A risk index for COVID-19 severity is associated with COVID-19 mortality in New York City

Wil Lieberman-Cribbin¹, Naomi Alpert¹, Raja Flores² and Emanuela Taioli¹,²*

Abstract

Background: New York City (NYC) was the epicenter of the COVID-19 pandemic, and is home to underserved populations with higher prevalence of chronic conditions that put them in danger of more serious infection. Little is known about how the presence of chronic risk factors correlates with mortality at the population level. Here we determine the relationship between these factors and COVID-19 mortality in NYC.

Methods: A cross-sectional study of mortality data obtained from the NYC Coronavirus data repository (03/02/2020–07/06/2020) and the prevalence of neighborhood-level risk factors for COVID-19 severity was performed. A risk index was created based on the CDC criteria for risk of severe illness and complications from COVID-19, and stepwise linear regression was implemented to predict the COVID-19 mortality rate across NYC zip code tabulation areas (ZCTAs) utilizing the risk index, median age, socioeconomic status index, and the racial and Hispanic composition at the ZCTA-level as predictors.

Results: The COVID-19 death rate per 100,000 persons significantly decreased with the increasing proportion of white residents (βadj = −0.91, SE = 0.31, p = 0.0037), while the increasing proportion of Hispanic residents (βadj = 0.90, SE = 0.38, p = 0.0200), median age (βadj = 3.45, SE = 1.74, p = 0.0489), and COVID-19 severity risk index (βadj = 5.84, SE = 0.82, p < 0.001) were statistically significantly positively associated with death rates.

Conclusions: Disparities in COVID-19 mortality exist across NYC and these vulnerable areas require increased attention, including repeated and widespread testing, to minimize the threat of serious illness and mortality.

Keywords: Mortality, Coronavirus, Comorbidities

Background

New York City was once considered the epicenter of the coronavirus disease (COVID-19) epidemic in the United States. From the first reported case on March 1, 2020, to April 2020, nearly 600 people were dying each day, new daily cases numbered in the thousands, and new hospitalizations were roughly 1500 per day [1]. The mortality count peaked on April 7, where there were 598 reported deaths, 6042 cases and 1578 people hospitalized. Reflecting on this time has highlighted racial and socioeconomic disparities in COVID-19 testing patterns that are indicative of gaps in initial COVID-19 response policy [2–5]. Analysis of COVID-19 case identification and mortality rates in New York City has also underscored variations in the amount of COVID-19 testing performed over time, delays in receiving testing results, and lags in the reporting of death data [6, 7]. To provide a more complete picture of mortality data, the New York City Department of Health and Mental Hygiene
COVID-19 Response Team issued a publication on the topic and reported that over 5000 excess deaths occurred from the COVID-19 pandemic from March 11th through May 2nd [8]. Within NYC, cumulative mortality rates are highest in Hispanic/Latino and African American/Black populations, as well as older age groups [1].

Clinical studies across the US and NYC have shown that people with underlying health conditions and risk factors, including diabetes, chronic lung disease, and cardiovascular disease, are at higher risk for severe outcomes and mortality from COVID-19 [1, 9–11]. However, little is known about how the presence of these chronic risk factors correlates with mortality at the population level [12–14]. Such information is key to developing effective public health interventions to mitigate COVID-19 spread and adverse outcomes, and population studies can guide policy at the community-level. In addition, population level studies are enriched with minorites and the whole spectrum of socioeconomic status, thus reflecting reality more closely. A study of US data looked at the prevalence of risk factors for COVID-19 severity according to CDC guidelines, and asserted that people living in low-income households and those who are Black or American Indian are more likely to have chronic conditions that place them in a higher risk category for COVID-19 complications [12]. The study however did not correlate this high-risk profile with mortality data. Within NYC, the distribution of behavioral risk factors and chronic conditions is highly variable from neighborhood to neighborhood; higher obesity rates in Brooklyn and the Bronx have been posited as being partially responsible for the increased COVID-19 mortality in these areas [15]. A separate study reported older age and severe obesity as risk factors for COVID-19 mortality and worse in-hospital outcomes in a Bronx cohort [16]. However, the research published so far only partially tackles a more complex reality when studying COVID-19. Minorities and underserved populations have higher prevalence of chronic conditions that put them in danger of a more serious infection; at the same time, they also have higher risk of exposure to the virus, as they frequently live in neighborhoods with higher population density and proportions of essential workers, with fewer opportunities to exercise social distancing [17–19]. Here we study risk factors associated with COVID-19 mortality at the neighborhood level in NYC. We hypothesized that areas with the highest prevalence of risk factors, minority residents and lower SES experienced the highest mortality in NYC.

**Methods**

**Risk index**

Risk factors were identified based on the CDC criteria for risk of severe illness and complications from COVID-19 [20]. Census tract level crude prevalence (%) of diabetes, asthma, chronic obstructive pulmonary disease (COPD)/emphysema/chronic bronchitis, cancer (besides skin), angina/coronary heart disease, chronic kidney disease, obesity, and hypertension were extracted from the 500 Cities Project [21], which sourced this data from the 2017 Behavioral Risk Factor Surveillance System (BRFSS) [22]. The New York City Department of Health and Mental Hygiene online tool EpiQuery [23] was used to extract United Hospital Fund (UHF) neighborhood level information on prevalence of hepatitis B and C from 2017 Communicable Disease Surveillance Data. ZIP code level rates, based on the population of each zip code, of alcohol related hospitalizations in 2017 were extracted from New York State’s hospital discharge database [24] based on the following ICD-10 diagnosis codes: F10, G62.1, I42.6, K29.20, K29.21, K70.10, K70.11, K70.30, K70.2, K70.31, K70.4, K70.9, K73.0-K73.9, K74.0, K74.2, K76.9, K76.6, K86.0, K86.1 (diagnoses of alcohol-related morbidity), R78.0, and T51 (alcohol poisoning). Hepatitis B, C, and alcohol related hospitalizations were used as proxies for liver disease. Epiquery was also used to assess 2015 birth rates by UHF, as a proxy for pregnancy, while heart attack hospitalizations by UHF in 2016 were extracted from New York City Environment and Health Data Portal [25]. The percent of the population 65 years and older was extracted from American Community Survey 2018 5-year estimates. The median age, proportion of white, and proportion of Hispanic residents were also extracted from 2018 census data. Median age was extracted despite the inclusion of age as a risk factor in order to adjust for different age distributions across NYC, as the prevalence of risk factors were not age adjusted. Data for all indicators were converted to ZIP code, using crosswalks provided by the census [26] and the New York City Department of Health (NYC DOH)[27]. A summary table of the risk factor measures and their citywide prevalence is included in Supplementary Table 1.

Each risk factor was then divided into quartiles and assigned a score of 1 to 4, with the lowest quartile (score 1) representing areas with the lowest severe COVID-19 risk associated with a given variable, and the highest quartile (score 4) representing the highest risk. Risk quartiles for Hepatitis B, C, and alcohol related hospitalizations were summed to create one variable capturing liver disease. These scores were then summed across all risk factors into a single score for each ZCTA, with higher values corresponding to higher prevalence of risk factors. NYC ZCTAs were classified into quartiles of this risk score.

**Socioeconomic status index**

Zip Code Tabulation Area (ZCTA) level data on median household income in the past 12 months (Table B19013), median gross rent (B25064), percent living
below 150% of the poverty line (Table C17002), education (B15002), percent working class (C24010), percent unemployed (B23025), and > 1 occupants per room (B25014) was downloaded from 2018 American Community Survey (ACS) 5-year estimates [28]. These variables were combined into an index of socioeconomic status using principal component analysis (PCA), as previously described [2]. A linear combination was then created to obtain an SES index score assigned to each ZCTA. A lower SES score corresponds to a lower SES, while a higher score represents higher SES.

COVID-19 mortality by zip code tabulation area

The COVID-19 related death count and the COVID-19 death rate were downloaded from the NYC Coronavirus (COVID-19) data repository [1] hosted by the NYCDOH from March 2nd (when the repository began publishing data) to July 6th, when Phase 3 of reopening started in NYC. The death rate, the risk sum, each component of the risk index, median age, SES index, and racial and Hispanic composition by ZCTA were mapped in ArcGIS v10.7.1. Deaths counts reflect people’s ZCTA of residence at the time of reporting, not the location of death.

Statistical analysis

The geographic unit of analysis was the ZCTA, where each ZCTA has a risk index score, SES index score, and mortality information. Wilcoxon rank-sum tests were performed to assess differences in COVID-19 death rate, risk index components, racial and ethnic composition, and median age according to risk index quartiles. Spearman correlations were performed to assess the association between the COVID-19 death rate per 100,000 residents, the risk index, median age, SES index, and the racial and Hispanic composition of the ZCTA. Stepwise linear regression was performed to predict the COVID-19 death rate per 100,000 residents through July 6th utilizing the risk index, median age, socioeconomic status index, and the racial and Hispanic composition at the ZCTA level as predictors. Variables that were significant at the $\alpha = 0.25$ level were considered for entry into the multivariable model and those that were significant at the $\alpha = 0.15$ level after adjustment remained in the final model. All analyses were performed in SAS v9.4.

Results

Risk index

Each component of the risk index, as well as the proportion of white and Hispanic residents, median age, SES index and COVID-19 death rate were analyzed according to quartiles of the risk index (Table 1). Across risk index quartiles representing increasing risk, the COVID-19 death rate per 100,000 persons significantly decreased with the increasing proportion of white residents in the ZCTA.
(β_adj = −0.91, SE = 0.31, p = 0.0037), while increasing proportion of Hispanic residents (β_adj = 0.90, SE = 0.38, p = 0.0200), median age (β_adj = 3.45, SE = 1.74, p = 0.0489), and risk index (β_adj = 5.84, SE = 0.82, p < 0.001) were statistically significantly positively associated with death rates (Table 3).

**Discussion**

This is the first analysis to incorporate CDC-defined risk factors for severe COVID-19 illness, NYC mortality data, racial and ethnic composition, and SES to predict COVID-19 death rates across NYC at the neighborhood level. Here we identify that areas with fewer white residents, more Hispanic residents, older residents, and with higher prevalence of a greater number of risk factors had increased COVID-19 mortality. These findings expand on existing literature that has partially addressed the potential risk of severe COVID-19 illness deriving from the complex intersection of chronic risk factors, SES and race [12] and can be useful in plans to address and reduce excess mortality in NYC5.
Within NYC, these areas of greatest risk correspond to the Bronx, southeastern and southern Brooklyn, and parts of southern Queens; areas that are also of lower socioeconomic status with larger resident minority populations. Past neighborhood analysis of NYC communities has reported that these areas have limited capacity to social distance, as measured by NYC subway data, and therefore have higher infection risk [29]. It is possible that these areas contain a large proportion of essential workers, such as healthcare, transportation and public-service workers, who are unable to work from home. It has been cited that social distancing, including working from home, is a privilege [18], and this is not a reality in communities characterized by higher proportions of comorbidities, risk factors for COVID-19, and adverse social determinants of health, including potential overcrowding and food insecurity. Further disparities in SES and healthcare access would only exacerbate adverse circumstances in these communities and contribute to increased risk of infection and mortality.

It has been noted that blanket public health recommendations in regards to social distancing for COVID-19 do not take local contexts and populations into account [17], and may be applied less effectively in socioeconomically disadvantaged and minority communities due to several factors, including overcrowded households and higher proportions of essential workers. Instead, understanding local contexts is central in creating effective public health solutions. One avenue to begin to rectify disparities is to perform increased testing in these communities [2], and more specifically in sub-groups at higher risk of serious COVID-19 complications. However, there remains a gap in appropriate testing and attention given to these vulnerable communities [2]. Here, we emphasize that many of the same communities that are more susceptible to increased COVID-19 spread also face higher risks of severe illness and mortality if infected. The data point to the need for more aggressive actions in preventing COVID-19 spread and detecting early infections in specific areas of NYC.

One limitation of this analysis is that the data used are aggregate ZCTA-level data, which limits our ability to draw individual-level conclusions. For instance, we cannot comment on the interaction between being non-white and having chronic risk factors on the likelihood of mortality, information that could help identify the most vulnerable populations. This study was ecological in nature and there are more individual level factors that may influence COVID-19 mortality that could not be taken into account here, although an accepted and comprehensive risk factor measure from the CDC was utilized. Another limitation is the cross-sectional nature of the analysis. Future studies should investigate these relationships with mortality using a larger study period, representing a larger number of COVID-19 related deaths, to validate these findings.

Conclusions
This analysis identifies that there are behavioral and racial/ethnic disparities in COVID-19 mortality, and highlights the most vulnerable areas in NYC that require increased focus from COVID-19 response policies. This analysis represents the first to our knowledge to investigate risk factors for COVID-19 mortality at the population-level in New York City, and supports the need for repeated and widespread testing in these communities with many risk factors for severe illness and mortality from COVID-19. Additionally, this analysis

| Table 2 Correlations among predictors and COVID-19 death rate per 100,000 residents |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                  | Statistic       | Death Rate      | Risk Index      | White Proportion (%) | Hispanic Proportion (%) | Median Age (years) |
| Death Rate up to                | p               | 0.67            | −0.56           | 0.42             | 0.03             | −0.63           |
| July 6                          | p-value         | <0.0001         | <0.0001         | <0.0001          | 0.6441          | <0.0001         |
| Risk Index                      | p               | −               | −0.48           | 0.25             | 0.08             | −0.67           |
| p-value                         | <0.0001         | 0.0012          | 0.27            | <0.0001          | 0.74             |
| White Proportion (%)            | p               | −               | −               | −0.40            | 0.39             |
| p-value                         | <0.0001         | <0.0001         | <0.0001         | <0.0001          |
| Hispanic Proportion (%)         | p               | −               | −               | −               | −0.39            |
| p-value                         | <0.0001         | <0.0001         | <0.0001         |
| Median Age (years)              | p               | −               | −               | −               | −0.37            |
| p-value                         | <0.0001         |

Note: Spearman correlations were performed
The socioeconomic status index was constructed from principal component analysis of 2018 American Community Survey estimates of median household income in the past 12 months, median gross rent, percent living below 150% of the poverty line, education, percent working class, percent unemployed, > 1 occupants per room. The education index was calculated on the population ≥ 25 years and is a weighted combination of the percent high school graduate, high school only, and more than high school, with a greater value indicating higher educational attainment.
shines light on vulnerable communities that require equitable resources to rebound from COVID-19 related hardships and mortality in these communities. Future analyses could explore how the presence of comorbidities influences the immune response following COVID-19 infection, as this could ultimately be a major determinant of mortality outcomes [30]. Future vaccination efforts should understand the vulnerability to
COVID-19 and mortality in NYC communities and direct attention to these areas accordingly. A successful public health response to future COVID-19 outbreaks will need to provide more attention to communities at the greatest risk in order to minimize the threat of serious illness and mortality in NYC.

Abbreviations
ACS: American Community Survey; CDC: Centers for Disease Control and Prevention; COPD: chronic obstructive pulmonary disease; COVID-19: coronavirus disease; NYC: New York City; NYCDOH: New York City Department of Health; PCA: principal component analysis; SES: Socioeconomic status; UHF: United Hospital Fund; US: United States; ZCTA: Zip code tabulation area

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12889-021-11498-x.

Additional file 1: Supplementary Table 1. Description of Risk Factor Measures. Supplementary Table 2. Correlations among socioeconomic status index and risk index components. Note: Spearman correlations were performed

Acknowledgments
Not applicable.

Authors’ contributions
ET conceptualized the study. NA provided statistical expertise. WL-C performed data analyses. All authors conducted the literature search. RF provided context for results within the ET conceptualized the study. NA provided statistical expertise. WL-C conducted the literature search. All authors have read and approved the final manuscript.

Authors’ information
Wil Lieberman-Cribbin: wil.lieberman-cribbin@icahn.mssm.edu
Naomi Alpert: naomialpert@mountsinai.org
Raja Flores: raja.flores@mountsinai.org
Emanuela Taioli: emanuela.taioli@mountsinai.org

Funding
No funding was received for this work.

Availability of data and materials
Data used in this analysis is available for open public access at https://github.com/nychealth/coronavirus-data. Individual data was not used.

Declarations
Consent to publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 10 December 2020 Accepted: 14 July 2021
Published online: 24 July 2021

References
1. NYC. Coronavirus disease 2019 (COVID-19) data. New York City Department of Health. https://github.com/nychealth/coronavirus-data. Updated daily. Accessed 1 Aug 2020.
2. Lieberman-Cribbin W, Turninello S, Flores RM, Taioli E. Disparities in COVID-19 testing and positivity in New York City. Am J Prev Med. 2020;59(3):326-32.
3. Wadhera RK, Wadhera P, Gaba P, Figueras JF, Maddox KE, Yeh RW, Shen C. Variation in COVID-19 Hospitalizations and Deaths Across New York City Boroughs. JAMA. 2020;323(21):2192-5.
4. Bojias GI. Demographic determinants of testing incidence and COVID-19 infections in new York City neighborhoods. Natl Bureau Econ Res. 2020, No. W26952.
5. Cordes J, Castro MC. Spatial analysis of COVID-19 clusters and contextual factors in new York City. Spatia Temporal Epidem Biol. 2020;1:100355.
6. Bergman A, Sella Y, Agre P, Casadevall A. Oscillations in US COVID-19 incidence and mortality data reflect diagnostic and reporting factors. Myxstems. 2020;5(4). https://journalas.nsm.edu/doi/10.1128/mSystems.0044-20.
7. Rosa, Amanda. Why It’s taking so long for New Yorkers to get test results. New York Times Published July 24, 2020. Available https://www.nytimes.com/2020/07/24/nyregion/coronavirus-test-results-nytc.html.
8. Okonji DR, Huynh M, Fine A, Baumgartner J, Castro A, Char HT, Daskalakis D, Deviney K, Guerra K, Harper S, Kennedy J. Preliminary estimate of excess mortality during the COVID-19 outbreak—New York City, March 11–May 2, 2020.
9. Chow N, Fleming-Dutra K, Gierke R, Hall A, Hughes M, Plishivtli T, Ritchey M, Roguski K, Skoff T, Ussery E. Preliminary estimates of the prevalence of selected underlying health conditions among patients with coronavirus disease 2019—United States, February 12–March 28, 2020.
10. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, Barnaby DP, Becker LB, Chehico JD, Cohen SL, Cokkingham J. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. JAMA. 2020;323(20): 2052-9.
11. Goyal P, Choi J, Pinheiro LC, Schenck EJ, Chen R, Jabri A, Satlin MJ, Campion Jr TR, Nahid M, Ringel JB, Hoffman KL. Clinical characteristics of Covid-19 in New York city. New England Journal of Medicine. 2020;382(24): 2372-4.
12. Raffman MA, Raffman JR. Disparities in the population at risk of serious illness from COVID-19 by race/ethnicity and income. Am J Preventive Med. 2020;59(1):137-9.
13. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. Nature. 2020;28:1-1.
14. Koma W, Neuman T, Claxton G, Rae M, Kates J, Michaud J. How many adults are at risk of serious illness if infected with coronavirus? Updated data. San Francisco: Kaiser Family Foundation-KFF, 2020.
15. El Chaar M, King K, Galvez Lima A. Are black and Hispanic persons disproportionately affected by COVID-19 because of higher obesity rates? Surg Obes Relat Dis. 2020;16(8):1096–1099. https://doi.org/10.1016/j.soard.2020.04.038. Epub 2020 May 11. PMID: 32522406; PMCID: PMC721681.
16. Palaiodimos L, Kokkidis DG, Li W, Karamanis D, Ognibene J, Arora S, et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a
17. Hooper MW, Napoles AM, Perez-Stable EJ. COVID-19 and racial/ethnic disparities. Jama. 2020;323(24):2466-7.
18. Yancy CW. COVID-19 and african americans. Jama. 2020;323(19):1891-2.
19. van Dorn A, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. Lancet (London, England). 2020;395(10232):1243.
20. Centers for Disease Control and Prevention. Coronavirus Disease 2019 (COVID-19) People Who Are at Increased Risk for Severe Illness. Updated June 25, 2020. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-at-increased-risk.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fneed-extra-precautions%2Fpeople-at-higher-risk.html.
21. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Population Health. 500 Cities Project Data [online]. 2018 [Accessed 23 October 2018]. URL: https://www.cdc.gov/500cities.
22. Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Data. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, [2017].
23. New York City Department of Health and Mental Hygiene. EpiQuery. Accessed 29 July 2020. Available at https://nyc.gov/health/epiquery.
24. New York State Department of Health. Statewide Planning and Research Cooperative System (SPARCS). https://www.health.ny.gov/statistics/sparcs/
25. New York City Department of Health. Environment & Health Data Portal. Heart Attack Hospitalizations. http://a816-dohbesp.nyc.gov/IndicatorPublic/VisualizationData.aspx?id=90,4466a0,13,Summarize.
26. United States Census Bureau. Relationship Files. Available at: https://www2.census.gov/geo/docs/maps-data/data/rel/ezctatractrel_10.txt.
27. New York City United Hospital Fund Codes. Available at https://www1.nyc.gov/assets/doh/downloads/pdf/ah/zipcodetable.pdf.
28. American Community Survey. 2018 American Community Survey 5-Year estimates. U.S. Census Bureau. https://data.census.gov/cedsci/. Updated January 23, 2020. Accessed 28 April 2020.
29. Carrion D, Colicino E, Pedretti NF, Arfer KB, Rush J, DeFelice N, Just AC. Assessing capacity to social distance and neighborhood-level health disparities during the COVID-19 pandemic. medRxiv. 2020.
30. de Lucena TM, da Silva Santos AF, de Lima BR, de Albuquerque Borborema ME, de Azevedo Silva J. Mechanism of inflammatory response in associated comorbidities in COVID-19: Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 2020;14(4):597-600.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.