NORC at the University of Chicago

NIC

TECHNICAL REPORT:

The Impact of COVID-19 on Seniors Housing

June 3, 2021

Research You Can Trust™

Contents

Study Objective and Overview

Our study objective was to estimate the average number of deaths per 1,000 residents by senior congregate living (also known as seniors housing) level of care. We then compared those rates to the rates observed among older adults living in what is referred to in this report as non-congregate living seniors (e.g., those living in private homes as opposed to a seniors housing property) in similar areas. Seniors housing encompasses a wide range of care settings serving the different needs of older adults. This study focuses on estimating differences in COVID-19-related deaths across skilled nursing facilities (referred to in this report as skilled nursing or SNF) as compared to those experienced in the different levels of care in memory care properties (MC), assisted living properties (AL), and independent living properties (IL). To achieve this objective we needed to:

- Identify states for inclusion in the analysis.
- Identify sources of data to use in the analysis: For each state we had to find sufficient information to assign the name and location of unique seniors housing properties, the predominant level of care of the property, the number of residents living in each property as of the first quarter of 2020, and the cumulative number of deaths in the property that were attributed to COVID-19 between the start of the pandemic and the data collection date. We also used secondary analyses to evaluate COVID-19 cases and other issues.
- Match property information contained in different data sources into a combined file.
- Estimate the total number of deaths that occurred in each property.
- Estimate the number of residents in each property.
- Define selection and exclusion criteria to support valid comparisons across levels of care
- Conduct analyses.
- Perform sensitivity analyses to test the impacts of our assumptions.

Conceptually, we used different data sources to create an analytic dataset that included the following variables for each unit of analysis:

- Name
- State
- County
- Level of care, defined as older adults living in non-congregate settings (e.g., private homes) in counties (CN), independent living, assisted living, memory care, or skilled nursing facility
- Residents or population at risk as of the first quarter of 2020
- Cumulative deaths from COVID-19 among residents between the start of the pandemic and the closest date measured to December 31, 2020
- Diagnosed cases of COVID-19 in the property
- Cases of diagnosed COVID-19 per 100,000 people in the county of the observation

The units of analysis were individual properties or, when data were available, the county in which properties were located, excluding residents who lived in seniors housing settings. To support secondary analysis, we calculated indicators about whether a property was a continuing care retirement community (CCRC), also known as a life plan community. We also obtained information on whether a property was newly opened and in the process of leasing in the first

quarter of 2020, because newly opened properties likely increased their number of residents after the first quarter of 2020.

We analyzed data at the state level and across all included states using descriptive statistics and mixed binomial logistic regression. Information on the data, methods, and assumptions used in this analysis can be found below, followed by data documentation and sensitivity analyses of our main results.

1.0 Selection of States for Inclusion: We selected states for inclusion in the analysis based on the public availability of state-reported data on the number of deaths that occurred in seniors housing properties. These properties do not report deaths to the federal government unless those deaths occurred in their Centers for Medicare and Medicaid Services (CMS) qualified skilled nursing facility subunits. This means that state-reported data were the only available data source for this information. We gave preference to states in which we could also identify deaths that occurred in each county by age group, to allow us to estimate mortality rates among non-congregate living seniors (i.e., those who did not live in a seniors housing property). After reviewing data from all 50 states, we selected the following states for inclusion in the analysis: Connecticut (CT), Colorado (CO), Florida (FL), Georgia (GA), and Pennsylvania (PA). We chose CT based on the quality and comprehensiveness of property-level data, although state reports do not provide adequate information to determine the number of deaths that occurred among people aged 75 and older in each county.

2.0 Data Sources:
We developed state and cross-state data files using data from five sources:

- The CMS COVID-19 Nursing Home Dataset
- The NIC MAP® data, powered by NIC MAP Vision, an affiliate of the National Investment Center for Seniors Housing & Care (NIC)
- Publicly available state reports of deaths in seniors housing settings
- Publicly available state reports of all deaths by county and age group
- American Community Survey (ACS) data on population
- New York Times-reported information on COVID-19 cases

In general, we used CMS data to estimate deaths in skilled nursing facilities; NIC MAP data to assign property level of care and to estimate the size of the population at risk; state data to estimate deaths in seniors housing properties; state-reported county COVID-19 and ACS data to estimate deaths that occurred in non-congregate living adults (i.e., those who did not live in a seniors housing property); and New York Times information on reported COVID-19 case rates as a control variable to obtain adjusted results using a statistical model. However, there are exceptions to these general uses that are described in the sections below. The sections describe each data source and the data fields obtained from them. Subsequent sections contain information on how these data were combined and analyzed.

2.1 CMS COVID-19 Nursing Home Dataset: The Nursing Home COVID-19 Public File contains information reported by Medicare and Medicaid skilled nursing facilities to the CDC's National Healthcare Safety Network system. Documentation of the inclusion criteria and limitations of these data are reported elsewhere.[1] Data included in the system contain those

reported by skilled nursing facilities over whom CMS has regulatory authority. CMS notes that for a variety of reasons, data reported in this dataset may be inconsistent with those reported at the state level.

From this file, we selected all properties with skilled nursing beds in each state included in our analysis. The data system contains multiple reports for each skilled nursing facility. For use in this analysis, we selected either the report submitted for January 4, 2021, or the most recent report before January 4 if a property did not submit a January 4 report. From these data, we used the following variables in at least one step in our analysis:

Exhibit 1. Variables from the CMS COVID-19 Nursing Home Dataset Used in Analysis

COVID-19 Nursing Home Dataset Variable Name	Description	Use in Analysis
Week Ending	Indicates the week for which the report is applicable	Used to identify the data report for January 4, 2021, or the most recent report previous to January 4, 2021, if the January 4 report was missing
Provider Name	Name of the skilled nursing facility	Used to match the property to information on other lists
Provider Address	Address of the skilled nursing facility	Used to adjudicate matches of properties across file sources
Provider State	U.S. state	Used to locate the property in a state
County	State county	Used to locate the property in a county
Residents Total Confirmed COVID-19	Number of residents with laboratory-confirmed COVID-19 since January 1, 2020, as reported by the entity	Used to estimate cumulative COVID-19 cases in a property
Residents Total COVID-19 Deaths	Number of residents with suspected positive or laboratory-confirmed COVID-19 who died in the property or another location (COVID-19 DEATHS) since January 1, 2020 as reported by the entity	Used to estimate the deaths that occurred in the property

Number of All Beds	Total number of resident beds in the property as reported by the entity	Used in some cases in combination with other information to estimate the number of residents in a property

2.2 NIC, NIC MAP Property Inventory: NIC is a nonprofit 501(c)(3) organization whose mission includes developing objective, time-series data on seniors housing for use by investors and other organizations.[2] NIC is also the sponsor of this research. NIC maintains the NIC MAP data file and shared certain aspects of these data with the research team for independent evaluation and use in the analysis.[3] The NIC MAP data cover more than 15,000 seniors housing properties in 140 U.S. metropolitan areas and contains information on the following property attributes which we used in this analysis: property name; address, state and county; level of care used by the majority of its residents; number of studio, one bedroom, two bedroom, and three bedroom units at each level of care; operational status (i.e., under construction, newly opened and in leasing process, or established); whether a property is considered a CCRC; and estimated occupancy rates.

2.3 State Data on Deaths and Cases that Occurred in Seniors Housing Properties: We obtained information on deaths and cases that occurred in seniors housing properties from publicly available state data files. Reporting standards and formats are different in each state. We report the details of each state file below.

2.3.1 Colorado: We obtained data for Colorado from the Colorado State Emergency Operations Center's Outbreak Data Page.[4] Colorado's data are different from other states used in this analysis in that they contain information associated with specific identified outbreaks across many different setting types, including seniors housing properties. Outbreaks are divided into those considered to be active and resolved, and the same property can have multiple entries reflecting different outbreaks that occurred and were investigated at different times. We included all data for both resolved and active outbreaks reported through January 4, 2021. For each outbreak, Colorado's data contained the property name, county, setting type, and COVID-19 cases and deaths among residents and staff, and other data not included in our analysis.

We identified seniors housing properties that were of interest to this analysis as those defined with a setting type equal to assisted living, independent living, skilled nursing, or memory care (setting field codes all contained the suffix "health care" prior to these entries). Data contained information on the cumulative number of cases and deaths among residents and staff as well as other information not used in this analysis. We summed cases and deaths across outbreaks in instances where multiple outbreaks occurred.

2.3.2 Connecticut: We obtained data for Connecticut from the state's Assisted Living Facilities with Residents Positive for COVID-19 database available publicly online.[5] The data are obtained from self-reports from individual assisted living properties to the Connecticut Long Term Care Mutual Aid Plan web-based reporting system. Similar to CMS skilled nursing facility data, Connecticut's system contains individual reports filed on different dates that contain cumulative information on total cases and deaths that occurred up until each date. For this analysis, we

selected records corresponding to the report ending January 5, 2021, or the nearest date preceding it if a report for January 5, 2021, was not available. Connecticut reports data on the property name, town, and cases and deaths among staff and residents that we extracted and used in this analysis. Connecticut also reports data on the estimated census in each property, which we did not use to maintain consistency with other states (see Section 6.0 Estimation of Residents). Unlike other states, Connecticut does not report cumulative numbers across the entire pandemic. Instead, because of a reporting change occurring July 15, 2020, it reports cumulative cases and deaths through July 14, 2020, and cumulative cases and deaths between July 15, 2020, and the report date. We extracted each of these fields for use in analysis (see Section 5.0 Estimation of Deaths). Because our analyses required measurement at the county level, we assigned counties to each Connecticut town using public information. Connecticut provides no state-reported information for skilled nursing facilities, instead referring the public to the CMS COVID-19 Nursing Home Dataset.

2.3.3 Florida: We obtained data on Florida seniors housing properties using the publicly available Florida Long Term Care Facility Deaths Associated with COVID-19 Report, which contained data through January 8, 2021.[6] The report provides information regarding cumulative COVID-19 deaths among residents and staff associated with seniors housing properties. Florida does not provide information on cases and does not differentiate reporting by level of care.

2.3.4 Georgia: We obtained data for the Georgia Department of Community Health's Long Term Care Facility COVID-19 Report, with cumulative data through January 4, 2021.[7] Georgia reports data on COVID-19 activity for all licensed skilled nursing, all licensed assisted living communities, and licensed personal care homes of 25 beds or more. The report differentiates properties into personal care homes (PCH) and skilled nursing facilities. PCHs are defined as "a residence offering protective care and oversight of a person who needs a watchful environment but does not have an illness, injury or disability requiring chronic or convalescent care, such as medical or nursing services," including assisted living and memory care properties.[8] The data contains information on property type (nursing or PCH); the name, county, and address of the property; the resident census; cumulative COVID-19 positive residents and staff; and cumulative COVID-19 deaths among residents. As with Connecticut, in Georgia we used another method to estimate residents in each property (see Section 6.0 Estimation of Residents).

2.3.5 Pennsylvania: We obtained data for Pennsylvania from the state's publicly available Department of Human Services' COVID-19 Personal Care and Assisted Living Data.[9] Data contained cumulative information as of January 21, 2021. The Pennsylvania data contain information on PCHs, assisted living, and assisted living special care properties. The state defines PCHs as "residences that provide shelter, meals, supervision and assistance with personal care tasks, typically for older people, or people with physical, behavioral health, or cognitive disabilities who are unable to care for themselves but do not need nursing home or medical care."[10] Thus, PCHs on the Pennsylvania list may contain properties for persons other than the aged. However, we removed these properties from consideration in the analysis in a later state (see Section 3.0 Matching of Properties). Pennsylvania data contain information on cumulative resident and staff cases, and deaths among residents. Pennsylvania data contain many entries labelled "No data." At the guidance of seniors housing experts in Pennsylvania, we assumed these values were equal to zero. The state also redacts exact numbers for values between 1 and 4, labelling these as "Redacted (1-4 deaths)." We reset the value of these cells to 2.5, assuming an equal probability that a property could have reported 1, 2, 3, or 4 deaths.

2.4 American Community Survey Data on Population: To estimate mortality rates among non-congregate living adults (i.e. those who did not live in a seniors housing property) we used data from the American Community (ACS - accessed through the U.S. Census "Explore Census Data" portal) to estimate the number of persons ages 75 and older who lived in each county at the start of the pandemic based on the 5-year ACS estimate for 2019.[11]

2.5 County Cases Per 100,000 Residents: We obtained estimates of the number of diagnosed cases of COVID-19 per 100,000 residents in counties used in our analysis from the New York Times "Coronavirus in the U.S.: Latest Map and Case Count Data" page.[12] That page allows the user to select by state and obtain cumulative cases per 100,000 persons at the county level for each date. The data updates daily, and we obtained data for each state between January 15, 2021, and February 25, 2021.

3.0 Matching Properties: We combined data from different sources into a single analytic record using fuzzy matching techniques. For each state, we obtained data from three sources: 1) CMS COVID-19 Nursing Home Dataset, 2) NIC MAP property inventory, and 3) state-reported deaths in seniors housing properties. Combining data required matching records from different lists. Fuzzy matching was used to overcome issues of non-standardized recording of property name across different lists, for example, when a property was noted as "Property Name" in one data source and "Property Name, Inc." in another data source.

To do so, first we compared combined records; each of the three sources are compared to all other records from both the same data source and any other data source that we located within the same state and county. In a first matching step, we linked pairs with identical names. For remaining possible pairs we converted all property name characters to capital letters and then calculated variables measuring the generalized edit distance (GED) and the soundex differences between names. The GED measures dissimilarity between two text strings calculated using the COMPGED function in SAS 9.4 (Cary, N.C.).[13] The soundex is an alternative string conversion function that can also be converted to a numeric index to measure the similarity/dissimilarity between two text strings.[14]

We selected any possible pair within the same state and county with a GED of 600 or less, or a distance between soundex values of 75 or less for manual review. Manual review consisted of manual comparison of matched names—and property addresses when available—to determine which pairs should be linked. When reviewers determined that most reasonable people would accept two properties as being the same, the pair was assigned as matched. When a match was indeterminate, the pair was reviewed in tandem by two reviewers and in most instances were not considered a match. We then linked records from all three sources into a combined single record per property for analysis. Matching also resulted in identifying duplicate properties with multiple entries that had slight and inconsequential differences in character recording within the same data source. In these instances, we selected the record with the highest number of deaths and deleted the duplicate entry.

4.0 Defining Property Level of Care: We needed to identify the seniors housing property level of care (LOC) to estimate potential differences in deaths rates associated with LOC. In this analysis, we used LOC to designate the intensity of services offered in a property in addition to housing that supports people who are aging. LOC can also be understood as a proxy indicator of the health needs of the residents, as people with greater health services needs will require a

higher intensity LOC. We defined four levels of care for analysis in this study; independent living, assisted living, memory care, and skilled nursing facilities. We assigned LOC using NIC MAP data and CMS data. We used NIC MAP data for all states to support consistent measurement/identification across all areas included in the state. One state (Colorado) differentiated independent living and memory care properties, and data from three other states (Connecticut, Georgia, and Pennsylvania) allowed for differentiation between skilled nursing, assisted living, and memory care seniors housing properties.

4.1 Defining Level of Care Using NIC MAP Data: We used NIC MAP data to assign LOC designations to properties identified in state and CMS lists that were matched to properties contained in the NIC MAP data. We assigned LOC to properties based on the level of care used by the majority of their residents using the designations independent living, assisted living, or skilled nursing. NIC MAP also includes a variable flagging whether a property is classified as a CCRC. NIC MAP data define independent living as properties that typically include services such as communal dining, housekeeping, transportation, emergency call, and social programming services in the monthly fee. They define assisted living properties as those where residents receive personal care services such as bathing, dressing, eating, walking, and toileting. Twenty-four-hour protective oversight is provided, but 24-hour medical care is not. Finally, NIC MAP defines skilled nursing as a licensed long-term health care and residential property that serves people who require constant medical supervision and/or who require significant physical assistance in transferring, management of continence, and use of medical devices. Memory care was not a designated value in the NIC MAP data, as all memory care properties were designated as assisted living using the NIC MAP level of care designation. Memory care refers to properties who specialize in the care of patients with Alzheimer's disease, Alzheimer's Related Dementia, or other dementias of aging. We assigned properties to a memory care designation if 50 percent or more of their estimated residents live in memory care units (see Section 6.1 Estimation of Residents).

Note that NIC MAP data assigned LOC based on the resident status of the majority of its residents. Many properties include residences associated with different LOCs located within the same community under the same property name. Properties designated as independent living, assisted living, or memory care residences may also contain skilled nursing facilities, and reports regarding their skilled nursing facility residents may appear in the CMS data. When a property appeared in both the NIC MAP and the CMS COVID-19 data, we continued to designate a property's LOC based on its NIC MAP designation.

4.2 Defining Level of Care for Properties Not Included in the NIC MAP Data: There are several reasons why a property might not be included in the NIC MAP data. The NIC MAP only contains data for properties in metro areas tracked by NIC, that NIC considers to be market rate, and that have 25 or greater units. This excludes many properties contained in both the CMS data and in state lists of COVID-19 deaths. We assigned property types to those without a match in the NIC MAP data in different ways for different property LOC. We assigned the LOC for skilled nursing facilities identified in CMS COVID-19 data that were not found in NIC MAP data as skilled nursing facilities. Assisted living properties that were not in the NIC MAP data were classified as having an unknown LOC because there was no information that allowed us to differentiate their LOC. Deaths from these properties were used in our estimate of mortality rates for non-congregate living seniors (see Section 5.2 Estimating Deaths among Non-Congregate Living Adults), but they were otherwise excluded from the analysis.

4.3 Defining Level of Care for Older Adults Living in Non-Congregate Settings: In addition to the four LOCs described above, we defined a fifth LOC category to describe adults aged 75 and older residing in non-congregate living settings. Estimation of residents and deaths that occurred among these people are described below.

5.0 Estimating Deaths: We used data from state lists, CMS, and other state-reported data to estimate deaths in each unit of analysis (property or county), using different sources and methods for deaths in properties than in seniors living in non-congregate settings. We assumed that the comparable non-congregate living older adult population were people aged 75 and older who were not living in a seniors housing and setting. This assumption was based on data indicating that the average age of people in independent living was 82.2, the average age of people in assisted living was 85.3, and the average age of people in skilled nursing facilities was 83.5.[15] We acknowledge that some people living in seniors housing are younger than 75 and also that potentially a more appropriate non-congregate living population would be aged 80 and older.

5.1 Estimating Deaths in Properties: For properties with matches in one of either the state-reported list or the CMS COVID-19 data, we estimated deaths as the reported cumulative deaths reported for the property. For properties with entries in both the state file and the CMS COVID-19 file, the number of cumulative deaths reported often disagreed with no discernable pattern regarding whether state- or CMS-reported totals were higher. In these instances, we selected the maximum number of reported deaths across the two data sources under the rationale that we did not want to under-report deaths, and also assumed that it was unlikely that a property would report deaths that did not occur. State-reported deaths may exceed CMS reported deaths when a property contains residents at LOC outside their nursing unit. We do not know why deaths reported to CMS would exceed those reported to the state. We assumed that no deaths occurred in properties that appeared in the NIC MAP property inventory that had no match in either the state-reported data or the CMS file.

In Connecticut, deaths are reported as the cumulative total that occurred between the start of the pandemic and July 14, 2020, and those that occurred on July 15, 2021, and after. The state advises against summing these values, arguing that summing may result in some deaths being counted twice. In their summary of Connecticut Long Term Care data, the COVID Tracking Project lists summed totals from both periods and provides this warning “Connecticut’s system for reporting resident deaths changed on July 15, 2020, from the Office of the Chief Medical Examiner to Long-Term Care Mutual Aid Plan. Each system has differing definitions of COVID-19 associated deaths. As a result, there is a possibility of duplication of death counts. Interpret this data with caution.”[16] In our baseline analysis, we present Connecticut deaths as the sum of two periods.

5.2 Estimating Deaths among Non-Congregate Living Adults: To estimate COVID-19 deaths among non-congregate living adults, we first estimated the total number of COVID-19 deaths that occurred in each county among adults aged 75 and older, and next subtracted deaths that we observed in all seniors housing properties identified in the same county. We estimated total deaths among those aged 75 and older in different ways (described below) by state and county, reflecting differences in state and county death reporting across states. To estimate total deaths that occurred in seniors housing settings, we summed all observed deaths

in any property, whether we were able to include these properties in our final analysis of deaths rates or not. The purpose of this was to account for all deaths in seniors housing settings and exclude those deaths from non-congregate living adult totals so as not to inflate the number of deaths that we estimated occurred in the community. Caution should be used in interpreting county mortality rates, as properties with less than 25 residents are not included in CMS data or in many state reports. These deaths would be attributed to county residents as there was no way to subtract them from the total.

5.2.1: Estimating Non-Congregate Living Adult Deaths in Florida and Georgia: Both Florida and Georgia publicly supply individual-level data on all deaths reported to the state, and each observation includes information on the individual's age at death and county of residence.[17,18] These data supported easy estimation of deaths that occurred in each county in the state from the start of the pandemic through the same "as-of" dates as indicated for state long-term care property data (January 8, 2021, for Florida, January 4, 2021 for Georgia). To estimate these, we summed reported deaths among people aged 75 and older by county and then subtracted deaths that occurred in properties in those same counties from these totals.

5.2.2. Estimating Non-Congregate Living Adult Deaths in Colorado: Colorado does not report deaths by age group at the county level statewide. However information needed to construct this information was reported on county health department websites in seven counties (Adams, Arapaho, Boulder, Denver, Douglas, El Paso, and Jefferson). Adams, Arapaho, Boulder, and Douglas counties directly reported cumulative deaths among people aged 75 and older.[19,20] Data collected for Adams, Arapaho, and Douglas were cumulative through February 6, 2021, and data for Boulder were cumulative through February 11, 2021. Data for Denver, El Paso, and Jefferson were not reported in uniform categories. For Denver and El Paso, we estimated deaths among people aged 75 and older as one-half the deaths among people aged 70–79 plus the deaths among people aged 80 and older.[21,22] For Denver, we estimated deaths as the total number of deaths in Denver multiplied by the website-reported percentage of deaths that occurred among people aged 75 and older (as described above). Data for Denver were cumulative as of February 6, 2021, and data for El Paso as of February 15, 2021. Jefferson County reported the total number of deaths that occurred across all ages and the percentage of deaths that occurred among those aged 60–79 and among those 80 and older. To estimate deaths in Jefferson County among those aged 75 and older, we used data from Denver and El Paso to estimate the percent of all deaths in those counties that occurred among those aged 60–79 that were attributable to people aged 70–79, and assumed this same percentage applied to deaths in Jefferson. Using that estimate, we then estimated deaths in Jefferson as one-half the estimated deaths among people aged 70–79, plus deaths that occurred among those 80 and older. As in Georgia and Florida, after we estimated the total number of deaths among people aged 75 and older in the county, we subtracted the number of deaths observed in seniors housing properties in that county to estimate the number of deaths that were attributable to non-congregate living adults.

5.2.3 Estimating Non-Congregate Living Adult Deaths in Pennsylvania: Like Colorado, Pennsylvania does not report deaths by age group at the state level for individual counties. However, this information was available on county websites for Allegheny, Lancaster, Montgomery, and Philadelphia counties. Philadelphia County reported deaths among people aged 75 and older directly on its website.[23] Lancaster and Montgomery counties reported

deaths similarly to Denver and El Paso (i.e., deaths among those aged 70–79 and deaths among those aged 80 and older), so we estimated deaths using the same method as we did for those counties.[24,25] Allegheny reported the total number of deaths and the percentage that occurred among those aged 70 and older.[26] To estimate deaths for Allegheny, we used the estimated percentage of all deaths among people aged 70 and older in Lancaster and Montgomery that occurred among people aged 70–74 to adjust the Allegheny number downward to reflect only adults aged 75 and older. All data in Pennsylvania were updated as of February 19, 2021. As in other states, estimates of deaths among non-congregate living adults were created by subtracting the number of deaths observed in seniors housing properties from the estimated county total.

5.2.4 Estimating Non-Congregate Living Adult Deaths in Connecticut: The state of Connecticut did not provide data on deaths by age group at the county level and no such information was available from any county health department.

6.0 Estimating Residents: We estimated the population at risk (residents) in each unit of analysis with the intent of creating a consistently measured estimate of the number of persons living in each property as of the first quarter of 2020. We based this estimate on the estimated total occupancy of a property adjusted by its occupancy rate. This is an inexact measure of the population at risk because we measure deaths that occurred between the first quarter of 2020 and the as-of dates reported above, and people may have entered or existed the property since the first quarter. Two of the five states (Connecticut and Georgia) provide data on the resident census as of the report date, although information from Connecticut was often missing. The CMS COVID-19 Nursing Home Dataset also provides resident Census information as of the report date.

We used the approach described below instead of using this information because resident census information was not available for all states and we wanted to create a consistently measured estimate. Additionally, the resident census information were reported for the date of report, which was usually in early January 2021, and people in properties may have left the property or died between the first quarter of 2020 and the first week of January 2021. An alternative method could have been to select resident census data from the first available report for each property. However, this still would not provide resident information for three states, and the date of first report would be inconsistent across properties. Additionally, resident information for LOC other than skilled nursing facilities were not available in most states.

6.1 Estimation of Residents in Independent Living, Assisted Living, Memory Care, and Skilled Nursing Properties that Were Listed in the NIC MAP Property Inventory: We used data on residences and occupancy rates contained in the NIC MAP property inventory for the subset of properties that were listed in that source. The NIC MAP property inventory contains information on the number of studio, one-bedroom, two-bedroom, and three-bedroom residences at each LOC within each property listed. For independent living, assisted living, and memory care we converted this to each property's total possible occupancy using the assumption of one person living in each studio and one-bedroom unit and two people living in each two- and three-bedroom unit. For skilled nursing facilities, we estimated total possible occupancy as the total number of skilled nursing facilities beds in each property. Many properties contained residents living in different LOC. We then adjusted the estimated total

occupancy by multiplying the total by the estimated occupancy rate for first quarter of 2020, to estimate the number of residents living in the property as of the first quarter of 2020.

A small number of properties reported only the total number of units and did not break these units out by unit mix. For these properties, we estimated the total possible occupancy as the number of units multiplied by the average number of occupants per unit seen in the same LOC in the same state. We excluded units that were listed units as being under construction as of the first quarter of 2020 from the analysis. For the small subset of properties with missing occupancy information, we applied the average occupancy information for other properties at the same LOC in the same state.

6.2 Estimation of Residents in Properties that Were Not Listed in the NIC MAP: We estimated the number of residents of skilled nursing facilities that were not listed in the NIC MAP property inventory based on the CMS COVID-19 Nursing Home Dataset-reported available beds multiplied by the average occupancy rate for NIC MAP-reported skilled nursing facilities in the same state. We did not estimate residents for properties at other LOC reported on state lists that were not also included in the NIC MAP inventory because we had no consistent way to do so across the states.

6.3 Estimation of Non-Congregate Living Adults: We estimated the number of comparable non-congregate living adults in each county as the number of people aged 75 and older within each county as reported in the five-year ACS estimates for 2019, minus the sum of the number of residents estimated to be living in each seniors housing property located in each county.

7.0 Analysis Inclusion Criteria: For analysis, we restricted our sample to properties and counties located in counties in each state for which data were collected for at least one property in the NIC MAP property inventory. We further excluded properties identified in state long-term care datasets that were not matched to properties in the NIC MAP property inventory or the CMS COVID-19 data because of an inability to estimate the number of residents living in those properties (Colorado, Florida, Pennsylvania) or an inability to determine the LOC of the property (all states except Colorado). The NIC MAP property inventory tracks only senior housing properties that are considered to be investment grade. Neither the NIC MAP nor CMS data track properties that have 25 or fewer residents. Properties listed on state lists that are not included in the NIC MAP property inventory may not be seniors housing (i.e., group homes for persons with disabilities), house fewer than 25 residents, or be state subsidized or charity properties. These exclusions should be considered when evaluating our results. We included skilled nursing facilities not listed on the NIC MAP property inventory that were located in included counties, because the CMS COVID-19 Nursing Home Dataset allowed us to estimate residents and LOC (skilled nursing).

We additionally excluded properties that NIC MAP data identified as being under construction or in what they term as a “lease-up” period during the first quarter of 2020. Lease-up refers to a newly opened property that is filling residences. We excluded these properties because the number of residents in them likely changed substantially over the observation period. Additionally, we excluded properties in either the NIC MAP property inventory or the CMS COVID-19 Nursing Home Dataset which had missing data for residents. Any deaths identified in properties that were excluded from our analysis dataset were included when creating the estimate of deaths that occurred among non-congregate living seniors.

For non-congregate living adults, we created observations for counties for all counties in areas covered by the NIC MAP property inventory for which it was possible to estimate deaths among people aged 75 and older.

Exhibits 2 through 7 provide the number of counties included in the analysis overall and in each state, and the corresponding number of observations at each LOC in each state and within each county at each state.

Exhibit 2. Observations[1] Included in the Analysis by State and Level of Care

State	Included Counties[2]	CN[3]	IL[4]	AL[5]	MC[6]	SNF[7]	Total
Colorado	12	7	59	88	23	177	354
Connecticut	7	0	36	72	14	260	382
Florida	30	30	194	500	89	766	1,579
Georgia	34	34	52	166	28	181	461
Pennsylvania	30	4	141	300	26	645	1,116
All States	**113**	**75**	**482**	**1,126**	**180**	**2,029**	**3,892**

NOTES: 1. An observation refers to a seniors housing property or the estimated population of non-congregate living adults aged 75 and older for which we identified sufficient data to estimate deaths from COVID-19. 2. Included Counties refers to counties that were eligible for the analysis based on containing at least one seniors housing property listed in the NIC MAP property inventory. 3. CN stands for county residents and is the abbreviation used to refer to observations of county populations with sufficient data to estimate deaths from COVID-19. 4. IL, Independent Living. 5. AL, Assisted Living. 6. MC, Memory Care. 7. SNF, Skilled Nursing.

Exhibit 3. Colorado: Observations by County and Level of Care

State	County	CN	IL	AL	MC	SNF	Total
Colorado	ADAMS	1	4	7	0	22	34
	ARAPAHOE	1	9	11	7	30	58
	BOULDER	1	10	12	2	15	40
	BROOMFIELD	0	2	1	1	2	6
	DENVER	1	11	14	2	33	61
	DOUGLAS	1	5	4	2	9	21
	EL PASO	1	8	17	5	28	59
	ELBERT	0	0	0	0	2	2
	JEFFERSON	1	10	22	4	30	67
	MORGAN	0	0	0	0	3	3
	TELLER	0	0	0	0	2	2
	WASHINGTON	0	0	0	0	1	1
	Total	**7**	**59**	**88**	**23**	**177**	**354**

Exhibit 4. Connecticut: Observations by County and Level of Care

State	County	CN	IL	AL	MC	SNF	Total

Connecticut	FAIRFIELD	0	8	21	7	54	90
	HARTFORD	0	16	21	3	81	121
	MIDDLESEX	0	1	7	1	21	30
	NEW HAVEN	0	10	17	2	70	99
	NEW LONDON	0	1	3	1	20	25
	TOLLAND	0	0	1	0	6	7
	WINDHAM	0	0	2	0	8	10
	Total	**0**	**36**	**72**	**14**	**260**	**382**

Exhibit 5. Florida: Observations by County and Level of Care

State	County	CN	IL	AL	MC	SNF	Total
Florida	BAKER	1	0	0	0	2	3
	BREVARD	1	5	23	6	29	64
	BROWARD	1	15	36	6	53	111
	CHARLOTTE	1	1	12	0	10	24
	CLAY	1	3	7	0	13	24
	COLLIER	1	12	8	5	13	39
	DUVAL	1	6	26	6	44	83
	FLAGLER	1	1	5	1	3	11
	HERNANDO	1	2	7	0	8	18
	HIGHLANDS	1	0	6	1	6	14
	HILLSBOROUGH	1	11	42	8	40	102
	INDIAN RIVER	1	5	8	1	7	22
	LAKE	1	5	16	1	24	47
	LEE	1	14	19	6	22	62
	MANATEE	1	6	13	5	19	44
	MARTIN	1	3	9	2	9	24
	MIAMI-DADE	1	6	17	0	76	100
	NASSAU	1	2	2	1	3	9
	ORANGE	1	4	22	2	43	72
	OSCEOLA	1	2	5	0	13	21
	PALM BEACH	1	26	35	8	67	137
	PASCO	1	1	21	1	24	48
	PINELLAS	1	16	46	10	96	169
	POLK	1	10	19	3	32	65
	SAINT JOHNS	1	5	7	2	7	22
	SAINT LUCIE	1	2	7	1	12	23
	SARASOTA	1	13	30	7	36	87

	SEMINOLE	1	5	19	3	12	40
	SUMTER	1	6	6	1	2	16
	VOLUSIA	1	7	27	2	41	78
	Total	**30**	**194**	**500**	**89**	**766**	**1579**

Exhibit 6. Georgia: Observations by County and Level of Care

State	County	CN	IL	AL	MC	SNF	Total
Georgia	BARROW	1	0	3	0	1	5
	BARTOW	1	0	4	2	5	12
	BURKE	1	0	0	0	4	5
	BUTTS	1	0	3	0	2	6
	CARROLL	1	0	5	0	6	12
	CATOOSA	1	0	1	0	3	5
	CHEROKEE	1	3	14	1	4	23
	CLAYTON	1	1	1	2	7	12
	COBB	1	3	25	4	18	51
	COLUMBIA	1	2	4	2	4	13
	COWETA	1	2	3	0	5	11
	DADE	1	0	0	0	1	2
	DAWSON	1	0	1	0	0	2
	DEKALB	1	6	15	2	23	47
	DOUGLAS	1	0	3	0	2	6
	FAYETTE	1	2	5	1	3	12
	FORSYTH	1	3	8	2	4	18
	FULTON	1	20	30	6	24	81
	GWINNETT	1	7	18	3	13	42
	HARALSON	1	0	0	0	5	6
	HEARD	1	0	0	0	1	2
	HENRY	1	0	3	2	4	10
	LAMAR	1	0	0	0	2	3
	MCDUFFIE	1	0	0	0	1	2
	MERIWETHER	1	0	1	0	2	4
	MORGAN	1	0	0	0	1	2
	NEWTON	1	0	3	0	2	6
	PAULDING	1	1	3	1	1	7
	PICKENS	1	0	1	0	4	6
	RICHMOND	1	1	4	0	15	21
	ROCKDALE	1	1	1	0	3	6
	SPALDING	1	0	1	0	5	7
	WALKER	1	0	3	0	4	8

	WALTON	1	0	3	0	2	6
	Total	**34**	**52**	**166**	**28**	**181**	**461**

Exhibit 7. Pennsylvania: Observations by County and Level of Care

State	County	CN	IL	AL	MC	SNF	Total
Pennsylvania	ADAMS	0	2	3	0	6	11
	ALLEGHENY	1	17	49	5	79	151
	ARMSTRONG	0	0	6	0	4	10
	BEAVER	0	0	8	0	8	16
	BERKS	0	5	18	0	20	43
	BUCKS	0	12	14	4	34	64
	BUTLER	0	4	8	1	16	29
	CARBON	0	0	3	0	4	7
	CHESTER	0	12	17	0	26	55
	COLUMBIA	0	0	0	0	6	6
	CUMBERLAND	0	7	5	1	18	31
	DAUPHIN	0	3	6	1	12	22
	DELAWARE	0	10	7	0	39	56
	FAYETTE	0	1	7	0	9	17
	JEFFERSON	0	0	0	0	4	4
	LACKAWANNA	0	0	12	2	24	38
	LANCASTER	1	18	14	1	30	64
	LEBANON	0	3	8	0	11	22
	LEHIGH	0	3	12	2	20	37
	LUZERNE	0	2	13	1	37	53
	MERCER	0	1	5	0	14	20
	MONTGOMERY	1	21	23	4	70	119
	NORTHAMPTON	0	6	13	2	15	36
	PERRY	0	0	0	0	4	4
	PHILADELPHIA	1	5	12	0	67	85
	PIKE	0	0	2	0	1	3
	WASHINGTON	0	1	6	0	17	24
	WESTMORELA	0	3	16	2	29	50
	WYOMING	0	0	1	0	2	3
	YORK	0	5	12	0	19	36
	Total	**4**	**141**	**300**	**26**	**645**	**1116**

8.0 Estimating Residents, Deaths, and Raw Mortality Rates by Property Type and State:
We summed estimated residents and deaths by level of care, state, and across states. We used

these summed results to estimate the number of deaths per 1,000 residents at each LOC. We used the SAS Survey Means procedure (SAS 9.4, Cary N.C.) to estimate the mean death rate per 1,000 residents at each LOC, and their associated 95 percent confidence intervals accounted for clustering of observations within county. We estimated the statistical significance of pairwise comparisons of mortality rates between levels of care among all observations using t-tests with Bonferroni adjustments for multiple comparisons.

Limitations of the data should be considered when evaluating these raw results, in particular: differences in the estimated COVID-19 cases diagnosed per 100,000 observed; differences in potential risk for COVID-19 mortality among residents in different LOC; the inclusion/exclusion criteria for state selection and selection of counties and properties; and the fact that properties at each LOC were not equally distributed across the counties included. Exhibit 8 presents the average estimate of surrounding county COVID-19 cases per 100,000 residents observed for each LOC within each state for the properties that were included in the analysis. We found similar average case rates across LOC within each state, and some differences (notably, a lower rate of reported cases in Florida counties) across states. Subsequent sections use statistical modeling to re-estimate deaths per 1,000 residents by LOC, adjusting for differences in the COVID-19 case rates between counties.

Exhibit 8. Average Case Rate per 100,000 Residents in the Counties that Surrounded Each Level of Care in Each State

	CO		CT		FL		GA		PA	
LOC	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
CN	7,299	1,300			4,940	1,547	7,721	1,175	6,865	805
IL	7,105	1,107	7,875	410	5,092	1,322	7,338	927	7,169	1,130
AL	7,159	1,065	7,743	689	4,946	1,413	7,723	1,001	7,241	1,241
MC	6,991	783	7,895	610	4,877	906	7,597	1,068	7,047	720
SNF	7,409	1,172	7,682	702	5,441	1,937	7,695	1,106	7,165	1,020

Exhibit 9 provides the estimated deaths per 1,000 residents and 95% confidence intervals by LOC across all included states. Exhibit 10 provides the estimated total number of residents and deaths, estimated death rate per 1,000 residents, and 95% confidence intervals for each rate stratified by state and LOC. Recall that we did not identify data in Connecticut with which we were able to estimate COVID-19 deaths by county among people aged 75 and older. Because of this, we do not provide a total estimated death rate across LOC for Connecticut, as this measure would appear artificially high when compared to other states because of the exclusion of county data.

Exhibit 9. Estimated COVID-19 Deaths per 1,000 Residents Observed in 2020 among Non-Congregate Living Adults Ages 75 and Older and Adults Living in Differing Levels of Care of Seniors Housing

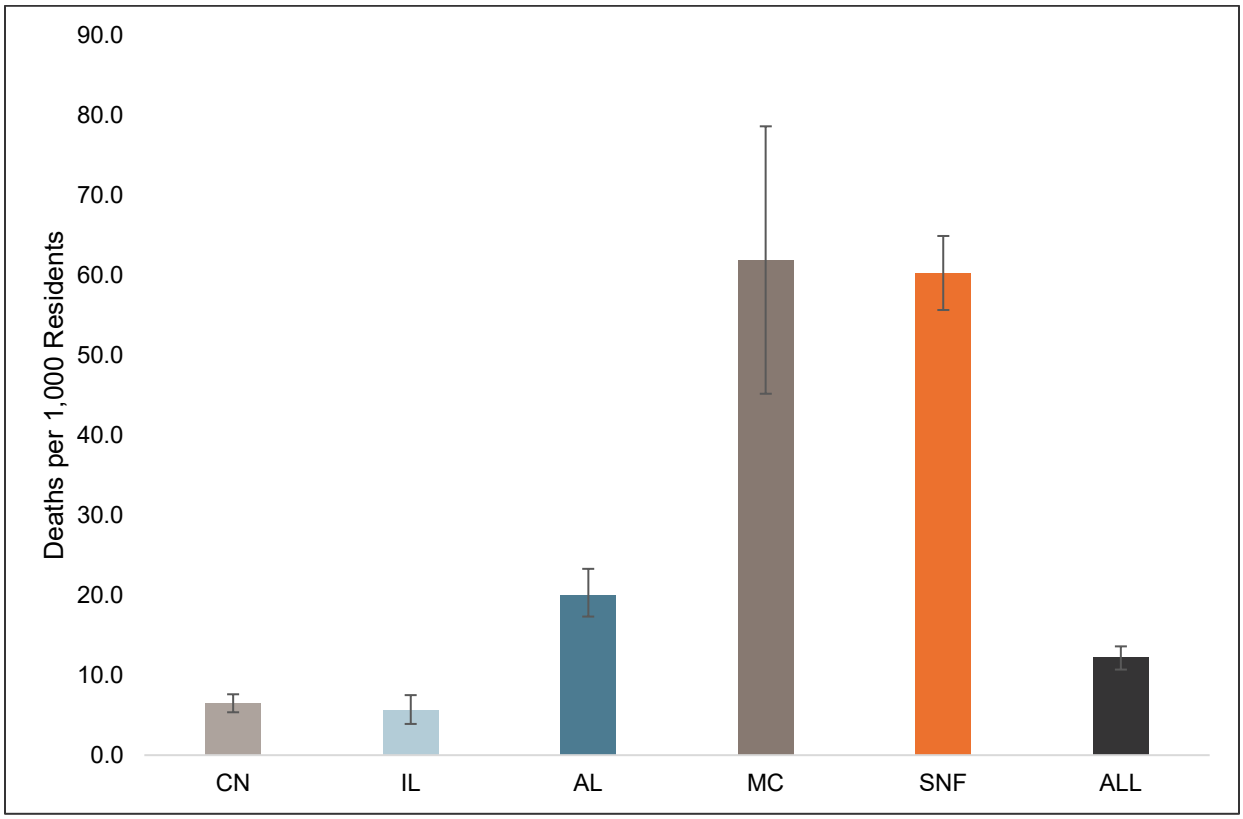


Exhibit 10. Estimated Residents, Deaths, Deaths per 1,000 Residents, and 95% Confidence Interval Observed in Each State and Overall

State/LOC	Residents	Deaths	Deaths Per 1,000	Lower 95% CI	Upper 95% CI
Colorado					
CN	143,247	962	6.7	5.3	8.1
IL	14,141	77	5.4	1.5	9.4
AL	7,737	107	13.8	6.2	21.4
MC	1,101	78	70.8	40.1	101.6
SNF	16,644	1,233	74.1	62.1	86.1
ALL	182,870	2,457	13.4	11.9	15.0
Connecticut					
CN	No Data				

IL	7,286	196	26.9	11.9	41.9
AL	6,225	245	39.4	26.5	52.2
MC	852	72	84.5	4.6	164.4
SNF	30,879	1,905	61.7	49.6	73.8
Florida					
CN	1,449,212	8,943	6.2	4.6	7.7
IL	68,767	304	4.4	2.7	6.1
AL	41,449	728	17.6	13.7	21.4
MC	4,075	256	62.9	37.7	88.2
SNF	88,446	5,296	59.9	51.2	68.6
ALL	1,651,950	15,527	9.4	8.3	10.5
Georgia					
CN	240,201	1422	5.9	5.1	6.8
IL	11,642	27	2.3	0.0	5.0
AL	10,821	135	12.5	8.4	16.5
MC	1,108	13	11.7	0.0	24.2
SNF	21,122	1242	58.8	48.4	69.2
ALL	284,894	2,839	10.0	9.0	11.0
Pennsylvania					
CN	229,480	2,051	8.9	5.8	12.0
IL	63,081	331	5.2	3.4	7.1
AL	21,525	536	24.9	19.4	30.4
MC	1,256	97	77.2	21.0	133.4
SNF	85,093	4,898	57.6	48.8	66.3
ALL	400,435	7,913	19.8	14.0	25.5
ALL Included					
CN	2,053,582	13,360	6.5	5.4	7.6
IL	164,921	935	5.7	3.9	7.5
AL	87,755	1,751	20.0	17.3	23.3
MC	8,391	516	61.8	45.2	78.6
SNF	242,185	14,574	60.2	55.6	64.9
ALL	2,556,833	31,135	12.2	10.7	13.6

NOTES: CN stands for county and is used to refer to older adults living in non-congregate settings in the broader counties, IL refers to independent living, AL refers to assisted living, MC refers to memory care, and SNF refers to skilled nursing facility. Total combines all estimates of residents and deaths. Note that non-congregate living

residents and deaths in Connecticut and for some counties in Colorado and Pennsylvania were not estimated and are not included in this total. In the "All Included" results, differences between COVID-19 mortality rates by LOC were significant at the p<.01 level for all LOC except CN and IL, and MC and SNF. All other pairwise comparisons between LOC were statistically significant at the 1% level using Bonferroni adjustments for multiple comparisons.

9.0 Expected Mortality Rates per 1,000 Adjusting for Differences in Case Rates: Differences in mortality rates among individual properties and across states are likely driven in part by differences among the counties in which they are located and differences in the severity of the COVID-19 epidemic in those counties. To account for these differences, we used a mixed binomial marginal model to estimate the death rate and 95% confidence interval for each LOC controlling for case count and clustering of residuals among properties at the county level using an R-sided random effect.[27,28]

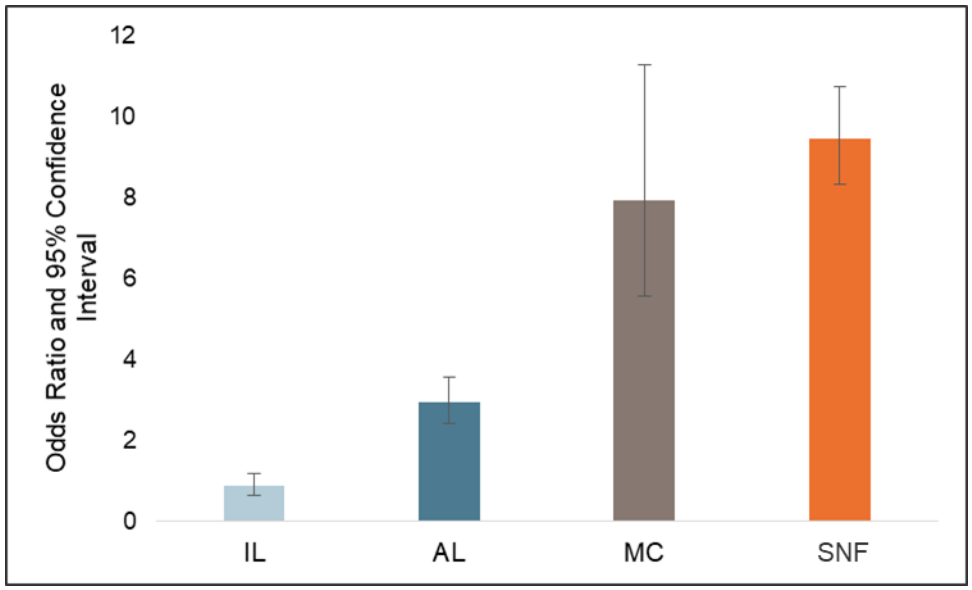
In this model the independent variables was LOC (CN, IL, AL, MC, and SNF) which was treated as an indicator variable with SNF as the reference, the interaction of level of care and state, and the county case rate among all residents observed in the county the property is located. The random effect was specified as observations clustering within county estimated using an assumption of compound symmetry covariance matrix. In this model formulation. The random effect is incorporated into the residual to correct for over dispersion and clustering in the outcome variable.

We estimated the model in SAS 9.4 (Cary, N.C.) using the Glimmix procedure and the Event/Trial syntax, with bias corrected sandwich estimates of variance. We used the LSMEANS statement with the iLink option to estimate the conditional mean death rate at each level of care and their corresponding 95% confidence intervals, setting the expected case rate to the average case rate seen across levels of care among all observations (6,496), and used Tukey-Kramer-adjusted confidence intervals to adjust for multiple comparisons to make pair-wise tests of differences in the expected least squared means. Because these confidence intervals are derived from the inverse log-link function to the least squared means, they are not symmetrical.[27] Finally, for presentation, we transformed mean mortality rates and their CIs into a deaths per 1,000 scale by multiplying final estimates through by 1,000.

9.1 Results: The estimated model showed decent statistic properties and significant fixed and random effects. The generalized Chi-Square statistic divided by degrees of freedom was equal to 1.51, indicating that modelled R-Side effects controlled for over dispersion and the random effect was significant (p<.0001). The Type III Tests of the fixed effects were significant for LOC (p<.0001); LOC interacted with state (p<.0001); and estimated COVID cases per 100,000 residents in the surrounding community (p<.01).

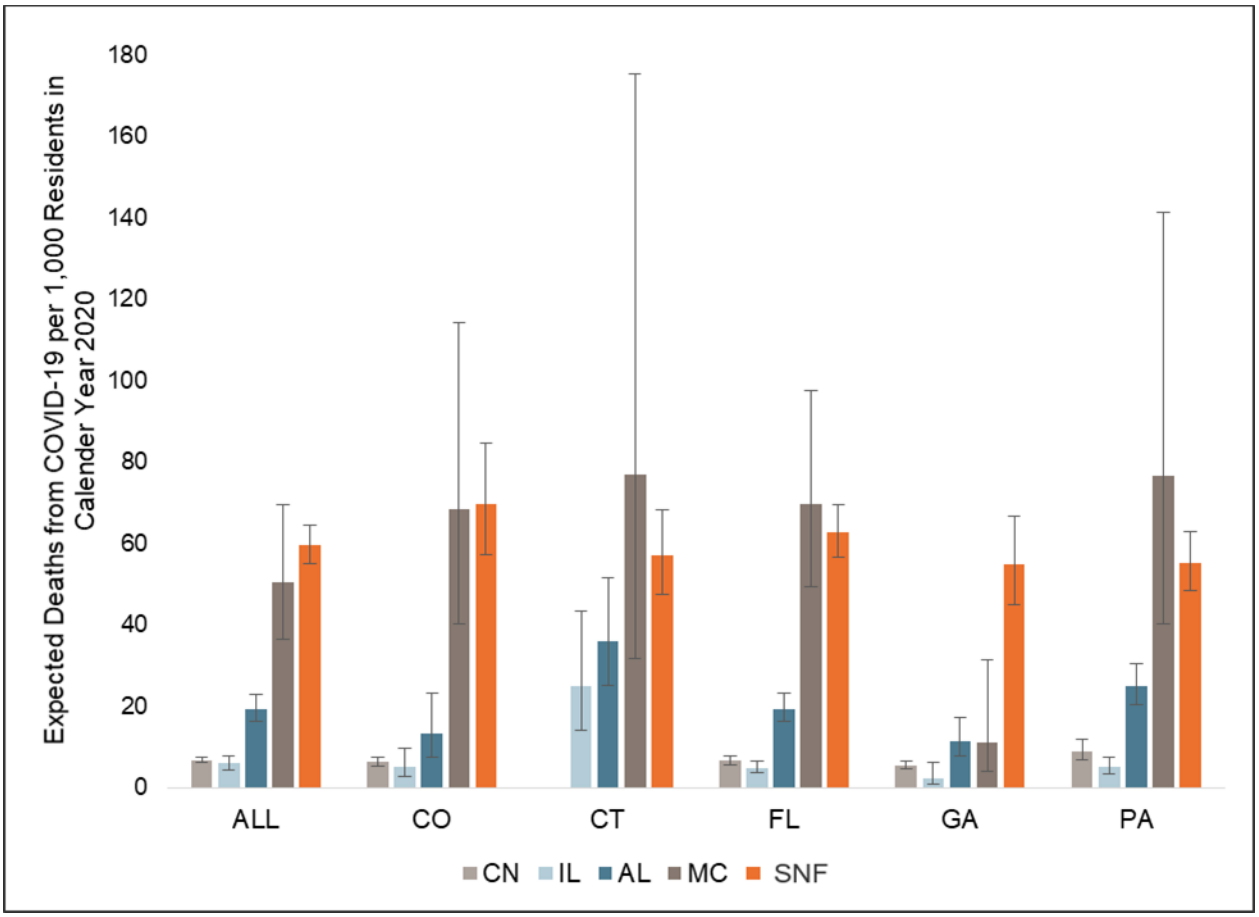
9.1.1 Comparative Odds of Death: We computed the adjusted odds ratio (aOR) of the risk of death in each LOC as compared to non-congregate living adults controlling for the case rate observed across observations. The aORs for AL (2.92, 95% CI 2.40–3.56); MC (7.92, 95% CI 5.54–11.30); and SNF (9.45, 95% CI 8.33–10.73) were statistically significantly higher than for non-congregate living adults. The aOR (0.88, 95% CI 0.65–1.19) for a person in IL was lower than for non-congregate living adults, but this result was not statistically significant.

Exhibit 11. Odds Ratio of the Probability of Death in Each Level of Care as Compared to Non-Congregate Living Adults



9.1.2 Expected Rate of COVID-19 Death: We used the model to estimate the expected rate of dying from COVID-19 if they lived in each LOC across all states and in each state individually. We used the LSMEANS statement to estimate the mean death rate and 95% confidence interval assuming the average COVID-19 infection rate seen across all counties included in the analysis (6,496). We estimated results across all observations (i.e., the averaged effect across all states) and within each state (the estimated rate for each LOC in that state). The overall results adjust for differences in state reporting and differences in epidemic impact across states. The model also accounts for correlation of observation outcomes nested within each unique county and adjusts the confidence intervals to reflect this. We present these estimates in rates per 1,000 residents, which involved uniformly multiplying model estimates by the constant 1,000. As a reminder from earlier sections, Connecticut had no data from which to estimate deaths among non-congregate living adults aged 75 and older.

Exhibit 12. Expected Deaths per 1,000 Residents by Level of Care and State



Mean estimates across states demonstrate a consistent pattern of increased probability of death given an increased intensity of LOC. Across all states, CN and IL were statistically equivalent, with both showing lower mortality rates than AL. Residents in AL experienced lower mortality rates than those in either MC or SNF, and mortality rates in MC and SNF were higher in MC and SNF than in other LOC. The one notable exception to this pattern were mortality rates observed in MC in Georgia. Statistical significance of these differences are presented in Section 9.1.3 Pairwise Comparisons.

Exhibits 13 through 18 provide the numeric results for each estimate above.

Exhibit 13 All Areas: Numeric Estimates of Expected Deaths per 1,000 Residents by Level of Care and State

State/LOC	Deaths Per 1,000	Standard Error	Lower 95% CI	Upper 95% CI
ALL				
CN	6.7	0.3	6.0	7.3

IL	5.9	0.9	4.4	7.8
AL	19.3	1.6	16.3	22.8
MC	50.4	8.4	36.3	69.6
SNF	59.6	2.4	55.1	64.5

Exhibit 14 Colorado: Numeric Estimates of Expected Deaths per 1,000 Residents by Level of Care and State

State/LOC	Deaths per 1,000	Standard Error	Lower 95% CI	Upper 95% CI
CO				
CN	6.3	0.5	5.4	7.5
IL	5.2	1.7	2.8	9.7
AL	13.1	3.9	7.3	23.3
MC	68.4	18.3	40.1	114.2
SNF	69.6	7.0	57.1	84.6

Exhibit 15 Connecticut: Numeric Estimates of Expected Deaths per 1,000 Residents by Level of Care and State

State/LOC	Deaths per 1,000	Standard Error	Lower 95% CI	Upper 95% CI
CT				
CN	NA	NA	NA	NA
IL	24.8	7.2	14.0	43.5
AL	36.1	6.7	25.1	51.7
MC	76.9	33.9	31.6	175.3
SNF	57.0	5.3	47.5	68.3

Exhibit 16 Florida: Numeric Estimates of Expected Deaths per 1,000 Residents by Level of Care and State

State/LOC	Deaths per 1,000	Standard Error	Lower 95% CI	Upper 95% CI
FL				
CN	6.5	0.5	5.5	7.6
IL	4.8	0.8	3.5	6.6
AL	19.3	1.8	16.2	23.1
MC	69.7	12.1	49.4	97.4
SNF	62.8	3.3	56.7	69.6

Exhibit 17 Georgia: Numeric Estimates of Expected Deaths per 1,000 Residents by Level of Care and State

State/LOC	Deaths per 1,000	Standard Error	Lower 95% CI	Upper 95% CI
GA				
CN	5.5	0.4	4.7	6.3
IL	2.2	1.1	0.8	6.1
AL	11.5	2.4	7.7	17.2
MC	11.0	5.9	3.8	31.2
SNF	54.8	5.5	45.1	66.6

Exhibit 18 Pennsylvania: Numeric Estimates of Expected Deaths per 1,000 Residents by Level of Care and State

State/LOC	Deaths per 1,000	Standard Error	Lower 95% CI	Upper 95% CI
PA				
CN	8.8	1.3	6.5	11.8
IL	5.1	1.0	3.4	7.5
AL	25.0	2.5	20.5	30.5
MC	76.7	24.7	40.3	141.3
SNF	55.1	3.7	48.2	62.9

9.1.3 Pairwise Comparisons: We used Tukey-Kramer adjustment to test for significant differences in the estimated mean death rate between paired comparisons and present these results in table form and as diffograms. Diffograms graphically present pairwise combinations in the model, with comparisons that are significant colored blue and comparisons that are not significant colored red. Additional information on the interpretation of diffograms is available online.[29] We present results for comparisons of each LOC compared to other LOC across all observations, and then more complex results comparing pairwise differences across states. Exhibit 19 shows that 8 of the 10 possible differences in mortality rates seen between the LOC across all states are statistically significant at the p<.0001 level. The differences between IL residents and non-congregate living adults, and the differences between memory care and skilled nursing residents are not statistically significant.

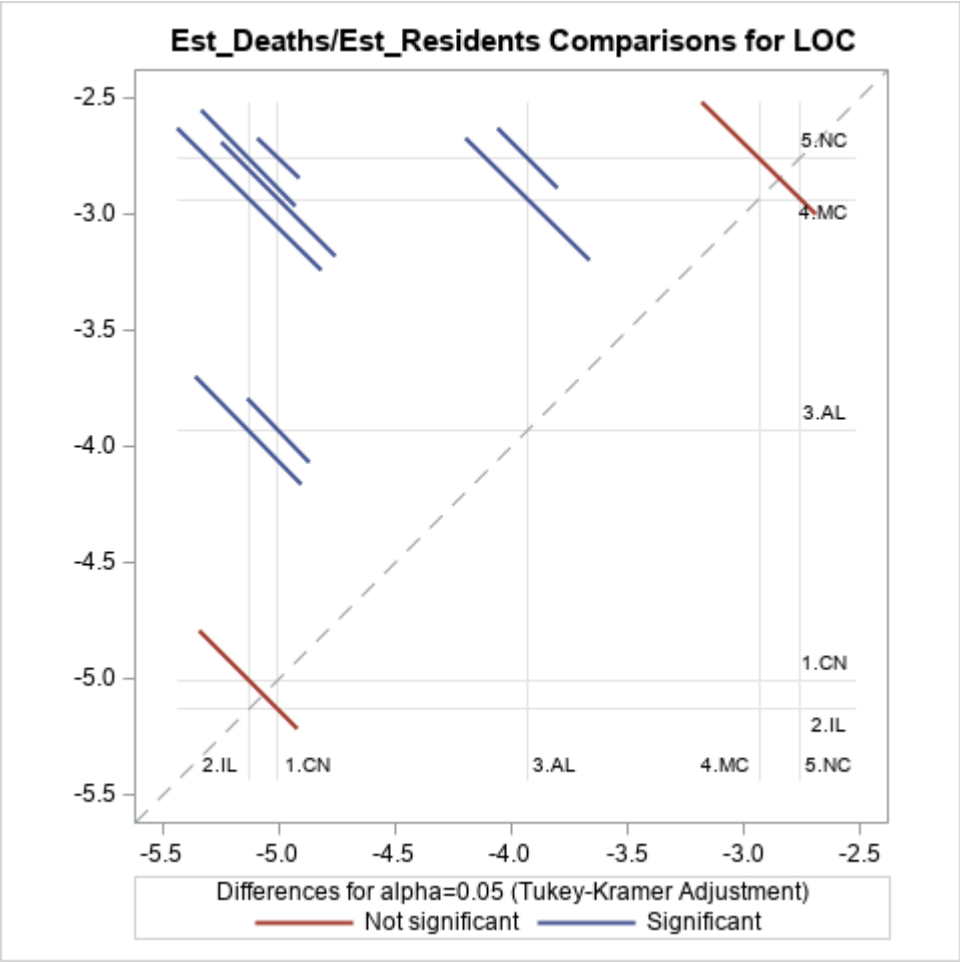
Exhibit 19. Statistical Significance of Differences between Pairwise Combinations of Level of Care

LOC 1	LOC 2	Adjusted P-Value	Significance Level
IL	AL	<.0001	***
IL	MC	<.0001	***
IL	SNF	<.0001	***
IL	CN	0.923	NS
AL	MC	<.0001	***
AL	SNF	<.0001	***
AL	CN	<.0001	***
MC	SNF	0.8585	NS
MC	CN	<.0001	***
SNF	CN	<.0001	***

NOTES: LOC, level of care; CN, non-congregate living adults; IL, independent living; AL, assisted living; MC, memory care; SNF, skilled nursing facility. Adjusted p-value indicates the probability that the difference is attributable to random chance after adjusting for the number of comparisons in the model. A significance value of *** indicates a very high level of confidence that the difference is not attributable to random chance. A value of NS indicates that the difference between the pairs is not statistically significant.

Exhibit 20 presents the diffogram for results of all observations. The figure contains five horizontal and five vertical bars representing each LOC used in the analysis. The intersection of two lines represents the comparison of differences for those two LOCs. For example, the lower left corner of the diffogram represents the comparison of IL and CN. The slanted, red bars are centered on the logit-scaled expected value for the mean for IL. The length of the bar represents the 95% confidence interval of the difference between means, again in logit scale. Difference bars comparing pairs that intersect the center diagonal are not statistically significant and are coded in red, while differences comparing pairs that do not intersect the center diagonal are statistically significant and are coded in blue.

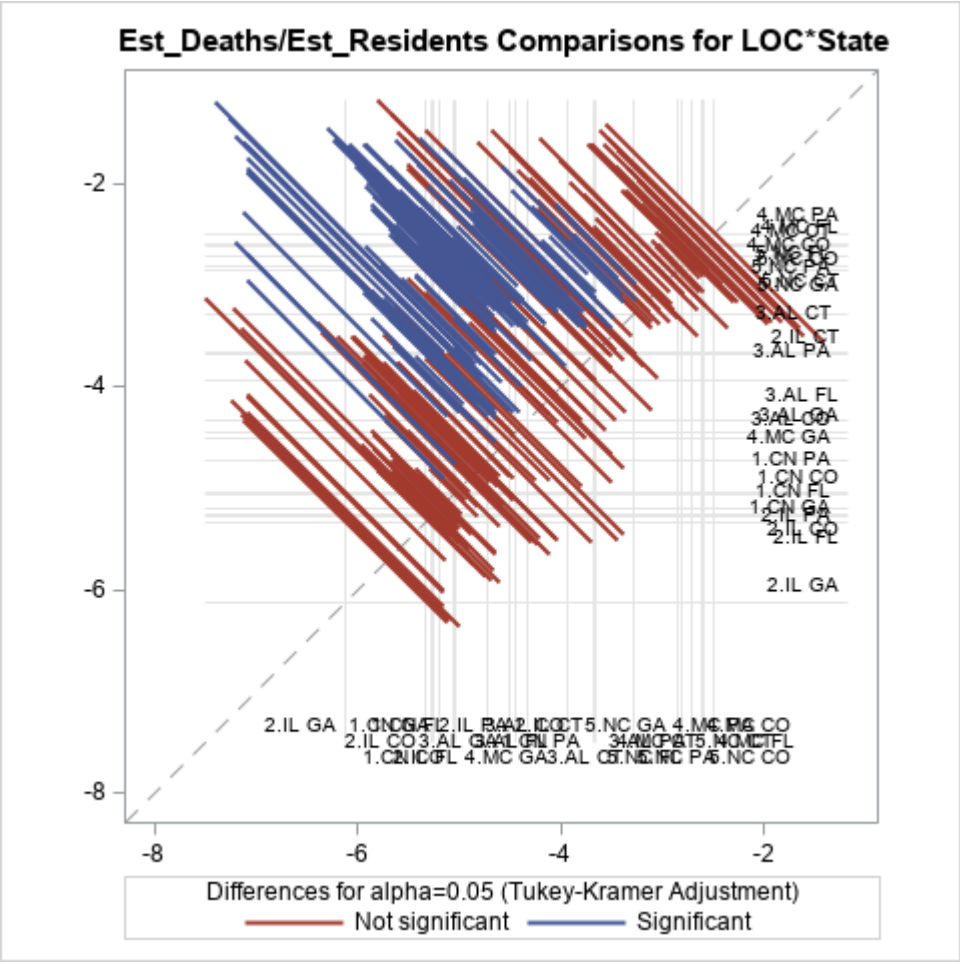
Exhibit 20. Differences of Expected Mean Death Rate Values between Pairwise Comparisons of Level of Care For Overall Results



NOTES: Axes values are modelled logit scale estimates of the Estimated Covid-19 Deaths (Est_Deaths) divided by the estimated residents (EST_Residents) for each LOC used in the analysis.

Exhibit 21 presents the diffogram for all possible comparisons of LOC across states. This demonstrates that in general most comparisons at the state level followed the same pattern of significance as in the overall results; there was no significant difference between CN and IL (left side of the figure) or between MC and SNF (right side of the figure), with primarily significant differences within state results comparing differences between CN and IL to AL, and AL to MC and SNF. There were some notable exceptions to this pattern discussed below.

Exhibit 21. Differences of Expected Mean Death Rate Values between Pairwise Comparisons of Level of Care For Individual State Results



Exhibits 22 through 26 present the Tukey-Kramer Adjusted statistical significance levels comparing LOC within each state. In Colorado, in addition to the model not detecting significant differences between CN and IL, or between MC and SNF, the difference between IL and AL LOC was not significant.

Exhibit 22. Colorado: Statistical Significance of Differences between Pairwise Combinations of Level of Care

Colorado

LOC	State 1	LOC 2	State 2	Adjusted P-Value	Significance Level
IL	CO	AL	CO	0.9027	NS
IL	CO	MC	CO	<.0001	***
IL	CO	SNF	CO	<.0001	***
IL	CO	CN	CO	1	NS
AL	CO	MC	CO	0.008	***

AL	CO	SNF	CO	<.0001	***
AL	CO	CN	CO	0.7587	NS
MC	CO	SNF	CO	1	NS
MC	CO	CN	CO	<.0001	***
SNF	CO	CN	CO	<.0001	***

In Connecticut, the model did not detect any significant statistical differences between any LOC. Caution should be taken in interpreting these results, as the estimated mean values in Connecticut are quite different and data in Connecticut did not include estimates for non-congregate living adults. We explore the impact of excluding data for all non-congregate living adults across all states in our sensitivity analyses.

Exhibit 23. Connecticut: Statistical Significance of Differences between Pairwise Combinations of Level of Care

Connecticut

LOC	State 1	LOC 2	State 2	Adjusted P-Value	Significance Level
IL	CT	AL	CT	1	NS
IL	CT	MC	CT	0.9071	NS
IL	CT	SNF	CT	0.4332	NS
AL	CT	MC	CT	0.9975	NS
AL	CT	SNF	CT	0.8284	NS
MC	CT	SNF	CT	1	NS

Comparisons in Florida mirror those seen across all states. There are no significant differences between CN and IL, or between MC and SNF, but all other differences are significant.

Exhibit 24. Florida: Statistical Significance of Differences between Pairwise Combinations of Level of Care

Florida

LOC	State 1	LOC 2	State 2	Adjusted P-Value	Significance Level
IL	FL	AL	FL	<.0001	***
IL	FL	MC	FL	<.0001	***
IL	FL	SNF	FL	<.0001	***
IL	FL	CN	FL	0.9892	NS
AL	FL	MC	FL	<.0001	***
AL	FL	SNF	FL	<.0001	***
AL	FL	CN	FL	<.0001	***
MC	FL	SNF	FL	1	NS
MC	FL	CN	FL	<.0001	***
SNF	FL	CN	FL	<.0001	***

Reported mortality rates in Georgia AL and MC properties were lower than in other states, and this difference affected the significance tests of pairwise comparisons. In Georgia, mortality rates compared among CN, IL, AL, and MC were statistically equivalent after adjusting for multiple comparisons. Mortality rates in Georgia SNF were significantly higher than in all other LOC, although this difference was only marginally significant at the 10% level when comparing Georgia AL to Georgia SNF.

Exhibit 25. Georgia: Statistical Significance of Differences between Pairwise Combinations of Level of Care

Georgia

LOC	State 1	LOC 2	State 2	Adjusted P-Value	Significance Level
IL	GA	AL	GA	0.32	NS
IL	GA	MC	GA	0.8958	NS
IL	GA	SNF	GA	<.0001	***
IL	GA	CN	GA	0.9893	NS
AL	GA	MC	GA	1.00	NS
AL	GA	SNF	GA	<.0001	NS
AL	GA	CN	GA	0.0956	*
MC	GA	SNF	GA	0.3002	NS
MC	GA	CN	GA	0.9999	NS
SNF	GA	CN	GA	<.0001	***

Pairwise comparisons in Pennsylvania generally followed the pattern of the overall results with the exception that the difference between AL and MC was not significant. As in overall results, CN and IL and differences between MC and SNF were not statistically different.

Exhibit 26. Pennsylvania: Statistical Significance of Differences between Pairwise Combinations of Level of Care

Pennsylvania

LOC	State 1	LOC 2	State 2	Adjusted P-Value	Significance Level
IL	PA	AL	PA	<.0001	***
IL	PA	MC	PA	<.0001	***
IL	PA	SNF	PA	<.0001	***
IL	PA	CN	PA	0.8624	NS
AL	PA	MC	PA	0.164	NS
AL	PA	SNF	PA	<.0001	***
AL	PA	CN	PA	<.0001	***
MC	PA	SNF	PA	1.00	NS
MC	PA	CN	PA	<.0001	***

SNF	PA	CN	PA	<.0001	***

8.2 Sensitivity Analyses: We ran additional models to test the sensitivity of our results to our sample inclusion criteria. Specifically, we were concerned about:

- The limitations of our non-congregate living adults' death data and the impact of including observations from counties without that population in our analyses
- Inclusion of observations from non-congregate living adults
- The effect of including skilled nursing facilities that were not included in the NIC MAP data along with other observations
- Model specification of random effect

We adjusted the inclusion criteria or model structure to address each of these primary concerns and re-estimated the overall results to compare the impacts of our model assumptions. Each alternative scenario tested in the sensitivity analysis is discussed and labelled below.

8.2.1. Scenario 1, Baseline Results: Scenario 1 is our baseline results presented above. This model estimates the average death rate by LOC among all included observations using a model that accounts for differences in LOC at the state level, and that controls for correlations between observations at the county level as part of the residual.

8.2.2. Scenario 2, Omitting Observations from Counties with No Death Data on Non-Congregate Living Adults: Of the 113 counties included in the analysis, only 75 had sufficient data to estimate deaths among non-congregate living adults. The 38 remaining counties included in the baseline analysis have observations of long-term care properties but no corresponding data for non-congregate living adults in the broader community. These include the entire state of Connecticut, five counties in Colorado, and 26 counties in Pennsylvania. We were concerned that including properties in counties without corresponding death data may influence our results. To test for this, we ran a scenario in which we included only properties and non-congregate living adult observations from counties where we had death information on non-congregate living adults.

8.2.3. Scenario 3, Omitting Observations from Non-Congregate Living Adults: We were likewise concerned that our process for estimating deaths among non-congregate living adults might be inaccurate. Additionally, the estimates of non-congregate living residents are much larger than the estimate of residents living in seniors housing properties. For this reason, we ran a scenario in which we omitted all observations of non-congregate living adults (LOC=CN) from the analysis.

8.2.4 Scenario 4, Omitting SNF Properties that Were Not Included in the NIC MAP Dataset: In the baseline analysis we included all properties for which we could develop a consistently measured estimate of residents to pair with the state- or CMS-reported estimate of deaths. This had the effect of excluding all IL, AL, and MC properties that were not in the NIC MAP data, but included SNF properties not listed in the NIC MAP. This was because the CMS estimate of available beds allowed us to estimate their residents in the same way that we estimated residents in SNF for properties listed in the NIC MAP dataset. We were concerned that all properties listed in the NIC MAP might be systematically different from properties not

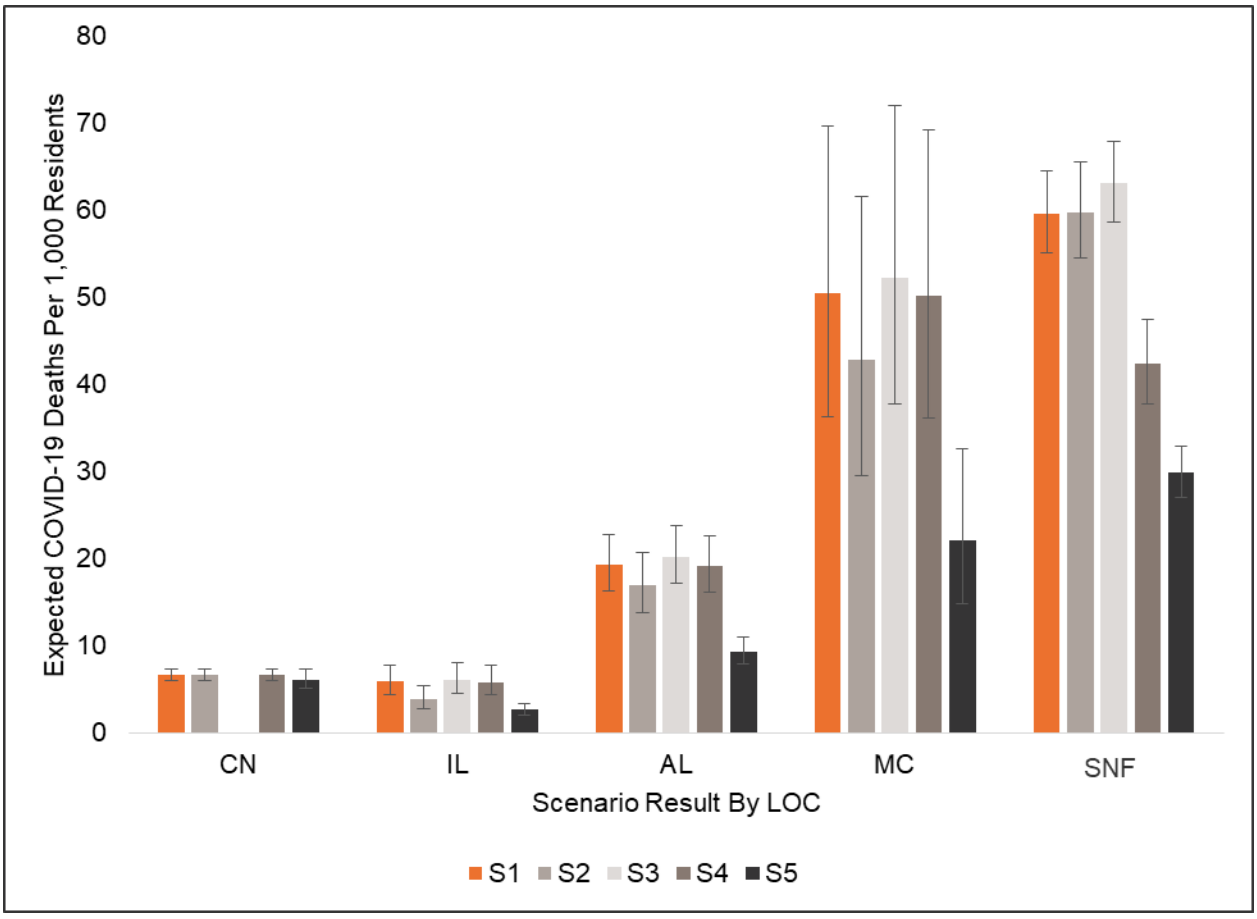
listed. However, we were also concerned about undercounting deaths in seniors housing and for this reason we included SNFs not listed in the NIC MAP data in our baseline analysis. In this scenario, we omit these properties and evaluate the impact of that omission on our results.

8.2.5 Scenario 5, Specification of Random Effect: Following published guidance, we specified an R-sided random effect that provides the population marginal (i.e., population-averaged) value at each LOC, a model that we thought was appropriate for a public health application. An alternative specification is the conditional model in which the random effect is modelled as a G-side random intercept term instead of an R-sided component of the residual. In this scenario, we tested the impact of this model assumption on the overall results.

We estimated results for each scenario using the same model structure. Although the average case rate per 100,000 residents seen across observations changed given inclusion or exclusion of different properties, we estimated the LSMEAN at community case rate of 6,496, the same value used in our baseline result.

Exhibit 27 shows the estimated mean death rate and 95% confidence interval corresponding to each scenario. The figure is organized by LOC to facilitate comparisons among scenario estimates within the same LOC.

Exhibit 27. Death Rate by Sensitivity Scenario and Level of Care across All Included Observations in Each Scenario



Across all scenarios, we observed a consistent rank order of severity of mortality rates across the scenario. However, there are some differences to note. First, in both scenario 2 (where we excluded observations that were not in communities with corresponding death data) and scenario 5 (where we controlled for county-level differences as a G-sided effect), the mean death rate among IL residents was significantly lower than non-congregate living adults at the p<.0001 level. It could be argued that both these scenarios better control for county-level variation, in scenario 2 by removing observations that cannot be directly attributed to non-congregate living adults, and in scenario 5 by controlling for county-level effects at the individual property level instead of in the residual. Estimates from scenario 2 are somewhat lower than the baseline, which resulted primarily from excluding Connecticut IL, AL, MC, and SNF properties from the analysis.

Finally, scenario 5 yields similar rank order results as the baseline model, but results in lower estimated mortality rates in IL, AL, MC, and SNF. Scenario 5 is a different functional form of the model than the other scenarios and its interpretation is not identical. The baseline interpretation of the baseline model is that the results estimate the average population death rate at each LOC accounting for differences in the county case rate and state-level differences, and adjusting the standard errors of the estimate for county clustering effect. We chose the marginal effect model used in our baseline scenario, because modelling clustering an R-sided effect is recommended for estimating population average values, and we were seeking to estimate the average death rate associated with LOC across all included observations, as opposed to the incremental impact within each county cluster.[30,31] Scenario 1 provides a stronger estimate of the mean death rate by LOC seen in these data and their confidence intervals. It could be argued that scenario 5 provides a better estimate of the incremental impact of LOC within each county cluster. While the interpretation of scenario 5 is different than for the baseline model, we include it to demonstrate that the relationship of LOC to severity of outcome is consistent even when using the conditional model.

9.0 Supplemental Analyses: In addition to analyzing deaths by LOC, we conducted supplemental analyses on two areas of interest. The first looked at observed differences in reported COVID-19 cases by LOC. The second looked at the impact (if any) of a property being a CCRC on COVID-19 deaths. Additional methods and results for these analyses are provided below.

9.1 Impact of Continuing Care Retirement Community Status on Expected Death Rate: CCRCs, also known as life plan communities, offer a range of senior care services that can be modified as a patient ages to meet their health care needs. These communities often offer a range of possible levels of care, from independent living through skilled nursing services housed within the same campus. For this study, the NIC MAP property inventory data provided information about the LOC utilized by the majority of residents living in a property, additional information on the number of units at each LOC, and information regarding whether the property was classified as a CCRC. CCRCs may have experienced differences in their COVID-19 mortality outcomes compared to other properties based on their mixed LOC, because their generally larger size better equipped them to manage the pandemic, or because they were better able to segregate patients in different LOC. In this secondary analysis, we modified the marginal model described in Section 9.0 to include a dichotomous measure of CCRC status and an interaction between CCRC and state. We removed the interaction term between LOC and

state to avoid over-specification. In this model, the ratio of deaths over residents is equal to a function of LOC, CCRC, the interaction between CCRC and state, and the county COVID-19 case rate. Clustering effects at the county level were managed as an R-sided effect.

We then estimated the mean expected mortality rates for CCRCs overall and within each state observations. Because there are no CCRCs that are defined as have a majority of residents who were community dwelling adult or MC patients, we dropped these observations in LOC CN and MC from the analysis. Thus, our comparisons of CCRC to non-CCRC can be interpreted as the effect of a property being a CCRC versus not being a CCRC among the pooled group of IL, AL, and SNF properties, controlling for the overall effect of LOC, and controlling for state differences in the impact of CCRCs and differences in the county COVID-19 case rate. We compared the statistical significance of pairwise differences between CCRC and non-CCRC properties overall and in each state using Tukey-Kramer-adjusted confidence intervals.

9.1.1 Results: Evaluating CCRCs required us to limit our sample to properties observed in the NIC MAP property inventory and non-congregate living adults. We identified 2,827 properties in the NIC MAP data that met the criteria of having met previous inclusion criteria, additionally being represented in the NIC MAP data, and not being in CN or MC observations. Of these, 277 (9.8%) were classified as CCRCs. Of CCRC properties, 74.0% were classified as majority IL, 4.3% were classified as majority AL, and 21.6% were classified as majority SNF. Additionally, 6.9% of CRCC properties were located in Colorado, 7.6% in Connecticut, 33.9% in Florida, 4.0% in Georgia, and 47.7% in Pennsylvania. From the model, type-III fixed effects were significant for LOC ($p<.0001$), CCRC ($p<.01$), and CCRC interacted with state ($p<.05$). County COVID-19 case rate was not significant ($p=.17$).

Across all observations, a property classified as a CCRC was significantly associated with a lower expected mean death rate. Across states, this effect is most observable in Colorado, Florida, and Georgia. When not adjusted for multiple comparisons, the effect is significant across all states ($p<.01$), in Colorado ($p<.05$), and Florida ($p<.0001$). After adjusting for multiple comparisons, the effect is significant overall ($p<.0$) and in Florida ($p<.01$). The mean expected mortality rates in CCRC properties across all areas was 10.02 per 1,000 (95% CI, 6.6–15.1) as compared to 19.9 in non-CCRC properties (95% CI, 17.9–22.3).

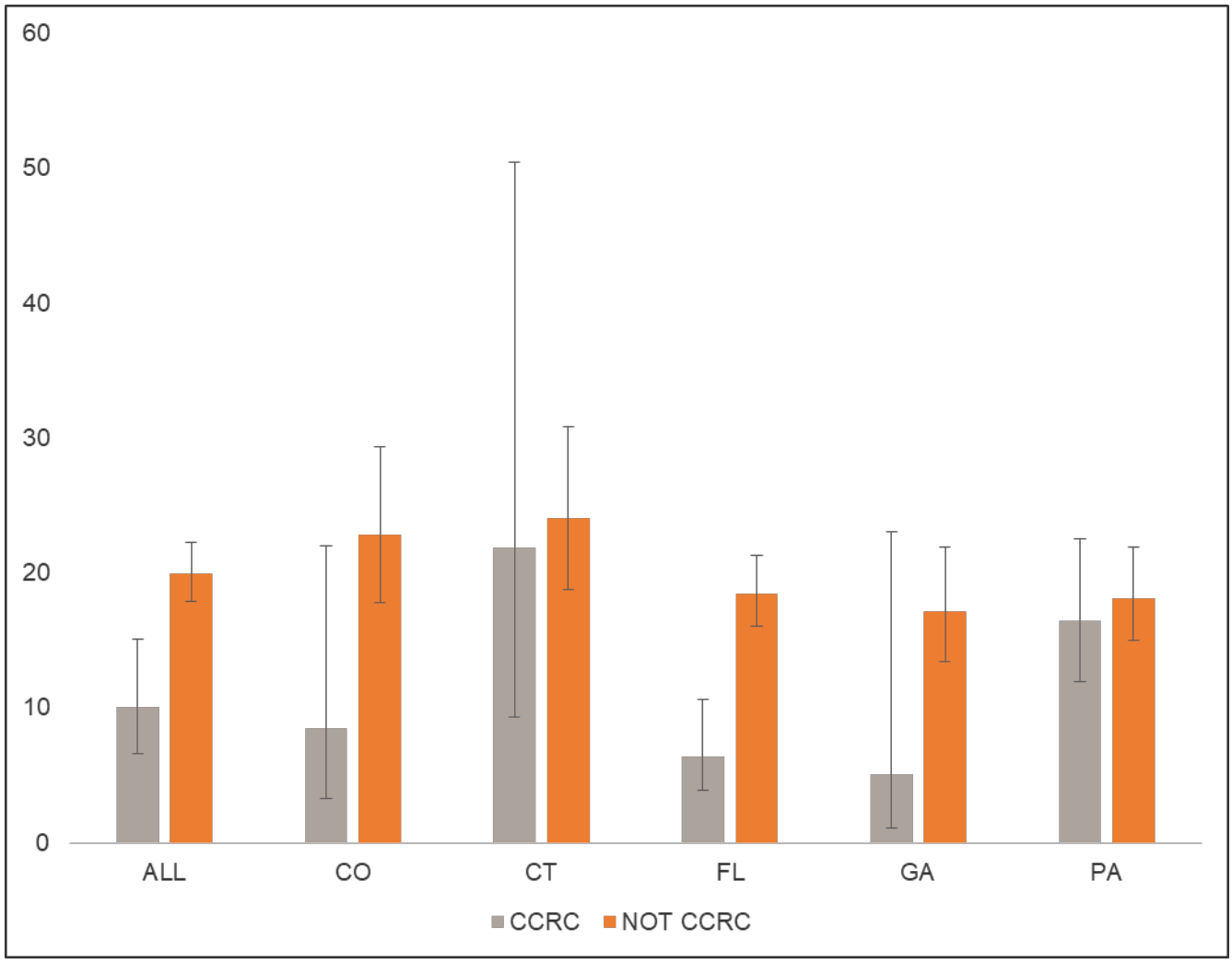
Exhibit 28. Statistical Significance of Pairwise Comparisons between CCRC and Non-CCRC Properties Overall and in Each State

State	Comparison	P-Value	Adjusted P-Value	Significance Level
CO	CCRC TO NON	0.0491	0.6218	NS
CT	CCRC TO NON	0.8245	1	NS
FL	CCRC TO NON	<.0001	0.0028	***
GA	CCRC TO NON	0.1202	0.8701	NS
PA	CCRC TO NON	0.6119	1	NS

ALL	CCRC TO NON	0.0011	0.0011	***

Note: NS, Not Significant, ***, Significant at the alpha = .01 level.

Exhibit 29. Estimated COVID-19 Mortality Rates per 1,000 Residents among Residents Living in CCRC and Non-CCRC Properties across States and Overall



10.0 Limitations

This study is affected by several limitations that are related to data quality and decisions made to use the data for the purpose of analysis. Readers should be aware of these limitations in their interpretations of these results.

10.1 Sample Limitations: This study represents a convenience sample of properties from which data were available. As of December 2020, when this study began, only five U.S. states provided data necessary to support the analysis. The five included Colorado, Connecticut, Florida, Georgia, and Pennsylvania, and provided some level of variation in their geographic region, demographics, timing of pandemic introduction, and political leadership. However, they are not representative of the U.S. population.

The sample was further limited by the exclusion of areas that are not monitored and contribute no data to the NIC MAP property inventory. The counties included in the analysis encompassed 63.3% of the population of Colorado; 95.2% of the population of Connecticut; 84.1% of the population of Florida; 60.0% of the population of Georgia; and 79.9% of the population of Pennsylvania. However, included counties tended to be more urbanized/less rural, meaning that this analysis has poor representation of rural and exurban areas.

Within areas selected for inclusion, we were only able to include properties where we had information from the NIC MAP or the CMS COVID-19 dataset that we could use to estimate the number of residents living in a property as of the first quarter of 2020. This had the effect of excluding IL, AL, and MC properties that were found on state lists but were not also contained in the NIC MAP property inventory. The NIC MAP tracks properties with more than 25 residents that it considers to be market grade. While market grade is not defined in the NIC MAP glossary, it is typically understood to include properties that would be suitable for investment, excluding all state-run and most state-subsidized properties. Additionally, CMS COVID-19 data did not contain information for properties with less than 25 residents. As such, this study provides no empirical information with which to draw conclusions about properties with less than 25 residents. Further, it includes no market rate properties for LOC other than SNF.

The sample was further limited by varying state reporting requirements for different LOC. While some states reported deaths that occurred at IL properties, it was not a requirement for every state. Since the sample used for this analysis reported properties by the LOC used by the majority of its residents, many of the IL-related deaths were captured through this determination.

We examined the number of properties in the dataset that include IL only without any other LOC. In total, 130 number of properties (26.9% of all IL properties) served only IL residents with no other LOC. Of these, 7.7% reported at least some deaths to state health agencies. Our analyses indicate that the mean average estimated mortality rates in IL are lower than among non-congregate living adults. This result is not statistically significant in our baseline analysis but is significant in some sensitivity analyses. We would urge caution in interpreting these results and would argue that mortality rates among those living in IL settings were most likely indistinguishable from non-congregate living adults. However, data limitations preclude us from stating this with certainty.

Sensitivity analysis—scenario 4—evaluated the impact of excluding SNF properties that were not included in the NIC MAP data in the analysis. It demonstrated that SNF properties with records in the NIC MAP experienced lower mortality rates than those that were only found in the CMS COVID-19 data. This finding should be considered when considering the mortality rates presented for IL, AL, and MC. If we had been able to include non-market grade properties, the mortality rates for these categories of seniors housing would likely have been higher.

For this study, we sought to create a reasonable convenience sample of properties for analysis given the limitations in available data. The results of this study show a fairly consistent and strong relationship between increased LOC and risk of death within the sample. However, our ability to infer that relationship to properties outside the sample is limited by the factors above.

10.2 Deaths from Other Causes: Our analysis estimates residents in each property as of the first quarter of 2020 and then estimates mortality rates based on the assumption that these individuals remained in the same property over the course of 2020. This assumes that no residents left a property for another with another care level designation and that none died for reasons other than COVID-19 over the course of 2020. Both assumptions are almost certainly false. A recent study estimated that in years prior to the pandemic, as many as one-third of all residents of skilled nursing die over the course of one year.[32] Additionally, people may move from one facility to another, especially from lower-intensity LOC to higher-intensity LOC due to deteriorating health due to aging. If people died or switched LOC at the same rate across LOC type and new residents took their place, then the true number of residents exposed at each LOC would be higher than our estimate, and the death rate at each LOC would be lower, but the relationships comparing LOC would be the same. However, if people in higher LOC died at a faster rate than residents at lower LOC, then the effect of attrition might influence the relationship between LOC that we observed. To test the impact of potentially differential mortality rates across property type, we ran a model in which we differentially increased the number of residents at each LOC to account for higher mortality and replacement in higher LOC. Specifically, we inflated the number of IL residents by 5%, the number of AL residents by 15%, and the number of MC and SNF residents by 35%, and reran the baseline model under this assumption. We did not inflate the number of non-congregate living adults (CN). In this model, the estimated mortality rates per 1,000 residents were 6.6 (95% CI, 6.1–7.2) for CN; 5.6 (95% CI, 4.0–7.7) for IL; 16.8 (95% CI, 13.7–20.5) for AL; 37.2 (95% CI, 23.9–57.4) for MC; and 44.1 (95% CI, 41.3–47.1) for SNF. As in the baseline model, the differences between LOC were all significant at the $p<.01$ level except the difference between CN and IL and the difference between MC and SNF, which were not significantly different. Although the actual rate of mortality from other causes in the selected properties used in the analysis are unknown, these results show that differential mortality from other causes and higher resident turnover are not the drivers of differences observed between LOC.

10.3 Limitations in Estimating Deaths among Non-Congregate Living Adults: We estimated deaths among non-congregate living adults by combining data from state- and county-reported deaths among all people with other state reports of property level deaths. We assumed that all deaths in the seniors housing property reports were also included in the state reports of all deaths, but we were not able to verify this. We assumed that we could reasonably estimate the number of non-congregate living adults aged 75 and older who died in each county by subtracting the number of people who died in seniors housing in that county from the

estimated total number of people aged 75 and older who died in that county. There are several reasons why this may be inaccurate:

- Reporting deaths is likely more complete in properties that are required to report COVID-19 outcomes to the state than reporting for non-congregate living adults. If this were the case, the estimated death rate for non-congregate living adults is too low.
- Most states and CMS did not report deaths from properties with less than 25 residents. Because of this, deaths that occurred in these properties are included in the death rate for non-congregate living adults. This resulted in a higher death rate among non-congregate living adults than if we had been able to account for them. Thus, the death rate for non-congregate living adults should be interpreted as that associated with non-congregate living adults and those living in seniors housing with less than 25 residents.
- There were a number of properties on state lists for which we were unable to estimate the number of residents. We were unable to include these in our analysis. We did include their deaths in our county totals of deaths that occurred in properties. However, we were unable to subtract their residents from the broader county population total, as we had no estimate of their residents. This had the effect of slightly increasing the estimated non-congregate living adults, which in turn resulted in a slightly lower death counts than what we would have obtained if we had been able to account for these residents. This effect is likely to be small given the relatively large size of most non-congregate living adult populations.

10.4 Omitted Variables: Because of the nature of the data (pooled residents from state reports) outside of LOC, we were unable to adjust for differences in chronic conditions and frailty across LOC. Additionally, the marginal model that we used estimated the average death rate seen across properties adjusting for community-reported COVID-19 case rates. LOC itself is a proxy measure of health status and our results are strongly suggestive that mortality rates were not uniform across LOC. However, our results as presented here cannot tell us if a person with the same underlying health risk would have been more at risk for death in any particular LOC than any other. Additional explanatory variables measuring poverty, race/ethnicity, and other measures likely explain some of the variance in outcomes seen in our baseline model. Additional analyses with these data, potentially linking with data on socioeconomic status from the American Community Survey, is warranted.

10.5 Data Quality and Data Transformation Assumptions: Data for this study were generated in different ways from different reporting organizations and using different assumptions. Inconsistencies in state reporting are captured in our modelled result with the interaction of LOC and state and our statistical model adjusts the confidence intervals for each mean value based on correlations between observations within counties in the same state. However, county-level differences in reporting are not captured in the mean results, but only in the confidence intervals.

In our observations of death data reported for properties that appeared on both state and CMS lists, the total number of deaths reported across sources often disagreed. In instances where the state list contained more deaths than the CMS list, this is easily explained by the fact that

many properties include residents in multiple LOC (for example, an AL property with an SNF unit). In these properties, deaths that occurred in SNF would be reportable to CMS and deaths that occurred AL would not. We and others have also observed instances where the state-reported total for a property was less than the total reported to CMS.[33] We have no explanation for this. In this analysis, if the property appeared on both a state list and in the CMS data, we were able to correct for any under-reporting by taking the maximum value seen across both lists. However, we had no method or data to correct for under-reporting for IL, AL, and MC properties that appeared only on state lists, or for any under-reporting for any properties that only appeared in the CMS data.

We estimated residents using NIC MAP and CMS data on property capacity and property occupancy. This is not equivalent to a census of residents, and (as stated earlier) residents may have died or left the property following the first quarter of 2020. However, we consistently applied the estimation method across LOC so errors in the estimates beyond differential mortality would not likely lead to differences in relative differences in estimated mortality rates from COVID-19.

All COVID-19-related data used in this analysis were generated during a pandemic. Methods and requirements for reporting differed in each state and over time. Data are subject to errors from both under-reporting, duplicate reporting, incomplete reporting, and inaccurate reporting. Additionally, death data used in this analysis do not include cases yet to be reported as of the date the data were accessed.

The data in this analysis were developed to provide information on the impact of the pandemic during calendar year 2020. Significant surges in COVID-19 infections that occurred nationwide in January and February 2021 are not reflected in these data, and including those data may lead to different results.

Finally, because of the difficulty of compiling data from multiple sources, this analysis provides cumulative information starting at the beginning of the pandemic through the end of 2020. This period encompasses a dynamic time during which the nation's understanding of the pandemic and its capacity to manage it changed dramatically. At the start of April 2020, there was still uncertainty regarding whether masking was necessary, whether COVID-19 was transmitted by air, and whether asymptomatic infection was possible. Medical treatment of COVID-19 was in its earliest stages, and testing and personal protective equipment was in critically short supply. By the end of 2020, our scientific knowledge of COVID-19 had grown enormously and many seniors housing operators had made substantial changes to their practices, which likely led to large declines in mortality relative to earlier in the pandemic. In Connecticut's data, we were able to observe differences in mortality rates before and after July 15, 2020, and these reflect large declines in mortality rates in the later quarter as compared to early in the pandemic.

Glossary

Activities of Daily Living (ADLs): Activities related to the care of and moving of the body, including: walking, bathing, dressing, toileting, transferring, and eating.

Assisted Living (AL): Properties where assisted living units comprise the largest share of inventory. Residents receive personal care services such as assistance with bathing, dressing, eating, walking, and toileting. Twenty-four-hour protective oversight is provided, but 24-hour medical care is not.

Congregate Living Settings: Residences that are seniors housing (independent living, assisted living, memory care, nursing care, or CCRC), where a group of people reside, meet, or gather in close proximity for an extended period of time.

Continuing Care Retirement Community (also referred to as CCRC or Life Plan Community): Age-restricted properties that offer at least independent living and nursing care, and may include a full continuum of care, including assisted living, memory care, and other supportive services to residents generally all on one campus.

Independent Living (IL): Properties where independent living units comprise the largest share of inventory. Independent living properties typically include services such as communal dining, housekeeping, transportation, emergency call, and social programming services in the monthly fee.

Long-Term Care: In this study, long-term care comprises the care settings for seniors of independent living, assisted living, memory care, and skilled nursing. It is often used generically and interchangeably with seniors housing. In this study, we use the term “seniors housing.”

Market Rate: Market rate seniors housing does not include affordable housing options that may be subsidized or rate capped.

Memory Care (MC): Properties where memory care units comprise the largest share of inventory are considered a specialty type of assisted living, serving residents with Alzheimer’s disease and other dementias, and can include standalone properties as well as properties with a minority of other care settings.

Non-Congregate Dwelling Adults: Adults living in residences (in the broader county) that are not seniors housing (independent living, assisted living, memory care, nursing care, or CCRC).

Non-Congregate Living Settings: Residences (in the broader county) that are not seniors housing (independent living, assisted living, memory care, nursing care, or CCRC). Examples include single-family homes, apartments, and condominiums.

Non-Continuing Care Retirement Community (Non-CCRC): Any combination of care settings that does not include both independent living and nursing care on the same campus.

Skilled Nursing Facility (SNF): For purposes of this analysis, properties where skilled nursing or nursing home beds comprise the largest share of inventory. A SNF is generally a licensed residential property that serves people who require 24-hour nursing and/or medical care and are often considered the frailest elderly. In this study, skilled nursing generally encompasses locations that provide post-hospital skilled nursing services to seniors. This term collectively includes skilled nursing properties, skilled nursing facilities, and nursing homes.

Seniors Housing: In this study, seniors housing includes properties with independent living, assisted living, memory care, and/or skilled nursing settings. Seniors housing encompasses a wide range of congregate care settings for seniors and includes nursing care, independent living, assisted living, and memory care.

1. Centers for Medicare & Medicaid Services. (2012). *COVID-19 nursing home dataset*. https://data.cms.gov/Special-Programs-Initiatives-COVID-19-Nursing-Home/COVID-19-Nursing-Home-Dataset/s2uc-8wxp.
2. The National Investment Centers for Seniors Housing & Care. (2021). *The NIC mission*. https://www.nic.org/our-mission/.
3. The National Investment Centers for Seniors Housing & Care. (2021). *The NICMAP data service*. https://www.nic.org/senior-housing-investment-nic-map.
4. Colorado State Emergency Operations Center. (2021). *Outbreak data*. https://covid19.colorado.gov/covid19-outbreak-data.
5. Connecticut Department of Health and Human Services. (2021). *Assisted living facilities with residents positive for COVID-19*. https://data.ct.gov/Health-and-Human-Services/Assisted-Living-Facilities-with-Residents-Positive/wjua-euxh.
6. Florida Department of Health. (2021). *Long-term care facility deaths associated with COVID-19 report*. http://ww11.doh.state.fl.us/comm/_partners/covid19_report_archive/long-term-care-facilities/ltcf-deaths-reports/ltcf_deaths_20210115.pdf.
7. Georgia Department of Community Health. (2021). *Long-term care facility COVID-19 report*. https://dch.georgia.gov/announcement/2021-04-01/long-term-care-facility-covid-19-report.
8. Division, G.C.P. (2021). *Personal care homes*. https://consumer.georgia.gov/consumer-topics/personal-care-homes#:~:text=A%20%E2%80%9Cpersonal%20care%20home%E2%80%9D%20is,as%20medical%20or%20nursing%20services.
9. Pennsylvania Department of Health. (2021). *COVID-19 long-term care facilities data for Pennsylvania*. https://www.health.pa.gov/topics/disease/coronavirus/Pages/LTCF-Data.aspx.
10. Pennsylvania Department of Human Services. (2021). *Personal care home information*. https://www.dhs.pa.gov/Services/Disabilities-Aging/Pages/Personal-Care-Homes.aspx.
11. U.S. Census Bureau. (2021). *2015-2019 American community survey 5-year estimates*. https://data.census.gov/cedsci/.
12. Coronavirus in the U.S.: Latest map and case count. (2021, April 5). *The New York Times*. https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html?action=click&module=Top%20Stories&pgtype=Homepage#states.
13. The SAS Institute. (2021). *SAS Viya programming documention, COMPGED function*. https://documentation.sas.com/?cdcId=pgmsascdc&cdcVersion=v_006&docsetId=lefunctionsref&docsetTarget=p1r4l9jwgatggtn1ko81fyjys4s7.htm&locale=en.
14. Sloan, S., & Lafler, K. P. (2018). *Fuzzy matching programming techniques using SAS® Software* https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2886-2018.pdf.
15. ATI Advisory. (2020). *2020 Seniors housing data book*. https://atiadvisory.com/wp-content/uploads/2020/05/ATI-Advisory-Seniors-Housing-Data-Book_May-2020.pdf.
16. COVID Tracking Project. (2021). *Connecticut: Long-term care historical data*. https://covidtracking.com/data/state/connecticut/long-term-care/history.
17. Florida Department of Health Open Data. (2021). *Florida COVID19 case line data*. https://open-fdoh.hub.arcgis.com/datasets/florida-covid19-case-line-data-3/data.
18. Georgia Department of Public Health. (2021). *Georgia department of public health daily status report: Download the data (CSV)*. https://dph.georgia.gov/covid-19-daily-status-report.
19. Tri-County Health Department Serving Adams, Arapahoe and Douglas Counties. (2021). *COVID-19 case reporting*. https://data.tchd.org/covid19/.
20. Boulder County. (2021). *COVID-19 illness data*. https://www.bouldercounty.org/families/disease/covid-19/covid-19-illness-data/.
21. Denver COVID-19 Data Summary. (2021). https://storymaps.arcgis.com/stories/50dbb5e7dfb6495292b71b7d8df56d0a.
22. El Paso County Public Health Department. (2021). *Public health data and analytics, COVID-19 deaths (panel 6 of 15)*. https://www.elpasocountyhealth.org/covid19data-dashboard.

23. City of Philadelphia. (2021). *Coronavirus disease 2019 (COVID-19) (death tab)*. https://www.phila.gov/programs/coronavirus-disease-2019-covid-19/data/testing/#/.
24. Lancaster County. (2021). *Lancaster county COVID-19*. https://covid-19-lancastercountypa.hub.arcgis.com/.
25. Montgomery County. (2021). *ARCGIS age and race distribution of deaths*. 2021https://insights.arcgis.com/index.html#/embed/b3ece1810a30426dac0224b2cba0f5d0.
26. Allegheny County Health Department. (2021). *COVID-19*. https://www.alleghenycounty.us/health-department/resources/covid-19/covid-19.aspx.
27. Schabenberger, O. (2005). *Introducing the GLIMMIX procedure for generalized linear mixed models, paper 196-30*. https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/196-30.pdf.
28. Kiernan, K. (2018). *Insights into using the GLIMMIX procedure to model categorical outcomes with random effects, paper SAS2179-2018*. https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/2179-2018.pdf.
29. Wicklin, R. (2017). *The diffogram and other graphs for multiple comparisons of means*. https://blogs.sas.com/content/iml/2017/10/18/diffogram-multiple-comparisons-sas.html.
30. Muff, S., Held, L., & Keller, L. F. (2016). Marginal or conditional regression models for correlated non-normal data? *Methods in Ecology and Evolution, **7***(12), 1514-24.
31. Carriere, I., & Bouyer, J. (2002). Choosing marginal or random-effects models for longitudinal binary responses: Application to self-reported disability among older persons. *BMC Medical Research Methodology, **2***, 15.
32. Vossius, C., et al. (2018). Mortality in nursing home residents: A longitudinal study over three years*. PLoS One*, ***13***(9), e0203480.
33. Teegardin, C. (2021, April 5). Warnock, Johnson seek investigation of Georgia's nursing home death data. *Atlanta Journal Constitution*. https://www.ajc.com/news/investigations/warnock-johnson-seek-investigation-of-georgias-nursing-home-death-data/C6AYC3DFGJENZE6LTA3T4OCVEE/.