medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Full Title: Radiomic Profiling of Lung CT in a Cohort of Sarcoidosis Cases

Authors: Nichole E Carlson, PhD[1], William Lippitt, MS, PhD[1], Sarah M Ryan, PhD[1], Margaret Mroz, MS[2], Zachary Buchalski, MS[1], Briana Barkes, MPH[2], Shu-Yi Liao, MD, MPH,ScD[2,3,4], Lisa A Maier, MD[2,3,4*], Tasha E Fingerlin, PhD[1,5*].

*These authors contributed equally to this work.

Institutions: [1]Colorado School of Public Health, Department of Biostatistics and Informatics; [2]National Jewish Health, Department of Medicine; [3]University of Colorado, Department of Medicine; [4]Colorado School of Public Health, Department of Environmental and Occupational Health; [5]National Jewish Health, Department of Department of Immunology and Genomic Medicine;

Corresponding Author:

Nichole E Carlson, PhD
Professor
Department of Biostatistics and Informatics
University of Colorado-Anschutz Medical Campus
13001 E 17th PL, MS B-119
Aurora, CO 80045
Tel: 303-724-4354
Fax:303-724-4491
e-mail:nichole.carlson@cuanschutz.edu

Author contributions: N.E.C. Substantial contributions to the conception or design of the work, lead biostatistical analysis and interpretation of the data, and wrote manuscript. W.L. developed the filtering method for this analysis, performed the sensitivity and validation analyses and edited of the manuscript. S.M.R. performed

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

image processing, conducted statistical analysis, prepared figures and table, interpreted the data, and contributed to the manuscript. M.M. coordinated the study, cleaned data, and contributed to the manuscript. Z.B assisted with the sensitivity and validation analyses and conducted further data analysis and created the tables and figures. B.B. data acquisition, data cleaning, project management, and contributed to the manuscript. S.L. contributed to the scientific rationale and interpretation of the data and contributed to the manuscript. L.A.M. Substantial contributions to the conception or design of the work, interpretation of the data and wrote the manuscript. T.E.F. Substantial contributions to the conception or design of the work, interpretation of the data and wrote the manuscript.

Sources of Support: This work was supported by the National Institutes of Health (R01 HL114587; R01 HL142049; U01 HL112695). Data from the GRADS study was used, which was funded by the NIH grant U01 HL112707 entitled “Sarcoidosis and A1AT Genomics and Informatics Centre”, as well as others (U01 HL112707, U01 HL112694, U01 HL112695, U01 HL112696, U01 HL112702, U01 HL112708, U01 HL112711, U01 HL112712). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health.

Supplementary Material Statement: This article has an online supplement, which is accessible from this issue's table of contents online.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Question Addressed: High resolution computed tomography (HRCT) of the chest is increasingly used in clinical practice for sarcoidosis. Visual assessment of chest HRCTs in patients with sarcoidosis has high inter- and intra-rater variation. Radiomics offers a reproducible quantitative assessment of HRCT lung parenchyma and could be useful as an additional summary measure of disease. We develop radiomic profiles on HRCT and map them to clinical and patient reported outcomes.

Patients and Methods: Three-dimensional radiomic features were calculated on chest HRCT for the left and right lung from sarcoidosis cases enrolled in the Genomic Research in Alpha-1 Antitrypsin Deficiency and Sarcoidosis study (N=321). Sparse K-means was used to group sarcoidosis cases using their radiomic profiles. Linear regression investigated how pulmonary function tests and patient reported outcomes differed between groups. Resampling approaches were used to validate the robustness of the findings.

Results: Five groups were identified. The new radiomic-based grouping was associated with Scadding stage ($p<0.001$), but each radiomic group had patients from all Scadding stages. All pulmonary function testing measures significantly differed between radiomic groups ($p<0.001$). Radiomic group remained significantly associated with pulmonary function even after adjusting for Scadding stage ($p<0.0001$). Individual radiomic measures explained 9-18% of the variation in pulmonary function testing. Only two patient reported outcomes (shortness of breath and physical health) differed by radiomic group ($p<0.013$).

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Answer to Question: Radiomic quantification of sarcoidosis identifies subgroups associated with pulmonary function and patient reported outcomes. These associations provide additional evidence that radiomics may be useful for quantifying new disease phenotypes.

Key words: quantitative imaging, texture analysis, pulmonary function, phenotyping, patient outcomes

Abstract word count: 250

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Sarcoidosis is a granulomatous interstitial lung disease which affects ~ 110 thousand individuals in the United States (1). Typical diagnoses is between 30-50 years of life, resulting in a decrease in quality of life and productivity (2). Pulmonary disease occurs in over 90% of those with sarcoidosis (3) with significant morbidity and mortality. Currently, visual assessment of chest radiography (CXR) is used to quantify lung abnormalities standardized via the Scadding staging system (4) which groups CXR findings into five groups from 0 to 4. Substantial variation exists in the radiographic patterns within each Scadding stage. Scadding stage has limited utility in predicting prognosis even in the extremes of the scale (5).

Chest high resolution computed tomography (HRCT) is increasingly used in clinical practice of sarcoidosis to monitor disease as it offers more detailed visualization of parenchymal abnormalities compared to CXR (4, 6–8). As with CXR, visual assessment of chest HRCT is used to evaluate abnormalities although there are limited standardized scoring metrics (4, 9, 10). This is due in part to the diverse and heterogeneous patterns present on chest HRCT in sarcoidosis, often with multiple patterns noted on one CT. These complexities result in high inter- and intra-rater variation (11). More automated systems that quantify sarcoidosis chest HRCT could decrease these sources of variation.

Radiomics is a field of study in which large numbers of quantitative features are extracted from medical images (12). A radiomics panel computes summary measures of

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the distribution of the Hounsfield units (HU) along with summary measures of the spatial relationships of neighbouring voxels (13). The result is a panel of quantitative measures characterizing the texture of the image.

Radiomic analysis has proved useful for quantifying HRCT in emphysema (7), idiopathic pulmonary fibrosis (14, 15), interstitial lung disease (16, 17), diffuse lung disease (18) and cancer (19). Ryan et al. (20) showed the potential utility of radiomics in sarcoidosis, comparing radiomic and other textural based measures between sarcoidosis patients and controls. It remains unclear whether radiomics also has the potential to differentiate varied phenotypes *within* sarcoidosis patients. Radiomic profiles within sarcoidosis patients that also correlate with pulmonary function or patient reported outcomes could indicate that radiomics may serve as a useful measure to track change in the lung parenchyma over time.

The goal of this research study is to develop a radiomic profile of sarcoidosis chest HRCT using a large, phenotypically-diverse population of sarcoidosis cases from the Genomic Research in Alpha-1 Antitrypsin Deficiency and Sarcoidosis (GRADS) study (21) using statistical clustering techniques and to investigate the clinical utility of the clusters by quantifying their association with pulmonary function testing (PFT) and patient reported outcomes (PROs). PFTs and PROs were chosen to capture both clinician and patient focused measures of disease (26,27).

Methods

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The study design and participants

The sarcoidosis population was recruited in the multicentre NHLBI-funded GRADS study (N=368) (21). This investigation has GRADS ancillary study approval and all participants provided informed consent (IRB approval HS-2779 and HS-2780;N=365). More details of this cohort can be found in the online supplement.

Pulmonary function testing included pre-bronchodilator (BD) forced expiratory volume at one second (FEV1), forced vital capacity (FVC), the ratio of FEV1 to FVC, and diffusing capacity of the lungs for carbon monoxide (DLCO). The PROs included the gastroesophageal reflux disease questionnaire (GERDQ) (24), the University of California San Diego Shortness of Breath Questionnaire (SOBQ) (25), two measures of fatigue (the Fatigue Assessment Scale [FAS] (26) and Patient-Reported Outcomes Measurement Information System fatigue profile [PROMIS] (27)), the Cognitive Failure Questionnaire (CFQ) (28), and the physical and mental subscales of the SF-12 (33,34). The final analysis dataset included N=321 patients who each had both an analysable CT and clinical data (online supplement Figure 1).

Radiomic Analysis

Details of the imaging acquisition and lung segmentation can be found in the online supplement. Radiomic features were calculated on the left and right lungs using RIA_lung from the lungct R package. These features include 44 first-order features and 239 GLCM features for each lung, for a total of 566 features. To calculate the GLCM

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

features, the Hounsfield units (HU) from each HRCT scan were first discretized into 16 bins with equal relative frequencies; then, the features were calculated in each of 26 directions, assuming a voxel distance of one; finally, these features from all directions were summarized using the mean statistic. The exact computation for each radiomic measure can be found elsewhere (31).

The ComBat function (https://github.com/Jfortin1/ComBatHarmonization) was used to harmonize the radiomic measures across scanners (32–34), while preserving the biological variability in the following covariates: age, height, BMI, sex, race and clinical phenotype.

As radiomic features are high-dimensional and repetitive (Figure 1), we first used a decorrelation filter (see online supplement), to identify a subset of features for analysis. This reduced the number of variables from 566 to 97. We then used robust k-means (35) to cluster patients (R package, RSKC). For outlier robustness, the proportion of cases trimmed was set at 0.1; the optimal bound on feature weights was found to be 7.5, using the permutation approach observed in sparcl R package (36). We used a standardized Gap statistic to determine the optimal number of clusters using the cluster R package (37). A standardized Gap statistic is the typical gap statistic divided by its standard deviation. The variables with the top five weights from the RSKC clustering algorithm were then selected for further investigation (discriminative features).

Statistical Analysis

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Linear regression was used to assess associations between cluster group and Scadding stage and each of the outcomes controlling for age, sex, race, height, and BMI. We fitted a model with only cluster group and only Scadding stage and then put both together in the model to determine if the association with radiomics remained significant in the presence of Scadding stage. Each outcome was modelled separately using a complete case analysis for that outcome. Linear regression was used quantify the associations between the selected discriminative radiomic features and the outcomes controlling for age, sex, race, height, and BMI.

Validation

The statistical analysis is an unsupervised learning problem, which makes traditional training and test validation approaches difficult because the true groupings are unknown. Instead, we investigated the effect of repeatedly applying our analysis pipeline under various conditions and with bootstrapped samples. The details of the validation can be found in the online supplement.

Results

Tables 1 and 2 show the characteristics of the study population, which included N=321 sarcoidosis cases, including 147 (45.8%) males and 233 (72.6%) whites, with an average age of 53 (SD=10) years, average height of 67.0 in (SD=4.2) and average BMI of 30.6 kg/m^2 (SD=6.5). Participants were spread across Scadding stages, with 43 (13.5%) in stage 0, 63 (19.8%) in stage 1, 92 (28.9%) in stage 2, 44 (13.8%) in stage 3 and 76 (23.9%) in stage 4. The average FVC was 2.62 L (SD=0.93), FEV1 was 3.57 L

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(SD=1.14), and DLCO 80.15 (SD=23.37). The population was predominantly of the non-obstructive type (72.7%).

Radiomic Based Clustering of Sarcoidosis

Five groups (clusters) of patients were identified (Table 1). Ordered from highest FVC to lowest, 75 (23.4%) patients were in radiomic group 0, 57 (17.8%) in group 1, 51 (15.9%) in group 2, 58 (18.1%) in group 3 and 80 (24.9%) in group 4. The new radiomic-based grouping was associated with Scadding stage ($p<0.001$), but not a direct reflection of Scadding stage (Table 1).

PFT differed between radiomic groups (Figure2; $p<0.0001$). For FVC, radiomic groups 3 and 4 had between 0.5 and 0.7 L lower average FVC compared to radiomic groups 0, 1, and 2. For FEV1, radiomic groups 2 and 3 and had an average FEV1 0.5 L lower than radiomic groups 0 and 1, which were similar. Radiomic group 4 had the lowest average FEV1 and was approximately 0.8 L lower than groups 0 and 1. DLCO has a similar pattern to FVC. Groups 2 and 4 demonstrated more obstructive disease compared to groups 0, 1, and 3 (57.4% and 38.0% vs. 13.5%-17.2%, respectively; $p<0.001$).

Some PROs also differed between groups, but not consistently (Figure 3). Average shortness of breath (SOBQ) score differed between radiomic groups ($p<0.0001$). Radiomic group 4 had higher average SOBQ compared to radiomic groups 0 and 1 with radiomic group 4 being approximately 17.5 units higher than radiomic groups 0 and 1. In

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

addition, average physical health (SF-12) was lowest for radiomic group 4 and linearly increased to radiomic group 0.

The new radiomic groups remained significantly associated with PFT after adjusting for Scadding stage (Figure 2; $p<0.0001$). Scadding stage also remains significant for all PFT after controlling for radiomic group ($p<0.0041$; Figure 2). The radiomic group also remained associated with SOBQ score ($p=0.032$; Figure 3), while Scadding stage was no longer associated with SOBQ ($p=0.072$; Figure 3). None of the other PROs maintained a significant association with radiomic group after adjustment for Scadding stage ($p>0.38$).

The five most discriminatory radiomic features included kurtosis, which is a measure of shape of the distribution of the HU from an image, as well as four summary measures from the gray level co-occurrence matrix (GLCM). The GLCM measure spatial correlation and similarity of the HU in image voxels near each other. Image visualization of cases with different values of kurtosis and two GLCM measures are shown in Figure 4 with distribution of kurtosis noted in Figure 5 for maximum, median and minimum kurtosis levels in our population. The images visibly show the increase in observable parenchymal abnormalities (left panel Figure 4) with lower kurtosis.

The discriminatory radiomic measures were jointly associated with FVC, FEV1, FEV1/FVC and DLCO ($p<0.001$; Table 3). The radiomic measures explained between 9 and 18% more variation in PFT than adjustment for demographics (age, race, sex,

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

height and BMI) only. For comparison, Scadding stage explained between 0 and 11% more variation in PFT. Kurtosis was associated with FVC, FEV1/FVC and DLCO ($p<0.0001$). A 1-unit increase in kurtosis was associated with a 0.4 L (SE=0.07; $p<0.0001$) increase in FVC. Kurtosis was also positively associated with increasing DLCO (1.59, SE=0.68, $p=0.020$). Kurtosis was negatively associated with FEV1/FVC. A 1-unit increase in kurtosis was associated with a 7% (SE=0.01, $p<0.0001$) decrease in FEV1/FVC ratio.

Validation of Clustering

Pairwise ARI values ranged from 0.3 to 1 and peaking around 0.5 (online supplement Figure 2 left). The maximum ARI of fit clusters with Scadding stage was 0.07. In the linear models, the maximum p-values for the significant of cluster in the demographic adjusted and demographic plus Scadding models were <0.0001 for both models. The corresponding proportions of p-values less than 0.01 were 100% for both models.

Bootstrap samples contained between 184 and 219 unique observations and contained 203 on average. Pairs of bootstrap samples contained between 100 and 159 unique overlapping observations and contained 129 on average or about 63% of the sample is used in each ARI calculation. Pairwise ARI values computed from unique overlapping observations ranged from 0.2 to 1 and had distribution peaking around 0.5 (online supplement Figure 2 right). The maximum ARI of fit clusters with Scadding stage was 0.17. For significance of cluster label in linear models fit to bootstrap samples, maximum p-values for the significant of cluster in the demographic adjusted and

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

demographic plus Scadding models were <0.0001 for both models. The corresponding proportions of p-values less than 0.01 were 100%.

Taken together the validation analysis suggest fairly mild sensitivity of clustering to the analysis pipeline. There is more sensitivity to bootstrapped samples; however cluster groups remain a significant predictor of FVC across all sources of random perturbations in the pipeline.

Discussion

Radiographic manifestations in sarcoidosis are protean. As a result, image characterization is traditionally done visually and not standardized. The utility of visual assessment in sarcoidosis is limited by the intra-observer and inter-observer variability. Radiomics is a more reproducible and computationally efficient approach to characterize HRCT. We used radiomics to quantitatively characterize images from a large, phenotypically diverse cohort of sarcoidosis subjects and demonstrated that radiomics are associated with clinical and patient reported outcomes of disease. Using a common unsupervised learning approach (35), we identified five clusters of cases based on radiomic characterization of their chest CT. These five radiomic-based groups differed significantly by PFTs and several PROs. Notably, each radiomic group included a range of Scadding stages and radiomic group remained significantly associated with PFT after adjusting for Scadding stage. These data suggest that radiomics represents radiographic abnormalities that differ from Scadding stage.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Demographic characteristics of the individual explained a sizable amount of the variation observed in PFT (R^2=50-60%). Radiomics also explained a significant amount of additional variation in PFT (8% to 18%). This additional amount of variation explained is consistent with other quantitative imaging approaches such as CALIPER (16) and those used for investigations of other lung conditions such as systemic sclerosis (38) and diffuse interstitial lung diseases (39).

Except in lung cancer research, much of the work on quantitative CT image analysis has focused only on the HU density (17, 38). Our radiomic panel included summarization of both the density and the spatial characterization. We note that more measures of the GLCM appear as discriminative in the cluster analysis than first-order densitometry measures. This implies that the texture (the spatial information) is perhaps more useful for differentiating various parenchymal abnormalities seen in sarcoidosis.

Although it remains speculative to the reasons for the association between decreased kurtosis and decreased FVC and DLCO, the following explanation is plausible. As kurtosis decreases, more severe parenchymal abnormalities are present that impact the function of the lung, as is shown in Figure 4. Severe parenchymal abnormalities in sarcoidosis are observed as higher HU values. In Figure 5 we observe the presence of a higher frequency of moderately high HU values dampening the peak in the HU distribution and leading to a lower kurtosis in this case. We also found a negative association between lung function and Gaussian weighted GLCM. Figure 4 highlights

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

more abnormalities present with a higher Gaussian weighted GLCM value, as might be expected with this association.

The associations with PROs were less consistent, although important patterns emerged. Radiomic group 4 demonstrated decreased quality of life based on physical function in the SF-12 and worse shortness of breath on the SOBQ, findings that are consistent with the worse lung function in this group. Radiomic group explained substantially less variation in PROs (1 to 15%) compared to PFTs. While the tools we used to measure PROs are validated and used widely, they are not generally correlated with objective measures of lung function (40). As an example, fatigue is a known multi-factorial PRO in sarcoidosis and may impact a number of other PRO (22, 40). For example, fatigue can be associated with shortness of breath and the SOBQ but also with other organ involvement; these two measurements can be confounded by physical function and cardiac or even neurological involvement. Our finding that radiomic group is associated with decreased physical function and increased shortness of breath provides encouragement that more detailed characterization of lung abnormalities, like we developed here, could contribute to a better understanding of the current disconnect between objective measures of lung function and patient experiences.

This study had several strengths. To our knowledge, this was the largest group of sarcoidosis patients with research grade HRCT available allowing for a full 3-D based quantitative analysis. In addition to the high-quality and robust clinical data, the GRADS study also provides a comprehensive and consistently collected set of patient reported

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

outcomes, which reflect important aspects of treatment decision-making processes (23). The results were also internally validated. The validation results taken together suggest our approach is not over fitting to these data.

This study is not without limitations. The GRADS study relied on enrolment from academic centres and could be skewed to a population that was referred for worse disease, was near one of the centres, which were primarily localized to the Eastern US. The demographic is predominantly white and of higher SES as a result and thus did not provide full representation of the spectrum of disease. In addition, many subjects had disease for a decade or longer and were treated. While the same protocol was used across study sites for obtaining the CT images we studied, the scanners themselves differed. We performed harmonization of the scans to mitigate the potential for systematic large differences in scans due to scanner type. In addition, the distribution of PFT measures is not dependent on scanner type such that we expect any differences in the radiomic measures due to scanner type are non-differential as they relate to PFTs. Finally, the number of radiomic clusters (groups) we identified may not directly translate to other populations of sarcoidosis patients or even be the optimal solution for this group of patients. The optimization algorithm we used is common, however, there is no single way to choose the optimal number of clusters.

Radiomic and other quantitative imaging approaches have a major strength in that they can be computed in a time efficient and reliable, reproducible automated procedure. The radiomics package in R took only 3 minutes to compute all the 3-D radiomics

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

measures. This means large quantities of scans can have radiomic profiles computed for aiding in decisions on what scans the radiologist might prioritize for further consideration for visual or clinician evaluation.

In summary, this work provides evidence that in sarcoidosis radiomic quantification is useful for classifying abnormality of the lung along with pulmonary function and to a lesser degree patient reported outcomes. Future work should evaluate the potential of radiomics to capture small changes over time not easily detected by radiologic assessment to further evaluate the potential of radiomics to serve as a clinically useful quantitative measure of sarcoidosis presentation and progression.

Table 1: Patient demographics by radiomic grouping, ordered from least severe (0) to most severe (4) based on average FVC. Unless otherwise noted values are mean (SD).

	N Miss	**Overall**	**0**	**1**	**2**	**3**	**4**	**P-value**
Sample Size		321	75	57	51	58	80	
Age (yr)	0	53 (10)	52 (10)	49.7 (10.5)	52.6 (9.9)	54.2 (9.4)	55.2 (8.9)	0.017
Male, N(%)	0	147 (45.8)	39 (52.0)	19 (33.3)	19 (37.3)	31 (53.4)	39 (48.8)	0.096
White, N(%)	0	233 (72.6)	63 (84.0)	44 (77.2)	40 (78.4)	39 (67.2)	47 (58.8)	0.005
Height (in)	0	67.0 (4.2)	67.7 (4.5)	66.3 (4.2)	66.5 (4.2)	67.2 (4.2)	66.9 (3.7)	0.333
BMI (kg/m^2)	0	30.6 (6.5)	32.6 (5.7)	30.0 (7.4)	27.4 (6.1)	32.5 (6.4)	29.9 (5.9)	<0.001
Scadding	3							<0.001
0		43 (13.5)	17 (23.0)	9 (15.8)	4 (8.0)	8 (14.0)	5 (6.2)	
1		63 (19.8)	27 (36.5)	17 (29.8)	5 (10.0)	6 (10.5)	8 (10.0)	
2		92 (28.9)	19 (25.7)	17 (29.8)	16 (32.0)	22 (38.6)	18 (22.5)	
3		44 (13.8)	9 (12.2)	12 (21.1)	6 (12.0)	10 (17.5)	7 (8.8)	
4		76 (23.9)	2 (2.7)	2 (3.5)	19 (38.0)	11 (19.3)	42 (52.5)	
FVC (L)	6	3.57 (1.14)	4.04	3.90 (1.16)	3.65	3.33	3.01	<0.001

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

			(1.11)		(1.22)	(0.90)	(1.01)	
Scanner Model[a]	0							0.97
Discovery		4 (1.2)	1 (1.3)	1 (1.8)	0 (0.0)	1 (1.7)	1 (1.2)	
iCT128		34 (10.6)	11 (14.7)	8 (14.0)	2 (3.9)	6 (10.3)	7 (8.8)	
LightSpeed		109 (34.0)	25 (33.3)	19 (33.3)	18 (35.3)	22 (37.9)	25 (31.2)	
Optima		3 (0.9)	0 (0.0)	1 (1.8)	1 (2.0)	0 (0.0)	1 (1.2)	
Sensation		8 (2.5)	2 (2.7)	2 (3.5)	1 (2.0)	0 (0.0)	3 (3.8)	
SOM Def		38 (11.8)	10 (13.3)	8 (14.0)	6 (11.8)	5 (8.6)	9 (11.2)	
SOM Def AS		91 (28.3)	20 (26.7)	15 (26.3)	17 (33.3)	15 (25.9)	24 (30.0)	
SOM Def Flash		34 (10.6)	6 (8.0)	3 (5.3)	6 (11.8)	9 (15.5)	10 (12.5)	

[a]Discovery = Discovery CT750 HD; iCT = iCT 128; LightSpeed = LightSpeed VCT; Optima = Optima CT660; Sensation = Sensation 64; SOM Def = SOMATOM Definition; SOM Def AS = SOMATOM Definition AS; SOM Def Flash = SOMATOM Definition Flash

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2: Summary measures of pulmonary function testing and self-reported outcomes by radiomic grouping

	N Miss	Overall	0	1	2	3	4	p-value
Sample Size		321	75	57	51	58	80	
FEV1 (L)	6	2.62 (0.93)	3.11 (0.85)	2.98 (0.95)	2.45 (1.05)	2.50 (0.69)	2.09 (0.74)	<0.001
FVC (L)	6	3.57 (1.14)	4.04 (1.11)	3.90 (1.16)	3.65 (1.22)	3.33 (0.90)	3.01 (1.01)	<0.001
FEV1/FVC (%)	6	0.73 (0.11)	0.77 (0.06)	0.76 (0.10)	0.66 (0.14)	0.75 (0.07)	0.70 (0.11)	<0.001
Obstructive Type[a]; N (%)	6	86 (27.3)	10 (13.5)	9 (15.8)	27 (57.4)	10 (17.2)	30 (38.0)	<0.001
DLCO (mL/min/mmHg)	17	80.2 (23.4)	92.3 (17.2)	86.6 (21.4)	80.2 (22.5)	77.0 (23.4)	65.5 (22.5)	<0.001
FAS[a]	141	29.8 (5.5)	29.3 (5.7)	32.5 (6.6)	29.6 (5.4)	29.0 (4.8)	29.3 (4.7)	0.11
GERDQ[b]	2	7.1 (2.1)	7.0 (1.7)	7.1 (2.2)	6.7 (2.0)	7.2 (2.2)	7.3 (2.3)	0.57
CFQ[c]	4	32.8 (17.0)	35.4 (17.9)	34.8 (14.0)	31.4 (15.5)	31.8 (16.1)	30.4 (19.2)	0.32
SOBQ[d]	18	26.3 (22.8)	19.5 (19.2)	22.1 (22.6)	24.9 (20.8)	27.3 (20.3)	35.8 (26.2)	<0.001
PROMIS Fatiguee[e]	39	25.5 (8.7)	23.2 (7.8)	26.2 (9.4)	25.9 (9.4)	26.1 (8.4)	26.3 (8.7)	0.24
SF12-Physical[f]	7	41.7 (11.2)	43.1 (11.9)	44.5 (11.52)	43.1 (10.8)	39.9 (10.1)	38.8 (10.7)	0.015

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

SF12-Mental[h]	7	48.1 (10.0)	48.0 (9.1)	46.7 (9.51)	48.2 (10.3)	48.3 (10.8)	49.1 (10.5)	0.74

[a]FAS range 5 – 50, higher values indicate more fatigue; [b]GERDQ range 0 – 18, higher values indicate a higher likelihood of GERD; [c]CFQ range 0 -100, higher indicates more cognitive failure; [d] SOBQ range 0 – 120, higher values indicate more shortness of breath; [e] PROMIS Fatigue range 10-50, higher values indicate more fatigue; [f] SF12-Physical-For general U.S. population, mean = 50, SD = 10 and higher values indicate better physical QOL; [g]SF12-Mental-For general U.S. population, mean = 50, SD = 10 and higher values indicate better mental QOL.

Table 3: Results of the regression analysis of the five discriminatory radiomic measures for PFT.[d]

	GLCM Gaussian	**GLCM Inv Gaussian**	**GLCM Sum Entropy**	**Kurtosis**	**GLCM Min**	**P-value**	**Rsq**	**Rsq - base**
FVC	**-0.15 (0.06)[b]**	**-0.21 (0.08)[b]**	0.09 (0.07)	**0.40 (0.07)[a]**	-0.04 (0.09)	<0.0001	0.724	0.611
FEV1	**-0.20 (0.06)[a]**	**-0.41 (0.08)[a]**	**0.15 (0.06)[b]**	0.04 (0.07	**0.34 (0.08)**	<0.0001	0.639	0.502
FEV1/FVC	**-0.02 (0.01)[c]**	**-0.06 (0.01)[a]**	0.01 (0.01)	**-0.07 (0.01)[a]**	**0.11 (0.01)[a]**	<0.0001	0.242	0.048
DLCO	-0.41 (0.58)	-1.35 (0.74)	**2.00 (0.62)[b]**	**1.59 (0.68)[b]**	-0.01 (0.79)	<0.0001	0.599	0.508

[a] P-values <0.001; [b] P-values<0.02; [c] P-values<0.03. [d]Each linear regression model included all five radiomic features and was additionally adjusted for age, gender, race, BMI and height. Bolded cells are statistically significant (p-values in the footnote). GLCM - Gaussian is the sum of the GLCM with a Gaussian weight applied. GLCM-IGE is the sum of the GLCM for the inverse Gaussian weighting scheme. GLCM Sum-Entropy is a measure of the disorder of the GLCM. GLCM Min is the minimum of the GLCM. P-value is for the partial F-test of significance of any of the five radiomic measures.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figures

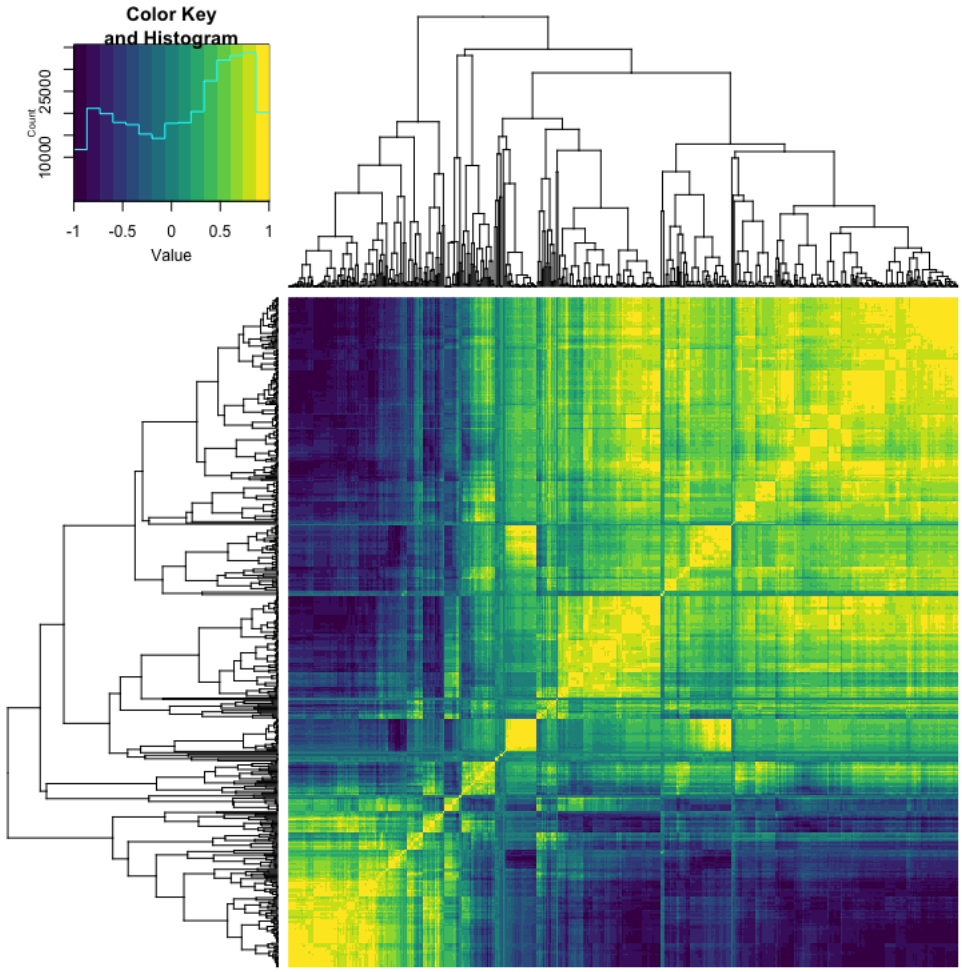

Figure 1: Heat map of the correlation between different radiomic measures for the entire population.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

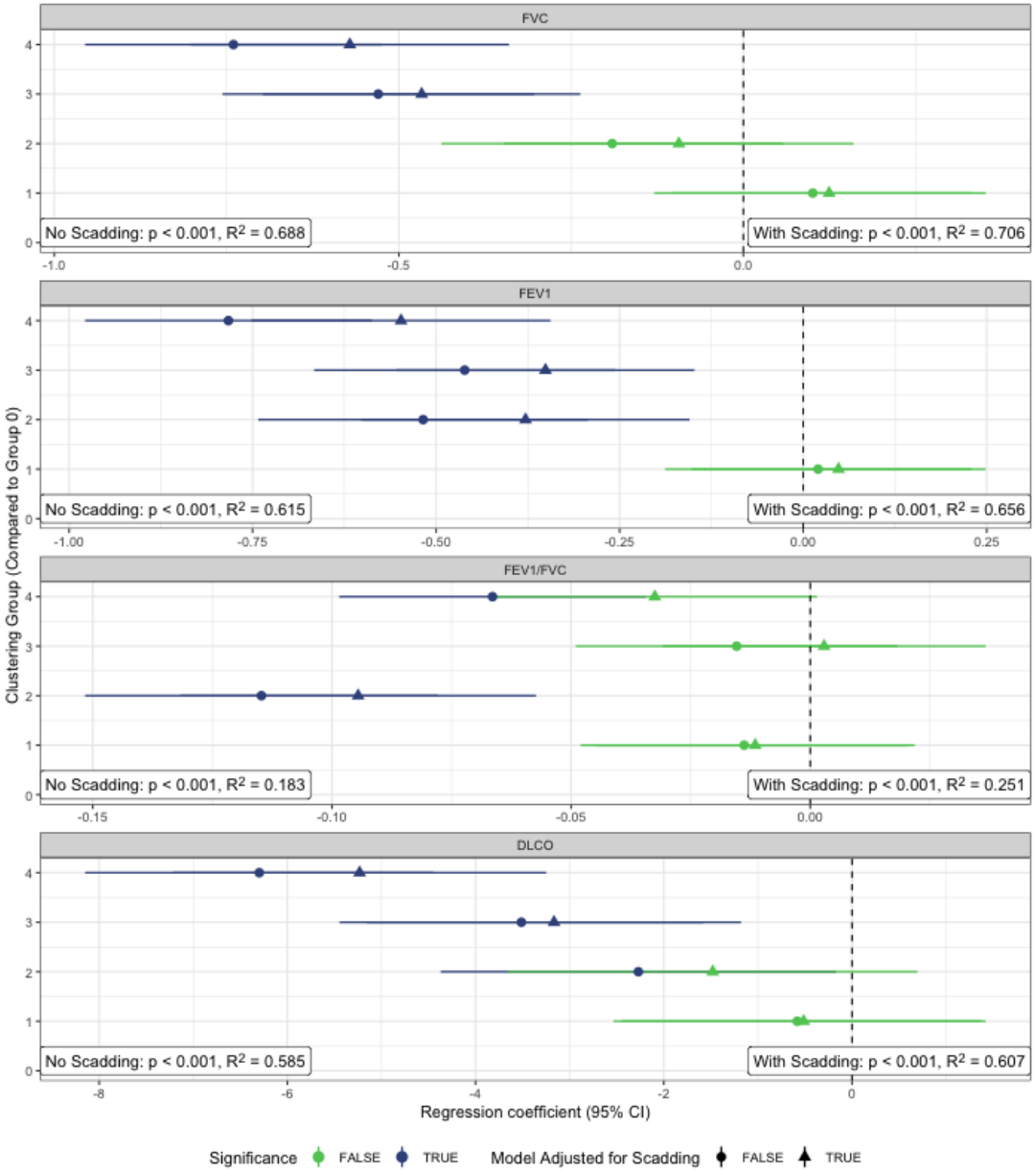

Figure 2: Group differences in average PFT values from regression models with adjustment for demographics (circles) and then also Scadding stage (triangle). The significance of the radiomic group differences from group 0 are in green (p>0.05) and blue (p<0.05). The bottom left text in each panel is the overall p-value for the association between radiomic grouping and PFT outcome along with the R^2. The bottom right text if the overall p-value for the association between radiomic grouping and PFT after additional adjustment for Scadding stage along with the R^2.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

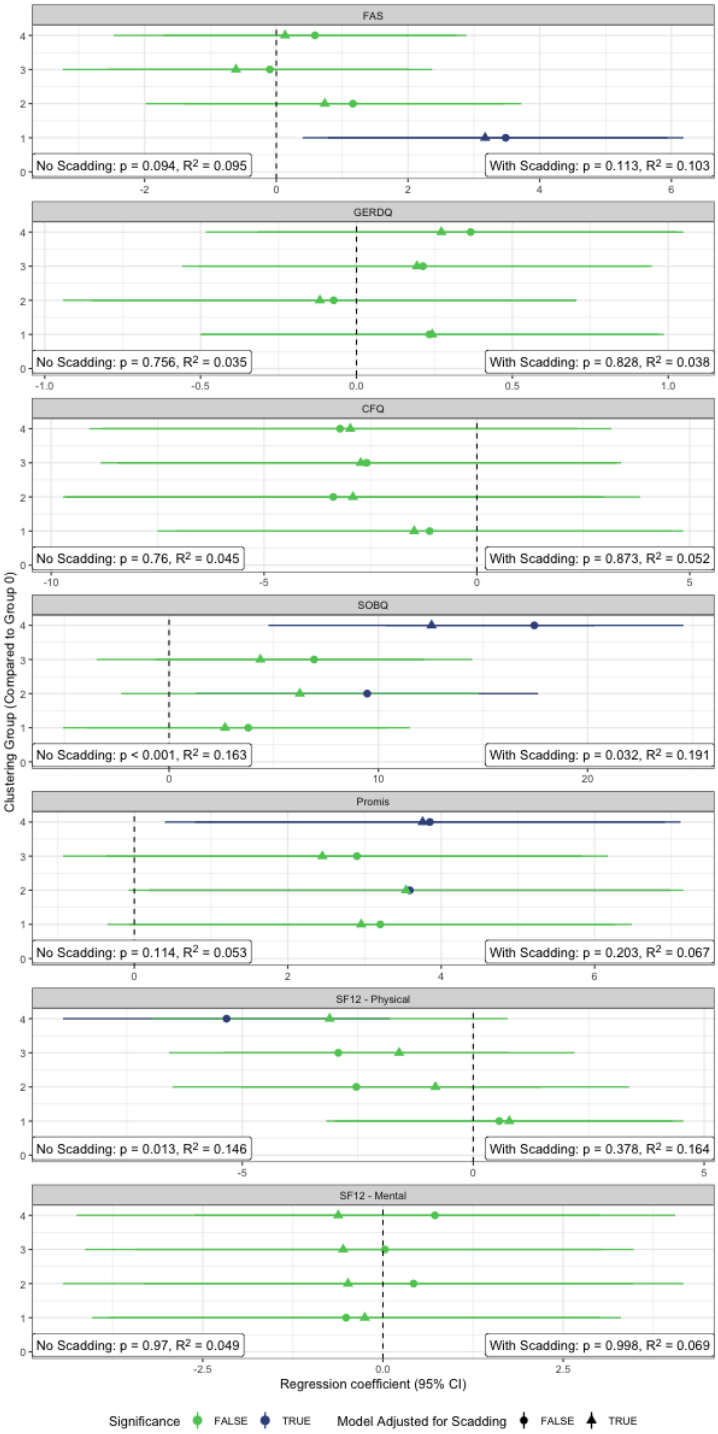

Figure 3: Group differences in average PRO values from regression models with adjustment for demographics (circles) and then also Scadding stage (triangle). The significance of the radiomic group differences from group 0 are in green (p>0.05) and blue (p<0.05). The bottom left text in each panel is the overall p-value for the association between radiomic grouping and PFT outcome along with the R^2. The bottom right text if the overall p-value for the association between radiomic grouping and PFT after additional adjustment for Scadding stage along with the R^2.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

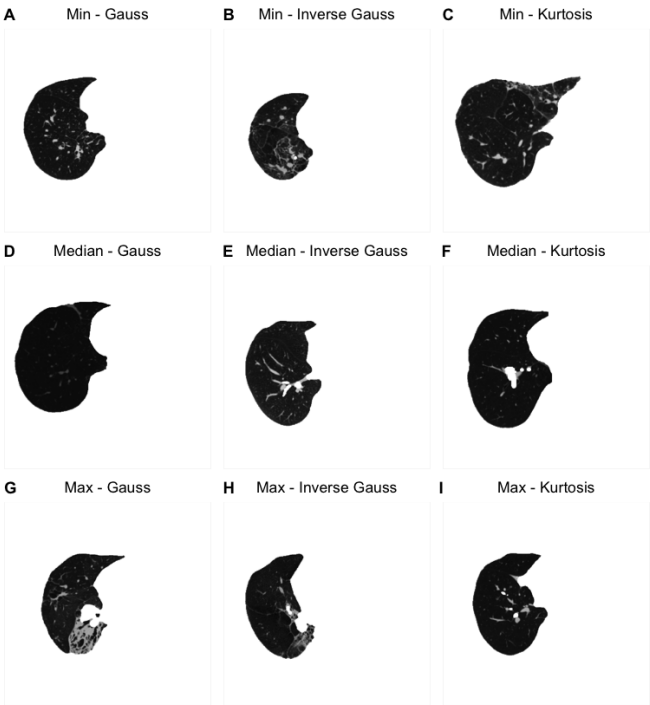

Figure 4: CT images in axial orientation for three patients with minimum, median, and maximum values for GLCM Gaussian (left column), the GLCM Inverse Gaussian, and the kurtosis.

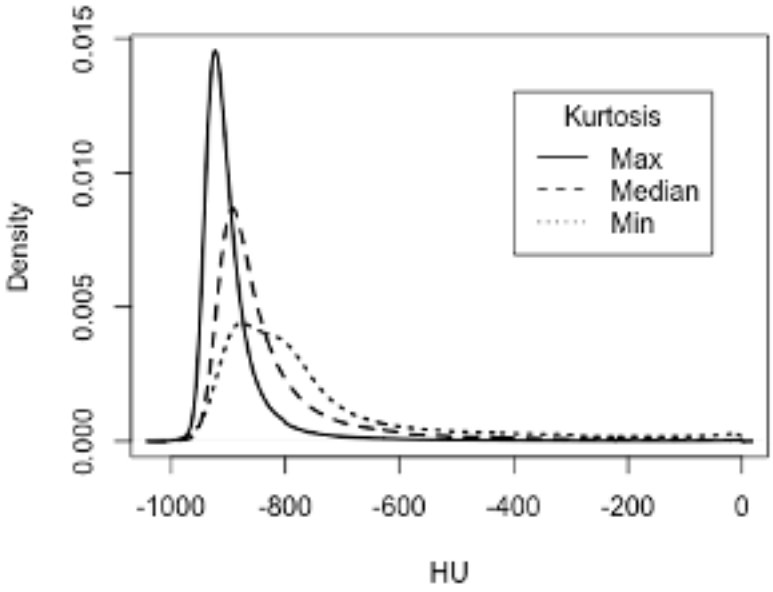

Figure 5: Distribution of HU from HRCT images with the maximum (32.2) (solid), median (9.6) (dashed) and minimum (0.4) (dotted) percentiles of the kurtosis distribution, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Erdal BS, Clymer BD, Yildiz VO, Julian MW, Crouser ED. Unexpectedly high prevalence of sarcoidosis in a representative U.S. Metropolitan population. *Respir Med* 2012;106:893–899.
2. Cox CE, Donohue JF, Brown CD, Kataria YP, Judson MA. The Sarcoidosis Health Questionnaire: a new measure of health-related quality of life. *Am J Respir Crit Care Med* 2003;168:323–329.
3. Wasfi YS, Rose CS, Murphy JR, Silveira LJ, Grutters JC, Inoue Y, Judson MA, Maier LA. A new tool to assess sarcoidosis severity. *Chest* 2006;129:1234–1245.
4. Scadding JG. Prognosis of intrathoracic sarcoidosis in England. A review of 136 cases after five years' observation. *Br Med J* 1961;2:1165–1172.
5. Oberstein A, von Zitzewitz H, Schweden F, Müller-Quernheim J. Non invasive evaluation of the inflammatory activity in sarcoidosis with high-resolution computed tomography. *Sarcoidosis Vasc Diffuse Lung Dis Off J WASOG* 1997;14:65–72.
6. Drent M, De Vries J, Lenters M, Lamers RJS, Rothkranz-Kos S, Wouters EFM, van Dieijen-Visser MP, Verschakelen JA. Sarcoidosis: assessment of disease severity using HRCT. *Eur Radiol* 2003;13:2462–2471.
7. Sluimer I, Schilham A, Prokop M, van Ginneken B. Computer analysis of computed tomography scans of the lung: a survey. *IEEE Trans Med Imaging* 2006;25:385–405.
8. Keijsers RGM, Heuvel DAF van den, Grutters JC. Imaging the inflammatory activity of sarcoidosis. *Eur Respir J* 2013;41:743–751.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

9. Moller DR. Negative clinical trials in sarcoidosis: failed therapies or flawed study design? *Eur Respir J* 2014;44:1123–1126.
10. Nunes H, Uzunhan Y, Gille T, Lamberto C, Valeyre D, Brillet P-Y. Imaging of sarcoidosis of the airways and lung parenchyma and correlation with lung function. *Eur Respir J* 2012;40:750–765.
11. Van den Heuvel DA, de Jong PA, Zanen P, van Es HW, van Heesewijk JP, Spee M, Grutters JC. Chest Computed Tomography-Based Scoring of Thoracic Sarcoidosis: Inter-rater Reliability of CT Abnormalities. *Eur Radiol* 2015;25:2558–2566.
12. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJWL, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, Balagurunathan Y, Gatenby RA, Gillies RJ. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012;30:1234–1248.
13. Haralick RM. Statistical and structural approaches to texture. *Proc IEEE* 1979;67:786–804.
14. Park YS, Seo JB, Kim N, Chae EJ, Oh YM, Lee SD, Lee Y, Kang S-H. Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: comparison with density-based quantification and correlation with pulmonary function test. *Invest Radiol* 2008;43:395–402.
15. Humphries SM, Yagihashi K, Huckleberry J, Rho B-H, Schroeder JD, Strand M, Schwarz MI, Flaherty KR, Kazerooni EA, van Beek EJR, Lynch DA. Idiopathic Pulmonary Fibrosis: Data-driven Textural Analysis of Extent of Fibrosis at Baseline and 15-Month Follow-up. *Radiology* 2017;285:270–278.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

16. Ungprasert P, Wilton KM, Ernste FC, Kalra S, Crowson CS, Rajagopalan S, Bartholmai BJ. Novel Assessment of Interstitial Lung Disease Using the "Computer-Aided Lung Informatics for Pathology Evaluation and Rating" (CALIPER) Software System in Idiopathic Inflammatory Myopathies. *Lung* 2017;195:545–552.
17. Ash SY, Harmouche R, Vallejo DLL, Villalba JA, Ostridge K, Gunville R, Come CE, Onieva Onieva J, Ross JC, Hunninghake GM, El-Chemaly SY, Doyle TJ, Nardelli P, Sanchez-Ferrero GV, Goldberg HJ, Rosas IO, San Jose Estepar R, Washko GR. Densitometric and local histogram based analysis of computed tomography images in patients with idiopathic pulmonary fibrosis. *Respir Res* 2017;18:.
18. Wang J, Li F, Doi K, Li Q. Computerized detection of diffuse lung disease in MDCT: the usefulness of statistical texture features. *Phys Med Biol* 2009;54:6881–6899.
19. Lee G, Lee HY, Park H, Schiebler ML, van Beek EJR, Ohno Y, Seo JB, Leung A. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. *Eur J Radiol* 2017;86:297–307.
20. Ryan SM, Fingerlin TE, Mroz M, Barkes B, Hamzeh N, Maier LA, Carlson NE. Radiomic measures from chest high-resolution computed tomography associated with lung function in sarcoidosis. *Eur Respir J* 2019;54:.
21. Moller DR, Koth LL, Maier LA, Morris A, Drake W, Rossman M, Leader JK, Collman RG, Hamzeh N, Sweiss NJ, Zhang Y, O'Neal S, Senior RM, Becich M, Hochheiser HS, Kaminski N, Wisniewski SR, Gibson KF, GRADS Sarcoidosis Study Group. Rationale and Design of the Genomic Research in Alpha-1 Antitrypsin Deficiency and Sarcoidosis (GRADS) Study. Sarcoidosis Protocol. *Ann Am Thorac Soc* 2015;12:1561–1571.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

22. Kampstra NA, Grutters JC, Beek FT van, Culver DA, Baughman RP, Renzoni EA, Wuyts W, Kouranos V, Wijsenbeek MS, Biesma DH, Wees PJ van der, Nat PB van der. First patient-centred set of outcomes for pulmonary sarcoidosis: a multicentre initiative. *BMJ Open Respir Res* 2019;6:e000394.
23. Wijsenbeek MS, Culver DA. Treatment of Sarcoidosis. *Clin Chest Med* 2015;36:751–767.
24. Jones R, Junghard O, Dent J, Vakil N, Halling K, Wernersson B, Lind T. Development of the GerdQ, a tool for the diagnosis and management of gastro-oesophageal reflux disease in primary care. *Aliment Pharmacol Ther* 2009;30:1030–1038.
25. Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a New Dyspnea Measure: The UCSD Shortness of Breath Questionnaire. *Chest* 1998;113:619–624.
26. Michielsen HJ, De Vries J, Van Heck GL, Van de Vijver FJR, Sijtsma K. Examination of the Dimensionality of Fatigue: The Construction of the Fatigue Assessment Scale (FAS). *Eur J Psychol Assess* 2004;20:39–48.
27. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, Cook K, Devellis R, DeWalt D, Fries JF, Gershon R, Hahn EA, Lai J-S, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays R, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010;63:1179–1194.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

28. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br J Clin Psychol* 1982;21:1–16.

29. Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Med Care* 1996;34:220–233.

30. Ware J, Kosinski M, Turner-Bowker D, Gandek B. How to score SF-12 items. *SF-12 V2 Score Version 2 SF-12 Health Surv* 2002;29–38.

31. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, Mitra S, Shankar BU, Kikinis R, Haibe-Kains B, Lambin P, Aerts HJWL. Robust Radiomics feature quantification using semiautomatic volumetric segmentation. *PloS One* 2014;9:e102107.

32. Fortin J-P, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, Roalf DR, Satterthwaite TD, Gur RC, Gur RE, Schultz RT, Verma R, Shinohara RT. Harmonization of multi-site diffusion tensor imaging data. *NeuroImage* 2017;161:149–170.

33. Fortin J-P, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, Adams P, Cooper C, Fava M, McGrath PJ, McInnis M, Phillips ML, Trivedi MH, Weissman MM, Shinohara RT. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* 2018;167:104–120.

34. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostat Oxf Engl* 2007;8:118–127.

35. Kondo Y, Salibian-Barrera M, Zamar R. RSKC□: An R Package for a Robust and Sparse K-Means Clustering Algorithm. *J Stat Softw* 2016;72:.

medRxiv preprint doi: https://doi.org/10.1101/2022.10.01.22280365; this version posted October 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

36. Witten DM, Tibshirani R. A framework for feature selection in clustering. *J Am Stat Assoc* 2010;105:713–726.

37. Maechler M, Roussseeuw P, Struyf A, Hubert M, Hornik K. cluster: Cluster Analysis Basics and Extensions. Rpackage version 2.1.0. 2019;

38. Camiciottoli G, Orlandi I, Bartolucci M, Meoni E, Nacci F, Diciotti S, Barcaroli C, Conforti ML, Pistolesi M, Matucci-Cerinic M, Mascalchi M. Lung CT Densitometry in Systemic Sclerosis: Correlation With Lung Function, Exercise Testing, and Quality of Life. *Chest* 2007;131:672–681.

39. Shin KE, Chung MJ, Jung MP, Choe BK, Lee KS. Quantitative Computed Tomographic Indexes in Diffuse Interstitial Lung Disease: Correlation With Physiologic Tests and Computed Tomography Visual Scores. *J Comput Assist Tomogr* 2011;35:266–271.

40. Thunold RF, Løkke A, Cohen AL, Hilberg O, Bendstrup E. Patient Reported Outcome Measures (PROMs) in Sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2017;34:2–17.