Demographic and Sex Differences in the Likelihood of COVID-19 Mortality amongst COVID-19 Positive Pediatric Patients

Author: Keanu Renne-Glover

Academic Advisor: Dr. Levi Waldron

Capstone Mentor: Matthew Ramos

Target Journal: Clinical Infectious Diseases

Abstract

This study is a retrospective cross-sectional study on the differences in the odds of mortality between laboratory-confirmed COVID-19 positive male and female pediatric patients ages 0-19 years from April 2020 to February 2021. The purpose of this study is to determine if the likelihood of mortality for pediatric patients (ages =< 19) with positive laboratory-confirmed COVID-19 is affected by sex assigned at birth, adjusting for comorbidity status and socio-demographic characteristics.

Exploratory analysis of the missing values was performed on age group, sex assigned at birth, the comorbidity status, hospital admissions status, and race/ethnicity stratified on the outcome variable, death status. Missing data analysis was performed to determine the mechanism of missingness: Missing Completely at Random (MCAR), Missing at Random (MAR), Missing Not at Random (MNAR). Descriptive statistical analysis was performed on socio-demographic characteristics – age group, sex assigned at birth, the comorbidity status, hospital admissions status, and race/ethnicity – and stratified on the outcome variable, death status. Odds ratios were calculated via simple logistic regression for each demographic exposure variable on the outcome variable (mortality), assessing the significance with alpha = 0.05. Multivariate Logistic regression was used to examine the associations between mortality from COVID-19 and demographic characteristics.

Complete case analysis was performed because of the inability to determine the mechanism of missingness. After adjusting for hypothesized mediators, there was no significant association between male pediatric patients and mortality when compared to female patients (OR = 0.993, 95% CI 0.964 - 1.023). Additionally, the mediators “Comorbidity Status” and “Hospitalization Admissions Status” were not shown to have a mediating effect on “Sex assigned at birth” as the primary exposure.

Given the incompleteness of the CDC dataset, non-differential bias of the outcome may be present within the statistical analysis because of the omission of all missing observations from the study based on the outcome variable, Death Status; thus, resulting in study results biased towards the null.

Introduction

COVID-19 is considered a relatively mild illness for infected pediatric patients (ages =< 18 years) with approximately 83% of confirmed cases presenting with mild to moderate infection severity; however a small proportion of pediatric patients develop and suffer from severe COVID-19 complications and outcomes requiring ICU admission (Gotzinger et al., 2020; Leidman, 2021; Tsankov et al., 2021). Gotzinger et al. performed a multi-national, multi-center cohort study to explore the relationship between pediatric age (ages =< 18) and COVID-19 illness severity for hospitalized children and adolescent patients. Although their sample population biases their results to represent patients with more severe COVID-19 symptoms because participants were sampled from hospitalized patients, Gotzinger et al. showed that pediatric patients (ages =< 18) generally suffered from a less severe COVID-19 disease outcome when compared to adults (Gotzinger et al., 2020). According to Dhochak et al., one possible pathophysiological explanation for the decreased severity of COVID-19 infection in pediatric patients (ages < 18), is that the ACE-2 counter regulatory enzyme exhibits lung protective effects

via the limitation of angiotensin-2 mediated pulmonary inflammatory response (Dhochak, Singhal, Kabra, & Lodha, 2020). The presence of comorbidities such as diabetes and hypertension results in a decreased concentration in the ACE-2 enzyme and/or down-regulation of the ACE-2 enzyme activity, resulting in more severe COVID-19 prognosis (Dhochak et al., 2020). Pediatric patients (ages <18), however, have naturally high levels of ACE-2; therefore, contributing to less severe disease for this age group when compared to older adults and the elderly (Dhochak et al., 2020).

Data on the symptoms and prognosis of COVID-19 infection in pediatric patients is more sparse than adult individual-level data; however, the trends between illness severity and the presence of comorbidities in adults can be seen in COVID-19 positive pediatric patients with underlying childhood comorbidities when compared to pediatric patients without comorbidities (Ludvigsson, 2020; Zhang, Peres, Silva, & Camargos, 2020). Tsankov et al. reveals in their meta-analysis on the severity of COVID-19 infection in pediatric patients with childhood comorbidities (i.e. childhood obesity, cardiovascular disease, chronic respiratory disease, renal disease, immune disorders, metabolic disease, neurological disorders, etc.) that COVID-19 positive pediatric patients with underlying illnesses have a higher risk of severe COVID-19 outcomes and associated mortality than pediatric patients without comorbidities (Tsankov et al., 2021). Examining 42 studies containing 275,661 pediatric patients without any comorbidities and 9353 pediatric patients with comorbidities, Tsankov et al. showed that pediatric patients with underlying conditions had 2.81 times (95% CI 1.31 - 6.02) higher risk of COVID-19 mortality (Tsankov et al., 2021). Additionally, when examining the risk of pediatric patients with obesity on COVID-19 severity in relation to pediatric patients without comorbidities, Tsankov et al. derived a relative risk ratio of 2.87 (95% CI 1.16 – 7.07) (Tsankov et al., 2021) – indicating that childhood obesity increases the likelihood of COVID-19 infection severity (Tsankov et al., 2021).

Mohamed et al. reviewed literature on the role that androgen sensitivity plays as a risk factor for COVID-19 disease severity. Androgens are hormones that, although present in different concentrations in both men and women, play a crucial role in the production of male physiological traits as well as male reproductive development and maintenance. Androgen receptors facilitate viral entry for the COVID-19 pathogen; thus, providing a clinical explanation for the observed low fatalities in pediatric patients as well as the differences in COVID-19 infection prevalence between the sexes in adults (Mohamed, Moulin, & Schioth, 2021). Ultimately, although comorbidities in pediatric COVID-19 patients have moderate exploration in the literature, the literature does not consider the affects that socio-demographic characteristics may have on mortality risk for male and female pediatric patients (ages =<19). Therefore, the purpose of this study is to determine if the likelihood of mortality for pediatric patients (ages =< 19) with positive laboratory-confirmed COVID-19 is affected by sex assigned at birth, adjusting for comorbidity status and socio-demographic characteristics.

Methods

Data Collection Procedures

The COVID-19 Case Surveillance Public Use Data is a 12 data element public use data set contained within the COVID-19 case surveillance system database. It is comprised of individual-level data reported to the state and territory health departments, including the New York City and District of Columbia Health Department ("COVID-19 Case Surveillance Public Use Data," 2021). COVID-19 is a disease that is required to be reported in all state health departments, and multiple territorial health departments in addition to both the New York City Health Department and the District of Columbia Health Department ("COVID-19 Case Surveillance Public Use Data," 2021). The health departments in these jurisdictions confirm cases of COVID-19 based on national standardized criteria and may gather additional information at their discretion in accordance with their own local guidelines ("COVID-19 Case Surveillance Public Use Data," 2021). The de-identified data collected by the health departments in these jurisdictions is voluntarily shared with the Center for Disease Control and Prevention (CDC).

Ethical Considerations

The COVID-19 Case Surveillance Public Use Data is a publicly available de-identified and anonymized dataset; therefore, Human Subject Research determination by the CUNY SPH IRB was not required. Data cells with low frequency records (<5) and indirect identifiers such as "date of first positive specimen" have been suppressed and anonymized to protect individual privacy. The deidentified data within the public use data set include the following elements: sex, race/ethnicity, 10-year age group, laboratory confirmed COVID-19 status, comorbidity status, ICU admission status, hospital admission status, and death status. A national standardized case definition was used to the probably, confirmed, and suspected COVID-19 cases and deaths.

Study Design

This study is a retrospective cross-sectional study on the differences in the odds of mortality between laboratory-confirmed COVID-19 positive male and female pediatric patients ages 0-19 years from April 2020 to February 2021.

Study participants (Inclusion / Exclusion Criteria)

Study participants include all completed laboratory confirmed COVID-19 cases from the voluntary-reporting publicly available data set.

Multicollinearity between Hospitalization Status and ICU Admissions Status variables was visually determined by a model matrix. If Hospital Admissions Status and ICU Admissions status show a multicollinear relationship, then ICU Admissions Status will be excluded from the statistical analysis and Hospital Admissions Status will be used.

Variables and Measures:

All COVID-19 case surveillance data was obtained from voluntary reporting by local, state, and territorial jurisdictions to the CDC using the COVID-19 case report form (Center for Disease Control and Prevention).

All variables were available as categorical variables and did not need to be recoded for this analysis.

The original dataset was sub-grouped to include only pediatric patients ages <= 19 Years. Therefore, the age categories used to sub-group the data into pediatric patients (ages =< 19) are "0 – 9 Years" and "10 – 19 Years".

The variable "Sex assigned at Birth" is the primary variable of interest in this study. Jurisdictions provided this information voluntarily to the CDC via their COVID-19 case report forms.

The variable "Death Status" is the outcome variable of interest in this study. Jurisdictions provided this information voluntarily to the CDC via their COVID-19 case report forms.

The prevalence of childhood comorbidities is not evenly distributed across the sexes. According to Shah et al., the prevalence of obesity is higher in males ages 5 – 19 than in females within the same age range in wealthy countries (Shah, 2020). Additionally, the presence of comorbidities in COVID-19 infected pediatric patients is more likely to result in more severe COVID-19 disease severity and/or mortality when compared to pediatric patients without comorbidities (Tsankov et al., 2021). Therefore, the Comorbidity Status variable is a mediator in this analysis.

The variable "Race/Ethnicity" is an potential effect measure modifier in this analysis because it serves as a proxy for historical, socio-economic issues have disproportionately contributed to health disparities for people of color in the United States, and that ultimately make people of color more susceptible to COVID-19 infection, severe illness, and/or mortality (Price-Haywood, Burton, Fort, & Seoane, 2020).

The Hospitalization Status variable is a mediator in this statistical analysis because the male sex has been shown to be independently associated with higher rates of hospitalization and ICU admission (Gomez & Du-Fay-de-Lavallaz, 2021). Additionally, hospital admissions is a proxy for physiological conditions (i.e., percentage of lymphocytes, lactate level, labored breathing, low oxygen levels, etc.) that require hospitalization due to the subsequent increase COVID-19 disease severity (Wang, 2021).

Statistical Analysis:

All analysis were performed using R version 4.3.

Exploratory analysis of the missing values was performed on Sex assigned at Birth, Race/Ethnicity, Hospital Admissions Status, ICU Admissions Status, Comorbidity Status, Death Status, and Age Group variables to determine the proportion of missing values.

Missing data analysis was performed to determine the mechanism of missingness: Missing Completely at Random (MCAR), Missing at Random (MAR), Missing Not at Random (MNAR). Variables whose missing values are MCAR are independent variables whose pattern of missingness is completely random: the likelihood of a missing value being missing is the same as a missing value having an assigned value (Harrison). Variables whose missing data is MNAR are

variables whose missing data have a relationship with another variable and whose true values are missing at random (Harrison).Variables whose missing data is MNAR are variables whose missing data have a relationship with another variable and whose missing data values are not missing completely at random.

Comparisons between the values of the non-missing observations and the missing observations on the dependent variable, “death status”, given the following explanatory variables were performed: comorbidity presence, age group, sex assigned at birth, race/ethnicity, hospital admissions status.

Confirmation of the MCAR mechanism of missingness was determined for the categorical data via chi-squared tests. If the relationship between the missing data of the explanatory and dependent variables is significant at alpha = 0.05, then MCAR can be excluded as a mechanism of missingness for that variable. Missing data with MAR and/or MNAR as a mechanism of missingness require collecting and examining a sample of the missing data, and are therefore, unable to be determined due to the data collection methods done by the CDC (Harrison).

If exploratory analysis of the missing data shows that the proportion of missing data for any of the variables, especially the dependent variable, is too large (at least 40% missing data) and/or if MAR assumption that the missing values of a variable can be influenced by the other variables is not substantively verifiable, then multiple imputation will be ruled out as a solution for the missing data (Jakobsen, 2017) . Ultimately, if MCAR is confirmed to not be the mechanism of missingness for a variable, then a complete-case statistical analysis will be performed because of the inability to determine MAR and MNAR due to data collection limitations.

The descriptive statistical analysis was performed on age group, sex assigned at birth, the comorbidity status, hospital admissions status, and race/ethnicity stratified on the outcome variable, death status. This analysis was limited to laboratory-confirmed COVID-19 positive pediatric cases ages <= 19 years.

Odds ratios were calculated via simple logistic regression for each demographic exposure variable on the outcome variable (mortality), assessing the significance with alpha = 0.05.

Mediation effects in the causal pathway between the primary exposure variable, Sex assigned at Birth, and the outcome variable, Death Status, are determined using the Baron & Kenny Mediation Model. Simple logistic regressions between the primary exposure variable and the outcome variable were performed. Then simple logistic regressions between the primary exposure variable and the theoretical mediators, Hospitalization Status and the Comorbidity Status, were performed. Lastly, multivariate logistic regression was performed for each mediator to determine the extent of mediation. Statistical significance was assessed at alpha = 0.05.

Multivariate Logistic regression was used to examine the associations between mortality from COVID-19 and demographic characteristics: sex assigned at birth, race/ethnicity, hospital admission status, comorbidity status. Statistical significance was assessed at alpha = 0.05

Interaction between the primary exposure variable, Sex assigned at Birth, and the potential effect measure modifier, Race/Ethnicity, was conducted. If shown to be significant, a multivariate logistic regression stratified by Race/Ethnicity will be conducted. Statistical significance was assessed at alpha = 0.05.

Results

Missing Data Analysis

Table 0a showed the results of exploratory missing data analysis. ICU Admissions Status (84.3%) and Comorbidity Status (81.5%) had the highest percentages of missing data; whereas, the Age Group variable for patients within the "0 – 9 Years" or "10 – 19 Year" age bracket had no missing data. A small quantity of missing data existed for both the "Sex assigned at Birth" (0.3%) and "Race/Ethnicity" (4%) variables.

Table 0b results showed statistically significant p-values (<0.001) for all explanatory variables – indicating that the missing data for all explanatory variables are not missing completely at random (MCAR) in relation to the dependent variable. Thus, the mechanism of missingness present in the COVID-19 Case Surveillance Public Use Data dataset is either Missing at Random (MAR) or Missing Not at Random (MNAR). Differentiation between MNAR and MAR is not possible because it requires knowledge of the missing variable values for the missing variables. Additionally, due to the voluntary nature of COVID-19 case reporting by state and local jurisdictions to the CDC, it is impossible to assume the mechanism of missingness as either MAR or MNAR. Multiple imputation was excluded as a possible solution for the missing data because of the large proportion of missing data for the outcome variable, Death Status, Hospital Admissions variable, and the Comorbidity Status variable (Table 0a). Lastly, substantive circumstantial associations between covariates and missing data were unable to be provided because of the voluntary data collection methods used by the CDC to create this surveillance dataset. Ultimately, a complete case analysis was performed because of the inability to determine the mechanism of missingness.

Multicollinearity Analysis

The visual analysis of heatmap model matrix showed an extremely similar relationship between "Yes" for Hospital Admission Status and "Yes" for ICU Admissions Status (Figure 1). For this reason, in addition to the higher percentage of missing observations within the ICU Admissions Status variable (84.3%), ICU Admission Status was excluded from the statistical model and only Hospitalization Status was used (Figure 1).

Descriptive Statistics

Table 1 shows the demographic and health status breakdown of the voluntarily reported COVID-19 infected individuals as of February 2021. A total of 742,461individual-level laboratory-confirmed COVID-19 positive pediatric patients were included in the analysis, with 28.9% of individuals between the ages of 0 and 9 years and approximately 71.1% being between 10 and 19 years of age; the distribution between females and males within this age bracket was relatively even with approximately 50.5% of females and 48.5% of males between the ages of 0 years and 19 years. The most common race/ethnicity presented in the data were "Whites" (27.7%). Hispanic/Latino" was the second most common reported race/ethnicity at approximately 12.6%

of overall observations; "Blacks, Non-Hispanic" was the third most common race/ethnicity comprising of approximately 6.2% of all individuals between the ages 0 and 19 years.

Both hospital admissions and the presence of a comorbidity were extremely low within the subgroup of COVID-19 positive pediatric patients at approximately 0.99% and 0.001%, respectively.

Logistic Regression Models
Table 2a shows the results of simple linear regression for the independent variables of interest to model their association between the death status outcome.

In the simple logistic regression, male pediatric patients (ages <= 19 years) were shown to not have statistically significant difference in the odds of mortality when compared to female pediatric patients (OR = 0.99, 95% CI 1 – 1.01). COVID-19 positive pediatric patients who identified as non-Hispanic Black (OR = 0.59, 95% CI 0.57 – 0.6), Asian (OR = 0.45, 95% CI 0.43 – 0.5), Native Hawaiian/Pacific Islander, Hispanic/Latino (OR = 0.99, 95% CI 0.92 – 1.1), or multiple race/ethnicities (OR = 0.37, 95% CI 0.36– 0.4) all showed statistically significant higher odds of mortality. Additionally, both hospital admission and the presence of comorbidities resulted in statistically significant higher odds of mortality (Table 2a).

In the multivariable model, after adjusting for hypothesized mediators, there was no significant association between male pediatric patients and mortality when compared to female patients (OR = 0.993, 95% CI 0.964 - 1.023). Additionally, the mediators “Comorbidity Status” and “Hospitalization Admissions Status” were not shown to have a mediating effect on “Sex assigned at birth” as the primary exposure variable because there was no change between the crude and adjusted odds ratio estimates derived in the simple and multivariate logistic regression models (Table 2b).

Multivariate regression analysis of the interaction between the effect measure modifier, Race/Ethnicity, and the primary exposure of interest, Sex assigned at Birth, showed no significant interaction between Sex and Race/Ethnicity.

Discussion

Reports have shown that there is a greater prevalence of males for COVID-19 hospitalizations, mortality, and morbidity than females (Gomez & Du-Fay-de-Lavallaz, 2021; Mohamed et al., 2021).Both the crude odds ratio (OR = 0.99, 95% CI 1 – 1.007) and the adjusted odds ratio (OR = 0.993, 95% CI 0.964 - 1.023) suggest that sex differences in pediatric patients aged 19 and below do not significantly contribute to the likelihood of an individual dying from COVID-19 infection: COVID-19 positive males ages =< 19 years do not have statistically significant greater odds of death when compared to females. This result is not consistent with the current literature regarding the ways that sex differences in COVID-19 infected individuals, most likely due to bias resulting from the missing observations within the COVID-19 Case Surveillance Public Use Data dataset. The nearly equal percentages of missing data between males (39.6%) and females (39.3%) in the missing mortality data and the subsequent nearly equal distribution of mortality for males and females (Table 1) resulted in non-differential misclassification of the health

outcome; Ultimately, given the incompleteness of the CDC dataset, as shown in the results, non-differential bias of the outcome may be present within the statistical analysis because of the omission of all missing observations from the study based on the outcome variable, Death Status; thus, resulting in study results biased towards the null – suggesting that there is no difference in the odds of mortality between COVID-19 positive male and female pediatric patients and contradicting emerging literature.

Data Limitations

The COVID-19 Case Surveillance Public Use Data is the aggregation of individual-level patient data voluntarily provided by local and state jurisdictions as well as United States controlled territories. Therefore, the consistency with which accurate data is recorded and shared with the CDC is depended on the guidelines of each jurisdiction; thus, resulting in an incomplete dataset. An initial missing data analysis was performed on the raw dataset to attempt to determine the mechanism of missingness. Comparisons between the values of the non-missing observations and the missing observations on the dependent variable, "death status", given the following explanatory variables were performed: comorbidity presence, age group, sex, race/ethnicity, hospital admissions status. Results showed statistically significant p-values (<0.05) for all explanatory variables – indicating that the missing data for all explanatory variables are not missing completely at random (MCAR) in relation to the dependent variable. Thus, the mechanism of missingness present in the COVID-19 Case Surveillance Public Use Data dataset is either Missing at Random (MAR) or Missing Not at Random (MNAR). However, due to the voluntary submission of COVID-19 case data by local, state, and territorial jurisdictions, we are unable to determine whether the mechanism of missingness is MNAR or MAR.

Additionally, the variable for comorbidities is aggregated into a dichotomous variable; therefore, we are unable to stratify our analysis by specific comorbidities and, thus, cannot control for the varying mediation effects that specific comorbidities may have on COVID-19 pediatric patients depending on sex assigned at birth. For example, asthma is one of the most common childhood comorbidities and was once posited to be a serious risk factor for COVID-19 disease severity and mortality in pediatric patients (Chatziparasidis & Kantar, 2021); however, current studies are revealing that phenotypical allergic asthma is no longer considered to be a major risk factor in pediatric patients , but is considered a risk factor in elderly adults (Chatziparasidis & Kantar, 2021); therefore, phenotypical allergic asthma would not be considered as a confounder for an analysis of the association between sex assigned at birth and mortality in COVID-19 pediatric patients, and would be excluded from the analysis. Obesity, however, increases risk for hospitalization and mortality for COVID-19 infected patients; therefore, it may be more appropriate to consider obesity as a theoretical mediator for COVID-19 disease severity within pediatric patients as oppose to a confounder (Huang et al., 2020). Ultimately the polychotomization of the comorbidity variable would allow for a more accurate analysis of the affects those individual comorbidities may have on male and female COVID-19 infected pediatric patients that is reflective of the most recent findings in the literature.

Conclusion

The results of this study contradict emerging literature by suggesting that there is no statistically significant difference in the odds of mortality between male and female pediatric patients with positive-COVID-19 diagnosis. Despite these results, however, the public health implications of this study are that for ongoing, federally mandated, reportable communicable illness, protocols must be established to allow for adequate systematic individual-level data collection and reporting. Given the high prevalence of missing data and the aggregation of unique patient data into categories, future research regarding sex differences in COVID-19 mortality for pediatric patients should focus on data collection and analysis of specific comorbidities as well as varying degrees of COVID-19 severity. Lastly, if sufficient future information on data collection and sharing methodology is available to justify MAR assumptions, future research endeavors should include multiple imputation analysis of the COVID-19 Case Surveillance Public Use Data in tangent with the inclusion of a polychotomized comorbidity variable to confirm the direction of non-differential bias in this study as well as the unbiased odds of mortality between the sexes in COVID-19 positive pediatric patients.

Tables and Figures

Table 0a: Proportion of Missing Data for Independent and Dependent Variables

Variable	N	N (missing)	Missing Percentage
Sex assigned at Birth	1226362	3245	0.3
Race/Ethnicity	1181007	48600	4
Hospital Admission Status	681038	548569	44.6
Comorbidity Status	227763	1001844	81.5
Death Status	742461	487146	39.6
Age group	1229607	0	0
ICU Admission Status	193067	1036540	84.3

Table 0b: Missing data analysis of independent variables on dependent variable, Death Status

Variable		Not missing	Missing (%)	p-value
Age Group	0 - 9 Years	214873 (60.8)	138596 (39.2)	<0.001
	10 - 19 Years	527588 (60.2)	348550 (39.8)	
Sex assigned at Birth	Female	374801 (60.7)	242400 (39.3)	<0.001
	Male	360097 (60.4)	236399 (39.6)	
	Unknown	673026 (53.2)	5923 (46.8)	
Race/Ethnicity	American Indian/Alaska Native, Non-Hispanic	5925 (47.7)	6499 (52.3)	
	Asian, Non-Hispanic	14659 (68.2)	6823 (31.8)	<0.001

	Black, Non-Hispanic	45938 (59.6)	31163 (40.4)	
	Hispanic/Latino	93228 (55.0)	76258 (45.0)	
	Multiple/Other, Non-Hispanic	41706 (73.1)	15372 (26.9)	
	Native Hawaiian/Other Pacific Islander, Non-Hispanic	1967 (67.7)	940 (32.3)	
	Unknown	328765 (64.9)	177940 (35.1)	
	White, Non-Hispanic	205442 (61.5)	128382 (38.5)	
Comorbidity Status	No	79481 (82.0)	17478 (18.0)	<0.001
	Unknown	75053 (79.6)	19182 (20.4)	
	Yes	28204 (77.1)	8365 (22.9)	
Hospital Admission Status	No	385041 (81.6)	86944 (18.4)	<0.001
	Unknown	191918 (96.4)	7225 (3.6)	
	Yes	7356 (74.4)	2533 (25.6)	
ICU Admission Status	No	68309 (84.8)	12234 (15.2)	<0.001
	Unknown	111450 (99.6)	460 (0.4)	
	Yes	547 (89.1)	67 (10.9)	

Table 1: Socio-demographic characteristics of Laboratory-Confirmed OCIVD-19 Positive Pediatric (ages =< 19 Years) patients, CDC COVID-19 Case Surveillance Public Use Data

			Mortality Status						
	Total		**No**		**Yes**		**Unknown**		**p-value**
	742461	n(%)	638,002	n(%)	190	n(%)	104,269	n(%)	
Variable									
Age Group									<0.001
0 - 9 Years	214873	28.941	189099	29.6	55	28.9	25719	24.7	
10 - 19 Years	527588	71.059	448903	70.4	135	71.1	78550	75.3	
Sex assigned at Birth									<0.001
Male	360097	48.5	309390	48.6	96	50.8	50611	48.6	
Female	374801	50.481	321723	50.5	93	49.2	52985	50.8	
Unknown	6726	0.9059	6120	1	0	0	606	0.6	
Race/Ethnicity									<0.001
White, Non-Hispanic	205442	27.67	174576	27.6	67	35.6	30799	29.6	
Black, Non-Hispanic	45938	6.1873	41616	6.6	30	16	4292	4.1	
Asian, Non-Hispanic	14659	1.9744	13571	2.1	1	0.5	1087	1	
Hispanic/Latino	93228	12.557	77661	12.3	43	22.9	15524	14.9	
American Indian/Alaska Native, Non-Hispanic	5925	0.798	5045	0.8	2	1.1	878	0.8	
Native Hawaiian/Other Pacific Islander, Non-Hispanic	1967	0.2649	1773	0.3	0	0	194	0.2	
Multiple/Other, Non-Hispanic	41706	5.6173	39118	6.2	7	3.7	2581	2.5	
Unknown	328765	44.28	280155	44.2	38	20.2	48572	46.7	
Hospital Admission Status									<0.001
No	385041	51.86	361081	73.8	74	45.4	23886	25.1	
Yes	7356	0.9908	6326	1.3	79	48.5	951	1	

Unknown	191918	25.849	121555	25	10	6.1	70353	73.9	
Comorbidity Status									<0.001
No	79481	10.705	77139	53.2	43	47.8	2299	6.1	
Yes	28204	3.7987	26414	18.2	37	41.1	1753	4.6	
Unknown	75053	10.109	41382	29	10	11.1	33661	89.3	

Table 2a: Crude associations of Mortality, Sex Assigned at Birth, and Socio-demographic characteristics, CDC COVID-19 Case Surveillance Public Use Data

		Simple Logistic Regression			
Variable	**Estimate (B)**	**Crude Odds Ratio**	**95% Wald Confidence Interval**		**p-value**
Sex assigned at Birth					
Male	-0.01	0.99	1	1.01	0.3
Unknown	-0.51	0.60	0.6	0.652	<0.001
Female	Reference Group				
Race/Ethnicity					
Black, Non-Hispanic	-0.53	0.59	0.6	0.61	<0.001
Asian, Non-Hispanic	-0.79	0.45	0.4	0.48	<0.001
Hispanic/Latino	0.126	1.13	1.1	1.16	<0.001
American Indian/Alaska Native, Non-Hispanic	-0.01	0.99	0.9	1.06	0.7
Native Hawaiian/Other Pacific Islander, Non-Hispanic	-0.48	0.62	0.5	0.72	<0.001
Multiple/Other, Non-Hispanic	-0.98	0.37	0.4	0.39	<0.001
Unknown	-0.02	0.98	1	0.99	0.017

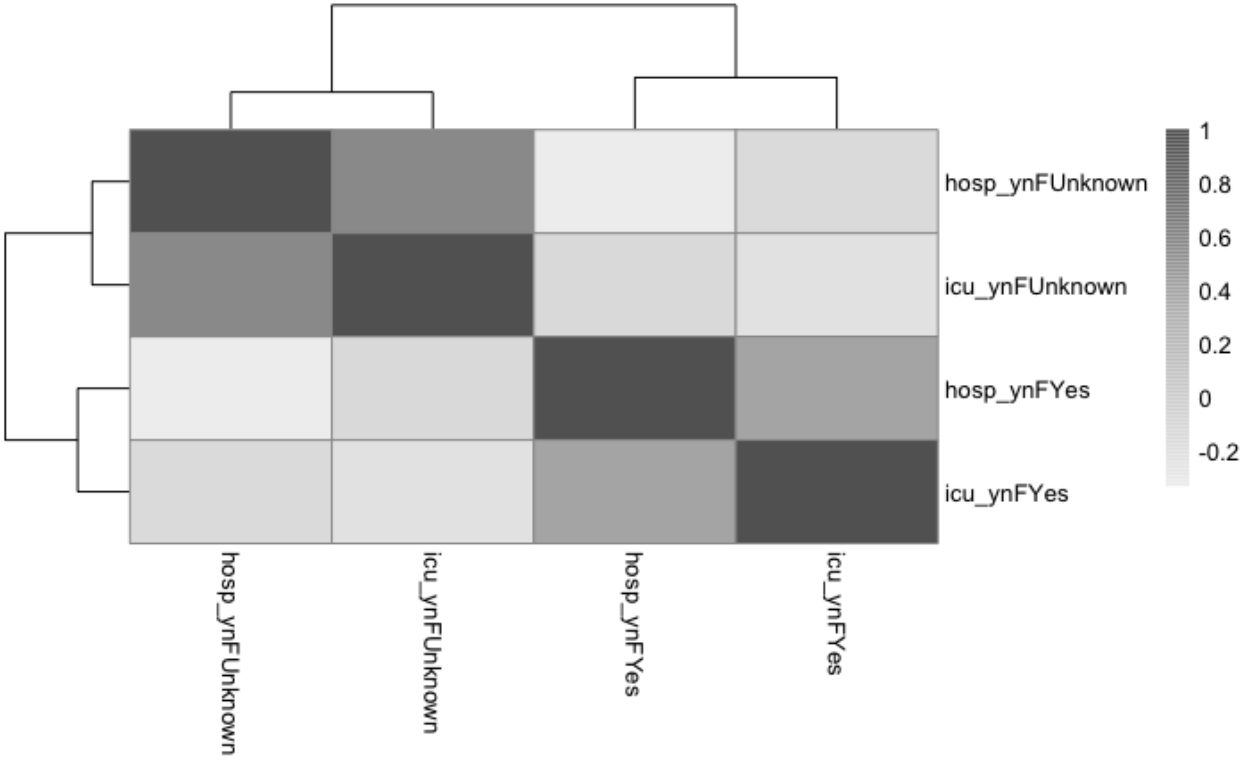
White, Non-Hispanic	Reference Group				
Hospital Admission Status					
Yes	0.898	2.45	2.3	2.6	<0.001
Unknown	2.166	8.72	8.6	8.87	<0.001
No	Reference Group				
Comorbidity Status					
Yes	0.803	2.23	2.1	2.377	<0.001
Unknown	3.288	26.80	26	28	<0.001
No	Reference Group				

Table 2b: Associations of Mortality and Sex Assigned at Birth adjusting for mediators, CDC COVID-19 Case Surveillance Public Use Data

		Multivariate Logistic Regression			
Variable	**Estimate (B)**	**Adjusted Odds Ratio**	**95% Wald Confidence Interval**		**p-value**
Sex assigned at Birth					
Male	-0.007	0.993	0.964	1.023	0.6355
Unknown	1.378019	3.967	2.973	5.328	<0.001
Female	Reference Group				
Race/Ethnicity					
Black, Non-Hispanic	-0.474858	0.622	0.575	0.673	<0.001
Asian, Non-Hispanic	-0.171714	0.842	0.700	1.008	0.0643
Hispanic/Latino	0.183926	1.201	1.142	1.265	<0.001

American Indian/Alaska Native, Non-Hispanic	-0.008909	0.991	0.705	1.371	0.9581
Native Hawaiian/Other Pacific Islander, Non-Hispanic	0.022689	1.023	0.644	1.549	0.9189
Multiple/Other, Non-Hispanic	0.591165	1.806	1.647	1.980	<0.001
Unknown	0.88752	2.4195	2.254	2.597	<0.001
White, Non-Hispanic	Reference Group				
Hospital Admission Status					
Yes	1.974324	7.202	6.452	8.0263	<0.001
Unknown	3.362423	28.86	27.182	30.657	<0.001
No	Reference Group				
Comorbidity Status					
Yes	0.8836	2.42	2.254	2.597	<0.001
Unknown	0.353202	1.424	1.336	1.517	<0.001
No	Reference Group				

Figure 1: Heatmap of the model matrix to determine the extent of multicollinearity Between Hospital Admissions and ICU Admissions variables



Bibliography

Center for Disease Control and Prevention. Human Infection with 2019 Novel Coronavirus
Case Report Form. Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf

Chatziparasidis, G., & Kantar, A. (2021). COVID-19 in Children with Asthma. *Lung, 199*(1), 7-12. doi:10.1007/s00408-021-00419-9

COVID-19 Case Surveillance Public Use Data. (2021). Retrieved from https://healthdata.gov/dataset/COVID-19-Case-Surveillance-Public-Use-Data/knt4-7efa

Dhochak, N., Singhal, T., Kabra, S. K., & Lodha, R. (2020). Pathophysiology of COVID-19: Why Children Fare Better than Adults? *Indian J Pediatr, 87*(7), 537-546. doi:10.1007/s12098-020-03322-y

Gomez, J. M. D., & Du-Fay-de-Lavallaz, J. M. (2021). Sex Differences in COVID-19 Hospitalization and Mortality. *Journal of Women's Health*. Retrieved from liebertpub.com/doi/10.1089/jwh.2020.8948#:~:text=Results%3A%20There%20were%208108%20positive,vasopressor%20support%20(4%25%20vs.

Gotzinger, F., Santiago-Garcia, B., Noguera-Julian, A., Lanaspa, M., Lancella, L., Calo Carducci, F. I., . . . ptbnet, C.-S. G. (2020). COVID-19 in children and adolescents in Europe: a multinational, multicentre cohort study. *Lancet Child Adolesc Health, 4*(9), 653-661. doi:10.1016/S2352-4642(20)30177-2

Harrison, E. Missing data. Retrieved from https://cran.r-project.org/web/packages/finalfit/vignettes/missing.html

Huang, Y., Lu, Y., Huang, Y. M., Wang, M., Ling, W., Sui, Y., & Zhao, H. L. (2020). Obesity in patients with COVID-19: a systematic review and meta-analysis. *Metabolism, 113*, 154378. doi:10.1016/j.metabol.2020.154378

Jakobsen, J. C. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology*. Retrieved from https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0442-1#Sec7

Leidman, E. (2021). COVID-19 Trends Among Persons Aged 0–24 Years — United States, March 1–December 12, 2020. Retrieved from https://www.cdc.gov/mmwr/volumes/70/wr/mm7003e1.htm

Ludvigsson, J. F. (2020). Systematic review of COVID-19 in children shows milder cases and a better prognosis than adults. *Acta Paediatr, 109*(6), 1088-1095. doi:10.1111/apa.15270

Mohamed, M. S., Moulin, T. C., & Schioth, H. B. (2021). Sex differences in COVID-19: the role of androgens in disease severity and progression. *Endocrine, 71*(1), 3-8. doi:10.1007/s12020-020-02536-6

Price-Haywood, E. G., Burton, J., Fort, D., & Seoane, L. (2020). Hospitalization and Mortality among Black Patients and White Patients with Covid-19. *N Engl J Med, 382*(26), 2534-2543. doi:10.1056/NEJMsa2011686

Shah, B. (2020). Sex and gender differences in childhood obesity: contributing to the research agenda. *BMJ Nutrition, Prevention & Health*. Retrieved from https://nutrition.bmj.com/content/early/2020/09/07/bmjnph-2020-000074

Tsankov, B. K., Allaire, J. M., Irvine, M. A., Lopez, A. A., Sauve, L. J., Vallance, B. A., & Jacobson, K. (2021). Severe COVID-19 Infection and Pediatric Comorbidities: A Systematic Review and Meta-Analysis. *Int J Infect Dis, 103*, 246-256. doi:10.1016/j.ijid.2020.11.163

Wang, Z. (2021). Identification of risk factors for in-hospital death of COVID - 19 pneumonia -- lessions from the early outbreak. Retrieved from https://pubmed.ncbi.nlm.nih.gov/33494706/

Zhang, L., Peres, T. G., Silva, M. V. F., & Camargos, P. (2020). What we know so far about Coronavirus Disease 2019 in children: A meta-analysis of 551 laboratory-confirmed cases. *Pediatr Pulmonol, 55*(8), 2115-2127. doi:10.1002/ppul.24869