical centers and five teaching hospitals) in the Netherlands were eligible for inclusion in that study. Pneumonia was defined as a new or progressive infiltrate on a chest X-ray plus at least two of the following criteria: cough, sputum production, rectal temperature $>38^{\circ}C$ or $<36^{\circ}C$, ausculatory finding consistent with pneumonia, leucocytosis ($>10.000/mm^3$ or $>15\%$ bands), C-reactive protein >3 times the upper limit of normal, positive blood culture, or positive culture of pleural fluid. Patients with cystic fibrosis, a history of colonization with Gram negative bacteria because of structural damage to the respiratory tract, malfunction of the digestive tract, life expectancy of less than 1 month because of underlying diseases, infections other than pneumonia needing antibiotic treatment, severe immunosuppression (neutropenia ($<0.5 \times 10^9/L$) or a CD4 count $<200/mm^3$), and needing mechanical ventilation in an intensive care unit were excluded.

2.1. Identification of pathogen and patient categorization

For each patient, sputum samples and blood samples were collected, cultured, and evaluated following standard procedures. In addition, Binax NOW-tests were used to detect urinary antigen for *Legionella pneumophila* and *Streptococcus pneumoniae*. Acute and convalescent serology samples were collected and evaluated for *Mycoplasma pneumoniae*, *L. pneumophila*, and *Chlamydophila pneumoniae*. Based on microbiological analyses performed, patients were categorized patients as having pneumococcal pneumonia (defined as *S. pneumoniae* isolated from a blood sample or adequate sputum sample containing >25 polymorphonuclear neutrofils and <10 epithelial cells per high power field), pneumonia with other pathogen specified, or pneumonia with no organism specified. When in a patient more than one pathogen was specified, *S. pneumoniae* was defined as the pathogenic organism.

2.2. Coding accuracy measurement

In the Netherlands, the principal diagnosis is defined as the medical condition most responsible for admission to a hospital. The principal diagnosis should be coded in the primary position on the hospital discharge record. In addition, different secondary diagnoses and complications can be coded in the secondary positions [14]. All records are coded according to the ICD-9-CM.

For all patients, we retrieved the coded hospital discharge record from the medical registration department of the participating hospitals. This was done more than 2 years after closure of the previously mentioned randomized clinical trial (March 2006). The following ICD-9-CM codes were considered to be correct for each different patient category: 481 for pneumococcal pneumonia (the ICD-9-CM specifically excludes coding pneumococcal pneumonia as 482.3); 482.x, and 483.x for pneumonia with other organism specified; and 485−486 for pneumonia with no organism specified. For each category of pneumonia, we evaluated the proportion of patients for whom the correct ICD-9-CM code was listed in the primary or secondary positions on the hospital discharge record and calculated the sensitivity. The sensitivity was defined as the number of correctly coded patients divided by total number of patients in that category. Analyses were conducted for all patients together and for the seven hospitals separately. For subjects without an ICD-9-CM code of 481−486, we identified the other codes assigned in the primary position.

2.3. Determinants for misclassification

To assess possible determinants for misclassification, we identified and analyzed the following characteristics between correctly and incorrectly coded subjects: age, gender, comorbidities (heart failure, history of stroke, liver disease, malignancy, renal disease), pneumonia severity index [15], and whether the patient was a nursing home resident. The reason we identified the previously mentioned characteristics was because they have been linked to pneumonia outcome in different prediction rule models [16−18] and that, especially, differences in prevalence of these items are important to predict possible selection bias in etiological or prognostic studies using cases of CAP identified through ICD-9-CM codes. In addition, we evaluated the following pneumonia outcome measures: duration of hospital stay and in-hospital mortality.

To study determinants for misclassification, multiple variable logistic regression analysis was conducted with incorrect code assigned as dependent. All possible determinants were included in the multivariate model when they were retained after backward stepwise elimination ($P < 0.10$).

3. Results

Of the 293 patients hospitalized for CAP, 40 (14%) had confirmed pneumococcal pneumonia, 82 (28%) had pneumonia with another organism specified, and 171 (58%) had pneumonia with no organism specified (Table 1). In

Table 1
(Overall) sensitivity for ICD-9-CM code assignment for hospitalized CAP

	Principal diagnosis code assigned					
CAP diagnosis	*n* (%)	481	482–483	485–486	Other	Sensitivity (%)
All diagnoses	293 (100)	32	30	150	81	72.4[a]
Pneumococcal pneumonia	40 (14)	14	7	13	6	35.0[b]
Pneumonia with other organism specified	82 (28)	10	15	30	27	18.3[b]
Pneumonia, organism unspecified	171 (58)	8	8	107	48	62.6[b]

[a] Overall sensitivity for any pneumonia-related ICD-9-CM code (481–486).
[b] Sensitivity for individual diagnoses.

total, 212 patients had any pneumonia-related ICD-9-CM code (481–486) as principal diagnosis, yielding an overall sensitivity for any pneumonia-related ICD-9-CM code of 72.4%. The overall sensitivity for six of the seven participating hospitals separately ranged from 61.5 to 82.0% (one hospital excluded for including only one patient). Expanding the criteria to a correct ICD-9-CM code as principal diagnosis or as any secondary diagnosis increased the overall sensitivity to 79.5%. The sensitivities for all three categories individually are shown in Table 1. When both ICD-9-CM codes 481 and 482.3 were considered valid for pneumococcal pneumonia, the sensitivity for that category increased from 35% to 47.5%. For cases without an ICD-9-CM code 481–486 as principal or secondary diagnosis ($n = 60$), the most frequently occurring other ICD-9-CM codes were 496 (chronic airway obstruction, not classified), 507.0 (pneumonitis due to inhalation of food or vomitus), and 162.9 (malignant neoplasm of bronchus or lung), respectively assigned to 11, 3, and 3 patients. Other codes assigned as principal diagnosis were diverse and occurred not more than once. None of the patients had an ICD-9-CM code for viral pneumonia (480.x) or influenza with pneumonia (487.x).

The characteristics of the correctly and incorrectly coded patients were not significantly different except duration of hospital stay, which was significantly associated with risk of misclassification in the multivariable regression analysis (Table 2; $P = 0.01$). Although not significant, malignancies were surprisingly more prevalent in incorrectly coded subjects. Fig. 1 shows the relation between duration of hospital stay and sensitivity of ICD-9-CM coding.

4. Discussion

Our study showed that in patients hospitalized with confirmed CAP, overall, only 72% was assigned any ICD-9-CM code for pneumonia (481–486) as the principal diagnosis on the hospital discharge record. For "pneumococcal pneumonia" and "pneumonia with other organism specified", sensitivity was as low as 35% and 18.3%, respectively.

Table 2
Characteristics of correctly and incorrectly coded patients

	Incorrectly coded ($n = 81$)	Correctly coded ($n = 212$)
Mean age (SD)	70.0 (13.1)	69.5 (14.1)
Sex		
Male	54 (66.7)	139 (65.6)
Female	27 (33.3)	73 (34.4)
Comorbidities		
Heart failure (%)	9 (11.1)	27 (12.7)
History of stroke (%)	8 (9.9)	19 (9.0)
Liver disease (%)	1 (1.2)	2 (0.9)
Renal disease (%)	7 (8.6)	20 (9.4)
Malignancy (%)	24 (29.6)	41 (19.3)
Nursing home resident (%)	3 (3.7)	8 (3.8)
Mean length of hospital stay, days (SD)*	15.2 (11.6)	11.7 (8.6)
In-hospital mortality (%)	4 (4.9)	16 (7.5)
Pneumonia severity index (%)		
Risk class 1	0	0
Risk class 2	1 (1.2)	15 (7.1)
Risk class 3	6 (7.4)	18 (8.5)
Risk class 4	57 (70.4)	143 (67.5)
Risk class 5	17 (21.0)	36 (17.0)

* $P = 0.01$.

Ideally, the ICD-9-CM code in the primary position (principal diagnosis) on the hospital discharge record always represents the medical condition that is chiefly responsible for the admission of the patient to a hospital. In the present study, however, we observed that this was the case in only 72%. Errors in classification can occur in any stage of the long chain of events leading to assignment

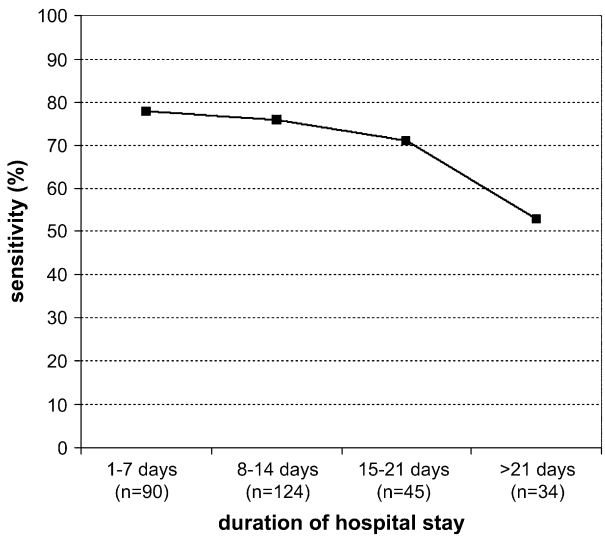

Fig. 1. Duration of hospital stay and overall sensitivity.

of an ICD-9-CM code to a hospital discharge record: hospitalization, diagnosing, medical record keeping, filling the discharge abstract form by the treating physician, and interpretation by the coding clerk. In this study, all patients were hospitalized with a principal diagnosis of CAP as confirmed according to a standardized exam. Therefore, lack of information on hospitalization or diagnosing are probably less plausible explanations for the errors in classification. This makes medical record keeping, filling the discharge abstract form by the treating physician, and interpretation by the coding clerk more probable causes for the inaccurate coding observed in this study. But also complications or additional diagnoses could decrease the likelihood that the diagnosis at admission remains the evident principal diagnosis in the medical record. Our findings are supportive to such an effect because the true principal diagnosis and the assigned code were most likely to coincide in brief (and probably uncomplicated) admissions and diverged during prolonged admissions.

Concerning the pneumonia-related patient characteristics evaluated in the present study, except for duration of hospital stay, there were no significant differences between correctly and incorrectly coded subjects. This important finding implies that when cases of pneumonia are identified through ICD-9-CM codes no specific patient categories remain undetected except for an underestimation of patients with prolonged and probably complicated hospital stay. Especially when studying quality of care, outcomes, resource use, or the development of predictive instruments, inclusion of cases representing severe pneumonia is essential, as they represent the pneumonia cases that are most likely to result in death, intensive care admission, and prolonged hospital length of stay. The latter appear less represented when cases of hospitalized CAP are selected through ICD-9-CM coded hospital discharge records.

Advantages of our study are the possibility to include pneumonia cases from seven different sites decreasing the chances of coding bias, the availability of microbiological data, and the ability to study differences in sensitivity between primary and secondary discharge codes. Because CAP occurs frequently in combination with other diseases, such as congestive heart failure, chronic bronchitis, or exacerbation of asthma, the pneumonia code may appear in a secondary position. In our study, extending the assessment from the first to secondary positions resulted in only 21 (7.1%) additional patients with a pneumonia-related code in a secondary position instead of the primary position. This important finding indicates that case selection based on principal diagnosis codes alone does not exclude large numbers of patients with CAP. Principal diagnosis codes may also have greater positive predictive value (PPV) for identifying pneumonia, although we were unable estimate the PPV in the present study because the population was limited to confirmed cases of pneumonia.

Our study also has some limitations, which need to be discussed. First, we used patients from a prospective clinical trial in which patients with pneumonia were included using a strict protocol. However, because the data collected in the CAP study were not available to medical records staff, we do not believe that the study had any effect on the coding practices. If the study had any effect on accuracy of coding practices, this was likely to lead to an underestimation of the real problem of inaccurate code assignment. Secondly, this study was conducted in seven hospitals in the Netherlands, which could prevent extrapolation of the findings to other countries due to potential differences in coding of hospital discharge records. In addition, our study provides only information on the sensitivity of ICD-9-CM code assignment for hospitalized CAP and not on pneumonia treated in primary care.

The findings from this study could have an impact on the validity of studies using coded hospital discharge records for case selection. For example, when studying the effects of pneumococcal vaccination on pneumococcal pneumonia incidence, a considerable underestimated incidence has to be faced. Or, in case–control studies, as a result of disease misclassification, selection bias could occur when not all observed cases (or controls) are true cases (or controls). Especially in hospital-based case–control studies, this can cause dilution of any association under study.

In conclusion, ICD-9-CM codes showed modest sensitivity for detecting CAP in hospital administrative databases, leaving at least one quarter of pneumonia cases undetected. Sensitivity decreased with longer duration of hospital stay.

Acknowledgments

We gratefully thank all hospitals that participated in this project.

References

[1] Segreti J, House HR, Siegel RE. Principles of antibiotic treatment of community-acquired pneumonia in the outpatient setting. Am J Med 2005;118(Suppl 7A):21S–8S.

[2] Woodhead MA, Macfarlane JT, McCracken JS, Rose DH, Finch RG. Prospective study of the aetiology and outcome of pneumonia in the community. Lancet 1987;1:671–4.

[3] Bi P, Whitby M, Walker S, Parton KA. Trends in mortality rates for infectious and parasitic diseases in Australia: 1907–1997. Intern Med J 2003;33:152–62.

[4] Mortensen EM, Restrepo MI, Anzueto A, Pugh J. The effect of prior statin use on 30-day mortality for patients hospitalized with community-acquired pneumonia. Respir Res 2005;6:82.

[5] Almirall J, Bolibar I, Balanzo X, Gonzalez CA. Risk factors for community-acquired pneumonia in adults: a population-based case-control study. Eur Respir J 1999;13:349–55.

[6] van de Garde EM, Souverein PC, van den Bosch JM, Deneer VH, Leufkens HG. Angiotensin-converting enzyme inhibitor use and pneumonia risk in a general population. Eur Respir J 2006;27:1217–22.

[7] O'Meara ES, White M, Siscovick DS, Lyles MF, Kuller LH. Hospitalization for pneumonia in the Cardiovascular Health Study:

incidence, mortality, and influence on longer-term survival. J Am Geriatr Soc 2005;53:1108−16.

[8] Etminan M, Zhang B, Fitzgerald M, Brophy JM. Do angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers decrease the risk of hospitalization secondary to community-acquired pneumonia? A nested case−control study. Pharmacotherapy 2006; 26:479−82.

[9] Aronsky D, Haug PJ, Lagor C, Dean NC. Accuracy of administrative data for identifying patients with pneumonia. Am J Med Qual 2005;20:319−28.

[10] Guevara RE, Butler JC, Marston BJ, Plouffe JF, File TM Jr, Breiman RF. Accuracy of ICD-9-CM codes in detecting community-acquired pneumococcal pneumonia for incidence and vaccine efficacy studies. Am J Epidemiol 1999;149:282−9.

[11] Movig KL, Leufkens HG, Lenderink AW, Egberts AC. Validity of hospital discharge International Classification of Diseases (ICD) codes for identifying patients with hyponatremia. J Clin Epidemiol 2003;56:530−5.

[12] De Bruin ML, van Hemel NM, Leufkens HG, Hoes AW. Hospital discharge diagnoses of ventricular arrhythmias and cardiac arrest were useful for epidemiologic research. J Clin Epidemiol 2005;58: 1325−9.

[13] Oosterheert J. Diagnosis and treatment of community-acquired lower respiratory tract infections. Utrecht, The Netherlands: Utrecht University; 2005.

[14] Prismant. LMR users manual edition March 2005 [in Dutch]. Available at. http://www.prismant.nl/prismant/download/content/LMR_Gebruikershandleiding.pdf. Accessed July 10, 2006.

[15] Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med 1997;336:243−50.

[16] Bartlett JG, Dowell SF, Mandell LA, File TM Jr, Musher DM, Fine MJ. Practice guidelines for the management of community-acquired pneumonia in adults. The Infectious Diseases Society of America. Clin Infect Dis 2000;31:347−82.

[17] Niederman MS, Bass JB Jr, Campbell GD, Fein AM, Grossman RF, Mandell LA, et al. Guidelines for the initial management of adults with community-acquired pneumonia: diagnosis, assessment of severity, and initial antimicrobial therapy. American Thoracic Society. Medical Section of the American Lung Association. Am Rev Respir Dis 1993;148:1418−26.

[18] McKean MC. Evidence based medicine: review of BTS guidelines for the management of community acquired pneumonia in adults. J Infect 2002;45:213−8.