

:: Abstract

Since the onset of the COVID-19 pandemic, the question of racial differences has been at the forefront of prognostic thought in determination of high-risk groups. We identified a total of 410 unique Medical Record Numbers (MRNs) in a retrospective data collection of patients in the Ochsner LSU Health System (LSUH) in Shreveport and Monroe, Louisiana. Data was collected from a mix of rapid and regular PCR nasal swabs collected at the two locations from 4/1/2020 to 4/30/2020. Data collected included symptoms, race, ethnicity, occupation, gender and age. to determine if there was a statistical difference in presenting symptoms based on race.

Black/African American patients were the most represented race (74%) with COVID-19 and most patients were non-Hispanic (70.15%) with a female predominance amongst all races. Gender was evaluated as a covariable, which demonstrated a slight female predominance in symptoms, with females representing 58.73%, and males representing 41.27% of the African American cohort. Fever was frequently accompanying cough, shortness of breath, headache, arthralgias/myalgias, non-Hispanic ethnicity, and male sex. In the grouped data, the top 10 rules predicted either shortness of breath or fever with COVID-19, and five of the rules were associated with African American race. Caucasian and African American patients had similar rates of anosmia.

While evaluating the racial distribution of COVID-19 as it pertained to symptoms, gender, and occupation, African American patients were statistically more affected by COVID-19 in Northern Louisiana, which was not observed in Southern Louisiana.

:: Introduction

SARS-CoV-2 infection is the cause of severe acute respiratory syndrome, commonly known as COVID-19. COVID-19 has a poor prognosis and has been associated with approximately 4.9 million deaths worldwide since December 2019, with the global number of cases exceeding 242 million [1]. The COVID-19 pandemic has been known to affect the African American/Black (AA) community in the United States disproportionately since its beginnings [2].

Shreveport, LA represents a unique opportunity to study outcomes of COVID-19 in the AA population because its population is majority AA. AA individuals represent 13.4% of the United States population as of 2019; however, the AA population makes up an estimated 57.1% of inhabitants in Shreveport as of 2019 [3]. Additionally, Shreveport serves as a large catchment area which includes the the surrounding region (known as the Ark-La-Tex), with Ochsner Shreveport being LERN Region 7's sole Level I Trauma Center [4].

We obtained de-identified data from COVID-positive patients arriving to the Ochsner facility in April of 2020. Additionally, we assessed the prevalence of both common COVID-19 symptoms (such as cough or anosmia) and other symptoms that are thought to be atypical (such as cardiac or GI events). Our primary goal was to determine a) if racial disparities manifested themselves in the Shreveport population, and if so, b) the nature of those disparities. We also look at how race relates to discharge disposition and admission status. Our study excluded pregnant women, prisoners, and pediatric patients.

Race	Count	Percent	Mean Age	% Female	% Male	% Hispanic	% Not Hispanic
AI	10	2.43	32.30	80.00	20.00	60.00	30.00
AS	4	0.97	45.25	25.00	75.00	0.00	75.00
AA	306	74.27	43.86	62.75	36.93	0.33	69.28
CC	79	19.17	47.46	45.57	54.43	1.27	84.81

Table 1. Distribution by race and gender, for AI=American Indian/Native American, AS = Asian, AA = African American/Black, and CC = Caucasian/White. Nearly 75% of the patient population was AA and nearly 2/3rds female. Note that only 399 patients have race data available, and not all percentages add up to 100%.

Racial Disparities in COVID-19 Symptoms

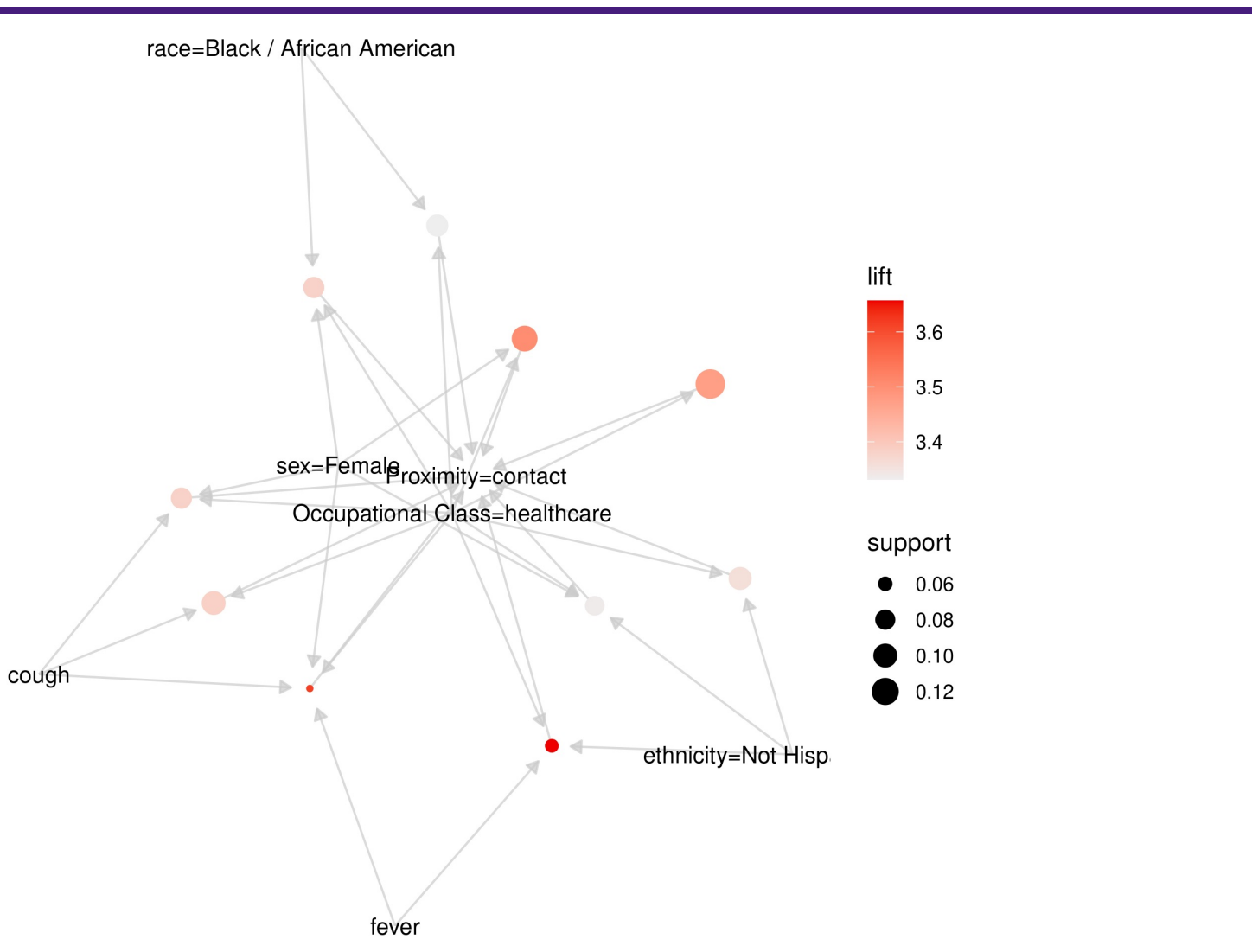


Figure 1. arulesViz visualization for top 10 association rules for the refactored input set. Race, ethnicity, gender, occupation class, and proximity were frequently associated with the association rules that were mined for the data.

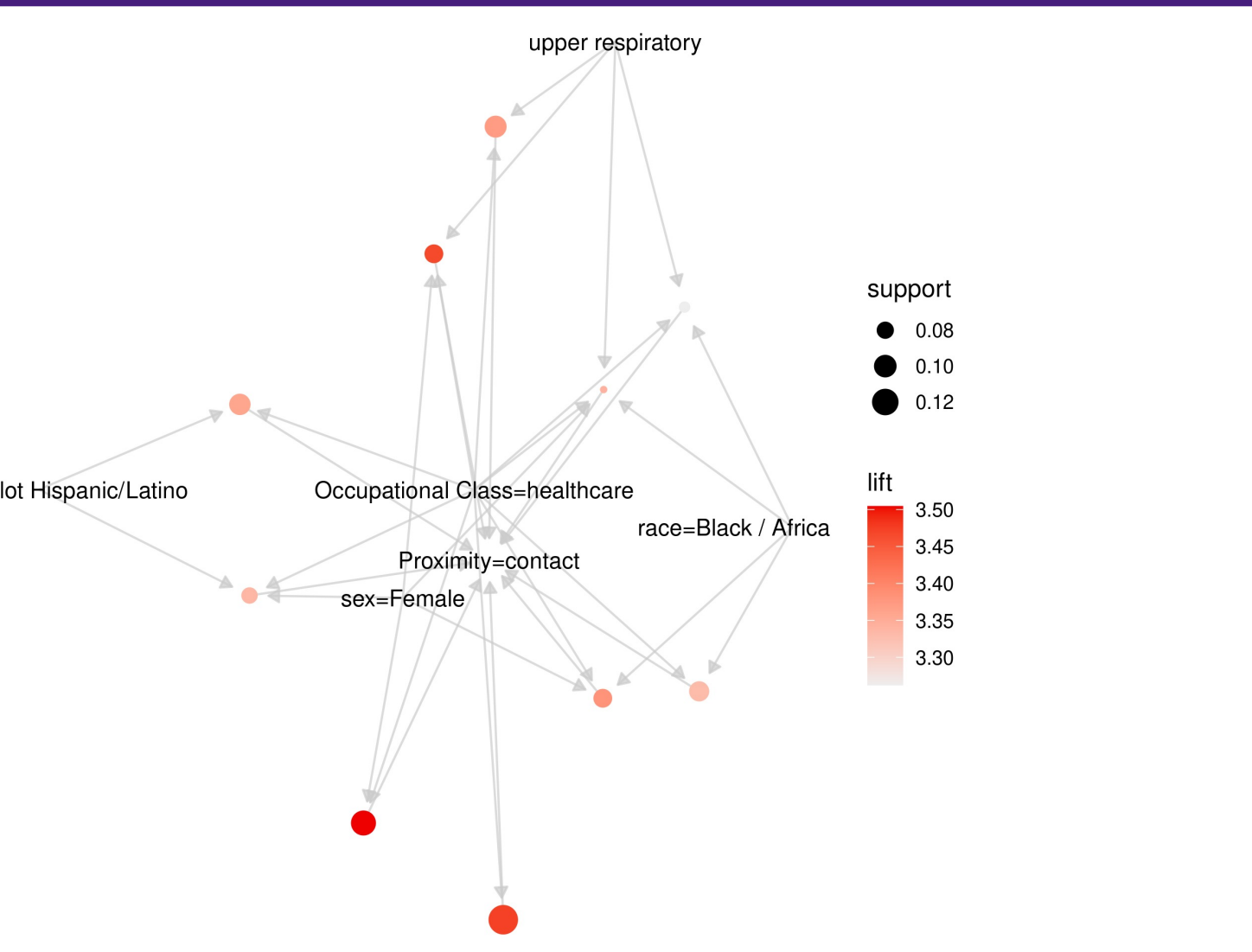


Figure 2. arulesViz visualization for top 10 association rules for the collapsed input set. A variety of features are associated. Race, ethnicity, gender, occupation class, and proximity were frequently associated with the association rules that were mined for the data.

:: Methods

We received data for 2,484 encounters (visits) for 453 de-identified patients; 7 patients fell within our exclusion criteria and were removed. We also removed records with obviously erroneous values (for instance, BMI well in excess of the largest ever recorded) when they were irreconcilable with other data, resulting in 1,961 encounters over 412 patients. Race was defined for 399 of these patients.

Hypothesis testing was performed on the basis of a) domain and b) normality of the dependent variable (if numeric). For non-numeric data, we used the χ^2 test. Numeric variables were first tested for normality using Shapiro-Wilk's normality test for $p > \alpha = 0.1$. This limit is standard practice with Shapiro-Wilk because the null hypothesis is that the distribution is normal and a larger α value indicates a greater caution towards accepting normality. For normal data, we used either Student's t -test for a binary independent variable, or One-way Analysis of Variance (ANOVA) for others. For non-normal data, we used the Wilcoxon rank sum test and Kruskal-Wallis respectively. For all hypothesis testing, we rejected the null hypothesis when $p \leq \alpha = 0.05$.

We considered the data set in two forms: *refactored* (where atypical symptoms were broken into granular categories) and *collapsed*, where collapsed data was grouped into GI, cardiovascular and upper respiratory symptoms. We performed association rule mining (ARM) using the "arules" package for R [5]. Association rules are directional rules that suggest that if one attribute is present (the antecedent), other specific attributes (the subsequent) can also be expected to be present. We visualized the rules using the "arulesViz" package for R and selected rules which had a lift ≥ 1 .

Count of...	<i>p</i>	AA	CC	Δ
<i>Typical Symptoms*</i>	0.033	2.92 ± 2.07	2.37 ± 2.01	0.56
<i>Refactored Symptoms</i>	0.589	0.43 ± 0.77	0.47 ± 0.75	-0.04
<i>Collapsed Symptoms*</i>	0.026	1.04 ± 0.77	0.82 ± 0.84	0.21

Table 2. Hypothesis testing for a number of typical, refactored, and collapsed symptoms. "*" denotes a statistically significant difference in the number of such symptoms. Typical and collapsed symptoms were found to be significantly different between AA and CC.

Symptom	% of AA	% of CC	$\Delta\%$	Symptom	% of AA	% of CC	$\Delta\%$
<i>Shortness of breath*</i>	39.22	21.52	17.70	<i>Weakness</i>	10.78	8.86	1.92
<i>Cough</i>	66.34	54.43	11.91	<i>Rash</i>	0.33	0.00	0.33
<i>Arthralgias or myalgias</i>	26.47	16.46	10.01	<i>Nausea</i>	11.44	11.39	0.05
<i>Ageusia</i>	7.52	1.27	6.25	<i>Hyposmia</i>	1.31	1.27	0.04
<i>Coryza</i>	11.76	6.33	5.44	<i>Dizziness</i>	1.96	2.53	-0.57
<i>Headache</i>	18.95	13.92	5.03	<i>Dysgeusia</i>	2.94	3.80	-0.86
<i>Anosmia</i>	8.17	5.06	3.11	<i>Fever</i>	46.41	48.10	-1.70
<i>Hypotension</i>	6.21	3.80	2.41	<i>Diarrhea</i>	12.75	15.19	-2.44
<i>Sore throat</i>	13.40	11.39	2.01	<i>Altered mental status</i>	6.54	11.39	-4.86

Table 3. Comparison of "typical" symptoms measured in COVID-19 patients. The $\Delta\%$ column depicts the relative difference between the AA and CC cohorts. Occurrence of shortness of breath, cough, and arthralgias or myalgias were at least 10% greater in the AA than in the CC population. Only shortness of breath was significantly different ($p=0.005$).

:: Results

74.27% of patients in our data set were AA (Table 1), of which about 63% were female and 37% were male; the vast majority (84.81%) were non-Hispanic. We also captured data for Asian Americans (AS) and American Indian/Alaskan Natives (AI), but because of the relatively small size of these cohorts, they are not considered for subsequent processing. The AA cohort was statistically significantly different in both, the number of typical symptoms, and the number of collapsed, atypical symptoms, in both cases having greater incidence (Table 2).

Concerning the typical symptoms of COVID-19, the AA population had a greater relative incidence of shortness of breath, cough, and either arthralgias or myalgias (Table 3); of this, only shortness of breath was statistically significantly different. When we look at the collapsed atypical symptoms, only upper respiratory symptoms were statistically significantly different (Table 4).

Our ARM analysis showed that the most frequent association rules involved race, ethnicity, gender, occupational class, and proximity (Figures 1 and 2). We mined 124 rules which had a lift of at least 1. Lift is a measure of the performance of a targeting model (association rule) at predicting or classifying cases as having an enhanced response (with respect to the population as a whole), measured against a random choice targeting model. The top 5 mined rules included:

- {Sex = Female, Occupational Class = Healthcare} \rightarrow {Proximity = Contact} (lift = 3.504)
- {Occupational Class = Healthcare} \rightarrow {Proximity = Contact} (lift = 3.474)
- {Sex = Female, upper respiratory, Occupational Class = healthcare} \rightarrow {Proximity = contact} (lift = 3.466)
- {Sex = Female, Race = AA, Occupational Class = healthcare} \Rightarrow {Proximity = contact} (lift = 3.381)
- {upper respiratory, Occupational Class = healthcare} \Rightarrow {Proximity = contact} (lift = 3.751)

Rule 10 was also instructive: {race = AA, upper respiratory, Occupational Class = healthcare} \Rightarrow {Proximity = contact} (lift = 3.262).

Value	% of AA	% of CC	$\Delta\%$
<i>GI symptoms</i>	23.86	26.58	-2.73
<i>Cardiovascular</i>	12.75	7.59	5.15
<i>Upper respiratory*</i>	66.99	48.10	18.89

Table 4. Comparison of collapsed or grouped "atypical" symptoms measured in COVID-19 patients. The $\Delta\%$ column depicts relative difference between the AA and CC cohorts. Only upper respiratory symptoms had a greater than 10% relative difference and it was also significantly different ($p=0.002$).

:: Discussion and Conclusions

Our analysis shows that the AA and CC cohorts respond differently to COVID-19 infection, at least in the period of April 2020. AA cohort was more prone to upper respiratory symptoms than the CC cohort (including shortness of breath). During this interval, the α variant of SARS-CoV-2 was the only variant known to exist. As several variants have been identified since this point, it is not known whether or not the symptoms associated with these variants show a similar disparity. These results not only confirm previous findings of disparity [2] but also provide more clarity to the medical impact of this disparity.

Due to the population distribution in Shreveport, insufficient data was available for the study of the AI and AS populations. Our data trend towards disparity in the AI population, but low sample size suggests reduced statistical power and we have left this to a future study.

Our ARM analysis strongly implied that female AA healthcare workers were likely to have contact proximity with COVID-19 infected individuals, and that they were associated with upper respiratory symptoms. The top 10 association rules predicted contact proximity and this fact in isolation is not particularly interesting as it is a function of their job; however, they also included instances where upper respiratory symptoms were part of the antecedent, possibly indicating the female healthcare workers may be at greater risk of symptomatic infection than their male counterparts. This should be studied further before any conclusions can be made.

:: Acknowledgment

Research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103424-20.

:: References

- H. Ritchie, E. Matheiu, L. Rhodes-Guirao, C. Appel, C. Giattino, E. Ortiz-Ospina, J. Hasell, B. Macdonald, D. Beltekian, M. Roser. "Coronavirus Pandemic (COVID-19)". <https://ourworldindata.org/coronavirus>. Last accessed Oct 26, 2021.
- R. Kullar, J.R. Marcelin, T.H. Swartz, D.A. Piggott, R.M. Gil, T.A. Mathew, T. Tan. "Racial Disparity of Coronavirus Disease 2019 in African American Communities". *J. Infect Dis.* 222(6), pp 890-893. Jun 30, 2020.
- United States Census Bureau. "QuickFacts: Shreveport city, Louisiana". Last accessed Oct 26, 2021.
- Louisiana Emergency Response Network. "State Designated Trauma Centers". <https://learn.la.gov/trauma/>. Last accessed Oct 26, 2021.
- M. Hahsler, C. Buchta, B. Gruen, K. Hornik, I. Johnson, C. Borgelt. "Package 'arules'". Version 1.6.8.