Assessing the Severity of COVID-19 Lung Injury in Rheumatic Diseases Versus the General Population Using Deep Learning–Derived Chest Radiograph Scores

Naomi J. Patel,¹ Kristin M. D'Silva,¹ Matthew D. Li,¹ Tiffany Y. T. Hsu,² Michael Diliorio,² Xiaqing Fu,¹ Claire Cook,¹ Lauren Prisco,² Lily Martin,² Kathleen M. M. Vanni,² Alessandra Zaccardelli,² Yuqing Zhang,¹ Jayashree Kalpathy-Cramer,¹ Jeffrey A. Sparks,² and Zachary S. Wallace¹

Objective. COVID-19 patients with rheumatic disease have a higher risk of mechanical ventilation than the general population. The present study was undertaken to assess lung involvement using a validated deep learning algorithm that extracts a quantitative measure of radiographic lung disease severity.

Methods. We performed a comparative cohort study of rheumatic disease patients with COVID-19 and ≥1 chest radiograph within ±2 weeks of COVID-19 diagnosis and matched comparators. We used unadjusted and adjusted (for age, Charlson comorbidity index, and interstitial lung disease) quantile regression to compare the maximum pulmonary x-ray severity (PXS) score at the 10th to 90th percentiles between groups. We evaluated the association of severe PXS score (>9) with mechanical ventilation and death using Cox regression.

Results. We identified 70 patients with rheumatic disease and 463 general population comparators. Maximum PXS scores were similar in the rheumatic disease patients and comparators at the 10th to 60th percentiles but significantly higher among rheumatic disease patients at the 70th to 90th percentiles (90th percentile score of 10.2 versus 9.2; adjusted P = 0.03). Rheumatic disease patients were more likely to have a PXS score of >9 (20% versus 11%; P = 0.02), indicating severe pulmonary disease. Rheumatic disease patients with PXS scores >9 versus ≤9 had higher risk of mechanical ventilation (hazard ratio [HR] 24.1 [95% confidence interval (95% CI) 6.7, 86.9]) and death (HR 8.2 [95% CI 0.7, 90.4]).

Conclusion. Rheumatic disease patients with COVID-19 had more severe radiographic lung involvement than comparators. Higher PXS scores were associated with mechanical ventilation and will be important for future studies leveraging big data to assess COVID-19 outcomes in rheumatic disease patients.

INTRODUCTION

Patients with rheumatic diseases and COVID-19 may have higher risk of mechanical ventilation and other poor outcomes versus general population comparators who also have COVID-19; factors driving these associations may include differences in comorbidity burden and immunosuppressive medication use (1–6). These associations suggest that rheumatic disease patients have more severe COVID-19 lung disease, but the pulmonary manifestations of COVID-19 remain poorly understood in this population.

¹Naomi J. Patel, MD, Kristin M. D'Silva, MD, MPH, Matthew D. Li, MD, Xiaqing Fu, MS, Claire Cook, MPH, Yuqing Zhang, ScD, Jayashree Kalpathy-Cramer, PhD, Zachary S. Wallace, MD, MSc; Massachusetts General Hospital and Harvard Medical School, Boston; ²Tiffany Y. T. Hsu, MD, PhD, Michael Diliorio, MD, Lauren Prisco, BA, Lily Martin, BS, Kathleen M. M. Vanni, BA, Alessandra Zaccardelli, MS, Jeffrey A. Sparks, MD, MMSc; Brigham and Women’s Hospital and Women’s Hospital and Harvard Medical School, Boston, Massachusetts.

Dr. Patel and D’Silva contributed equally to this work. Drs. Sparks and Wallace contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24883&file=acr24883-sup-0001-Disclosureform.pdf.

Address correspondence via email to Zachary S. Wallace, MD, MSc, at zswallace@mgh.harvard.edu. Submitted for publication November 17, 2021; accepted in revised form March 15, 2022.
SIGNIFICANCE & INNOVATIONS

• Patients with rheumatic disease and COVID-19 may be at higher risk of poor outcomes, including mechanical ventilation and death, but differences in the severity of imaging findings of COVID-19-associated lung disease that may contribute to these associations remain poorly understood.
• Using the pulmonary x-ray severity (PXS) score, a validated measure of lung disease severity in COVID-19, we found that patients with rheumatic disease were more likely to have more severe lung disease compared to general population comparators with COVID-19.
• Among rheumatic disease patients, those with higher PXS scores (representing severe radiographic lung disease) were more likely to require mechanical ventilation or die than those with lower PXS scores.
• PXS scores may be leveraged in future studies using big data to assess outcomes of COVID-19 in patients with rheumatic disease.

Among patients with symptomatic COVID-19 in the general population, radiographic severity correlates with severe COVID-19 outcomes (7). However, there can be significant variability in the interpretation and grading of disease severity on radiographs. A convolutional Siamese neural network-based algorithm was previously validated to estimate COVID-19 radiographic severity on portable chest radiographs, yielding pulmonary x-ray severity (PXS) scores (8). PXS scores correlate with manual assessment of lung severity (combined evaluation of density and extent of lung opacities) by multiple radiologists, and PXS scores from hospital admission radiographs were predictive of subsequent mechanical ventilation or death within 3 days (8).

Higher radiographic severity of COVID-19, as indicated by higher PXS scores, may indicate more severe subsequent pulmonary complications of COVID-19 in rheumatic disease patients, explaining some of the observed associations between these conditions and higher risk of mechanical ventilation and/or death than the general population. Additionally, since rheumatic disease patients are more likely than the general population to have baseline chronic lung disease, it is unclear whether PXS scores may reliably predict risk of severe COVID-19 outcomes in this population. If so, deep learning–derived radiograph scores may be leveraged for use in larger studies evaluating the relationships between rheumatic disease, immunosuppression, imaging results, and COVID-19 outcomes. In this study, we examined whether patients with rheumatic disease and COVID-19 have worse PXS scores than general population comparators with COVID-19 and aimed to validate the use of the PXS score in rheumatic disease patients.

PATIENTS AND METHODS

Study population and rheumatic disease identification. From a previously described cohort of patients at Mass General Brigham (MGB) with confirmed rheumatic disease and COVID-19 (1,2), we identified patients ≥18 years of age seen at Massachusetts General Hospital or Brigham and Women’s Hospital (tertiary care hospitals in Boston, Massachusetts) between January 31, 2020 and January 31, 2021 who had chest radiographs performed within ±2 weeks of their initial positive COVID-19 test result. This study was approved by the MGB Institutional Review Board (2020P000833). Patients were not involved in the design, conduct, or reporting of this study.

Comparator identification. Each rheumatic disease patient was matched to up to 10 comparators without rheumatic disease; comparators also had confirmed COVID-19 (by polymerase chain reaction, or less commonly antigen testing) and a chest radiograph performed within ±2 weeks of COVID-19 diagnosis. Rheumatic disease patients were matched based on age (±5 years), sex, hospital site at which the chest radiograph was performed (to account for technical differences), and the date of positive COVID-19 test result (±14 days, to control for calendar time since testing criteria, hospital capacity, and treatment strategies changed over time). The date of the earliest positive COVID-19 test was used as the index date.

PXS score. All patients included in the study had at least 1 chest radiograph within ±2 weeks of COVID-19 diagnosis. Chest radiographs were assessed using a previously validated convolutional Siamese neural network algorithm to assess COVID-19 lung disease severity, which had been further tuned and tested for both anteroposterior and posteroanterior frontal chest views (8,9). In this model, the inputs are pixel-level image data from frontal chest radiographs, and the output is a quantitative score of COVID-19 lung severity: the PXS score (https://github.com/QTIM-Lab/PXS-score). The PXS score quantifies the extent as well as the density of any pulmonary opacities and is not specific for COVID-19–related pulmonary findings. PXS scores range from approximately 0 to 24, with higher PXS scores indicative of more severe radiographic lung disease (Figure 1).

We defined a PXS score of ≤9 as indicating nonsevere disease and a score of >9 as indicating severe disease based on previously determined thresholds (10). Prior data have shown 5-fold higher odds of subsequent intubation or death within 3 days of hospital admission in the general population with a PXS score of >9 compared to all patients with COVID-19 (10). For those patients with multiple radiographs, we used the maximum PXS score within ±14 days of COVID-19 diagnosis for each patient for the primary analysis; we also performed
secondary analyses using initial PXS scores. For those who were mechanically ventilated, we used the maximum PXS score prior to the time of mechanical ventilation in the analysis assessing the risk of mechanical ventilation with high versus low maximum PXS score.

Covariates and clinical outcome assessments. For the patients with rheumatic disease, clinical variables of interest were extracted from the electronic health record (EHR) by manual record review. These included the rheumatic disease diagnosis and duration, immunomodulatory medications at the time of COVID-19 diagnosis (including specific dose of glucocorticoids when applicable), and disease activity level (based on global assessment by the treating provider as documented in the EHR).

For rheumatic disease patients and comparators, additional variables were extracted from the COVID-19 Data Mart, an EHR-based data enclave established by MGB that includes all patients diagnosed with COVID-19. Variables extracted from the COVID-19 Data Mart included location (outpatient, emergency department, or inpatient) of the maximally scored chest radiograph, site of COVID-19 test, demographic characteristics (age, sex, and self-identified race/ethnicity), smoking status, and medical comorbidities (11). Baseline characteristics including comorbidities and body mass index were assessed in the 1 year prior to the index date, and the Charlson comorbidity index was calculated from diagnosis codes in the EHR using any data available prior to the index date (12). COVID-19 clinical outcomes, specifically mechanical ventilation and death, were also extracted from the COVID-19 Data Mart.

Statistical analysis. Categorical variables were presented as number (%), and continuous variables were presented as mean ± SD or median ± interquartile range (IQR), as appropriate. Continuous variables were compared using a 2-sample t-test for continuous normally distributed variables or Wilcoxon's test for

Figure 1. Examples of chest radiographs from patients with COVID-19 and their associated pulmonary x-ray severity (PXS) scores.
continuous nonnormally distributed variables. Categorical variables were compared using chi-square tests.

For each group (rheumatic disease patients and comparators), we determined the PXS score at the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th percentiles. We then compared the PXS scores between matched rheumatic disease patients and comparators using unadjusted and adjusted quantile regression at each percentile. We initially used unadjusted analyses because rheumatic disease patients and comparators were matched on a number of relevant potential confounders (e.g., age, sex, time of COVID-19 infection), and other covariates (e.g., comorbidity scores) are likely mediators of the association between the exposure (rheumatic disease versus no rheumatic disease) and the outcome (PXS score). We also performed adjusted quantile regression, adjusting for age, Charlson comorbidity index (dichotomized by <2 or ≥2), and interstitial lung disease (ILD) in the 12 months prior to COVID-19 diagnosis. We used Poisson regression to compare the number of chest radiographs per subject. For both the time from COVID-19 diagnosis to initial or maximum PXS score and median initial PXS score, we used a Wilcoxon’s test for unadjusted analyses and quantile regression for adjusted analyses. We used a chi-square test to compare the site where the maximally scored chest radiograph was performed in unadjusted analyses and multinomial regression in adjusted analyses. In unadjusted and multivariable adjusted quantile regression, we added an “id” variable to indicate matched sets in the MODEL statement to account for the correlation among subjects within the matched set.

In the analysis comparing the outcomes of mechanical ventilation and death between those with a maximum PXS score of >9 (indicating severe lung disease) to those with a maximum PXS score of ≤9 within ±14 days of COVID-19 diagnosis, the index

---

Table 1. Clinical characteristics of COVID-19 patients with rheumatic disease and comparators without rheumatic disease who had chest radiographs performed*

| Characteristic                        | Patients with rheumatic disease (n = 70) | Comparators without rheumatic disease (n = 463) | P   |
|--------------------------------------|------------------------------------------|-------------------------------------------------|-----|
| Age, mean ± SD years†                | 62 ± 16                                  | 62 ± 15                                         | 0.77|
| Female†                              | 48 (69)                                  | 315 (68)                                        | 0.93|
| Race                                 |                                          |                                                 |     |
| White                                | 33 (47)                                  | 227 (49)                                        |     |
| Black or African American            | 13 (19)                                  | 56 (12)                                         |     |
| Asian                                | 5 (7)                                    | 18 (4)                                          |     |
| Other                                | 19 (27)                                  | 162 (35)                                        |     |
| Hispanic or Latinx ethnicity         | 4 (6)                                    | 40 (9)                                          | 0.41|
| Body mass index, mean ± SD kg/m²     | 29.9 ± 6.7                               | 29.7 ± 6.4                                      | 0.72|
| Smoking status                       |                                          |                                                 |     |
| Never                                | 41 (59)                                  | 289 (62)                                        |     |
| Former                               | 27 (39)                                  | 129 (28)                                        |     |
| Current                              | 2 (3)                                    | 36 (8)                                          |     |
| Unknown                              | 0 (0)                                    | 9 (2)                                           |     |
| Charlson comorbidity index, median (IQR) | 4 (2–7)                               | 1 (0–3)                                         | <0.0001|
| Comorbidities                        |                                          |                                                 |     |
| Hypertension                         | 38 (54)                                  | 146 (32)                                        | <0.01|
| Diabetes mellitus                    | 22 (31)                                  | 71 (15)                                         | <0.01|
| Coronary artery disease              | 15 (21)                                  | 47 (10)                                         | <0.01|
| Heart failure                        | 8 (11)                                   | 36 (8)                                          | 0.30|
| Asthma                               | 13 (19)                                  | 32 (7)                                          | <0.01|
| Chronic obstructive pulmonary disease| 6 (9)                                    | 20 (4)                                          | 0.12|
| Obstructive sleep apnea              | 7 (10)                                   | 31 (7)                                          | 0.32|
| Chronic kidney disease               | 21 (30)                                  | 46 (10)                                         | <0.0001|
| Interstitial lung disease            | 4 (6)                                    | 5 (1)                                           | 0.01|
| Site where COVID-19 test was performed |                                      |                                                 | 0.08|
| Outpatient                           | 43 (61)                                  | 232 (50)                                        |     |
| Emergency department or inpatient    | 27 (39)                                  | 231 (50)                                        |     |
| Site where initial chest radiograph was performed |     |                                                 | 0.47|
| Outpatient                           | 21 (30)                                  | 126 (27)                                        |     |
| Emergency department                 | 40 (57)                                  | 295 (64)                                        |     |
| Inpatient                            | 9 (13)                                   | 42 (9)                                          |     |
| Time from COVID-19 diagnosis to initial chest radiograph, median (IQR) days | 0 (0–2)                                | 0 (0–0)                                         | 0.01|

* Values are the number (%) unless indicated otherwise. IQR = interquartile range.
† Matching factors.
date was considered to be the time of either the initial or maximum PXS score. Person-months of follow-up were determined for each subject from the index date to the first of any of the following events: occurrence of the outcome of interest, death, or date of the last encounter in MGB at the time of data analysis. We calculated incidence rates as events/1,000 person-months. Cox proportional hazards regression models were used to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs) for the outcomes of mechanical ventilation and death, comparing patients with rheumatic diseases with a maximum PXS score of \(>9\) to those with a maximum PXS score of \(\leq 9\). We also performed a secondary analysis among rheumatic disease patients evaluating those with an initial PXS score of \(>9\) to those with an initial PXS score of \(\leq 9\) as well as a secondary analysis among general population comparators evaluating those with either an initial or maximum PXS score of \(>9\) to those with an initial or maximum PXS score of \(\leq 9\), respectively. The level of significance was set as a 2-tailed \(P\) value of <0.05, and statistical analyses were completed using SAS statistical software, version 9.4.

### RESULTS

#### Cohort characteristics. Of 624 patients with rheumatic disease and confirmed COVID-19 who were seen at either Massachusetts General Hospital or Brigham and Women’s Hospital, 70 patients had at least 1 chest radiograph available for review. Thus, 70 patients with rheumatic disease and 463 matched comparators, all with confirmed COVID-19, were included in our analyses. Patients with rheumatic disease and their comparators had similar demographic characteristics; the mean age was 62 years in each group, with 69% female patients in the rheumatic disease group and 68% female comparators in the comparators group (Table 1). The distribution of race/ethnicity and smoking status was similar between groups. The Charlson comorbidity index was higher in the rheumatic disease patients than comparators (median of 4 versus 1, \(P < 0.001\)). Four patients (6%) with rheumatic diseases had ILD compared with 5 (1%) in the comparator group (\(P = 0.005\)). Among the patients with rheumatic disease, the most common diagnosis was rheumatoid arthritis (n = 26 [37%]), followed by other inflammatory arthritis (n = 12 [17%]) and systemic lupus erythematosus (n = 10 [14%]) (Table 2). Many patients (n = 27 [39%]) were taking oral glucocorticoids (median prednisone-equivalent daily dose of 5 mg), and a majority were taking other immunomodulatory medications (Table 2). Most COVID-19 tests were performed in the outpatient setting (n = 43 [61%] versus n = 232 [50%] in rheumatic disease patients and comparators, respectively) rather than the emergency department or inpatient setting (n = 27 [39%] versus n = 231 [50%] in rheumatic disease patients and comparators, respectively) (\(P = 0.08\) for difference between groups across categories).

#### Table 2. Detailed characteristics of patients with rheumatic disease and COVID-19 who had chest radiographs performed within 2 weeks of COVID-19 diagnosis*

| Characteristic                              | Patients with rheumatic disease (n = 70) |
|--------------------------------------------|----------------------------------------|
| Rheumatic disease diagnosis                |                                        |
| Rheumatoid arthritis                       | 26 (37)                                |
| Other inflammatory arthritis†              | 12 (17)                                |
| Systemic lupus erythematosus               | 10 (14)                                |
| Vasculitis‡                                | 7 (10)                                 |
| Inflammatory myopathy                      | 5 (7)                                  |
| Polymyalgia rheumatica                     | 4 (6)                                  |
| Other rheumatic conditions§                 | 8 (11)                                 |
| Autoimmune disease duration, median (IQR) years | 9 (3–16)                              |
| Autoimmune disease status                  |                                        |
| Remission                                  | 21 (30)                                |
| Low activity                               | 28 (40)                                |
| Moderate/high activity                     | 15 (21)                                |
| Unknown                                   | 6 (9)                                  |
| Immunomodulatory medication at the time of COVID-19 diagnosis |                                        |
| Biologic DMARDs                            |                                        |
| CD20 inhibitor                             | 10 (14)                                |
| TNF inhibitor                              | 3 (4)                                  |
| IL-6 receptor inhibitor                    | 4 (6)                                  |
| IL-12/IL-23 inhibitor                      | 1 (1)                                  |
| CTLA-4 immunoglobulin                      | 2 (3)                                  |
| Targeted synthetic DMARDs                 |                                        |
| JAK inhibitor                              | 2 (3)                                  |
| Conventional synthetic DMARDs             |                                        |
| Hydroxychloroquine                         | 18 (26)                                |
| Methotrexate                               | 11 (16)                                |
| Mycophenolate mofetil                      | 4 (6)                                  |
| Azathioprine                               | 3 (4)                                  |
| Other¶†                                    | 4 (6)                                  |
| Oral glucocorticoid                        | 27 (39)                                |
| Prednisone-equivalent daily dose, median (IQR) mg | 5 (5–10)                              |

* Values are the number (%) unless indicated otherwise. DMARDs = disease-modifying antirheumatic drugs; IL = interleukin; IQR = interquartile range; TNF = tumor necrosis factor.
† Includes other inflammatory arthritis (n = 8), psoriatic arthritis (n = 2), nonsystemic juvenile idiopathic arthritis (n = 1), and ankylosing spondylitis (n = 1).
‡ Includes other vasculitides (n = 3), giant cell arteritis (n = 2), granulomatosis with polyangiitis (n = 1), and Behçet’s disease (n = 1).
§ Includes antiphospholipid syndrome (n = 2), mixed connective tissue disease (n = 2), IgG4-related disease (n = 1), Sjögren’s syndrome (n = 1), systemic sclerosis (n = 1), and undifferentiated connective tissue disease (n = 1).
¶ Includes leflunomide (n = 1), cyclophosphamide (n = 1), sulfasalazine (n = 1), and tacrolimus (n = 1).

#### Comparisons of PXS scores between rheumatic disease patients and comparators. In 70 rheumatic disease patients, there were a total of 206 chest radiographs analyzed within 2 weeks of COVID-19 diagnosis date compared with 463 comparators who had 936 chest radiographs in the same timeframe (Table 3). The majority of initial radiographs in each group were performed in the emergency department (n = 40 [57%] versus n = 295 [64%] in rheumatic disease patients and
comparators, respectively) compared to 21 (30%) versus 126 (27%) in the outpatient setting and 9 (13%) versus 42 (9%) in the inpatient setting ($P = 0.16$ for difference between groups across all categories) (see Supplementary Table 1, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24883). The median (IQR) number of chest radiographs per person was 1 (1–3) in the rheumatic disease patients and 1 (1–2) in the comparators. The median (IQR) time from COVID-19 diagnosis to initial chest radiograph was 0 days (0–2) in rheumatic patients and 0 days (0–0) in comparators, and the time from COVID-19 diagnosis to the chest radiograph with the maximum PXS score was 1 day (0–6) in rheumatic patients and 0 days (0–4) in comparators (see Supplementary Figure 1, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24883). PXS scores ranged from 0.8 to 16.5 in rheumatic disease patients and 0.7 to 15.9 in comparators. Using unadjusted quantile regression, the maximum PXS scores for each patient were not significantly different between rheumatic disease patients and comparators at the 10th to 60th percentiles in each respective group (e.g., 50th percentile scores of 3.1 versus 3.2, respectively; difference of −0.1 [95% CI −1.8, 1.7], $P = 0.95$) (Table 3). However, PXS scores were significantly higher in rheumatic disease patients at the 70th percentile (6.8 versus 5.6; difference of 1.2 [95% CI 0.4, 2.0], $P = 0.01$), 80th percentile (8.9 versus 7.3; difference of 1.5 [95% CI 1.0, 2.2], $P < 0.001$), and 90th percentiles (10.2 versus 9.2; difference of 1.0 [95% CI 0.12, 1.96], $P = 0.01$). After adjusting for age, Charlson comorbidity index (dichotomized by <2 or ≥2), and ILD, maximum PXS scores were still significantly higher at the 80th and 90th percentiles in rheumatic disease patients (e.g., 90th percentile scores of 10.2 versus 9.2; $P = 0.03$) (Table 3). A higher proportion of patients with rheumatic disease had maximum PXS scores indicating severe lung disease (>9) than comparators (14 [20%] versus 49 [11%]; $P = 0.03$). These findings persisted when we excluded patients with ILD from the rheumatic disease patient and comparator groups. Figure 2 illustrates the frequency of different PXS scores in both the rheumatic disease patients and comparators. PXS scores from the initial chest radiograph were similar across the 10th to 90th percentiles between rheumatic disease patients and comparators (e.g., 50th percentile score of 2.2 versus 2.5, respectively; unadjusted $P$ value 0.56, and adjusted $P$ value of 0.81) (see Supplementary Table 2 and Supplementary Figure 2, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24883).

**Table 3.** Pulmonary x-ray severity (PXS) scores from chest radiographs in rheumatic disease patients and matched comparators using PXS scores from within 2 weeks before or after COVID-19 diagnosis*

| Characteristic                        | Patients with rheumatic disease (n = 70) | Comparators without rheumatic disease (n = 463) | Unadjusted $P$ | Adjusted $P$‡ |
|---------------------------------------|----------------------------------------|-----------------------------------------------|----------------|---------------|
| Total chest radiographs, no.          | 206                                    | 936                                           | <0.001         | <0.001        |
| No. of chest radiographs per subject, median (IQR) | 1 (1–3) | 1 (1–2) | <0.001 | <0.001 |
| PXS score ≤9†                        | 56 (80)                                | 414 (89)                                      | 0.03           | 0.01          |
| PXS score >9†                        | 14 (20)                                | 49 (11)                                       |                |               |
| Maximum PXS score†                   |                                        |                                               |                |               |
| 10th percentile                       | 0.9                                    | 0.9                                           | 0.60           | 0.31          |
| 20th percentile                       | 1.0                                    | 1.1                                           | 0.38           | 0.76          |
| 30th percentile                       | 1.5                                    | 1.5                                           | 0.90           | 0.61          |
| 40th percentile                       | 2.1                                    | 2.2                                           | 0.75           | 0.56          |
| 50th percentile                       | 3.1                                    | 3.2                                           | 0.95           | 0.93          |
| 60th percentile                       | 4.4                                    | 4.1                                           | 0.76           | 0.41          |
| 70th percentile                       | 6.8                                    | 5.6                                           | 0.01           | 0.18          |
| 80th percentile                       | 8.9                                    | 7.3                                           | <0.001         | <0.001        |
| 90th percentile                       | 10.2                                   | 9.2                                           | 0.01           | 0.03          |

* Values are the number (%) unless indicated otherwise.
† Adjusted for age, Charlson comorbidity index (dichotomized as <2 or ≥2), and interstitial lung disease.
‡ Based on the highest PXS score per patient within ±2 weeks of COVID-19 diagnosis.

**PXS scores and severe COVID-19 outcomes.** Among patients with rheumatic disease (n = 70), those with a maximum PXS score >9 (n = 14) compared to those with PXS scores ≤9 (n = 56) had higher rates of mechanical ventilation (578.9 versus 2.7 events/1,000 months; HR 24.1 [95% CI 6.7, 86.9], $P < 0.001$) (Table 4). Those with PXS scores >9 had numerically higher rates of death compared to those with PXS scores ≤9, although this was not statistically significant (7.5 versus 0.9 events/1,000 months; HR 8.2 [95% CI 0.7, 90.4], $P = 0.09$). These differences in outcomes between those with higher and lower PXS scores also persisted when we excluded patients with ILD from the rheumatic disease patient and comparator groups.

Similar results were seen when using initial instead of maximum PXS score. Among patients with rheumatic diseases, those with an initial PXS score of >9 compared to those
with PXS scores of ≤9 had significantly higher rates of mechanical ventilation (36,000 versus 9.7 events/1,000 months; HR 26.4 [95% CI 7.5, 92.8], \( P < 0.001 \)) and death (18.4 versus 1.5 events/1,000 months; HR 11.7 [95% CI 0.95, 142.8], \( P = 0.045 \)) (see Supplementary Table 3, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24883).

These associations with both mechanical ventilation and death were similarly seen in general population comparators. General population comparators with a maximum PXS score of >9 had higher rates of mechanical ventilation (134.7 versus 5.9 events/1,000 months; HR 11.4 [95% CI 6.9, 18.9], \( P < 0.001 \)) and death (23.5 versus 4.7 events/1,000 months; HR 3.8 [95% CI 1.9, 7.6], \( P < 0.001 \)) (see Supplementary Table 4, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24883). General population comparators with an initial PXS score of >9 also had higher rates of mechanical ventilation (116.6 versus 7.8 events/1,000 months; HR 6.6 [95% CI 3.9, 11.3], \( P < 0.001 \)) and death (31.1 versus 5.1 events/1,000 months; HR 4.8

Figure 2. Frequency of maximum pulmonary x-ray severity (PXS) scores in patients with rheumatic disease and comparators. PXS score is based on the highest score within 14 days of COVID-19 diagnosis. A, Overlap in the percentages of patients with rheumatic disease and comparators (purple bars), excess percentage of patients with rheumatic disease over comparators (red bars), and excess percentage of comparators over patients with rheumatic disease (blue bars). B, Cumulative probability curves for rheumatic disease patients (red line) and comparators (blue line).
Table 4. Risk of mechanical ventilation or death in patients with rheumatic diseases and COVID-19 with high versus low maximum PXS scores*

| Outcomes          | PXS score >9† | PXS score ≤9‡ | P  |
|-------------------|---------------|---------------|----|
| Mechanical ventilation |               |               |    |
| Patients in PXS category | 14           | 56            |    |
| Mechanical ventilation‡ | 11           | 3             |    |
| Person-time, months | 19           | 1,102         |    |
| Rate (events/1,000 months) (95% CI) | 578.9 (236.8, 921.1) | 2.7 (0.0, 5.8) | <0.001 |
| Hazard ratio (95% CI) | 24.13 (6.7, 86.9) | Ref.          |    |
| Death             |               |               |    |
| Patients in PXS category | 13           | 57            |    |
| Death             | 2             | 1             |    |
| Person-time, months | 266.6        | 1,144.33      |    |
| Rate (events/1,000 months) (95% CI) | 7.5 (0.0, 17.9) | 0.9 (0.0, 2.6) |          |
| Hazard ratio (95% CI) | 8.2 (0.7, 90.4) | Ref.          | 0.087 |

* Values are the number unless indicated otherwise. 95% CI = 95% confidence interval; PXS = pulmonary x-ray severity; Ref. = reference.
† Based on the highest PXS score within ±2 weeks of COVID-19 diagnosis.
‡ For mechanical ventilation outcome, the highest PXS score on radiographs prior to intubation was used.

[95% CI 2.1, 11.2], P < 0.001) (see Supplementary Table 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24883).

DISCUSSION

Using a previously validated deep learning–derived scoring algorithm, we found that patients with rheumatic disease and COVID-19 had more severe radiographic lung disease than matched comparators with COVID-19. Our results provide validation of the use of both the initial and maximum PXS score among rheumatic disease patients and corroborate previous studies suggesting increased risk for mechanical ventilation. As rheumatic disease patients may be more susceptible to severe lung involvement from COVID-19 than the general population (1,2), additional studies are needed to determine how PXS scores may be helpful in clinical research studies studying risk factors for poor COVID-19 outcomes in rheumatic disease and other patients. Our results emphasize the importance of COVID-19 risk mitigation strategies, such as vaccination with an additional dose and postexposure prophylaxis with monoclonal antibodies, in patients with rheumatic diseases due to their increased susceptibility to severe pulmonary involvement from COVID-19.

PXS scores have been previously validated as a prognostic tool in general population patients with COVID-19. In both training and validation data sets, PXS scores correlated well with radiologist-determined radiographic lung disease severity (Pearson’s correlation coefficient of 0.86), and the initial PXS score was predictive of subsequent intubation or death (receiver operating characteristic area under the curve = 0.80) (8). The PXS algorithm has been further tuned using inpatient and outpatient COVID-19 data sets in the US and Brazil (9). After an intervention in which radiologists were exposed to machine learning–derived PXS scores of different chest radiographs, the interrater agreement among subsequent diagnostic radiologists improved (Fleiss k coefficient 0.40–0.66), indicating that PXS scores could be integrated into the radiology workflow to standardize interpretation of COVID-19 chest radiographs without requiring additional training of radiologists (10). PXS scores have also been used to examine the association of radiographic lung severity of COVID-19 with other factors such as acute neurologic disease and right ventricular strain. Lang et al examined PXS scores in patients with acute neuroimaging findings (e.g., intracranial hemorrhage, infarction, or leukoencephalopathy) and found that they had significantly higher mean PXS scores than comparators (9.2 versus 5.0; P < 0.001) (13). Gibson et al found that right ventricular strain is common in intubated patients with COVID-19 but is not associated with radiographic lung disease (14).

We found that initial PXS scores were similar between rheumatic disease patients and general population comparators but that maximum PXS scores were higher in rheumatic disease patients; this underscores the similar nature of both groups at baseline as well as the natural history of acute COVID-19, which often does not reach peak severity until a week or more into the clinical course (15). Several factors may explain why rheumatic disease patients are at higher risk for worsened lung disease over the course of acute COVID-19. First, patients with rheumatic diseases are commonly treated with immunosuppressive medications that might increase the severity of infection. Second, rheumatic diseases are often associated with chronic systemic inflammation, which might predispose these patients to more severe hyperinflammatory responses to COVID-19. Indeed, our prior work demonstrated that, compared with general population patients, those with rheumatic diseases and COVID-19 had a more severe inflammatory state (16). Third, some rheumatic diseases are associated with inflammatory lung disease (e.g., ILD), and the trigger underlying this manifestation is unknown. Infectious etiologies, like COVID-19, may be responsible for precipitating inflammatory lung disease in patients predisposed for rheumatic disease. These hypotheses deserve additional study. Notably, the differences in maximum PXS scores between groups...
have more severe COVID-19 and corroborates previous findings suggesting increased risk for mechanical ventilation and mortality. The higher risk of severe pulmonary involvement from COVID-19 provides additional support for COVID-19 vaccination and risk mitigation among rheumatic disease patients.

**AUTHOR CONTRIBUTIONS**

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Wallace had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study conception and design.** Patel, D’Silva, Li, Fu, Zhang, Kalpathy-Cramer, Sparks, Wallace.

**Acquisition of data.** Patel, D’Silva, Li, Hsu, Dilorio, Cook, Prisco, Martin, Vanni, Zaccardelli, Kalpathy-Cramer, Sparks, Wallace.

**Analysis and interpretation of data.** Patel, D’Silva, Li, Fu, Zhang, Kalpathy-Cramer, Sparks, Wallace.

**REFERENCES**

1. D’Silva KM, Serling-Boyd N, Wallwork R, et al. Clinical characteristics and outcomes of patients with coronavirus disease 2019 (COVID-19) and rheumatic disease: a comparative cohort study from a US ‘hot spot’. Ann Rheum Dis 2020;79:1156–62.

2. Serling-Boyd N, D’Silva KM, Hsu TY, et al. Coronavirus disease 2019 outcomes among patients with rheumatic diseases 6 months into the pandemic. Ann Rheum Dis 2021;80:660–6.

3. D’Silva KM, Jorge A, Cohen A, et al. COVID-19 outcomes in patients with systemic autoimmune rheumatic diseases compared to the general population: a US multicenter, comparative cohort study. Arthritis Rheumatol 2021;73:914–20.

4. Bower H, Frisell T, Di Giuseppe D, et al. Impact of the COVID-19 pandemic on morbidity and mortality in patients with inflammatory joint diseases and in the general population: a nationwide Swedish cohort study. Ann Rheum Dis 2021;80:1086–1093.

5. Kroon FP, Najm A, Alunno A, et al. Risk and prognosis of SARS-CoV-2 infection and vaccination against SARS-CoV-2 in rheumatic and musculoskeletal diseases: a systematic literature review to inform EULAR recommendations. Ann Rheum Dis 2022;81:422–32.

6. Conway R, Grimshaw AA, Konig MF, et al. SARS-CoV-2 infection and COVID-19 outcomes in rheumatic disease: a systematic literature review and meta-analysis [review]. Arthritis Rheumatol 2022;74:76–75.

7. Tousie D, Voutsinas N, Finkelestein M, et al. Clinical and chest radiography features determine patient outcomes in young and middle-aged adults with COVID-19. Radiology 2020;297:E197–206.

8. Li MD, Arun NT, Gidwani M, et al. Automated assessment and tracking of COVID-19 pulmonary disease severity on chest radiographs using convolutional Siamese neural networks. Radiol Artif Intell 2020;2:e200079.

9. Li MD, Arun NT, Aggarwal M, et al. Improvement and multi-population generalizability of a deep learning-based chest radiograph severity score for COVID-19. medRxiv 2020 doi: 10.1101/2020.09.15.20196453. E-pub ahead of print.

10. Li MD, Little BP, Akkasab TK, et al. Multi-radiologist user study for artificial intelligence-guided grading of COVID-19 lung disease severity on chest radiographs. Acad Radiol 2021;28:572–6.

11. Mass General Brigham. Research Information Science & Computing. New COVID-19 tools for researchers. URL: https://rc.partners.org/about/projects-initiatives/new-covid-19-research-tools-researchers.
12. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chron Dis 1987;40:373–83.

13. Lang M, Li MD, Jiang KZ, et al. Severity of chest imaging is correlated with risk of acute neuroimaging findings among patients with COVID-19. AJNR Am J Neuroradiol 2021;42:831–7.

14. Gibson LE, Fenza RD, Lang M, et al. Right ventricular strain is common in intubated COVID-19 patients and does not reflect severity of respiratory illness. J Intensive Care Med 2021;36:900–9.

15. Zirpe KG, Tiwari AM, Gurav SK, et al. Timing of invasive mechanical ventilation and mortality among patients with severe COVID-19-associated acute respiratory distress syndrome. Indian J Crit Care Med 2021;25:493–8.

16. Hsu TY, D’Silva KM, Patel NJ, et al. Laboratory trends, hyperinflammation, and clinical outcomes for patients with a systemic rheumatic disease admitted to hospital for COVID-19: a retrospective, comparative cohort study. Lancet Rheumatol 2021;3: e638–47.