

Income and Other Contributors to Poor Outcomes in US Sarcoidosis Patients

Logan J Harper MD, Alicia K Gerke MD, MBA, Xiao-Feng Wang PhD, Manuel L Ribeiro Neto MD,

Robert P Baughman MD, Kelli Beyer, Marjolein Drent MD PhD, Marc A Judson MD,

Lisa A Maier MD, MSPH, Leslie Serchuck MD, Noopur Singh, and Daniel A Culver DO

ONLINE DATA SUPPLEMENT

Online Data Supplement

METHODS

Predictors

We sought to determine the effect of household income on health-related quality of life, health outcomes and social outcomes for patients with sarcoidosis who participated in the FSR-SARC questionnaire. Household income was reported as a multiple-choice response with 18 levels, ranging from 0 to >\$250,000. We collapsed the responses to 3 equal tertiles: <\$35,000, \$35,000-\$85,000 and >\$185,000 to create 3 similar sized cohorts. A yearly income of \$35,000 was roughly equivalent to the Medicaid income cut off for a family of 4 in 2015 (\$33,465), the midpoint of our study period. Complete income data were available for 2057 of the included 2318 US respondents.

A priori we identified several potential confounding variables including markers of sarcoidosis extent and markers of social status. Social/demographic factors were based on self-report. Continuous variables included age at time of survey and household size. Categorical variables included gender, race, insurance status, education, organ involvement, medication history and comorbidity development. Following initial review, we divided comorbidity development into “steroid associated comorbidities” and “sarcoidosis associated comorbidities”.

Patients were asked to self-identify for their race with the options “American Indian or Alaska Native”, “Asian Indian”, “Black or African American”, “Chamorro”, “Chinese”, “Filipino”, “Guamanian”, “Japanese”, “Korean”, “Native Hawaiian”, “Other”, “Other Asian”, “Other Pacific Islander”, “White”, “Vietnamese”, “Samoan”, “Ashkenazi Jewish”, “Central American”, “Cuban”, “Dominican (Republic)”, “Hispanic or Latino”, “Mexican”, “Non-Hispanic or Latino”, “Other Hispanic/Latino/Spanish”, “Other Latin American”, “Puerto Rican”, “South American”. Patients were collapsed into “white” (if only selected “white”), “black” (if selected “Black or African American” regardless of if they selected other

racess) and “other” (if they selected any race besides “Black or African American”). This was done due to past studies demonstrating worse outcomes in black sarcoidosis patients.

For insurance status, available responses were “private health insurance”, “single service plan (e.g. dental, vision, prescription)”, “Medicare”, “Medigap”, “Medicaid”, “military health care”, “Indian health service”, “state sponsored health plan”, “other government program”, “No coverage” and “unsure”. We collapsed responses collapsed into “private”, “government” and “uninsured” due to sample size constraints and small numbers of responses in multiple categories.

Patient education included the responses “Never went to school”, “eighth grade or less”, “more than eighth grade but did not complete high school”, “completed a GED”, “high school graduate”, “went to a business, trade, or vocational school instead of high school”, “Went to college, but did not graduate”, “went to a business, trade, or vocational school after high school”, “graduated from a college or university”, “professional training beyond a four-year college or university”. These responses were collapsed down to “high school or less”, “more than high-school without college degree”, “completed college degree”, “more than college degree” due to sample size constraints.

Organ involvement was based on survey response data. Pulmonary involvement was defined as “have you ever been diagnosed with pulmonary (lung) involvement”. Skin involvement was defined as “sarcoidosis of the skin”, “abnormal skin pigmentation”, “lupus pernio”, “nodules under the skin”, “papules”, “plaques”, “other skin lesions diagnosed as caused by sarcoidosis”. Ocular involvement was defined as “sarcoidosis of the eye”, “anterior uveitis”, “posterior uveitis”, “nodules”, “keratoconjunctivitis”, “optic neuritis”, “vasculitis of the retina”. Cardiac involvement was defined as any of the following conditions “that a cardiologist diagnosed as related to the patient’s sarcoidosis”: “atrial arrhythmia”, “ventricular arrhythmia”, “cardiac sarcoidosis”, “cardiomyopathy”, “congestive heart failure”, “heart block”, “heart valve abnormality”, “pericarditis”, “syncope”. Neurologic

involvement was defined as “have you ever been diagnosed with neurosarcoidosis by a medical provider” and included “brain or lining of the brain”, “peripheral neuropathy” or “spinal cord involvement”. Patients with only small fiber neuropathy were not included in neurologic involvement for this study due to the subjectivity of this diagnosis.

Medication use was collected from survey response. Medications included in the survey were prednisone, methylprednisolone, hydroxychloroquine, chloroquine, methotrexate, azathioprine, leflunomide, mycophenolate, infliximab, adalimumab, certolizumab, golimumab, etanercept, rituximab, cyclophosphamide, pentoxifylline, IVIG, thalidomide, and corticotropin. Patients were asked if they were currently taking these medications or had taken them in the past. These responses were collapsed into “currently on medication”, “medication in past” and “never on medication”.

We included development of any new comorbidity following diagnosis of sarcoidosis in our model. Comorbidities were self-reported and included “cancer other than basal cell or squamous cell skin cancer”, “cataracts”, “chronic fatigue syndrome”, “chronic pain syndrome”, “congestive heart failure”, “diabetes”, “depression”, “fibromyalgia”, “glaucoma”, “hypertension”, “lymphoma”, “obesity”, “osteoporosis/osteopenia”, and “sleep apnea”. These comorbidities were separated into sarcoidosis associated comorbidities and steroid associated comorbidities, as described in the manuscript.

Outcomes

We investigated the association of predictors across a range of patient reported outcomes. Health related quality of life was assessed with the Sarcoidosis Health Questionnaire(14), a validated 29 item survey designed to assess the impact of sarcoidosis on patients daily lives.

Health outcomes included ever hospitalized, use of supplemental oxygen, need for mobility assist device (assessed by patient response on the FSR-SARC questionnaire). “Ever hospitalized” was assessed by the question “Have you ever been admitted to the hospital for your sarcoidosis (Not for diagnostic

procedures)?". Patient responses were dichotomous ("yes" or "no"). Patients were asked if they used any oxygen devices including "ventilator", "BiPAP with sip ventilation", "oxygen", "CPAP", "BiPAP" and "inhalers". This was dichotomized into "oxygen", including "ventilator, "BiPAP with sip ventilation" and "oxygen" vs. "no oxygen" which included "CPAP", "BiPAP", "inhalers" and "none". We did this to separate patients who required respiratory support from patients who had no needs or likely OSA. Patients were asked if they used mobility assist devices including "cane", "walker", "wheelchair" or "none". We dichotomized this question into "use mobility device" which included those who selected "wheelchair" or "walker" vs. "no mobility device" which included "cane" and "none". We did this to split the respondents into a group who had significant mobility difficulties vs a group with minimal mobility challenges.

Social outcomes were "effect on family finances" and "need to quit job" (assessed by patient response on the FSR-SARC questionnaire). For "effect on family finances", we dichotomized the 4 responses into two groups: little effect which included "no financial impact" and "slightly affected" and severe effect which included "greatly affected" and "severely affected". We did this to simplify the model as there was no clear distinction between "greatly" and "severely" affected, and we saw little potential difference between "No financial impact" and "slightly affected". "Need to quit job" was assessed by the survey question "have you needed to quit your job due to health effects from sarcoidosis?". Potential responses were "yes", "no", "not applicable, do not work" and "unsure".

Models

Health Outcomes – Predictors of development of new sarcoidosis associated comorbidities.

Our outcome was categorical and dichotomous: "presence of any sarcoidosis associated comorbidity" vs. "absence of any sarcoidosis comorbidity". We investigated association with our predictors using multivariate logistic regression analysis utilizing lasso for variable selection. Our predictors were income

(ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race (categorical, 3 levels), insurance type (categorical, 3 levels), household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous) and medication history (categorical, 3 levels). Final model is reported (Table E4). Variables in the final model were then analyzed via random forest to determine relative importance.

Health Outcomes – Development of Steroid Associated Comorbidity

Our outcome was categorical and dichotomous: “presence of any steroid associated comorbidity” vs. “absence of any steroid comorbidity”. We investigated association with our predictors using multivariate logistic regression analysis utilizing lasso for variable selection. Our predictors were income (ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race (categorical, 3 levels), insurance type (categorical, 3 levels), household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous) and medication history (categorical, 3 levels). Final model is reported in Table E5. Variables in the final model were then analyzed via random forest to determine relative importance.

Health Outcomes – Prevalence of Patients Who Have Ever Been Hospitalized for Sarcoidosis

Our outcome was dichotomous: “patient had ever been hospitalized for sarcoidosis” vs. “patient had never been hospitalized for sarcoidosis”. We investigated association with our predictors using multivariate logistic regression analysis utilizing lasso for variable selection. Our predictors were income (ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race

(categorical, 3 levels), insurance type (categorical, 3 levels), household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous), medication history (categorical, 3 levels), presence of any sarcoidosis associated comorbidities (dichotomous), and presence of any steroid associated comorbidities (dichotomous). Final model is reported in Table E6. Variables in the final model were then analyzed via random forest to determine relative importance.

Quality of Life - SHQ

Our outcome was patient score on the 29 point SHQ (continuous). We investigated association with our predictors using multivariate linear regression analysis utilizing lasso for variable selection. Our predictors were income (ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race (categorical, 3 levels), insurance type (categorical, 3 levels), household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous), medication history (categorical, 3 levels), presence of any sarcoidosis associated comorbidities (dichotomous), and presence of any steroid associated comorbidities (dichotomous). Final model is reported in Table E7. Variables in the final model were then analyzed via random forest to determine relative importance.

Social Outcome - Effect on Family Finances

Our outcome was reported financial impact on patient finances (ordinal, 3 levels). We investigated association with our predictors using multivariate logistic regression analysis utilizing lasso for variable selection. Our predictors were income (ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race (categorical, 3 levels), insurance type (categorical, 3 levels),

household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous), medication history (categorical, 3 levels), presence of any sarcoidosis associated comorbidities (dichotomous), and presence of any steroid associated comorbidities (dichotomous). The final model is reported in table E8. Variables in the final model were then analyzed via random forest to determine relative importance.

Other Outcomes – Oxygen Use

Our outcome was use of supplemental oxygen (dichotomous). We investigated association with our predictors using multivariate logistic regression analysis utilizing lasso for variable selection. Our predictors were income (ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race (categorical, 3 levels), insurance type (categorical, 3 levels), household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous), medication history (categorical, 3 levels), presence of any sarcoidosis associated comorbidities (dichotomous), and presence of any steroid associated comorbidities (dichotomous). The final model is reported in Table E9. Variables in the final model were then analyzed via random forest to determine relative importance.

Other Outcomes – Mobility Device Use

Our outcome was use of mobility device (dichotomous). We investigated association with our predictors using multivariate logistic regression analysis utilizing lasso for variable selection. Our predictors were income (ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race (categorical, 3 levels), insurance type (categorical, 3 levels), household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic

involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous), medication history (categorical, 3 levels), presence of any sarcoidosis associated comorbidities (dichotomous), and presence of any steroid associated comorbidities (dichotomous). The final model is reported in table E10. Variables in the final model were then analyzed via random forest to determine relative importance.

Other Outcomes – Joblessness

Our outcome was “have not lost job due to sarcoidosis”, “lost job due to sarcoidosis”, “never worked”, (categorical, 3 levels). We investigated association with our predictors using multivariate logistic regression analysis utilizing lasso for variable selection. Our predictors were income (ordinal, 3 levels), patient age at time of survey (continuous), gender (categorical, 2 levels), race (categorical, 3 levels), insurance type (categorical, 3 levels), household size (continuous), education (ordinal, 4 levels), pulmonary involvement (dichotomous), neurologic involvement (dichotomous), cardiac involvement (dichotomous), ocular involvement (dichotomous), skin involvement (dichotomous), duration of symptoms (continuous), medication history (categorical, 3 levels), presence of any sarcoidosis associated comorbidities (dichotomous), and presence of any steroid associated comorbidities (dichotomous). The final model is reported in table E11. Variables in the final model were then analyzed via random forest to determine relative importance.

Institutional Review Board

The SARC survey was approved by Advarra institutional review board number CR00126348.

RESULTS

Supplemental Oxygen

In univariate analysis, lower income was strongly associated with use of supplemental oxygen (Table E1). In multivariate analysis, income remained a significant predictor of supplemental oxygen use [$< \$35,000$ OR 1.7 (1.04-2.6), $\$35,000-\$85,000$ OR 1.5 (1.04-2.3)] (Table E1). Random Forest analysis suggested that income was the second major driver of oxygen use, following duration of symptoms (Figure E2).

Mobility Device

In univariate analysis, lower income was also strongly predictive of use of a mobility device (Table E4). In multivariate analysis lower income remained predictive of mobility device use [$< \$35,000$ OR 3.3 (2.0-5.4), $\$35,000-\$85,000$ OR 1.8 (1.2-2.8)] (Table E2). In random forest analysis higher age and neurologic involvement, predictably, were the two major determinants of mobility device use. Of the other drivers, income was the most important (Figure E3).

Need to Quit Job

“Need to quit job due to sarcoidosis” and “Not applicable – did not work” were analyzed together as nominal variables. “Need to quit a job related to sarcoidosis” (Table E3a) and “Did not work” (Table E3b) were associated with lower income in univariate analysis. In multivariate analysis, lower income remained significantly associated with losing a job due to sarcoidosis [$< \$35,000$ OR 4.7 (3.1-7.2), $\$35,000-\$85,000$ OR 1.7 (1.2-2.3)] (Table E3a) or not having worked prior to sarcoidosis diagnosis [$< \$35,000$ OR 4.6 (2.7-8.0), $\$35,000-\$85,000$ OR 1.4 (0.9-2.2)] (Table E3b). In random forest analysis, income was a leading predictor of having to quit a job due to sarcoidosis or not working prior to diagnosis, along with older age and longer duration of symptoms (Figure E4).

Missing Data

Analysis of the missing participants demonstrates a similar composition to our study population based on our predictor variables except for age, medication use, pulmonary involvement and education (online data supplement Table E4a and Table E4b). Regarding outcomes, there was a difference in SHQ, financial impact, lose job for respondents who did not provide their income but no difference in hospitalization, oxygen use or mobility device use (online data supplement Table E5).

DISCUSSION

The self-reported nature of our data resulted in a data set with moderate amounts of missing data. While this has the potential to introduce selection bias, we observed minimal difference in rates of our prediction variables between patients with and without income data (online data supplement Table E12a and Table E12b). We did note difference between those with and without income data in some of our outcomes (New sarcoidosis comorbidity, SHQ, job loss and effect on family finance) (online data supplement Table E13). However, the consistency of our results across all outcomes argues against systematic differences between those with and without missing data.

Legends

Table E1: Predictors of use of supplemental oxygen

Data are expressed as absolute number with percentages, or mean with SD if appropriate.

Variables which were significant ($p < 0.05$) in multivariate analysis are bolded.

Table E2: Predictors of using wheelchair or walker.

Data are expressed as absolute number with percentages, or mean with SD if appropriate.

Variables which were significant ($p < 0.05$) in multivariate analysis are bolded.

Table E3a: Predictors of never working

Data are expressed as absolute number with percentages, or mean with SD if appropriate.

“Never Worked” and “Lost Job” were analyzed together as 3 separate categorical outcomes.

Variables which were significant ($p < 0.05$) in multivariate analysis are bolded.

Table E3b: Predictors of having lost job due to sarcoidosis

Data are expressed as absolute number with percentages, or mean with SD if appropriate.

“Never Worked” and “Lost Job” were analyzed together as 3 separate categorical outcomes.

Variables which were significant ($p < 0.05$) in multivariate analysis are bolded.

Table E4: Full Model – Multivariate Logistic Regression for Developing New Sarcoidosis Associated Comorbidity (Dichotomous)

Table E5: Final Model - Multivariate Logistic Regression for Developing Steroid Associated Comorbidities (Dichotomous)

Table E6: Final Model - Multivariate Logistic Regression for Ever Being Hospitalized for Sarcoidosis
(Dichotomous)

Table E7: Supplemental Table E7 - Full Model - Multivariate Linear Regression for Sarcoidosis Health
Questionnaire (Continuous)

Table E8: Final Model - Multivariate Logistic Regression for Effect on Family Finances (Dichotomous)

Table E9: Final Model - Multivariate Logistic Regression for Use of Supplemental Oxygen (Dichotomous)

Table E10: Final Model - Multivariate Logistic Regression for Use of Mobility Device (Dichotomous)

Table E11: Final Model - Multivariate Logistic Regression for Quit Job vs. Never Worked vs. Still Working
(Categorical)

Supplemental Table E12a – Differences in covariates based on missing income data

Supplemental Table E12b – Differences in covariates based on missing income data

Supplemental Table 13 – Difference in outcome based on missing income data

Figure E1: Comorbidity Frequency

Rate of each measured comorbidity in our cohort.

Figure E2: Random Forest Model –Supplemental Oxygen Use

Random Forest plots of analysis of variables included in multivariate model selected by LASSO for prediction of patient using supplemental oxygen. Higher values of “Mean Decrease Gini” reflect increasing proportional importance of the individual variables to predicting this categorical outcome.

Figure E3: Random Forest Model –Walker or Wheelchair Use

Random Forest plots of analysis of variables included in multivariate model selected by LASSO for prediction of patient using a walker or wheelchair. Higher values of “Mean Decrease Gini” reflect increasing proportional importance of the individual variables to predicting this categorical outcome.

Figure E4: Random Forest Model –Lose Job or Never Worked

Random Forest plots of analysis of variables included in multivariate model selected by LASSO for prediction of patient never having worked or losing their job after diagnosis of sarcoidosis. Higher values of “Mean Decrease Gini” reflect increasing proportional importance of the individual variables to predicting this categorical outcome.

SUPPLEMENTAL TABLES

Table E1: Supplemental Oxygen Use

	Univariate OR	Multivariate - Lasso n= 1091 OR
Income		
>\$85,000	Ref	Ref
\$35,000-\$85,000	1.8 (1.3-2.3)	1.5 (1.04-2.3)
<\$35,000	2.5 (1.8-3.3)	1.7 (1.04-2.6)
Age	1.03 (1.02-1.04)	
Male	0.9 (0.7-1.1)	
Race		
White	Ref	
African American	1.7 (1.3-2.2)	1.4 (0.97-2.1)
Other	0.9 (0.6-1.3)	0.8 (0.4-1.4)
Insurance Type		
Private	Ref	
Government	2.1 (1.7-2.5)	1.5 (1.1-2.2)
No coverage	1.4 (0.8-2.4)	0.7 (0.2-1.9)
Household Size	0.9 (0.8-0.97)	
Patient Education		
>College	Ref	
College	1.5 (1.1-2.1)	
Post HS education	1.6 (1.2-2.3)	
HS	2.0 (1.4-2.9)	
Pulmonary Involvement	3.4 (2.3-4.9)	2.7 (1.5-4.7)
Neurologic Involvement	1.8 (1.4-2.4)	1.7 (1.2-2.5)
Cardiac Involvement	2.2 (1.7-2.8)	2.1 (1.5-2.9)
Ocular Involvement	1.7 (1.4-2.1)	
Skin Involvement	1.6 (1.3-2.1)	
Duration of Symptoms	1.03 (1.03-1.04)	1.01 (1.0-1.03)
Medication History		
Never	Ref	
Past	2.7 (1.7-4.2)	
Current	3.4 (2.3-5.3)	
New Sarcoidosis Comorbidity	2.6 (2.1-3.3)	1.4 (0.966-2.0)
New Steroid Comorbidity	2.7 (2.1-3.3)	1.7 (1.2-2.4)

Table E2: Use of Mobility Device (Wheelchair/Walker)

	Univariate OR	Multivariate – Lasso n= 1307 OR
Income		
>\$85,000	Ref	Ref
\$35,000-\$85,000	2.4 (1.7-3.4)	1.8 (1.2-2.8)
<\$35,000	5.5 (3.9-7.7)	3.3 (2.0-5.4)
Age	1.01 (1.004-1.03)	1.02 (1.01-1.04)
Male	0.8 (0.6-1.001)	
Race		
White	Ref	
AA	1.7 (1.3-2.2)	
Other	1.9 (1.3-2.7)	
Insurance Type		
Private	Ref	Ref
Government	3.2 (2.6-4.1)	1.8 (1.2-2.6)
No coverage	1.6 (0.8-2.9)	0.7 (0.3-2.0)
Household Size	0.9 (0.9-1.0)	
Patient Education		
>College	Ref	
College	1.8 (1.3-2.5)	
Post HS education	1.5 (1.1-2.1)	
HS	1.7 (1.1-2.5)	
Pulmonary Involvement	0.8 (0.6-1.1)	
Neurologic Involvement	6.2 (4.8-8.1)	4.7 (3.4-6.7)
Cardiac Involvement	1.4 (1.1-1.8)	
Ocular Involvement	2.3 (1.8-2.9)	1.3 (0.9-1.8)
Skin Involvement	1.6 (1.2-2.1)	
Duration of Symptoms	1.02 (1.01-1.03)	
Medication History		
Never	Ref	
Past	1.9 (1.2-3.1)	
Current	3.1 (2.0-4.9)	
New Sarcoidosis Comorbidity	4.1 (3.1-5.4)	2.4 (1.7-3.5)
New Steroid Comorbidity	2.8 (2.2-3.6)	

Table E3a: Need to Quit Job due to Sarcoidosis

	Univariate OR	Multivariate – Lasso n=1317 OR
Income		
>\$85,000	Ref	Ref
\$35,000-\$85,000	2.2 (1.7-2.8)	1.7 (1.2-2.3)
<\$35,000	8.3 (6.2-10.9)	4.7 (3.1-7.2)
Age	1.0 (0.99-1.01)	1.0 (0.98-1.01)
Male	0.6 (0.5-0.8)	
Race		
White	Ref	
African American	1.9 (1.5-2.4)	
Other	2.2 (1.6-3.1)	
Insurance Coverage		
Private	Ref	Ref
Government	4.4 (3.6-5.4)	2.3 (1.7-3.3)
No coverage	2.5 (1.5-4.1)	0.9 (0.4-1.9)
Household Size	1.0 (0.9-1.02)	1.0 (0.9-1.1)
Patient Education		
>College	Ref	
College	1.5 (1.1-2.0)	
Post HS education	2.6 (2.0-3.5)	
HS	2.5 (1.8-3.5)	
Pulmonary Involvement	1.3 (0.995-1.7)	
Neurologic Involvement	4.3 (3.3-5.5)	3.8 (2.7-5.4)
Cardiac Involvement	1.7 (1.3-2.1)	
Ocular Involvement	2.4 (1.9-2.9)	
Skin Involvement	2.1 (1.7-2.6)	
Duration of Symptoms	1.03 (1.02-1.04)	1.01 (1.0-1.3)
Medication History		
Never	Ref	
Past	5.3 (3.3-8.5)	
Current	7.4 (4.7 -11.6)	
New Sarcoidosis Comorbidity	4.9 (3.9-6.1)	2.8 (2.1-3.9)
New Steroid Comorbidity	3.1 (2.5-3.8)	1.5 (1.1-2.1)

Table E3b: Did Not Work Prior to Sarcoidosis Diagnosis

	Univariate OR	Multivariate - Lasso OR
Income		
>\$85,000	Ref	Ref
\$35,000-\$85,000	1.7 (1.2-2.5)	1.4 (0.9-2.2)
<\$35,000	6.2 (4.2-9.2)	4.6 (2.7-8.0)
Age	1.04 (1.02-1.05)	1.03 (1.01-1.06)
Male	0.5 (0.3-0.7)	
Race		
White	Ref	
African American	1.3 (0.9-1.9)	
Other	1.8 (1.1-2.8)	
Insurance Coverage		
Private	Ref	Ref
Government	4.7 (3.5-6.2)	3.4 (2.2-5.2)
No coverage	1.7 (0.8-3.8)	0.4 (0.1-2.0)
Household Size	0.9 (0.8-0.99)	1.1 (0.9-1.3)
Patient Education		
>College	Ref	
College	1.6 (1.04-2.6)	
Post HS education	3.5 (2.3-5.5)	
HS	3.6 (2.2-5.8)	
Pulmonary Involvement	0.8 (0.6-1.1)	
Neurologic Involvement	1.2 (0.8-1.9)	1.3 (0.7-2.3)
Cardiac Involvement	0.9 (0.6-1.3)	
Ocular Involvement	1.5 (1.2-2.1)	
Skin Involvement	1.2 (0.9-1.7)	
Duration of Symptoms	1.0 (0.997-1.02)	1.0 (0.98-1.02)
Medication History		
Never	Ref	
Past	0.8 (0.6-1.3)	
Current	0.8 (0.5-1.1)	
New Sarcoidosis Comorbidity	1.3 (0.99-1.8)	0.9 (0.6-1.4)
New Steroid Comorbidity	1.4 (1.02-1.8)	0.7 (0.5-1.1)

Supplemental Table E4

Full Model – Multivariate Logistic Regression for Developing New Sarcoidosis Associated Comorbidity (Dichotomous)

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4517	0.1544	88.3934	<.0001
Income					
\$0-\$34,999	1	0.8624	0.1708	25.4891	<.0001
\$35,000-\$84,999	1	0.3486	0.1507	5.3497	0.0207
>\$85,000					
Neurologic Involvement - Yes	1	0.9651	0.1742	30.7122	<.0001
Cardiac Involvement - Yes	1	0.8077	0.1528	27.9389	<.0001
Skin Involvement - Yes	1	0.4734	0.137	11.9401	0.0005
Duration of Symptoms	1	0.0387	0.00619	39.0598	<.0001

Supplemental Table E5

Final Model - Multivariate Logistic Regression for Developing Steroid Associated Comorbidities (Dichotomous)

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.36	0.5277	68.2701	<.0001
Age	1	0.0362	0.00787	21.1432	<.0001
Income					
\$0-\$34,999	1	0.2855	0.2197	1.6889	0.1937
\$35,000-\$84,999	1	0.4138	0.1703	5.9083	0.0151
>\$85,000					
Insurance					
Government	1	0.4209	0.186	5.1228	0.0236
None	1	-0.2061	0.4265	0.2335	0.629
Private					
Race					
African American	1	0.7059	0.1986	12.6334	0.0004
Other	1	0.654	0.2813	5.4036	0.0201
White					
Neurologic Involvement - Yes	1	0.8622	0.1982	18.9277	<.0001
Ocular Involvement - Yes	1	0.4759	0.1617	8.6616	0.0032
Lung Involvement - Yes	1	0.7533	0.2015	13.9776	0.0002
Cardiac Involvement - Yes	1	0.3004	0.1726	3.0288	0.0818
Medication History					
Current	1	0.947	0.2369	15.9772	<.0001
Past					
Never					
Male	1	-0.3833	0.1712	5.0121	0.0252
Duration of Symptoms	1	0.0379	0.00771	24.1639	<.0001

Supplemental Table E6

Final Model - Multivariate Logistic Regression for Ever Being Hospitalized for Sarcoidosis
(Dichotomous)

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4128	0.0973	211.0159	<.0001
Neurologic Involvement - Yes	1	0.8601	0.1495	33.0885	<.0001
Cardiac Involvement - Yes	1	1.1462	0.1343	72.845	<.0001
Develop Sarcoidosis Comorbidity	1	0.77	0.1217	40.0561	<.0001

Supplemental Table E7 - Full Model - Multivariate Linear Regression for Sarcoidosis Health Questionnaire (Continuous)

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	4.650213	0.211847	21.95	<.0001
Age	0.018234	0.002986	6.11	<.0001
<i>Income</i>				
\$0-\$34,999	-0.50758	0.084475	-6.01	<.0001
\$35,000-\$84,999	-0.25063	0.065851	-3.81	0.0002
>\$85,000	0	.	.	.
<i>Insurance</i>				
Government	-0.15653	0.069728	-2.24	0.025
None	0.012982	0.172312	0.08	0.94
Private	0	.	.	.
Household Size	-0.05855	0.020696	-2.83	0.0048
Neurologic Involvement	-0.19454	0.071927	-2.7	0.007
No	0	.	.	.
Cardiac Involvement	-0.11723	0.065268	-1.8	0.0728
No	0	.	.	.
Lung Involvement	-0.34891	0.076643	-4.55	<.0001
No	0	.	.	.
Ocular Involvement	-0.08788	0.061984	-1.42	0.1566
No	0	.	.	.
Skin Involvement	-0.44755	0.060503	-7.4	<.0001
No	0	.	.	.
Medication History				
Current	-0.57511	0.091123	-6.31	<.0001
Past	-0.41361	0.098029	-4.22	<.0001
Never	0	.	.	.
Male	0.324641	0.066778	4.86	<.0001
No	0	.	.	.
Develop Steroid Comorbidity	-0.1829	0.062015	-2.95	0.0033
No	0	.	.	.
Develop Sarcoidosis Comorbidity	-0.61504	0.061225	-10.05	<.0001
No	0	.	.	.

Supplemental Figure E8

Final Model - Multivariate Logistic Regression for Effect on Family Finances (Dichotomous)

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.4301	0.547	6.836	0.0089
Age	1	-0.0367	0.00843	18.9272	<.0001
Income					
\$0-\$34,999	1	2.0634	0.2441	71.475	<.0001
\$35,000-\$84,999	1	0.9879	0.1899	27.0682	<.0001
>\$85,000					
Insurance					
Government	1	-0.0448	0.1971	0.0516	0.8204
None	1	0.8253	0.5486	2.2634	0.1325
Private					
Education					
High School or Less	1	-0.1735	0.2646	0.4301	0.512
>HS without College Degree	1	0.2113	0.2277	0.8612	0.3534
College	1	-0.1456	0.2245	0.4209	0.5165
> College					
Neurologic Involvement - Yes	1	0.8073	0.2045	15.5834	<.0001
Cardiac Involvement - Yes	1	0.498	0.1871	7.0831	0.0078
Ocular Involvement - Yes	1	0.2136	0.1748	1.4927	0.2218
Lung Involvement - Yes	1	0.5233	0.2214	5.5882	0.0181
Medication History					
Current	1	1.0404	0.2827	13.542	0.0002
Past	1	0.8808	0.305	8.3383	0.0039
Never					
Develop Steroid Comorbidity	1	0.3913	0.1735	5.0859	0.0241
Develop Sarcoidosis Comorbidity	1	0.804	0.168	22.9062	<.0001
Duration of Symptoms	1	-0.00891	0.00793	1.2624	0.2612

Supplemental Table E9

Final Model - Multivariate Logistic Regression for Use of Supplemental Oxygen (Dichotomous)

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
<i>Intercept</i>	1	-3.6966	0.328	127.0225	<.0001
Income					
\$0-\$34,999	1	0.5065	0.2378	4.537	0.0332
\$35,000-\$84,999	1	0.4333	0.1999	4.6998	0.0302
>\$85,000					
Insurance					
Government	1	0.4186	0.1901	4.8493	0.0277
None	1	-0.3812	0.5256	0.5261	0.4683
Private					
Race					
African American	1	0.3516	0.1935	3.3017	0.0692
Other	1	-0.2824	0.3227	0.7662	0.3814
White					
Neurologic Involvement - Yes	1	0.5332	0.1896	7.9054	0.0049
Lung Involvement - Yes	1	0.9747	0.2879	11.4629	0.0007
Cardiac Involvement - Yes	1	0.7224	0.1715	17.7403	<.0001
Develop Steroid Comorbidity	1	0.5168	0.1854	7.7668	0.0053
Develop Sarcoidosis Comorbidity	1	0.3173	0.1797	3.1158	0.0775
Duration of Symptoms	1	0.0132	0.00671	3.8851	0.0487

Supplemental Table E10

Final Model - Multivariate Logistic Regression for Use of Mobility Device (Dichotomous)

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.7322	0.514	84.7482	<.0001
Age	1	0.0218	0.00833	6.8415	0.0089
Income					
\$0-\$34,999	1	1.1953	0.2482	23.1861	<.0001
\$35,000-\$84,999	1	0.6022	0.2254	7.1385	0.0075
>\$85,000					
Insurance					
Government	1	0.5773	0.1881	9.4243	0.0021
None	1	-0.2582	0.4963	0.2706	0.6029
Private					
Neurologic Involvement - Yes	1	1.5535	0.174	79.6752	<.0001
Ocular Involvement - Yes	1	0.2517	0.1698	2.1979	0.1382
Develop Sarcoidosis Comorbidity	1	0.8858	0.1874	22.3448	<.0001

Supplemental Table E11

Final Model - Multivariate Logistic Regression for Quit Job vs. Never Worked vs. Still Working (Categorical)

Parameter	Quit Job vs. Never Worked	DF	Estimate	Standard Error	Wald Chi- Square	Pr > ChiSq
Intercept	Quit Job	1	-2.4092	0.5043	22.819	<.0001
Intercept	Never Worked	1	-4.8579	0.7094	46.8917	<.0001
Age	Quit Job	1	-0.00486	0.00802	0.3675	0.5444
Age	Never Worked	1	0.0417	0.0108	14.9593	0.0001
Income						
\$0-\$34,999	Quit Job	1	1.558	0.2106	54.7408	<.0001
\$0-\$34,999	Never Worked	1	1.5436	0.2728	32.0269	<.0001
\$35,000-\$84,999	Quit Job	1	0.5013	0.173	8.3982	0.0038
\$35,000-\$84,999	Never Worked	1	0.3259	0.2441	1.782	0.1819
>\$85,000						
Insurance						
Government	Quit Job	1	0.8483	0.1731	24.0254	<.0001
Government	Never Worked	1	1.2177	0.217	31.4898	<.0001
None	Quit Job	1	-0.088	0.3728	0.0558	0.8133
None	Never Worked	1	-0.8288	0.771	1.1558	0.2823
Private						
Household Size	Quit Job	1	-0.0254	0.0543	0.2191	0.6397
Household Size	Never Worked	1	0.0792	0.0748	1.1193	0.2901
Neurologic Involvement - Yes	Quit Job	1	1.3449	0.1715	61.4931	<.0001
Neurologic Involvement - Yes	Never Worked	1	0.2587	0.2853	0.8222	0.3645
Develop Steroid Comorbidity	Quit Job	1	0.4305	0.1587	7.3538	0.0067
Develop Steroid Comorbidity	Never Worked	1	-0.3036	0.2158	1.9779	0.1596
Develop Sarcoidosis Comorbidity	Quit Job	1	1.0468	0.1557	45.2033	<.0001
Develop Sarcoidosis Comorbidity	Never Worked	1	-0.056	0.2144	0.0683	0.7938
Duration of Symptoms	Quit Job	1	0.0137	0.00678	4.1008	0.0429
Duration of Symptoms	Never Worked	1	0.000849	0.00916	0.0086	0.9261

Supplemental Table E12a – Differences in covariates based on missing income data

	Complete Income Data n=2057	Missing Income Data n=261	p
Age (mean (SD))	51.8 (10.8)	55.0 (10.9)	<0.001
Gender = Male (%)	531 (25.8)	69 (26.4)	0.9
Race (%)			0.7
African American	370 (18%)	48 (19%)	
Other	167 (8%)	24 (19%)	
White	1506 (74%)	183 (72%)	
Duration (mean (SD))	11.63 (11.7)	11.92 (13.0)	0.7
Medication Use (%)			0.009
Current	1191 (59%)	121 (49%)	
None	271 (14%)	42 (17%)	
Past	550 (27%)	84 (34%)	
Pulmonary Involvement	1604 (83%)	190 (77%)	0.03
Skin Involvement	1217 (65%)	140 (60%)	0.2
Ocular Involvement	667 (37%)	89 (41%)	0.3
Cardiac Involvement	429 (26%)	58 (29%)	0.5
Neurologic Involvement	343 (20%)	38 (19%)	0.8
Insurance (%)			0.6
Government	606 (30%)	69 (27%)	
None	70 (4%)	8 (3%)	
Private	1348 (67%)	173 (70%)	
Patient Education			0.06
>College	398 (21%)	59 (26%)	
College	613 (32%)	58 (26%)	
Post HS education	596 (31%)	62 (28%)	
HS	323 (17%)	46 (20%)	
Household Size	2.7 (1.5)	2.7 (1.7)	0.5

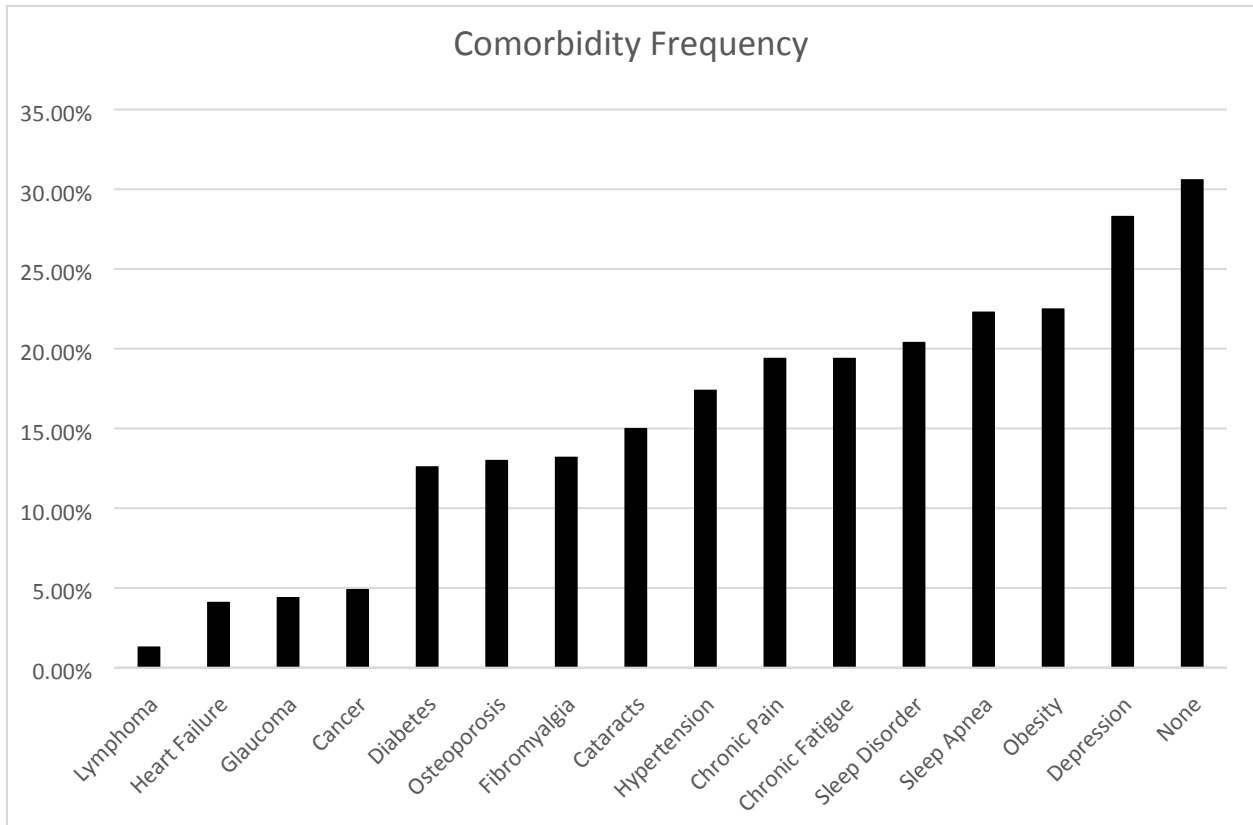
Supplemental Table E12b – Differences in covariates based on missing income data

	Complete Income Data n=2057	Missing Income Data n=261	p
Cancer = Yes (%)	85 (5%)	15 (7%)	0.3
Cataracts = Yes (%)	274 (15%)	34 (15%)	1.0
Chronic fatigue = Yes (%)	428 (23%)	36 (16%)	0.01
Chronic pain = Yes (%)	359 (20%)	39 (17%)	0.4
Heart Failure = Yes (%)	77 (4%)	8 (4%)	0.7
Diabetes = Yes (%)	234 (13%)	25 (11%)	0.5
Depression = Yes (%)	524 (29%)	57 (25%)	0.2
Fibromyalgia = Yes (%)	253 (14%)	18 (8%)	0.02
Glaucoma = Yes (%)	76 (4%)	14 (6%)	0.2
Hypertension = Yes (%)	323 (18%)	35 (15%)	0.4
Lymphoma = Yes (%)	25 (1%)	2 (1%)	0.7
Obesity = Yes (%)	420 (23%)	42 (18%)	0.1
Osteoporosis = Yes (%)	234 (13%)	33 (14%)	0.6
Sleep Apnea = Yes (%)	414 (23%)	46 (19%)	0.3
Sleep disorder = Yes (%)	377 (21%)	43 (19%)	0.5
none = Yes (%)	555 (30%)	75 (33%)	0.5

Supplemental Table 13 – Difference in outcome based on missing income data

	Complete Income n= 2057	Missing Income n= 261	p
New Sarcoidosis Comorbidity (Yes%)	979 (53%)	105 (46%)	0.04
New Steroid Comorbidity (Yes%)	1002 (55%)	121 (53%)	0.7
Ever Hospitalized	747 (38%)	84 (35%)	0.4
SHQ (mean (SD))	3.6 (1.1)	4.0 (1.2)	<0.001
Lost Job (%)			<0.001
No	1092 (56%)	135 (59%)	
Not applicable - did not work	220 (11%)	45 (20%)	
Yes	628 (32%)	50 (22%)	
Affect Finances (%) = No/Slight Effect	1059 (53%)	144 (63%)	0.005
Oxygen Device (%) = using Oxygen	419 (21%)	57 (23%)	0.6
Mobility Device= Yes (%)	348 (17%)	35 (14%)	0.2

Figure E1 – Comorbidity Frequency



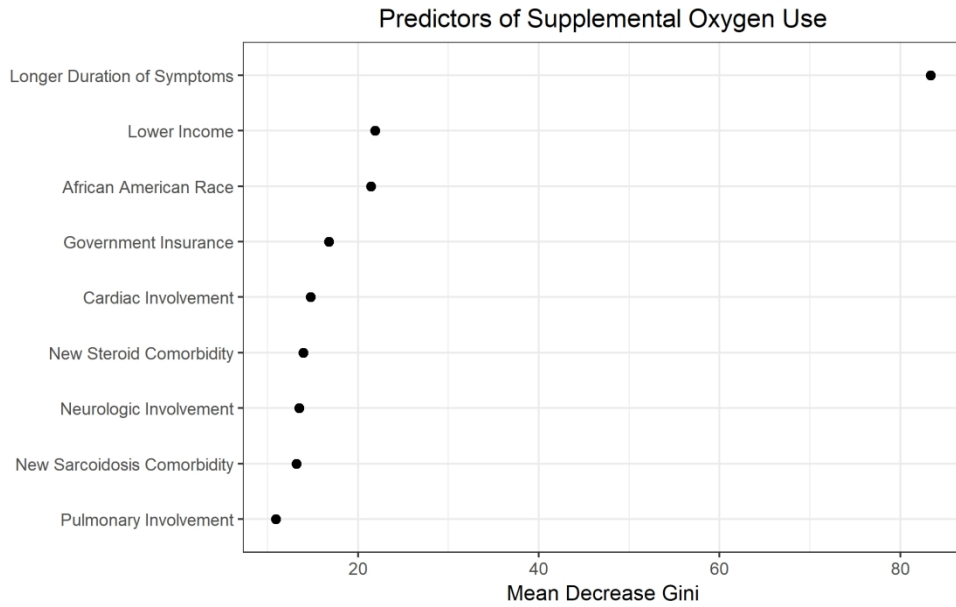


Figure E2: Random Forest Model –Supplemental Oxygen Use
 Random Forest plots of analysis of variables included in multivariate model selected by LASSO for prediction of patient using supplemental oxygen. Higher values of “Mean Decrease Gini” reflect increasing proportional importance of the individual variables to predicting this categorical outcome.

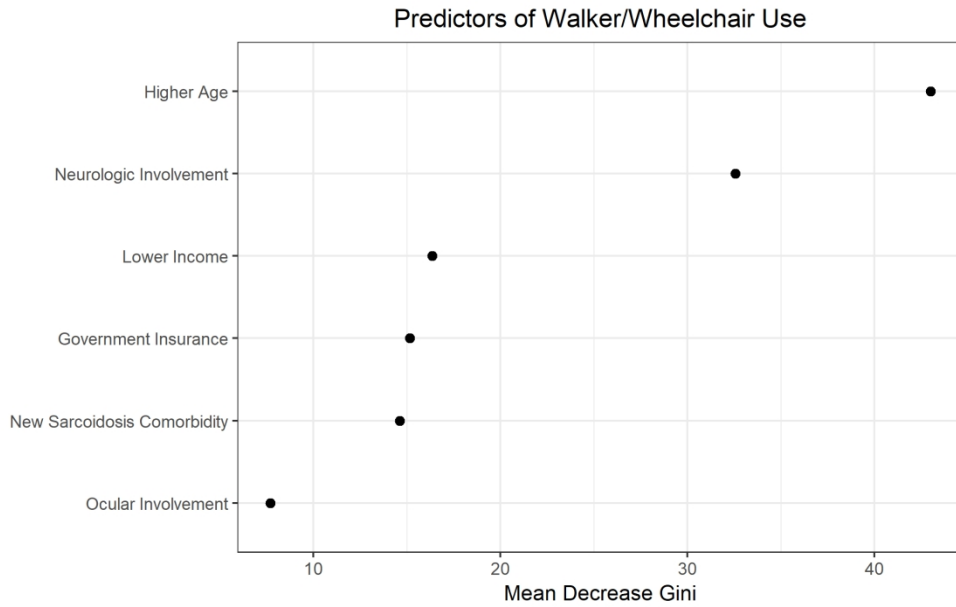


Figure E3: Random Forest Model –Walker or Wheelchair Use
Random Forest plots of analysis of variables included in multivariate model selected by LASSO for prediction of patient using a walker or wheelchair. Higher values of “Mean Decrease Gini” reflect increasing proportional importance of the individual variables to predicting this categorical outcome.

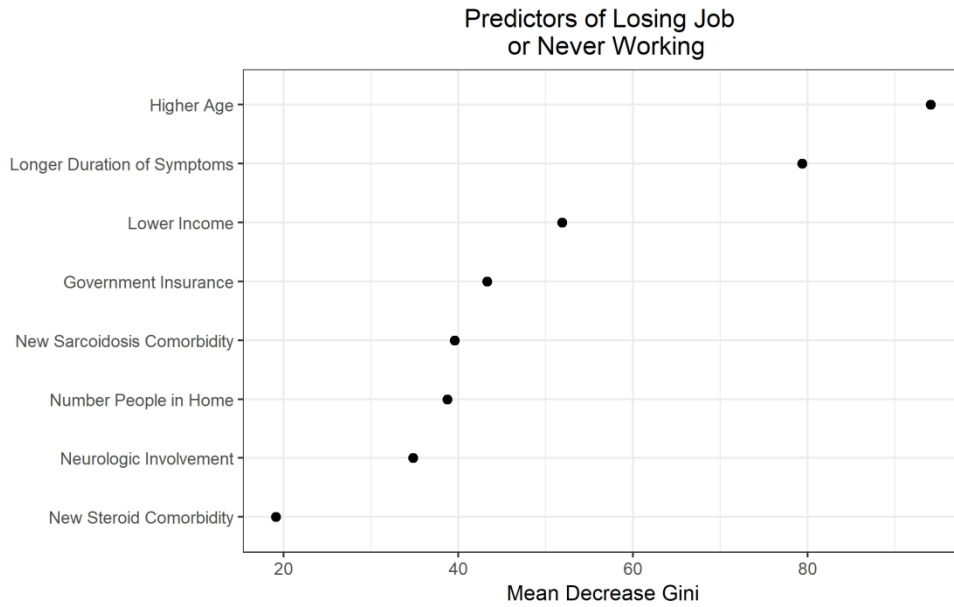


Figure E4: Random Forest Model –Lose Job or Never Worked
 Random Forest plots of analysis of variables included in multivariate model selected by LASSO for prediction of patient never having worked or losing their job after diagnosis of sarcoidosis. Higher values of "Mean Decrease Gini" reflect increasing proportional importance of the individual variables to predicting this categorical outcome.