1-24-2024

Social Determinants of Health, Comorbidities, and COVID-19 Among Adults Aged 18 and Older in Indiana

kandeh kamara
Walden University

Walden University

College of Health Sciences and Public Policy

This is to certify that the doctoral study by

Kandeh Kamara

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Nancy Rea, Committee Chairperson, Public Health Faculty
Dr. Patrick Dunn, Committee Member, Public Health Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2024

Abstract

Social Determinants of Health, Comorbidities, and COVID-19 Among Adults Aged 18 and Older in Indiana

by

Kandeh Kamara

PhD, Walden University, 2019

MPH, Indiana University School of Medicine, 2008

MS, Ball State University, 2003

BA, Indiana University Purdue University Indianapolis, 1993

BS, Indiana University Purdue University Indianapolis, 1992

Doctoral Study Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Public Health

Walden University

February 2024

Abstract

The SARS-CoV-2 coronavirus, also known as the COVID-19 pandemic, continues to be a major global health challenge and is responsible for about 762,201,169 confirmed cases of COVID-19, including 6,893,190 deaths globally. Social determinants of health and preexisting comorbidities increase the burden of COVID-19 among vulnerable adult populations. However, this population has limited empirical knowledge regarding the relationship between social determinants of health and preexisting comorbidities and COVID-19 infection. The purpose of this quantitative study was to examine the association between social determinants of health (exposure, county code of residence), preexisting comorbidities (e.g., obesity), and COVID-19 diagnosis, after controlling for age, sex, and race. This cross-sectional study, guided by the social-ecological model, provided an update on COVID-19 infection and its associated factors among adults ages 18 years and older in Indiana. Descriptive statistics, chi-squared tests for associations, and backward elimination and hierarchal logistic regression tests were used to identify associated factors after adjusting for other risk variables, using secondary data collected from the Centers for Disease Control and Prevention. Results indicated a significant association between age group, race, exposure, county code of residence, underlying medical conditions, and COVID-19 infection status. Multivariate analysis indicated that the underlying medical conditions predictor most uniquely contributed statistically to the model ($\beta = 0.059$, $p < .001$). Results may be used to target and promote awareness of, mainly, underlying medical conditions-related COVID-19 infection risk in attempts to reduce the burden of COVID-19 in adults in Indiana and prevent future outbreaks.

Social Determinants of Health, Comorbidities, and COVID-19
Among Adults Aged 18 and Older in Indiana

by

Kandeh Kamara

PhD, Walden University, 2019

MPH, Indiana University School of Medicine, 2008

MS, Ball State University, 2003

BA, Indiana University Purdue University Indianapolis, 1993

BS, Indiana University Purdue University Indianapolis, 1992

Doctoral Study Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Public Health

Walden University

February 2024

Dedication

I would like to dedicate this dissertation first and foremost to God Almighty, creator of Heavens and Earth and all that is in them, for allowing me to start and complete this remarkable journey. To my late grandmother, Ya. Boinkey Conteh, and my late grandfather, Pa. Kaprie Kamara, who raised me in a God-fearing manner and put me through primary and secondary school in Sierra Leone. To my outstanding father, Mr. Frank B. Kamara, and my outstanding mother, Ms. Susan P. Turay, who instilled in me the value of education during my childhood, making me a better person in life, and helping me accomplish my goal of writing this dissertation.

Acknowledgments

I would like to thank my chair, Dr. Nancy K. Rea, for her tremendous mentorship, comments, suggestions, and substantial guidance and support throughout this dissertation writing process. I would also like to express my profound gratitude to my committee member, Dr. Patrick Dunn, and my university research reviewer, Dr. German A. Gonzalez, for their valuable time, constructive feedback, and continuous encouragement during the dissertation.

I wish to acknowledge and sincerely thank my CDC Surveillance Review and Response Group (SRRG), Data, Analytics & Visualization Task Force | COVID-19 Emergency Response Centers for Disease Control and Prevention for allowing me to access the COVID-19 Case Surveillance Public Use Data with Geography data set for Indiana for my research. I am greatly indebted to them.

I give special thanks to my brother, Francis Kamara, and my sisters, Florence Kamara, Marian Kamara, and Victoria Crosby, for their remarkable love, support, and inspiration during this DrPH journey at Walden University. Finally, I am also grateful to my daughter, Kiara Kamara, for the valuable support she gave me during this dissertation writing. The highest honor goes to God almighty for his everlasting love and mercy to me and my family.

Table of Contents

List of Tables

List of Figures

Section 1: Foundation of the Study and Literature Review

Introduction

The ongoing Coronavirus Disease 2019 (COVID-19) pandemic is a pressing public health issue of tremendous magnitude (Abebe et al., 2020; Centers for Disease Control and Prevention [CDC], 2022; Iezadi et al., 2021; World Health Organization [WHO], 2021). A recent WHO report showed that there were about 762,201,169 confirmed cases of the disease and had been 6,893,190 deaths worldwide as of April 6, 2023 (WHO, 2023). Despite the improved and increasing distribution of the COVID-19 vaccine, the United States still bears the biggest burden of COVID-19, accounting for about 104,242,889 cases, 1,127,104 deaths, and 120,820 new cases as of April 9, 2023 (Abebe et al., 2020; CDC, 2023; Iezadi et al., 2021; WHO, 2022).

The COVID-19 disease is a disease that affects the upper and lower airways and causes severe damage to the lungs and vital organs (WHO, 2021). The COVID-19 disease emerged from poorly managed seafood and animal markets in Wuhan, China at the end of December 2019 and spread quickly around the world (Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; Roberts Kennedy, 2021; WHO, 2021). The WHO declared it a pandemic on March 11, 2020 (Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; Roberts Kennedy, 2021; WHO, 2021).

According to Abebe et al. (2020), Sun et al. (2020), and WHO (2020), the COVID-19 is caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) and is capable of causing an overall case-fatality rate of about 2.3%. Coronaviruses are capable of infecting humans, farm animals, birds, bats, mice, and many other wild animals (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; WHO, 2020).

The first known coronavirus disease infection, referred to as severe acute respiratory syndrome (SARS), emerged from China in 2002 and was responsible for 8,098 infections and 774 deaths spanning 29 countries. The reservoirs or hosts for the SARS-CoV viruses are believed to be bats and musk cats before the virus crossed over to human beings (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020). The second coronavirus, the Middle East respiratory syndrome (MERS), first emerged in Saudi Arabia in 2012 and resulted in 2,494 cases and 858 deaths spanning 27 countries (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020). The reservoirs for the MERS virus are believed to be bat and camel. The third and current coronavirus -nCoV (CoV) 2019, which caused the COVID-19 pandemic, a catastrophic pandemic with a case fatality rate of about 2.3%, is zoonotic and believed to emerge from bats (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020).

According to Abebe et al. (2020), Aldila et al. (2021), Gralinski and Menachery (2020), Hatipoğlu (2020), and WHO (2021), the COVID-19 disease is characterized by several symptoms such as dry cough, high fever, shortness of breath, and multiple organ failure. Some individuals may not exhibit any symptoms (Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020). Reports show that the incubation period of COVID-19 usually ranges from 3–14 days, and the disease is diagnosed through nasopharyngeal swab testing, blood samples, chest computed tomography (CT), and nucleic acid test, performed using the fast-yielding and reliable polymerase chain reaction (PCR) method (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020). Furthermore, Abebe et al., Aldila et al., Gralinski and Menachery, Hatipoğlu, and WHO (2001) noted that the COVID-19 virus is usually transmitted from person to person via contaminated airborne droplets of asymptomatic and symptomatic patients and other contaminated materials.

Thus, strategies such as contact tracing; 14-day social isolation; use of personal protective gear; good hygiene; treatments such as plasma exchange; ventilation support; use of disinfectants containing 75% ethanol, Lysol, and chlorine bleach; wearing protective face masks; and getting the COVID-19 vaccine are best means to combat the pandemic (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; WHO, 2021). Other proven control and prevention methods entail the utilization of community health programs to foster proper COVID-19 preventive health behaviors such as getting the COVID-19 vaccine (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; WHO, 2021).

Social determinants of health, which entail the overall conditions into which people are born, grow, live, work, and grow old (CDC, 2022; Salgado de Snyder et al., 2021; Singu et al., 2020), and preexisting comorbidities have been implicated in COVID-19 infection (CDC, 2021).

Although these studies have provided useful epidemiological information on COVID-19 characteristics and the potential social determinants of health and preexisting medical conditions in the general population, there are scarce data on the relationships between predictors, risk factors, and COVID-19 among men and women ages 18 and older in Indiana. This serious health problem continues to challenge researchers, health policy makers, health educators, clinicians, and public health experts, demonstrating a continued need to evaluate COVID-19 incidence rates in this setting and population. This study addressed this association in this population to clarify the relationship between factors associated with COVID-19 incidence and guide and improve COVID-19 prevention measures in the local communities in Indiana (WHO, 2021).

Hence, the results of the current study will promote positive social change by providing vital public health data and information for key populations that will effectively inform COVID-19 health policy and mitigation efforts in Indiana and elsewhere (Samuel et al., 2021). Moreover, a clear understanding of COVID-19 characteristics can foster effective COVID-19 health program planning and COVID-19 vaccine distribution in local communities in Indiana (Samuel et al., 2021), in addition to improving people’s knowledge and preventive behaviors towards COVID-19 and its related risk factors (Samuel et al., 2021).

This section of the study presents the background of the study. Also, I describe the dataset of laboratory-confirmed cases in the COVID-19 Case Surveillance Public Use Data With Geography (from the CDC), which contain data and independent variables such as age group from the current COVID-19 pandemic, and the dependent variable, COVID-19 diagnosis, to examine how social determinants of health (age group, sex, race, exposure, county code of residence) and underlying medical conditions (e.g., obesity) collectively influence COVID-19 infection incidence among adults aged 18 years and older in the state of Indiana (CDC, 2023; Indiana Department of Health [IDH], 2021). Furthermore, I highlight the problem statement, purpose of the study, research questions and associated hypotheses, nature of the study, and significance of the current study. I also present the theoretical framework that guides this study, and then I define the main terms utilized in this study and address the literature review strategy on the main variables used in this study.

Background

The COVID-19 pandemic burden remains very high worldwide. About 762,201,169 people have been infected, and about 6,893,190 have died globally from the disease (WHO, 2023).

Several reports including those from the CDC, the Commission on Social Determinants of Health (CSDH), and the WHO on social determinants of health have identified and established connection between factors such as age group, sex, exposure, race, and county code of residence and health inequities for diseases such as COVID-19 (CDC, 2021; CSDH, 2008; Roberts Kennedy, 2021; WHO, 2021). For example, Maroko et al.'s (2020) ecological study out of New York and Illinois showed that areas with higher COVID-19 incidence rates in New York City ZIP code/neighborhoods were characterized by low education and up to 60% non-Hispanic White population. Moreover, in Chicago, high COVID-19 areas were characterized by low income, high poverty rate, high unemployment rate, and high concentration of non-Hispanic Black residents (Maroko et al., 2020). Epidemiologic studies have also established associations between underlying medical conditions such as obesity (as measured by BMI > 30 kg/m^2

), cancer, and diabetes and COVID-19 (Hege et al., 2022; Roberts Kennedy, 2021). For example, Liu et al.'s (2020) systematic review and meta-analysis study of 24 studies that included 10,948 COVID-19 patients from China, the United States, and other countries indicated a strong connection between preexisting chronic conditions and COVID-19 severity (OR 3.50), including intensive care unit admitting (OR 3.36). Moreover, diabetes was present in 10.0% and hypertension was present in 20.0% of the study population.

A number of epidemiologic studies show age disparities in the COVID-19 burden with adult groups that are prone to the COVID-19 virus and its severity and death compared to younger individuals (Abebe et al., 2020; CDC, 2022; Dixon et al., 2021; Hamza et al., 2022; Timsina et al., 2020). For example, in a longitudinal study of 61,993 confirmed cases of COVID-19 in Spain, Casas-Deza et al. (2021) pointed out a higher number of COVID-19 cases among individuals ages 40–49 years old. Also, about 43% of study participants were over 50 years old, while 14.9% of participants were 75 years and older. In addition, adults aged 75 years and older accounted for about 46% of hospitalizations. Another COVID-19 study by Kawatake de Sousa et al. (2021) out of Brazil that assessed the association between patient's age and COVID-19 infection and mortality reported the highest incidence rates among patients in the age group 30–39 years (19,639.19 cases per 100,000 people) and 40–49 years (19,098.20 cases per 100,000 people). Older adults are highly vulnerable to contracting COVID-19 virus compared to other groups, likely because of their preexisting comorbidities such as obesity (as measured by body mass index [BMI] > 30 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension and extensive social determinants of health inequities, such as race and county code of residence (Abebe et al., 2020; Ahmad et al., 2020; Casas-Deza et al., 2021; CDC, 2022; Fielding-Miller et al., 2020; Gimeno-Miguel et al., 2021; Hamza et al., 2022; Jacobs et al., 2021; Khanijahani et al., 2021; Mankowski et al., 2021; Unruh et al., 2022). For instance, Parohan et al. (2020) noted that diabetic patients have a higher risk for COVID-19 infection as a result of defective immunity response. Furthermore, the expression of angiotensin-converting enzyme 2 (ACE2) is increased in Type 1 and Type 2 diabetes mellitus, and the coronaviruses bind with target cells through ACE2 (Parohan et al., 2020).

This trend is a serious source of concern for Indiana (Figure 1) because Indiana had one of the highest proportions of COVID-19 infection diagnoses among vulnerable adults in the Midwest region of the United States (1,683,505) as of April 4, 2023 (Indiana Department of Health, 2023; Marion County Public Health Department [MCPHD], 2023; U.S. Bureau of Labor Statistics, 2023), showing the need for characterizing COVID-19 and its risk factors and COVID-19 control and preventive measures for vulnerable older adults in this setting. Moreover, most of the population of Indiana consists of vulnerable adults, who number 5,116,486 (76%; Indiana Department of Health, 2022). Furthermore, the state has a high poverty rate of about 14.1%, an uninsured rate of 9.1%, and a median household income of only about $56,185, and only 27.8% of individuals have a bachelor’s degree or higher (Indiana Department of Health, 2022). Therefore, this subpopulation is paramount in this study. Additionally, recent epidemiologic studies (Cai et al., 2020; Cuomo et al., 2020) have linked Indiana to one of the highest burdens of

HIV infection in 2015 associated with opioid use and drug injection among vulnerable adults, which has devastated the healthcare system.

The current study built on studies that failed to adequately address age-specific social determinants of health factors and preexisting comorbidities factors in the ongoing COVID-19 pandemic (Dixon et al., 2021; Hanson et al., 2020; Samuel et al., 2021; Timsina et al., 2020). Furthermore, the few emerging but scant studies related to the connection between social determinants of health factors and preexisting comorbidities and COVID-19 infection incidence in the adult population that have been done in Indiana, as well as most of those that have been conducted in other areas outside Indiana, have reported inconsistent or conflicting results on the burden of COVID-19 infection and some social determinants of health factors and preexisting comorbidities, and they also lacked a clear theoretical framework to examine associations between these variables and to produce high-quality evidence and clearly understand the particular contributions of these factors to the COVID-19 outcome in this population (Dixon et al., 2021; Hanson et al., 2020; Khanijahani et al., 2021; Samuel et al., 2021; Timsina et al., 2020). Additionally, several of the COVID-19 studies in the region, including in Indiana (Dixon et al., 2021; H anson et al., 2020; Khanijahani et al., 2021; Samuel et al., 2021; Timsina et al., 2020), are broad and encompass people of all ages.

It is of paramount public health interest to know and understand how social determinants of health (age group, sex, race, exposure, county code of residence) and underlying medical conditions (such as obesity) relate to COVID-19 disease, particularly in adults aged 18 years and older in Indiana communities (Dixon et al., 2021; Hanson et al., 2020; Samuel et al., 2021; Timsina et al., 2020). Furthermore, the public health need to determine and characterize vulnerable groups that bear the brunt of the COVID-19 burden is urgent (Dixon et al., 2021; Hanson et al., 2020; Samuel et al., 2021; Timsina et al., 2020). Therefore, this current study was needed. This vital knowledge and information could be utilized by various state and federal entities to effectively monitor COVID-19, test for COVID-19, distribute and promote COVID-19 resources, guide various community COVID-19 strategies, and develop population health strategies (Samuel et al., 2021; Timsina et al., 2020).

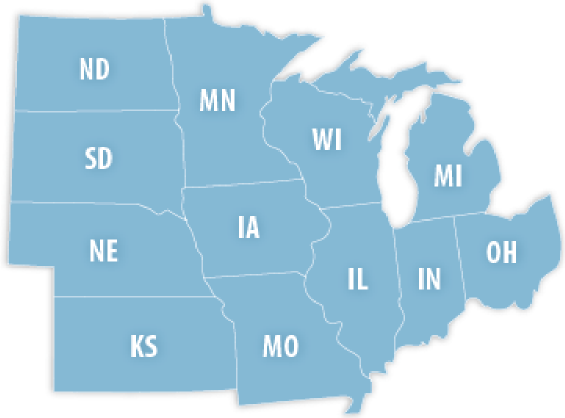
According to Bathina et al. (2021), further research to examine the link between poverty status and COVID-19 disease is warranted. Furthermore, in Brown et al.'s (2022) large rapid review and synthesis of urgent public health recommendations from 338 peer-reviewed studies related to equity and social determinants of health and COVID-19 disease, a substantial number of studies (50) recommended research to characterize the association between COVID-19 and social determinants of health factors. Moreover, an epidemiological update of association between various social determinants of health and preexisting comorbidity factors among adults, especially in contexts with elevated COVID-19 incidence such as Indiana, is imperative given the variations in social determinants of health factors and preexisting comorbidities and the incidence of COVID-19 disease among adults (Indiana Department of Health, 2023).

Given the tremendous vulnerability of adults to COVID-19 infection, the current study aimed at investigating whether there is an association between social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical condition (e.g., obesity), and COVID-19 infection diagnosis incidence, especially among adults aged 18 years and older in the state of Indiana, using the COVID-19 Case Surveillance Public Use Data with Geography from the CDC (2023).

The COVID-19 disease pandemic is still ongoing, even with the wide uptake of the approved COVID-19 vaccines and other effective treatments available to combat the spread of the disease in various communities. Furthermore, clear knowledge and preventive behaviors towards COVID-19 and its risk factors are still lacking (CDC, 2022; Khanijahani et al., 2021; Salgado de Snyder et al., 2021). These also created an urgent need for the current study. Thus, it was of paramount importance to fully understand the characteristics of the disease among adults ages 18 and older to effectively combat it. The results of the current study could be used to strategically guide and inform the COVID-19 public health agenda and improve people's decisions to be healthy.

Figure 1

Map of the Midwest Region of the United States Showing Indiana



Note. From *Midwest Census Region*, by the U.S. Bureau of Labor Statistics (https://www.bls.gov/regions/midwest/midwest.htm). In the public domain.

Problem Statement

According to Abebe et al. (2020), CDC (2022), Iezadi et al. (2021), and WHO (2022, 2023), the ongoing COVID-19 pandemic that emerged out of China in November 2019 is one of the world's most serious public health crises in history, and the United States has the largest population of COVID-19-infected people in the world, estimated at 104,242,889 cases, 1,127,104 deaths, and 120,820 new cases as of April 9, 2023. From March 15, 2023, to April 4, 2023, Indiana alone reported 5,900 new cases of COVID-19 and 75 deaths (Indiana Department of Health, 2023; MCPHD, 2023). Also, Indiana had one of the highest burdens and number of cases of COVID-19 diagnoses among vulnerable adults in the Midwest region of the United States (1,683,505) as of April 4, 2023 (Indiana Department of Health, 2023). COVID-19 disease is a highly contagious and transmittable respiratory disease caused by the SARS-CoV-2 infection, with a case fatality rate of about 2.3% (Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; WHO, 2020).

It has long been documented that social determinants of health and preexisting comorbidities play a significant role in the transmission and severity of infectious diseases including COVID-19 in the United States, including Indiana (CSDH, 2008; Indiana Department of Health, 2023; Smith et al., 2018). The social determinants of health include sex, race, exposure, and age (CDC, 2022; CSDH, 2008; Smith et al., 2018). Furthermore, adults and those with preexisting comorbidities are more likely to get sick and hospitalized with COVID-19 than younger individuals (Khanijahani et al., 2021; Mankowski et al., 2021). The United States have a high burden of COVID-19 posing a major health burden for adults aged 18 years and older, who are more likely to have preexisting comorbidities and have serious complications and severity, including social determinants of health inequities (CDC, 2021).

Although leading public health entities including the CDC and the WHO have stressed the significance of diligently practicing social distancing, hand hygiene, wearing a mask in public, getting fully vaccinated, and promoting COVID-19 health awareness as the best strategies for combating and controlling COVID-19, failure to quickly and fully address social determinants of health and preexisting comorbidities effectively and also enlighten individuals about the pandemic continue to worsen the burden of COVID-19 (CDC, 2022; CSDH, 2008; Smith et al., 2018, WHO, 2021) in various communities and vulnerable populations (CDC, 2022). The CDC (2022) and Chopel et al. (2015) emphasized the significance of studying and addressing the social determinants of health and preexisting comorbidities because they are critical for prevention and improving health equity in health outcomes such as COVID-19 and promoting a more equitable society. Researchers have explored the association between individual-level factors (e.g., age) and community-level factors (e.g., county code of residence) on COVID-19 infection among adults in a number of countries (Khanijahani et al., 2021; Samuel et al., 2021). A few studies (Dixon et al., 2021; Hanson et al., 2020; Samuel et al., 2021; Timsina et al., 2020) have been conducted on the association between various social determinants of health factors and comorbidities and COVID-19 infection in Indiana. However, most of these studies, including most of those that have been conducted in other areas outside Indiana, have reported inconsistent or conflicting results on the incidence of COVID-19 infection and some social determinants of heath factors and preexisting comorbidities, and they have also lacked a clear theoretical framework to examine associations between these variables and to produce high-quality evidence and to clearly understand the particular effects of these factors on COVID-19 outcomes in this population (Dixon et al., 2021; Hanson et al., 2020; Khanijahani et al., 2021; Samuel et al., 2021; Timsina et al., 2020). Additionally, most of the COVID-19 studies in the region, including in Indiana (Dixon et al., 2021; Hanson et al., 2020; Khanijahani et al., 2021; Samuel et al., 2021; Timsina et al., 2020), were found to be targeting the general population as a whole, leaving specifically the population of adults ages 18 and older in Indiana inadequately targeted or investigated, despite the mounting rate of COVID-19 infection among this group during the ongoing pandemic in Indiana. These limitations are important because scarcity of research data fully characterizing COVID-19-infection-associated social determinants of health and preexisting comorbidities for older adults makes them an invisible at-risk population, hinders appropriate allocation and distribution of needed resources and prevention of COVID-19 infection among at-risk populations, and hinders a full understanding of the continual rapid spread of the COVID-19 disease (CDC, 2021; CSDH, 2008; Dixon et al., 2021; Hanson et al., 2020; Khanijahani et al., 2021; Maness et al., 2021; Timsina et al., 2020) among the population of at-risk adults aged 18 and older in the United States, particularly Indiana. Notably, in Brown et al.’s (2022) rapid review and synthesis of public health recommendations from 338 peer-reviewed studies on addressing equity and social determinants of health factors during the COVID-19 pandemic, a substantial number of articles emphasized the need for

COVID-19 research on the social determinants of health and underlying conditions. Establishing a thorough understanding of the factors that promote the ongoing spread of the disease is an essential step in limiting the spread of this disease and curtailing mortality in Indiana and elsewhere. Thus, in the present study, I investigated the association between social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (e.g., severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension), and the likelihood of contracting COVID-19 among adults 18 years and older in Indiana using the social ecological model (SEM) framework and a cross-sectional study of secondary data (COVID-19 Case Surveillance Public Use Data with Geography) collected by the CDC during the ongoing pandemic in Indiana.

The findings from my study can contribute to filling a gap in knowledge and provide a clearer understanding of social determinants of health and preexisting comorbidities related to COVID-19 incidence. It can strengthen the use of a sound theoretical framework (e.g., SEM) to support and increase opportunity for individuals to adhere to and follow evidence-based research and data on social determinants of health and preexisting comorbidities related to COVID-19, which can result in increased prioritization in COVID-19 vaccination, efficient allocation of scarce COVID-19 resources, increased trust in COVID-19 vaccine uptake, and targeted COVID-19 intervention measures specifically focused on vulnerable individuals aged 18 years and older. It can also foster evidence-based decision making that improves population health

(CDC, 2022; España et al., 2021; Khanijahani et al., 2021; Salgado de Snyder et al., 2021).

Purpose of the Study

The purpose of this quantitative study was to examine the association between social determinants of health (exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, after controlling for age group, sex, and race. I also evaluated the predictive ability of these independent variables on the current status of COVID-19 infection among adults 18 years and older in Indiana. These findings will help inform potential public health initiatives to reduce COVID-19 disparity among adults ages 18 and older in Indiana and elsewhere.

Research Questions and Hypotheses

Research Question 1 (RQ1): Is there an association between individuals' age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana?

Null hypothesis ($H_0 1$): There is no association between individuals' age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana.

Alternative hypothesis ($H_a 1$): There is a statistically significant association between individuals' age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana.

Research Question 2 (RQ2): Is there an association between individuals' exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race?

Null hypothesis ($H_0 2$): There is no association between individuals' exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

Alternative hypothesis ($H_a 2$): There is a statistically significant association between individuals' exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

Research Question 3 (RQ3): What is the association between underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana?

Null hypothesis ($H_0 3$): There is no association between underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana.

Alternative hypothesis (H_a3): There are statistically significant associations between underlying medical conditions (severe obesity [as measured by BMI $\geq$ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana.

Research Question 4 (RQ4): What are the significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI $\geq$ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana?

Null hypothesis (H_04): There are no statistically significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI $\geq$ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

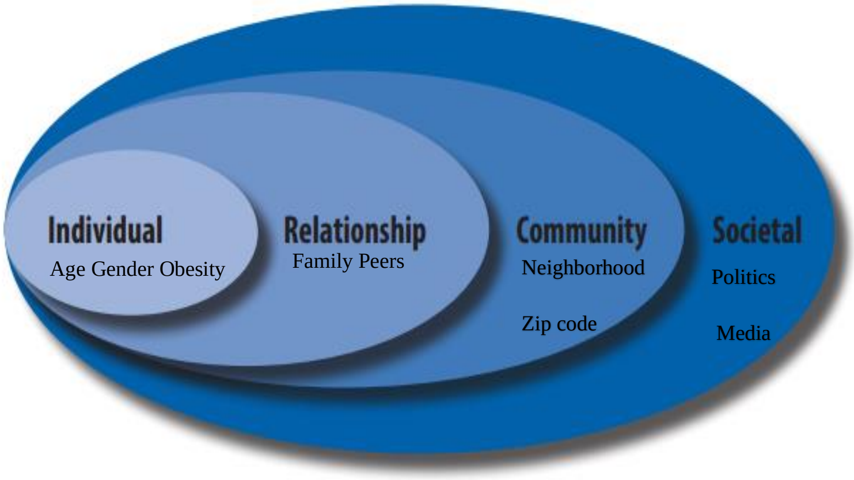
Alternative hypothesis (H_a4): There are statistically significant predicting social determinants of health (age group, sex race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

Theoretical and/or Conceptual Framework

According to Glanz et al. (2008), the SEM framework by McLeroy et al. (1988), first proposed by Bronfenbrenner (1979), is one of the most commonly used health behavioral models that fully explain and describe the multiple dimensions and factors such as social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) that influence health, including COVID-19 infection/incidence in a community. The model can also effectively help to evaluate the predictive ability of these independent variables on COVID-19 infection among community groups, including adults 18 years and older in Indiana (Bronfenbrenner, 1979; Glanz et al., 2008; McLeroy et al., 1988). In epidemiological studies, COVID-19 has been found to be associated with social determinants of health and underlying medical conditions at multiple levels (Figueroa et al., 2020; Hains et al., 2020; Hanson et al., 2020; Khanijahani et al., 2021; Lundon et al., 2020; Maness et al., 2021; Salgado de Snyder et al., 2021; Smith et al., 2018; Upshaw et al., 2021). Therefore, effectively tackling COVID-19 infection incidence and preventing the disease requires addressing social determinants of health and preexisting comorbidities factors at each of the noted levels of the SEM (Glanz et al., 2008; Samuel et al., 2021). These factors can be grouped, understood, and thoroughly explained utilizing the major levels of influence conceptualized by the SEM framework (Figure 2). These levels will be explained in more detail later in this section.

Figure 2

The Social Ecological Model Diagram



Intrapersonal/Individual Level

Intrapersonal/individual-level social determinants of health include age group, sex, race, exposure, and underlying medical conditions (e.g., severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension; Alam et al., 2021; Fielding-Miller et al., 2020; Glanz et al., 2008). According to this model, for example, sex may increase the possibility of contracting COVID-19 partly due to established cultural roles and gender norms in various communities that influence risk for contracting the disease, such as greater chance of women working in healthcare, service industries, and caregiving to the sick at home compared with men (Glanz et al., 2008;

Ortolan et al., 2020). This increases women’s likelihood of exposure to viruses such as SARSCoV-2 (Glanz et al., 2008; Ortolan et al., 2020). In terms of age group, for example, the model postulates that the high COVID-19 vulnerability in adults may likely result from issues such as underlying medical conditions/comorbidities and weak immune response to viral infection that worsens disease progression (Casas-Deza et al., 2021; Glanz et al., 2008).

The model further postulates that, for example, unfavorable socioeconomic circumstances may be responsible for the racial disparities in COVID-19 outcomes that minorities such as Blacks and Hispanics suffer from. Such circumstances may subject these individuals to working service jobs outside the home, living in congested housing, and using public transportation, which increase the chance of contacts with potential sick people and acquisition of infection (Glanz et al., 2008; Unruh et al., 2022). More educated individuals are more likely to have higher paying jobs, increased social status, and health literacy, which can increase health understanding, and compliance with health interventions such as social distancing and other COVID-19 pandemic interventions to prevent the spread of the virus (Glanz et al., 2008; Samuel et al., 2021). In addition, lower educational attainment may lead to failure to perceive the threat of COVID-19 and easily overlook one’s health, making individuals prone to higher risk of virus exposure and transmission (Glanz et al., 2008; Samuel et al., 2021). For individuals’ exposure related to workplace or occupation, the model presents those bad conditions at the individual and community level, including racism, which may promote COVID-19 infection (Glanz et al., 2008; Upshaw et al., 2021).

Additionally, according to the model, for example, obese individuals may express higher levels of viral entry factors such as ACE2, which increase SARS-CoV-2 virus binding and entry in the human lungs and other organs in the body and thus may increase susceptibility to SARS-CoV-2 infection. Furthermore, obese individuals may have other comorbidities such as chronic lung malfunction that can hinder breathing and reduce oxygen level in the lungs, which can lead to obese patients developing more severe COVID-19 (Glanz et al., 2008; Zhang et al., 2021). The model also postulates that, for example, poor immune response caused by underlying medical conditions such as diabetes may increase and worsen COVID-19 situations in such patients (Glanz et al., 2008; Parohan et al., 2020), thus also promoting COVID-19 progression and possible organ damage (Shah et al., 2021).

Community Level

Community-level social determinants of health include, for example, county code of residence/neighborhood and housing elements (Glanz et al., 2008; Hanson et al., 2020). According to this model, county code of residence/neighborhood—characterized, by, for example, high poverty rate, lower educational attainment individuals, disadvantaged vulnerable minorities such as Black and Hispanic, residential segregation, and poor infrastructure—can foster closer and more prolonged interactions in crowded environments that can increase exposure to the COVID-19 virus and hinder access to preventive measures such as early and frequent COVID-19 testing and vaccination, which may increase COVID-19 spread (Clouston et al., 2021; Glanz et al., 2008; Samuels-Kalow, 2021).

The model can help in using rigorous and sound statistical analysis methods such as hierarchical regression to determine the explanatory power of the model as a whole as well as the contribution of each of the factors and comorbidities in order to fully examine various social determinants of health elements and comorbidities linked to COVID-19 and the role of these factors in a person's health status and subsequent COVID-19 disparities based on the SEM framework (Bronfenbrenner, 1979; Glanz et al., 2008; Green & Salkind, 2014; McLeroy et al., 1988).

It is possible to better understand the COVID-19 infection incidence characteristics of adults ages 18 years and older in Indiana and come up with sound public health problem-solving strategies using the SEM (Glanz et al., 2008). The model can help in answering the research questions effectively and increase and improve the current data on the ongoing COVID-19 pandemic. The devastating COVID-19 pandemic has generated major public health concern around the world (WHO, 2021). Certain social determinants of health and underlying medical conditions including age and race have been implicated in the COVID-19 issue and disparity in various communities in the United States, including Indiana (WHO, 2021). The current quantitative and cross-sectional study using secondary data, the COVID-19 Case Surveillance Public Use Data with Geography (from the CDC), to characterize COVID-19 and risk factors in adults aged 18 and above in Indiana needed a sound theoretical framework. In my review of the literature, I found that other theories such as the epidemiologic triad of disease utilize a host, an agent, and an enabling environment to promote transmission of a disease (Ana et al., 2018). The epidemiological theory of disease highlights that poor health or disease results from various interconnected factors instead of just one factor (Glanz et al., 2008). However, even though these causal theories have provided valuable information about disease in vulnerable populations around the world, they focus less on the interdependent nature of these factors and have consistently failed to fully integrate multilevel perspectives (Glanz et al., 2008; Wilcox & Echaubard, 2017). They have also failed to demonstrate that an infectious disease such as COVID-19 also encompasses social determinants of health and preexisting comorbidities in a population and community (Glanz et al., 2008; Wilcox & Echaubard, 2017; WHO, 2021). Well-established social determinants of health elements (e.g., race and sex) and comorbidities have been linked to the COVID-19 infection disparity in various populations and communities in the United States (CDC, 2021; WHO, 2021). These factors also have links to many constructs of the SEM, particularly the individual- and community-level influences, which can be properly understood and described using McLeroy's SEM (Glanz et al., 2008; Samuel et al., 2021). So, for this study, the ecological model by McLeroy et al. (1988) was most appropriate.

A great number of epidemiological studies have utilized the SEM to promote public health agendas and intervention programs due to its sound ability to explain poor health conditions based on multiple interacting factors such as the individual and community instead of just looking at individual-level health factors such as obesity and age. The COVID-19 disease and disparities are characterized by multiple social determinants of health and comorbidity elements, so gaining proper understanding and effectively combating the disease will require this model (CDC, 2017; Glanz et al., 2008; Samuel et al., 2021).

Hence, for this study, the ecological model by McLeroy et al. (1988) was most appropriate. Glanz et al. (2008) noted that the model components demonstrate a reciprocal relationship among themselves via a complex interaction with no specific cause–effect link relationship between model components. So, from this, it is likely that the interconnectivity between components may promote differential effects on the subpopulation of adults 18 years and older in Indiana (Glanz et al., 2018). Furthermore, this was vital because my study was not geared towards establishing cause-and-effect relationship between the study's independent variables and the dependent variable COVID-19 diagnosis, but to examine whether there was an association between social determinants of health (age group, sex, race, exposure, county code of residence) and underlying medical conditions (e.g., obesity [as measured by BMI $\geq$ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 disease status among adults aged 18 years and older during the COVID-19 pandemic in Indiana. Additionally, the model is well renowned in public health research and disease interventions to improve people's lives in various communities (Glanz et al., 2008). For this study, I focused mainly on the individual and community levels of the model.

Nature of the Study

To address the four research question(s) in this current quantitative study, the specific research design used secondary data using the COVID-19 Case Surveillance Public Use Data with Geography (from the CDC) with a cross-sectional design to examine how social determinants of health (age group, sex, race, exposure, county code of residence) and underlying medical conditions (e.g., obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) collectively influence COVID-19 infection incidence among adults 18 years and older in the state of Indiana (ISDH, 2021). A cross-sectional study design can help in critically evaluating association at a single point in time (snapshot) and is therefore capable of measuring the prevalence of health outcomes such as COVID-19 infection and exposure prevalence at a single and specific point in time and place (Creswell, 2009). It can also describe demographics of a population (Creswell, 2009).

Literature Search Strategy

I searched online for relevant studies using the Walden Library database, PubMed, Web of Science, Embase, Medline, PsycINFO, Cochrane Library, ERIC, CINAHL, and Google Scholar. I obtained additional papers or articles by reviewing bibliographies. I also specifically searched for relevant studies on relevant theoretical frameworks that address infectious diseases, including the COVID-19 pandemic. From this review, I identified the relevant and needed theoretical framework for this research. Furthermore, embarking on this review contributed to a robust study design for empirical research.

The key literature search terms for this study included various combinations of *COVID-19 risk*, *COVID-19 pandemic*, *COVID-19 transmission*, *social determinants of health*, *preexisting comorbidities*, *COVID-19 control or prevention*, *COVID-19* *epidemiology*, *COVID-19 disparity*, *adults and COVID-19*, *COVID-19 vaccine,* and *social and health behavior theories*.

The review entailed only studies that were published in English and contained the following: fever, risky behavior, dry cough, comorbidity, and COVID-19—basically, all the relevant studies that met the CDC and WHO COVID-19 disease case definition and its association with comorbidities and social determinants of health (CDC, 2020; WHO, 2020). Furthermore, specific studies that characterized most of the study population as adult COVID-19 patients with comorbidity, healthcare workers, essential workers, or minorities were reviewed for this study. Various studies were excluded from this study if they only addressed the seasonal flu because previous research had indicated that COVID-19 exhibited similar symptoms as the seasonal flu, which had resulted in COVID-19 misdiagnosis. Additionally, I excluded studies that were not written in the English language and studies more than 3 years old, as the COVID-19 pandemic in Indiana and elsewhere started in 2019, although a few studies older than 3 years were included in the study if they contained pertinent COVID-19 epidemiological background. Also, I excluded articles that only addressed comorbidities such as obesity because COVID-19 is considered a novel infectious disease. I utilized only peer-reviewed journals and credible reports from government agencies or other well-respected organizations for the literature review. These types of materials are well written and organized, and they also provide high-quality, reliable, credible, and evidence-based data that can be effectively analyzed and used for COVID-19 study. Moreover, they provide in-depth information on specific study topics, and they are widely used in epidemiologic, health promotion, and disease control and intervention studies and programs.

From the studies and reports that met my literature review criteria, I extracted information about the author, the year, the sample used, the study methodology used, the guiding theories, reported strengths, limitations, and findings that encompass the sample characteristics and COVID-19 infection disparities.

Theoretical Framework

McLeroy et al.'s (1988) SEM (Glanz et al., 2008; Samuel et al., 2021) served as a theoretical framework for this study. The model was first utilized by Bronfenbrenner (1979) to mainly study and understand the reciprocal interaction between individuals and their natural environment. The SEM encompasses multiple levels, including individual-level factors such as biological traits (gender, race); interpersonal-level factors such as friends and family; organizational-level factors such as the CDC; community/ neighborhood-level factors; and political-level factors such as policies that influence disease transmission and health behaviors such as COVID-19 (Glanz et al., 2008; Samuel et al., 2021). According to Glanz et al. (2008) and Samuel et al, (2021), the SEM depicts an interplay among the various levels.

Studies of the Social-Ecological Model

In a recent qualitative study out of Kenya on truck drivers published in the *African Journal of AIDS Research,* Christian et al. (2020) adopted the SEM framework and found that a greater number of the participants agreed to using an HIV self-test compared to the control group. Moreover, at the intrapersonal level, individuals mentioned perceived barriers such as a lack of time to test or obtain a self-test kit. At the interpersonal level, participants mentioned facilitators such as engaging in self-testing with a partner. At the policy level, participants offered various HIV/AIDS health programs and policy recommendations. This study clearly showed that the model can help explain diseases such as COVID-19 at multiple levels of interactions (Glanz et al., 2008).

In another recent study out of the United States, Noh and Min (2020) used 2,054 9-year-old children to study obesity within the individual and community levels of the SEM framework. They reported that at the individual level, for example, the child’s race (for Black and Hispanic children) was linked to a higher risk of obesity compared to their White counterparts (b = 1.56 and b = 1.84, respectively. Also, at the community level, children whose teachers perceive the school environment as a low threat for obesity had low risk of obesity when compared to children whose teachers perceived the environment as a major threat for obesity (b = −0.59). The model was able to effectively predict social-ecological factors that were associated with child obesity/overweight (BMI).

Clearly, these studies show that McLeroy et al.’s (1988) SEM can effectively address health research and interventions at various levels, including individual- and community-level health factors such as race, obesity, and the neighborhood. Moreover, the studies demonstrate the utility of the model in providing opportunities for exploring multiple and interdependent levels of influence exerted on diseases such as COVID-19.

For the current study, the SEM provided better guidance towards clearly understanding the multiple factors that impact COVID-19 infection and addressing the social determinants of health and underlying medical conditions associated with the spread of COVID-19 among the study population in Indiana (Christian et al., 2020; Glanz et al., 2008; Noh & Min, 2020).

The model helped provide evidence-based data and pertinent information for combating COVID-19 and properly guiding COVID-19 vaccine distribution in various communities to curtail the spread of the disease in the United States (Samuel et al., 2021). It also helped to answer this study's research questions and provided better understanding and knowledge on COVID-19 disease and related social determinants of health and comorbidity risk factors among the study population (Christian et al., 2020; Glanz et al., 2008; Noh & Min, 2020; Samuel et al., 2021). Multiple social determinants of health and preexisting medical conditions such as age and obesity contribute to COVID-19 characteristics and disparity; thus, characterizing and effectively addressing the disease requires McLeroy et al.'s (1988) SEM (Glanz et al., 2008; Samuel et al., 2021).

Literature Review Related to Key Variables and/or Concepts

In the following literature review, I summarize prior studies relevant to my study of COVID-19 infection incidence within McLeroy's SEM framework.

Individual Level of Social Ecological Model

At the individual level, an adult's age group, sex, race, exposure, and underlying medical conditions such as severe obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension significantly correlate with COVID-19 infection/incidence.

Age Group and COVID-19

Recent large epidemiologic studies have suggested a link or correlation between adult age group and COVID-19 disease. For example, Casas-Deza et al. (2021) recently conducted a large epidemiologic study that applied a longitudinal study method to 61,993 confirmed COVID-19 cases in Aragon, Spain from March to November 2020. In this study, the authors performed Kruskal-Wallis and $\chi2$ statistical analysis tests to assess the relationship between age group and COVID-19 incidence and death among the study group and found a higher incidence of COVID-19 cases among individuals aged 40–49 years. Also, 43% of participants were over 50 years of age, while 14.9% of them were above 75 years. Furthermore, the disease was more severe in adults aged 75 years and older.

A large systematic review and meta-analysis study by Biswas et al. (2021) that included 20 studies comprising 64,676 COVID-19 patients examined the association of sex, age, and comorbidities with mortality in COVID-19 patients and found that people aged 50 years or older had 3.45 times significantly higher risk of COVID-19 compared to individuals younger than 50 years (RR 3.45; $p = 0.0008$). Furthermore, in an ecological study by Kawatake de Sousa et al. (2021) on all ages using data on 524,496 COVID-19 cases and 11,516 deaths from the Health Department of the State of EspíritoSanto, Brazil to assess COVID-19 infection and mortality from March 2020 to June 2021, the researchers found the highest incidence rates among people aged 30 to 39 years (19,639.19 cases per 100,000 people) and those aged 40 to 49 years (19,098.20 cases per 100,000 people), while deaths were highest among people aged 60 to 69 years (24.85%)

and people 80 years and older (23.11%). Similarly, in a retrospective, observational study out of Spain, Gimeno-Miguel et al. (2021) used 68,913 individuals to investigate COVID-19 infection based on age. They found that COVID-19 risk was 6.19 times higher for men and 8.19 times higher for women aged 80 and over compared to individuals aged 45–64 years. Also, in this study, patients ages 80 years and older had 8.23 times higher risk than those aged 45–64 years.

Furthermore, in a recent study, Ramírez-Soto et al. (2021) used 961,894 COVID-19 cases from the national population-based surveillance system of Peru to examine the association between age and COVID-19 incidence, mortality, and fatality per 100,000 people. The authors demonstrated that, among the study population, the highest number of COVID-19 cases were in women aged 40 to 49 years (12,592 per 100,000 women) and men 80 years or older (12,571 per 100,000 men). Also, fatality rates were higher in men aged 50 and older compared with women aged 50 or older. The highest fatality rates were reported in men and women 80 years or older. The highest mortality rates were found in individuals ages 80 years or older (2,590 per 100,000 men and 1,288 per 100,000 women). The fatality rate was 2 times higher in men than in women (4.96% vs. 2.41%, respectively), and was high in men 50 years of age or older.

Also, recently Casillas et al. (2021) conducted a retrospective cross-sectional study of 278 Latinx patients with COVID-19 at a large tertiary public safety-net hospital in Los Angeles County to assess the impact of age in COVID-19-related mortality. They found that mortality increases with age, starting from the 4th until the 9th decade of life (2.99%, 11.76%, 15.22%, 31.58%, 62.50%, and 100% mortality, respectively).

Additionally, in Hamza et al.'s (2022) study out of the Kashmiri community in India that examined associations between age and COVID-19 infection and mortality rate, severity of COVID-19 disease increased with increasing age, and the severity of illness was found to be 35.74% in individuals 60 years of age and older. Moreover, in Al Mutair et al.'s (2021) study, age had a significant association with mortality (*OR* = 1.080), and the odds of nonsurvivors were 1.069 times higher among older patients as compared to younger patients.

Different elements may help explain this trend seen in COVID-19 incidence in these populations and communities. Casas-Deza et al. (2021) concluded that the high COVID-19 vulnerability in adults probably resulted from long-established structural systemic inequities and social determinants of health issues in the United States, in addition to preexisting comorbidities, and poor and weak immune response to viral infection associated with age.

However, in a recent study, Hamza et al. (2022) investigated the association between age and COVID-19 infection in the Kashmiri community, India, using 957 patients diagnosed with SARS-CoV-2 infection. They found no significant association between the age group and COVID-19 incidence, severity of the illness, and deaths, even though COVID-19 incidence was high among older individuals. Two hundred eighty-six patients 40 years and younger had COVID-19, with severity of illness 30.42% (P = 0.428) and a mortality rate of 11.54% (P = 0.93). There were 366 cases for individuals 40–60 years; the severity of the illness was found to be 32.51% (P = 0.889), and the mortality rate was 12.84% (P = 0.57). There were 305 COVID-19 cases in individuals 60 years and older, with a severity of the illness of 35.74% (*P* = 0.36) and a mortality rate of 10.49% (*P* = 0.56; Hamza et al., 2022). Similarly, a cross-sectional study conducted on 57,958 new cases of COVID-19 and associated factors including age in Fars province, Iran, found that patients ages 57 years and older had a significantly higher mortality rate than individuals ages 57 years and younger (*OR* = 17.56, *P* < 0.001). However, no significant association was observed between COVID-19 infection and death among children aged ≤ 4 years and those between 5 and 15 years old (Semati et al., 2022). Additionally, in El Rifay et al.'s (2021) study, there was an association between antibody levels and COVID-19 (*P* = .039). However, people over 70 years were seronegative. Also, in a recent study published in *Science*, Mena et al. (2021) analyzed the association between age, income status, and COVID-19 incidence and mortality in Greater Santiago, Chile. They found higher COVID-19 death rates in young people and those in low-income areas.

Although these studies have presented information on the relationship between age and COVID-19 infection, overall, contradictory, and conflicting epidemiological information has been presented on the association between age and the ongoing COVID-19 pandemic, and the association between age and COVID-19 infection remains unclear. Furthermore, most of these studies were conducted outside Indiana, and Indiana has a scarcity of data on the relationship between age and COVID-19 infection. Additionally, some of the studies even included patients younger than 18 years of age. This is disturbing and of major concern, particularly for Indiana, where the characterization of the pandemic has also received less attention, despite the severe impact on the state during the ongoing pandemic. Clearly understanding the age-related risk element that makes some populations, particularly older people in Indiana, vulnerable to COVID-19 is vital to ensure tailored and more effective interventions for this and future pandemic prevention and control in Indiana and elsewhere. Recently, Casas-Deza et al. (2021) recommended further research to characterize the course of the disease in elderly patients, who seem to be the most vulnerable population group.

I addressed this lack of knowledge and clear understanding gap by conducting a cross-sectional study using secondary data (COVID-19 Case Surveillance Public Use Data with Geography) from the CDC. I also used the social-ecological framework to comprehensively investigate the association between participant’s age group and COVID-19 infection among individuals ages 18 years and older in Indiana. Understanding the factors that influence COVID-19 infection is important. My results can guide COVID-19 prevention and control policies and help in the effective allocation of COVID-19 vaccines and resources in Indiana.

Sex and COVID-19

A range of epidemiologic studies have noted the gender burden in the ongoing COVID-19 disease pandemic in the United States and other countries. Abate et al. (2020) conducted a systematic review and meta-analysis of 57 studies with 221,195 COVID-19-confirmed participants in various countries to examine COVID-19 susceptibility between men and women. They found a higher number of COVID-19 infections among men 55.00 (51.43–56.58), $p < 0.001$), compared to women 45.00 (41.42–48.57). Furthermore, Ronquillo De Jesús et al. (2022), in a study published in the *Journal of Infection in*

Developing Countries, used 1,037,567 confirmed COVID-positive cases in Mexico and analyzed the association between gender and COVID-19 infection. They demonstrated that males accounted for 60% or higher COVID-19 cases. Similarly, in another recent systematic review and meta-analysis of 85 studies from the United States, China, and other countries, Ortolan et al. (2020) examined the association between gender and COVID-19 mortality, severity, and disease outcomes and found an association between male and mortality (*OR* = 1.81) and poor recovery in men (*OR* = 0.72). Males had a greater chance of experiencing severe COVID-19 (*OR* = 1.46).

Also, Unruh et al. (2022) conducted a recent retrospective cohort study out of Cook County, Illinois, and examined the association between gender and COVID-19. Unruh et al. found more COVID-19 mortality among males than their female counterparts 6,227 (58%), and 4,527 (55%) respectively with $P < 0.01$; 2022). In another study, Green et al. (2021) conducted a meta-analysis using national data from Denmark, England, Israel, Italy, Spain, Canada, and Mexico to examine the association between gender and COVID-19 infection and case-fatality rates. They found higher death rates in males of various age groups such as 20–29 (1.71), 60–69 (1.84), and 80 years and older (1.49). Also, males aged 40–60 had more than twice the death rate compared to females. In Hamza et al.’s (2022) recent retrospective study that included 957 COVID-19 patients, most individuals 639 were males while 318 patients were females. Moreover, the number of males admitted was twice that of the females (2022).

Additionally, in a large analysis of a secondary database study that used data from the national population-based surveillance system of Peru and included 961,894 COVID-

19 cases and deaths, Ramírez-Soto et al. (2021) examined association between gender and COVID-19 incidence, mortality and fatality and found that more than half 498,568 (51.8%) of the COVID-19 cases were men, and 24,730 (68.9%) deaths occurred in men. In addition, the COVID-19 infection rate was greater in men (p< .001) compared to women, and men had twice as high risk of COVID-19 death (OR = 2.16; p<.001) compared to women (2021). The authors also, found that the number of COVID-19 diseases was higher among men compared to women (3079; 2819 per 100,000 people, respectively) and the death rate was twice as high in males than in females (153; 68 per 100,000 people, respectively).

Furthermore, Shakor et al. (2021) conducted a rapid review of 60 studies or papers from various countries to examine the epidemiological factors including gender which highly contributed to the morbidity and mortality rates of COVID-19. The authors demonstrated that being male was associated with a high incidence of COVID-19 and most of the studies show that 60% of COVID-19 patients were male. Moreover, Sundaram *et al.* (2022) conducted a cohort study of 100,902 hospitalized COVID-19 patients using the Cerner Real World Data (CRWD) COVID-19 Database to examine the association between sex and COVID-19 infection in the United States. They found that although about half of the 100,902 total cohort were females (50.7%), of the 8574 patients who died, 58.2% were male and 41.3% were female. Also, the death rate was 10.11% for males and 6.93% for females (p<0.001; 2022).

Also, Jin et al.'s (2020) case series study using public data on 43 COVID-19 hospitalized patients, 37 patients who died of COVID-19 and 1,019 patients who survived in China, compared with data of 524 patients with SARS, including 139 deaths, from Beijing in early 2003 to examine the association between gender and COVID-19 infection, the results showed that men and women have the same prevalence COVID-19, however, men encountered a more severe COVID-19 than women($P = 0.035$). Also, of the 37 patients who died, 70.3% were men and 29.7% were women, and the number of men who died from COVID-19 is 2.4 times that of women ($P = 0.016$; 2020). Moreover, men had a higher risk of dying (χ 2 test, $P = 0.016$; 2020). Similarly, in Al Mutair et al. (2021) study, gender had a significant relationship with COVID-19 infection (OR = 5.930). Moreover, the odds of non survivors among male patients were found to be 2.084 times as compared to female patients.

Importantly, Cruz-Arenas et al. (2021); Fortunato et al. (2021); and Ramírez-Soto et al. (2021) concluded that the higher incidence and death rates associated with COVID-19 for men are likely due to the presence of preexisting comorbidities such as cardiovascular diseases, obesity or diabetes, including genetic and hormonal differences between males and females with greater amounts of angiotensin-converting enzyme 2 (ACE2) receptors in males which provides effective binding site for the COVID-19 virus (SARS-CoV-2) to access the patients compared to women. The authors also, mentioned that COVID-19 susceptibility of men could be attributed to social behaviors or lifestyles such as smoking that increase men’s risk for comorbidities associated with COVID-19. Additionally, according to Cruz-Arenas et al. (2021), some COVID-19 patients such as women can produce or generate stronger and sufficient serum SARS-CoV-2 IgG antibodies and stronger immune responses to effectively combat the COVID-19 virus (SARS-CoV-2) infection as compared to men.

On the other hand, Gimeno-Miguel et al. (2021) retrospective, observational study out of Spain that investigated COVID-19 infection based on sex showed an increased average risk of 6.19 in men and 8.19 in women aged 80 and over compared to those 45-64-year-old. Similarly, Alam et al. (2021) conducted a web-based self-reported survey to examine the association between comorbidities such as diabetes disease and obesity, and COVID-19 infection in 780,961 male and female participants from 183 different countries and territories. They found that although most of the participants were male (67.39%), there was a strong significant association between Females and COVID-19 infection with females 1.69 times (AOR: 1.69) more likely to get infected by Covid-19 compared to male participants. Also, Fortunato et al. (2021) conducted a retrospective epidemiological study using 1175 COVID-19 confirmed cases from surveillance data in the Apulian District of Foggia, Italy from 29 February to 30 June 2020 to examine the association between gender and COVID-19 infection. The authors found an association and a similar number of women and men with COVID-19 incidence (50.7%; 49.3%), (χ^2 $p>0.05$).

Even though, these specific studies have provided pertinent information on the relationship between sex and COVID-19 infection in various setting and also suggests that these men and women are vulnerable to COVID-19, and demonstrate noticeable sex differences in COVID-19 risk, there is still lacking and conflicting information assessing the influence of gender status on the incidence of COVID-19 patients within the diverse communities, particularly among adults ages 18 years and older in Indiana. So clearly there is an urgent need to clarify and understand the impact of sex/gender on COVID-19 infection incidence in various settings particularly, among adults ages 18 years and older in Indiana. To address this knowledge gap, I conducted a cross-sectional study to examine association between sex/gender and COVID-19 infection among adults ages 18 years and older in Indiana to provide clear epidemiological data that can be used to tailor COVID-19 intervention according to gender and guide the incorporation of gender analysis into preparedness and response efforts of health interventions. In addition, the findings can provide gender and COVID-19 characteristics data that will help clinicians make appropriate patient-tailored medical decisions. Recently, Abate et al. (2020) recommended further studies on the discrepancies in severity and mortality rate due to COVID-19 among men and women and the associated factors.

Race and COVID-19 Infection

Recent epidemiologic evidence indicates that racial minorities are more likely to test positive for SARS-CoV-2 and have an increased risk of hospitalization and death for COVID-19 compared to white individuals.

Boserup et al. (2020) conducted a cross-sectional study examining the impact of the COVID-19 pandemic on racial/ethnic groups in the United States. They found that COVID-19 infection rates for blacks were 27.0 times higher than whites in Maine. Also, Hispanics were 2.1 times higher than whites in New Mexico. Moreover, death rates were 6.3 times higher in Blacks than whites in the District of Columbia. Also, a recent study out of Mexico that used 1,037,567 confirmed COVID-positive cases and analyzed the association between ethnic groups (indigenous and nonindigenous) and COVID-19 infection and deaths by Ronquillo De Jesús (2022) discovered that overall cumulative deaths per 100,000 population were substantially higher for the nonindigenous population (87.1) than for the indigenous population (13.9).

Similarly, a recent study out of the United Kingdom that used the Biobank data by Hastie et al. (2020) conducted a large retrospective cohort study to examine the association between race/ethnicity and socioeconomic factors and the incidence of COVID-19. They used 499 cases and found that Black people and South Asian people had a higher chance of testing positive for COVID-19 after adjustment for elements such as socioeconomic, lifestyle, and health-related factors s (Black people: OR = 4.30, $p<0.001$) South Asian: OR = 2.42, $p<0.001$) compared to White individuals. Moreover, in a Cross-sectional study that examined the association between race/ethnicity, and COVID-19 incidence that used 587 cases out of the United Kingdom, De Lusignan et al. (2020) found increases in the likelihood of testing COVID-19 positive among Black compared to White adults (OR = 4.75, 95% CI = 2.65–8.51). In another recent retrospective cohort study of 1,052 adult cases out of California, Azar et al. (2020) examined the association between Race/ethnicity and COVID-19 incidence and found that Blacks had a higher chance of testing positive for COVID-19 and also likely to be admitted in hospital than White (OR = 2.67, $p < 0.01$) after adjusting for age, sex, comorbidities, and income variables. Additionally, De Lusignan et al. (2020) conducted a Cross-sectional study on 587 cases out of the United Kingdom to assess Race/ethnicity, socioeconomic hardship, and COVID-19 infection incidence. They found that Blacks had a higher chance of testing positive for COVID-19 compared to White adults after adjustment (OR = 4.75). The authors also noted that people in poor communities had a greater chance of testing positive for COVID-19 (OR = 2.03, $p<0.0001$) compared to their counterparts in wealthy areas.

Selden and Berdahl (2020) used data from the Medical Expenditure Panel Survey to examine the association between employment/occupation, racial/ethnic and COVID-19 incidence, hospitalizations, and mortality among United States adults. They reported that Blacks had a greater chance of experiencing severe COVID-19 and were 1.6 times more likely than Whites to live in households with health workers. Also, adults of Hispanic race had a higher risk for COVID-19 and were 64.5% times more likely to reside in households with people that are compelled to work in the community, compared to Black adults (56.5%), and 46.6% for White adults.

According to Goldstein et al (2022), and Samuels-Kalow (2021) importantly the racial disparities exhibited in COVID-19 disease could be linked to the tremendous socioeconomic disadvantage and poverty that increase COVID-19 risk among people such as Blacks and Hispanics. Low-income earners are more likely to work essential jobs such as service jobs that subject them to work outside the home and live in crowded houses, as compared with White patients. Furthermore, Goldstein et al. (2022); Samuels-Kalow (2021). emphasized that taking up high-risk occupations, such as caring for the sick in hospitals, as well as dependence on public transportation can increase these individual’s chances of encountering COVID-19-infected patients and therefore, acquisition of infection.

However, in Gold et al. (2020) cross-sectional study that used 305 COVID-19 patients in Georgia hospitals United States to assess the relationship between Race and COVID-19 incidence, 247 (83%) of patients were Black. There were no significant differences in the number of people who were treated with mechanical ventilation or died among White, Hispanic, and Pacific Islander 50 (17%). Furthermore, in a recent retrospective cohort study of 338 COVID-19 patients out of New York City that examined the association between White, Asian, Hispanic, Black, and unknown races, Toussie et al. (2020) did not find significant differences in COVID-19 incidence by race, although being Hispanic independent predicted severe chest x-ray findings among those who were hospitalized p = 0.03). Moreover, Unruh et al. (2022) conducted a recent retrospective cohort study out of Cook County, Illinois, and examined the association between race group and COVID-19 infection and found that Black patients made up 28% of COVID-19 mortality in the county. However, of the 10,813 deaths, the majority 4,527 (42%) were White.

Despite these data showing that race may contribute to COVID-19 infection, there remains little research investigating the relationship between race and COVID-19 incidence particularly, within adults 18 years and older in Indiana, furthermore, due to the inconsistent results, it is still not yet understood whether such disparities are driven by race. Thus, the influence of race factor on COVID-19 incidence deserves urgent research, to further elucidate the epidemiological burden of SARS-CoV-2 within Indiana and elsewhere and inform public health prevention measures (Ronquillo De Jesús, 2022;

Unruh et al., 2022). Therefore, I investigated the association between race and COVID-19 incidence among adults ages 18 and older in Indiana.

Exposure and COVID-19

Previous studies indicate that exposure related to for example, occupation/employment status, and workplace is associated with risk of COVID-19 incidence/infection. For example, Wang et al. (2020) conducted a case series study out of Liaocheng, China using 26 COVID-19 cases to assess the association between an individual's occupation and COVID-19 incidence. They found that most of the first cases of COVID-19 were retail workers at a supermarket 16 (61.5%). Moreover, some of the patients were retirees (15.4%), students (11.5%), agricultural workers (7.7%), and self-employed (3.9%). In another study, Fan et al. (2020), conducted a retrospective cohort study out of Gansu Province, China, using 54 individuals and assessed the association between occupation and COVID-19 incidence. In their study of COVID-19, they found that most COVID-19 cases were migrant laborers/workers (29.0%) returning from Wuhan, and retirees (47.0%; $p = 0.009$).

Moreover, Fielding-Miller et al. (2020) assessed county-level associations between COVID-19 mortality and the percentage of individuals engaged in farm work. They found that a higher percentage of farmworkers was associated with higher COVID-19 related deaths and each additional percentage unit of farmworkers in a county was connected to 5.79 more deaths ($p < 0.001$).

Additionally, Selden and Berdahl (2020) used data from the Medical Expenditure Panel Survey to examine the association between employment/occupation, racial/ethnic, and COVID-19 incidence, hospitalizations, and mortality among United States adults. They found that Blacks who were highly prone to getting COVID-19 were 1.6 times more likely to reside with health workers than Whites. Also, Hispanic adults who were prone to getting COVID-19 were 64.5% times more likely to reside with people who worked in the community compared to 56.5% for Black adults and lower 46.6% for older Whites. Furthermore, in Cruz-Arenas et al.'s (2021) recent cross-sectional study of 2974 hospital employees out of Mexico City, individuals who worked in the security and janitorial work sectors had substantially higher rates of COVID-19 incidence (62.5% and 45.4%, respectively) compared to other work sector groups. They also found the strongest association between working in the security sector (OR 12.35, $p = 0.027$) and janitorial sector (OR 13.70, $p<0.001$) groups and COVID-19 infection.

Upshaw et al. (2021) summed it up by saying that harsh conditions at the individual and community level, including racism may play a pivotal role in increasing the COVID-19 incidence in certain groups. Moreover, individuals with low earning capacity have a greater chance of taking up for example essential service jobs and reside in more crowded housing conditions that may hinder their efforts to practice valuable preventive measures such as physical distance from others, thus increasing their risk of contracting COVID-19 (Upshaw et al., 2021).

However, some studies did not find significant differences in COVID-19 symptom severity by occupation. For example, Ouyang et al. (2020) conducted a Cross-sectional study using 11 patients from Bejing, China, and investigated the relationship between occupation and COVID-19 infection. They found no significant association between COVID-19 severity and an individual’s occupation. Also, in this study, most of the people (54.6%) reported their occupation as retired or working as farmers, with approximately 83.3% of them experiencing severe COVID-19 symptoms. In another case-control study out of Wenzhou, China that used (11 cases; and 62 controls), to assess the association between occupation and COVID-19 disease incidence, Yu et al. (2020) found that most of the patients reported working in the agriculture sector (48.9%), followed by self-employed workers (22.8%). However no significant association was found between the type of occupation and the seriousness of COVID-19.

Even though these studies provided pertinent information on the relationship between exposure related to for example, workplace, occupation factor, and COVID-19 infection in various populations, they provided conflicting results and were also, conducted outside Indiana, leaving the state with scarce data and a lack of clear knowledge on the relationship between exposure and related occupation/employment status and risk of COVID-19 incidence/infection for the State and elsewhere (Upshaw et al., 2021). I conducted a cross-sectional study to fill this gap in understanding to improve COVID-19 health in the study population.

Underlying Medical Conditions and COVID-19 Infection

Recent epidemiological studies conducted in the United States and elsewhere have demonstrated an association between underlying medical conditions or comorbidities such as obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, hypertension, and COVID-19 infection/incidence.

Obesity (BMI ≥ 30 kg/m2)/Comorbidity. A few epidemiological studies have noted an association between obesity and COVID-19 characteristics among certain populations globally. For example, in a recent retrospective cohort research out of New Orleans, LA that was published in the *Annals of Epidemiology,* Mankowski et al. (2021) looked at the relationship between certain comorbidity with COVID-19 infection using 309 COVID-19 adults from March 9 to May 29, 2020. They also, discovered that Black participants were more prone to obesity (55%) compared to their Whites counterparts (36%). Moreover, obese patients had a higher chance of experiencing COVID-19 problems compared to nonobese patients.

In another recent study, Gimeno-Miguel et al. (2021) conducted a retrospective, observational study out of Spain using 68,913 individuals to investigate COVID-19 infection based on comorbidity. They found that COVID-19 incidence was more likely severe in individuals with comorbidity (OR) men, OR women) cardiovascular disease (1.28, 1.39), diabetes (1.37, 1.24), and obesity (1.21, 1.26) respectively.

Furthermore, Hamza et al. (2022) examined associations between comorbidities and COVID-19 infection and mortality rate in the Kashmiri community, in India. They found high COVID-19 severity in patients with comorbidities including cancer, chronic obstructive pulmonary disease (COPD), heart disease, diabetes mellitus, and hypertension nearly 54.54% and death up to 20.7%. Furthermore, the severity of COVID-19 disease increased with increasing age, for example in the age group 60 years, the severity of the illness was found to be 35.74%. Similarly, the Sajjad et al.'s (2022) case series study of 17 patients out of Ireland, spanning March to May 2020 to examine the association between obesity and COVID-19 infection reported that all the patients were obese. They also found that Obesity was highly connected with severe COVID-19 disease to the extent that five patients needed critical hospitalization and subsequently used mechanical ventilation.

Moreover, a retrospective epidemiological study by Fortunato et al. (2021) using 1175 cases of COVID-19 from Foggia District, Italian from the months of 29 February to 30 June 2020, showed that 373 cases (31.7%) had chronic diseases including obesity (6.7%). Similarly, a recent systematic review and meta-analysis of 13 studies that examined the link between metabolic risk factors and COVID-19 incidence found a high prevalence of obesity in COVID-19 patients 29% (Moazzami et al., 2020). Also, in Karmakar et al.’s (2021) cross-sectional study of US counties, which used data sets from the Johns Hopkins University, obesity rate was associated with COVID-19 diagnosis incidence (IRR, 1.02 $P < .001$) and death (IRR, 1.02, $P < .001$).

According to Moazzami et al. (2020); and Zhang et al. (2021), this high-risk COVID-19 disparity from obesity can be attributed to the fact that obese individuals may express higher levels of a key receptor such as ACE2, a receptor that promotes the effective binding of SARS-CoV-2 virus to enter the human lungs and other organs in the body and thus, may increase susceptibility to SARS-CoV-2 infection. The authors also emphasized that obese individuals may have other comorbidities such as chronic lung malfunction which can hinder breathing and reduced blood oxygen levels in the lungs which can lead to obese patients developing more severe COVID-19 outcomes. Also, obesity can trigger the production and release of detrimental elements in the body, such as TNF-α, and leptin which causes inflammation and poor immunity response to infection (Moazzami et al., 2020; Zhang et al., 2021).

However, in a recent study published in the journal *Preventive Medicine Reports,* Alam et al. (2021) conducted a web-based survey using 780,961 participants from various countries to examine the association between preexisting comorbidities including obesity and COVID-19 infection. They found no significant association between obesity and COVID-19 infection OR (0.94), p = 0.28.

Even though these studies provided strong data to suggest associations between obesity and COVID-19 in various populations, they were conducted outside Indiana, leaving Indiana with scarce data. These studies also presented inconsistent and conflicting results (Moazzami et al., 2020; Zhang et al., 2021), causing ambiguity in knowledge on the association between obesity and COVID-19 among adults ages 18 and older in Indiana, and this can stifle COVID-19 intervention in this community and elsewhere.

I conducted a cross-sectional study to critically examine the association between obesity comorbidity and COVID-19 incidence among adults ages 18 years and older in Indiana to clarify the relationship and provide high-quality evidence or data for policymakers. A clear understanding of the link between obesity and COVID-19 would improve COVID-19 mitigation efforts and help protect people that are most vulnerable to severe COVID-19. Furthermore, the findings can help us to better understand the disparities within different settings.

Diabetes Comorbidity. Many epidemiological studies have linked COVID-19 risk increases with Diabetes comorbidity in certain populations. For example, in Gimeno-Miguel et al.'s (2021) retrospective, observational study out of Spain using that used 68,913 individuals to investigate COVID-19 infection based on comorbidity, COVID-19 incidence was more likely severe in individuals with comorbidity (OR) men, OR women) diabetes (1.37, 1.24). Similarly, Hamza et al.'s (2022) study that characterized comorbidities and COVID-19 infection and mortality rate in the Kashmiri community, in India, showed high COVID-19 severity in patients with comorbidities including diabetes mellitus nearly 54.54% and death up to 20.7%. Also, severity of COVID-19 disease increased with the increasing age, for example in the age group 60 years, the severity of illness was found to be 35.74%.

Moreover, in a recently published retrospective case–control study, Assaad et al. (2022) used 696 patients in a largest health care network in New York state to examine association between Type 2 diabetics and mortality in covid-19 infection. They found that more cases than controls had diabetes characteristics including hypertensivity (74% vs 67%, $p = 0.03$), and hemoglobin A1c > 8.1 (71% vs 44%, $p < 0.001$). The studies also showed that home insulin use was independently linked to increased risk of death (Hazard ratio: 1.92). Moreover, 33% were on home insulin and 51% were on metformin. For the duration of their hospital stay. Also, Ernesto Noyola et al. (2021) conducted an ecological study using COVID-19 public data from the Mexican Ministry of Health from 1, 2020 through December 20, 2020, in each of the 32 states in Mexico to examine the association between obesity and diabetes and COVID-19 incidence and mortality rates in

Mexico; (c) they found significantly higher association between diabetes and COVID-19 incidence (adjusted beta 0.2; p < 0.001). Also, a substantial number of older individuals with diabetes had higher mortality rates (p < 0.001).

Additionally, Parohan et al. (2020) conducted a Systematic Reviews and Meta-Analyses of 14 retrospective studies with 29,909 COVID-19 infected patients and 1,445 cases of deaths to examine the association between age, gender, diabetes, and hypertension and COVID-19 infection in various countries. They found a significant association between diabetes and risk of mortality (OR 2.41, p =.037). Similarly, in another recent systematic review and meta-analysis study published in the *Journal of Diabetes,* Bradley et al. (2022) used 24 studies including 10 648 patients to examine the link between diabetes and COVID-19 infection. The findings revealed that Diabetes significantly increased the odds of severe COVID-19 (OR 3.39; P < .0001). Also, most of the patients needed mechanical ventilation (OR 3.03; P < .0001), the prevalence of diabetes was about 31% in hospitalized COVID-19 patients, and the increased death risk was (OR 2.44; P < .0001). In a similar study, Kaminska et al. (2021) recently, conducted a systematic review and meta-analysis using Nineteen studies with 10,801 patients to examine the association between diabetes mellitus and COVID-19 infection and noted that diabetes was an independent risk factor of the severity of COVID-19 in-hospital settings, and COVID-19 severity was 34.8% in Diabetic patients (OR = 1.43; P = 0.20) compared to 22.8% the control group. Also, in a recent systematic review and meta-analysis of 13 studies that examined the link between metabolic risk factors and risk of Covid-19 found a high Diabetes prevalence in Covid-19 patients 22% (95% CI: 12%

33%; Moazzami et al., (2020). Furthermore, Al Mutair et al. (2021) conducted a retrospective study to assess the link between variables such as (age, gender, Diabetes) and COVID-19 incidence in Saudi Arabia. They found that the chances of death among diabetic patients were 5.17 times higher compared to their counterparts who do not have Diabetes mellitus.

According to Parohan et al. (2020); Shah et al. (2021), diabetes patients may also carry the angiotensin converting enzyme 2 (ACE2), that can bind to SARS-CoV and promote entry of the virus into the human body and organs and weakens the immune system enabling the SARS-CoV2 to foster COVID-19 problems in diabetes patients including death from COVID-19. Moreover, Impaired glycemic control can weaken the immune system in individuals with diabetes making it hard to combat the virus (Parohan et al., 2020; Shah et al., 2021). Also, according to Kaminska et al. (2021), Diabetes induces immunosuppression which leaves patients susceptible to developing a far more severe course of infection.

On the contrary, Azam et al.'s (2020) Systematic Review, and Meta-Analysis study of 14 studies and 2568 individuals suffering from COVID-19 infection and diabetes in China and various other countries found that individuals with diabetes had a lower chance of experiencing recurrent SARS-CoV-2 positivity (RR 0.5, I 2 = 53%).

Although these studies these studies provided some insight into the association between diabetes and COVID-19, they presented conflicting results or data, and there is still lack of clear data on COVID-19 infection. Also, most of these studies were conducted solely in other states. Therefore, there is scarce data on the relationship between diabetes and COVID-19 infection among adults ages 18 and older in Indiana. For my study, I assessed whether there is an association between diabetes and COVID-19 infection in my study population using a secondary data set (COVID-19 Case Surveillance Public Use Data with Geography) from the CDC to fully inform COVID-19 prevention policy and programs in Indiana and elsewhere.

Bradley et al. (2022) recommended further studies on the association between diabetes and COVID-19 infection to gain a better insight on the COVID-19 characteristics and improve COVID-19 prevention and control strategies.

Cancer Diagnosis. A number of studies have suggested an association between Cancer diagnosis comorbidity and COVID-19 infection disparity and mortality in certain populations. For example, in Shakor et al.'s (2021) study, COVID-19 severity rate was slightly higher among cancer patients than diabetic patients (53% and 45%, respectively). Additionally, Parohan et al. (2020) conducted a Systematic Reviews and Meta-Analyses of 14 retrospective studies with 29,909 COVID-19 infected patients and 1,445 cases of deaths to examine the association between cancer, age, gender, diabetes, and hypertension and COVID-19 infection in various countries around the world. They found a significant association between cancer rates and higher risk of COVID-19 infection and mortality (pooled ORs 3.04, p < .001).

Similarly, in a recent multicenter, retrospective, cohort study published in the *Lancet Oncology*, Tian et al. (2020) used older patients (aged 18 and older) with cancer in China to examine the link between COVID-19 risk factors and COVID-19 severity in these patients. They found that, individuals with cancer had a greater chance of experiencing severe COVID-19 than their counterparts without cancer (OR = 3·61, p<0·0001). Moreover, Najjar et al. (2022) conducted a multicenter retrospective study to examine association between cancer and COVID-19 infection using 114 patients in Syria from March 28, 2020, to March 29, 2021. They conducted multiple logistic regression analysis and found that individuals with underlying medical factors had a higher risk for severe COVID-19 (OR = 2.814, *P* = .044). Moreover, Ssentongo's (2020) systematic review and meta-analysis study that assessed a link between preexisting comorbidities such as cancer and COVID-19 death using 25 studies with 65, 484 COVID-19 individuals in China Africa, and other countries found a higher probability of death in patients with cancer (1.47) than their counterparts without cancer. According to Parohan et al. (2020); Shah et al. (2021); Shakor et al. (2021), this is likely due to the fact that cancerous genes promote the destruction of white blood cells and therefore triggers poor immune response that permit COVID-19 progression and possible organ damage.

However, Li et al. (2022) conducted a Mendelian Randomization Study using 1,388,342 participants of whom 5,101 were very serious respiratory confirmed patients and 1,383,241 controls to assess the associations between cancers and COVID-19 outcomes or severity. They found no statistically significant causal association between cancer and COVID-19 infection risk for example, lung cancer (p = 0.60), breast cancer (p = 0.43, colorectal cancer (p = 0.85), and prostate cancer (*p* = 0.54).

Even though these studies have provided information on COVID-19 transmission and severity risk from Cancer diagnosis, cancer infection and COVID-19 infection characterization studies are limited in Indiana. In this study, I will examine the association between cancer diagnosis and COVID-19 infection among adults ages 18 and older in Indiana to provide better knowledge and clear information to improve COVID-19 prevention strategies.

Hypertension. Previous studies have found correlation between hypertension comorbidity and COVID-19 infection disparity in certain populations. For example, Parohan et al. (2020) conducted a Systematic Reviews and Meta-Analyses using 14 retrospective studies that entailed 29,909 COVID-19 individuals and 1,445 cases of deaths to assess the association between age, gender, and hypertension and COVID-19 infection in China and various countries. They found a higher association between hypertension and COVID-19 incidence and deaths (ORs 2.70, $p = .003$). Similarly, in Shakor et al. (2021)' s study, Hypertension was greatly linked with COVID-19 in most countries and spanned between (11.9% - 67.3%).

Moreover, Ssentongo (2020) did a systematic review and meta-analysis study that examined the association between preexisting medical conditions and COVID-19 deaths using 25 studies with 65, 484 COVID-19 patients in China and other nations. They found that people with hypertension had a greater chance of COVID-19 death than their counterparts without hypertension (RR 1.82).

Also, in Kinge et al.'s (2022) cross-sectional retrospective study that examined association between hypertension and COVID-19 disease in India using COVID-19 patients. They found a substantial significant link between hypertension and length of hospital admission in COVID cases ($P < 0.05$). Furthermore, 13 (3.8%) required intensive care.

According to Shah et al. (2021), these findings could be attributed to a poor sympathetic nervous system (SNS) functioning and defective parasympathetic nervous system which can trigger high blood pressure and promote inefficient immune response to combat COVID-19. This mechanism is known to induce oxidative stress, inflammation, rapid disease progression and vital organ damage.

Nevertheless, a recent meta-analysis study of 42 studies and 14138 patients published in the *International Journal of Hypertension*, that evaluated the link between hypertension and COVID-19 by Li et al. (2021) found lower incidence of hypertension in COVID-19 patients in China (OR 17.7) compared to the general population of the country. Also, the number of deaths in COVID-19 patients suffering from hypertension treated with ACEIs was much lower compared to that in nonuse patients treated with ACEIs/ARBs. The authors concluded that hypertension may have the potential to inhibit the COVID-19 risk but increase the risk of encountering severe clinical outcomes.

Overall, even though the independent/explanatory variable underlying medical conditions for example, **(**obesity (as measured by BMI $\geq$ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) has been shown to be a very common risk factor for serious COVID-19 disease in certain population, underlying medical conditions factor in COVID-19 patients has been less investigated in Indiana. Also, there are many controversies in the relationship between this factor and COVID-19 infection, and it is still unclear how underlying medical conditions contribute to COVID-19 infection among adults 18 years and older in Indiana, and it remains largely unknown whether this factor contributed to high risk of COVID-19 infection among my study population.

In this current study, I conducted cross sectional study to assess association between the independent/explanatory variable underlying medical conditions (for example, obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 infection to clarify this relationship and provide updated data related to the predictive ability of underlying medical conditions on COVID-19 infection.

The findings can help increase understanding of the COVID-19 and underlying medical conditions characteristics, the importance of controlling preexisting co-morbid risk factors in guiding and prioritizing COVID-19 vaccine resource allocation and interventions (Moazzami et al., (2020).

Community Level of Social-Ecological Model

Community level factors include county code of residence or neighborhood factors (Glanz et al.,2008).

County Code of Residence and COVID-19

Epidemiologic studies have linked various county code/Zip code/neighborhood-level risk factors to COVID-19 burden and potential disparities in COVID-19 incidence worldwide. For instance, Cuadros et al. (2021) conducted a comparative study to assess the COVID-19 incidence and deaths between urban and rural areas in the United States using credible COVID-19 data from Johns Hopkins University that entailed 3108 counties and 10,143,327 COVID-19 cases and 234,186 COVID-19 deaths information. They found a substantially higher number (8,613,630) of COVID-19 incidence cases in urban areas, and (1,529,697) incidence cases in rural areas. But COVID-19 incidence in rural counties increased dramatically by over 180% from the first to the second period of the pandemic onset (RR]: 2.86, $P < .001$), and by more than 80% spanning the second to the third period (RR: 1.88 $P < .001$). whereas in urban counties, the incidence rate only went up by 90% from the first to the second period (RR: 1.92, $P < .001$). Furthermore, in the third period, COVID-19 incidence rate jumped by over 25% in rural places compared to urban counties (RR: 1.27, $P < .001$). Additionally, a recent ecological cross-sectional study by Marko et al. (2020) that used data from New York City, Illinois, and the American Community Survey to assess the link between various social determinants of health elements and COVID-19 incidence rates in ZIP code/neighborhoods found that areas with substantially higher COVID-19 incidence rates in New York City were characterized as individuals in working-class and middle-income communities and comprised mainly of service workers and occupations such as essential fast food restaurant services for individuals with low education. However, areas with higher COVID-19 incidence rates in Chicago were characterized as Zip codes/neighborhoods with most vulnerable, poor, jobless, and non-Hispanic Black residents.

Furthermore, recently, Azar et al. (2020) performed a large retrospective cohort study of 1,052 adult cases out of California to evaluate the link between Race and socioeconomic factors and COVId-19 disease incidence in various areas and found that COVID-19 patients residing in wealthier neighborhoods had a lower chance than those residing in lower income areas to be hospitalized from COVID-19 (OR = 0.55, $p<0.001$). Similarly, in De Lusignan et al.'s (2020) Cross sectional study on 587 cases out of the United Kingdom that assessed race, poverty, and COVID-19 issues reported that individuals living in poor areas had a greater probability of getting COVID-19 (OR = 2.03, p<0.0001) than their counterparts from wealthier areas. Moreover, Oluyomi et al. (2021) conducted geospatial (GIS) analysis of neighborhood-level/zip code to census track analysis to examine the association between 29 neighborhood-level variables and COVID-19 incidence in Harris County, Texas. They found substantial associations between Blacks, foreign-born population areas, poor places, and people 65 years and older and COVID-19 infection incidence.

Similarly, Palacio and Tamariz (2021) used multiple logistic regression analysis to examine the association between race and COVID-19 by zip code in Miami-Dade County, South Florida. They found that places with high concentrations of Blacks had higher COVID-19 incidence compared with areas characterized by larger number of non-Hispanic White residents (38%). Also, in Unruh et al.'s (2022) recent retrospective cohort study out of Cook County, Illinois, a large number (164) Cook County Zip codes reported COVID-19 incidence and deaths. Moreover, 76 (46%) of the areas were inhabited mainly by Minorities. Likewise, in another recent cross-sectional study of United State counties published in the Journal of American Medical Association, Karmakar et al. (2021) used data sets from the Johns Hopkins University to look at the link between various county-level sociodemographic risk factors and COVID-19 incidence and deaths. They found that counties occupied by higher number of minorities encountered higher COVID-19 incidence (IRR, 1.03; $P < .001$) and death rates (IRR, 1.03; $P < .001$) compared to their counterparts. Moreover, substantial incidence rates were noted in counties with higher percentages of minorities including African American

(IRR, 1.02; $P < .001$), Hispanic or Latinx (IRR, 1.02; $P < .001$), American Indian or Alaskan Native (IRR, 1.01; $P < .001$), and Asian (IRR, 1.03; $P < .001$). The authors also found higher mortality rates in counties with higher percentages of African American groups (IRR, 1.02; $P < .001$), Hispanic or Latinx (IRR, 1.02; $P < .001$), and American Indian or Alaskan Native (IRR, 1.02; $P < .001$).

Moreover, Liao and De Maio's (2021) cross-sectional ecological study of 3,141 US counties that evaluated association between race and COVID-19 incidence and death found COVID-19 incidence rates to be substantially higher among counties with higher number of Blacks and a very large number of Hispanic residents (RR, 1.042). Hanson et al. (2020) examined the link between race and COVID-19 diagnosis in various zip codes in Indiana. They found that African Americans were four times more likely to experience COVID-19 incidence than their non-African Americans counterparts (OR = 4.00). Moreover, COVID-19 rates increased per 10,000 people in zip codes with substantially higher number of African American or Hispanic residents.

Furthermore, Samuels-Kalow et al. (2021) conducted negative binomial regression models on 9898 COVID-19 patients, using health care secondary data from Boston spanning February 5–May 4, 2020, to assess independent link between poverty, and the type of people that live in the neighborhood and the incidence of COVID-19. They found substantial association between Hispanic individuals (IRR = 1.25) and the number of poor people (IRR = 1.25) and substantial COVID-19 cases (Samuels-Kalow et al., 2021). In addition, areas with high COVID-19 cases had substantially higher number of Hispanic population (n = 72 census tracts), Black (n = 36), lack of insurance (n = 33), SNAP benefit recipients (n = 39), and individuals living in poverty (n = 23), ($P < .05$; Samuels-Kalow et al., 2021).

According to Emeruwa et al. (2020), this is likely because certain COVID-19 risk factor elements might be concentrated within geographic areas. For example, persons living predominantly in Black neighborhoods with poor infrastructure and high poverty might be more likely to work in low paying high-risk areas where they could become infected and crowded homes household members could become infected by others. Similarly, Woolley et al. (2022) summed it up by saying residing in a poor area and an overcrowded home can inhibit public health interventions for COVID-19 such as social distancing and therefore this can increase individual's vulnerability to contracting the virus.

However, Little et al.'s (2021) retrospective cohort study in New York City used 3,528 COVID-19 confirmed individuals between March 1, 2020 and April 1, 2020 to look at the linkage between poverty level and COVID-19 hospitalization and other characteristic such as death. They found that individuals with COVID-19 who resided in high poverty places seem to be substantially younger, suffer from underlying medical conditions and were more likely to be of female relative to their counterparts residing in low poverty areas. They also did not find a relationship between living in a high poverty area and higher COVID-19 hospital admission. In this study the authors also found a linkage between poverty and lower hospital death.

Although these epidemiologic studies have yielded some information on association between county code/Zip code/neighborhood factor and COVID-19 infection, some of them provides conflicting results. Also, the studies did not look at county code/Zip codes /neighborhood and COVID-19 infection among adults 18 years and older in Indiana, so clear information on association between county code of residence and COVID-19 infection studies in Indiana are limited (Liao and De Maio, 2021; Little et al., 2021). Little et al. (2021) suggested further studies to fully characterize Zip code/neighborhood and COVID-19 infection in various settings. Therefore, I conducted a cross sectional study to investigate the link between county code of residence and COVID-19 disease incidence among adults ages 18 years and older in Indiana to effectively inform the ongoing public health response to the COVID-19 pandemic.

Definitions

COVID-19 infection: A respiratory disease resulting from activities of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2; CDC, 2021).

COVID-19 positive case: Refers to Covid positive Patients who were confirmed for COVID 19 infection through RT-PCR, RT-PCR swab (CDC, 2021; WHO, 2021)

COVID-19 diagnosis/current status (outcome or dependent variable): Includes any individual with laboratory confirmation of SARS-CoV-2 infection by real time reverse transcription–polymerase chain reaction (RT-PCR) test, and probable case (i.e., no known positive SARS-CoV-2 test result; (CDC, 2021; Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; WHO, 2021). It is categorical and categorized as 1= laboratory confirmed case, 0 = probable case.

Age group: Defined as the time the study participants were born to the time of their admission in a hospital setting and received a diagnosis for COVID-19. Age group is in years, categorical and categorized as 0 = 18-49 years; 1 = 50-64 years; 2 = 65 years plus; 3 = unknown; 4 = missing (when not provided); 5 = NA. Age was measured as a categorical ordinal variable (Casas-Deza et al., 2021, CDC, 2021).

Sex: Sex of the patient is defined as male or female including the social norm and roles that influence the patient's health behavior and exposure to diseases. This was categorized as 0 = female; 1 = male; 2 = other; 3= unknown; 4 = missing (when not provided); 5 = NA in the study. Sex was categorical or dichotomous variable (Abate et al., 2020; CDC, 2021; Ronquillo De Jesús et al., 2022).

Race: Race of patient refers to the patient's reported identified race at the time of data collection. This was grouped as: 0 = Black, 1 = White, 2 = Asian, 3 = American Indian/Alaskan native, 4 = missing (when not provided), 5 = unknown, 6 = Native Hawaiian/other Pacific, 7 = multiple/other, 8 = Islander, and 9 = NA (Boserup et al., 2020; CDC. 2023).

Exposure status: The reported exposure due to employment or workplace status of the patient at time of data collection. This was stratified as 0 = unknown, 1 = yes, 2 = missing (when not provided; CDC, 2023; Fan et al., 2020; Wang et al., 2020).

Underlying medical conditions: The reported comorbidity for example, obesity: BMI > 30 kg/m^2 , diabetes, the malfunction of the pancreas and insulin malfunction in the patients; cancer diagnosis, the patients diagnosed with abnormal cell growth and proliferation; Hypertension entail high blood pressure (CDC, 2023; Gimeno-Miguel et al., 2021; Hamza et al., 2022; Mankowski et al., 2021; (Shakor et al., 2021; Parohan et al., 2020), and was assessed as 0 = no, and 1= yes (CDC, 2023).

County code of residence/ neighborhood of residence (county FIPS code): Refers to the county area where the COVID-19 disease was contracted or transmitted either by direct or indirect contact with a case, contaminated object or surface, or some other mechanisms (CDC, 2023; Maroko et al., 2020). This community level factor was coded as Lake County = 18089, Marion County = 18097 and was continuous variable (CDC, 2023).

Social determinants of health: Entail conditions into which an individual is born, grows, lives, works, and ages, and they look at the person (CDC, 2022; Salgado de Snyder et al., 2021; Singu et al., 2020).

Health equity: Refers to the opportunity for everyone to have fair, just opportunity or equal access health amenities and to be as healthy as possible, through the elimination of various health obstacles or disparities for example, poverty, discrimination, lack of health care, and stigma (CDC, 2022).

Secondary data analysis: Refers to the comprehensive analysis of data that was collected by another entity of someone other person for another primary purpose (Creswell, 2009).

Assumptions

I assumed that the current COVID-19 disease data in the COVID-19 dashboard that was obtained from hospitals in the local communities represented the scope of COVID-19 disease in Indiana during the ongoing pandemic and were not misdiagnosed as seasonal flu disease. Since this is the first time, we have encountered both the flu seasonal disease and a novel worldwide COVID-19 pandemic (CDC, 2022). Some previous epidemiologic studies have emphasized the presence of similar symptoms between COVID-19 disease and the common seasonal flu, and thus performing similar tests for the two, many result in COVID-19 misdiagnosis (CDC, 2022; Dunbar, 2020; Ossei et al., 2021; Sieber et al., 2021). The second assumption was that the number of the COVID-19 confirmed cases was truly derived from individuals confirmed by COVID-19 RT-PCR laboratory testing, COVID diagnosed through nasopharyngeal swab testing, sputum, feces or blood samples, chest computed tomography (CT), and nucleic acid test, performed using the fast-yielding and reliable polymerase chain reaction (PCR) method (CDC, 2021; Abebe et al., 2020; Aldila et al., 2021; Gralinski & Menachery, 2020; Hatipoğlu, 2020; WHO, 2021). A third assumption is that people who did not get tested and confirmed through the RT-PCR, RT-PCR swab were not included in this particular data set, since certain disadvantaged individuals may encounter health barriers such as lack of proper COVID-19 knowledge and widespread misinformation, and lack of health insurance, (Dunbar, 2020; Ossei et al., 2021; Sieber et al., 2021). Fourthly, since the data have been validated by the Marion County Health Department, The Indiana Department of Health, the CDC, and the WHO, I also assumed that timely and correct mechanisms of data collection from patients and recording processes was employed.

Scope and Delimitations

I only focused on social determinants of health (age group, sex, race, exposure, county code/zip code of residence), underlying medical conditions (obesity (as measured by BMI > 30 kg/m$^{2)}$, prostate cancer diagnosis, diabetes, and hypertension) risk factors for COVID-19 characteristics among adults ages 18 and older in Indiana and omitted information on other populations in the state of Indiana. It was not within the parameters of this study to draw cause and effect relationship, so I did not conclude that the social determinants of health and the preexisting comorbidity disparities exhibited among the study population resulted in the COVID-19 disease. Also, since I used COVID-19 data collected during the ongoing pandemic in Indiana, the study findings will only be generalized to the population of Indiana and may likely not be generalized to outside states or nations that have suffered from the COVID-19 pandemic. The other states may vary in some social determinants of health and preexisting comorbidity prevalence and influences. Additionally, due to the detrimental effects of the COVID-19 disease and the widespread uncertainty and fear linked to the disease because of a lack of epidemiology knowledge of the characteristics of COVID-19 disease, including the critical and urgent need to enlighten local community residents about the disease and COVID vaccine, the COVID-19 data were collected under harsh conditions. Therefore, the data may not be complete. Moreover, as Maroko et al. (2020) noted the COVID-19 data may not fully reflect the incidence of COVID-19 cases due to inherent biases in COVID-19 testing problems. During the pandemic, testing for SARS-CoV-2 has also produced high numbers of false negatives or false positives, which may be attributed to the lack of a more sensitive rapid testing at the beginning of the pandemic. Furthermore, some individuals may have been asymptomatic and remained as potential carriers of the disease in local communities unknowingly. Thus, such problems can hinder the reliability and accuracy of the data and hinder area comparison.

Limitations

Some notable limitations of the study are that it employed secondary data which encompass participant self-report that can promote biased results of major social determinant factors (Alam et al., 2021; Parker et al., 2021). Also, utilizing secondary data has the potential to hamper the researcher's ability to choose from a wide range of vital variables and so can hinder the strength of the data analysis (Parker et al., 2021) Moreover, employing preexisting data may have the ability to produce results that may not foster generalizability of the data to other population and therefore could vitally threaten the study's external validity (Creswell, 2009; Palacio & Tamariz, 2021). Additionally, the credibility of the result is contingent on the accuracy and quality of the primary data that was collected for the studies (Moazzami et al., 2020). Moreover, COVID-19 disease is novel and had never been diagnosed in Indiana before the current pandemic, so the data at the start of the pandemic may be missing accurate information on some of the people in the database (Abedi et al., 2021). Data on racial demographics may be hard to navigate since some patients may have been reluctant to identify their race which may hinder data on race. During the pandemic, COVID-19 data were reported by Indiana County code/ZIP codes and not by block groups or census tracks and therefore this could lead to misclassification (Marion County Public Health Department (MCPHD), (2023); Palacio & Tamariz, 2021). Another noticeable limitation of the study stems from sample selection bias due to solely focusing on adults aged 18 years and older. Thus, this may hinder generalizability of the result to the whole of Indiana's population (Parker et al., 2021). Additionally, the cross-sectional design cannot establish cause and effect linkage, so I could not conclude that the social determinants of health factors, and the underlying medical conditions led to COVID-19 infection (Creswell, 2009). The COVID-19 primary data were gathered during high COVID-19 emergency conditions while the primary focus was on disease scrambling to produce a viable COVID-19 vaccine, treatment regimen, contact tracing, and controlling the spread of the disease in various local community in Indiana and elsewhere, therefore the COVID-19 information may be half hazard (CDC, 2021; WHO, 2021). Though I used public data, I needed permission from the CDC to access the dataset for use, after obtaining approval from the Walden University Institutional Review Board (IRB). Furthermore, it was vital to carefully examine subject's data to ensure that it does not entail participant's personal data that can publicly endanger their COVID-19 confidential records (Creswell, 2009). Creswell (2009) recommend taking measures such as increasing the sample size of a study to effectively address sample representativeness, generalizability, and threat to the study's external validity.

Significance

The current study is significant since a clear data is scarce on the link between key social determinants of health factors (age, sex, race, exposure, county code of residence), underlying medical conditions (for example, obesity) and COVID-19 disease infection incidence in adults aged 18 years and older in Indiana. Moreover, most of the studies have provided inconsistent or conflicting results and reports on the incidence of COVID-19 infection and some social determinants of heath factors and preexisting comorbidities in the general population including adults 18 years and older in the United

States and around the world (Abedi et al., 2021; Alam et al., 2021; Khanijahani et al., 2021; Upshaw et al., 2021). The study findings may increase COVID-19 characteristics understanding and knowledge in local communities in Indiana state and clarify the relationship between social determinants of health factors and preexisting comorbidities associated with COVID-19 incidence that can contribute to better targeted disease control measures, and also help guide pandemic mitigation strategies for Indiana adults aged 18 years and older (Hanson et al., 2020; Maness et al., 2021). In terms of public health practice, the sound evidence provided in this study may enhance policymakers, and public health professionals' ability to combat the COVID-19 challenge in vulnerable adults 18 years and older in Indiana and elsewhere (Hanson et al., 2020; Maness et al., 2021). Also, the results may be utilized to foster multidisciplinary collaboration among various public health among health care, and state entities to address COVID-19 health issues (Hanson et al., 2020; Maness et al., 2021; Tipirneni, 2021). Because social determinants of health factors such as age and preexisting comorbidities such as obesity have been shown to impact COVID-19 outcomes, it is important to fully comprehend and characterize the relationship between these predictors and COVID-19 infection (Samuel et al., 2021). By doing so, potential COVID-19 interventions may support positive patient outcomes at the individual, community, and societal levels, contributing to positive social change. Additionally, the findings of this study can contribute uniquely, by the application of the multiple levels of influence found in McLeroy's SEM to multiple COVID-19 incidence in Indiana and elsewhere that can promote evidence-based testing, prevention, and control strategies and health policy recommendations that also have tremendous implications for improvements in population health and reducing COVID-19 during this pandemic or future disease outbreaks.

Summary and Conclusions

From the literature, I identified pertinent research methodologies such as qualitative techniques, quantitative techniques, and various systematic review and meta-analysis studies employed to investigate health outcomes such as COVID-19 infection (De Lusignan et al., 2020; Karmakar et al., 2021; Liao and De Maio, 2021; Ssentongo, 2020)

In addition, the review showed epidemiologic studies that successfully used the SEM, established and proposed by McLeroy, and colleagues' (1988) as a robust framework for explaining the multiple factors that influence people's health in a community at the individual, community, and policy levels rather than just focusing on the factors that affect individuals at the individual level (Christian et al., 2020; Glanz et al., 2008; Noh and Min, 2020). The various results from this extensive review confirmed and strengthened Glanz et al.'s (2008) recommendation for employing the socio ecological model framework for epidemiological and evidence-based research to comprehend diseases such as COVID-19 infection characteristics among a population such as adults ages 18 years and older in Indiana. Additionally, the robust review of the pertinent and credible literature helped guide the scope of the four defined research questions for the current study.

Also, as discovered in the literature review, it could be acknowledged that an independent variable could affect COVID-19 incidence on different levels within

McLeroy et al.’s (1988) SEM framework (for example, individual, community levels social determinants of health and underlying medical conditions, such as age group, obesity, and county code/zip code of residence) in the United States, particularly Indiana (Glanz et al., 2008; Hanson et al., 2020; Maness et al., 2021; McLeroy et al., 1988; Tipirneni, 2021). Previous epidemiologic studies including Khanijahani et al. (2021); Mankowski et al. (2021), mentioned that adults and those with preexisting comorbidities are more likely to get sick and hospitalized with COVID-19 than younger individuals.

This section covered characteristics of the COVID-19 pandemic globally, including Indiana, an area that has one of the highest burdens of the disease in the Midwest region of the USA. I also clearly and systematically presented relevant information characterizing COVID-19 infection incidence and its associated social determinants of health and preexisting comorbidities in this review that encompassed keywords, and the pertinent inclusion and exclusion criteria. Furthermore, this section addressed the pressing study problem at hand, the purpose of the study, the theoretical foundation, the four research questions and hypotheses. The section also included the relevant definition of major terms, associated assumptions, the scope and delimitations, notable limitations, and the significance of the study and implications for positive social change.

The next section, research design and data collection described the relevant methodology that was employed to analyze the secondary data from the CDC. The section also discussed the specific study population, sampling and designated sampling procedures, study design, and the ethical considerations in conducting this current study.

Section 2: Research Design and Data Collection

Introduction

The purpose of this quantitative study was to examine relationships between social determinants of health (age group, sex, race, exposure, county code of residence) and underlying medical conditions (obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, and hypertension) and COVID-19 diagnosis among adults 18 years and older in Indiana, using the SEM framework. Gaining better knowledge and information on the relevant elements that are linked to COVID-19 infection and its spread is crucial, causing the need to embark on this study. The findings could help to properly improve the COVID-19 public health control and intervention program agenda and policy in Indiana and various areas of the globe.

In this section, I address an overview of the study design and rationale, the study population and locations, and the specific sampling and sampling procedures utilized to gather the secondary data, which entailed the calculation of the sample size or power analysis. I also present the research questions and hypotheses, variables (dependent and independent variables), and data management. Furthermore, I provide information pertaining to the instrumentation and operationalization of the current study variables, the specific data analysis plan, the various threats to validity, and the ethical procedures of the study.

Research Design and Rationale

I conducted a quantitative, descriptive, retrospective cross-sectional study to analyze secondary data, the COVID-19 Case Surveillance Public Use Data with

Geography from the CDC collected during the current COVID-19 pandemic on patients admitted to various hospitals and health clinics in Indiana, to fully investigate the association between social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension), and COVID-19 disease diagnosis status among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

A cross-sectional study design served my study best because such a design increases the robustness of a study and provides the opportunity to examine associations between predictors and outcome at a unique juncture in time while gauging a population parameter such as the proportion of some disease, including COVID-19, and its affiliated social determinants of health in a community utilizing robust descriptive statistics, chi-square, and appropriate logistic regression statistical methods (Creswell, 2009). Furthermore, a cross-sectional study design provided the opportunity to answer the four evaluative and descriptive research questions for characterizing COVID-19 in this designated study population, demonstrate the usefulness of the various aspects of the ecological model, and produce more generalizable results (Creswell, 2009). Using a cross-sectional study design can save time and money when the primary data are already available through the collection organization or site. Moreover, the design strengthens study reliability and validity, as Creswell (2009) and Frankfort-Nachmias and Nachmias (2008) noted.

Additionally, previous epidemiological studies (Cruz-Arenas et al., 2021; Kinge et al., 2022) have successfully utilized a quantitative retrospective cross-sectional study and secondary data mechanisms to assess the contributions of risk factor elements to various infectious diseases such as COVID-19 in vulnerable populations and communities.

Methodology

Population

The target population for this study included all cases of women and men 18 years and older in the CDC COVID-19 Case Surveillance Public Use Data with Geography data set with a confirmatory COVID-19-positive reverse-transcriptase-polymerase-chain-reaction (RT-PCR) blood test result, and probable case during the current COVID-19 pandemic in Indiana. These specific populations were at a greater risk for COVID-19 disease, likely due to poor COVID-19 health practice and their experience with unfavorable social determinants of health and underlying comorbidities (Dixon et al., 2021; Hanson et al., 2020; Khanijahani et al., 2021; Samuel et al., 2021; Timsina et al., 2020). For this study, I used all the individuals in the data set who met the selection criteria (target population size approximately 170,124 cases). I conducted a cross-sectional retrospective research study in Indiana, one of the smallest states in the Midwest region of the United States, with a population of about 6,785,528 (Figure 1), in which most of the population consists of vulnerable adults, who number 5,116,486 (76%; Indiana Department of Health, 2023; MCPHD, 2023; U.S. Bureau of Labor Statistics, 2023). Indiana's population can be characterized as 2.7% Asian, 7.7% Hispanic, 84.2%

White, 10.2% Black, and 2.3% two or more race groups (Indiana Department of Health, 2022). The state has a high poverty rate of about 14.1% and a median household income of only about $56,185. Only 27.8% of individuals have a bachelor's degree or higher (Indiana Department of Health, 2022). Indiana also has a large, underserved population with a high uninsured rate of 9.1% and a CDC Social Vulnerability Index (SVI) of 0.7971 (The SVI measures factors that hinder a community's ability to combat issues such as poverty, poor transportation, and congested housing), and normally higher SVI scores indicate increased vulnerability. (CDC, 2021; Indiana Department of Health, 2022; Samuels-Kalow et al., 2021). The COVID-19 Community Vulnerability Index (CCVI) is 0.94 (CDC, 2022; Indiana Department of Health, 2022). Finally, the COVID-19 virus has never been diagnosed in Indiana before the current and ongoing pandemic. These trends may contribute to the high burden of COVID-19 infection in the state and make this area and population paramount focus (Indiana Department of Health, 2022).

Sampling and Sampling Procedures

The utilization of a secondary data analysis method to address pertinent research questions in health science research has gained prevalence over the years due to the availability of vast amounts of data collected and preserved by various organizations and researchers (Creswell, 2009). However, Creswell (2009) recommended that researchers carefully evaluate already-existing data for data integrity and appropriateness to ensure that the variables can logically answer their research questions and that the study hypotheses are suitable (for example, by carefully examining the purpose of this study and who was directly responsible for collecting the information) (Creswell, 2009). This information can be obtained from the investigators from the primary study. Secondary data can be provided by private or public entities, and criteria for accessing a secondary data set are determined by these data-collecting entities (Creswell, 2009).

For this study sample, I included all the available respondents who met the designated inclusion criteria for the study because the entire data set was available for rigorous analysis. Moreover, using all the appropriate respondents who met the study criteria yielded and strengthened the methodological rigor, power, and population validity and increased the statistically significant difference between groups, as noted by Creswell (2009) and Frankfort-Nachmias and Nachmias (2008).

Sample Frame

The sampling frame for this study was the list of all men and women ages 18 and older who participated in the COVID-19 Case Surveillance Public Use Data with Geography surveillance during the current COVID-19 pandemic surveys. For the inclusion criteria, a participant was selected for the sample if they identified as being female or male, were aged 18 and older, had been tested through (PCR) for the COVID-19 virus and had a COVID-19 diagnosis confirmatory result, probable case, and had been treated in one of the hospitals or clinics in Indiana during the current COVID-19 pandemic. Inclusion of all the eligible and high-risk participants aged 18 and older in this study helped increase the methodological rigor of this study due in part to their experience with unfavorable social determinants of health and the presence of comorbidities (Hanson et al., 2020; Maness et al., 2021; Samuel et al., 2021). I restricted my analysis to all COVID-19 laboratory-confirmed cases and probable cases reported for which data were available on the specific variables needed for this study (age group, sex, race, exposure, county code of residence, and underlying medical conditions). The exclusion criteria applied to all patients with hospital or clinic data before the start of the pandemic in 2019, since the COVID-19 pandemic was first noted in 2019.The exclusion criteria also pertained to information on data regarding seasonal flu and COVID-19 misdiagnosis for individuals who visited hospitals or clinic from 2019 to 2022. Since the common flu and COVID-19 exhibit similar symptoms which may lead to COVID-19 misdiagnosis problem for some patients, and this could hinder the result (WHO, 2021). The study sample encompassed all the participants who met all the pertinent inclusion criteria established for this study. Creswell (2009) noted that inclusion of all eligible participants in a study can improve the study results' methodological rigor and thus the study results.

This study was a secondary analysis of data from the 2019–2023 COVID-19 Case Surveillance Public Use Data with Geography, which was provided by the CDC. A cross-sectional survey was conducted from 2019 to 2023 to enable Indiana to monitor key indicators such as those related to COVID-19 health status and the linked social determinants of health and comorbidities. In the COVID-19 Case Surveillance Public Use Data with Geography, data were collected from a representative sample of women and men using a standardized survey questionnaire through face-to-face interviews.

Power Analysis/Sample Size Calculation

Conducting power analysis enables researchers to determine the smallest sample size that is suitable to detect the effect of a given test at the desired level of significance and can increase the methodological rigor of epidemiological studies (Creswell, 2009). Creswell (2009) noted that in general, increasing sample size potentially increases the power of a study or the ability to detect a specific effect size within a sample. Furthermore, the sample size increases when the power of the study is increased from 80% to 90% or 95%, and this also enhances the likelihood of detecting important differences between study groups. It is important to note that larger samples may not improve robustness (Creswell, 2009).

Based on my quantitative research questions and the designated cross-sectional study design, the widely utilized open-source statistical power application or software G*Power software (3.1.9.7; Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany; https://www.gpower.hhu.de/) was the most appropriate for the current study (Faul et al., 2007; Kang, 2021). With it, I was able to calculate the study's sample size and conduct the relevant power analysis of the study population to yield the correct and needed sample size for the study. G*Power is profoundly useful for sample size and power calculations since it also takes into consideration (F, t, χ^2, Z, and exact tests). It is free and very easy to use (Faul et al., 2007; Kang, 2021). Moreover, it provided the opportunity to present results of the link between the independent variables and the dependent variable, COVID-19 diagnosis, via graphs, in addition to calculating the appropriate effect size and required minimum sample size based on the effect size as well as the number of predictors in the study considering the alternative hypothesis (Faul et 2007). Also, importantly, I used backward elimination, logistic regression, and hierarchical regression model statistical tests for my study, contingent on the number of dependent (COVID-19 diagnosis) and independent variables (variables exerting influence on COVID-19 infection) utilized in my study and their appropriate levels of measurement dichotomous/binary, so it was vital to calculate the sample size using methods that encompassed effect size and helped to ensure sufficient power and sample size (Faul et al., 2007).

I calculated the required minimum sample size and power analysis for this study, applying effect size (0.15), alpha (0.05), power (0.80 or 80%), and designated number of predictors (6), yielding about 582 as the total sample size needed for my study and thus, fostering methodological rigor and credible and reliable study results, in addition to ensuring efficient resource and time use (Faul et al., 2007; Kang, 2021). Eventually, this led to and ensured a rigorous study design that was free of biased and misleading results (Faul et al., 2007; Kang, 2021).

According to Creswell (2009), quantitative methods require obtaining appropriate samples to improve study population representativeness, provide unbiased results, and increase study robustness. Therefore, sample size selection is vital for successful research (Creswell, 2009).

For this study, it was appropriate to use a systematic sampling method to choose the sample. It required selecting every Kth individual or sampling unit of the population after the first subject or sampling unit was selected at random from the total of sampling units. The designated formula is K = N/n, that is, population size/sample size (Creswell, 2009). Preliminary data from the CDC indicated that 2,065,880 confirmed cases of COVID-19 and 4,666,339 probable cases had occurred in Indiana by April 12, 2023

(CDC, 2023). I utilized these figures as my population size for this study. Therefore, for example, I wished to select a sample size of 150,000 persons from the population size of 6,732,219. Because my G-power analysis showed that a total sample size of about 582 was required, I could do numbers 582 and above. To select 150,000 people, my calculation was the following: K= *N*/*n* = 6,732,219/150,000 = 45. Therefore, I picked every 45th individual. The first selection is generally determined by a random process, such as the use of a table of random digits or a prepared data sheet. So, if the 10th person is selected; the sample would then consist of individuals numbered 10, 55, 100, 145, 190, and so on. The results from these calculations indicated that my sample size estimate compared to the codebook for my secondary data set was sufficient to be able to answer my research questions. A sample size of approximately 150,000 sufficed to establish a link between the predictors and the outcomes. A random sample size of approximately 150,000 respondents could be drawn for the study.

A systematic sampling method is conducive to use with very large populations such as mine or when large samples are to be picked (Creswell, 2009). Each case in SPSS is already numbered so it is easier to pick each Kth individual or case.

According to Creswell (2009) and Kang (2021), five key elements should be noted during sample size analysis, notably the statistical test for the study, independent variables, effect size, power (1-β), and significance level (α), because they are pivotal for sample size calculation and the power analysis process. The one dependent and multiple independent variables that influenced the outcome variable used in the study and their various levels of measurement helped determine the appropriate statistical test for this study (Creswell, 2009; Frankfort-Nachmias & Nachmias, 2008; Faul et al., 2007). Additionally, the effect size was based on Cohen's suggestions. The main dichotomous predictor age group was based on literature (Casillas et al., 2021; Creswell, 2009; Kang, 2021).

Procedures for Recruitment, Participation, and Data Collection

Data Accessibility and Permission

After my proposal defense, I requested permission from the Walden University IRB office to use the secondary dataset (COVID-19 Case Surveillance Public Use Data with Geography) from the CDC for the COVID-19 pandemic in Indiana.

The dataset (COVID-19 Case Surveillance Public Use Data with Geography) from the CDC is readily available for public use and can be downloaded from the page https://data.cdc.gov/ directly using the Export function on the page. I did not need to complete and submit the Registration Information and Data Use Restrictions Agreement (RIDURA) to provide the necessary information to proceed with my request for access. However, I submitted a written request to the CDC for access to the dataset. I was allowed through email to utilize the dataset. Once the data was available, I abstracted individual data for all eligible participants.

Data Sources

I used a secondary dataset (COVID-19 Case Surveillance Public Use Data with Geography) from the CDC for this study. It is the most current, accurate, and credible data available for public use (CDC, 2023). I extracted the variables (age group, sex, race, exposure, county code of residence, underlying medical conditions, and outcome

(COVID 19 Diagnosis/current status) from the database used in this study. The CDC established the COVID Data Tracker in 2020 to track COVID-19 and monitor its effect on the various states in the United States and other countries (CDC,2023). It also provides government entities, researchers, public health officials, and communities members with the COVID-19 data in one simple and consolidated web-based system (CDC, 2023). Since the beginning of the pandemic, the COVID Data Tracker has widened to include vast entities such as county, state, and global data into over 70 webpages of pandemic-related information (CDC, 2023). COVID-19 identified at state hospitals or clinics are reported to the CDC through this web-based system. Demographics such as age, gender, place of residency, and risk factor–specific information were collected using standardized case report forms through interviews conducted face to face in private by states for example (CDC, 2023). Then the CDC evaluates the data to ensure high data quality and integrity and to also ensure there are no duplication of reported cases. If they find any duplicate records, a single record was kept for each duplicated subject (CDC, 2023).

Instrumentation and Operationalization of Constructs

I utilized secondary data collected for COVID-19 pandemic public health response and surveillance efforts to determine the risk factors associated with the COVID-19 pandemic in Indiana. Data for the COVID-19 Case Surveillance Public Use Data with Geography secondary dataset survey was collected using trained personnel who complete case investigation forms. Individuals including patients or family members were interviewed utilizing standardized questioner form and measurement procedures that entailed similar survey questions and gathered pertinent information on patient's health records and other vital information such as age group, sex, race, exposure, county code of residence/Zip code, underlying medical conditions, and COVID-19 laboratory test results (Indiana Department of Health, 2022). This COVID-19 information are then reported by the state to the CDC through the COVID Data Tracker web-based system (CDC, 2023; Indiana Department of Health, 2022). The COVID-19 infection testing was performed utilizing COVID-19 specific RT-PCR testing in hospitals and clinics in the state of Indiana to diagnose COVID-19 (Indiana Department of Health, 2022). The dependent variable for the study, COVID 19 Diagnosis is binary, thus, the current study's reliability and validity was assessed utilizing Chi-square Cramer's V correlations test that can provide appropriate strength of the relationship between the explanatory and the outcome variables.

Operationalization for Each of the Variables

Tables 1 and 2 show the variables that were utilized in this cross-sectional secondary analysis of the COVID-19 Case Surveillance Public Use Data with Geography dataset from CDC (CDC, 2023; Indiana Department of Health, 2022), showing the designated variable names, variables labels, and level of measurement. The independent variables employed in this study included social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (for example, obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension). The dependent/ outcome variable is the COVID-19 Diagnosis.

Dependent/Outcome Variable: COVID-19 Diagnosis

From the secondary data that I obtained for the current study, this included laboratory-confirmed cases for COVID-19 RNA, and probable cases (i.e., no known positive SARS-CoV-2 test result; CDC, 2023). The Dependent/ Outcome Variable was dichotomized into laboratory-confirmed case, and probable case (CDC, 2023; Samuel et al., 2021) and was coded as 1 = laboratory-confirmed case, and 0 = Probable case.

Independent/Explanatory Variables

Individual-Level Factors

Based on the literature (Khanijahani et al., 2021; Samuel et al., 2021), the individual level was coded as follows:

Sex. Sex of the patient was categorized as 0 = female, 1= male, 2 = other, 3 = unknown, 4 = missing (coded to missing when not provided), 5 = NA in the study, and sex was recorded as a categorical variable.

Age Group. Age is for the patient's years of life at the time data collection for surveillance and research purposes. Age of the patient was categorized into six groups, 0 = 18-49 years; 1 = 50-64 years; 2 = 65 + years, 3 = unknown, 4 = missing (coded to missing when not provided), 5= NA. Age was measured as a categorical and ordinal variable.

Race. Race of patient refers to the patient's reported identified race at the time of data collection. This was grouped as: 0 = Black, 1 = White, 2 = Asian, 3 = American Indian/Alaskan native, 4 = missing (coded to missing when not provided), 5 = unknown, 6 = Native Hawaiian/other Pacific, 7 = multiple/other, 8 = islander, 9 = NA. It was nominal.

Exposure. Exposure is the reported exposure due to, for example, employment or workplace status of the patient at time of data collection. This was stratified as 0 = unknown, 1= yes, 2 = missing (coded to missing when not provided). It was nominal.

Underlying Medical Conditions. Underlying medical conditions refer to the reported presence of comorbidity for example, obesity: BMI $\geq$ 40 kg/m^2, diabetes, cancer diagnosis, Hypertension, and was assessed as 0 = no, and 1= yes (CDC, 2023).

Community-Level Factor: County Code of Residence

County code of residence entails the specific location where the COVID-19 disease was contracted through direct or indirect mechanism with a case or contaminated material. Based on the data set, the community level factor was categorized as Lake County = 18089 and Marion County = 18097 (CDC, 2023).

Table 1

Independent Variables Analyzed in This Cross-Sectional Study

Name of variable	Variable definition/coding	Level of measurement
Sex (Gender)	Gender of patient 0 = female, 1 = male, 2 = other, 3 = unknown, 4 = missing, and 5 = NA	Nominal
Age group	Age group of patients 0 = 18–49 years, 1 = 50–64 years, 2 = 65+ years, 3 = unknown, 4 = missing, 5 = NA	Ordinal
Race	Race of patient 0 = Black, 1 = White, 2 = Asian, 3 = American Indian/Alaskan Native, 4 = missing, 5 = unknown, 6 = Native Hawaiian/other Pacific, 7 = multiple/other, 8 = Islander, 9 = NA	Nominal
Exposure	Exposure status of the patient due to, for example, employment or workplace 0 = unknown, 1 = yes, 2 = missing	Nominal
County code	County ZIP code/neighborhood of residence 18089 = Lake County 18097 = Marion County	Continuous/scale
Underlying medical conditions	The reported comorbidity, including obesity: BMI ≥ 40 kg/m^2 0 = no, 1 = yes	Nominal/dichotomous

Table 2

Definition and Measurement of Dependent Study Variable

Name of variable	Variable definition/coding of variable	Measure of variable
COVID-19 diagnosis	COVID-19 diagnosis 1 = laboratory-confirmed case and 0 = probable case	Categorical/dichotomous/nominal

Data Analysis Plan

Secondary data from the CDC (COVID-19 Case Surveillance Public Use Data with Geography) database was utilized as an updated and credible source to conduct this study analysis. Then I analyzed the designated variables in (Table 1) employing IBM© SPSS© Statistics version 27 (IBM© Corp., Armonk, NY, USA). (Green & Salkind, 2014; Wamer, 2021) to thoroughly examine the dataset and assess whether there was a linkage between COVID-19 diagnosis incidence and the independent variables social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (obesity (as measured by BMI $\geq 40\ \text{kg/m}^2$)), prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 Diagnosis among adults ages 18 years and older during the COVID-19 pandemic in the state of Indiana.

Descriptive statistics in the form of frequencies and percentages were used to describe the sample against the independent/explanatory variables of the study and show the sociodemographic characteristics of respondents. Chi-square tests for the categorical variables were used to examine differences by county code of residence/ Zip code/neighborhood. The association between (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity (as measured by BMI $\geq$ $40\ \text{kg/m}^2$)), prostate cancer diagnosis, diabetes, and hypertension), and COVID-19 diagnosis was assessed using logistic regression models. All the variables that demonstrated statistical significance in the univariable analysis were adjusted for in the applicable multivariable logistic regression models ($p < 0.05$), and the results were documented as odds ratio (OR) with 95% confidence intervals (95% CI) (Green & Salkind, 2014).

I further analyze the data using a backward logistic regression model and a two-block hierarchical logistic regression test to examine associations of COVID-19-related variables (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity (as measured by BMI $\geq 40\ \text{kg/m}^2$)), prostate cancer diagnosis, diabetes, and hypertension), and COVID 19 Diagnosis (Green & Salkind, 2014).

Data Management

After receiving the COVID-19 database, I identified the designated variables for this study. Then I used the statistical software IBM© SPSS© Statistics version 27 (IBM© Corp., Armonk, NY, USA). (Green & Salkind, 2014; Wamer, 2021) to carry out data cleaning and check for data accuracy, and missing data, and to ensure that the dataset was useful for this study. After this, I performed a study analysis on the complete data set that met my study criteria. In addition, I created dummy variables employing the transform method in SPSS to categorize variable levels for variables with more than two levels to be able to conduct backward stepwise elimination multiple logistic, and hierarchical logistic regression analysis appropriately. I recoded and created a dummy variable for the dependent variable COVID-19 diagnosis into a new variable current_status1. The new variable had two groups and was coded as 0 for probable cases and 1 for laboratory-confirmed cases (Field, 2013; Green & Salkind, 2014). Furthermore, I recoded and created a dummy variable for the independent variable age group into a new variable neagegroup with two groups (that included two age groups 18-64 years and 65 + years (Hamza et al., 2022). These two age group variables were coded as 0 = 18-64, and 1 = (65 + years). I also recode the independent variables sex into a new variable nesex with two groups female and male (Hamza et al., 2022). These two sex group variables were coded as 0 = female, and 1 = male; race was recoded into a new variable nerace with three groups that included Black, White, and other races (that is Asian and the other races) and were coded as 0 = Black, 1 = White, and 2 = other race (Sundaram et al., 2022); the exposure was recoded into a new variable nexposure with two groups that included missing and yes, and were coded as 0 = missing and 1= yes; county code of residence was recoded into a new variable newcountyzpc with two groups Lake and Marion (Hamza et al., 2022). These two groups were coded as 0 = Lake, and 1 = Marion; then, underlying medical conditions was recoded into a new variable neundmedcon with two groups that included no and yes, and were coded as 0 = no and 1 = yes. After that, I used the new recoded variable for this cross-sectional study analysis on the final data. I utilized frequencies, graphs, and charts for further analysis of the data. Moreover, I ensured strict confidentiality of the data to protect respondents, and data was shared only to my chair and committee members.

Research Questions and Hypotheses

The current analysis in this quantitative study sought to answer four research questions and assess their associated hypothesis:

RQ1: Is there an association between individuals’ age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana?

$H_0$1: There is no association between individuals’ age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana.

H_a1: There is a statistically significant association between individuals’ age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana.

RQ2: Is there an association between individuals’ exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race?

$H_0$2: There is no association between individuals’ exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

H_a2: There is a statistically significant association between individuals’ exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

RQ3: What is the association between underlying medical conditions (severe obesity [as measured by BMI $\geq$ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana?

$H_0$3: There is no association between underlying medical conditions (severe obesity [as measured by BMI $\geq$ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana.

H_a3: There are statistically significant associations between underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana.

RQ4: What are the significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana?

H_04: There are no statistically significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

H_a4: There are statistically significant predicting social determinants of health (age group, sex race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

Fully understanding the answers to these four questions can clarify the relationship between social determinants of health factors (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity (as measured by BMI≥ 40 kg/m$^{2)}$, prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 disease among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

Statistical Tests for the Study Outcome

Data analyses were performed using IBM© SPSS© Statistics version 27 (IBM© Corp., Armonk, NY, USA). (Green & Salkind, 2014; Wamer, 2021). The SPSS statistical software can analyze secondary datasets to determine associations between social determinants of health elements, preexisting comorbidities, and COVID-19 health outcome abstracted data because it can produce more accurate tabulated reports, charts, and various plots of distributions and trends, perform descriptive and inferential statistical analysis to address the research questions (Green & Salkind, 2014). Furthermore, it can calculate the effect size measures (Green & Salkind, 2014). For descriptive statistics, the independent and dependent variables were described as numbers (%) for categorical variables or data will be presented as proportions and percentages for categorical data (Green & Salkind, 2014).

To examine research question 1 (RQ1), I employed a bivariate chi-square test to determine the link between the independent variables, social determinants of health factors (age group, sex, race) and the dependent or outcome variable (COVID-19-Diagnosis). I utilized the chi-square, degrees of freedom (df), sample size (N), chi-square value, and the probability value ($P \leq 0.05$), to assess the link between proportions, and reporting association with COVID-19 status. Comparisons between groups were made using Pearson's chi-square for trend, and statistical significance was set at alpha or ($P \leq 0.05$). Additionally, I performed Binary logistic regression to assess the association between the independent variables and the dependent or outcome variable (Green and Salkind, 2014).

For (RQ2), I performed a bivariate Chi-square test of independence between the independent variables, social determinants of health factors (exposure, county code of residence), and the dependent variable (COVID-19- Diagnosis). I used the chi-square, degrees of freedom (df), sample size (N), chi-square value, and probability value ($P \leq 0.05$) to evaluate the associations between groups and the link with COVID-19 status. Pearson's chi-square was used for the trend, and statistical significance was set at alpha or ($P \leq 0.05$). Furthermore, Binary logistic regression was conducted to assess the association between the independent variables and the dependent or outcome variable while controlling for age, sex, and race, and significance was set at $p = .05$, Odds Ratio (OR) was used for association (Green and Salkind, 2014).

To evaluate (RQ3), I employed the Chi-square test of independence to examine the association between the independent variable underlying medical conditions (obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) and the dependent variable (COVID 19- Diagnosis). I employed the chi-square, degrees of freedom (df), sample size (N), chi-square value, and probability value ($P \leq 0.05$), to measure associations between proportions, and the link with COVID-19 status. Pearson's chi-square for trend was utilized to make comparisons, and statistical significance was set at alpha or ($P \leq 0.05$).

In addition, I did a Binary logistic regression test to assess the link between the independent variables and the dependent variable COVID-19 diagnosis while controlling for age, sex, and race. Significance was determined at a p-value equals to or <0.05, and Odds Ratio (OR) was utilized for association (Green and Salkind, 2014).

The Chi-square test was useful for this study because of the categorical nature of the study variables and thus provided an opportunity for testing the study hypothesis and determining if the categorical dependent and independent variables are linked. Furthermore, the study is composed of 1 dependent variable that is nominal in nature, and independent variables with 2 or more levels (independent groups), nominal or ordinal. The chi-square test also helped provide variables that demonstrated statistical significance for inclusion in the backward stepwise elimination multiple logistic regression model analysis.

To assess Research 4 (RQ4), I did a backward stepwise elimination logistic regression test to find the best predicting social determinants of health factors (age group, sex, race, exposure, county code of residence), underlying medical conditions (obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) for COVID 19- Diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana. To carry out the test, I simultaneously entered all six independent variables, and variables were removed in a stepwise manner sequentially, beginning with the highest p-value and ending with the best and statistically significant independent variable (s) (Green & Salkind, 2014). This also identified the model containing the fewest individual predictors (Green & Salkind, 2014). A p-value of 0.05 or less determined statistical significance for the study, and the change in R-squared determined predictability of the independent variables for COVID-19 incidence.

In addition, I conducted a hierarchical logistic regression test to examine whether predictors of social determinants of health factors (age group, sex, race, exposure, county code of residence), underlying medical conditions (obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) influence COVID-19 disease status among adults ages 18 years and older during the COVID-19 pandemic in Indiana. To carry out the test, I entered the data into the regression system in a block wise manner based on theory or logic and a predetermined manner created by me and controlling for age, sex, and race (Green and Salkind, 2014). In step 1 (block 1), I enter my control demographic covariates (i.e., age group, sex, and race, after entering COVID-19 Diagnosis in the dependent variable box. Then in step 2 (block 2), I enter the predictor variables (exposure, county code of residence), underlying medical conditions (obesity (as measured by BMI ≥ 40 $kg/m^{2)}$, prostate cancer diagnosis, diabetes, and hypertension). The adjusted odds ratios (AOR), 95% confidence intervals (CI), and p-values were calculated with a statistical significance level set at p-value<0.05 (See Table 3).

Table 3

Description of Research Questions and Variables by Level of Measurement and Statistical Analysis

Research questions	Independent variables (IV) and measurement	Dependent variables (DV) and measurement	Statistical analysis
RQ1	Age group (Categorical) Sex (Categorical) Race of patient (Categorical)	COVID-19 diagnosis (Categorical)	Descriptive (frequency distribution) Chi-square Multivariate logistic regression (binary logistic regression)
RQ2	Exposure (Categorical) County code (Categorical) Age (Categorical) Sex (Categorical) Race (Categorical)	COVID-19 diagnosis (Categorical)	Descriptive (frequency distribution) Chi-square Multivariate logistic regression (binary logistic regression)
RQ3	Underlying medical conditions (obesity, prostate cancer diagnosis, diabetes, hypertension, age, sex, race; Categorical)	COVID-19 diagnosis (Categorical)	Descriptive (frequency distribution) Chi-square Multivariate logistic regression (binary logistic regression)
RQ4	Age group (Categorical) Sex (Categorical) Race (Categorical) Exposure (Categorical) County code (Categorical) Underlying medical conditions (obesity, prostate cancer diagnosis, diabetes, hypertension; Categorical)	COVID-19 diagnosis (Categorical)	Multivariate logistic regression (backward stepwise elimination, hierarchical logistic regression)

Note. The statistical significance level was set at $p < .05$.

Based on Glanz et al. (2008); MacLory (1988) s’ SEM, and past research (Creswell, 2009), the associated covariates are age, sex, and race. The rationale for adding these potential covariate variables is that they are known to exert an effect on the outcome of the study since COVID-19 Diagnosis (dependent variable) depends on these elements.

Before conducting the logistic regression models, I first assessed and tested the pertinent regression assumptions to ensure that no deviations from linearity, normality, or homoskedasticity were evident (Green and Salkind, 2014). Furthermore, I assessed multicollinearity using variance inflation factor (VIF) to ensure that the multicollinearity diagnostics test is met and shows that all variable inflation factors are less than 3.00 (Green & Salkind, 2014). Multicollinearity of independent variables was detected if the variance inflation factor (VIF) was more than 3. Then Hosmer and Lemeshow test was done to test the goodness of the logistic regression mode.

Multiple regression statistical method provides an opportunity for examining the predictive power of my independent variables (age group, sex, race, exposure, county code of residence), underlying medical conditions (for example obesity) on a single binary dependent variable COVID 19 Diagnosis (Green & Salking, 2008). It also has the potential to assess the significant influence a particular predictor exerts on the dependent variable while controlling for all other predictors in the model. Moreover, it can yield confidence intervals and foster generalization of the study result to a population, for example, to all of Indiana (Green & Salkind, 2014). Furthermore, a vast number of credible epidemiological studies (Palacio and Tamariz, 2021; Khanijahani et al., 2021; Upshaw et al., 2021) have successfully demonstrated the utility of multiple regression techniques with similar variables and operational definitions for those variables.

Threats to Validity

Validity and reliability are critical constructs and quality criteria of quantitative research (Creswell, 2009). Validity entails an overall judgment of the soundness of a research design and method. Creswell (2009) noted *External Validity and Internal Validity* as vital elements to consider during a study. According to Creswell (2009), *External Validity entails* the generalizability of a study result to other populations and settings. While *Internal* Validity on the other hand pertains to the degree to which the independent variable influences the dependent variable outcome. So, it is vital that researchers identify threats that accompany validity and appropriately address them to improve study results and integrity.

Threats to External Validity

This study's threat to external validity, as Creswell (2009) mentioned entails elements inherent in the study that hinder the generalizability of my study results to larger, other populations, or various settings. It also deals with the specificity of the study variables and the characteristics of the selected group. For example, choosing adults ages 18 and older in an Indiana setting during the current pandemic timing could make it difficult to generalize my result to other populations (Creswell, 2009). According to Creswell (2009), selection biases pose a major threat to external validity.

Threats to Internal Validity

For this study, as Creswell (2009) noted, threats to internal validity relate to the causal link or other factors such as the quantitative and cross-sectional aspects of the research design which may hamper cause and effect relationship between the independent and dependent variable. Another threat to internal validity entails selection bias during sampling when the study and control group have very different characteristics (Creswell, 2009). This threat can be addressed through random assignment. Furthermore, Creswell (2009) mentioned that a very small sample size for a study can contribute to insufficient statistical power to produce results that we can trust or findings attributable to chance. Such a threat can be addressed by employing a large sample to improve the power of the research (Creswell, 2009). Moreover, confounds that occur when an observed result between two variables may not be due to the independent variable under investigation, but due to a third factor that was not controlled for can lead to internal validity (Creswell, 2009). This concern can be properly addressed by identifying and controlling for confounding variables to confirm that COVID-19 incidence was connected to the independent variables and did not occur by chance (Creswell, 2009). Moreover, internal validity is common in correlation design and so as Creswell (2009) noted, my study applying a correlational design suffers from internal validity threat.

Ethical Procedures

The data that was used for this study included routinely collected COVID-19 case data collected during the current COVID-19 pandemic surveillance and response activities in Indiana. Authorized data staff from the CDC anonymized the data before sharing it with me to conceal participants’ identity for presentation (CDC, 2023; Indiana Department of Health, 2022). I obtained ethics clearance from the Walden University IRB before embarking on my study. After my study had Walden IRB approval (IRB approval number is 06-28-23-0598302), then accordingly, I obtained approval from the

CDC by submitting a written request via email indicating details of the study and the variables being requested even though the (COVID-19 Case Surveillance Public Use Data with Geography) dataset is for public use, and can be viewed and downloaded from the data page directly using the Export function on the page, or through API. Then I received a link to the dataset, prefiltered for all of Indiana.

Even though I used existing secondary data (COVID-19 Case Surveillance Public Use Data with Geography) with nonidentifiable information from the CDC, I checked the data set thoroughly to ensure that no personal participant's information was present in other to safe guide their COVID-19 personal confidentiality. I also adhered to strict Ethical standards and IRB guidelines for this research (CDC, 2023; Creswell, 2009; Indiana Department of Health, 2022). Moreover, I revealed the study limitations in the discussion section of the study to foster transparency and improve decision making by stakeholders (Creswell, 2009).

Summary

In this section (section 2), I outlined the methodology employed for this research and discussed the study design and the specific setting and the research population. Furthermore, I outlined the data collection and analysis procedures data analysis processes, including a discussion of the inclusion and exclusion criteria of the sampling framework. Also, I presented information on the operationalization for each of the variables and their appropriate levels of measurement. I furthermore restated the research questions and hypotheses and highlighted the appropriate statistical tests utilized to analyze the data set to address the research questions. Moreover, I addressed the threats to external and internal validity and how to fix these threats to improve the cross-sectional study result. Finally, I presented the ethical procedures including the Walden IRB approval and the CDC agreements to gain access to the secondary data and maintain patients' information confidentially.

The result of this quantitative cross-sectional study is presented in Section 3. It can be expected to be used in different ways to inform public health leaders to promote efficient and equitable COVID-1 resource allocation, such as vaccine allocation.

Section 3: Presentation of the Results and Findings

Introduction

The purpose of this quantitative study was to examine the association between social determinants of health (exposure, county code of residence), preexisting comorbidities (obesity [as measured by BMI > 30 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension), and COVID-19 diagnosis, after controlling for age, sex, and race among adults ages 18 years and older during the COVID-19 pandemic in Indiana. The independent variables (age group, sex, race, exposure, county code of residence) and underlying medical conditions were included in the study to evaluate their relationship with COVID-19 diagnosis outcome (confirmed case or probable case). Four research questions and their associated hypotheses were addressed in the current analysis to determine whether there was any statically significant link between the independent and dependent variables. Understanding the answers to the four questions can clarify the relationship between social determinants of health elements and underlying medical conditions linked to COVID-19 diagnosis and foster better age-, area-, and comorbidity-relevant prevention and intervention strategies by public health officials to improve health in older adults affected by COVID-19 in Indiana and elsewhere.

RQ1: Is there an association between individuals' age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana?

$H_0$1: There is no association between individuals' age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana.

H_a1: There is a statistically significant association between individuals' age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana.

RQ2: Is there an association between individuals' exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race?

H_02: There is no association between individuals' exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

H_a2: There is a statistically significant association between individuals' exposure, county code of residence, and COVID-19 diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

RQ3: What is the association between underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana?

H_03: There is no association between underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana.

H_a3: There are statistically significant associations between underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, controlling for age group, sex, and race, among adults 18 years and older in Indiana.

RQ4: What are the significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana?

H_04: There are no statistically significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

H_a4: There are statistically significant predicting social determinants of health (age group, sex race, exposure, county code of residence), underlying medical conditions (severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) for COVID-19 diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

Section 3 includes results of statistical analyses (descriptive statistics, chi-square, binary logistic regression, backward stepwise logistic regression, and hierarchical logistic regression) of the CDC COVID-19 Case Surveillance Public Use Data with Geography data set with RT-PCR laboratory confirmed case and probable case (i.e., no known positive SARS-CoV-2 test result) results during the current COVID-19 pandemic in Indiana (CDC, 2023; Khedr et al., 2020). I furthermore provide brief descriptions of the COVID-19 data set. My study descriptive and inferential analysis findings are presented in detail to answer each of the four research questions. I conclude with a summary of the results for the four research questions.

Data Collection

I conducted a secondary data analysis using the COVID-19 Case Surveillance Public Use Data with Geography data set from the CDC collected from 2020 to 2023 for Indiana. I assessed the database from the CDC in Microsoft Excel file format and transported it into SPSS file format and obtained information from the designated female, male, and other category questionnaire, and the COVID-19-virus-specific RT-PCR test result file for all the patients who visited hospitals in Indiana. Then I obtained the variables of interest for this study. I employed the statistical software SPSS version 27.0 (IBM) to analyze the data and ensured that all of the data that met the research criteria were utilized for this research—namely, participants must have reported their gender as female or male, must have been aged 18 and older, and must have had a confirmed

COVID-19-specific RT-PCR test result or been noted as probable case. The outcome data COVID-19 diagnosis (current_status) included laboratory-confirmed cases and probable cases (when COVID-19 PCR was negative or not done; CDC, 2023; Khedr et al., 2020). Based on previously published methods (CDC, 2023; Khedr et al., 2020; Samuel et al., 2021), probable case was also selected as a COVID-19 incidence measure because it captured the full characteristics of the COVID-19 incidence, and probable case data were collected during COVID-19 surveillance while "not a case" was not collected due to the nature of pandemic surveillance. Moreover, these COVID-19 data were just public health data gathered for the pandemic response through contract tracing and case reporting mechanisms, and so noncases were not included in this dataset, but instead probable case data were collected (CDC, 2020; ISDH, 2020; Khedr et al., 2020; MCH, 2020; Samuel et al., 2021). All participant personal identifiers had already been removed from the database before the CDC released it to me in order to keep participants' private information confidential and protect their privacy.

Results

Descriptive Statistics

The secondary data set initially comprised 1,048,575 participants with COVID-19 laboratory-confirmed cases and probable cases from the hospitals in Indiana. During data management, cleaning, and review, I noted that there was no missing data; therefore, I did not address missing data through defining them as discrete missing value, but I utilized the exclude cases pairwise method using SPSS to ensure that any potential missing data were excluded from any analysis. I deleted participants aged 0 to 17 years from the data set and selected individuals from Marion County and Lake County after selecting specific cases for analysis in SPSS. After that, I included in my study all the individuals in the data set who met the criteria, representing 171,205 cases. The G*Power analysis required a sample size of 587 (power = 0.80, alpha = 0.05, and effect size/odds ratio 0.15).

I conducted statistical analysis on the COVID-19 Case Surveillance Public Use Data with Geography data set. To describe the population of these data sets, six categorical variables for social determinants of health (age group, sex, race, exposure, county code of residence) and underlying medical conditions such as obesity and COVID-19 diagnosis, measured at the nominal level or ordinal level, were selected, and frequency was determined. Table 4 indicates the overall characteristics of the sample of adults aged 18 years and older. Most of the study population was younger (111,351; 65%) and in the age group 18–49 years; only 22,791 (13.3%) of the respondents were in the age group 65 years and over. Of those who were in the age group 18–49 years, 97,394 (56.9%) were laboratory-confirmed cases, and 13,957(8.2%) were probable cases (Table 4). Of those who were in the age group 65 years and over, 20,926 (12.2%) were laboratory-confirmed cases and 1,865 (1.1%) were probable cases (Table 4).

In my sample, more than half of the participants were female (93,007; 54.3%), and approximately 77,117 (45.0%) were male; .7% were of unknown gender or belonged to other groups. Among female participants, 82,325 (48.1%) were laboratory-confirmed cases, and 10,682 (6.2%) were probable cases (Table 4). Of those who were male, 68,149

(39.8%) were laboratory-confirmed cases, and only 8,968 (5.2%) were probable cases (Table 4).

Additionally, most of the respondents were White (85,068; 49.7%), one fifth (36,312; 21.2%) were Black, and the remaining 24,307 (29.1%) were in the unknown category (Table 4). Of those who were White, 73,127 (42.7%) were laboratory-confirmed cases, and 11,941 (7.0%) were probable cases (Table 4). For Blacks, 33,190 (19.4%) were laboratory-confirmed cases, and only 3,122 (1.8%) were probable cases (Table 4). In the other races, 4,710 (2.8%) were probable cases, and 45,114 (26.4%) were confirmed cases.

For exposure, more than 96% (164,994) of the respondents in the dataset did not provide information about how they acquired COVID-19 infection or what the exposure mechanisms were (e.g., employment, occupation). Only 6,211 (3.6%) provided information about the mechanism of acquisition of COVID-19 (e.g., employment, occupation). Moreover, of those who were missing this information, a great number (145,247; 84.8%) were laboratory-confirmed cases and 19,747(11.5%) were probable cases (Table 4). Among those who reported exposure, 6,185 (3.6%) were laboratory-confirmed cases, and only 26 (0.0%) were probable cases (Table 4). Only 8,010 (4.7%) of the respondents reported that they had underlying medical conditions such as obesity. Of those who reported underlying conditions, 7,997 (4.7%) were laboratory-confirmed cases and 13 (0.0%) were probable cases (Table 4). For respondents who reported no underlying medical conditions, 142,543 (83.8%) were laboratory-confirmed cases and 19,637 (11.5%) were probable cases (Table 4).

The study participants were found to be unevenly distributed between the two counties in Indiana. Most of the participants (113,037; 66%) were from Marion County; only 58,168 (34%) of the respondents were Lake County residents (Table 4). Of those who came from Marion County, 98,854 (57.7%) were laboratory-confirmed cases and 14,183 (8.3%) were probable cases (Table 4). For those from Lake County, 52,578 (30.7%) were confirmed cases and 5,590 (3.3%) were probable cases (Table 4).

Table 4 shows the number and percentage of COVID-19 probable cases compared to confirmed infection cases by demographic variables among participants ages 18 and older in Indiana. Between 2020 and 2022, 151,432 (88.4%) COVID-19 diagnosed cases (laboratory-confirmed cases) in participants aged 18 and older were reported; COVID-19 infection incidence was higher among women (54.3%) than among men (45%). A lesser number 19,773 (11.5%) was reported for probable cases.

Table 4

Descriptive Characteristics of Variables

Variables	Probable Cases Frequency	Total Percent (%)	Confirmed Cases Frequency	Total Percent (%)	Total Frequency	Total Percent (%)
Age group						
18 to 49 years	13,957	8.2	97,394	56.9	111,351	65
50 to 64 years	3,899	2.3	32,787	19.2	36,686	21.4
65+ years	1,865	1.1	20,926	12.2	22,791	13.3
NA	52	0.0	325	0.2	377	.2
Sex						
Female	10,682	6.2	82,325	48.1	93,007	54.3
Male	8,968	5.2	68,149	39.8	77,117	45
NA	92	0.1	536	0.3	628	.4
Unknown	31	0.0	422	0.2	453	.3
Race						
American Indian/Alaska Native	2	0.0	68	0.0	70	.0
Asian	322	0.2	3,305	1.9	3,627	2.1
Black	3,122	1.8	33,190	19.4	36,312	21.2
Multiple/Other	208	0.1	1,137	0.7	1,345	.8
NA	1,714	1.0	18,721	10.9	20,435	11.9
Native Hawaiian/Other Pacific Islander	1	0.0	40	0.0	41	.0
Unknown	2,468	1.4	21,844	12.8	24,307	14.2
White	11,941	7.0	73,127	42.7	85,068	49.7
Exposure						
Missing (when not provided)	19,747	11.5	145,247	84.8	164,994	96.4
Yes	26	0.0	6,185	3.6	6,211	3.6
Underlying Medical Conditions						
No	19,637	11.5	142,543	83.8	162,180	95.3
Yes	13	0.0	7,997	4.7	8,010	4.7
County Code of Residence						
Lake County	5,590	3.3	52,578	30.7	58,168	34
Marion	14,183	8.3	98,854	57.7	113,037	66
COVID-19 Infection (Probable cases/Laboratory confirmed cases)	19,773	11,5%	151,432	88.5%	171,205	100%

Note. *N* = 171,205.

Assumption Testing

The next subsection assessed each of the four questions and hypothesis by determining the patterns of association between the dependent and independent variables using the COVID-19 Case Surveillance Public Use Data with Geography data set for Indiana to improve COVID-19 health. I utilized statistical analysis techniques including chi-square test and logistic regression for association to measure the differences versus relationships between the main variables after evaluating the statistical assumptions.

The assumptions of cross-tabulation and chi-square test were assessed and met (there was an adequate sample size of over 40 cases and adequate cell count, a minimum of five cases or counts per cell, no cells in the cross-tabulations had an expected count of less than five or zero cells had expected count less than 5; Green & Salkind, 2014). Furthermore, before I conducted full logistic regression analysis, I performed a preliminary regression analysis and regression assumptions were tested, and no deviations from linearity, normality, or homoskedasticity were discovered. Multicollinearity diagnostics indicated that all variable inflation factors for the variables age group, sex, race, county code of residence, and underlying medical conditions were less than 3.00 (acceptance VHF = 3 or less), and also tolerance criteria were met with values 1.0 (Table 5; Green & Salkind, 2014).

Multiple logistic regression enabled me to properly examine the effects of the five independent variables on the dependent variable while also controlling for the effect of one variable and assessing the contribution of the other variable. This provides the opportunity for the researcher to assess the independent contribution of each independent variable to, for example, COVID-19 disease (Frankfort-Nachmias & Nachmias, 2008; Green & Salkind, 2014).

Table 5

Results to Determine Multicollinearity Between Study Variables

		Collinearity statistics	
Coefficient model	Variable	Tolerance	VIF
2	Age group	.982	1.019
	Sex	.996	1.004
	Race	.984	1.016
	Exposure	.712	1.404
	Underlying medical conditions	.707	1.414
	County code of residence	.995	1.005

Note. Dependent variable: COVID-19 diagnosis.

Hypothesis Testing and Data Analysis

The chi-square tests and logistic regression model results of each independent variable are presented in Tables 6, 7, 8, and 9 for the current study.

RQ1: Is there an association between individuals' age group, sex, race, and COVID-19 Diagnosis in adults 18 years and older in Indiana?

$H_0$1: There is no association between individuals' age group, sex, race, and COVID-19 Diagnosis in adults 18 years and older in Indiana.

H_a1: There is a statistically significant association between individuals' age group, sex, race, and COVID- 19 Diagnosis in adults 18 years and older in Indiana.

To examine the association between age group, sex, race and COVID- 19 diagnosis, I conducted a chi-square test of independence between the three categorical variables and COVID- 19 diagnosis. Table 6 presents the results of the chi-square test and cross-tabulation for the variables. The chi-square test revealed a statistically significant association between age group and COVID-19 diagnosis $\chi2$ = (1, N = 170,124) = 289.670, p < .001, phi = .041; and race $\chi2$ = (2, N = 170124) = 1043.200, p = .001, phi = .078. However, the test indicated no statistically significant association between sex and COVID-19 diagnosis $\chi2$= (1, N = 170124) = 0.855, p = 0.355, phi = -.002. Since the majority or categories demonstrated an association, therefore, I rejected the null hypothesis for RQ. 1. I utilized the chi-square test of independence formula x2 =∑(Oi-Ei)/Ei, in SPSS (Green & Salkind, 2014).

Table 6

Crosstabulation and Chi-Square of Age Group, Race, Sex, Exposure, Underlying Medical Conditions, County Code of Residence, and COVID- 19 Diagnosis

	COVID-19 diagnosis				
Variable	Probable case	Laboratory-confirmed case	$\chi2$	*df*	*p*
Age group			289.670	1	0.001
18–64 years	17,788 (10.5%)	129,606 (76.2%)			
65+ years	1,862 (1.1%)	20,868 (12.3%)			
Race			1043.200	2	0.001
Black	3,122 (1.8%)	33,190 (19.5%)			
White	11,927 (7.0%)	73,022 (42.9%)			
Other race	4,601 (2.7%)	44,262 (26.0%)			
Sex			0.855	1	0.355
Female	10,682 (6.3%)	82,325 (48.4%)			
Male	8,968 (5.3%)	68,149 (40.1%)			
Exposure			770.834	1	0.001
Missing	19,624 (11.5%)	144,368 (84.9%)			
Yes	26 (0.0%)	6,106 (3.6%)			
Underlying medical conditions			1057.583	1	0.001
No	19,637 (11.5%)	142,543 (83.8%)			
Yes	13 (0.0%)	7,931 (4.7%)			
County code of residence			317.380	1	0.001
Lake County	5,553 (3.3%)	52,150 (30.7%)			
Marion County	14,097 (8.3%)	98,324 (57.8%)			

Note. *N = 17,0124.* Dependent variable: COVID-19 diagnosis.

I also conducted binary logistic regression analysis to determine the association between the independent variables age group, sex, race, and COVID- 19 Diagnosis in adults 18 years and older in Indiana. Table 7 indicated the binary logistic regression analysis between all the variables in the model. The logistic regression model was statistically significant, Wald = 13692.175, p < .001. The model explained between 0.9% (Cox and Snell R2) and 1.7% (Nagelkerke R2) of the variance in COVID-19 infection and correctly or accurately classified 88.4% of the cases. Sensitivity was 100% and specificity was 0%. While a Ombibus test of model coefficients test confirmed the goodness of fit of this model (p < .001), the Hosmer and Lemeshow did not (p < .001), therefore, a statistically significant overall fit model was not found, and the model did not adequately describe the data. Also, participants aged 65 years and over were 1.625 times more likely of being diagnosed with COVID-19 compared to those in the age group 18 years to 64 years old (OR = 1.625, 95% CI [1.546, 1.709]). The results of this analysis (table 7) demonstrated that age group was significantly associated with COVID-19 diagnosis among adults ages 18 and older in Indiana, Wald = 360.301, p < .001, therefore, the null hypothesis was rejected, and the alternative hypothesis was accepted. Furthermore, when comparing individual's race, Whites were 0.565 times less likely of being diagnosed with COVID-19 than Blacks (OR = 0.565, 95% CI [0.542, 0.589]), and participants of other race group were 0.909 times less likely of being diagnosed with COVID-19 than Blacks (OR = 0.909, 95% CI [0.867, 0.953]). Race was significantly associated with COVID-19 diagnosis among adults ages 18 and older in Indiana, Wald = 1110.973, p < .001. Therefore, the null hypothesis was rejected, and the alternative hypothesis was accepted. When comparing sex, males were 0.996 times less likely than females of being diagnosed with COVID-19 (OR = 0.996, 95% CI [0.966, 1.026]). However, the results of this analysis (table 7) demonstrated that sex was not significantly associated with COVID-19 diagnosis among adults ages 18 and older in Indiana, Wald = 0.079, p = .778. Therefore, the null hypothesis was accepted, and the alternative hypothesis was rejected. Overall, since most categories demonstrated an association, therefore, I rejected the null hypothesis for RQ1.

Table 7

Binary Logistic Regression Predicting the Likelihood of COVID-19 Diagnosis Based on Age Group, Sex, Race, Exposure, and Underlying Medical Conditions With Constant, OR, 95% CI, Wald, and p-values

Variable	Wald	Sig.	Exp(B) *OR*	95% confidence interval Lower	 Upper
Age group					
Ref: 18–64 years					
65+ years	360.301	0.001	1.625	1.546	1.709
Sex					
Ref: Female					
Male	0.079	0.778	0.996	0.966	1.026
Race					
Ref: Black					
White	723.994	0.001	0.565	0.542	0.589
Other race	15.359	0.001	0.909	0.867	0.953
Exposure					
Ref: Missing					
Yes	295.948	0.001	29.514	20.070	43.401
County code of residence.					
Ref: Lake County					
Marion County	316.200	0.001	0.741	0.717	0.766
Underlying medical conditions					
Ref: No					
Yes	242.731	0.001	75.736	43.942	130.532

Note. *N* = 170,124. Reference category: not observed; variable controlled for in the model: age group, sex, and race.

RQ2: Is there an association between individuals’ exposure, county code of residence, and COVID-19- Diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race?

$H_0$2: There is no association between individuals’ exposure, county code of residence, and COVID 19- Diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

H_a2: There is a statistically significant association between individuals’ exposure, county code of residence, and COVID 19- Diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race.

I performed a chi-squared test for independence to determine the association between exposure, county code of residence, and COVID 19- Diagnosis. Table 6 presents the results of the chi-square test and cross-tabulation for the variables. Results of the chi-square test of association between the variables indicated statistical significance between the variables, individual’s exposure status and COVID-19 infection $\chi2$= (1, N = 170,124) = 770.834, *p* = .001, phi = 0.067 (Table 6); county code of residence and COVID-19 diagnosis $\chi2$ = (1, N = 170124) = 317.380, *p* = .001, phi = -.043 (Table 6). Therefore, I rejected the null hypothesis for RQ2 that says there is no association between individuals’ county code of residence and COVID- 19 Diagnosis in adults 18 years and older in Indiana. The chi-squared test for independence formula utilized in SPSS was x2 =∑(Oi-Ei)/Ei.

To further assess the relationship between the variables exposure, county code of residence and COVID 19 diagnosis, I conducted a binary logistic regression analysis on the variables while controlling for age group, sex, and race. First, I examined the association between the independent variable, exposure and COVID 19 diagnosis. The model was statistically significant, Wald = 310.072, p < .001, and correctly classified 88.4% of the cases. The odds ratios indicated that those who reported been exposed were more likely to be diagnosed with COVID 19 than individuals who reported not been exposed (OR = 31.923, 95% CI [21.711, 46.937]). The control variables, age group, sex, and race were then added to the model. The logistic regression model was statistically significant, Wald = 12990.682, p < .001. The model explained between 1.5% (Cox and Snell R2) and 3.0% (Nagelkerke R2) of the variance in COVID 19 and correctly classified 88.4% of the cases. Sensitivity was 100% and specificity was 0%. The results of this analysis indicated a significant association between exposure status and COVID 19 diagnosis, Wald = 295.948, p < .001. Participants who were exposed were more likely to be diagnosed with COVIS 19 than those who were not exposed (OR = 29.514, 95% CI [20.070, 43.401]). Because there was a significant association between exposure status and COVID 19 diagnosis, the null hypothesis was rejected, and the alternative hypothesis was accepted.

In addition, I performed univariate logistic analysis for county code of residence and COVID-19 diagnosis. The model revealed statistically significant association result, Wald = 315.638, p < .001, and correctly classified or accurately classified 88.4% of the cases. In this model, the odds ratio showed that those from Marion County were 0.743 times less likely to be COVID-19 diagnosed than those from Lakw County (OR = 0.743, 95% CI [0.719, 0.767]; Table 7), Nex, the control variables, age group, sex, and race were fed into the model with the significant risk factor county code of residence. After controlling for age group, sex, and race in the model, the model was significant, Wald =

11449.186, $p < .001$. The model explained between 1% (Cox and Snell R2) and 2% (Nagelkerke R2) of the variance in COVID-19 diagnosis and correctly or accurately classified 88.4% of the cases. Sensitivity was 100% and specificity was 0%. Also, while a Ombibus test of model coefficients test confirmed the goodness of fit of this model ($p < .001$), the Hosmer and Lemeshow did not show a good fit ($p < .001$), therefore, the overall model did not show goodness of fit and the model did not adequately describe the data. The results of this analysis indicated a significant association between County code of residence and COVID-19 diagnosis, Wald = 316.2000, $p < .001$ (Table 7), and indicated that participants that were from Marion County were less likely of being COVID-19 diagnosed when compared to those from Lake County (OR = 0.741, 95% CI [0.717, 0.766]). Because the results indicated a significant relationship, the null hypothesis was rejected, and the alternative hypothesis was accepted.

RQ3: What is the association between underlying medical conditions severe obesity (as measured by BMI $\geq$ 40 $kg/m^{2)}$, prostate cancer diagnosis, diabetes, and hypertension and COVID 19- Diagnosis controlling for age group, sex, and race among adults 18 years and older in Indiana?

$H_0$3: There is no association between underlying medical conditions severe obesity (as measured by BMI $\geq$ 40 $kg/m^{2)}$, prostate cancer diagnosis, diabetes, and hypertension and COVID 19- Diagnosis controlling for age group, sex, and race among adults 18 years and older in Indiana.

H_a3: There is statistically significant associations between underlying medical conditions severe obesity (as measured by BMI ≥ 40 kg/m$^{2)}$, prostate cancer diagnosis, diabetes, and hypertension and COVID-19- Diagnosis controlling for age group, sex, and race among adults 18 years and older in Indiana.

To examine the association between underlying medical conditions and COVID-COVID-19 diagnosis, I conducted a chi-square test between the two variables. Table 6 presents the results of the chi-square test and cross-tabulation for the variables. Results of the chi-square test of association between the variables indicated statistical significance between the variables $\chi 2$ = (1, N = 170124) = 1057.583, *p* = 0 .001, Phi = -.043. Because the results indicated a significant relationship, the null hypothesis was rejected, and the alternative hypothesis was accepted. The chi-squared test for independence formula utilized in SPSS was x2 =∑(Oi-Ei)/Ei.

To further assess the relationship between the variables, a binary logistic regression analysis was carried out in the study population when controlling for age group, sex, and race. The first step examined the association between the independent variable, underlying medical conditions and COVID-19 diagnosis. The model was statistically significant, Wald = 254.671, p = 0.001, and correctly classified 88.4% of the cases. In this model, the odds ratio showed that those with underlying medical conditions were 84.045 times more likely to be COVID-19 diagnosed than those without underlying medical conditions (OR = 84.045, 95% CI [48.770, 144.836]; Table 7).

Nex, the control variables, age group, sex, and race were fed into the model with the significant risk factor underlying medical conditions. After controlling for age group, sex, and race, the independent variable underlying medical condition's logistic regression model was statistically significant, Wald = 12662.687, p < .001 (Table 7). The model explained between 1.8% (Cox and Snell R2) and 3.5 % (Nagelkerke R2) of the variance in COVID-19 diagnosis and correctly or accurately classified 88.4% of the cases. Sensitivity was 100% and specificity was 0%. Also, while a Ombibus test of model coefficients test confirmed the goodness of fit of this model (p < .001), the Hosmer and Lemeshow test did not show a good fit (p < .001), therefore, the overall model did not show goodness of fit, and the model did not adequately describe the data. Moreover, participant's underlying medical conditions was statistically associated with COVID-19 infection diagnosis, Wald = 242.731, p < .001. Also, participants with underlying medical conditions were 75.736 times more likely of being COVID-19 diagnosed than those without underlying medical conditions (OR = 75.736, 95% CI [43.942,130.532]). Because the results indicated a significant relationship, the null hypothesis was rejected, and the alternative hypothesis was accepted.

RQ4: What are the significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity (as measured by BMI≥ 40 kg/m$^{2)}$, prostate cancer diagnosis, diabetes, and hypertension) for COVID 19- Diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana?

$H_0$4: There are no statistically significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) for COVID 19- Diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

H_a4: There are statistically significant predicting social determinants of health (age group, sex race, exposure, county code of residence), underlying medical conditions (severe obesity (as measured by BMI ≥ 40 kg/m^2)), prostate cancer diagnosis, diabetes, and hypertension) for COVID 19- Diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana.

To evaluate research question four and to test my hypothesis 4, I performed Backward elimination multiple regression analyses using model that included all variables that were statistically significantly associated in the univariate and bivariate analysis (age group, race, underlying medical conditions, and county code of residence; $p \leq 0.05$). Furthermore, based on agreement with other analyzes and the congruence with the determinants of ecological model, I decided to also include the variable that was not found to be statistically significant in the chi-square analysis and binary logistic regression (sex $p >= 0.05$; Al Mutair et al., 2021; Brown et al., 2022; Glanz et al., 2008; Green et al., 2021; Khanijahani et al., 2021; McLeroy et al., 1988). Then I entered all the dependent variable and the predictor variables simultaneously in the model (Green & Salkind, 2014).

Table 8 showed the results of the backward stepwise elimination multiple regression analysis best model predicting individuals' COVID-19 infection diagnosis based on the SEM framework. The results of the test indicate that the use of these six independent variables (age group, sex, race, exposure, underlying medical conditions, and county code of residence) created statistically significant model predicting COVID-19 infection diagnosis F (4, 170124) = 421.368, $p = .001$, (adjusted $R^2 = 0.010$ or 1%) and account for a good portion of the variance for each. In model 3 the final and best model, age group (t = 11.951, VIF =1.019, $p <.001$); exposure (t = 11.155, VIF =1.404, $p < .001$); county code of residence (t = -17.367, VIF =1.005, $p<.001$), and underlying medical conditions (t = 20.549, VIF =1.414, $p < .001$) were identified as significant independent predictors of COVID-19 diagnosis incidence. These model, also indicates that (R^2) .010 or 1% of the variance in COVID-19 infection diagnosis status or in predicting COVID-19 can be explained by this model (variables). Cohen in (1988) noted this as a small effect but however has major practical and clinical implication given the deadliness of the COVID-19 virus disease in adults aged 18 and over years in Indiana (Creswell, 2009; Faul et al., 2007; Green & Salkind, 2014). The results also, indicates that County-level social determinants of health and comorbidities, particularly age group, exposure, underlying medical conditions, and county code of residence, offer significant explanatory power regarding COVID-19 incidence rates. However, after backward regression, underlying medical conditions emerged as the most unique significant risk factor associated with being a COVID-19 infected case-patient and that statistically significantly contributed uniquely to the final model (b/beta/ β = 0.059 or 5.9%, t = 20.549, p <.001). The equation for the model or nature of the relationship was described in the regression equation Y = b0 + b1X1i + b2X2i + b3X3i, b4X4i, b5X5i where Y = COVID-19 diagnosis, X1i represented the value age group, X2i represented the value exposure, X3i represented the value county zip code, and X4i represented the value underlying medical conditions.

Table 8

Backward Multiple Regression Analysis Summary Predicting COVID-19 Diagnosis From Predictor Variables

Model 3 variables	B (coefficient)	SEB	β	*t*	VHF	*P*
Age group	0.027	0.002	0.029	11.951	1.019	< .001
Exposure	0.55	0.005	0.032	11.155	1.404	< .001
County code of residence	-0.028	0.002	-0.042	-17.367	1.005	< .001
Underlying medical conditions	0.089	0.004	0.059	20.549	1.414	< .001
Adjusted R^2	0.010					
F	421.368					

Note. N =170,124.

*p < .05; **p < .01; ***p < .001.

Moreover, to assess the specific effects of exposure, county code of residence, and underlying medical conditions on COVID-19 diagnosis status, I conducted a two-block hierarchical logistic regression model (Table 9). The first step block 1 model 1, I examined the association between the independent variable, (control variables) age group, sex, race, and COVID-19 diagnosis incidence without the independent or predictor variables exposure, underlying medical conditions and county code of residence.

In the first model, ($R^2 = 0.002$, $p = .001$). Age group ($b = 0.039$, $\beta = 0.041$, $p = 0.001$). was associated with more COVID-19 diagnosis while participant's sex ($b = -.001$, $\beta = -.002$, $p = 0.417$), and race status ($b = 4.362$, $\beta = 0.000$, $p = 0.968$) were not associated with COVID 19 diagnosis. The addition of exposure, county code of residence, and underlying medical conditions in Model 2 added 0.010 or 1% ($p< .001$) to the variance in Model 1, bringing the total variance explained to 1.2% ($p < .001$). Also, in Model 2, Age group status remained significantly associated with more COVID-19 diagnosis ($b= 0.027$, $\beta = 0.029$, $p = .001$), while sex ($b = -.001$, $\beta = -.001$, $p < 0.601$) and race ($b = 0.002$, $\beta = 0.003$, $p = .172$) were negatively associated with COVID-19 infection. Furthermore, table 9 revealed that in the final model, (model 2), underlying medical conditions made the most unique contribution statistically to the model ($\beta = 0.059$, $p < .001$) and was the most significant predictors of COVID-19 diagnosis, even though, exposure ($b = 0.055$, $\beta = 0.032$, $p < .001$) and county code of residence ($b= -.028$, $\beta = -.042$, $p < .001$) were also associated with COVID-19 diagnosis. The final model explained 0.8% of the overall variance ($p < .001$) and represented a 0.6% increase in the variance explained in Model 2, suggesting that the independent variables exposure, county code of residence and underlying medical conditions explained 0.8% additional variance in outcome even after controlling for the effects of age group, sex, and race. The ANOVA result revealed that the model is a statistically significant predictor of COVID-19 while controlling for age group, sex, and race and adding the predictor variables exposure, county code of residence and underlying medical conditions ($P < 0.001$). The equation for the model or nature of the relationship was described in the regression equation $Y = b0 + b1X1i +$

b2X2i + b3X3i, b4X4i, b5X5i where Y = COVID-19 diagnosis, X1i represented the value age group, X2i represented the value sex, X3i represented the value race, X4i represented the value exposure, X5i represented county zip code of residence, and X6i represented underlying medical conditions.

Table 9

Hierarchical Logistic Regression Analysis Summary Predicting COVID-19 Diagnosis From Predictor Variables

Model 1 (Block 1)	B (coefficient)	SEB	β	*P*	
Predictor					
Age group	0.39	0.002	0.041	0.001	
Sex	-.001	0.002	-.002	0.417	
Race	4.362	0.001	0.000	0.968	
R^2					0.002
P					0.001
ΔR^2					0.002
Model 2 (Block 2)	B (coefficient)	SEB	β	*P*	
Predictor					
Age group	0.027	0.002	0.029	0.001	
Sex	-.001	0.002	-.001	0.601	
Race	0.002	0.001	0.003	0.172	
Exposure	0.055	0.005	0.032	0.001	
Underlying medical conditions	0.090	0.004	0.059	0.001	
County code of residence	-.028	0.002	-.042	0.001	
R^2					0.010
ΔR^2					0.008
P					0.001

Note. *N* = 170,124.

*$p < .05$; **$p < .01$; ***$p < .001$.

Summary

In Section 3, I provided univariate descriptives of the study’s six categorical variables. First, I diligently cleaned and recoded the variables of this study and then conducted descriptive analysis and presented the descriptive results in a table containing the frequencies and percentages of each variable assessed. Furthermore, before embarking on the full regression models, I performed logistic regression test to determine multicollinearity and other significant logistic regression assumption tests to ensure that the assumptions were properly met for logistic regression. The secondary data set initially contained 1,048,575 participants with COVID-19 laboratory confirmed cases and probable cases from the hospitals in Indiana. After selecting the target age group of the male and female participants aged 18 years and older that were from Marian County and Lake County, I included in my study all the individuals in the data set that fully met the criteria representing 170124 cases. Of these participants, 150,474 (88.4%) COVID-19 diagnosed cases (laboratory confirmed cases) in both women and men aged 18 and older were reported. A lesser number 19,650 (11.6%) was reported as probable cases. COVID-19 infection incidence was higher among women (54.7%) than among men (45.3%). Chi-square tests of independent and logistic regression models were conducted to determine if there were significant relationships between COVID-19 diagnosis and six social determinants variables. Results of the analyses found significant associations between COVID-19 diagnosis and five of the independent variables: age group, race, county code of residence, and underlying medical conditions, but showed no significant association between COVID-19 diagnosis and sex for the chi-square and binary logistic regression tests. In both backward stepwise logistic and hierarchical logistic regression final models, underlying medical conditions made the most unique contribution statistically to the model ($\beta = 0.059$, $p < .001$). Section 4 will entail a review and interpretation of major findings and a thorough discussion of how the SEM theoretical framework guided the interpretation of findings. Then the study’s limitations, professional application, recommendations for further research, and the potential positive social change impact of the study will be presented also.

Section 4: Application to Professional Practice and Implications for Social Change

Introduction

The purpose of this quantitative study was to examine the association between social determinants of health (exposure, county code of residence, underlying medical conditions [severe obesity (as measured by BMI $\geq$ 40 kg/m^2)], prostate cancer diagnosis, diabetes, and hypertension) and COVID-19 diagnosis, after controlling for age group, sex, and race. I also evaluated the predictive ability of these independent variables on the current status of COVID-19 infection among adults 18 years and older in Indiana. These findings will help inform potential public health initiatives to reduce COVID-19 disparity among adults ages 18 and older in Indiana and elsewhere.

To address the four research questions in this quantitative study, the specific research design utilized secondary data using the COVID-19 Case Surveillance Public Use Data with Geography (from the CDC) with a cross-sectional design to examine how social determinants of health (age group, sex, race, exposure, county code of residence) and underlying medical conditions (such as obesity [as measured by BMI $\geq$ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension) collectively influence COVID-19 infection incidence among adults 18 years and older in Indiana (ISDH, 2021). A cross-sectional study design can provide a better opportunity for assessing the association between independent and dependent variables at a unique point in time and properly measuring prevalence of health outcomes such as COVID-19 infection at a specific place and time (Creswell, 2009). It can also describe the demographics of a population (Creswell, 2009).

Studies are scarce (Hanson et al., 2020; Samuel et al., 2021) on the association between various social determinants of health factors and comorbidities and COVID-19 infection in Indiana. Also, most of the few studies, including most of those that have been conducted in areas outside Indiana, have reported inconsistent or conflicting results on the incidence of COVID-19 infection and some social determinants of health factors and preexisting comorbidities, and they have also lacked a clear theoretical framework to examine associations between these variables and to produce high-quality evidence and to clearly understand the particular effects of these factors on COVID-19 outcomes in this population. Furthermore, most of the COVID-19 studies in the region, including in Indiana (Dixon et al., 2021; Hanson et al., 2020; Khanijahani et al., 2021; Samuel et al., 2021; Timsina et al., 2020), were found to be targeting the general population as a whole, leaving specifically the population of adults ages 18 and older in Indiana inadequately targeted or investigated despite the heavy burden of COVID-19 infection among this population during the ongoing pandemic in Indiana. These limitations are critical and certainly essential because the scarcity of research data fully characterizing COVID-19-infection-associated social determinants of health and preexisting comorbidities for older adults makes them an invisible at-risk population, hinders appropriate allocation and distribution of needed resources and prevention of COVID-19 infection among at-risk populations, and hinders a full understanding of the continual rapid spread of the COVID-19 disease (CDC, 2021; CSDH, 2008; Dixon et al., 2021; Hanson et al., 2020; Khanijahani et al., 2021; Maness et al., 2021; Timsina et al., 2020) among at-risk adults ages 18 and older in the United States, particularly Indiana. Notably, in Brown et al.’s

(2022) rapid review and synthesis of public health recommendations from 338 peer-reviewed studies on addressing equity and social determinants of health during the COVID-19 pandemic, a substantial number of articles (50) recommended COVID-19 research on the social determinants of health and underlying conditions. A comprehensive understanding of the factors that contribute to the ongoing spread of the disease is an essential step in limiting the spread of this disease and curtailing mortality in Indiana and elsewhere. Thus, in the present study, I investigated the association between social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (e.g., severe obesity [as measured by BMI ≥ 40 kg/m^2], prostate cancer diagnosis, diabetes, and hypertension), and the likelihood of contracting COVID-19 disease among adults 18 years and older in Indiana using the SEM framework and cross-sectional analysis of secondary data (COVID-19 Case Surveillance Public Use Data with Geography) collected by the CDC during the ongoing pandemic in Indiana. Understanding the association between the variables can reduce the COVID-19 burden in this group.

Regarding the findings from this study, the chi-square test of independence revealed a statistically significant association between age group and COVID-19 diagnosis $\chi 2$ = (1, *N* = 170,124) = 289.670, *p* < .001, phi = .041; race and COVID-19 diagnosis $\chi 2$ = (2, *N* = 170,124) = 1043.200, *p* = .001, phi = .078; individual’s exposure status and COVID-19 diagnosis $\chi 2$= (1, *N* = 170,124) = 770.834, *p* = .001, phi = 0.067; county code of residence and COVID-19 diagnosis $\chi 2$ = (1, *N* = 170,124) = 317.380, *p* = .001, phi = -.043; and underlying medical conditions and COVID-19 diagnosis $\chi 2$ = (1, *N*

= 170,124) = 1057.583, p = 0 .001, phi = -.043. However, the sex variable was not statistically significantly associated with COVID-19 infection diagnosis $\chi 2$ = (1, N = 170,124) = 0.855, p = 0.355, phi = -.002. Moreover, the binary logistic regression model showed a statistically significant association between the variables and indicated that participants aged 65 years and over were 1.625 times more likely to be diagnosed with COVID-19 compared to those in the age group 18 years to 64 years (*OR* = 1.625, 95% CI [1.546, 1.709]), Wald = 360.301, p < .001. In terms of the individual’s race, Whites were 0.565 times less likely to be diagnosed with COVID-19 than Blacks (*OR* = 0.565, 95% CI [0.542, 0.589]), and participants of other race groups were 0.909 times less likely to be diagnosed with COVID-19 than Blacks (*OR* = 0.909, 95% CI [0.867, 0.953]), Wald = 1110.973, p < .001. Sex was not significantly associated with COVID-19 diagnosis, and when comparing sex, males were 0.996 times less likely than females to be diagnosed with COVID-19 (*OR* = 0.996, 95% CI [0.966, 1.026]), Wald = 0.079, p = .778. Furthermore, when age group, sex, and participant’s race were controlled for, exposure status, individual’s county code of residence, and underlying medical conditions remained significantly associated with COVID-19 infection diagnosis, indicating that participants who were exposed were more likely to be diagnosed with COVID-19 than those who were not exposed (*OR* = 29.514, 95% CI [20.070, 43.401]), Wald = 295.948, p < .001. In terms of county code of residence, participants from Marion County were less likely to be diagnosed with COVID-19 when compared to those from Lake County (*OR* = 0.741, 95% CI [0.717, 0.766]), Wald = 316.2000, p < .001. Regarding underlying medical conditions, participants with underlying medical conditions were more likely to be diagnosed with COVID-19 than those without underlying medical conditions (*OR* = 75.736, 95% CI [43.942,130.532]), Wald = 242.731, $p < .001$.

In the backward logistic regression model, the full regression yielded a significant model for predicting COVID-19 diagnosis incidence, $F(4, 170124) = 421.368$, $p = .001$, (adjusted $R^2 = 0.010$ or 1%) and accounts for a good portion of the variance for each. Also, it produced a more efficient and parsimonious model—Model 3, the final and best model—containing only four variables, age group ($t = 11.951$, VIF =1.019, $p < .001$), exposure ($t = 11.155$, VIF =1.404, $p < .001$), county code of residence ($t = -17.367$, VIF = 1.005, $p < .001$), and underlying medical conditions ($t = 20.549$, VIF = 1.414, $p < .001$), which were identified as significant independent predictors of COVID-19 diagnosis incidence. This model improved the explanatory power of the variables and explains (R^2) .010 or 1% of the variance in COVID-19 infection diagnosis status incidence and predicting COVID-19. However, only underlying medical conditions emerged as the most unique significant predictor of COVID-19-infected case-patient and statistically significantly contributed uniquely to the final model (b/beta/ $\beta = 0.059$ or 5.9%, $t = 20.549$, $p <.001$). In Hierarchical Logistic Regression Model 1, ($R^2 = 0.002$, $p = .001$), age group (b = 0.039, $\beta = 0.041$, $p = 0.001$). was associated with more COVID-19 diagnosis while participant's sex (b = -.001, $\beta = -.002$, $p = 0.417$) and race status (b = 4.362, $\beta = 0.000$, $p = 0.968$) were not. The addition of exposure, county code of residence, and underlying medical conditions in Model 2 added 0.010 or 1% ($p < .001$) to the variance in Model 1, bringing the total variance explained to 1.2% ($p < .001$). Also, in Model 2, age group status remained significantly associated with more COVID-19 diagnoses (b = 0.027, β = 0.029, p = .001), while sex (b = -.001, β = -.001, p < 0.601) and race (b = 0.002, β = 0.003, p = .172) were negatively associated with COVID-19 infection. Moreover, Model 2 indicated that underlying medical conditions made the most unique contribution statistically to the model (β = 0.059, p < .001) and was the most significant predictor of COVID-19 diagnosis, even though exposure (b = 0.055, β = 0.032, p < .001) and county code of residence (b = -.028, β = -.042, p < .001) were also associated with COVID-19 diagnosis. The final model explained 0.8% of the overall variance (p < .001) and represented a 0.6% increase in the variance explained in Model 2, suggesting that the independent variables exposure, county code of residence, and underlying medical conditions explained 0.8% additional variance in outcome even after controlling for the effects of age group, sex, and race. The ANOVA result revealed that the model was a statistically significant predictor of COVID-19 while controlling for age group, sex, and race and adding the predictor variables exposure, county code of residence, and underlying medical conditions (P < 0.001). In Model 2, underlying medical conditions made the most unique contribution statistically to the model (β = 0.059, p < .001). In this section, I outline the key findings that emerged from my study, interpret the results, address the limitations of the study, and provide recommendations for future studies. Then I address implications for professional practice and positive social change, followed by a conclusion.

Interpretation of the Findings

Utilizing the SEM variables for the current study answered my four research questions that an adult aged 18 years and older’s age group, race, exposure, county code of residence, social determinants of health elements, and underlying medical conditions elements were most significantly associated with the study population's COVID-19 diagnosis incidence. I present the interpretation of findings from the current study and a vivid comparison with existing literature in relation to the four research questions.

Research Question 1

RQ1 was as follows: Is there an association between individuals' age group, sex, race, and COVID-19 diagnosis in adults 18 years and older in Indiana?

Concerning the individual-level determinant age group, the results of the chi-square test of association between age group and COVID-19 infection diagnosis and binary logistic regression model indicated a statistically significant relationship between the individual's age group and COVID-19 infection diagnosis status. The binary logistic regression model demonstrated that participants aged 65 years and over were more likely to be diagnosed with COVID-19 compared to their younger counterparts in the age group 18 years to 64 years. This finding is consistent with those of many previous studies that have indicated that having older adult status could increase the risk of COVID-19 infection diagnosis (Al Mutair et al., 2021). Similarly, one large systematic review and meta-analysis of 20 studies and 64,676 COVID-19 patients by Biswas et al. (2021) found that adults aged 50 and older had a 3.45 greater chance of obtaining a SARS-CoV-2 positive result than those younger than 50 years (RR 3.45, $p = 0.0008$). Furthermore, in a retrospective observational study out of Spain, Gimeno-Miguel et al. (2021) utilized 68,913 individuals to study age-related COVID-19 infection. They reported an increased average risk of 6.19 in men and 8.19 in women aged 80 and over compared to individuals aged 45–64 years. Also, in this study, patients aged 80 years and older had 8.23 times higher risk of COVID-19 than those in the 45- to 64-years age group.

From the perspective of the SEM theory and previous literature, different elements and factors may help explain this trend seen in COVID-19 incidence in this study population considering age group. A plausible explanation for my finding according to Casas-Deza et al. (2021) could be that the high COVID-19 vulnerability in adults probably resulted from long-established structural systemic inequities and social determinants of health issues in the United States, in addition to preexisting comorbidities, and poor and weak immune response to viral infection associated with age. Due to these inequities and other characteristics of the older adult population, members of this group have urgent and special social and health needs. My results indicate the need for the strengthening of COVID-19 health promotion and behavior educational programs among adults aged 18 and older in Indiana, including well-targeted COVID-19 prevention and continuous care and support services designed to target the older adult population. Moreover, older adult populations and the COVID-19 issue may trigger enormous public health challenges a few years from now that could overwhelm healthcare and social amenities, so prevention efforts should diligently focus on this vulnerable group (CDC, 2021; WHO, 2021). Also, because my study and studies elsewhere have shown significant age differences in COVID-19 risks (Al Mutair et al., 2021; Casas-Deza et al., 2021; Gimeno-Miguel et al., 2021; Kawatake de Sousa et al., 2021), there is a need to study key contextual factors that influence how older adults’ activity risks overlap with other risks such as unemployment and poverty factors among adults ages 18 years and older.

Concerning the individual-level determinant sex, the results of the chi-square test of association between sex and COVID-19 infection diagnosis status and the binary logistic regression model indicate that there was no association between the two variables. Furthermore, in the binary logistic regression model, men were less prone to COVID-19 disease infection, demonstrating that male participants were less likely than females to be diagnosed with COVID-19. This finding is surprising; however, it is corroborated by a similar COVID-19 infection diagnosis and sex relationship study that found that women were more prone to COVID-19 infection. For example, Gimeno-Miguel et al.'s (2021) retrospective observational study out of Spain that assessed COVID-19 infection based on sex showed a COVID-19 risk of 6.19 in men and 8.19 in women aged 80 and over compared to their counterparts who were 45–64 years old. Similarly, Alam et al. (2021) conducted a web-based self-reported survey to examine the link between underlying medical conditions such as diabetes and COVID-19 infection in 780,961 male and female participants from 183 different countries. They found that even though most of the participants were male (67.39%), there was a higher link between females and COVID-19 infection with females being 1.69 times (AOR: 1.69) more likely to get COVID-19 compared to their male counterparts. Also, Fortunato et al. (2021) conducted a retrospective epidemiological study using 1,175 confirmed COVID-19 cases in Italy spanning February 29 to June 30, 2020, to examine the connection between gender and COVID-19 infection. The authors found a similar number of women and men with confirmed COVID-19 (50.7% and 49.3%, respectively) and COVID-19 infection rates (2.02%, 2.25%, respectively, $p > 0.05$).

In contrast to my study results, one previous study found that sex was associated with COVID-19 infection incidence and males were more prone to COVID-19 disease infection, demonstrating that male participants were more likely than females to be diagnosed with COVID-19 (Ramírez-Soto et al., 2021).

Studies by Cruz-Arenas et al. (2021), Fortunato et al. (2021), and Ramírez-Soto et al. (2021) pointed out that the trend seen for sex may be because of the presence of preexisting comorbidities such as cardiovascular diseases and obesity, including differences in genetic response and various hormones between males and females. Also, greater release of angiotensin-converting enzyme 2 (ACE2) receptors in males can alter binding ability for the COVID-19 virus (SARS-CoV-2) and access compared to women (Cruz-Arenas et al., 2021; Fortunato et al., 2021; Ramírez-Soto et al., 2021). The authors also mentioned that the COVID-19 susceptibility of men could be attributed to certain detrimental social behaviors or lifestyles such as smoking that heighten men's risk for chronic disease linked to COVID-19. Additionally, according to Cruz-Arenas et al. (2021), some COVID-19 patients such as women can express large amounts of antibodies that can promote stronger immune responses to effectively combat COVID-19 as compared to men. Although my study found no association between sex and COVID-19 infection diagnosis, it demonstrated that sex status factor is vital to explain the disparities that exist among males and females. However, the mechanisms through which the enzyme 2 (ACE2) works to promote COVID-19 disease in males and females also remain unclear and conflicting, so the findings of the present study warrant further clinical studies assessing sex-specific effects of ACE2 expression about COVID-1 infection incidence. Moreover, the contrasting findings of my study relative to the literature warrant more honest debate on the COVID-19 dilemma to help tailor COVID-19 interventions according to sex and promote educational programs among male and female adults aged 18 and older in Indiana.

Regarding the individual level determinant, race status, the results of the chi-square test of association between race and COVID-19 infection diagnosis status and the results of the binary logistic regression model indicated that there was a statistically significant association between the two variables. In addition, the logistic regression model indicated that Whites were less likely to be diagnosed with COVID-19 than Blacks, and participants of other race groups were also less likely to be diagnosed with COVID-19 than Blacks. These findings are comparable to studies done in the United States and elsewhere that found that having a minority status could increase the risk of COVID-19 infection diagnosis (Azar et al., 2020; Selden & Berdahl, 2020). Similarly, a recent retrospective cohort study out of the United Kingdom by Hastie et al. (2020) used the Biobank data to discover that Black and South Asian individuals had a higher chance of testing positive for COVID-19 compared to their White counterparts (Black: OR = 4.30, $p<0.001$) South Asian: OR = 2.42, $p<0.001$). also, in a cross-sectional study of 587 cases out of the United Kingdom, De Lusignan et al. (2020) found higher COVID-19 positive results among Black compared to White adults (OR = 4.75). Additionally, a 2020 Cross-sectional study out of the United Kingdom found that Blacks were more likely to test positive for COVID-19 compared to White adults after adjustment (OR = 4.75) (De Lusignan et al., 2020). The researchers also reported that people living in poor conditions and areas were more likely to test positive for COVID-19 (OR = 2.03, $p<0.0001$).

Contrary to my study results, Gold et al. (2020)'s cross-sectional study in Georgia, United States found no statistically significant differences by race in the number of people who received mechanical ventilation or died. As Creswell (2009) noted, this discrepancy in findings for race and COVID-19 association could be related to methodological differences and so warrants urgent further study.

This current study demonstrated that race status is a risk factor for COVID-19 infection among the study population. Goldstein et al. (2022); and Samuels-Kalow (2021) explained that the racial disparities seen in COVID-19 outcomes may be due in part to the wider poverty and other deuterating socioeconomic situations that place poor people such as Blacks and Hispanics at a greater risk of acquiring COVID-19. Individuals who earn low income have a greater chance of working essential jobs such as gas station attendant service jobs that subject them to work outside the home and live in crowded houses, as compared to their White counterparts. The reason may also be as Goldstein et al. (2022); and Samuels-Kalow (2021) noted that securing high-risk occupations, such as caring for sick people in hospitals, and traveling via public transportation can heighten individual's chances of interacting with COVID-19-infected patients and thus may expose them to COVID-19 virus and disease. These study results indicate that there may be a need for timely and continuing public health intervention that improves poverty and disturbing social conditions for vulnerable minority groups and provide targeted support to reduce

COVID-19 spread and incidence among older adult minorities in Indiana and elsewhere. Furthermore, these results emphasize that more individual-level data is warranted to carefully examine occupational and work environment-related exposure risks (Ronquillo De Jesús, 2022; Unruh et al., 2022) to be able to combat and reduce COVID-19 incidence among older adults in Indiana and elsewhere.

The current study utilized the individual-level determinants for COVID-19 infection diagnosis status including, age group, sex, and race which are believed to promote COVID-19 disease in vulnerable populations (Glanz et al., 2008). My study results are consistent with the SEM theory (Glanz et al., 2008) in showing that age group and race are associated with a high incidence of COVID-19, suggesting they are relevant to disparities in the pandemic.

Research Question 2

RQ2 asks the following question: Is there an association between individuals' exposure, county code of residence, and COVID-19- Diagnosis in adults 18 years and older in Indiana while controlling for age group, sex, and race?

The results of the chi-square test of association and binary logistic regression model analysis indicated a significant relationship between the individual level factor exposure and COVID-19 infection diagnosis incidence. Also, after controlling for (age group, sex, and race) variables, the model showed that respondents who reported having been exposed were more likely to have COVID-19 infection compared to those who were not exposed. These results are corroborative with patterns seen in previous studies (Fielding-Miller et al., 2020). Similarly, in the study by Selden and Berdahl (2020) in the

United States, the researchers found that Blacks who had higher chance of experiencing severe COVID-19 were 1.6 times more likely than their White counterparts to live in households containing health workers. They also reported that Hispanic adults at high risk for COVID-19 had 64.5% more chance to reside in households with individuals who work from outside the home, compared to 56.5% for Black adults and just 46.6% when it comes to White adults. Furthermore, in Cruz-Arenas et al.'s (2021) recent cross-sectional study of 2974 hospital employees out of Mexico City, the security and janitorial work groups had substantially higher rates of positive results (62.5% and 45.4%, respectively) than the other work groups. The researchers also found a stronger link between working in the security job (OR 12.35, $p = 0.027$) and janitorial job (OR 13.70 $p<0.001$) groups and COVID-19 infection.

According to Upshaw et al. (2021), a plausible explanation for the finding could be that the poor social conditions at the individual and community level, including racism may increase the likelihood of COVID-19 infection. This may also be attributed to the low-income earning problem that forces these individuals to likely hold for example occupations such as fast-food restaurant servers or live with individuals in crowded housing conditions that hinder the opportunity to practice physical distancing from others, heighten risk of COVID-19 disease exposure incidence (Upshaw et al., 2021).

This study showed that individuals exposed to COVID-19 had a greater chance of getting COVID-19 disease diagnosis, so current and future prevention efforts should focus on this at-risk group, create a conducive work and living environment, and strengthen COVID-19 health promotion and behavior educational programs among older adults aged 18 and older in Indiana and elsewhere. Also, because my study and previous studies have demonstrated serious exposure risk in COVID-109 (Upshaw et al., 2021), there is an urgent need to critically examine certain contextual factors that contribute to how for example healthcare workers, and food restaurant service activity risks overlap with other risk factors among individuals in service workers and caretakers.

Regarding the community-level social determinant of health factor county code of residence, results of the chi-square test of association between the variables indicated a statistical significance association between the variables county code of residence and COVID-19 diagnosis. Furthermore, after controlling for confounders (age group, sex, and race), the county code of residence variable was found to be statistically associated with COVID-19 infection and showed that participants that were from Marion County were less likely to have COVID-19 infection diagnosed compared to those from Lake County. My results are consistent with studies done in the United States and elsewhere that found an association between community level county code of residence and COVID-19 infection diagnosis (Azar et al., 2020; De Lusignan et al., 2020; Karmakar et al., 2021; Oluyomi et al., 2021; Palacio & Tamariz, 2021; Unruh et al., 2022). For example, the studies by Liao and De Maio (2021) used a cross-sectional ecological study design to discover that COVID-19 incidence rates were higher among counties for each percentage increase of Black residents and especially of Hispanic residents (RR, 1.042). Also, in the study by Hanson et al. (2020), that assessed the link between race and COVID-19 diagnosis by zip code in Indiana, the researchers found four times higher odds of getting COVID-19 in African Americans than in their non-African American counterparts, and higher COVID-19 cases per 10,000 people were exhibited in zip codes with higher percentage of African American or Hispanic residents. In another study conducted by Samuels-Kalow et al. (2021) in Boston spanning February 5–May 4, 2020, the researchers found a significant association between the percentage of the population that is Hispanic (IRR = 1.25), household poverty rate (IRR = 1.25) and high COVID-19 cases. They also found that places with high COVID-19 cases were characterized by a high percentage of Hispanic people (n = 72), non-Hispanic Black (n = 36), uninsured (n = 33), SNAP benefits recipients (n = 39), and poor people (n = 23), ($P < .05$) (Samuels-Kalow et al., 2021).

My result that participants that were from Marion County were less likely to have COVID-19 infection diagnosed compared to those from Lake County is surprising, however, it is in line with a similar study that examined the U.S. urban and rural areas community level factor association with COVID-19 infection diagnosis by Cuadros et al. (2021) using COVID-19 data from Johns Hopkins University. The researchers found a higher COVID-19 incidence rate in rural counties by over 180% from the first to the second period of the pandemic (RR = 2.86, $P < .001$), and by over 80% from the second to the third period (RR = 1.88, $P < .001$). But in urban counties, the incidence rate went up by only 90% from the first to the second period (RR = 1.92, $P < .001$). Moreover, they reported more than 25% higher COVID-19 incidence rate in rural in the third period compared to urban counties (RR = 1.27, $P < .001$).

The high COVID-19 infection intensity or risk found in my study for the nonmetropolitan rural county zip code areas of Lake County in Indiana is very concerning. A plausible explanation for the finding could be that this area is characterized by a high number of older people with more underlying medical conditions, high poverty status, low educational level, and high number of minorities, compared to Marion County, the largest county in the metropolitan urban city areas of Indianapolis, although Lake County has lower population density (CDC, 2021; Indiana Department of Health, 2023). Studies have shown that these characteristics are linked to high COVID-19 vulnerability (Karmakar et al., 2021). Another important explanation for the finding could be that residents in Lake County areas may have perceived their risk of contracting the COVID-19 virus as low or not severe and so local policymakers were relatively slow to put in place rigorous mitigation, intervention programs such as wearing face mask and social distancing to reduce COVID-19 exposure and contain the spread of COVID-19 even though, they have low population density (Glanz et al., 2008). The CDC (2021) noted that wearing a mask was one of the best ways people could prevent transmission of COVID-19 disease in different settings and populations. Furthermore, according to the health belief model, if an individual perceived risk and severity of acquiring COVID-19 disease is high they will be motivated to take the necessary preventive measures such as wearing face masks and social distancing to prevent COVID-19 infection (Glanz et al., 2008). My study results have demonstrated geographic patterns of COVID-19 incidence, including showing nonmetropolitan counties such as Lake County are at higher risk for COVID-19 infection. I did not investigate the relation between an individual’s perception of the severity of COVID-19 and the likelihood of these individuals contracting COVID-19 among older adults aged 18 years and older. So, I recommend future research in this area to get the full scope of the spread of the disease in nonmetropolitan counties such as Lake County and metropolitan counties such as Marion County, Indiana. Additionally, my results indicate the need for the strengthening of COVID-19 health promotion and behavior educational programs among older adults aged 18 and older in Lake County and Indiana as a whole. This will help inform public health decision-makers and policy makers know when and whether the burden of COVID-19 affected each state's nonmetropolitan or metropolitan population. It will also improve engagement in health-seeking behaviors (CDC, 2021). The CDC (2021); Little et al. (2021); and WHO (2021) suggested further studies to fully characterize Zip code/neighborhood and COVID-19 infection in various settings.

Research Question 3

RQ3 asks the following question: What is the association between underlying medical conditions (severe obesity (as measured by BMI $\geq$ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) and COVID- COVID-19 diagnosis controlling for age group, sex, and race among adults 18 years and older in Indiana? The results of the chi-square test of association and binary logistic regression model analysis indicated a significant relationship between the individual level factor underlying medical conditions and COVID-19 infection diagnosis incidence. The binary logistic regression indicated after controlling for (age group, sex, and race variables), that participants who reported having underlying medical conditions were more likely to be COVID-19 infection diagnosed than those without underlying medical conditions. My results are in line with previous study done in the United States and elsewhere that found that having underlying medical conditions such as obesity disease were associated with significantly increased risk of COVID-19 infection incidence (Al Mutair et al., 2021; Fortunato et al., 2021; Hamza et al., 2022; Kaminska et al., 2021; Kinge et al., 2022; Moazzami et al., 2020; Parohan et al., 2020; Sajjad et al., 2022; Shakor et al., 2021). Similarly, Mankowski et al.'s (2021) retrospective cohort study used 309 confirmed COVID-19 adult people in New Orleans to discover that Black COVID-19 patients were more likely to be obese (55%) than their white counterparts (36%). The researchers also pointed out that obese patients had a higher COVID-19 risk compared to nonobese patients. Also, another recent retrospective, observational study out of Spain by Gimeno-Miguel et al. (2021) found that people with underlying medical conditions such as heart disease had severe COVID-19 incidence (OR) men OR women) cardiovascular disease (1.28, 1.39) respectively. Moreover, in a recent cross-sectional study of US counties by Karmakar et al. (2021), the researchers found that obesity rate was significantly associated with COVID-19 incidence (IRR = 1.02, $P < .001$) and mortality (IRR =1.02 $P < .001$). Another retrospective observational study conducted by Gimeno-Miguel et al. (2021) reported that COVID-19 incidence was more likely severe in individuals with underlying medical conditions (OR) men, OR women) diabetes (1.37, 1.24) respectively. Additionally, a retrospective case–control study by Assaad et al. (2022) in New York State measuring if comorbidity influenced COVID-19 diagnosis discovered that more cases than controls had diabetes problems (74%), (67%), p=0.03) respectively. Similarly, Bradley et al.'s (2022) systematic review and meta-analysis study on 24 studies that included 10 648 patients assessed the link between diabetes and COVID-19 infection and found that Diabetes greatly increased the chance of severe COVID-19 (OR 3.39, P < .0001) and greater number of the patients needed mechanical ventilation (OR 3.03; P < .0001). In this same study, the prevalence of diabetes was about 31% in hospitalized COVID-19 patients, and the increased death risk was (OR 2.44; P < .0001). Furthermore, an ecological study conducted by Ernesto Noyola et al. (2021) in Mexico to measure if obesity and diabetes influenced COVID-19 incidence and mortality rates found a significantly higher association between diabetes and COVID-19 incidence (beta 0.2; p < 0.001). The researchers also found substantially higher COVID-19 mortality rates in older people suffering from diabetes (p < 0.001). Another Systematic review and Meta-Analyses study conducted by Parohan et al. (2020) using 14 retrospective studies to examine the association between age, gender, diabetes, and hypertension and COVID-19 infection in various countries. found a significant association between diabetes and risk of mortality (OR 2.41, p=.037). Moreover, similar studies by Bradley et al. (2022); Kaminska et al. (2021); Moazzami et al. (2020); Najjar et al. (2022); Ssentongo (2020); Tian et al. (2020) discovered an association between comorbidity factors and COVID-19 issues.

According to Kaminska et al. (2021); Moazzami et al. (2020); Parohan et al. (2020); Shah et al. (2021); and Zhang et al. (2021), people with underlying medical conditions may have angiotensin-converting enzyme 2 (ACE2) which promotes COVID-19 virus binding to the human cells and release of dangerous elements such as TNF-α and leptin in the body which destroy the immunity system and cause SARS-CoV2 infected patients to encounter more serious COVID-19 complications or death.

The current study demonstrates that underlying medical condition factors are vital to explain the disparities that exist among older adults in Indiana. A clear understanding of the associations between underlying medical conditions and COVID-19 would improve risk stratification and help protect those most vulnerable to severe COVID-19. My results also indicate the need for the strengthening of COVID-19 health promotion and behavior educational programs among older adults aged 18 and older in Indiana. Recently, Bradley et al. (2022) recommended further studies on the link between comorbidity such as diabetes and COVID-19 infection to gain a better insight into the COVID-19 characteristics and improve COVID-19 prevention and control strategies.

Research Question 4

RQ4 asks the following question: What are the significant predicting social determinants of health (age group, sex, race, exposure, county code of residence), underlying medical conditions (severe obesity (as measured by BMI ≥ 40 kg/m^2), prostate cancer diagnosis, diabetes, and hypertension) for COVID- 19 Diagnosis among adults ages 18 years and older during the COVID-19 pandemic in Indiana?

The results from the analysis that examined the predictive ability of the six independent variables (age group, sex, race, exposure, underlying medical conditions, and county code of residence) for COVID-19 infection diagnosis incidence using backward stepwise logistic regression demonstrated that the use of these six independent variables created statistically significant model predicting COVID-19 infection diagnosis and contributed to the variance for each. In model 3 the final and best model, age group; exposure; county code of residence, and underlying medical conditions were identified as significant independent predictors of COVID-19 diagnosis incidence. This model also indicates that (R^2) .010 or 1% of the variance in COVID-19 infection diagnosis status or in predicting COVID-19 can be explained by this model (variables). Cohen (1988) noted this as a small effect but has major practical and clinical implications given the deadliness of the COVID-19 virus disease in adults aged 18 and over years in Indiana (Creswell, 2009). The results also, indicates that the social determinants of health and underlying medical condition, particularly age group, exposure, underlying medical conditions, and county code of residence, offer significant explanatory power regarding COVID-19 incidence rates. However, after backward regression, underlying medical conditions emerged as the most unique significant risk factor associated with being a COVID-19-infected case-patient, and that statistically significantly contributed uniquely to the final model (b/beta/ $\beta = 0.059$ or 5.9%, $t = 20.549$, $p < .001$). In the hierarchical logistic regression model, the addition of the predictors exposure, county code of residence, and underlying medical conditions in Model 2 added 0.010 or 1% ($p < .001$) to the variance in Model 1, bringing the total variance explained to 1.2% ($p < .001$). Also, in Model 2, Age group status remained significantly associated with more COVID-19 diagnoses ($b = 0.027$, $\beta = 0.029$, $p = .001$), while sex ($b = -.001$, $\beta = -.001$, $p < 0.601$) and race ($b = 0.002$, $\beta = 0.003$, $p = .172$) were negatively associated with COVID-19 infection. However, in the final model, (model 2), underlying medical conditions made the most unique contribution statistically to the model ($\beta = 0.059$, $p < .001$) and was the most significant predictors of COVID-19 diagnosis, even though, exposure ($b = 0.055$, $\beta = 0.032$, $p < .001$) and county code of residence ($b = -.028$, $\beta = -.042$, $p < .001$) were also associated with COVID-19 diagnosis. The final model explained 0.8% of the overall variance ($p < .001$) and represented a 0.6% increase in the variance explained in Model 2, suggesting that the independent variables exposure, county code of residence, and underlying medical conditions explained 0.8% additional variance in outcome even after controlling for the effects of age group, sex, and race. The ANOVA result revealed that the model was a statistically significant predictor of COVID-19 while controlling for age group, sex, and race.

My results are consistent with previous research done in the United States and elsewhere that found that having underlying medical conditions status could increase the risk of COVID-19 infection diagnosis (Fortunato et al., 2021; Hamza et al., 2022; Moazzami et al., 2020; Sajjad et al., 2022; Shakor et al., 2021). Similarly, one large systematic review and meta-analysis study by Kaminska et al. (2021) unearthed that diabetes independently contributed statistically to the severity of COVID-19 in hospitals and that Diabetic patients were more likely to encounter more severe COVID-19 problems that their nondiabetic counterparts (OR=1.43; P=0.20). Moreover, an ecological study by Ernesto Noyola et al. (2021) in Mexico that evaluated the linkage between obesity and diabetes and COVID-19 incidence and death rates in the country revealed a very significant association between diabetes and COVID-19 incidence (beta 0.2; $p < 0.001$), and the death rate was also higher in this diabetic group ($p < 0.001$). Additionally, my result was supported by Parohan et al.'s (2020) recent Systematic Reviews and Meta-Analyses of 14 retrospective studies with 29,909 COVID-19-infected patients and 1,445 cases of deaths which reported a significant association between diabetes and risk of death (OR 2.41, p=.037).

Importantly, Kaminska et al. (2021); Moazzami et al. (2020); Parohan et al. (2020); Shah et al. (2021); Zhang et al. (2021) concluded that people suffering from underlying medical conditions may also harbor angiotensin converting enzyme 2 (ACE2) which promotes COVID-19 virus binding and entry into the human cells and organs to activate inflammation agents such as TNF-α, and leptin leading to poor immune response to combat COVID-19 effectively. Thus, this may cause patients to be more susceptible to developing severe cases and death from COVID-19 (Kaminska et al., 2021; Moazzami et al., 2020; Parohan et al., 2020; Shah et al., 2021; Zhang et al., 2021). The current study revealed that at the individual level, underlying medical condition status was associated with COVID-19 infection diagnosis, indicating the need for the strengthening of COVID-19 health promotion and behavior educational programs among older adults aged 18 and older in Indiana.

Study Findings in Context of the Theoretical Framework

Collectively, my findings indicate that multiple levels of the SEM framework for human health outcomes come to action when it comes to explaining COVID-19 disease. According to the SEM, the individual level variables and factors age group, sex, race, exposure, and underlying medical conditions, and the community level including county code of residence/neighbor variable or factor are believed to be the primary determinant of various health patterns and strongly influence and impact an individual's health and health behaviors (Glanz et al., 2008). Further, Glanz et al. (2008) concluded that these variables have been shown to link with health status such as the COVID-19 disease, through various interplay. Therefore, clearly understanding how these factors influence the transmission of the COVID-19 disease among older adults aged 18 and older in Indiana will be critical in controlling and preventing ongoing and future outbreaks (CDC, 2021; Glanz et al., 2008; WHO, 2021). COVID-19 prevention programs, such as health education and policies in Indiana may also benefit by influencing change at various levels of the SEM. Glanz et al. (2008) explained and emphasized that several health promotions, behavior education, and disease prevention programs utilized the multiple levels of influence of the model to combat and reduce diseases such as COVID-19 (Glanz et al., 2008).

Limitations of the Study

Although the current study highlighted significant effects of SEM variables/factors on an adult's COVID-19 infection, as with most cross-sectional studies of secondary data sets, this study had some noticeable limitations so the results should be interpreted with care. First, the study used a secondary data analysis from the COVID-19 Case Surveillance Public Use Data with Geography data set from the CDC which limits researcher's ability to define variables. Although I attempted to include main predictors from previous studies to investigate variables/factors of COVID-19 infection, there were still other potential variables excluded from the study, such as education, a likely predictor of COVID-19 infection. Second, causality cannot be established between the designated variables and the COVID-19 infection due to my cross-sectional design. Third, sample selection bias comes into play due to my study specifically focusing on adults aged 18 years and older, and Lake County and Marion County, Indiana, as these adult's social determinants of health and comorbidities variable differ from those in other age groups and places in the U.S. Thus, my findings are not necessarily applicable or generalizable to all population in Indiana, adults outside the sampled area. But still, some of my general findings and vital recommendations are likely to be important to any older adults in the whole of Indiana and elsewhere. Fourth, The COVID-19 infection information was obtained through self-response from the respondents, which can result in certain recall bias (Creswell, 2009, CDC, 2021). Fourth, the study used the 2019-2023 COVID-19 Case Surveillance Public Use Data with Geography data set from the CDC collected for Indiana. Fifth, although the data are the most current/credible COVID-19 data available from the 2019-2023 (COVID-19 Case Surveillance Public Use Data with Geography) CDC data set recent data, COVID-19 disease is new and thus some cases at the beginning of the pandemic may not be properly represented or fully captured in the database in Indiana. However, since I mainly set out to examine the effect of variables at the individual and community levels on COVID-19 infection, the data in the study were the best data available. Sixth, another major limitation of this study is that existing comorbidities (for example obesity, diabetes) were lumped under one umbrella, underlying medical conditions in the data set. So, it was difficult for me to determine the specific effect of any of the existing comorbidities on COVID-19 in this present study. Further, it made it difficult to determine whether individuals with certain comorbidities will be observed to have more risks of this disease than participants who did not have the conditions irrespective of their gender and age in Indiana. For instance, some previous epidemiologic studies have revealed a link between example, obesity or diabetes and COVID-19 incidence in certain populations (Al Mutair et al., 2021; Fortunato et al., 2021; Hamza et al., 2022; Kinge et al., 2022; Parohan et al., 2020) Finally, my study also failed to consider other epidemiological variables/factors, such as COVID-19 participants' vaccination status, types of jobs they do, and their educational level, which might affect COVID-19's transmission.

Even with these limitations, my study provides evidence of an increase in the COVID-19 infection incidence among adults ages 18 and older in Indiana. I also determined the presence of key predictors of the disease and underlying medical conditions. Furthermore, this study is one of the first to document the COVID-19 infection among adults aged 18 years and older in Indiana applying the SEM framework. Another strength of my study is based on the use of a large sample size and access to study participant records from a wide area which may have improved study power and effect size.

Recommendations

My cross-sectional and secondary data analysis study showed that underlying medical conditions variable made the most unique contribution statistically to the model ($\beta = 0.059$, $p < .001$). Therefore, in this context there is an urgent need for public health policy, COVID-19 health programs, and social interventions to: 1. Use it as a health indicator and target these underlying potential risk factors to improve COVID-19 infection prevention in adults aged 18 years and older in Indiana. This may contribute to risky behavior changes such as eating healthy, exercising to maintain a healthy weight, and avoid being for example obese, particularly in the study population. Studies by Kaminska et al. (2021); Moazzami et al. (2020); Parohan et al. (2020); Shah et al. (2021); and Zhang et al. (2021) have reported a link between underlying medical conditions and COVID-19 incidence; 2. The results of this study indicated no significant association between sex and COVID-19 infection diagnosis, so it may be useful to explore other possible factors related to sex; 3. My study results have demonstrated geographic patterns of COVID-19 incidence, including showing nonmetropolitan counties such as Lake County are at higher risk for COVID-19 infection. Likely, an individual's poor perception of the COVID-19 severity may have contributed to the high risk of the disease among older adults in Lake County (Cuadros et al., 2021). To get a better picture of this trend, I recommend future research to explore the relationship between individual's perception of COVID-19 severity and COVID-19 spread, including COVID-19 vaccine uptake status among adults in Lake County to be able to fully understand the spread of the disease in such rural county; 4. I recommend future studies in Indiana and elsewhere to examine the specific effects of individual comorbidities (for example obesity, diabetes) on COVID-19 risk among older adults in Lake County and Marion County to understand the specific effect of any of the existing comorbidities on COVID-19 spread in the adult population in these counties. I was unable to look at the specific effects of each comorbidity on COVID-19 health because the collectors of my secondary data lumped all comorbidities under one group (underlying medical conditions). Finally, further research in Indiana and elsewhere can use the COVID-19 vaccine variable as a cofactor when examining the link between social determinants of health and comorbidity factors and

COVID-19 infection, since studies show that COVID-19 vaccine uptake decision also influence COVID-19 infection (Brian, 2021). The data can provide valuable insight into the COVID-19 pandemic in the adult population in Indiana.

Implications for Professional Practice and Social Change

Professional Practice

With regards to professional practice, the results of this study revealed a significant association between age group, race, exposure, county code of residence, and underlying medical conditions and COVID-19 infection diagnosis incidence in adults aged 18 years and older in Indiana. However, the study showed that underlying medical conditions variable made the most unique contribution statistically to the model ($\beta = 0.059$, $p < .001$) and independently predicted and influenced COVID-19 infection incidence among adults aged 18 years and older in Indiana. The CDC (2021) and the WHO (2021) over the years have consistently stressed social determinants of health and underlying medical conditions research to understand and explain health disparities. So, COVID-19 health promotion, behavior health educational programs, and interventions should diligently address the underlying medical conditions and improve among older adults aged 18 years and older in Indiana. Because of the high explanatory power and unique predictability insight indicated by backward logistic and hierarchical models, COVID-19 health policies and social interventions targeting underlying medical conditions could produce profound and valuable results (Hanson et al., 2020; Smith et al., 2018). Most importantly, recent epidemiologic studies supporting social interventions highlight this approach, as this factor has been found to hinder immune response and therefore increase COVID-19 risk and severity (Bayartai et al., 2022). My findings that older adults were associated with COVID-19 risk is vital. Studies from Al Mutair et al. (2021) and Bayartai et al. (2022) have shown that older adults are likely to suffer from comorbidities like obesity and so may have a higher risk of COVID-19 disease burden. Thus, my result could assist doctors, community health workers, public health practitioners, and other health professionals to quickly identify these high-risk patients for developing severe SARS-COV-2 disease and provide them with updated COVID-19 health information which may increase their knowledge and awareness of the disease, and risk factors and help reduce their burden in Indiana. Moreover, to control, reduce or prevent COVID-19 spread among older adults in the vulnerable community such as Lake County Indiana, researchers, COVID-19 health education programmers, and health policymakers can now utilize the SEM’s individual and community levels variables to inform the design and critically evaluate interventions and programs geared towards COVID-19 risk-reduction and also helping individuals obtain community resources (Al Mutair et al., 2021; Bayartai et al., 2022; Glanz et al., 2008).

Positive Social Change

My study answered the four questions. Hence, the findings from my study can contribute to filling a knowledge gap and provide a clearer understanding of social determinants of health and preexisting comorbidities related to COVID-19 incidence and burden. Furthermore, the results of this study will contribute to positive social change by providing high-quality epidemiologic data and vital information for vulnerable populations that will better inform COVID-19 health education intervention programs and guide COVID-19 vaccine distribution uptake, help meet the essential needs of the study population, and stop the spread of the disease in Indiana and elsewhere (Samuel et al., 2021). The finding that Lake County had a higher COVID-19 risk than Marian County is crucial, and may help inform effective mitigation strategies, programs, and policies, more equitable future public health response strategies, and targeting of resources such as COVID-19 vaccine distribution to prevent or reduce the transmission and burden of COVID-19 properly and speedily in Indiana. Furthermore, positive social change can be achieved if the results of my study are utilized to foster targeted COVID-19 interventions specifically by age group, neighborhood, and comorbidity problems (Etowa et al., 2022; Hanson et al., 2020; Lamb et al., 2021; Maness et al., 2021). My results can help focus attention and increase COVID-19 and social determinants of health and comorbidity elements future research to combat COVID-19 (Hanson et al., 2020; Maness et al., 2021; Samuel et al., 2021). Furthermore, my study result indicated that the underlying medical conditions variable made the most unique contribution statistically to the model predicted and influenced COVID-19 infection incidence, so this finding has implications for positive social change if well-respected and influential public figures in the local community can be engaged to endorse and advocate underlying medical conditions risk-reduction behavior change and their benefits to peers and individuals at high-risk. Once my study findings are published and disseminated through avenues such as journals, it would provide community members, particularly older adults at high risk for COVID-19 virus infection in Indiana with concrete data and information and an improved understanding of the critical link between underlying medical conditions and

COVID-19 disease. Consequently, this would guide proper COVID-19 risk management geared towards the target population and ultimately contribute to positive social change, especially among Indiana residents aged 18 years and older. Moreover, the finding that underlying medical conditions uniquely predict poor COVID-19 health status can contribute to positive social change by providing evidence based and reliable COVID-19 data that can contribute to the understanding of COVID-19 infection trend, enhance existing data, and help strengthen ongoing comprehensive surveillance of COVID-19 in Indiana to provide a better understanding of the magnitude of COVID-19 disease. This can also inform COVIS-19 prevention programming to prioritize high-risk populations target for allocating scarce resources for COVID-19 prevention, such as COVID-19 vaccination. Further, this can help monitor and track trends in COVID-19 knowledge levels, prevention efforts, in COVID-19 vulnerable communities and populations, especially older adults aged 18 and older periodically in Indiana and evaluate COVID-19 prevention effectiveness and care activities, to ensure a quality evidence base for COVID-19 infection prevention and control policies and programs (CDC, 2021, WHO, 2021). Also, it clarifies and adds to the scarce body of literature regarding the relationship between social determinants of health and underlying medical conditions and COVID-19 in adults aged 18 years and older in Indiana, and also, highlights some of the key factors that are influencing COVID-19 disease incidence in Indiana (Hanson et al., 2020; Dixon et al., 2021; Samuel et al., 2021; Timsina et al., 2020). Moreover, behavioral change to improve underlying medical conditions issues, for example maintaining a healthy diet and exercise to reduce obesity strategies and programs should target both the social context such as discrimination and poverty, and the individual factors such as age since there have been reports of known severity of COVID-19 in elderly, poor, and marginalized patients in Indiana (Hanson et al., 2020; Dixon et al., 2021; Samuel et al., 2021; Timsina et al., 2020). Thus, this can bring about positive social change. COVID-19 programs and health promotion education should take place at the individualistic model and the community level to contribute to positive social change.

Conclusion

The ongoing COVID-19 disease pandemic is a public health issue that requires much attention and intervention by public officers and policymakers. Research is scarce in Indiana on the potential link between COVID-19 diagnosis and social determinants of health and underlying medical conditions variables/factors (Hanson et al., 2020; Dixon et al., 2021; Samuel et al., 2021; Timsina et al., 2020). Furthermore, the few studies that have been conducted in Indiana and elsewhere have reported inconsistent results and have also focused on the general population (Hanson et al., 2020; Dixon et al., 2021; Samuel et al., 2021; Timsina et al., 2020). The current study examined these potential links and discovered that when controlling for age group, sex, and race, COVID-19 diagnosis incidence is significantly associated with exposure, county code of residence, and underlying medical conditions. However, underlying medical conditions emerged as the variable that made the most significant unique contribution to the variance in COVID-19 risk. These findings build on previous evidence of the link between social determinants of health variables and underlying medical conditions and COVID-19 disease incidence (Hanson et al., 2020; Dixon et al., 2021; Samuel et al., 2021; Timsina et al., 2020), and should be taken into account when considering how to design and implement targeted COVID-19 intervention and prevention policies and health education programs in Indiana (CDC, 2021; WHO, 2021). Although these significant links were demonstrated in this study, additional research is warranted to create targeted COVID-19 intervention and prevention policies and programs that will fully benefit older adults aged 18 years and older in Indiana and elsewhere. The CDC (2021) and the WHO (2021) have continually emphasized a continued need to assess predictive factors of diseases such as COVID-19. My study used the social determinants of health and underlying medical conditions to accomplish this. My study emphasizes how there are the social determinants of health variables and underlying medical condition elements that heighten some people's vulnerability to COVID-19 infection more than other people. The current study results support this claim as variables at the individual level of the SEM such as underlying medical conditions have had a statistically significant impact on COVID-19 infection diagnosis incidence among adults aged 18 years and older in Indiana. Also, a significant variation in the risk of COVID-19 incidence was observed at the county zip code community level. Even though my study utilized a secondary dataset analysis and has its limitations, I believe that it provides sound evidence that secondary analysis of data using backward elimination logistic regression, hierarchal logistic regression models, and the SEM framework can yield tremendous opportunity for determining the link between individual and community level variables and COVID-19 among adults aged 18 and older in Indiana. Moreover, it high lights the significance of using social determinants of health and underlying medical conditions variables as COVID-19 health indicators and shed light on areas for COVID-19 health policy, health education program, and social interventions. In addition, the data from my study contributes to a clear understanding of COVID-19 characteristics and also adds to the knowledge of COVID-19 health educators, policymakers, and practitioners who are eager to stop the ongoing COVID-19 pandemic and prevent future crises in Indiana and elsewhere by highlighting the key variable (underlying medical conditions) that uniquely influence COVID-19 incidence in Indiana, hence promotes positive social change also. Public health officers, health education programmers, and practitioners in Indiana should use the results and my study recommendations to guide future research and to create and implement effective COVID-19 health programs and policies that will reduce the spread, incidence, and overall burden of COVID-19 in Indiana and eventually result in the improvement of people's lives in Indiana and elsewhere.

References

Abate, B. B., Kassie, A. M., Kassaw, M. W., Aragie, T .G., & Masresha, S. A. (2020). Sex difference in coronavirus disease (COVID-19): A systematic review and meta-analysis. *BMJ Open*, *10*, Article e040129. https://doi.org/10.1136/bmjopen-2020-040129

Abebe, E. C., Dejenie, T. A., Shiferaw, M. Y., & Malik, T. (2020). The newly emerged COVID-19 disease: A systemic review. *Virology Journal*, *17*(1), 1–8. https://doi.org/10.1186/s12985-020-01363-5

Abedi, V., Olulana, O., Avula, V., Chaudhary, D., Khan, A., Shahjouei, S., Li, J., & Zand, R. (2021). Racial, economic, and health inequality and COVID-19 infection in the United States. *Journal of Racial and Ethnic Health Disparities*, *8*, 732–742. https://doi.org/10.1007/s40615-020-00833-4

Ahmad, K., Erqou, S., Shah, N., Nazir, U., Morrison, A. R., Choudhary, G., & Wu, W.-C. (2020). Association of poor housing conditions with COVID-19 incidence and mortality across US counties. *PLoS ONE*, *15*(11), 1–13. https://doi.org/10.1371/journal.pone.0241327

Alam, M. R., Kabir, R., & Reza, S. (2021). Comorbidities might be a risk factor for the incidence of COVID-19: Evidence from a web-based survey. *Preventive Medicine Reports*, *21*, Article 101319. https://doi.org/10.1016/j.pmedr.2021.101319

Aldila, D., Samiadji, B. M., Simorangkir, G. M., Khosnaw, S. H. A., & Shahzad, M. (2021). Impact of early detection and vaccination strategy in COVID-19 eradication program in Jakarta, Indonesia. *BMC Research Notes*, *14*(1), 1–7. https://doi.org/10.1186/s13104-021-05540-9

Al Mutair, A., Al Mutairi, A., Alhumaid, S., Maaz Abdullah, S., Zia Zaidi, A. R., Rabaan, A. A., & Al-Omari, A. (2021). Examining and investigating the impact of demographic characteristics and chronic diseases on mortality of COVID-19: Retrospective study. *PLoS ONE*, *16*(9), 1–8. https://doi.org/10.1371/journal.pone.0257131

Assaad, M., Hekmat-Joo, N., Hosry, J., Kassem, A., Itani, A., Dahabra, L., Abou Yassine, A., Zaidan, J., & El Sayegh, D. (2022). Insulin use in Type 2 diabetic patients: A predictive of mortality in Covid-19 infection. *Diabetology & Metabolic Syndrome*, *14*(1), 1–8. https://doi.org/10.1186/s13098-022-00857-2

Azam, M., Sulistiana, R., Ratnawati, M., Fibriana, A. I., Bahrudin, U., Widyaningrum, D., & Aljunid, S. M. (2020). Recurrent SARS-CoV-2 RNA positivity after COVID-19: A systematic review and meta-analysis. *Scientific Reports*, *10*, Article 20692. https://doi.org/10.1038/s41598-020-77739-y

Banks, D. E., Hensel, D. J., & Zapolski, T. C. B. (2020). Integrating individual and contextual factors to explain disparities in HIV/STI among heterosexual African American youth: A contemporary literature review and social ecological model. *Archives of Sexual Behavior*, *49*(6), 1939–1964. https://doi.org/10.1007/s10508-019-01609-6

Bathina, K. C., ten Thij, M., Valdez, D., Rutter, L. A., & Bollen, J. (2021). Declining well-being during the COVID-19 pandemic reveals US social inequities. *PLoS ONE*, *16*(7), Article e0254114. https://doi.org/10.1371/journal.pone.0254114

Bayartai, M., Luomajoki, H., De Micheli, R., Tringali, G., Marazzi, N., & Sartorio, A. (2022). Changes in the Oswestry Disability Index after a 3-week in-patient multidisciplinary body weight reduction program in adults with obesity. *Journal of Clinical Medicine*, *11*(11), Article 3175. https://doi.org/10.3390/jcm11113175

Biswas, M., Rahaman, S., Biswas, T. K., Haque, Z., & Ibrahim, B. (2021). Association of sex, age, and comorbidities with mortality in COVID-19 patients: A systematic review and meta-analysis. *Intervirology*, *64*(1), 36–47. https://doi.org/10.1159/000512592

Boserup, B., McKenney, M., & Elkbuli, A. (2020). Disproportionate Impact of COVID-19 pandemic on racial and ethnic minorities. *American Surgeon*, *86*(12), 1615–1622. https://doi.org/10.1177/0003134820973356

Bradley, S. A., Banach, M., Alvarado, N., Smokovski, I., & Bhaskar, S. M. M. (2022). Prevalence and impact of diabetes in hospitalized COVID-19 patients: A systematic review and meta-analysis. *Journal of Diabetes*, *14*(2), 144–157. https://doi.org/10.1111/1753-0407.13243

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

Brown, C., Wilkins, K., Craig-Neil, A., Upshaw, T., & Pinto, A. D. (2022). Reducing inequities during the COVID-19 pandemic: A rapid review and synthesis of public health recommendations. *Public Health Reviews*, *42*, Article 1604031. https://doi.org/10.3389/phrs.2021.1604031

Bradley, S. A., Banach, M., Alvarado, N., Smokovski, I., & Bhaskar, S. M. M. (2022). Prevalence and impact of diabetes in hospitalized COVID-19 patients: A systematic review and meta-analysis. *Journal of Diabetes*, *14*(2), 144–157. https://doi.org/10.1111/1753-0407.13243

Burton, K., Ayangeakaa, S., Kerr, J., Kershner, S., & Maticka-Tyndale, E. (2019). Examining sexual concurrency and a number of partners among African, Caribbean, and Black women using the social-ecological model: Results from the ACBY study. *Canadian Journal of Human Sexuality*, *28*(1), 46–56. https://doi.org/10.3138/cjhs.2018-0033

Cai, M., Shah, N., Li, J., Chen, W.-H., Cuomo, R. E., Obradovich, N., & Mackey, T. K. (2020). Identification and characterization of tweets related to the 2015 Indiana HIV outbreak: A retrospective infoveillance study. *PLoS ONE*, *15*(8), 1–13. https://doi.org/10.1371/journal.pone.0235150

Casas-Deza, D., Bernal-Monterde, V., Aranda-Alonso, A. N., Montil-Miguel, E., Julián-Gomara, A. B., Letona-Giménez, L., & Arbones-Mainar, J. M. (2021). Age-related mortality in 61,993 confirmed COVID-19 cases over three epidemic waves in Aragon, Spain. Implications for vaccination programmes. *PLoS ONE*, *16*(12), 1–13. https://doi.org/10.1371/journal.pone.0261061

Casillas Jr, E., Wu, G., Iantorno, S., Ning, W. V., Joon Choi, Chan, P., & Lee, M. M. (2021). COVID-19: Highlighting Health Disparities in the Los Angeles Latinx

Community. *Clinical Medicine & Research*, *19*(4), 161–168. https://doi.org/10.3121/cmr.2021.1654

Cuadros, D. F., Branscum, A. J., Mukandavire, Z., Miller, F. D., & MacKinnon, N. (2021). Dynamics of the COVID-19 epidemic in urban and rural areas in the United States. *Annals of Epidemiology*, *59*, 16–20. https://doi.org/10.1016/j.annepidem.2021.04.007

Centers for Disease Control and Prevention. Health Equity Considerations and Racial and Ethnic Minority Groups. Retrieved July 7, 2022, from https://www.cdc.gov/coronavirus/2019- ncov/community/health-equity/race-ethnicity.html

Centers for Disease Control and Prevention. (2021). Alzheimer’s disease and healthy aging. https://www.cdc.gov/aging/aginginfo/alzheimers.htm

Centers for Disease Control and Prevention. (2022). Alzheimer’s disease and healthy aging. https://www.cdc.gov/aging/aginginfo/alzheimers.htm

Centers for Disease Control and Prevention. (2023). COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker/#datatracker-home

Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases. (2022). Similarities and Differences between Flu and COVID-19. Retrieved December 13 from https://www.cdc.gov/flu/symptoms/flu-vs-covid19.htm

Chopel, A. M., Minkler, M., Nuru-Jeter, A., & Dunbar, M. (2015). Social Determinants of Late-Stage HIV Diagnosis and its Distributions among African Americans and

Latinos: A critical literature review. *Journal of Health Disparities Research & Practice*, *8*(4), 1–29.

Christian, S. N., Mantell, J. E., Romo, M. L., Grov, C., George, G., Mwai, E., Nyaga, E., Odhiambo, J. O., Govender, K., & Kelvin, E. A. (2020). Applying a social-ecological lens to opinions about HIV self-testing among Kenyan truckers who declined to test: a qualitative study. *African Journal of AIDS Research (AJAR)*, *19*(2), 147–155. https://doi.org/10.2989/16085906.2020.1764070

Christie, A., Brooks, J. T., Hicks, L. A., Sauber-Schatz, E. K., Yoder, J. S., Honein, M. A., & CDC COVID-19 Response Team (2021). Guidance for Implementing COVID-19 Prevention Strategies in the Context of Varying Community Transmission Levels and Vaccination Coverage. *MMWR. Morbidity and mortality weekly report*, *70*(30), 1044–1047. https://doi.org/10.15585/mmwr.mm7030e2

Christie, A., Henley, S. J., Mattocks, L., Fernando, R., Lansky, A., Ahmad, F. B., Adjemian, J., Anderson, R. N., Binder, A. M., Carey, K., Dee, D. L., Dias, T., Duck, W. M., Gaughan, D. M., Lyons, B. C., McNaghten, A. D., Park, M. M., Reses, H., Rodgers, L., Van Santen, K., … Beach, M. J. (2021). Decreases in COVID-19 Cases, Emergency Department Visits, Hospital Admissions, and Deaths Among Older Adults Following the Introduction of COVID-19 Vaccine - United States, September 6, 2020-May 1, 2021. *MMWR. Morbidity and mortality weekly report*, *70*(23), 858–864. https://doi.org/10.15585/mmwr.mm7023e2

Clouston, S. A. P., Natale, G., & Link, B. G. (2021). Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. *Social Science & Medicine*, *268*, N.PAG. https://doi.org/10.1016/j.socscimed.2020.113554

Commission on Social Determinants of Health. (2008). *Closing the gap in a generation: Health equity through action on the social determinants of health*. Final report of the Commission on Social Determinants of Health. Geneva: World Health Organization. [Google Scholar].

Creswell, J. (2009). Research design: Qualitative, quantitative, and mixed methods Approaches (Laureate Education, Inc., custom Ed.). Thousand Oaks, CA: Sage Publications.

Cruz-Arenas E, Cabrera-Ruiz E, LagunaBarcenas S, Colin-Castro CA, Chavez T, FrancoCendejas R, et al. (2021) Serological prevalence of SARS-CoV-2 infection and associated factors in healthcare workers in a “non-COVID” hospital in Mexico City. PLoS ONE 16(8): e0255916. https://doi.org/10.1371/journal.pone.0255916

Cuomo, R. E., Cai, M., Shah, N., Li, J., Chen, W., Obradovich, N., Mackey, T. K., & Chen, W.-H. (2020). Characterising communities impacted by the 2015 Indiana HIV outbreak: A big data analysis of social media messages associated with HIV and substance abuse. *Drug & Alcohol Review*, *39*(7), 908–913. https://doi.org/10.1111/dar.13091

Dalsania, A.K., Fastiggi, M.J., Kahlam, A. *et al.* The Relationship Between Social Determinants of Health and Racial Disparities in COVID-19 Mortality. *J. Racial* *and Ethnic Health Disparities* (2021). https://doi.org/10.1007/s40615-020-00952-y

Diego F. Cuadros, Adam J. Branscum, Zindoga Mukandavire, F. DeWolfe Miller, Neil MacKinnon (2021). Dynamics of the COVID-19 epidemic in urban and rural areas in the United States. Annals of Epidemiology, 59, 16-20. https://doi.org/10.1016/j.annepidem.2021.04.007

Dunbar, S. (2020). As the pandemic meets flu season, labs turn to rapid molecular testing. *MLO: Medical Laboratory Observer*, *52*(9), 8–14.

Eman M khedr, Enas Daef, Aliae Mohamed-Hussein, Ehab F Mostafa, Mohamed zein, Sahar M Hassany, Hanan Galal, Shimaa Abbas Hassan, Islam Galal, Amro A. Zarzour, Helal F Hetta, Hebatallah M. Hassan, Mariam Taher Amin, Maiada k Hashem (2020). Impact of comorbidities on COVID-19 outcome. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery.* https://doi.org/10.1101/2020.11.28.20240267

Emeruwa UN, Ona S, Shaman JL, Turitz A, Wright JD, Gyamfi-Bannerman C, et al. Associations Between Built Environment, Neighborhood Socioeconomic Status, and SARS-CoV-2 Infection Among Pregnant Women in New York City. *Jama*. 2020; 324(4):390–2. https://doi.org/10.1001/jama.2020.11370

Ernesto Noyola, D., Hermosillo-Arredondo, N., Ramírez-Juárez, C., & Werge-Sánchez, A. (2021). Association between obesity and diabetes prevalence and COVID-19 mortality in Mexico: an ecological study. *Journal of Infection in Developing Countries*, *15*(10), 1396–1403. https://doi.org/10.3855/jidc.15075

Etowa, J., Demeke, J., Abrha, G., Worku, F., Ajiboye, W., Beauchamp, S., Taiwo, I., Pascal, D., & Ghose, B. (2022). Social determinants of the disproportionately higher rates of COVID-19 infection among African Caribbean and Black (ACB) population: A systematic review protocol. *Journal of Public Health Research*, *11*(2), 20–24. https://doi.org/10.4081/jphr.2021.2274

Falcó, A., González-Rubio, F., Ioakeim-Skoufa, I., Pico-Soler, V., Aza-Pascual-Salcedo, M., Prados-Torres, A., & Gimeno-Feliu, L. A. (2021). Chronic diseases associated with increased likelihood of hospitalization and mortality in 68,913 COVID-19 confirmed cases in Spain: A population-based cohort study. *PLoS ONE*, *16*(11), 1–14. https://doi.org/10.1371/journal.pone.0259822

Faul, F., Erdfelder, E., Land, A., and Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences.

Fielding-Miller RK, Sundaram ME, Brouwer K (2020) Social determinants of COVID-19 mortality at the county level. PLoS ONE 15(10): e0240151. https://doi.org/10.1371/journal.pone.0240151

Figueroa, J. F., Wadhera, R. K., Lee, D., Yeh, R. W., & Sommers, B. D. (2020). Community-Level Factors Associated with Racial And Ethnic Disparities In COVID-19 Rates In Massachusetts. *Health Affairs*, *39*(11), 1984–1992. https://doi.org/10.1377/hlthaff.2020.01040

Fortunato F, Martinelli D, Lo Caputo S, Santantonio T, Dattoli V, Lopalco PL, Prato R. (2021). Sex and gender differences in COVID-19: an Italian local register-based study. *BMJ* Open. 7;11(10):e051506. https://doi.org/10.1136/bmjopen-2021-051506

Frankfort-Nachmias, C., & Nachmias, D. (2008). Research methods in the social sciences Worth publisherss, New York, N.Y.

Gimeno-Miguel, A., Bliek-Bueno, K., Poblador-Plou, B., Carmona-Pírez, J., Poncel-Moazzami, B., Chaichian, S., Kasaeian, A., Djalalinia, S., Akhlaghdoust, M., Eslami, M., & Broumand, B. (2020). Metabolic risk factors and risk of Covid-19: A systematic review and meta-analysis. *PLoS ONE*, *15*(12), 1–14. https://doi.org/10.1371/journal.pone.0243600

Gimeno-Miguel, A., Bliek-Bueno, K., Poblador-Plou, B., Carmona-Pírez, J., Poncel-Falcó, A., González-Rubio, F., Ioakeim-Skoufa, I., Pico-Soler, V., Aza-Pascual-Salcedo, M., Prados-Torres, A., & Gimeno-Feliu, L. A. (2021). Chronic diseases associated with increased likelihood of hospitalization and mortality in 68,913 COVID-19 confirmed cases in Spain: A population-based cohort study. *PLoS ONE*, *16*(11), 1–14. https://doi.org/10.1371/journal.pone.0259822

Glanz, K., Rimer, B. K., and Viswanath, K. (2008). Health behavior and health education Theory, Research, and Practice (4th ed). San Francisco, CA. Jossey-Bass.

Gralinski, L. E., & Menachery, V. D. (2020). Return of the Coronavirus: 2019 -nCoV. *Viruses (1999-4915), 12*(2), 135. https://doi.org/10.3390/v12020135

Goldstein, N. D., Webster, J. L., Robinson, L. F., & Welles, S. L. (2022). Disparities of COVID-19 and HIV Occurrence Based on Neighborhood Infection Incidence in Philadelphia, Pennsylvania. *American Journal of Public Health*, *112*(3), 408–416. https://doi.org/10.2105/ajph.2021.306538

Green, S., B., & Salkind, N., J. (2014). Using SPSS for windows and macintosh: analyzing and understanding data. Pearson, Edition 7.

Green, H., Fernandez, R., & MacPhail, C. (2021). The social determinants of health and health outcomes among adults during the COVID-19 pandemic: A systematic review. *Public Health Nursing*, *38*(6), 942–952. https://doi.org/10.1111/phn.12959

Green, M. S., Nitzan, D., Schwartz, N., Niv, Y., & Peer, V. (2021). Sex differences in the case-fatality rates for COVID-19—A comparison of the age-related differences and consistency over seven countries. *PLoS ONE*, *16*(4), 1–13. https://doi.org/10.1371/journal.pone.0250523

Hamza, A., Shah, N., Azad, A., Ghanshyam, O., & Khan, Z. (2022). Impact of age, gender and comorbidities affecting the severity of COVID-19 infection in Kashmir. *Journal of Family Medicine & Primary Care*, *11*(4), 1519–1524. https://doi.org/10.4103/jfmpc.jfmpc_278_21

Hanson, A. E., Hains, D. S., Schwaderer, A. L., & Starr, M. C. (2020). Variation in COVID-19 Diagnosis by Zip Code and Race and Ethnicity in Indiana. *Frontiers in public health*, *8*, 593861. https://doi.org/10.3389/fpubh.2020.593861

Hatipoğlu, N. (2020). The “New” Problem of Humanity: New Coronavirus (2019-nCoV / COVID-19) Disease. *Medical Journal of Bakirkoy*, *16*(1), 1–8. https://doi.org/10.5222/BMJ.2020.22931

He, J., Liu, L., Chen, X., Qi, B., Liu, Y., Zhang, Y., & Bai, J. (2021). The experiences of nurses infected with COVID-19 in Wuhan, China: A qualitative study. *Journal of Nursing Management (John Wiley & Sons, Inc.)*, *29*(5), 1180–1188. https://doi.org/10.1111/jonm.13256

Hege, A., Lane, S., Spaulding, T., Sugg, M., & Iyer, L. S. (2022). County-Level Social Determinants of Health and COVID-19 in Nursing Homes, United States, June 1, 2020–January 31, 2021. *Public Health Reports*, *137*(1), 137–148. https://doi.org/10.1177/00333549211053666

Hong Liu , Shiyan Chen , Min Liu , Hao Nie , Hongyun Lu. Comorbid Chronic Diseases are Strongly Correlated with Disease Severity among COVID-19 Patients: A Systematic Review and Meta-Analysis. *Aging and disease*. 2020, 11(3): 668-678. https://doi.org/10.14336/AD.2020.0502

Iacob, C. I., Ionescu, D., Avram, E., & Cojocaru, D. (2021). COVID-19 Pandemic Worry and Vaccination Intention: The Mediating Role of the Health Belief Model Components. *Frontiers in psychology*, *12*, 674018. https://doi.org/10.3389/fpsyg.2021.674018

Iezadi, S., Gholipour, K., Azami-Aghdash, S., Ghiasi, A., Rezapour, A., Pourasghari, H.,

& Pashazadeh, F. (2021). Effectiveness of non-pharmaceutical public health interventions against COVID-19: A systematic review and meta-analysis. *PLoS ONE*, *16*(11), 1–19. https://doi.org/10.1371/journal.pone.0260371

Indiana Department of Health. (2023). *Indiana COVID-19 home dashboard*. https://www.coronavirus.in.gov/indiana-covid-19-dashboard-and-map/

Jin J-M, Bai P, He W, Wu F, Liu X-F, Han D-M, Liu S and Yang J-K (2020) Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. *Frontiers in Public Health* 8:152. https://doi.org/10.3389/fpubh.2020.00152

Kaminska, H., Szarpak, L., Kosior, D., Wieczorek, W., Szarpak, A., Al-Jeabory, M., Gawel, W., Gasecka, A., Jaguszewski, M. J., & Jarosz-Chobot, P. (2021). Impact of diabetes mellitus on in-hospital mortality in adult patients with COVID-19: a systematic review and meta-analysis. *Acta Diabetologica*, *58*(8), 1101–1110. https://doi.org/10.1007/s00592-021-01701-1

Kang, H. (2021). Sample size determination and power analysis using the G*Power software *Journal of Educational Evaluation for Health Professions*; 18(17). https://doi.org/10.3352/jeehp.2021.18.17

Karmakar M, Lantz PM, Tipirneni R. (2021). Association of Social and Demographic Factors With COVID-19 Incidence and Death Rates in the US. *JAMA Netw Open.* ;4(1):e2036462. https://doi.org/10.1001/jamanetworkopen.2020.36462

Kawatake de Sousa, C. D., Cristina Morais, T., Guerrero Daboin, B. E., Portugal, I., Emidio Cavalcanti, M. P., de Oliveira Echeimberg, J., Cauê Jacintho, L., Daminello Raimundo, R., Elmusharaf, K., & Eduardo Siqueira, C. (2021).

Epidemiological profile of COVID-19 in the State of Espírito Santo, Brazil, from March 2020 to June 2021. *Revista Brasileira de Crescimento e Desenvolvimento Humano*, *31*(3), 507–520. https://doi.org/10.36311/jhgd.v31.12770

Khanijahani, A., Iezadi, S., Gholipour, K., Azami-Aghdash, S., & Naghibi, D. (2021). A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. *International Journal for Equity in Health*, *20*(1), 1–30. https://doi.org/10.1186/s12939-021-01582-4

Kinge, K., Chavhan, S., Adsul, B., Kumbhar, M., Gokhale, C., & Ingale, A. (2022). An observational study to find association between Hypertension and severe and fatal COVID-19 infection in COVID-19 dedicated hospital, Mumbai. *Journal of Family Medicine & Primary Care*, *11*(1), 277–280. https://doi.org/10.4103/jfmpc.jfmpc_254_21

Kjøllesdal, M., Skyrud, K., Gele, A., Arnesen, T., Kløvstad, H., Diaz, E., & Indseth, T. (2022). The correlation between socioeconomic factors and COVID-19 among immigrants in Norway: a register-based study. *Scandinavian Journal of Public Health*, *50*(1), 52–60. https://doi.org/10.1177/14034948211015860

Lee, H., Andrasfay, T., Riley, A., Wu, Q., & Crimmins, E. (2022). Do social determinants of health explain racial/ethnic disparities in COVID-19 infection? *Social Science & Medicine*, *306*, N.PAG. https://doi.org/10.1016/j.socscimed.2022.115098

Li, B., Zeng, L., Sun, N., Zhao, Y., Zhao, F., Bian, H., Yi, W., Yang, J., Li, B., & Su, G.

(2021). Hypertension May Reduce the Infection Risk but Increase the Severity of COVID-19: Based on the Current Data in China. *International Journal of Hypertension*, 1–12. https://doi.org/10.1155/2021/6594863

Li, Z., Wei, Y., Zhu, G., Wang, M., & Zhang, L. (2022). Cancers and COVID-19 Risk: A Mendelian Randomization Study. *Cancers*, *14*(9), N.PAG. https://doi.org/10.3390/cancers14092086

Liao TF, De Maio F. (2021). Association of Social and Economic Inequality with Coronavirus Disease 2019 Incidence and Mortality Across US Counties. *JAMA,* 4(1):e2034578. https://doi.org/10.1001/jamanetworkopen.2020.34578

Little, C., Alsen, M., Barlow, J., Naymagon, L., Tremblay, D., Genden, E., Trosman, S., Iavicoli, L., & van Gerwen, M. (2021). The Impact of Socioeconomic Status on the Clinical Outcomes of COVID-19; a Retrospective Cohort Study. *Journal of Community Health*, *46*(4), 794–802. https://doi.org/10.1007/s10900-020-00944-3

Maness, S. B., Merrell, L., Thompson, E. L., Griner, S. B., Kline, N., & Wheldon, C. (2021). Social Determinants of Health and Health Disparities: COVID-19 Exposures and Mortality Among African American People in the United States. *Public Health Reports*, *136*(1), 18–22. https://doi.org/10.1177/0033354920969169

Mankowski, N., Al-Qurayshi, Z., Souza, S., Campbell, B., Beighley, A., Denson, J., Mauldin, B., Bojanowski, C., Friedlander, P., & Zifodya, J. (2021). The effect of race, socioeconomic status, and comorbidity on patients afflicted with COVID 19:

A Local Perspective. *Annals of Epidemiology*, *64*, 83–87. https://doi.org/10.1016/j.annepidem.2021.09.013

Marion County Public Health Department (MCPHD) COVID-19 Dashboard (2023). Retrieved April 9 from https://marionhealth.org/homeslider/latest-on-coronavirus/

Maroko, A. R., Nash, D., & Pavilonis, B. T. (2020). COVID-19 and Inequity: a Comparative Spatial Analysis of New York City and Chicago Hot Spots. *Journal of Urban Health*, *97*(4), 461–470. https://doi.org/10.1007/s11524-020-00468-0

McLeroy, K. R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly*, *15*(4), 351–377. https://doi.org/10.1177/109019818801500401

Mena, G. E., Martinez, P. P., Mahmud, A. S., Marquet, P. A., Buckee, C. O., Santillana, M. (2021). Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science:* 372(6545). https://doi.org/10.1126/science.abg5298

Moazzami, B., Chaichian, S., Kasaeian, A., Djalalinia, S., Akhlaghdoust, M., Eslami, M., & Broumand, B. (2020). Metabolic risk factors and risk of Covid-19: A systematic review and meta-analysis. *PLoS ONE*, *15*(12), 1–14. https://doi.org/10.1371/journal.pone.0243600

Najjar, M., Albuaini, S., Fadel, M., Aljbawi, A., Mohsen, F., Sulaiman, S., & Koudsi, A. (2022). COVID-19 Disease in Syrian Patients With Cancer: Clinical Manifestations, Laboratory Findings, Treatment, and Outcomes. *JCO Global Oncology*, *8*, 1–14. https://doi.org/10.1200/GO.21.00283

Noh, K., & Min, J. J. (2020). Understanding School-Aged Childhood Obesity of Body Mass Index: Application of the Social-Ecological Framework. *Children (Basel, Switzerland)*, *7*(9), 134. https://doi.org/10.3390/children7090134

Ochieng, J. M., & Crist, J. D. (2021). Social Determinants of Health and Health Care Delivery: African American Women's T2DM Self-Management. *Clinical Nursing Research*, *30*(3), 263–272. https://doi.org/10.1177/1054773820916981

O'Donoghue, G., Perchoux, C., Mensah, K. *et al.* A systematic review of correlates of sedentary behaviour in adults aged 18–65 years: a socio-ecological approach. *BMC Public Health* **16,** 163 (2016). https://doi.org/10.1186/s12889-016-2841-3

Oliveira, A. R. S., Cohnstaedt, L. W., & Cernicchiaro, N. (2018). Japanese encephalitis virus: Placing disease vectors in the epidemiologic triad. *Annals of the Entomological Society of America*, *111*(6), 295–303. https://doi.org/10.1093/aesa/say025

Ortolan, A., Lorenzin, M., Felicetti, M,. Doria, A., Ramonda,. R. (2020). Does gender influence clinical expression and disease outcomes in COVID-19? A systematic review and meta-analysis *International Journal of Infectious Diseases*, 99, 496-504. https://doi.org/10.1016/j.ijid.2020.07.076

Ossei, P. P. S., Ayibor, W. G., Taylor, J., Agyemang, L., Aninkora, K. O., & Asante, B. O. (2021). COVID-19 Impact on Medical Practice in Sub-Saharan Africa; The Need to Guard Against Medical Negligence: A Case Report in a Health Care

Facility in Kumasi, Ghana. *Clinical Pathology*, 1–6. https://doi.org/10.1177/2632010X211025308

Palacio, A., & Tamariz, L. (2021). Social Determinants of Health Mediate COVID-19 Disparities in South Florida. *JGIM: Journal of General Internal Medicine*, *36*(2), 472–477. https://doi.org/10.1007/s11606-020-06341-9

Parohan, M., Yaghoubi, S., Seraji, A., Javanbakht, M. H., Sarraf, P., & Djalali, M. (2020). Risk factors for mortality in patients with Coronavirus disease 2019 (COVID-19) infection: a systematic review and meta-analysis of observational studies. *Aging Male*, *23*(5), 1416–1424. https://doi.org/10.1080/13685538.2020.1774748

Qeadan, F., VanSant-Webb, E., Tingey, B., Rogers, T. N., Brooks, E., Mensah, N. A., Winkfield, K. M., Saeed, A. I., English, K., & Rogers, C. R. (2021). Racial disparities in COVID-19 outcomes exist despite comparable Elixhauser comorbidity indices between Blacks, Hispanics, Native Americans, and Whites. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-88308-2

Quandt, S. A., LaMonto, N. J., Mora, D. C., Talton, J. W., Laurienti, P. J., and Arcury, T. A. (2020). COVID-19 Pandemic among Latinx Farmworker and Nonfarmworker Families in North Carolina: Knowledge, Risk Perceptions, and Preventive Behaviors. *International Journal of Environmental Research. Public Health*, 17, 5786. https://doi.org/10.3390/ijerph17165786

Ramírez-Soto, M. C., Arroyo-Hernández, H., & Ortega-Cáceres, G. (2021). Sex differences in the incidence, mortality, and fatality of COVID-19 in Peru. *PLoS ONE*, *16*(6), 1–10. https://doi.org/10.1371/journal.pone.0253193

Roberts Kennedy, B. (2021). African Americans and Covid-19: A Multifaceted Model of Biopsychosocial- Spiritual/Cultural Factors Addressing Disparities in Increased Covid-19 Infection. *Journal of Cultural Diversity*, *28*(4), 88–97.

Robinson, T. Applying the Socio-ecological Model to Improving Fruit and Vegetable Intake Among Low-Income African Americans. *J Community Health* **33,** 395–406 (2008). https://doi.org/10.1007/s10900-008-9109-5

Rodríguez del Águila, M. M., González-Ramírez, A. R. (2014). Sample size calculation, *Allergologia et Immunopathologia*, 42(5), 485-492. https://doi.org/10.1016/j.aller.2013.03.008

Rollston, R., & Galea, S. (2020). COVID-19 and the Social Determinants of Health. *American Journal of Health Promotion*, *34*(6), 687–689. https://doi.org/10.1177/0890117120930536b

Ronquillo De Jesús, E., Suárez, V., López-Perea, P., Suárez Quezada, M., & von Borries-Medrano, E. (2022). COVID-19 cases, hospitalizations, outpatients, and deaths in Mexico by ethnicity and state-level income. *Journal of Infection in Developing Countries*, *16*(6), 969–980. https://doi.org/10.3855/jidc.15618

Sajjad, M. M., Nasir, A., Yousaf, S., & Rahim, M. W. A. (2022). Obesity as A Risk Factor For Severe COVID-19 Disease. *Pakistan Armed Forces Medical Journal*, *72*(1), 51–53. https://doi.org/10.51253/pafmj.v72i1.4477

Salgado de Snyder, V. N., McDaniel, M., Padilla, A. M., & Parra-Medina, D. (2021).

Impact of COVID-19 on Latinos: A Social Determinants of Health Model and Scoping Review of the Literature. *Hispanic Journal of Behavioral Sciences*, *43*(3), 174–203. https://doi.org/10.1177/07399863211041214

Samuels-Kalow, M. E., Dorner, S., Cash, R. E., Dutta, S., White, B., Ciccolo, G. E., Brown, D. F. M., & Camargo Jr, C. A. (2021). Neighborhood Disadvantage Measures and COVID-19 Cases in Boston, 2020. *Public Health Reports*, *136*(3), 368–374. https://doi.org/10.1177/00333549211002837

Samuel, L. J., Gaskin, D. J., Trujillo, A. J., Szanton, S. L., Samuel, A., & Slade, E. (2021). Race, ethnicity, poverty and the social determinants of the coronavirus divide: U.S. county-level disparities and risk factors. *BMC Public Health*, *21*(1), 1–11. https://doi.org/10.1186/s12889-021-11205-w

Semati, A., Zare, M., Mirahmadizadeh, A., Hemmati, A., & Ebrahimi, M. (2022). Epidemiological Study of Infection and Death Due to COVID-19 in Fars Province, Iran, From February to September 2020. *Iranian Journal of Medical Sciences*, *47*(3), 219–226.

Shah, H., Khan, M. S. H., Dhurandhar, N. V., & Hegde, V. (2021). The triumvirate: why hypertension, obesity, and diabetes are risk factors for adverse effects in patients with COVID-19. *Acta Diabetologica*, *58*(7), 831–843. https://doi.org/10.1007/s00592-020-01636-z

Shakor, J. K., Isa, R. A., Babakir-Mina, M., Ali, S. I., Hama-Soor, T. A., & Abdulla, J. E.

(2021). Health related factors contributing to COVID-19 fatality rates in various communities across the world. *Journal of Infection in Developing Countries*, *15*(9), 1263–1272. https://doi.org/10.3855/jidc.13876

Shih, S.-F., Wagner, A. L., Masters, N. B., Prosser, L. A., Lu, Y., & Zikmund-Fisher, B. J. (2021). Vaccine hesitancy and rejection of a vaccine for the novel coronavirus in the United States. *Frontiers in Immunology*, *12*, Article 558270. https://doi.org/10.3389/fimmu.2021.558270

Sieber, P., Flury, D., Güsewell, S., Albrich, W. C., Boggian, K., Gardiol, C., Schlegel, M., Sieber, R., Vernazza, P., & Kohler, P. (2021). Characteristics of patients with Coronavirus Disease 2019 (COVID-19) and seasonal influenza at time of hospital admission: a single center comparative study. *BMC Infectious Diseases*, *21*(1), 1–9. https://doi.org/10.1186/s12879-021-05957-4

Singu, S., Acharya, A., Challagundla, K.,, Byrareddy, S.N. (2020). Impact of Social Determinants of Health on the Emerging COVID-19 Pandemic in the United States. *Frontiers in Public Health*. 8,406. https://doi.org/10.3389/fpubh.2020.00406

Sun, J., He, W.-T., Wang, L., Lai, A., Ji, X., Zhai, X., Li, G., Suchard, M. A., Tian, J., Ssentongo, P., Ssentongo, A. E., Heilbrunn, E. S., Ba, D. M., & Chinchilli, V. M. (2020). Association of cardiovascular disease and 10 other pre-existing comorbidities with COVID-19 mortality: A systematic review and meta-analysis. *PLoS ONE*, *15*(8), 1–16. https://doi.org/10.1371/journal.pone.0238215

Sun Y, Guan X, Jia L, Xing N, Cheng L, Liu B, et al. (2020). Independent and combined effects of hypertension and diabetes on clinical outcomes in patients with COVID-19: a retrospective cohort study of Huoshen mountain hospital and Guanggu Fangcang Shelter Hospital. *J Clin Hypertens*. doi: 10.1111/jch.14146. [Epub ahead of print]

Sundaram, Suneha S.; Melquist, Stephanie; Kalgotra, Pankush; Srinivasan, Sachin; Parasa, Sravanthi; Desai, Madhav; Sharma, Prateek (2022). Impact of age, sex, race, and regionality on major clinical outcomes of COVID-19 in hospitalized patients in the United States. *BMC Infect Dis* **22**, 659 https://doi.org/10.1186/s12879-022-07611-z

Tian, J., Yuan, X., Xiao, J., Zhong, Q., Yang, C., Liu, B., Cai, Y., Lu, Z., Wang, J., Wang, Y., Liu, S., Cheng, B., Wang, J., Zhang, M., Wang, L., Niu, S., Yao, Z., Deng, X., Zhou, F., & Wei, W. (2020). Clinical characteristics and risk factors associated with COVID-19 disease severity in patients with cancer in Wuhan, China: a multicentre, retrospective, cohort study. *Lancet Oncology*, *21*(7), 893–903. https://doi.org/10.1016/S1470-2045(20)30309-0

Unruh, L. H., Dharmapuri, S., Xia, Y., & Soyemi, K. (2022). Health disparities and COVID-19: A retrospective study examining individual and community factors causing disproportionate COVID-19 outcomes in Cook County, Illinois. *PLoS ONE*, *17*(5), 1–14. https://doi.org/10.1371/journal.pone.0268317

Upshaw TL, Brown C, Smith R, Perri M, Ziegler C, Pinto AD (2021) Social determinants of COVID-19 incidence and outcomes: A rapid review. *PLoS ONE* 16(3): e0248336. https://doi.org/ 10.1371/journal.pone.0248336

U.S. Bureau of Labor Statistics. (2023, April 14). U.S. Bureau of Labor Statistics. Retrieved April 14, 2023, from https://www.bls.gov/

Viswanath, K., Bekalu, M., Dhawan, D., Pinnamaneni, R., Lang, J., & McLoud, R. (2021). Individual and social determinants of COVID-19 vaccine uptake. *BMC Public Health*, *21*(1), 1–10. https://doi.org/10.1186/s12889-021-10862-1

Wang L, Duan Y, Zhang W, Liang J, Xu J, Zhang Y. (2020). Epidemiologic and Clinical Characteristics of 26 Cases of COVID-19 Arising from Patient-to-Patient Transmission in Liaocheng, China. *Clinical Epidemiol*, 12: 387–391.

Willis, D. E., Andersen, J. A., Bryant-Moore, K., Selig, J.P., Long, C. R., Felix, H. C.,. Curran. G. M., ,McElfish, P. A. (2021). COVID-19 vaccine hesitancy: Race/ethnicity, trust, and fear. *Clinical and Translational Science*. 2021; 00:1–8. https://doi.org/10.1111/cts.13077

Woolley, A. E., Dryden-Peterson, S., Kim, A., Naz-McLean, S., Kelly, C., Laibinis, H. H., Bagnall, J., Livny, J., Ma, P., Orzechowski, M., Gomez, J., Shoresh, N., Gabriel, S., Hung, D. T., & Cosimi, L. A. (2022). At-home Testing and Risk Factors for Acquisition of SARS-CoV-2 Infection in a Major US Metropolitan Area. *Open Forum Infectious Diseases*, *9*(11), 1–8. https://doi.org/10.1093/ofid/ofac505

World Health Organization (WHO) (2021). Coronavirus (COVID-19) Dashboard. https://covid19.who.int/

World Health Organization (WHO) (2022). Coronavirus (COVID-19) Dashboard. https://covid19.who.int/

World Health Organization (WHO) (2023). WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/

Zhang, X., Lewis, A. M., Moley, J. R., & Brestoff, J. R. (2021). A systematic review and meta-analysis of obesity and COVID-19 outcomes. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-86694-1

Zhou, J., Veit, M., & Su, S. (2020). COVID-19: Epidemiology, Evolution, and Cross-Disciplinary Perspectives. *Trends in Molecular Medicine, 26*(5), 483–495. https://doi.org/10.1016/j.molmed.2020.02.008