Published in final edited form as:

Arthritis Care Res (Hoboken). 2020 October ; 72(10): 1392–1403. doi:10.1002/acr.24043.

Performance of Administrative Algorithms to Identify Interstitial Lung Disease in Rheumatoid Arthritis

Bryant R. England, MD^{1,2}, Punyasha Roul, MS², Tina D. Mahajan, MD², Namrata Singh, MBBS, MSCI^{3,4}, Fang Yu, PhD⁵, Harlan Sayles, MS^{2,5}, Grant W. Cannon, MD⁶, Brian C. Sauer, PhD⁶, Joshua F. Baker, MD, MSCE⁷, Jeffrey R. Curtis, MD, MS, MPH⁸, Ted R. Mikuls, MD, MSPH^{1,2}

¹VA Nebraska-Western IA Health Care System, Omaha, NE

²Division of Rheumatology & Immunology, Department of Internal Medicine, University of Nebraska Medical Center, Omaha, NE

³Center for Comprehensive Access Delivery Research, Iowa City VA, Iowa City, IA

⁴University of Iowa Hospitals and Clinics, Iowa City, IA

⁵Department of Biostatistics, College of Public Health, University of Nebraska Medical Center, Omaha, NE

⁶VA Salt Lake City & University of Utah, Salt Lake City, UT

⁷Corporal Michael J. Crescenz VA & University of Pennsylvania, Philadelphia, PA

⁸University of Alabama at Birmingham, Birmingham, AL

Abstract

Objective: To determine the performance of administrative-based algorithms for classifying interstitial lung disease (ILD) complicating rheumatoid arthritis (RA).

Methods: Participants in a large, multicenter RA registry were screened for ILD using International Classification of Diseases (ICD) codes. Medical record review confirmed ILD among participants screening positive and a random sample of those screening negative. ICD and procedure codes, provider specialty, and dates were extracted from Veterans Affairs administrative data to construct ILD algorithms. Performance of these algorithms against medical record review was assessed by sensitivity, specificity, positive predictive value (PPV), negative predictive value, and Kappa using inverse probability weighting to account for sampling methods.

Results: Medical records of 536 RA patients were reviewed, confirming 182 (stringent definition) and 203 (relaxed definition) ILD cases. Initially, we identified 2 ICD codes from inpatient or outpatient encounters as optimal discriminating factors (specificity 96.0%, PPV 65.5%, Kappa 0.70). Subsequently, we constructed a set of ICD-9/10 codes that improved algorithm specificity (specificity 96.8%, PPV 69.5%, Kappa 0.72). Algorithms that included a

Address correspondence: Bryant R. England, MD; 986270 Nebraska Medical Center, Omaha, NE 68198-6270. (P) 402.559.7288 (Fax) 402.559.6788 bryant.england@unmc.edu.

Disclosures: None

pulmonologist diagnosis or chest CT plus pulmonary function testing or lung biopsy had improved performance (specificity 98.0%, PPV 77.4%, Kappa 0.75). PPV increased with exclusion of other ILD causes (78.5%), in comparisons with the relaxed ILD definition (82.4%), and in sensitivity analyses (83.4-86.3%). Gains in specificity and PPV with greater algorithm requirements were accompanied by declines in sensitivity.

Conclusion: Administrative algorithms with optimal combinations of ICD codes, provider specialty, diagnostic testing, and exclusion of other ILD causes accurately classify ILD in RA.

Keywords

rheumatoid arthritis; interstitial lung disease; administrative data

Introduction

Interstitial lung disease (ILD) clinically affects between 5-10% of rheumatoid arthritis (RA) patients resulting in poor long-term outcomes including reduced survival and greater functional disability (1–3). Given disease heterogeneity and lack of well-characterized classification criteria for RA-ILD, the case definitions used and prevalence estimates reported for RA-ILD are highly variable between studies (1, 3–5). Administrative data sources are increasingly being utilized in RA outcomes research, primarily to facilitate investigations examining predictors of disease-related outcomes or the safety and effectiveness of disease-modifying anti-rheumatic drugs (DMARDs) (6). Yet, only a few studies have begun to leverage these large administrative databases to study RA-ILD (7–11).

Prior studies utilizing administrative databases for RA-ILD research have constructed RA-ILD cohorts or identified ILD outcomes in RA patients using claims databases (7, 8, 10), death records (9), or national patient registries (11). All have used diagnostic codes for RA and ILD in combination, though additional requirements for RA diagnosis such as DMARD receipt and specific ILD diagnostic codes selected has varied between studies. To enhance specificity, authors have required ILD diagnostic tests (7) or excluded other causes of ILD (e.g. sarcoidosis, hypersensitivity pneumonitis, pneumoconioses, etc.) (8). However, the validity of these algorithms has received only limited attention (12), hampering wider adoption of these methods for studying RA-ILD. With validated ILD algorithms, large administrative datasets could be leveraged for comparative effectiveness research and epidemiologic analyses, while deployment in electronic health records could enhance recruitment into patient registries or clinical trials.

The objective of this study was to develop and evaluate the performance of several different administrative algorithms for the identification of ILD in a multi-center RA registry. We hypothesized that administrative algorithms that included multiple ILD diagnostic codes, a pulmonologist diagnosis, procedure codes for computed tomography (CT) of the chest, pulmonary function tests (PFTs), or lung biopsy, and exclusion of other causes of ILD would accurately classify RA-ILD compared to a comprehensive review of medical records.

Patients and Methods

Patient selection

We selected subjects enrolled in the Veterans Affairs Rheumatoid Arthritis (VARA) registry, a multicenter, prospective cohort study of U.S. Veterans with RA initiated in 2003 (13). All subjects fulfilled the 1987 American College of Rheumatology (ACR) criteria for RA (14). Participants provided informed consent prior to enrollment and all sites (n=13) obtained local institutional review board approval. This study obtained approval from the VARA Scientific Ethics and Advisory Committee.

To enrich the study sample with ILD cases, we performed stratified subsampling through initial ILD screening. We queried national VA data in the Corporate Data Warehouse (CDW) to identify VARA participants with 1 inpatient or 2 outpatient ILD diagnostic codes (>30 days apart) from health care providers (physicians, physician assistants, and advanced practice nurses). International Classification of Diseases, 9th and 10th revision, Clinical Modification codes were selected from those previously proposed to ascertain ILD status or closely related codes (Table S1) (7, 9–12, 15). We performed detailed, systematic medical record review on all subjects identified through initial screening (n=293) and a random sample of all VARA subjects not identified by the ILD screening method (n=243) so as to be able to comment on the sensitivity of the selected ICD codes and ILD algorithms.

ILD data abstraction

Data was abstracted from the electronic medical records using the Compensation and Pension Record Interchange in a standardized fashion by three rheumatologists (BRE, TDM, NS) blinded to the results of the administrative algorithms using Research Electronic Data Capture (REDCap) (16, 17). Regardless of ILD screening status, participants' outpatient and inpatient clinical notes, imaging reports, pathology reports, and PFT results from the earliest available date in the medical record were reviewed and recorded. Data abstracted included pulmonologist diagnoses, other physician diagnoses, chest CT results, chest x-ray results, PFT results, lung biopsy results as well as dates corresponding to the aforementioned items. To ensure consistency between reviewers, charts were reviewed in sets of 5 in duplicate until >95% agreement on abstracted data was obtained between reviewers. As our reference standard, participants were classified as RA-ILD by medical record review using both stringent and relaxed ILD definitions. The stringent definition classified participants as RA-ILD if they had a pulmonologist diagnosis and imaging (chest CT or x-ray) findings of ILD or if they had a non-pulmonologist provider diagnosis plus two of the following: chest CT or x-ray findings interpreted by the reading radiologist as ILD, pathology from a lung biopsy consistent with ILD, or interpretation of PFTs as restrictive by the reading pulmonologist. The relaxed ILD definition additionally classified subjects as ILD who had a provider diagnosis of ILD (pulmonologist or non-pulmonologist) and either imaging findings consistent with ILD or pathology demonstrating ILD.

Algorithm development

We queried National VA data within the CDW from January 1, 1999 to August 31, 2018 for all participants to obtain the necessary components of each ILD algorithm. Data queried

included inpatient and outpatient encounters in the VA, inpatient and outpatient encounters occurring outside the VA and billed to the VA, specialty of outpatient encounters, and outpatient and inpatient procedures.

We tested the characteristics of possible administrative ILD algorithms in four stages. In the first stage, we tested the performance of algorithms using different encounter types (inpatient vs. outpatient) and frequency of ILD diagnostic codes (1 vs. 2). In stage 2, we compared different ICD-9 and ICD-10 code sets (Table S1). These code sets were created by removing ICD-9 and ICD-10 codes with descriptions including “unspecified” or “other”, those pertaining to rheumatoid lung, and those not consistently included in prior studies. Stage 3 testing compared algorithms that incorporated additional data available in administrative datasets that may improve algorithm specificity. These additional data were provider specialty on the ICD-9/10 diagnoses, and procedure codes for chest CT, PFTs, and lung biopsy procedures. Current Procedural Terminology (CPT) and ICD Procedural Codes for chest CT and lung biopsy were adapted from those used in idiopathic pulmonary fibrosis (IPF) algorithms (Table S2) (18, 19). In addition to CPT and ICD procedure codes, we also identified PFTs through the use of stop codes in National VA data which designate clinical services provided in PFT labs. In the final stage (stage 4), we excluded other causes of ILD recorded after the final ILD diagnosis using codes for pneumoconioses, radiation pneumonitis, hypersensitivity pneumonitis, and other connective tissue diseases (Table S3).

Statistical analysis

Enrollment characteristics of the VARA patients selected for these analyses were assessed descriptively, stratified by medical record review ILD classification status. Agreement between ILD algorithms and medical record review classification was assessed with percent agreement and Kappa statistics. Levels of agreement based on the Kappa statistic were interpreted as near perfect (values of 0.8-1.0), substantial (0.6-0.8), moderate (0.4-0.6), fair (0.2-0.4), or slight (0.0-0.2) (20). We also calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) along with 95% confidence intervals for each algorithm, treating medical record classification as the reference standard. All analyses accounted for the subsampling from the overall VARA registry by the use of inverse probability weighting (R package CompareTests) (21). This ensured that the prevalence of ILD in weighted analyses was consistent with the overall cohort. Algorithm selection through each stage was based on optimal Kappa values. Several sensitivity analyses were performed testing variations of administrative ILD algorithms and using medical record ILD definitions with fewer requirements. Analyses were conducted using Stata v15 (StataCorp) and R version 3.5.1 within the VA Informatics and Computing Infrastructure (VINCI). We report our study in accordance with proposed reporting guidelines for assessing the quality of validation studies of health administrative data (22).

Results

Enrollment characteristics

We identified 293 subjects in the VARA registry who met the initial ILD screening criteria and randomly selected 243 VARA participants who did not screen positive for ILD (Figure

1). Detailed medical record review performed on all 536 of these subjects confirmed 182 and 203 ILD cases using stringent and relaxed ILD definitions. Patient characteristics were reflective of the overall VARA registry and the VA population with a male predominance and mean age at enrollment in the 7th decade of life (Table 1). Those with ILD were older, more frequently rheumatoid factor positive, less likely to be treated with methotrexate, and more likely to receive prednisone at enrollment.

The majority of ILD cases occurred among those who screened positive for ILD (97.3% stringent and 96.6% relaxed), had a pulmonologist diagnosis (94.5% stringent and 84.7% relaxed), and had CT evidence of ILD (98.4% stringent and 96.1% relaxed) (Table 1). Approximately half of the ILD cases were prevalent at the time of enrollment into the registry, and the initial date of ILD diagnosis occurred after implementation of ICD-10 in 17 cases (21 cases relaxed ILD definition). Among non-ILD cases, pulmonologist ILD diagnosis was present in 1.4-1.5%, non-pulmonologist ILD diagnosis was present in 8.5% (stringent) and 2.7% (relaxed), and CT evidence of ILD was present in 11.3% (stringent) and 7.2% (relaxed).

Stage I: Frequency of diagnosis codes and encounter types

Performance of eight different algorithms (1A to 1H) reflecting differences in frequency, encounter types, and date ranges for ILD diagnosis codes in classifying ILD is shown in Table 2. Kappa was greatest for algorithms 1D (0.71) and 1F (0.70). Performance was similar in classifying ILD with the relaxed definition (Kappa 0.71). Sensitivity ranged from 76.3-81.7% and specificity ranged from 96.0-97.1%, but PPV of these algorithms were modest (65.5-73.9%). Because of their equivalent performance, Algorithm 1F, which required 2 diagnosis codes 30 days apart from either inpatient or outpatient encounters, was selected for further testing.

Stage II: Diagnosis code selection

Exclusion of ICD-10 codes J84.2 and J99 did not result in any difference in ILD classification (Table 3). Exclusion of rheumatoid lung codes (ICD-9: 714.81; ICD-10: M05.1x) minimally attenuated sensitivity and NPV while improving specificity and PPV. Kappa was improved from the all-inclusive ICD code algorithm when rheumatoid lung codes were excluded. Medical record review identified RA-related pleural effusions and pulmonary nodules as reasons for these codes in the absence of ILD. Algorithm performance measured by Kappa worsened when “unspecified” and “other” ILD codes were excluded. Based on these performance characteristics, we constructed algorithm 2H with the following ICD codes: ICD-9 515.x, 516.3, 516.8, 516.9 and ICD-10 J84.1, J84.89, J84.9. This algorithm had the best Kappa (0.72), specificity (96.8%, 97.3% relaxed ILD definition), and PPV (69.5%, 75.3% relaxed ILD definition), with minimal attenuation of sensitivity (80.6%, 74.4% relaxed ILD definition). Algorithm 2H was thus used for further comparisons in Stage III testing.

Stage III: Provider specialty and diagnostic studies

The additional requirement of a pulmonologist diagnosis (algorithm 3A) increased the specificity from 96.8% (algorithm 2H) to 98.5%, PPV from 69.5% to 79.9%, and had

substantial agreement by Kappa (0.68; 0.63 relaxed ILD definition) (Table 4). Requiring a rheumatologist diagnosis (algorithm 3B) modestly improved specificity but reduced sensitivity and overall algorithm performance by Kappa. Algorithms requiring a CT or PFTs between 7 and 180 days prior to ILD diagnosis (algorithms 3C and 3D) also modestly improved specificity and PPV while reducing sensitivity and NPV. Kappa for these algorithms were 0.72-0.74 (0.69-0.73 relaxed ILD definition). The requirement of a lung biopsy (algorithm 3E) was highly specific (99.9%) but poorly sensitive (9.3%), resulting in a PPV of 87.5% and only slight agreement by Kappa (0.15). Requiring a CT, PFTs, or lung biopsy in addition to a pulmonologist diagnosis (algorithm 3F and 3G) modestly affected specificity, PPV, and Kappa. As sensitivity was reduced with the requirement of a pulmonologist diagnosis (63.7%), we tested an algorithm requiring a chest CT plus either PFTs or a lung biopsy (algorithm 3H) and an algorithm requiring either a pulmonologist diagnosis or chest CT plus either PFTs or a lung biopsy (algorithm 3I). Sensitivity improved in these algorithms to 70.1% and 76.4%. Algorithm 3H had a higher specificity and PPV, but sensitivity and agreement by Kappa were better for algorithm 3I (Kappa 0.75 vs. 0.73; 0.72 vs. 0.70 sensitive ILD definition). Algorithm performance for identifying ILD was similar by Kappa between algorithms 3A, 3C, 3D, 3G, 3H, and 3I (0.68-0.75, 0.63-0.72 relaxed ILD definition) indicating substantial agreement.

Stage IV: Exclusion of other ILD

Using the top-performing models (3A, 3C, 3D, 3G, 3H, 3I), we then excluded those with a diagnosis code for other causes of ILD that occurred on or after the date of the last ILD code recorded, following approaches in IPF (18, 19, 23). After excluding other causes of ILD, there was modest improvement in specificity and PPV (range 0.7-1.2% improvement, Table 5) for each algorithm. Specificity for these algorithms ranged from 97.9-98.8%) and PPVs ranged from 76.0-82.9% (80.3-86.6% relaxed ILD definition). Overall performance by Kappa was similar after excluding other causes of ILD (Kappa 0.67-0.74, 0.61-0.71 relaxed ILD definition). Algorithm 4I had the best agreement with medical record review (Kappa 0.74, 0.70 relaxed ILD definition), indicating substantial agreement. Performance metrics for this algorithm were sensitivity 73.2% (65.4% relaxed ILD definition), specificity 98.2% (98.5% relaxed ILD definition), and PPV 78.5% (82.4% relaxed ILD definition).

Sensitivity Analyses

Because ILD may be detected on different types of CT scans (e.g. high-resolution, CT-angiogram, low dose CT for lung cancer screening), we tested both broad CT codes and specific CT codes. Algorithms performed similarly regardless of the CT scan codes utilized (Table S4). Similarly, we tested algorithms with specific (open via thoracotomy and bronchoscopy) and broad (open, bronchoscopy, and percutaneous) lung biopsy codes. These algorithms also performed similarly, with excellent specificity but limited sensitivity. We tested algorithms that only required diagnostic testing (CT, PFT, and lung biopsy) to be completed at least 7 days prior to ILD diagnosis, rather than within a 7-180 day window. These algorithms had modestly improved sensitivity and Kappa values. We tested a broader time window for excluding other causes of ILD, excluding cases if a diagnostic code for other causes of ILD was ever recorded in national VA data. These algorithms reduced sensitivity and Kappa values. Because some non-ILD cases had clinical diagnoses or

diagnostic testing for ILD but did not fulfill primary ILD definitions, we compared algorithm 4I against two additional ILD definitions with fewer requirements. Specificity was 98.6% and PPV improved to 83.4% and 86.3% in these models, with Kappa values still suggesting substantial agreement (Kappa 0.67).

Discussion

To facilitate the use of administrative data for RA-ILD research, we have characterized the performance of administrative algorithms for identifying ILD among RA patients compared to detailed medical record review. Algorithms including specific ICD-9 and ICD-10 ILD codes attributed to multiple encounters, pulmonologist diagnosis or diagnostic testing, and exclusion of other causes of ILD were able to accurately classify ILD. The best performing algorithm (algorithm 4I, Table 5) requiring 2 ILD diagnosis codes at least 30 days apart, a single pulmonologist diagnosis for ILD or CT and either (PFTs or lung biopsy) 7-180 days prior to ILD diagnosis, and exclusion of other ILD causes after the last ILD diagnosis yielded substantial agreement to medical record review by Kappa (0.74). PPVs for this algorithm ranged from 78.5-86.3% depending on the requirements of the ILD reference standard. Because there is a trade-off between sensitivity and specificity with these ILD algorithms and differences in availability of algorithm components within different datasets, the choice of algorithm will depend on the purpose of the study and available data. Our results provide detailed data on the performance of several ILD algorithms that will support investigator selection of ILD case finding approaches in future studies.

Similar to administrative algorithms developed to identify RA (24) and other rheumatic conditions (25–27), the requirement of multiple diagnosis codes for ILD separated over time enhanced the specificity and PPV of administrative ILD algorithms. Importantly, these results demonstrate that some diagnosis codes incorporated into prior RA-ILD algorithms lack specificity for ILD (7–11). Most notably among these were ICD-9 (714.81) and ICD-10 (M05.1x) codes for “rheumatoid lung”. Because there are numerous pulmonary manifestations of RA including ILD, obstructive lung disease, nodules, and pleural effusions (28), these codes may be used for these other entities in the setting of RA. Indeed, pleural effusions and pulmonary nodules were reasons for these codes occurring in the absence of ILD. We recommend the following ICD-9 and ICD-10 codes for identifying ILD in RA patients: ICD-9 515.x, 516.3, 516.8 and ICD-10 J84.1, J84.89, J84.9 (bolded in Table S1).

Our results illustrate that requiring a pulmonologist diagnosis of ILD achieves excellent specificity (algorithm 3A, 98.5%), but at the expense of sensitivity (63.7%). This reduction in sensitivity may be exaggerated in our cohort because pulmonologist diagnoses outside the VA health care system would not be captured by our algorithms. Therefore, algorithms with pulmonologist diagnosis may actually perform better in other settings. At least in our sample, further requiring diagnostic testing such as chest CT, PFTs, and lung biopsy did not significantly improve the PPV or Kappa from algorithms that already included a pulmonologist diagnosis of ILD. Eliminating the requirement of a pulmonologist diagnosis, we found that requiring a recent CT plus PFTs or lung biopsy in the prior 6 months achieved a similar PPV (81.9% vs. 79.9%). Using broad vs. specific CPT codes for these diagnostic tests rendered little impact on model performance. Combining either a pulmonologist

diagnosis or the aforementioned diagnostic tests (algorithm 3I) optimized the sensitivity while preserving a reasonable specificity, leading to optimal algorithm performance by Kappa. Further refining this algorithm with exclusion of other ILD causes maintained overall algorithm performance while modestly increasing PPV (algorithm 4I).

Because we performed detailed medical record review on a random sample of VARA participants who did not screen positive for ILD, we were able to assess not only the specificity and PPV but also the sensitivity and NPV. Algorithms that incorporated pulmonologist diagnosis or diagnostic testing obtained the highest PPV (78.5%), but algorithms without these additional criteria had similar Kappa values (0.72), reflecting improvements in specificity at the expense of sensitivity. As the overall performance (measured by Kappa) did not significantly differ for several algorithms, choice should be directed by specific study needs for either greater sensitivity or specificity and data availability. For example, completion of RA-ILD comparative effectiveness and outcomes research in large administrative datasets will require specific ILD algorithms, such as algorithm 4I (PPV 78.5-86.3%). Epidemiologic studies of RA-ILD, rather, may implement both a specific (algorithm 4I) and more sensitive algorithm (algorithm 2H, NPV 98.2% and 97.9%), recognizing the “truth” lies between the estimates from the specific and sensitive algorithms.

The generalizability of our findings may be limited by male predominance of the VARA registry, unique exposures of the Veteran population, as well as the coding practices represented by the 13 VARA-associated VA medical centers at which this work was conducted. However, the Veterans Health Administration represents the largest integrated health care system in the US with reduced barriers to access among its beneficiaries and a single electronic health record. Patients may receive care outside the VA, which affects capture by administrative algorithms and medical record review. To mitigate this, we reviewed the clinical notes for mention of outside care and selected claims originating from non-VA care when constructing our administrative ILD algorithms. Supporting the validity of our findings is that limited testing in a prior study of Kaiser Permanente Northern California found a PPV of 63% for 2 ILD diagnosis codes using imaging reports for ILD validation (12). This is in agreement with the PPV of 65.5% for a similar algorithm (algorithm 1F) in our study. The derived ILD algorithms are also currently being externally validated in additional non-VA datasets.

Validation of ILD diagnoses through medical records was retrospective, with diagnostic testing dictated through regular clinical care and interpreted by the treating providers. Additionally, there are currently no widely accepted classification criteria for RA-ILD. Because of the potential of misclassification in our reference standard, we used stringent and relaxed primary ILD definitions for each stage of algorithm development and testing. Furthermore, because some participants not fulfilling the stringent or relaxed primary ILD definitions had clinical diagnoses or diagnostic testing suggestive of ILD, we performed sensitivity analyses comparing optimal algorithms to ILD definitions with fewer requirements. Overall algorithm performance was consistent between ILD definitions, with increased PPV using ILD definitions with fewer requirements. Because only approximately 10% of ILD cases were initially diagnosed after ICD-10 implementation, our findings may

underestimate ICD-10 code contribution to ILD classification, a possibility that will need to be addressed in future research. Finally, we assessed the accuracy of ILD algorithms within a cohort fulfilling 1987 ACR criteria (14), and the performance of these algorithms may vary if applied outside of this setting (e.g. in combination with administrative algorithms to identify RA). However, the results from our study will serve as a valuable benchmark for future efforts focused on external validation. Given the high specificity of administrative algorithms for RA (24), we would anticipate to observe minimal reductions in the specificity of these ILD algorithms.

In conclusion, we have demonstrated that administrative algorithms can be used to accurately identify ILD in a RA cohort. Our results detail the performance metrics of these different algorithms for ILD, which can be applied to large administrative data sources to perform further clinical and epidemiologic study of RA-ILD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: BRE is supported by the UNMC Physician-Scientist Training Program, UNMC Internal Medicine Scientist Development Award, UNMC Mentored Scholars Program, RRF Scientist Development Award, and the National Institute of General Medical Sciences (U54GM115458). TRM is supported by a VA Merit Award (BX004790) and grants from the National Institutes of Health: National Institute of General Medical Sciences (U54GM115458), National Institute on Alcohol Abuse and Alcoholism (R25AA020818), and National Institute of Arthritis and Musculoskeletal and Skin Diseases (2P50AR60772). Dr. Baker is supported by a VA Merit Award from Clinical Science Research & Development (I01 CX001703). Dr. Cannon is supported by Specialty Care Center of Innovation, Veterans Health Administration, Department of Veterans Affairs. Dr. Curtis is supported in part by the Patient Centered Outcomes Research Institute (PCORI).

References

1. Bongartz T, Nannini C, Medina-Velasquez YF, Achenbach SJ, Crowson CS, Ryu JH, et al. Incidence and mortality of interstitial lung disease in rheumatoid arthritis: a population-based study. *Arthritis Rheum.* 2010;62(6):1583–91. [PubMed: 20155830]
2. Kelly CA, Saravanan V, Nisar M, Arthanari S, Woodhead FA, Price-Forbes AN, et al. Rheumatoid arthritis-related interstitial lung disease: associations, prognostic factors and physiological and radiological characteristics—a large multicentre UK study. *Rheumatology.* 2014;53(9):1676–82. [PubMed: 24758887]
3. Koduri G, Norton S, Young A, Cox N, Davies P, Devlin J, et al. Interstitial lung disease has a poor prognosis in rheumatoid arthritis: results from an inception cohort. *Rheumatology.* 2010;49(8):1483–9. [PubMed: 20223814]
4. Carmona L, Gonzalez-Alvaro I, Balsa A, Angel Belmonte M, Tena X, Sanmarti R. Rheumatoid arthritis in Spain: occurrence of extra-articular manifestations and estimates of disease severity. *Ann Rheum Dis.* 2003;62(9):897–900. [PubMed: 12922967]
5. Gabbay E, Tarala R, Will R, Carroll G, Adler B, Cameron D, et al. Interstitial lung disease in recent onset rheumatoid arthritis. *Am J Respir Crit Care Med.* 1997;156(2 Pt 1):528–35. [PubMed: 9279235]
6. Hudson M, Tascilar K, Suissa S. Comparative effectiveness research with administrative health data in rheumatoid arthritis. *Nat Rev Rheumatol.* 2016;12(6):358–66. [PubMed: 27080692]
7. Curtis JR, Sarsour K, Napalkov P, Costa LA, Schulman KL. Incidence and complications of interstitial lung disease in users of tocilizumab, rituximab, abatacept and anti-tumor necrosis factor alpha agents, a retrospective cohort study. *Arthritis Res Ther.* 2015;17:319. [PubMed: 26555431]

8. Raimundo K, Solomon JJ, Olson AL, Kong AM, Cole AL, Fischer A, et al. Rheumatoid Arthritis-Interstitial Lung Disease in the United States: Prevalence, Incidence, and Healthcare Costs and Mortality. *J Rheumatol*. 2018.
9. Olson AL, Swigris JJ, Sprunger DB, Fischer A, Fernandez-Perez ER, Solomon J, et al. Rheumatoid arthritis-interstitial lung disease-associated mortality. *Am J Respir Crit Care Med*. 2011;183(3):372–8. [PubMed: 20851924]
10. Suissa S, Hudson M, Ernst P. Leflunomide use and the risk of interstitial lung disease in rheumatoid arthritis. *Arthritis Rheum*. 2006;54(5):1435–9. [PubMed: 16645972]
11. Hyldgaard C, Hilberg O, Pedersen AB, Ulrichsen SP, Lokke A, Bendstrup E, et al. A population-based cohort study of rheumatoid arthritis-associated interstitial lung disease: comorbidity and mortality. *Ann Rheum Dis*. 2017;76(10):1700–6. [PubMed: 28611082]
12. Herrinton LJ, Harrold LR, Liu L, Raebel MA, Taharka A, Winthrop KL, et al. Association between anti-TNF-alpha therapy and interstitial lung disease. *Pharmacoepidemiol Drug Saf*. 2013;22(4):394–402. [PubMed: 23359391]
13. Mikuls TR, Reimold A, Kerr GS, Cannon GW. Insights and Implications of the VA Rheumatoid Arthritis Registry. *Fed Pract*. 2015;32(5):24–9.
14. 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(3):315–24. [PubMed: 3358796]
15. Dixon WG, Hyrich KL, Watson KD, Lunt M, Consortium BCC, Symmons DP, et al. Influence of anti-TNF therapy on mortality in patients with rheumatoid arthritis-associated interstitial lung disease: results from the British Society for Rheumatology Biologics Register. *Ann Rheum Dis*. 2010;69(6):1086–91. [PubMed: 20444754]
16. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap Consortium: Building an International Community of Software Platform Partners. *J Biomed Inform*. 2019;103208. [PubMed: 31078660]
17. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377–81. [PubMed: 18929686]
18. Ley B, Urbania T, Husson G, Vittinghoff E, Brush DR, Eisner MD, et al. Code-based Diagnostic Algorithms for Idiopathic Pulmonary Fibrosis. Case Validation and Improvement. *Ann Am Thorac Soc*. 2017;14(6):880–7. [PubMed: 28355518]
19. Raghu G, Weycker D, Edelsberg J, Bradford WZ, Oster G. Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2006;174(7):810–6. [PubMed: 16809633]
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74. [PubMed: 843571]
21. Katki HA, Li Y, Edelstein DW, Castle PE. Estimating the agreement and diagnostic accuracy of two diagnostic tests when one test is conducted on only a subsample of specimens. *Stat Med*. 2012;31(5):436–48. [PubMed: 22139832]
22. Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttman A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol*. 2011;64(8):821–9. [PubMed: 21194889]
23. Esposito DB, Lanes S, Donneyong M, Holick CN, Lasky JA, Lederer D, et al. Idiopathic Pulmonary Fibrosis in United States Automated Claims. Incidence, Prevalence, and Algorithm Validation. *Am J Respir Crit Care Med*. 2015;192(10):1200–7. [PubMed: 26241562]
24. Chung CP, Rohan P, Krishnaswami S, McPheeters ML. A systematic review of validated methods for identifying patients with rheumatoid arthritis using administrative or claims data. *Vaccine*. 2013;31 Suppl 10:K41–61. [PubMed: 24331074]
25. Moores KG, Sathe NA. A systematic review of validated methods for identifying systemic lupus erythematosus (SLE) using administrative or claims data. *Vaccine*. 2013;31 Suppl 10:K62–73. [PubMed: 24331075]
26. Harrold LR, Saag KG, Yood RA, Mikuls TR, Andrade SE, Fouayzi H, et al. Validity of gout diagnoses in administrative data. *Arthritis Rheum*. 2007;57(1):103–8. [PubMed: 17266097]

27. Shrestha S, Dave AJ, Losina E, Katz JN. Diagnostic accuracy of administrative data algorithms in the diagnosis of osteoarthritis: a systematic review. *BMC Med Inform Decis Mak.* 2016;16:82. [PubMed: 27387323]
28. Shaw M, Collins BF, Ho LA, Raghu G. Rheumatoid arthritis-associated lung disease. *Eur Respir Rev.* 2015;24(135):1–16. [PubMed: 25726549]

Significance and Innovations

- Utilization of large data sources to study rheumatoid arthritis-interstitial lung disease (RA-ILD) has been limited by the lack of validated administrative based algorithms for case identification.
- We compared the performance of administrative based algorithms for ILD against detailed medical record review in a multi-center RA registry using stratified sampling.
- Administrative based algorithms with optimal diagnostic codes, inclusion of pulmonologist diagnosis and/or diagnostic tests, and exclusion of other causes of ILD accurately classify RA-ILD.
- Several ILD algorithms have similar overall performance, so selection should be based on the preference for specificity or sensitivity and the availability of algorithm components in the data source.

Figure 1. Derivation of study sample and classification of interstitial lung disease by medical record review.

The Veterans Affairs Rheumatoid Arthritis registry was screened for at least 2 outpatient or 1 inpatient discharge diagnosis of interstitial lung disease (ILD) using International Classification of Diseases, 9th and 10th edition codes. Detailed medical record review was performed for all subjects who screened positive and a random sample of those who screened negative to validate ILD diagnoses by abstracting physician diagnoses (pulmonologist and non-pulmonologist), imaging findings (computed tomography and chest

x-ray), pulmonary function tests, and pathology findings. Using the primary ILD definition, 182 cases were identified, while using a more sensitive definition resulted in 203 ILD cases. Abbreviations: CT, computed tomography; ILD, interstitial lung disease; PFTs, pulmonary function tests; VARA, Veterans Affairs Rheumatoid Arthritis registry. Study flow diagram.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Characteristics of study cohort at registry enrollment by ILD status

	Stringent ILD definition		Relaxed ILD definition	
	ILD Mean (SD) or N (%) [*]	No ILD Mean (SD) or N (%) [*]	ILD Mean (SD) or N (%) [*]	No ILD Mean (SD) or N (%) [*]
N	182	354	203	333
Patient Characteristics				
Age, years	66.2 (9.8)	63.6 (10.4) [†]	66.4 (9.8)	63.4 (10.4) [†]
Male sex	173 (95.6)	323 (91.2)	193 (95.5)	303 (91.0)
Caucasian	138 (75.8)	273 (77.3)	154 (76.2)	257 (77.2)
Smoking status				
Current	53 (29.6)	89 (25.9)	59 (29.8)	83 (25.5)
Former	100 (55.9)	187 (54.4)	109 (55.1)	178 (54.8)
Never	26 (14.5)	68 (19.8)	30 (15.2)	64 (19.7)
High-school education	147 (86.5)	273 (86.7)	161 (85.6)	259 (87.2)
RA duration, years	11.7 (12.6)	11.1 (11.1)	12.0 (12.7)	10.9 (10.9)
Anti-CCP antibody +	137 (83.0)	246 (80.4)	152 (83.5)	231 (79.9)
Rheumatoid factor +	144 (87.8)	243 (79.2) [†]	158 (87.3)	229 (79.0) [†]
DAS28	4.2 (1.4)	4.0 (1.6)	4.2 (1.4)	4.0 (1.6)
MDHAQ	1.0 (0.6)	1.0 (0.6)	1.0 (0.6)	1.0 (0.6)
Methotrexate	51 (30.7)	163 (52.2) [†]	59 (32.1)	155 (52.7) [†]
bDMARDs	52 (28.6)	79 (22.3)	56 (27.6)	75 (22.5)
Prednisone	103 (62.1)	123 (39.4) [†]	113 (61.4)	113 (38.4) [†]
Interstitial Lung Disease Status (by medical record review)				
Screened positive for ILD	177 (97.3)	116 (32.8) [†]	196 (96.6)	97 (29.1) [†]
Pulmonologist diagnosis ILD	172 (94.5)	5 (1.4) [†]	172 (84.7)	5 (1.5) [†]
Non-pulmonologist diagnosis ILD	175 (96.2)	30 (8.5) [†]	196 (96.6)	9 (2.7) [†]
Imaging consistent with ILD	182 (100.0)	48 (13.6) [†]	202 (99.5)	28 (8.4) [†]
CT evidence of ILD	179 (98.4)	40 (11.3) [†]	195 (96.1)	24 (7.2) [†]
Restrictive pattern on PFTs	98 (53.9)	35 (9.9) [†]	99 (48.8)	34 (10.2) [†]
Pathology suggesting ILD	22 (12.1)	3 (0.9) [†]	23 (11.3)	2 (0.6) [†]
Prevalent at enrollment	91 (50.0)	-	101 (49.8)	-

^{*}
% of non-missing

[†]
P < 0.05 by independent t-test or chi-square test

Abbreviations: anti-CCP, anti-cyclic citrullinated peptide antibody; bDMARDs, biologic disease-modifying anti-rheumatic drugs; CT, computed tomography; DAS28, 28-joint disease activity score; MDHAQ, multidimensional health assessment questionnaire; ILD, interstitial lung disease; PFTs, pulmonary function tests; RA, rheumatoid arthritis; SD, standard deviation

Table 2.

Performance of algorithms with various encounter type and diagnostic code frequency for classifying interstitial lung disease in rheumatoid arthritis (Stage I)

Algorithm	Description ^a	Sensitivity	Specificity	PPV	NPV	%Agreement	Kappa
<i>Stringent ILLD definition</i>							
1A	1 outpatient diagnosis	85.8 (68.8, 94.3)	91.2 (89.9, 92.4)	48.2 (40.1, 56.3)	98.5 (96.1, 99.5)	90.8	0.57 (0.48, 0.65)
1B	1 discharge diagnosis	48.4 (38.9, 58.0)	96.9 (96.3, 97.3)	58.9 (52.4, 65.2)	95.3 (93.2, 96.7)	92.7	0.49 (0.41, 0.57)
1C	1 outpatient or discharge diagnosis	87.3 (68.9, 95.6)	89.8 (88.6, 91.0)	44.7 (37.7, 52.0)	98.7 (96.0, 99.6)	89.6	0.54 (0.46, 0.52)
1D	2 outpatient diagnosis, >30 days apart	80.7 (65.7, 90.1)	96.5 (95.6, 97.3)	68.2 (59.6, 75.7)	98.2 (96.1, 99.2)	95.2	0.71 (0.62, 0.79)
1E	1 discharge diagnosis or 2 outpatient diagnosis, >30 days apart	85.2 (67.8, 94.0)	94.5 (93.4, 95.4)	58.7 (51.0, 66.0)	98.6 (96.2, 99.5)	93.7	0.66 (0.58, 0.74)
1F	^b 2 diagnoses, >30 days apart	81.7 (66.4, 91.0)	96.0 (95.0, 96.8)	65.5 (56.9, 73.3)	98.3 (96.1, 99.2)	94.8	0.70 (0.61, 0.78)
1G	^b 2 diagnoses, >30 days and 365 days apart	66.3 (55.0, 76.1)	97.0 (96.2, 97.7)	67.0 (58.0, 74.9)	96.9 (95.2, 98.1)	94.5	0.64 (0.55, 0.72)
1H	^b 2 diagnoses, >30 days and 730 days apart	72.1 (59.6, 81.9)	96.8 (95.9, 97.5)	67.2 (58.2, 75.1)	97.4 (95.6, 98.5)	94.7	0.67 (0.58, 0.75)
<i>Relaxed ILLD definition</i>							
1A	1 outpatient diagnosis	87.3 (72.5, 94.7)	92.8 (91.0, 94.2)	58.1 (47.9, 67.6)	98.4 (96.1, 99.4)	92.2	0.65 (0.56, 0.74)
1B	1 discharge diagnosis	44.8 (36.1, 53.7)	97.2 (96.7, 97.6)	63.8 (57.4, 69.7)	94.1 (91.6, 95.8)	91.9	0.48 (0.41, 0.56)
1C	1 outpatient or discharge diagnosis	89.3 (72.7, 96.3)	91.4 (89.7, 92.8)	53.8 (45.0, 62.4)	98.7 (96.0, 99.6)	91.1	0.62 (0.53, 0.71)
1D	2 outpatient diagnosis, >30 days apart	74.6 (60.5, 84.9)	97.1 (96.1, 97.8)	73.9 (65.3, 81.0)	97.2 (94.7, 98.5)	94.9	0.71 (0.62, 0.79)
1E	1 discharge diagnosis or 2 outpatient diagnosis, >30 days apart	80.0 (63.8, 90.0)	95.2 (94.1, 96.1)	64.6 (56.9, 71.6)	97.7 (95.0, 99.0)	93.7	0.68 (0.59, 0.76)
1F	^b 2 diagnoses, >30 days apart	76.3 (61.8, 86.5)	96.7 (95.6, 97.5)	71.6 (63.0, 78.9)	97.4 (94.8, 98.7)	94.6	0.71 (0.61, 0.79)
1G	^b 2 diagnoses, >30 days and 365 days apart	62.7 (51.4, 72.7)	97.6 (96.8, 98.3)	74.3 (65.4, 81.6)	96.0 (93.8, 97.5)	94.2	0.65 (0.56, 0.73)
1H	^b 2 diagnoses, >30 days and 730 days apart	68.0 (55.6, 78.2)	97.4 (96.5, 98.1)	74.4 (65.4, 81.7)	96.5 (94.2, 97.9)	94.5	0.68 (0.59, 0.76)

^a ICD-9: 515.x, 516.8, 516.9, 714.81; ICD-10: M05.1x, J84.1, J84.2, J84.89, J84.9, J99

^b outpatient or discharge diagnoses

Abbreviations: PPV, positive predictive value; NPV, negative predictive value; ILLD, interstitial lung disease; CT, computed tomography; PFT, pulmonary function test

Table 3.

Performance of algorithms with various interstitial lung disease diagnostic codes for classifying interstitial lung disease in rheumatoid arthritis (Stage II)

Algorithm	Description	Sensitivity	Specificity	PPV	NPV	%Agreement	Kappa
<i>Stringent ILD definition</i>							
2A	All ILD codes [*]	81.7 (66.4, 91.0)	96.0 (95.0, 96.8)	65.5 (56.9, 73.3)	98.3 (96.1, 99.2)	94.8	0.70 (0.61, 0.78)
2B	Exclude M05.1x (n=6 excluded)	81.0 (65.7, 90.5)	96.3 (95.4, 97.1)	67.0 (58.9, 74.2)	98.2 (96.1, 99.2)	95.1	0.71 (0.62, 0.78)
2C	Exclude J84.2 (n=0 excluded)	81.7 (66.4, 91.0)	96.0 (95.0, 96.8)	65.5 (56.9, 73.3)	98.3 (96.1, 99.2)	94.8	0.70 (0.61, 0.78)
2D	Exclude J84.89 & J84.9 (n=3 excluded)	76.9 (61.2, 87.5)	96.2 (95.3, 96.8)	64.9 (57.4, 71.7)	97.8 (95.5, 99.0)	94.5	0.67 (0.59, 0.75)
2E	Exclude J99 (n=0 excluded)	81.7 (66.4, 91.0)	96.0 (95.0, 96.8)	65.5 (56.9, 73.3)	98.3 (96.1, 99.2)	94.8	0.70 (0.61, 0.78)
2F	Exclude 714.81 (n=8 excluded)	81.3 (66.1, 90.6)	96.3 (95.3, 97.1)	67.1 (58.3, 74.9)	98.2 (96.1, 99.2)	95.0	0.71 (0.62, 0.78)
2G	Exclude 516.8, 516.9 (n=31 excluded)	74.6 (61.6, 84.3)	96.6 (95.6, 97.4)	67.3 (57.8, 75.5)	97.6 (95.7, 98.7)	94.8	0.68 (0.59, 0.76)
2H	ICD-9 515.x, 516.3, 516.8, 516.9; ICD-10 J84.1, J84.89, J84.9 (n=17 excluded)	80.6 (65.5, 90.1)	96.8 (95.8, 97.5)	69.5 (61.1, 76.7)	98.2 (96.1, 99.2)	95.4	0.72 (0.63, 0.80)
<i>Relaxed ILD definition</i>							
2A	All ILD codes [*]	76.3 (61.8, 86.5)	96.7 (95.6, 97.5)	71.6 (63.0, 78.9)	97.4 (94.8, 98.7)	94.6	0.71 (0.61, 0.79)
2B	Exclude M05.1x (n=6 excluded)	75.2 (60.9, 85.5)	97.0 (96.0, 97.7)	72.9 (64.9, 79.7)	97.3 (94.8, 98.6)	94.8	0.71 (0.62, 0.79)
2C	Exclude J84.2 (n=0 excluded)	76.3 (61.8, 86.5)	96.7 (95.6, 97.5)	71.6 (63.0, 78.9)	97.4 (94.8, 98.7)	94.6	0.71 (0.61, 0.79)
2D	Exclude J84.89 & J84.9 (n=3 excluded)	72.1 (57.7, 83.1)	96.8 (96.0, 97.5)	71.3 (63.9, 77.7)	96.9 (94.2, 98.4)	94.4	0.69 (0.59, 0.77)
2E	Exclude J99 (n=0 excluded)	76.3 (61.8, 86.5)	96.7 (95.6, 97.5)	71.6 (63.0, 78.9)	97.4 (94.8, 98.7)	94.6	0.71 (0.61, 0.79)
2F	Exclude 714.81 (n=8 excluded)	75.5 (61.2, 85.8)	96.9 (95.9, 97.7)	73.0 (64.1, 80.3)	97.3 (94.8, 98.6)	94.8	0.71 (0.62, 0.79)
2G	Exclude 516.8, 516.9 (n=31 excluded)	68.3 (56.0, 78.5)	97.1 (96.0, 97.8)	72.1 (62.5, 80.0)	96.5 (94.2, 97.9)	94.2	0.67 (0.57, 0.75)
2H	ICD-9 515.x, 516.3, 516.8, 516.9; ICD-10 J84.1, J84.89, J84.9 (n=17 excluded)	74.4 (60.3, 84.8)	97.3 (96.4, 98.0)	75.3 (67.0, 82.1)	97.2 (94.7, 98.5)	95.1	0.72 (0.63, 0.80)

^{*} ICD-9: 515.x, 516.3, 516.8, 516.9, 714.81; ICD-10: M05.1x, J84.1, J84.2, J84.89, J84.9, J99

Abbreviations: PPV, positive predictive value; NPV, negative predictive value; ILD, interstitial lung disease; CT, computed tomography; PFT, pulmonary function test

Table 4.

Performance of algorithms with various provider specialties and diagnostic testing for classifying interstitial lung disease in rheumatoid arthritis (Stage III)

Algorithm	Description*	Sensitivity	Specificity	PPV	NPV	%Agreement	Kappa
<i>Stringent ILD definition</i>							
3A	1 pulmonologist ILD diagnosis	63.7 (51.8, 74.1)	98.5 (97.7, 99.0)	79.9 (69.8, 87.2)	96.7 (94.7, 97.9)	95.5	0.68 (0.59, 0.77)
3B	1 rheumatologist ILD diagnosis	55.7 (45.3, 65.5)	98.0 (97.3, 98.5)	72.1 (62.9, 79.7)	96.0 (94.1, 97.3)	94.4	0.60 (0.51, 0.68)
3C	CT 7-180 days prior to ILD diagnosis	72.9 (58.8, 83.5)	97.7 (96.8, 98.4)	75.2 (65.3, 83.0)	97.5 (95.3, 98.6)	95.6	0.72 (0.62, 0.80)
3D	PFT 7-180 days prior to ILD diagnosis	75.0 (62.1, 84.6)	98.1 (97.5, 98.5)	78.1 (72.2, 83.1)	97.7 (95.8, 98.7)	96.1	0.74 (0.66, 0.81)
3E	Lung biopsy 7-180 days prior to ILD diagnosis	9.3 (7.3, 11.8)	99.9 (99.6, 99.9)	87.5 (67.6, 95.9)	92.2 (90.3, 93.7)	92.1	0.15 (0.12, 0.20)
3F	1 pulmonologist diagnosis and CT or lung biopsy 7-180 days prior to ILD diagnosis	58.2 (46.7, 68.9)	98.4 (98.0, 98.7)	77.5 (72.0, 82.2)	96.2 (94.0, 97.6)	95.0	0.64 (0.55, 0.72)
3G	1 pulmonologist diagnosis and CT or lung biopsy or PFTs 7-180 days prior to ILD diagnosis	63.0 (51.3, 73.4)	98.6 (97.9, 99.1)	81.0 (71.5, 87.9)	96.6 (94.6, 97.9)	95.6	0.69 (0.59, 0.77)
3H	CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis	70.1 (57.3, 80.4)	98.5 (98.0, 98.9)	81.9 (76.0, 86.6)	97.2 (95.1, 98.4)	96.0	0.73 (0.64, 0.81)
3I	Pulmonologist diagnosis or CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis	76.4 (62.7, 86.2)	98.0 (97.2, 98.5)	77.4 (69.1, 84.0)	97.9 (96.0, 98.9)	96.2	0.75 (0.66, 0.82)
<i>Relaxed ILD definition</i>							
3A	1 pulmonologist ILD diagnosis	55.5 (45.1, 65.4)	98.6 (97.8, 99.1)	81.6 (71.5, 88.6)	95.2 (92.9, 96.8)	94.3	0.63 (0.53, 0.72)
3B	1 rheumatologist ILD diagnosis	52.5 (42.7, 62.1)	98.5 (97.8, 99.0)	79.7 (70.7, 86.4)	95.0 (92.7, 96.6)	94.0	0.60 (0.51, 0.69)
3C	CT 7-180 days prior to ILD diagnosis	65.8 (53.2, 76.4)	98.1 (97.1, 98.7)	79.3 (69.4, 86.7)	96.2 (93.8, 97.8)	94.8	0.69 (0.59, 0.78)
3D	PFT 7-180 days prior to ILD diagnosis	68.2 (56.1, 78.3)	98.5 (98.0, 98.9)	83.3 (77.7, 87.6)	96.6 (94.3, 98.0)	95.5	0.73 (0.64, 0.80)
3E	Lung biopsy 7 days prior to ILD diagnosis	8.3 (6.7, 10.4)	99.9 (99.7, 99.9)	91.7 (72.1, 97.9)	90.8 (88.6, 92.6)	90.8	0.14 (0.11, 0.18)
3F	1 pulmonologist diagnosis and CT or lung biopsy 7-180 days prior to ILD diagnosis	50.9 (41.1, 60.6)	98.5 (98.1, 98.8)	79.3 (74.0, 83.7)	94.8 (92.3, 96.5)	93.8	0.59 (0.50, 0.67)
3G	1 pulmonologist diagnosis and CT or lung biopsy or PFTs 7-180 days prior to ILD diagnosis	55.0 (44.7, 64.8)	98.7 (98.0, 99.2)	82.8 (73.2, 89.4)	95.2 (92.9, 96.8)	94.4	0.63 (0.54, 0.72)
3H	CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis	62.8 (51.4, 72.9)	98.8 (98.3, 99.1)	85.4 (79.9, 89.7)	95.9 (93.5, 97.4)	95.1	0.70 (0.61, 0.78)
3I	Pulmonologist diagnosis or CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis	68.2 (55.8, 78.4)	98.3 (97.5, 98.8)	81.0 (72.8, 87.2)	96.6 (94.3, 98.0)	95.3	0.72 (0.62, 0.79)

* ICD-9: 515.x, 516.3, 516.8, 516.9; ICD-10: J84.1, J84.89, J84.9

Abbreviations: PPV, positive predictive value; NPV, negative predictive value; ILD, interstitial lung disease; CT, computed tomography; PFT, pulmonary function test

Table 5.
Performance of algorithms with exclusion of other causes of interstitial lung disease (Stage IV)

Algorithm	Description	Sensitivity	Specificity	PPV	NPV	%Agreement	Kappa
<i>Stringent ILD definition</i>							
4A	1 pulmonologist ILD diagnosis and exclusion other ILD ^a	60.4 (49.2, 70.7)	98.7 (97.9, 99.2)	81.1 (70.8, 88.4)	96.4 (94.4, 97.7)	95.4	0.67 (0.57, 0.75)
4C	CT 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	69.7 (56.5, 80.4)	97.9 (97.0, 98.6)	76.0 (65.9, 83.8)	97.2 (95.1, 98.4)	95.5	0.70 (0.60, 0.79)
4D	PFT 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	71.7 (59.6, 81.4)	98.2 (97.7, 98.6)	78.8 (72.7, 83.8)	97.4 (95.6, 98.5)	96.0	0.73 (0.65, 0.80)
4G	1 pulmonologist diagnosis and CT or lung biopsy or PFTs 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	59.8 (48.7, 70.0)	98.8 (98.1, 99.2)	81.9 (72.2, 88.7)	96.4 (94.4, 97.7)	95.5	0.67 (0.57, 0.75)
4H	CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	66.9 (54.9, 77.1)	98.7 (98.2, 99.0)	82.9 (76.9, 87.6)	96.9 (94.8, 98.1)	95.9	0.72 (0.63, 0.79)
4I	Pulmonologist diagnosis or CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	73.2 (60.3, 83.0)	98.2 (97.4, 98.7)	78.5 (70.1, 85.1)	97.6 (95.7, 98.6)	96.1	0.74 (0.65, 0.81)
<i>Relaxed ILD definition</i>							
4A	1 pulmonologist ILD diagnosis and exclusion other ILD ^a	52.8 (42.9, 62.4)	98.8 (98.0, 99.3)	82.9 (72.6, 90.0)	95.0 (92.7, 96.6)	94.2	0.62 (0.52, 0.70)
4C	CT 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	63.0 (51.1, 73.6)	98.3 (97.3, 98.9)	80.3 (70.1, 87.6)	96.0 (93.5, 97.5)	94.7	0.68 (0.58, 0.77)
4D	PFT 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	65.4 (53.9, 75.4)	98.7 (98.2, 99.0)	84.2 (78.6, 88.6)	96.3 (94.1, 97.7)	95.4	0.71 (0.62, 0.79)
4G	1 pulmonologist diagnosis and CT or lung biopsy or PFTs 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	52.2 (42.5, 61.7)	98.9 (98.2, 99.3)	83.7 (74.1, 90.3)	95.0 (92.7, 96.6)	94.3	0.61 (0.52, 0.70)
4H	CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	60.0 (49.2, 70.0)	98.9 (98.5, 99.3)	86.6 (81.0, 90.8)	95.6 (93.2, 97.2)	95.0	0.68 (0.59, 0.76)
4I	Pulmonologist diagnosis or CT and PFTs or lung biopsy 7-180 days prior to ILD diagnosis and exclusion other ILD ^a	65.4 (53.6, 75.5)	98.5 (97.7, 99.0)	82.4 (74.0, 88.5)	96.3 (94.1, 97.7)	95.2	0.70 (0.61, 0.78)

^a exclusion of other ILD using diagnostic codes for pneumoconioses, radiation, hypersensitivity pneumonitis, other connective tissue diseases on or after the last ILD diagnosis code date (see Table S3)

Abbreviations: PPV, positive predictive value; NPV, negative predictive value; ILD, interstitial lung disease; CT, computed tomography; PFT, pulmonary function test