Published in final edited form as:

J Rheumatol. 2015 November ; 42(11): 2071–2074. doi:10.3899/jrheum.150186.

The Capture of Rheumatic Disease through Hospital Diagnoses with Examples of Rheumatoid Arthritis and Giant Cell Arteritis: What are We Missing?

C. John Michet III¹, Cynthia S. Crowson, M.S.^{1,2}, Sara J. Achenbach, M.S.², and Eric L. Matteson, M.D., M.P.H.^{1,3}

¹Division of Rheumatology, Department of Internal Medicine, Mayo Clinic, Rochester, MN

²Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN

³Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, MN

Abstract

Objective—We examined hospitalizations for patients with known rheumatoid arthritis (RA) or giant cell arteritis (GCA) to evaluate whether hospitalization-related diagnoses accurately capture patients with rheumatologic diseases.

Methods—Diagnosis codes for hospitalizations in 1996–2012 among previously identified population-based cohorts of RA/GCA patients were examined for RA/GCA mention.

Results—RA/GCA mention occurred in only 55% of 2407 hospitalizations among RA patients and 31% of 502 hospitalizations among GCA patients. RA/GCA was mentioned more often in recent years, younger patients and rheumatic medication users.

Conclusion—Coding for RA/GCA during hospitalizations was often missed. Research utilizing hospital diagnoses alone could be biased.

Keywords

Rheumatoid arthritis; giant cell arteritis; hospitalization; diagnosis; International Classification of Disease; coding

INTRODUCTION

Electronic data capture is commonplace, but assessing the limitations of data obtained for one purpose (e.g., diagnosis coding) when using it for another purpose (e.g., research) can be challenging. Several studies have addressed the positive predictive value of International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM) codes and algorithms for the identification of patients with rheumatic diseases in electronic databases

Corresponding Author: Eric L. Matteson, MD, MPH, Mayo Clinic, 200 1st St SW, Rochester, MN 55905, Matteson.Eric@mayo.edu, Phone: 1-507-284-2511.

The authors do not have any financial conflicts of interest for the work reported herein.

[1–3]. These studies showed high positive predictive value for patients with multiple codings of the same rheumatic disease over a period of time. However, the National Inpatient Sample, a free and widely accessible database from the Healthcare Cost and Utilization Project (HCUP), is limited to information regarding hospitalizations without the ability to link those hospitalizations to other healthcare for the same patient, so multiple codings are not available [4]. Another issue not well addressed by these studies is whether there are hospitalized patients with rheumatologic diseases who are not identified due to an absence of a rheumatologic code from an inpatient stay. We examined primary and secondary diagnosis codes from hospitalizations for patients with known rheumatoid arthritis (RA) or giant cell arteritis (GCA) in order to evaluate whether hospitalization-related diagnoses accurately capture patients with rheumatologic diseases. We also examined whether certain reasons for hospitalization were more or less likely to be associated with mention of RA/GCA.

PATIENTS AND METHODS

This retrospective, population-based cohort study utilized the resources of the Rochester Epidemiology Project, a medical records linkage system containing the complete inpatient and outpatient medical records from all health care providers in Olmsted County, MN, including Mayo Clinic and its affiliated hospitals, Olmsted Medical Center, local nursing homes and private practitioners [5, 6].

Two previously identified cohorts of Olmsted County residents were analyzed. The RA cohort consisted of patients aged ≥ 18 years who first fulfilled the 1987 American College of Rheumatology (ACR) classification criteria for RA in 1980–2007 [7, 8]. The GCA cohort consisted of patients who fulfilled the 1990 ACR criteria for GCA in 1950–2009. [9, 10].

Hospitalization data were retrieved electronically from all Olmsted County medical providers. Follow-up began with the latter of RA/GCA diagnosis date or January 1, 1996, and ended at death, emigration from Olmsted County, or December 31, 2012. All primary and secondary diagnoses from admission to discharge were analyzed. RA mention was defined as ICD-9-CM 714.xx with expansion to include codes for closely related diseases [i.e., ICD-9-CM codes: 696.0, 715.9X, 716.5–716.59, 716.8X, 716.9X, 719.3X, 720.0, 720.2, 721.90]. GCA mention was defined as ICD-9-CM 446.5 with expansion to include 446.7 and 447.6.

In order to examine whether certain reasons for hospitalization were associated with mention of RA/GCA, the primary discharge diagnoses were grouped together using the Clinical Classifications Software (CCS) for ICD-9-CM from the Healthcare Cost and Utilization Project [11]. The CCS classifies diagnoses into 18 chapters: infections and parasitic diseases; neoplasms; endocrine, nutritional, and metabolic diseases and immunity disorders; diseases of the blood and blood-forming organs; mental illness; diseases of the nervous system and sense organs; diseases of the circulatory system; diseases of respiratory system; diseases of the digestive system; diseases of the genitourinary system; complications of pregnancy, childbirth and puerperium; diseases of the skin and subcutaneous tissue; diseases of the musculoskeletal system and connective tissue; congenital anomalies; certain

conditions originating in the perinatal period; injury and poisonings (which includes fractures); symptoms, signs and ill-defined conditions; and residual codes, unclassified.

Information on RA disease severity was collected previously by review of each patient's full medical record [12]. Characteristics included rheumatoid factor (RF) positivity and anti-citrullinated protein antibody (ACPA) positivity at the time of RA diagnosis, as well as repeatedly high erythrocyte sedimentation rates (ESR) (i.e., 3 ESR measures ≥ 50 mm/hr with at least 30 days between 2 measurements), the presence of joint erosions/destructive changes or rheumatoid nodules during the first year following diagnosis of RA. Information on the use of corticosteroids for both RA and GCA patients and use of disease modifying anti-rheumatic medications (DMARDs) and biologic response modifiers for patients with RA during each hospitalization was also collected.

Statistical Methods

Descriptive statistics (means, percentages, etc.) were used to summarize the data. Within each cohort, comparisons of hospitalizations with and without mention of RA/GCA were performed using generalized linear models with random intercepts for each subject to account for multiple hospitalizations in the same subject. Analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA).

RESULTS

The RA cohort contained 499 patients (68.1% female) who had 2407 hospitalizations with a mean age at hospital admission of 68.8 years. RA mention by ICD-9-CM 714.x in discharge diagnoses occurred in 1334 hospitalizations (55.4%). Using the expanded list of related codes, an additional 68 hospitalizations were identified for a total of 1402 hospitalizations with mention of RA (58.2%).

Hospitalization with RA mention occurred more often in younger patients (mean 67.2 years versus 70.9 years; $p < 0.001$; Table 1), in patients with long-standing RA (mean 11.2 years versus 10.0 years; $p < 0.001$) and in more recent hospitalizations ($p < 0.001$).

The mention of RA differed significantly in primary discharge diagnoses for three CCS chapters. RA was more likely to be mentioned when the primary discharge diagnoses of the hospitalization concerned diseases of the musculoskeletal system and connective tissue diseases compared to all other primary discharge diagnosis categories (78% vs 55%; $p < 0.001$). RA was less likely to be mentioned when the primary discharge diagnosis of the hospitalization concerned neoplasms (49% vs 59%; $p = 0.034$) or Symptoms, signs, and ill-defined conditions (47% vs 59%; $p = 0.020$).

RA was more likely to be mentioned if patients were RF/ACPA positive (Table 2). Other RA disease severity measures occurring in the first year of RA diagnosis, such as erosions, repeatedly elevated ESR levels, and rheumatoid nodules, were not associated with an increased likelihood of RA mention. Biologic use and corticosteroid use were both associated with higher likelihood of RA mention.

The GCA cohort consisted of 119 patients (79.0% female) who had 502 hospitalizations. GCA was mentioned in 155 hospitalizations (30.9%). Two additional hospitalizations (0.4%) were coded as unspecified arteritis (ICD-9-CM 447.6) and no hospitalizations were coded as Takayasu's disease (ICD-9-CM 446.7). GCA was more likely to be mentioned in slightly younger patients (mean 82.7 versus 84.8 years; $p=0.002$; Table 3), those with shorter disease duration (mean 4.9 versus 10.4 years; $p<0.001$) and those taking corticosteroids (60% vs 14%; $p<0.001$).

While no CCS chapter categorization of primary discharge diagnoses yielded statistically significant results, data from two chapters stood out. GCA was somewhat more likely to be mentioned when the primary discharge diagnosis concerned diseases of the circulatory system compared (37% vs 29%; $p = 0.07$) and GCA was less likely to be mentioned when the primary discharge diagnosis concerned diseases of the genitourinary system (14% vs 32%; $p = 0.06$).

DISCUSSION

This study evaluated how often hospital diagnoses included mention of two systemic rheumatologic diseases using cohorts of patients with known, and well characterized RA and GCA. Coding for RA in hospitalized patients with RA occurred only half the time. Hospitalizations with RA mention of RA occurred more frequently in recent years, younger patients and those taking rheumatic medications. In hospitalized patients with GCA, coding for GCA occurred only 30% of the time and generally in younger patients and those using corticosteroids.

These results raise the concern that a large number of patients with diagnosed rheumatic diseases may not be included in studies in which only hospital discharge diagnoses are used for the identification of patients with rheumatic diseases. This failure of capture could result in studies flawed by biases of selection and detection.

While use of hospital codes for identification of cases is a convenient and efficient strategy for epidemiologic and related healthcare utilization studies, this strategy has inherent potential shortcomings. The data are dependent on accurate diagnoses by physicians and complete coding by the hospital. The presence or absence of coding is influenced by a number of factors, including diagnoses or procedures considered primary for the hospitalization, and those which may increase reimbursement for services.

A study of accuracy of hospital diagnoses for RA in the Danish National Patient Register found that only 59% of RA mentions in the National Patient Register were confirmed in medical records [13]. The investigators were not able to evaluate how many patients with RA were not coded in their register.

Strengths of this study include the use of well-characterized cohorts of patients who are known to have RA/GCA. It includes every hospitalization the patients underwent in Olmsted County over a 16 year period. To our knowledge, there are no other studies that have utilized cohorts of patients with known rheumatic diseases to evaluate the capacity of hospital discharge diagnoses to identify these patients.

Study limitations include lack of data for hospitalizations occurring outside Olmsted County and the uncertainty regarding generalizability of results from one county to the entirety of the United States; however, it is unlikely that hospital diagnoses differ significantly from region to region. Generalizability of these results outside the United States is difficult or impossible due to differences in healthcare systems in other countries which may have different incentives for coding practices.

It is incumbent on researchers to be judicious in their use of hospital diagnoses and ICD-9-CM codes for the identification of patients with rheumatic diseases. It is equally important for reviewers and readers to be shrewd in their judgment of epidemiological studies based solely on these methods.

Acknowledgments

Funding Statement:

This work was supported by a grant from the National Institutes of Health, NIAMS (R01 AR46849), and made possible by the Rochester Epidemiology Project (R01 AG034676 from the National Institute on Aging).

References

1. Asgari MM, Wu JJ, Gelfand JM, Salman C, Curtis JR, Harrold LR, Herrinton LJ. Validity of diagnostic codes and prevalence of psoriasis and psoriatic arthritis in a managed care population, 1996–2009. *Pharmacoepidemiol Drug Saf.* 2013; 22:842–9. [PubMed: 23637091]
2. Kim SY, Servi A, Polinski JM, Mogun H, Weinblatt ME, Katz JN, Solomon DH. Validation of rheumatoid arthritis diagnoses in health care utilization data. *Arthritis Res Ther.* 2011; 13:R32. [PubMed: 21345216]
3. Ng B, Aslam F, Petersen N, Yu HJ, Suarez-Almazor ME. Identification of rheumatoid arthritis patients using an administrative database: a Veterans Affairs study. *Arthritis Care Res.* 2012; 64:1490–96.
4. Healthcare Cost and Utilization Project. HCUPnet; <http://hcupnet.ahrq.gov/>
5. StSauver JL, Grossardt BR, Yawn BP, Melton LJ 3rd, Rocca WA. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester epidemiology project. *Am J Epidemiol.* 2011; 173:1059–68. [PubMed: 21430193]
6. Rocca WA, Yawn BP, StSauver JL, Grossardt BR, Melton LJ 3rd. History of the Rochester epidemiology project: half a century of medical records linkage in US population. *Mayo Clin Proc.* 2012; 87(12):1202–1213. [PubMed: 23199802]
7. Myasoedova E, Crowson CS, Kremers HM, Thorneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising? Results from Olmsted County, Minnesota, 1955–2007. *Arthritis Rheum.* 2010; 62:1576–82. [PubMed: 20191579]
8. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* 1988; 31:315–24. [PubMed: 3358796]
9. Chandran A, Udayakumar P, Crowson C, Warrington K, Matteson E. The incidence of giant cell arteritis in Olmsted County, Minnesota, over a 60-year period 1950–2009. *Scand J Rheum.* 2015; 21:1–14.
10. Hunder GG, Bloch DA, Michel BA, Stevens MB, Arend WP, Calabrese LH, et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis Rheum.* 1990; 33:1122–28. [PubMed: 2202311]
11. Elixhauser, A.; Steiner, C.; Palmer, L. Clinical Classifications Software (CCS). US Agency for Healthcare Research and Quality; 2014. www.hcupus.ahrq.gov/toolssoftware/ccs/ccs.jsp [Accessed January 16, 2015]

12. Myasoedova E, Crowson CS, Nicola PJ, Maradit-Kremers H, Davis JM III, Roger VL, et al. The influence of rheumatoid arthritis disease characteristics on heart failure. *J Rheum.* 2011; 38:1601–6. [PubMed: 21572155]
13. Pedersen M, Klarlund M, Jacobsen S, Svendsen AJ, Frisch M. Validity of rheumatoid arthritis diagnoses in the Danish National Patient Registry. *Eur J Epi.* 2004; 19:1097–1103.

TABLE 1

Characteristics of Hospitalizations in Patients with Rheumatoid Arthritis (RA) according to mention of RA in hospitalization-related diagnoses

Characteristic	No Mention of RA (n=1005)	Mention of RA (n=1402)	p value
Age (years) at Hospital Admission, mean (SD)	70.9 (15.3)	67.2 (13.7)	<0.001
Sex, n (%)			0.40
Female	607 (38.1)	988 (61.9)	
Male	398 (49.0)	414 (51.0)	
Hospital Admission Year, mean (SD)	2003.3 (3.9)	2005.5 (4.4)	<0.001
Length of Stay (days), mean (SD)	4.8 (6.2)	4.3 (4.6)	0.26
RA duration at Hospital Admission (years), mean (SD)	10.0 (6.7)	11.2 (7.5)	<0.001

Variables in **bold** statistically significant at P <0.05, two-sided

TABLE 2

Association between Rheumatoid Arthritis (RA) Severity Measures and Mention of RA in Hospitalization-Related Diagnoses

Characteristic	No Mention of RA (n=1005)	Mention of RA (n=1402)	p value
Rheumatoid Factor and/or Anti- Citrullinated Protein Antibody Positivity			<0.001
No	456 (60.2)	302 (39.8)	
Yes	549 (33.3)	1100 (66.7)	
Presence of Erosions/destructive changes on radiographs *			0.07
No	772 (43.7)	993 (56.3)	
Yes	233 (36.3)	409 (63.7)	
Presence of Rheumatoid Nodules *			0.16
No	881 (43.5)	1143 (56.5)	
Yes	124 (32.4)	259 (67.6)	
3 Erythrocyte Sedimentation Rate Values 50mm/hr *			0.83
No	844 (41.6)	1183 (58.4)	
Yes	161 (42.4)	219 (57.6)	
DMARD use during hospitalization			0.14
No	732 (49.9)	734 (50.1)	
Yes	273 (29.0)	668 (71.0)	
Biologic use during hospitalization			<0.001
No	980 (43.9)	1251 (56.1)	
Yes	25 (14.2)	151 (85.8)	
Corticosteroid use during hospitalization			0.037
No	575 (50.3)	569 (49.7)	
Yes	430 (34.0)	833 (66.0)	

Values in the table are n (%). Variables in **bold** statistically significant at P <0.05. Abbreviation: DMARD=disease-modifying anti-rheumatic drug

* In the first year after RA diagnosis

TABLE 3

Characteristics of Hospitalizations in Patients with Giant Cell Arteritis (GCA) according to mention of GCA in hospitalization-related diagnoses

Characteristic	No Mention of GCA (n=345)	Mention of GCA (n=157)	p value
Age (years) at Hospital Admission, mean (SD)	84.8 (7.7)	82.7 (4.7)	0.002
Sex, n (%)			0.35
Female	272 (70.8%)	112 (29.2%)	
Male	73 (61.9%)	45 (38.1%)	
Hospital Admission Year, mean (SD)	2003.5 (4.8)	2005.1 (4.7)	0.07
Length of Stay (days), mean (SD)	4.3 (4.6)	4.6 (5.5)	0.56
GCA duration at Hospital Admission (years), mean (SD)	10.4 (6.7)	4.9 (4.5)	<0.001
Corticosteroid use during hospitalization, n(%)			<0.001
No	263 (86.2%)	42 (13.8%)	
Yes	75 (39.5%)	115 (60.5%)	

Variables in **bold** statistically significant at P <0.05