The Classification of COVID-19 Mortality and Evaluation of Deceased Characteristics by Identification Method in Allegheny County, Pennsylvania

by

Rheana Dee Lipscomb

BS, Lycoming College, 2018

Submitted to the Graduate Faculty of the

Department of Epidemiology

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This essay is submitted

by

Rheana Dee Lipscomb

on

November 19, 2021

and approved by

Essay Advisor: Wendy C King, PhD, Associate Professor, Department of Epidemiology,
Graduate School of Public Health, University of Pittsburgh

Marnie Bertolet, PhD, Assistant Professor, Department of Epidemiology,
Graduate School of Public Health, University of Pittsburgh

Jennifer Fiddner, MPH, CIC, Epidemiology Research Associate Supervisor,
Allegheny County Health Department

Ada O Youk, PhD, Associate Professor, Department of Biostatistics,
Graduate School of Public Health, University of Pittsburgh

The Classification of COVID-19 Mortality and Evaluation of Deceased Characteristics by Identification Method in Allegheny County, Pennsylvania

Rheana Dee Lipscomb, MPH

University of Pittsburgh, 2021

Abstract

COVID-19 mortality data are an important component of measuring the impact of the pandemic and can be used to identify disproportionately affected populations. Insights from COVID-19 case and mortality data are used to target interventions and mitigation strategies to prevent further COVID-19 related mortality. Passive surveillance systems used to identify COVID-19 mortality are limited by underreporting and timeliness (i.e., lag in reporting), which can result in an underestimation of COVID-19 mortality. The public health workforce has been overwhelmed by effects of the COVID-19 pandemic, and most local health departments have not yet had the opportunity to identify COVID-19 deaths missed by their surveillance systems. This descriptive analysis identified COVID-19 related deaths in Allegheny County between March 14, 2020 and February 19, 2021 in Allegheny County, Pennsylvania. Deceased characteristics were compared between deaths identified by the Allegheny County Health Department (ACHD) COVID-19 mortality surveillance and deaths not identified by ACHD surveillance, identified as part of the 2021 ACHD Pittsburgh Summer Institute. A multivariate logistic regression model was used to assess if age, sex, race/ethnicity, and vulnerability index independently predicted if COVID-19 deaths were identified by ACHD surveillance. Most COVID-19 deaths were identified by ACHD COVID-19 surveillance: of 1,462 total COVID-19 deaths, 1,325 (91%) deaths were identified through ACHD COVID-19 mortality surveillance, and 137 (9%) deaths were identified outside of ACHD COVID-19 mortality surveillance. The odds of being female sex (OR 1.52) and

race other than white or black (OR 3.49) were significantly higher among deaths not identified by ACHD surveillance, when controlling for age and vulnerability index, which were not related to identification method. ACHD COVID-19 mortality surveillance from March 14, 2020 to February 19, 2021 was largely accurate and indiscriminate and is likely subject to the general limitations of passive surveillance systems. The public health importance of these findings is that COVID-19 deaths have been identified outside of traditional surveillance methods, and these deaths differ from the deaths identified by passive surveillance regarding sex and race/ethnicity, although these differences are small.

Table of Contents

List of Tables

List of Figures

Disclaimer

These data were provided by the Pennsylvania Department of Health. The Department specifically disclaims responsibility for any analyses, interpretations, or conclusions.

1.0 Introduction

1.1 Overview of the COVID-19 Pandemic

Coronavirus disease 2019 (COVID-19) is an infectious respiratory illness caused by the SARS-CoV-2 virus. SARS-CoV-2, a single-stranded RNA virus in the family *Coronaviridae*, and has been responsible for three significant outbreaks of disease: SARS-CoV, MERS-CoV, and SARS-CoV-2. SARS-CoV-2 is transmitted person-to-person via droplets that contain the virus or contact of mucus membranes with a surface contaminated with the virus. The viral load in oropharyngeal secretions is highest during the early symptomatic period of infection. Infected individuals may continue to shed the virus after symptoms have ceased.[1] COVID-19 has an incubation period between 2-14 days, with a mean incubation period of 5.2 days.[2] R_0, a mathematical term that represents the reproductive number for a pathogen, of the original COVID-19 strain was estimated to be 2.79, while the R_0 of the now predominant delta strain is 5.08.[3] People infected with SARS-CoV-2 often exhibit mild symptoms, but the disease can progress to a severe form which can result in death. Older adults and people with certain comorbidities are at a higher risk for severe illness from COVID-19.[4] Mitigation efforts to prevent the spread of COVID-19 have been an evolving public health practice since the emergence of the SARS-CoV-2 virus and have been essential for managing the COVID-19 pandemic due to its high transmissibility and associated morbidity and mortality. Hand hygiene, masking, and physical distancing practices have been recommended throughout the pandemic to reduce the transmission of COVID-19. Isolation of COVID-19 cases and quarantine of close-contacts of confirmed COVID-19 cases have been utilized by public health agencies to reduce the spread of COVID-19.[5]

The availability of COVID-19 testing has been essential for mitigation strategies related to the pandemic. As of February 7, 2021, over 305 million COVID-19 tests were conducted in the United States (US).[6] The current gold standard COVID-19 tests for SARS-CoV-2 utilize real-time reverse transcriptase polymerase chain reaction (RT-PCR) from nasopharyngeal and oropharyngeal swab samples. The US Food and Drug Administration has issued Emergency Use Authorizations starting on February 4, 2020 for commercial and laboratory-developed tests, including molecular, antibody, and antigen tests.[7,8] The different forms of COVID-19 tests have varying utility based on their sample collection methods, time from sample collection to results, sensitivity and specificity, and cost.

The lack of a reliable gold standard COVID-19 tests at the beginning of the pandemic in the early months of 2020 made evaluating newly developed tests challenging. False negative COVID-19 cases are of particular concern for COVID-19 mitigation strategies. Incorrect test results can be influenced by the timing of the test in relation to disease progression; some studies have suggested that testing one to three days after symptom onset will decrease the chance of a false-negative test.[7] Testing for COVID-19 is currently recommended by the Centers for Disease Control and Prevention (CDC) for individuals who have symptoms of COVID-19, had close contact with someone with confirmed COVID-19, are unvaccinated and participated in activities that put them at high risk for contracting COVID-19, and people who have been referred to testing by a health care provider or healthcare agency.[9] The CDC also recommends that people who are not fully vaccinated with a negative COVID-19 test following an exposure should be tested again 5-7 days after the exposure or when symptoms develop. Fully vaccinated people are recommended to be tested 5-7 days after COVID-19 exposure.[9] The quality of COVID-19 tests and the access to

COVID-19 testing has improved since the start of the pandemic, although there is still significant room for improvement.

There has been substantial concern about the lack of equitable access to COVID-19 testing in the US, particularly following a shortage of COVID-19 tests at the beginning of the pandemic.[6] Federal, state, and local governments, along with health insurance companies, have worked to expand COVID-19 testing to more of the population through free-of-charge testing. Actions have also been taken to increase testing capacity and to improve access to testing for at-risk and low-resource populations. Despite these efforts, there remains inequities regarding access to testing among uninsured, racial and ethnic minorities, rural, and low-income populations.[6] Individual perceptions regarding COVID-19, COVID-19 testing, and mitigation strategies, such as isolation and quarantine, may influence someone's willingness to seek out COVID-19 testing. A study conducted through social media platforms for adults residing in the US sought to examine perceived barriers to COVID-19 testing. Approximately half (51.7%) of participants responded that they could access a COVID-19 test if needed. However, perceived access was notably lower among participants aged 18-39 years, participants with an annual household income of less than $30,000, and participants without health insurance.[6] This study quantifies perceived access and barriers to COVID-19 testing among young adults and low-income individuals. Current public health practice relies on testing to inform the public of guidance to reduce the transmission of COVID-19, as well as for surveillance efforts to identify disproportionally impacted areas and communities. Inequitable and limited access to testing remains a significant public health issue.

The burden of the COVID-19 pandemic on healthcare systems, coupled with limited resources for individuals seeking healthcare services, has had a devastating effect on health outcomes related to the pandemic. Access to care has been an ongoing issue since the start of the pandemic, both for individuals with COVID-19 and with other health conditions who cannot access healthcare services because of an over-burdened healthcare system. Studies on excess mortality during the pandemic indicate that some portion of excess deaths not attributed to COVID-19 may have resulted indirectly from COVID-19, as people have had difficulties accessing care or have delayed health care.[10,11,12]

Access to and quality of care is not equitable among minority groups in the US, and inequitable access impacts outcomes of patients with COVID-19. A study of adult Medicare beneficiaries admitted with COVID-19 at 1,188 US hospitals aimed to assess if race (Black versus White) and the hospital patients with COVID-19 were admitted to predicted worse outcomes.[13] The study found that, after adjusting for sociodemographic factors, including age and clinical characteristics, the odds of 30-day inpatient mortality or discharge to hospice were 11% higher for Black patients than for White patients. However, the difference in odds of mortality or hospice were almost eliminated when the authors adjusted for the hospital in which care was received.[13] These findings indicate that the difference in mortality rates among Black and White patients are at least in part attributable to the different hospitals to which Black and White patients were admitted, which had differences in the racial make-up of the patient populations. This study illustrates that COVID-19 associated mortality is complex, and further research is needed to fully assess direct and indirect mortality from COVID-19 related to access and quality of healthcare, particularly among minoritized populations.

1.2 COVID-19 Mortality Data and Reporting

As of October 1, 2021, 4.7 million deaths globally and approximately 700,000 deaths in the US had been directly attributed to COVID-19.[14,15] The case fatality rate of COVID-19 as of October 1, 20201, in US is 1.6%, with 213.35 deaths per 100,000 people.[16] Of the documented COVID-19 associated mortality in the US, at least 90% of deaths attributed to COVID-19 listed COVID-19 as the underlying cause of death, and the remaining deaths list COVID-19 as a contributing cause of death.[17]

At the national level, the CDC compiles provisional death data from the National Vital Statistics System on a daily and weekly basis. Provisional data are derived from death certificates submitted at the state and local level.[18] From April 1, 2020 to October 3, 2021, the highest number of deaths in a week occurred over the week ending January 9, 2021, with almost 26,000 COVID-19 attributed deaths occurring nationally. The lowest number of deaths in a week occurred over the week ending July 3, 2021, with 1,500 COVID-19 attributed deaths. Two-thirds (67%) of COVID-19 related deaths have occurred in a hospital or other inpatient healthcare setting, with nursing homes or long-term care facilities the site of approximately 17% of COVID-19 related mortality.[18]

In the US, those aged 65 and older have experienced the highest COVID-19 related mortality, accounting for 77% of COVID-19 related deaths versus only 3.5% in those 45 and younger.[18] The COVID-19 mortality rates per 100,000 by age group through April 9, 2021 are: 43 among ages less than 1, 24 among ages 1-4, 67 among ages 5-14, 587 among ages 15-24, 2,527 among ages 25-34, 6,617 among ages 35-44, 17,905 among ages 45-54, 44,631 among ages 55-64, 80,617 among ages 65-74, 104,212 among ages 75-84, and 120,648 among ages 85 and older.[19] The non-Hispanic White population accounts for 62% of COVID-19 related deaths, followed by the Hispanic population at 18%, and non-Hispanic Black population at 15%, while these populations account for 60%, 19%, and 13% of the population as of 2018 respectively.[18,20] Despite the percentage of COVID-19 related deaths being highest among the non-Hispanic White population, mortality rates are higher in minority groups compared to the White population: 178 among Black or African Americans, 172 among the Indian or Alaskan Native people and 154 among Hispanic or Latinos versus 124 among Whites per 100,000 through March, 7, 2021.[21]

Studies on excess mortality indicate there may be an underestimation of COVID-19 deaths nationally, which may in part be due to limited testing at the beginning of the pandemic and delays in COVID-19 death reporting.[22] Estimating excess death during the pandemic is an important indicator of the mortality caused directly or indirectly from COVID-19. Excess mortality can be attributed directly to COVID-19 as a significant contributing cause of death, or excess mortality from other causes can be attributed to the increased burden on the healthcare industry, limited access to care, or undocumented COVID-19 infection.[11, 23] A study that assessed excess all-cause mortality incidence rates from December 29, 2019, to January 2, 2021, using data from the National Vital Statistics System, identified that among adults aged 65 and older, Black and Hispanic persons experienced the highest excess mortality incidence rates. Among the population less than 65 years of age, Black and American Indian/Alaskan Native populations experienced the highest excess mortality incidence rates. The highest age-adjusted death rates were also observed among Black and American Indian/ Alaskan Native populations in 2020.[10] A study of mortality in the US between March 1, 2020, and January 2, 2021, compared mortality data from 2014-2019 to estimate excess death during the pandemic. Overall, the US experienced 22.9% more deaths than expected from previous years, which surpasses the average annual increases of mortality of greater than or equal to 2.5% observed in recent years. Like the previous study, the excess death rate was higher among non-Hispanic Black populations than among non-Hispanic White or Hispanic populations. The study did not adjust for population aging in the model and relied on provisional data for the analysis.[11] Despite these limitations, this study highlights the burden of COVID-19 and inequities in COVID-19 associated mortality.

Pennsylvania. By October 3, 2021, there were over 1.1 million confirmed COVID-19 cases and almost 30,000 COVID-19 related deaths reported in the state of Pennsylvania. Information from the State Department of Health's Vital Records Program is provisional and includes laboratory confirmed or clinically confirmed COVID-19 deaths.[24] Both White and Black Pennsylvanians are over-represented among COVID-19 deaths, with the White population accounting for 81% and Black population accounting for 13% of COVID-19 related deaths even though they represent only 76% and 10%, respectively, of the total Pennsylvania population. In contrast, Asian and Other races are under-represented.[25] In Pennsylvania, deaths in long term care facilities account for over 50% of deaths due to COVID-19.[26]

Allegheny County. By October 3, 2021, there were over 121,000 cases and over 2,100 deaths in Allegheny County, Pennsylvania, yielding a death rate of 177 per 100,000 individuals. The average COVID-19 test positivity rate is 6% in Allegheny County from the start of the pandemic through May 24, 2021, and ranges from 1% to 15% by month.[27] The Allegheny County Health Department (ACHD) considers deaths to be COVID-19 related if they were among confirmed or probable COVID-19 cases who died because of complications related to COVID-19. Sixty-two percent of individuals identified to have COVID-19 related deaths in Allegheny County were hospitalized at the time of their death or were discharged from a COVID-19 related hospital admission prior to death, and 52% were residents of long-term care facilities. As with the state, White and Black individuals were slightly over-represented in COVID-19 related deaths in

Allegheny County (White: 82% of deaths, 80% of population; Black: 13.5% of deaths, 13% of the population). Despite only accounting for 19% of the Allegheny County population, 88% of deaths in Allegheny County were among those aged 65 and older.[28]

1.3 COVID-19 Data Handling Overview

The National Center for Health Statistics (NCHS) is responsible for producing national health statistics based on the records compiled in the National Vital Statistics System (NVSS), which collects death certificates and other vital records from all States in the US and US territories. Death certificate data in the NVSS are coded and classified with the current revision of the International Classification of Diseases (ICD-10), as required of member nations of the World Health Organization.[29] Death certificate coding conducted by the NCHS is predominantly automated, although some causes of death such as COVID-19, are more likely to require manual coding due to the need for review, such as for novel diseases or varying terminology on the death certificates. This can cause delays in cause-specific death statistics reporting, which may result in a lag in data reporting of cause-specific mortality. Provisional death counts are based on the flow of mortality data in NVSS, and weeks can pass between when a death occurs and when it is submitted to NCHS (after a death has been processed, coded, and tabulated). Death counts are continuously revised and are subject to change with updated death certificate data sent by states. Provisional death counts from the NCHS may differ from other sources, such as state and local COVID-19 data dashboards that predominately utilize data from local health departments and state departments of health.[30] Discrepancies between COVID-19 mortality reports nationally from those at the state or local level are likely due to reporting times and to the classification of COVID-19 case status for COVID-19 related deaths.

Pennsylvania. In Pennsylvania, death information is provided to local and state health agencies by the medical certifier. Reporting deaths through the Electronic Death Registration System (EDRS) is mandated by law. Cause of death information that is certified on a report of death is registered by the state vital records office, the Bureau of Health Statistics and Registries (BHSR). A death is registered once the medical certifier and funeral home director have signed off on the death and it has been accepted by the BHSR. The BHSR then submits deaths to the CDC, which codes the cause of death information with ICD-10 codes. In Pennsylvania, all deaths from natural causes are required to be reported by the medical certifier in EDRS, including deaths with probable or pending laboratory results. The Pennsylvania Department of Health (PADOH) has altered medical certifiers such that COVID-19 should be reported for all deaths in which the disease caused or is assumed to have caused or contributed to the death, and COVID-19 should not be included if it did not cause or contribute to the death.[17] The CDC has noted that Pennsylvania has experienced reporting lags during the pandemic that are much longer in duration than in recent years.[23] Reporting delays may have resulted in a short-term underestimation of provisional total mortality and COVID-19 associated mortality in Pennsylvania for 2020.

Allegheny County. In Pennsylvania, when a death certificate is signed by a funeral director or coroner, it is received by the local registrar. These data are sent to Pennsylvania Department of Health local offices, and individual-level Allegheny County death certificate information is then sent to ACHD. The death certificate information is also received by the PADOH, which compiles the data for each county. Datasets of vital statistics data, including death certificate information, are then sent from the PADOH to the ACHD. After being coded by the NCHS, ICD-10 coded death certificates are matched to COVID-19 case data by Pennsylvania Electronic Death Registration System (PA-EDRS) though PADOH. Moving through this process is time consuming; this data lag is one reason why the ACHD and the PADOH can have different numbers for COVID-19 cases and deaths for the same time period.

1.4 Social Determinants of COVID-19 in Allegheny County

The RAND Cooperation identified that in the early months of COVID-19 pandemic Black residents of Allegheny County were more likely to die because of COVID-19 when compared to White residents; however, over time this gap has narrowed. The cumulative death rates among Black and White residents of Allegheny County were 35.4 and 26.4 deaths per 100,000 respectively as of August 30, 2020 and the difference narrowed starting in November 2020. As of October 18, 2021, the cumulative death rate of Black versus White residents of Allegheny County was 191 versus 189 deaths per 100,000 people.[27] Age adjusted mortality rates are not yet available for the same period, but indicate that through January 30, 2021 Black residents had a higher mortality rate than White residents when adjusted for age (157.3 versus 123.2 deaths per 100,000).[31] Stakeholders such as the Black Equity Coalition in Allegheny County have garnered substantial momentum in identifying and improving health outcomes of minority populations related to COVID-19; more COVID-19 tests per capita have been conducted for Black residents of Allegheny County than White residents. This rate of testing for Black residents in Allegheny County differs from the rate of testing for Black versus White populations nationally. A report conducted by stakeholders in Allegheny County identified that testing sites in Allegheny County were generally near or available to communities at increased vulnerability for COVID-19.[32]

Despite the increased and effective activism, Black residents of Allegheny County test positive and are hospitalized at a higher rate compared to White residents of Allegheny County.[32] Stakeholders in Allegheny County have contributed greatly to improving the outcomes of COVID-19 for minority populations, but further assessment and work is needed to continue to improve these outcomes.

Within the context of the COVID-19 pandemic, it is important to evaluate factors that may be related to susceptibility for increased rates of infection, morbidity, and mortality. A vulnerability index, created by SURGO Ventures, has been used to assess which communities may be more susceptible to the impacts of the COVID-19 pandemic. This vulnerability index metric, scored on a scale of 0 (low vulnerability) to 1 (high vulnerability), measures health, economic, and social factors at the census tract level to assess community susceptibility within the context of the COVID-19 pandemic.[33] Vulnerability index scores within Allegheny County are predominately driven by socioeconomic status, disability, and household composition.[27] While the average vulnerability index of Allegheny County is 0.31, 9.4% of the county population has a score of 0.6 or above, classified as having high or very high vulnerability.[34]

1.5 Gaps in Knowledge and Public Health Significance

Accurate reporting of COVID-19 mortality is essential for public trust at a time where public health statistics and messaging are met with significant scrutiny. Additionally, accurately capturing and assessing COVID-19 related mortality is essential for public health practice to identify the impact and severity of the COVID-19 pandemic. Studies assessing ICD-10 codes from death certificates that included the COVID-19 ICD-10 code indicate that 92% of death certificates included a plausible chain-of-event condition, a significant contributing condition, or both for COVID-19.[35] While there is room for improvement in thoroughly documenting COVID-19-attributeable deaths, this analysis supported the accuracy of documenting COVID-19 mortality nationally with the inclusion of factors contributing to death that are related to COVID-19.[35] However, deaths may be missed by surveillance because of underreporting and timeliness of reporting.[36] These issues are likely only exacerbated further by the strain the COVID-19 pandemic has put on the healthcare and public health workforces.

Surveillance systems aiming to capture COVID-19 mortality in the US vary in methodology at the state and county levels and are used to inform prevention efforts and assess the effectiveness of pandemic-related policies. Public health surveillance is limited in its capacity to fully identify cases of disease because the majority of those tested are symptomatic, biasing surveillance systems towards symptomatic infections and not fully capturing mild and sub-clinical cases. Current surveillance systems for infections and deaths of COVID-19 are estimated to have an incomplete case ascertainment.[37,38] An analysis on COVID-19 surveillance systems in the US found that these systems are limited by under-reporting, completeness of data, and the timeliness of data entry.[38] Surveillance systems are also limited if there are barriers to testing or access to care, preventing populations from being identified as cases.

Mortality data are used to identify populations disproportionately affected by COVID-19, and these insights can be used to target interventions for these populations. Underestimating COVID-19 associated mortality underestimates the impact and severity of the pandemic. Surveillance methods to capture COVID-19 mortality may not identify all deaths, and it is important to retroactively identify these missed deaths to accurately report COVID-19 mortality. Identifying if populations are disproportionately missed in the assessment of COVID-19 mortality can help to address and include those communities in future interventions and assessment. If COVID-19 deaths are missed in surveillance due to a lack of testing among certain populations, the assessment of missed COVID-19 deaths can be used to advocate for directed public health messaging and resources. Deaths not captured by COVID-19 surveillance may also be indiscriminate and could result from the limitations of passive surveillance systems. There is currently limited work assessing COVID-19 mortality not captured by surveillance systems. Due to the burden of the pandemic on public health agencies, health departments have not necessarily had the time or resources to retrospectively analyze the accuracy of these surveillance systems. It is of public health interest to identify deaths missed by COVID-19 mortality surveillance and assess if these deaths differ from deaths identified by traditional surveillance methods.

2.0 Objectives

The first objective of this research is to identify all COVID-19 related deaths in Allegheny County. All COVID-19 related deaths include those captured by ACHD's surveillance systems, as well as additional COVID-19 related deaths identified via death certificates. These additional COVID-19 related deaths include deaths among known cases with the death variable incorrectly documented as "no" and "unknown", and deaths identified from vital statistics data that were not in COVID-19 surveillance data. The second objective of this research is to determine if there is differential reporting by demographic characteristics of deaths that were missed versus deaths that were captured by ACHD's surveillance system. My hypothesis was that Black race and high vulnerability would be associated with COVID-19 deaths not being identified by Allegheny County COVID-19 surveillance, while sex and age group would not be related to identification method.

3.0 Methods

3.1 ACHD COVID-19 Data and COVID-19 Mortality Surveillance

The surveillance system used by the ACHD to identify COVID-19 related deaths is based on Pennsylvania's version of the National Electronic Disease Surveillance System (PA-NEDSS). The COVID-19 case definition from the Council of State and Territorial Epidemiologists is used.[39] Probable COVID-19 cases are defined as having met the clinical criteria for COVID-19 and either having an epidemiologic linkage with no confirmatory or presumptive laboratory evidence, having presumptive laboratory evidence, or meeting vital records criteria with no confirmatory testing. Confirmed COVID-19 cases are defined as meeting confirmatory laboratory evidence.[39] ACHD considers all probable or confirmed COVID-19 cases documented in PA-NEDSS that were identified from COVID-19 testing or were epidemiologically linked to a COVID-19 case with the death documented in PA-NEDSS to be a COVID-19 related death.

For the Pittsburgh Summer Institute COVID-19 Vital Statistics Project, the ACHD provided COVID-19 case and death data in a single dataset (referred to as ACHD COVID-19 surveillance dataset), derived from PA-NEDSS. It included all documented confirmed or probable COVID-19 cases in Allegheny County from March 14, 2020, through May 10, 2021. Deaths were denoted as "yes", "no", "unknown", or as a blank entry. Blank entries were combined with "unknown" deaths for the analysis, as blank entries indicated the data was not captured and therefore unknown. Deaths were reported to ACHD by long-term care facilities, hospital staff, relatives, or identified through COVID-19 case investigation.

3.2 Vital Statistics Data

Preliminary vital statistics data were provided to the ACHD by the PADOH (referred to as PADOH vital statistics dataset). The PADOH vital statistics dataset spanned from January 1, 2020, to February 19, 2021, and included individuals who were believed to be residents of Allegheny County. The vital statistics data consisted of demographics, death certificate information that included free-text information for the underlying cause of death and contributing causes of death, as well as ICD-10 codes that corresponded to each contributing cause of death. Vital statistics data are preliminary for three years following the year in question and includes incomplete death certificates or death certificate data that may be amended in the future. Data are approximately 75% complete within eight weeks of the death.[40] The degree of missingness increases closer in temporal proximity to February 19, 2021, the end data of the vital statistics data provided by PADOH.

3.3 Classification of COVID-19 Deaths in Allegheny County

COVID-19 deaths in the ACHD COVID-19 surveillance dataset were matched with the PADOH vital statistics dataset as part of the Pittsburgh Summer Institute COVID-19 Vital Statistics Project using Excel, Microsoft, Redmond, WA, USA. COVID-19 deaths were matched to assess the underlying causes of death from the PADOH vital statistics data for the COVID-19 deaths in the ACHD COVID-19 surveillance dataset. Records were considered a match and then merged if the death certificate number if available, first name, last name, date of birth, date of death, sex, and social security number if available matched. COVID-19 cases in the ACHD

COVID-19 surveillance dataset noted to have died that did not match with vital statistics records in the PADOH vital statistics dataset were manually searched to identify data entry errors that may have prevented a match. ACHD COVID-19 case data were merged with PADOH vital statistics records in SAS 9.4, SAS Institute, Cary, NC, USA.

Deaths in the PADOH vital statistics dataset found to be related to COVID-19 that were not included in the ACHD COVID-19 surveillance dataset were identified through the Pittsburgh Summer Institute's COVID-19 Vital Statistics Project. These deaths were either among COVID-19 cases with the death variable not documented as "yes," in the ACHD COVID-19 surveillance dataset, or among residents of Allegheny county who were not in the ACHD COVID-19 surveillance dataset but whose deaths were identified from PADOH vital statistics dataset. COVID-19 deaths from the PADOH vital statistics data were identified among COVID-19 cases from the ACHD COVID-19 surveillance dataset with the death variable incorrectly not documented as "yes" by identifying exact matches to the first name, last name, and date of birth to records in vital statistics data, and were verified by sex, social security number if available, and address. COVID-19 deaths in the PADOH vital statistics dataset that were not in the ACHD COVID-19 surveillance dataset were identified by first removing all death certificates that did not include COVID-19. Then, all deaths previously matched to PADOH vital statistics data (COVID-19 deaths identified by the ACHD and COVID-19 deaths identified with the death variable incorrectly documented) were removed from a copy of the PADOH vital statistics dataset. The remaining modified PADOH vital statistics dataset then contained all COVID-19 related deaths not previously identified.

Deaths among the ACHD COVID-19 surveillance dataset were determined to be related to COVID-19 if they included the ICD-10 code for COVID-19 (U07.1) on the death certificate, or if

COVID-19, coronavirus, “Covid”, or other COVID-19 keywords were included in the free-text section of the death certificate data. Additional COVID-19 deaths identified among cases with the death variable incorrectly documented as determined by a match with a vital statistics record, and among COVID-19 deaths in the PADOH vital statistics data that were not documented as COVID-19 cases by ACHD, were only included as COVID-19 deaths if they included the ICD-10 code for COVID-19 on the death certificate. The free text of the death certificate for COVID-19 deaths identified outside of ACHD surveillance was not searched due to the number of deaths among these subsets and time constraints.

Deaths were designated as having been identified through ACHD surveillance or not identified through ACHD surveillance. Deaths not identified through ACHD surveillance included COVID-19 cases with the death variable incorrectly documented and COVID-19 deaths found among the PADOH vital statistics dataset that were not among the ACHD COVID-19 surveillance dataset.

3.4 Additional Analysis by PADOH

Based on the Pittsburgh Summer Institute Project findings, epidemiologists with PADOH analyzed the COVID-19 related deaths that were not documented in the ACHD COVID-19 surveillance dataset against their state databases. This analysis was conducted to identify why the additional deaths were not accounted for within the Allegheny County COVID-19 datasets. PADOH looked at data outside of Allegheny County’s jurisdiction to determine if deaths were incorrectly attributed to Allegheny County and looked at suspect COVID-19 cases or records of individuals being deemed to not be a COVID-19 case in PA-NEDSS, which would not be included in ACHD's COVID-19 surveillance dataset from PA-NEDSS.

3.5 Demographic Variables

Age was derived directly from the age variable in the PADOH vital statistics dataset, calculated from date of birth to date of death, and was categorized in accordance with the CDC's reporting of COVID-19 deaths.[41] Sex, race and ethnicity and home address were reported on the death certificate. Race and ethnicity were categorized as non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, non-Hispanic Biracial, and Hispanic. The vulnerability index was obtained from Sugro ventures.[34] Using Excel, the mean COVID-19 Community Vulnerability Index by census tract was assigned to each death based on the census tract of the decedent's address.[34] Vulnerability index is reported according to Surgo ventures 5-level 0.2-incriment categorization, from very low vulnerability (0.00-0.19) to very high vulnerability (0.80-1.00).[34] The vulnerability index was also evaluated as a two-level variable of not high vulnerability (very low, low, and moderate vulnerability: 0.00-00.59) and high vulnerability (high and very high vulnerability: 0.60-1.00).

To test associations, high (0.60-0.79) and very high vulnerability (0.80-1.00) categories were collapsed because only 43 deaths were categorized as very high vulnerability. Age categories <65 years were collapsed in one category because only 3 deaths occurred among those ages 18-39 and 14 deaths occurred among those ages 40-49; race/ethnicity groups were collapsed to include all races and ethnicities aside from non-Hispanic White and non-Hispanic Black as "Other" because only 21 deaths occurred among the Non-Hispanic Asian, Non-Hispanic Native American, Non-Hispanic Biracial, and Hispanic groups.

3.6 Statistical Analyses

Frequencies and percentages were reported for all variables overall and by whether deaths were identified by ACHD COVID-19 surveillance or not. The Cochran–Armitage tests for ordinal variables (month in which the death occurred, age category and 5-level vulnerability index), and the Chi-squared test or Fisher's Exact test, as appropriate, for categorical variables (sex, race/ethnicity and high vulnerability [yes/no]) were used to evaluate associations with inclusion in the ACHD COVID-19 surveillance dataset.

A multivariable logistic regression model was used to determine the association of sociodemographic characteristics with the death identification method, controlling for all other covariates. The model included race/ethnicity, age, sex, and the 4-level vulnerability index. An interaction term of sex by age was tested but eliminated if not significant. All analyses were performed using SAS 9.4.

4.0 Results

4.1 ACHD Surveillance Dataset COVID-19 Deaths

Of the 1,690 deaths attributed to COVID-19 in the ACHD surveillance dataset through February 9, 2021, there were 197 deaths that did not match with vital statistics records (i.e., no death certificate available) and 89 deaths in which the death certificate was not ICD-10 coded with the underlying cause of death, such that COVID-19 related could not be verified. Among the remaining 1,404 deaths, 1,325 (94%) included the ICD-10 code for COVID-19 or included COVID-19 in the free text of the death certificate (**Figure 1**).

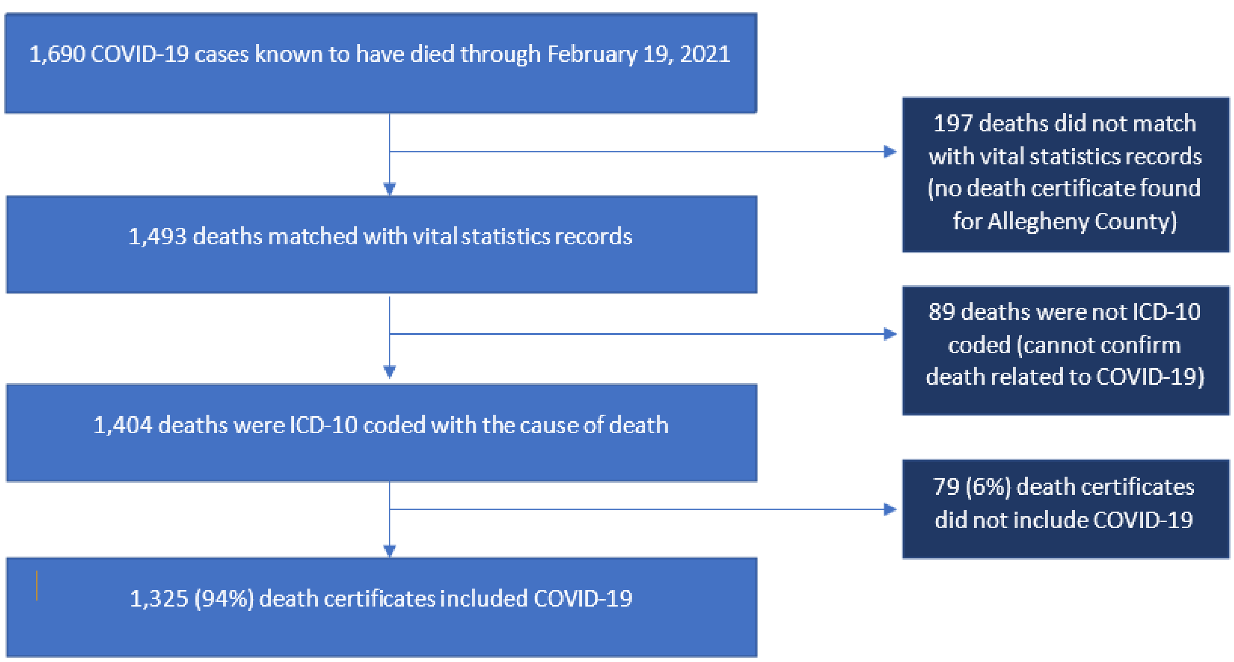

Figure 1 Determining if deaths identified by ACHD COVID-19 surveillance included COVID-19 on the death certificate

4.2 Additional ACHD COVID-19 Surveillance Deaths

There were 46,858 (47%) of COVID-19 cases with the death variable documented as "no" among the ACHD COVID-19 surveillance dataset through May 10, 2021. The exact number of cases through February 19, 2021 could not be determined from the dataset because there was not a specified diagnosis date for each case but could be determined through May 10, 2021 as that was the end data of the ACHD COVID-19 surveillance dataset. A total of 71 (1.5%) of COVID-19 cases in the ACHD COVID-19 surveillance dataset with the death variable documented as "no" matched with vital statistics records from the PADOH vital statistics dataset, of which 9 included the COVID-19 ICD-10 code U07.1 on the death certificate. There were 50,618 (51%) of COVID-19 cases with the death variable documented as "unknown" or blank. A total of 127 (0.3%) of COVID-19 cases in the ACHD COVID-19 surveillance dataset with the death variable documented as "unknown" matched with vital statistics records from the PADOH vital statistics dataset, of which 7 of were related to COVID-19 per ICD-10 codes. Thus, 16 COVID-19 related deaths were identified among COVID-19 cases in the ACHD surveillance dataset with the death variable incorrectly documented (**Figure 2**).

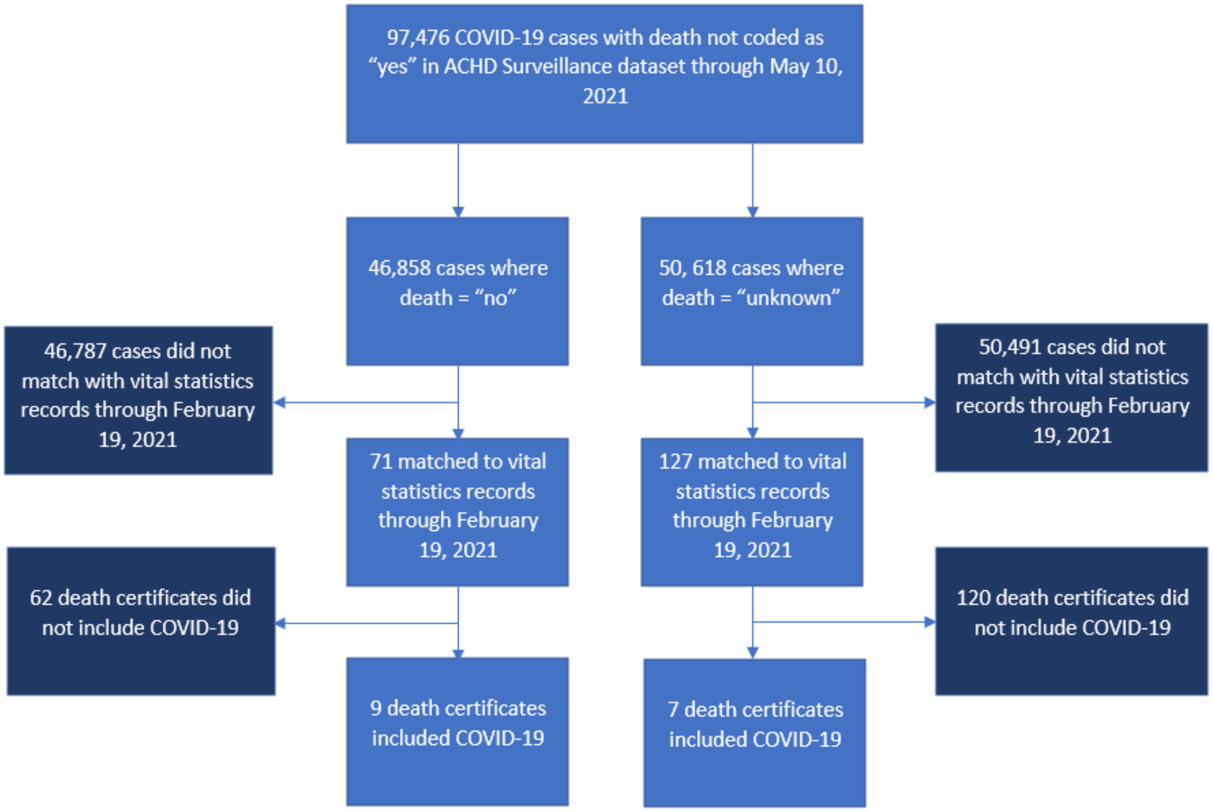

Figure 2 COVID-19 deaths identified with an incorrect value for the death variable in the ACHD COVID-19 surveillance dataset among known cases through May 10, 2021*[1]

4.3 Additional Vital Statistics COVID-19 Deaths

There were 235 total deaths identified in preliminary vital statistics records assigned to Allegheny County that were not in the ACHD surveillance dataset. Almost half (114) of the 235 deaths identified were determined to not be within ACHD's jurisdiction by PADOH (**Figure 3**).

---

*The exact date cases were identified could not be determined from the ACHD COVID-19 surveillance dataset. Because the PADOH vital statistics dataset ended on February 19, 2021, cases could not have matched beyond this date.

Of the 121 additional COVID-19 deaths identified that were within ACHD's jurisdiction, PADOH did not have a record within their state COVID-19 surveillance datasets for 38 of the 121 deaths. An additional 5 deaths were missing a PA-NEDSS record, but PADOH found that these deaths may belong to a jurisdiction outside of Pennsylvania but could not determine with certainty. Records were missing from the ACHD COVID-19 surveillance dataset because they were documented as suspect COVID-19 cases or as not being a COVID-19 case for 73 of the 121 deaths in Allegheny County's jurisdiction.

In summary, a total of 137 deaths were identified to be in Allegheny County's jurisdiction that were not captured by COVID-19 mortality surveillance through February 19, 2021, including the 16 deaths from the ACHD surveillance dataset with the death variable incorrectly documented and 121 deaths identified from preliminary vital statistics records that were not in the COVID-19 surveillance dataset.

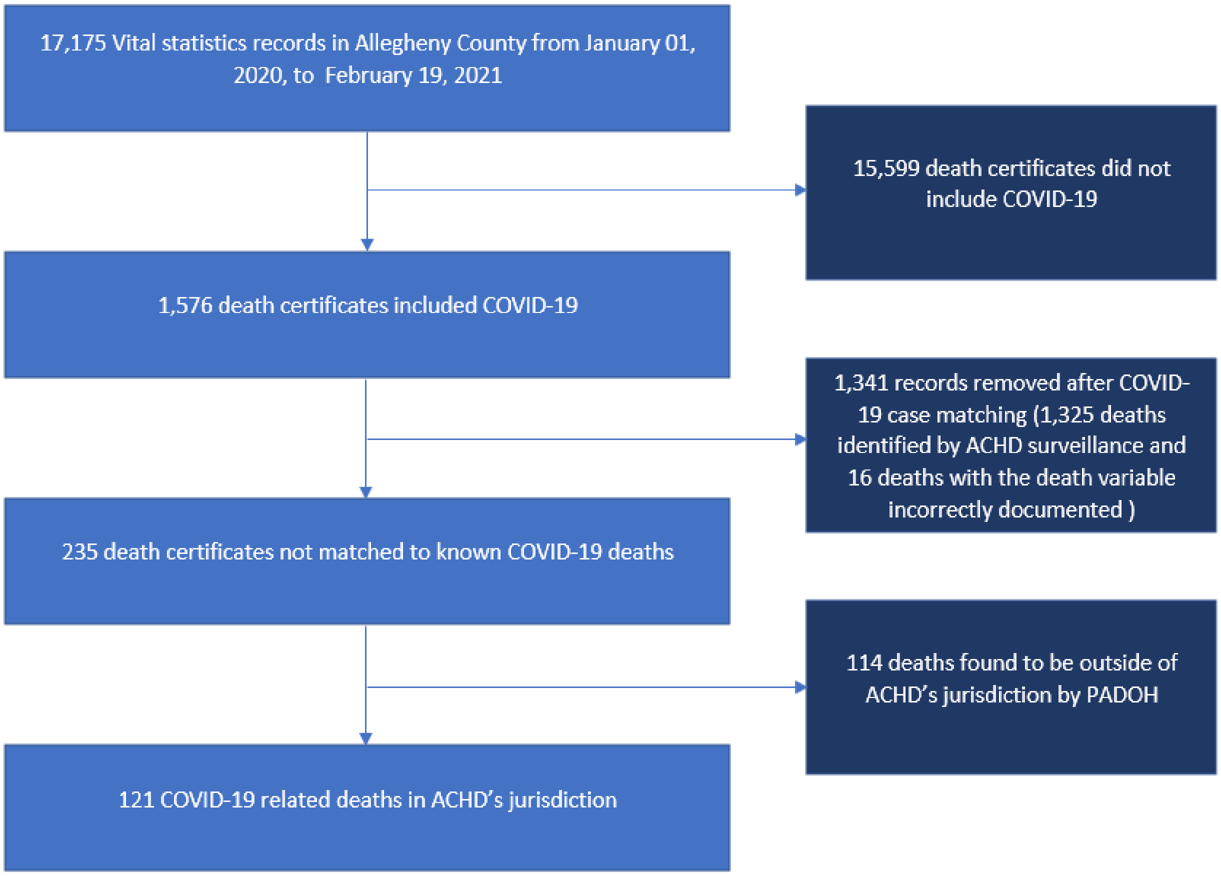

Figure 3 Determining additional COVID-19 deaths from vital statistics records not in the ACHD COVID-19 surveillance data

4.4 Long-Term Care Facilities

Among deaths identified by ACHD surveillance, 59% (786 of 1,325) were documented as having been in a long-term care facility. Although facility name is available in the PADOH vital statistics dataset, it does not differentiate hospitals and other non-long term care facilities from long-term care facilities and is not a reliable indicator of long-term care facility status. Therefore, long-term care facility status is unavailable among the additional COVID-19 deaths identified that were not in the ACHD surveillance data.

4.5 Associations Between Whether Deaths Were Identified by ACHD Surveillance or Not

The highest percentage of COVID-19 deaths occurred in December of 2020 among both deaths identified by and not identified by surveillance. The pattern of deaths among those identified by surveillance and not identified by surveillance follow approximately the same distribution from March 2020 to February 2021 (**Figure 4**). Whether COVID-19 deaths were identified by ACHD surveillance was not significantly associated with the month in which the death occurred. The largest percent difference by month of death between those identified and not identified by surveillance was 12% in December 2020 and the smallest percent difference was 0% in March, October, and November 2020 (p-value 0.24).

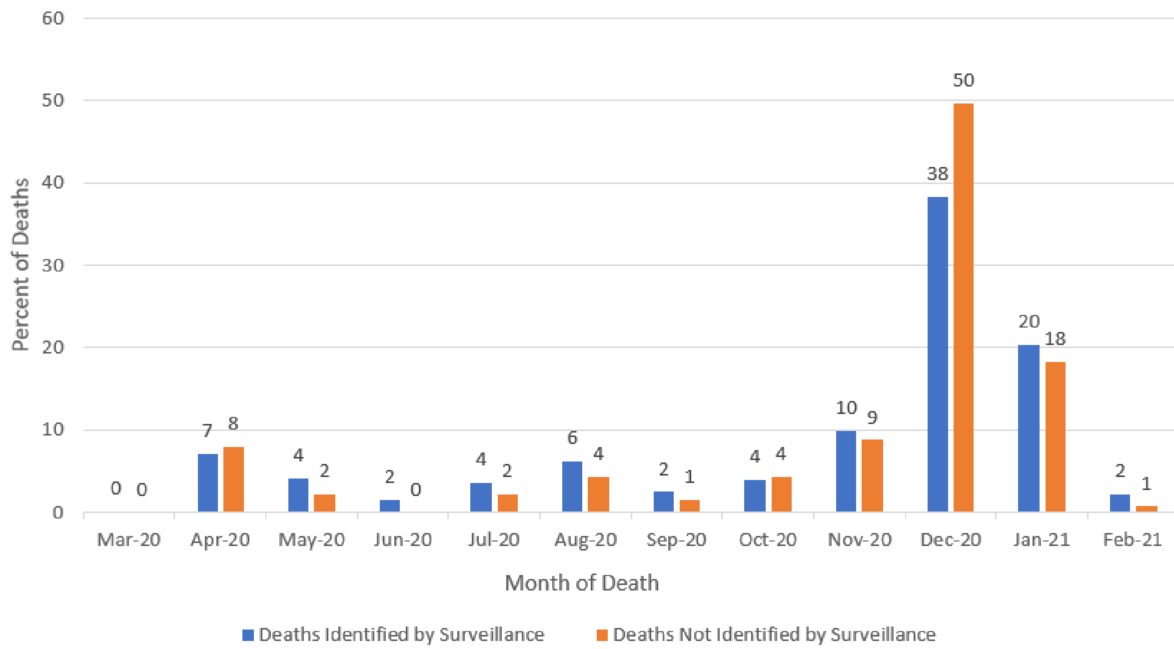

Figure 4 Percentage of deaths by month of death stratified by identification method

Sociodemographic information for deaths overall and stratified by identification method can be found in **Table 1**. Almost half (48%) of deaths were 85 years or older in age. Only three deaths occurred in persons <40 years (ages 28, 34, and 34). There was no significant difference in representation of age groups by how the deaths were identified. The largest percent difference by age between those identified and not identified by surveillance was 6% among those aged 85+ and the smallest percent difference was 0% among those aged 65-74 (p-value 0.30). There was not a significant difference across race/ethnicity groups by how the deaths were identified, with the largest percent difference by race between those identified and not identified by surveillance being 3% (for Other race/ethnicity) and the smallest percent difference was 1% (non-Hispanic White race/ethnicity) (p-value 0.07). There were more female than male deaths (55% versus 45%)

overall, and there was a significant difference in the identification method by sex; females were less likely to have been identified by ACHD surveillance (54% versus 65%; p-value 0.02).

Table 1 Characteristics of residents of Allegheny County with a COVID-19 related death between March 2020 and February 2021 overall and stratified by identification method

	Deaths Related to COVID-19			
	Overall N=1462	Identified by Surveillance n=1325	Not Identified by Surveillance n=137	P-value
Characteristics	%(n)*	%(n)	%(n)	
Age (years) ǂ**				0.30
18 – 64	9 (127)	9 (117)	7 (10)	
65 – 74	15 (222)	15 (201)	15 (21)	
75 – 84	29 (417)	29 (384)	24 (33)	
85 +	48 (696)	47 (623)	53 (73)	
Race/Ethnicity ǂ**				0.07
Non-Hispanic White	85 (1242)	85 (1127)	84 (115)	
Non-Hispanic Black	14 (198)	14 (181)	12 (17)	
Other	1 (21)	1 (16)	4 (5)	
Missing	(1)	(1)	(0)	
Sex §				**0.02**
Male	45 (654)	46 (606)	35 (48)	
Female	55 (808)	54 (719)	65 (89)	
COVID-19 Community Vulnerability Index ǂ**				0.20
Very Low (0.00 – 0.19)	55 (769)	56 (704)	49 (65)	
Low (0.20 – 0.39)	19 (269)	19 (238)	23 (31)	
Moderate (0.40 – 0.59)	14 (193)	14 (174)	14 (19)	
High – Very High (0.60 – 1.00)	12 (167)	12 (148)	14 (19)	
Missing	(64)	(61)	(3)	

Table 1. continued				
Month of Death ǂ				0.24
March 2020	0 (1)	0 (1)	0 (0)	
April 2020	7 (106)	7 (95)	8 (11)	
May 2020	4 (58)	4 (55)	2 (3)	
June 2020	1 (20)	2 (20)	0 (0)	
July 2020	3 (51)	4 (48)	2 (3)	
August 2020	6 (88)	6 (82)	4 (6)	
September 2020	2 (35)	2 (33)	1 (2)	
October 2020	4 (59)	4 (53)	4 (6)	
November 2020	10 (143)	9 (131)	9 (12)	
December 2020	39 (576)	38 (508)	50 (68)	
January 2021	20 (295)	20 (270)	18 (25)	
February 2021	2 (30)	2 (29)	1 (1)	

*Percentages may not total to 100% because of rounding

**Groups with less than 5% of the sample (age groups 18-54 and 55-64, Hispanic, non-Hispanic Asian, non-Hispanic Native American, and non-Hispanic Biracial, COVID-19 Community Vulnerability Index 0.60-0.79 and 0.80-1.00) were combined for testing associations

ǂ Cochran–Armitage test of association

Fisher's Exact test of association

§ Chi-squared test of association

Census tract data were missing for 64 of the 1,462 COVID-19 related deaths due to unverified and non-geocoded addresses in the vital statistics data; therefore, vulnerability index values are missing for 64 deaths. A map of mean vulnerability index by census tract in Allegheny County is shown in **Figure 5**. The majority (55%) of census tracks in Allegheny County had very low (<0.20) vulnerability. Five percent of deaths identified by surveillance and 2% of deaths not identified by surveillance did not have verified addresses, and thus could not be assigned the vulnerability index of their census tract. A large portion (45%) of the deaths missing vulnerability index data occurred more recently (in January or February of 2021). There was not a statistically significant association between vulnerability index and death identification method whether evaluated as ordinal (p-value 0.20) or binary (yes/no to high; p=0.40). The percent difference by categorical vulnerability index between the death identification methods was 2% for both low and high vulnerability. The largest difference within the ordinal vulnerability index by death identification method was 7% among low vulnerability and the smallest percent difference was 0% among moderate vulnerability. A sensitivity analysis was completed with very low vulnerability index (<0.20) and not very low vulnerability (>0.20), and there was not a significant association between this categorization of vulnerability and death identification method (p-value 0.11).

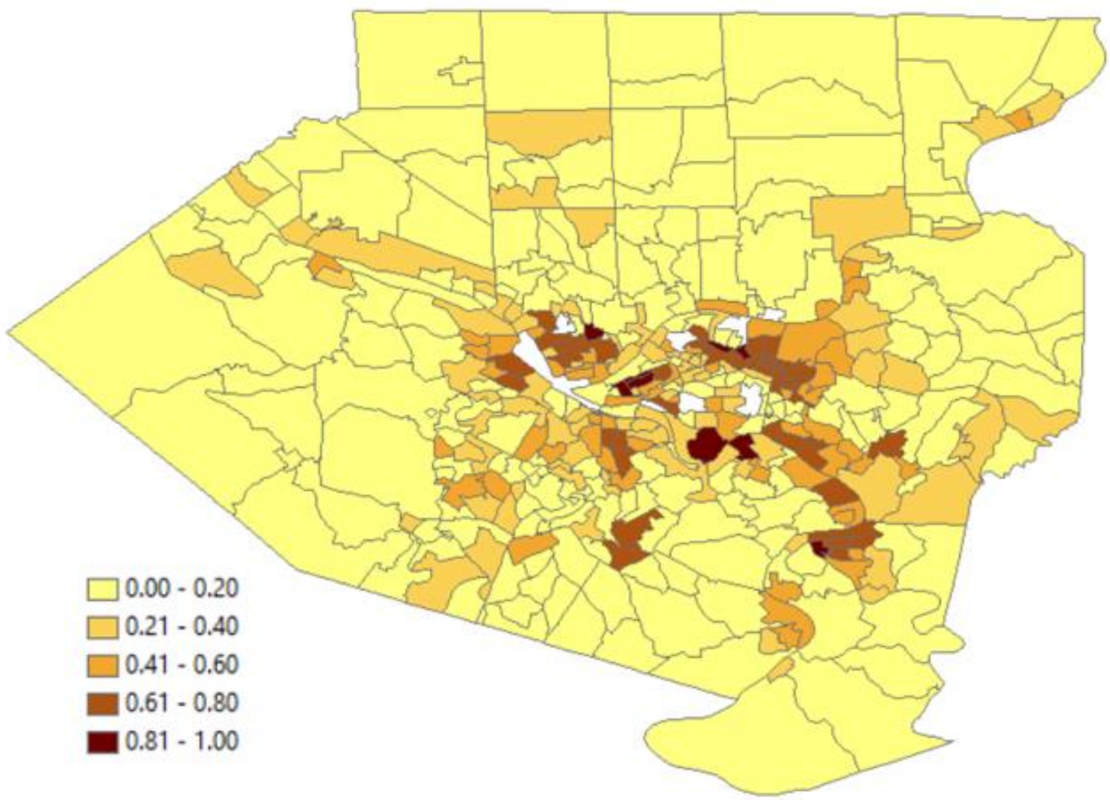

Figure 5 Mean vulnerability index by census tract in Allegheny County

Table 2 presents the multivariable logistic regression analysis including sex, age group, race/ethnicity, and ordinal vulnerability index. The interaction term of sex by age was not significant (p-value 0.78) so it was not included. The adjusted odds of being female versus male sex (OR=1.52, 95% CI 1.04-2.23) and race/ethnicity other than non-Hispanic White or Black versus non-Hispanic White (OR=3.49, 95% CI 1.22-10.01) were statistically significantly higher among those identified outside of ACHD surveillance In addition the odds of non-Hispanic Black deaths being identified outside of ACHD surveillance was 0.83 (95% CI 0.45-1.51) compared to non-Hispanic White deaths; however, this difference was not statistically significant. Age and vulnerability index were not significant predictors of identification method.

Table 2 Odds of a COVID-19 death not being identified by the ACHD COVID-19 surveillance system by age group, race/ethnicity, sex, and vulnerability index

Covariate	Deaths Related to COVID-19, n=1397		
	OR	95% CI	P-Value
Age (years)			0.54
18– 64	0.69	0.33-1.45	
65 – 74	0.95	0.56-1.63	
75 – 84	0.76	0.49-1.18	
85 +	Reference		
Race/Ethnicity			**0.04**
Non-Hispanic White	Reference		
Non-Hispanic Black	0.83	0.45-1.51	
Other	3.49	1.22-10.01	
Sex			**0.03**
Male	Reference		
Female	1.52	1.04-2.23	
COVID-19 Community Vulnerability Index			0.30
Very Low (0.00 – 0.19)	Reference	0.92-2.28	
Low (0.20 – 0.39)	1.45	0.72-2.16	
Moderate (0.40 – 0.59)	1.24	0.72-2.16	
High – Very High (0.60 – 1.00)	1.56	0.86-2.85	

My hypothesis was that Black race and high vulnerability would be associated with COVID-19 deaths not being identified by Allegheny County COVID-19 surveillance, while sex and age group would not be related to identification method. This hypothesis was shown to not be true, as female sex and Other race/ethnicity were associated with higher odds of deaths being identified outside of ACHD surveillance. Black race and high vulnerability were not associated with death identification method.

5.0 Discussion

There were 137 additional COVID-19 deaths identified during the Pittsburgh Summer Institute COVID-19 Vital Statistics Project. These deaths differed by sex and race/ethnicity when compared to the 1,325 COVID-19 related deaths identified by ACHD surveillance. There were no significant differences between deaths identified and not identified by ACHD surveillance when compared by vulnerability index, age, and month of death.

Among the deaths identified by ACHD COVID-19 surveillance, the 197 deaths that did not match to records in the PADOH vital statistics dataset may not have matched because the death certificates were not yet included in the PADOH vital statistics dataset, or the deaths were attributed to another county by PADOH. It is not likely that these deaths had not actually occurred and are much more likely attributed to the time needed to process vital statistics data. Deaths that were not ICD-10 coded were likely still being processed and should be analyzed in the future to identify if the deaths are related to COVID-19.

The majority of deaths (83%) attributed to COVID-19 by ACHD through February 19, 20201 were able to be merged with PADOH vital statistics data. Most (94%) deaths identified by ACHD surveillance that were ICD-10 coded included COVID-19 on the death certificate, which indicates that the methods ACHD employs to determine COVID-19 related mortality from COVID-19 surveillance data are largely effective. A small number (6%) of the deaths originally attributed to COVID-19 by ACHD did not include COVID-19 on the death certificate. The top causes of death for these deaths includes heart disease, dementia, Alzheimer's disease, sepsis, and pneumonia. These causes of death have been associated with COVID-19 illness, and these deaths may be related to COVID-19 despite COVID-19 not being listed on the death certificate.[42]

Almost half (114 of 235, 49%) of the additional deaths identified from vital statistics records that were not among COVID-19 surveillance data were found by PADOH to belong to a different jurisdiction. This is largely an effect of preliminary vital statistics data, as this analysis occurred prior to the cleaning of PADOH vital statistics data for Allegheny County to ensure the residency status for each death. The majority (73 of 121, 60%) of additional COVID-19 deaths identified from vital statistics data that were not among the ACHD COVID-19 surveillance data had the death certificate associated with "suspect" COVID-19 PA-NEDSS entries (suspect cases are defined as having met supportive laboratory evidence and no prior history of being a confirmed or probable case).[39] These entries would not have been included among the ACHD COVID-19 surveillance dataset, which explains why these deaths were not identified and counted by ACHD surveillance. The additional deaths found to be within ACHD's jurisdiction may have not been captured by ACHD surveillance because the individuals may not have been tested for or epi-linked to COVID-19 or may have been tested outside of Allegheny County or Pennsylvania. Deaths may have not been identified by surveillance simply because of the limitations of notifiable disease surveillance, such as underreporting, lack of representativeness as a result of underreporting, and lack of timeliness in reporting.[36]

The findings of this descriptive analysis indicate that COVID-19 mortality surveillance is largely indiscriminatory when accounting for age, race, ethnicity, sex, and vulnerability index. However, female sex was a significant predictor of the death being identified outside of ACHD surveillance. This is an unexpected finding, as I had hypothesized that sex would not be associated with the death identification method. Deaths among Black residents were less likely to be identified outside of ACHD surveillance methods compared to White residents, although this finding was not significant. However, Other race/ethnicity was a significant predictor for the deaths identified outside of surveillance. This finding is limited by a small sample size and diverse distribution of race and ethnicity within this categorization. Age and vulnerability index were not associated with how deaths were identified, which indicates that COVID-19 mortality surveillance in Allegheny County is fairly indiscriminatory regarding age and race/ethnicity aside from those classified as Other race/ethnicity for this analysis.

The distribution of the deaths that occurred towards the beginning of the pandemic were similar among the deaths identified by surveillance and those identified outside of surveillance. This is an interesting finding in that it could have been expected that more COVID-19 related deaths were missed at the beginning on the pandemic; however, this does not appear to be the case. The largest discrepancy pertaining to month of death was in December 2020, which is when the most deaths occurred among the deaths identified by ACHD surveillance and the deaths identified outside of ACHD surveillance. The spike in deaths identified outside of ACHD surveillance may have been due to the influx in total COVD-19 cases and deaths, which could have overwhelmed the public health workforce and contributed to the missed deaths.

An important strength of this analysis is the inclusion of all COVID-19 deaths that have been documented to have occurred within Allegheny County by ACHD through February 19, 2021. This provided a large sample size for comparison to the deaths missed by surveillance. Another strength of this analysis was the inclusion of findings from PADOH, as they determined that almost half of the deaths originally identified to be deaths missed by surveillance were outside of ACHD's jurisdiction, and the inclusion of those deaths in the analysis had the potential to influence the findings of this analysis. This analysis was able to identify and analyze deaths missed by surveillance, which has not been conducted for many counties so far because of the pandemic continuing to put strain on the public health workforce.

A limitation of this study is that the deaths with the death variable incorrectly documented in the COVID-19 surveillance dataset were only identified if there was an exact match by name and date of birth between vital statistics records and the COVID-19 surveillance dataset. This potentially misclassified deaths with errors in the name or date of birth within the COVID-19 surveillance dataset. To avoid this, deaths found to be among the PADOH vital statistics dataset that were not in the ACHD COVID-19 surveillance dataset were searched among the ACHD COVID-19 surveillance dataset to identify misclassified deaths. The deaths identified outside of surveillance were also determined to be COVID-19 related by ICD-10 codes, which may have missed some deaths with COVID-19 in the free-text of the death certificate but were missing the corresponding ICD-10 code, although this occurred at a very low frequency among the deaths identified by surveillance. Another limitation is that vital statistics records used to identify deaths related to COVID-19 were preliminary, incomplete, and are subject to be amended, limiting the number of deaths identified. A limitation in the analysis of vulnerability index was that 64 of the 1,462 total deaths in the analysis were missing values; however, 61 of the deaths missing values for vulnerability index were identified by surveillance. Analysis of vulnerability index was limited, as values were assigned to decedents based on their census tracts. Processing for decedent's addresses occurs at the state level, and these deaths may have been missing census tract information because the addresses were unverified, or the geocoding process had not yet occurred. This also means that these deaths could be assigned to another jurisdiction after this process and may be outside of ACHD's jurisdiction. Addresses for all decedents may be for either the long-term care facilities or for decedents' permanent addresses prior to living in a long-term care facility depending on the length of time the decent lived at the long-term care facility. This means that vulnerability index values may not be accurate based off the census tracts of some decedents.

Another limitation is that long term care facility could not be evaluated as a predictor of how deaths were identified, as long-term care facility residence was only known among the deaths identified from surveillance from the ACHD COVID-19 surveillance dataset. Facility is a variable in vital statistics records, but this includes hospital information in addition to long term care facilities and it could not be inferred whether a decedent was a resident of a long-term care facility from this information.

Future work should include an analysis by census tract or neighborhood between the groups identified and not identified by surveillance. This work could also include a spatial analysis to test for associations between neighboring groupings to potentially identify areas of Allegheny County with a disproportionate number of deaths missed by surveillance. Lastly, future work should identify deaths missed by surveillance after updated vital statistics records are provided by PADOH to include deaths that occurred past February 19, 2021.

In conclusion, ACHD COVID-19 mortality surveillance from March 14, 2020 to February 19, 2021 was largely accurate and indiscriminatory and is likely subject to the general limitations of passive surveillance systems. The identification of female sex and race/ethnicity other than White and Black as significant predictors of being missed by COVID-19 mortality surveillance, independent of age and vulnerability index were unexpected findings. With the identification of additional deaths missed by surveillance past February 19, 2021, these associations could be explored further, including the evaluation of more specific race/ethnicity categorizations. The public health implications of this work are that COVID-19 mortality surveillance conducted by ACHD for Allegheny County, Pennsylvania, is effective in capturing most COVID-19 related deaths; however, there are limitations to the surveillance system used and deaths identified outside of the surveillance system differ by sex and race/ethnicity compared to deaths identified by surveillance, although these differences are small.

Bibliography

1. Dhar Chowdhury S, Oommen AM. Epidemiology of COVID-19. *J Dig Endosc*. 2020;11(01):03-07. doi:10.1055/s-0040-1712187

2. Zhai P, Ding Y, Wu X, Long J, Zhong Y, Li Y. The epidemiology, diagnosis and treatment of COVID-19. *Int J Antimicrob Agents*. 2020;55(5):105955. doi:10.1016/j.ijantimicag.2020.105955

3. Liu Y, Rocklöv J. The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. *J Travel Med*. Published online August 9, 2021. doi:10.1093/jtm/taab124

4. Coronavirus (COVID-19) frequently asked questions | CDC. Accessed October 3, 2021. https://www.cdc.gov/coronavirus/2019-ncov/faq.html

5. Community Mitigation Framework | CDC. Accessed October 3, 2021. https://www.cdc.gov/coronavirus/2019-ncov/community/community-mitigation.html

6. Ali SH, Tozan Y, Jones AM, Foreman J, Capasso A, DiClemente RJ. Regional and socioeconomic predictors of perceived ability to access coronavirus testing in the United States: results from a nationwide online COVID-19 survey. *Ann Epidemiol*. 2021;58:7-14. doi:10.1016/j.annepidem.2021.03.001

7. Chau CH, Strope JD, Figg WD. COVID-19 Clinical Diagnostics and Testing Technology. *Pharmacotherapy*. 2020;40(8):857-868. doi:10.1002/phar.2439

8. Coronavirus Disease 2019 (COVID-19) Emergency Use Authorizations for Medical Devices | FDA. Accessed December 2, 2021. https://www.fda.gov/medical-devices/emergency-use-authorizations-medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices

9. COVID-19 Testing Overview. Accessed October 3, 2021. https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/testing.html

10. Rossen LM, Ahmad FB, Anderson RN, et al. Disparities in Excess Mortality Associated with COVID-19 - United States, 2020. *MMWR Morb Mortal Wkly Rep*. 2021;70(33):1114-1119. doi:10.15585/mmwr.mm7033a2

11. Woolf SH, Chapman DA, Sabo RT, Zimmerman EB. Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021. *JAMA*. Published online April 2, 2021. doi:10.1001/jama.2021.5199

12. Indirect Impact of COVID-19 Results in Higher Pandemic Death Toll. Accessed October 3, 2021. https://www.census.gov/library/stories/2021/02/indirect-impact-of-covid-19-results-in-higher-pandemic-death-toll.html

13. Asch DA, Islam MN, Sheils NE, et al. Patient and Hospital Factors Associated With Differences in Mortality Rates Among Black and White US Medicare Beneficiaries Hospitalized With COVID-19 Infection. *JAMA Netw Open*. 2021;4(6):e2112842. doi:10.1001/jamanetworkopen.2021.12842

14. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. Accessed October 3, 2021. https://covid19.who.int/

15. United States COVID: 44,490,464 Cases and 719,674 Deaths - Worldometer. Accessed October 3, 2021. https://www.worldometers.info/coronavirus/country/us/

16. Mortality Analyses - Johns Hopkins Coronavirus Resource Center. Accessed October 3, 2021. https://coronavirus.jhu.edu/data/mortality

17. www.health.pa.gov/topics/Documents/Reporting-Registries/SRN 2020-13 Additional Guidance for Medical Professionals - Reporting of Deaths Attributed to COVID-19.pdf.

18. NVSS - Provisional Death Counts for COVID-19 - Executive Summary. Accessed October 3, 2021. https://www.cdc.gov/nchs/covid19/mortality-overview.htm

19. Ahmad FB, Cisewski JA, Miniño A, Anderson RN. Provisional Mortality Data - United States, 2020. *MMWR Morb Mortal Wkly Rep*. 2021;70(14):519-522. doi:10.15585/mmwr.mm7014e1

20. Single-Race Population Estimates. Accessed November 1, 2021. https://wonder.cdc.gov/single-race-population.html

21. The COVID Racial Data Tracker | The COVID Tracking Project. Accessed November 1, 2021. https://covidtracking.com/race

22. Weinberger DM, Chen J, Cohen T, et al. Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Intern Med*. 2020;180(10):1336-1344. doi:10.1001/jamainternmed.2020.3391

23. Excess Deaths Associated with COVID-19. Accessed October 3, 2021. https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm

24. Pennsylvania Data. Accessed October 3, 2021. https://www.health.pa.gov/topics/disease/coronavirus/Pages/Cases.aspx

25. COVID-19 Deaths by Race/Ethnicity | KFF. Accessed November 3, 2021. https://www.kff.org/other/state-indicator/covid-19-deaths-by-race-ethnicity/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D

26. COVID-19 Has Claimed the Lives of 100,000 Long-Term Care Residents and Staff | KFF. Accessed November 3, 2021. https://www.kff.org/policy-watch/covid-19-has-claimed-the-lives-of-100000-long-term-care-residents-and-staff/

27. Vulnerability, Inequity, and COVID-19: A Portrait of the Pandemic in Allegheny County | RAND. Accessed October 3, 2021. https://www.rand.org/pubs/tools/TLA1080-1-v2/tool.html

28. Workbook: COVID-19 Summary. Accessed October 3, 2021. https://tableau.alleghenycounty.us/t/PublicSite/views/COVID-19Summary_16222279737570/COVID-19Summary?:showAppBanner=false&:display_count=n&:showVizHome=n&:origin=viz_share_link&:isGuestRedirectFromVizportal=y&:embed=y

29. National Research Council (US) Committee on National Statistics. The U.S. Vital Statistics System: A National Perspective. Published online 2009.

30. Technical Notes: Provisional Death Counts for Coronavirus Disease (COVID-19). Accessed October 3, 2021. https://www.cdc.gov/nchs/nvss/vsrr/covid19/tech_notes.htm

31. Siegel M, Critchfield-Jain I, Boykin M, Owens A, Nunn T, Muratore R. Actual Racial/Ethnic Disparities in COVID-19 Mortality for the Non-Hispanic Black Compared to Non-Hispanic White Population in 353 US Counties and Their Association with Structural Racism. *J Racial Ethn Health Disparities*. Published online August 30, 2021. doi:10.1007/s40615-021-01109-1

32. New Data Tool Will Help Track Inequities in COVID-19 Testing, Hospitalizations, and Deaths by UrbanKind Institute - issuu. Accessed October 3, 2021. https://issuu.com/urbankind/docs/press_release_new_data_tool_will_help_track_inequi

33. Surgo Ventures. COVID-19 Community Vulnerability Index (CCVI) Methodology. December 2020. Accessed January 12, 2021. https://www.precisionforcoviddata.org/

34. Surgo Precision For Covid. Accessed October 13, 2021. https://www.precisionforcoviddata.org/

35. Gundlapalli AV, Lavery AM, Boehmer TK, et al. Death Certificate-Based ICD-10 Diagnosis Codes for COVID-19 Mortality Surveillance - United States, January-December 2020. *MMWR Morb Mortal Wkly Rep*. 2021;70(14):523-527. doi:10.15585/mmwr.mm7014e2

36. Principles of Epidemiology: Lesson 5, Appendix E|Self-Study Course SS1978|CDC. Accessed November 1, 2021. https://www.cdc.gov/csels/dsepd/ss1978/lesson5/appendixe.html

37. Post LA, Issa TZ, Boctor MJ, et al. Dynamic Public Health Surveillance to Track and Mitigate the US COVID-19 Epidemic: Longitudinal Trend Analysis Study. *J Med Internet Res*. 2020;22(12):e24286. doi:10.2196/24286

38. Ibrahim NK. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. *J Infect Public Health*. 2020;13(11):1630-1638. doi:10.1016/j.jiph.2020.07.019

39. Coronavirus Disease 2019 (COVID-19) | CDC. Accessed October 1, 2021. https://ndc.services.cdc.gov/case-definitions/coronavirus-disease-2019-2021/

40. Provisional Death Counts for Coronavirus Disease 2019 (COVID-19). Accessed October 8, 2021. https://www.cdc.gov/nchs/nvss/vsrr/covid19/index.htm

41. Risk for COVID-19 Infection, Hospitalization, and Death By Age Group | CDC. Accessed October 19, 2021. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-age.html

42. Elezkurtaj S, Greuel S, Ihlow J, et al. Causes of death and comorbidities in hospitalized patients with COVID-19. *Sci Rep*. 2021;11(1):4263. doi:10.1038/s41598-021-82862-5