A Data Driven Approach for Prioritizing COVID-19 Vaccinations in the Midwestern United States

Greg Arling[1], Matthew Blaser[2], Michael D. Cailas[3]*, John R. Canar[4], Brian Cooper[4], Joel Flax-Hatch[3], Peter J. Geraci[5], Kristin M. Osiecki[6], and Apostolis Sambanis[5]

[1] Purdue University, School of Nursing, College of Health and Human Sciences

[2] United States Environmental Protection Agency, Research Associate under an inter-agency agreement with Oak Ridge Institute for Science and Education

[3] Environmental and Occupational Health Sciences, School of Public Health, University of Illinois Chicago

[4] United States Environmental Protection Agency Region V; and Health Policy and Administration, School of Public Health, University of Illinois Chicago

[5] Health Policy and Administration, School of Public Health, University of Illinois Chicago

[6] University of Minnesota, Rochester, Center for Learning Innovation

Abstract

Considering the potential for widespread adoption of social vulnerability indices (SVI) to prioritize COVID-19 vaccinations, there is a need to carefully assess them, particularly for correspondence with outcomes (such as loss of life) in the context of the COVID-19 pandemic. The University of Illinois at Chicago School of Public Health Public Health GIS team developed a methodology for assessing and deriving vulnerability indices based on the premise that these indices are, in the final analysis, classifiers. Application of this methodology to several Midwestern states with a commonly used SVI indicates that by using only the SVI rankings there is a risk of assigning a high priority to locations with the lowest mortality rates and low priority to locations with the highest mortality rates. Based on the findings, we propose using a two-dimensional approach to rationalize the distribution of vaccinations. This approach has the potential to account for areas with high vulnerability characteristics as well as to incorporate the areas that were hard hit by the pandemic.

*Corresponding author: mihalis@uic.edu

Introduction

This research brief summarizes the findings of an in-progress study conducted by the SPH-PHGIS research team, which aims to identify the limitations and potentials of SVIs for prioritizing

vaccination plans. Due to this issue's urgency and importance, the CDC's Social Vulnerability Index (CDC.SVI) will be used as a case study. Here we present preliminary findings, with more to be reported in the coming weeks. The objectives of our research are to: assess the performance of the CDC.SVI in classifying counties according to their COVID-19 mortality rates; and propose an alternative approach to prioritization that incorporates both social vulnerability and actual experience of losses, i.e., COVID-19 mortality. Our goal is to provide better information on a community's vulnerability to a pandemic, as well as to the impact of vaccinations or other mitigation efforts in reducing mortality from the pandemic.

Background

In October 2020, the National Academies of Sciences, Engineering, and Medicine (NASEM) released a consensus study recommending a four-phase framework for equitable COVID-19 vaccine allocation [1]. In December 2020, the Advisory Committee on Immunization Practices (ACIP) recommended a detailed phased implementation plan for vaccination, starting with health care personnel and residents of long-term care facilities [2]. Given the limited supply of vaccines, the ACIP recommends for the next phase, 1b, to vaccinate "persons aged ≥75 years and frontline essential workers." [2] Both of these public health institutions raise the issue of promoting justice and mitigating health inequalities, especially for the particular racial and ethnic minority groups that were disproportionally affected by COVID-19 [1,2]. On this issue the NASEM study made specific recommendations that "vaccine access should be prioritized for geographic areas identified through CDC's Social Vulnerability Index or another more specific index." [1]

The CDC's Social Vulnerability Index (CDC.SVI) is one of the many indices in use aiming to "help local officials identify communities that may need support before, during, or after disasters." [3] This SVI, which is constructed from census data at various scales of aggregation, seeks to classify the relative social vulnerability of a location to a hazard based on a combination of factors. The CDC.SVI has a separate set of rankings for census tracts and counties according to 15 social attributes, including unemployment, minority status, and disability obtained from the American Community Survey. The CDC.SVI further groups these attributes into four related themes: Socioeconomic Status; Household Composition & Disability; Minority Status & Language; and Housing Type & Transportation. Census tract or county rankings can be obtained by state or nationally based on the individual indicator rankings, either summed across all 15 indicators or within each of the four themes.

Although each of the indices (i.e., themes) attempts to represent the underlying construct of social vulnerability, Tapsell et al. have pointed out that "there is still no consensus on a) the primary factors that influence social vulnerability, b) the methodology to assess social vulnerability, or c) an equation that incorporates quantitative estimates of social vulnerability into either overall vulnerability assessment or risk." [4] The supporting documents for the ACIP recommendations raise a few of the issues that are likely to be exacerbated with the use of SVIs (see ACIP's Evidence Table for COVID-19 Vaccines Allocation in Phases 1b and 1c of the Vaccination Program).

A notable strength of the CDC.SVI is the ease with which public health (PH) agencies can obtain the CDC.SVI ranking of their location from the well-organized CDC portal [3]. From the perspective of planning and resource allocation, social vulnerability is presumed to be an indicator of a community's risk for COVID-19, along with the need for additional resources to mount mitigation efforts against the pandemic. The popularity of this rank-based SVI has spawned other

2

rank-based indicators taking the same approach but directly addressing the COVID-19 pandemic. For example, the COVID-19 Community Vulnerability Index (CCVI) "incorporates the latest evidence on COVID-19 risk factors, fine-tuned with data collected over the course of the pandemic." [5]

Considering the potential for widespread adoption of SVIs to prioritize COVID-19 vaccinations, we feel that this approach should be carefully assessed, particularly for correspondence with outcomes (such as loss of life) in the context of the COVID-19 pandemic. Most applications of the rank-based CDC.SVI have been to single events such as natural or environmental disasters [6]. The COVID-19 pandemic, in contrast, involves a series of events characterized by waves and extended now over more than a year. For the Midwest, we know that the two waves have different characteristics [7] that are likely to challenge the applicability of SVIs. Although cases have been distributed across different age groups, loss of life has been concentrated among older adults and disproportionately in nursing homes and other long-term care settings [8]. The modeling of risk and social vulnerability must take these unique circumstances into account.

Our earlier work in Cook County, Illinois, revealed differences in spatial patterns of COVID-19 deaths in private households compared to those in long-term care facilities [8]. Neighborhood characteristics were predictive of household deaths but not deaths in long-term care facilities. Also, vaccination rollouts have been rapid and comprehensive for residents of long-term care facilities [9]. Although people of advanced age and those from racial and ethnic groups living in the community have been assigned high priority for vaccinations, these groups have had substantially lower vaccination rates than whites [10]. Nearly all major national COVID-19 reporting portals disclose the total number of COVID-19 deaths, with some offering separate reports on deaths in long-term care facilities. Yet, none of them report deaths occurring in households only (i.e., persons not in long-term care or other group settings). Also, much of the modeling of COVID-19 incidence and mortality has failed to distinguish between these residential settings [11]. A strength of our research is the separation of mortality figures for the two settings [7,8,11]. In this analysis we examine deaths among individuals in private households, under the assumption that this is where the challenge of vaccine prioritization lies.

Assessment Methodology

The University of Illinois at Chicago School of Public Health Public Health GIS (UIC-SPH-PHGIS) team developed a methodology for assessing and deriving vulnerability indices based on the premise that these indices are, in the final analysis, classifiers. Within this context, a SVI represents categories of a societal state (i.e., defined by socioeconomic and environmental factors) that unrealized events such as a pandemic have the potential to harm and thereby cause losses. In essence, the index derivation approach (e.g., the ranking of the original variables and the additive model for the themes in the CDC.SVI) becomes a classifier for each one of the n locations in terms of a potential for loss. The potential for loss is a common construct for many social vulnerability definitions, for example, "social vulnerability to natural hazards is the potential for loss and is complex interaction among risk, mitigation, and the social fabric of a place." [12] To validate the index as a classifier, a realized disaster loss (DL) event must be used corresponding to an actual disaster loss for each one of the n locations (e.g., number of deaths for a census tract or county). Consequently, performance assessment of the SVI classifier is easily accomplished by comparing the SVI (i.e., potential for loss) and the DL severity (i.e., actual losses or harm) rankings of each

3

location. To simplify the assessment, a confusion or error matrix is used with m classes for each dimension; m << n and usually contains 3 to 5 classes to correspond to the common color visualization schemes seen in the plethora of SVI maps. These classes are derived with the application of a binning or discretization methodology which transforms the n numerical values of a variable into m categorical counterparts.

For our study, the m×m PHGIS performance assessment (PA) matrix has the structure in Figure 1.

		SVI Categories				
		1	2	...	m	Σ
Disaster Loss Categories	1	c_{11}	c_{12}	...	c_{1m}	c_{1+}
	2	c_{21}	c_{22}	...	...	c_{2+}
	⋮	⋮	⋮	⋮	⋮	⋮
	m	c_{m1}	...	...	c_{mm}	c_{m+}
	Σ	c_{+1}	c_{+2}	...	c_{+m}	n

Figure 1. Schematic of the performance assessment (PA) matrix proposed by the PHGIS team for evaluating the SVIs.

The rank-based CDC.SVI as a prioritization tool for COVID-19 vaccination programs needs to identify high-risk areas, which are likely to be those with the greatest losses. The implicit assumption is that socioeconomic and environmental conditions, as well as the health condition of the residents, are the underlying causes of this elevated risk that leads to losses. This is aptly expressed in a rank-based CCVI report focusing on loss data that "can help us understand where and how the disease is impacting vulnerable populations, in order to prioritize resources and rapid response accordingly." [5]

The matching areas in terms of the two classifiers are contained in the diagonal elements of the PHGIS PA matrix, c_{ii}, providing an overall classification performance measure (Figure 1). For example, the first element of the PA matrix, c_{11}, contains the areas (e.g., counties) classified as having the lowest vulnerability and realization of losses. The sum of the matching areas divided by the total number of areas, n, yields an overall classification performance (OCP) rate. Similarly, the off-diagonal elements, c_{ij}, of the PA matrix identify the misclassified areas. These are areas recording a discrepancy between the class of their vulnerability status and the severity of the losses from past or on-going events. For example, the last element of the 1st column, c_{m1}, indicates the number of areas with the highest actual disaster loss that were classified to be the lowest vulnerable areas by the SVI. The sum of these below off-diagonal elements divided by the total number of areas, n, yields an overall underestimation rate (OUR). On the other hand, the element on the top right-hand corner indicates the areas that were predicted to have the highest vulnerability but

4

experienced the lowest level of DL. The elements above the diagonal indicate overestimation error (OE). The sum of these upper off-diagonal elements divided by the total number of areas, n, yields an overall overestimation rate (OER). Ideally, a well-designed COVID-19 vaccination plan will have a minimum of both OUR and OER. A high OUR implies high-risk areas that are not accounted for by the SVI, whereas a high OER implies allocation of valuable vaccine resources in areas with low risk.

Case studies

The study areas for our assessment of the CDC.SVI are counties in Illinois(n=102) and Wisconsin (n=72). The index rankings for all of the counties were derived from the CDC.SVI portal. We focus on the second COVID-19 wave, because of its recency and because it has different patterns from the first wave. Illinois and Wisconsin were selected due to their differences in spatial mortality patterns between waves. The assessment variable (target) was the household mortality rate (per 100,000). Details about the wave dates and characteristics are provided by the authors at the [Midwest Comprehensive Visualization Dashboards: COVID-19 MCVD](#) [13].

The vulnerability status to the pandemic is accounted for with the use of the summary classification (i.e., sum of the four CDC.SVI themes). The realization of the pandemic is represented by the DL (i.e., COVID-19 related deaths as of January 17, 2021). For this study, DL is the mortality rate per 100,000 of the household (not nursing home or long-term care) population recorded in the counties during the second wave of the pandemic [11]. For this application four classes of severity were used (1 is the lowest, 4 the most severe).

Findings for counties in Illinois and Wisconsin are reported separately for the sum of the ranks across all four themes and for Theme 3, Minority Status and Language (Figures 2 to 5). If vaccine prioritization were to be based solely on vulnerability according to the CDC.SVI, it would channel vaccinations and other resources first into SVI category 4, followed by SVI category 3, SVI 2, SVI 1. Unfortunately, the CDC.SVI categories for the counties do not match well with COVID-19 mortality rates. The mismatch patterns are consistent across all four PA matrices. A good match between vulnerability status and DL severity, by county, is seen in the diagonal elements of the PA matrix. Counties in the lower left cells would have high DL combined with low vulnerability. This mismatch could be of special PH concern in the cells furthest from the diagonal (i.e., close to c_{41}), as they would have the lowest priority for vaccination even though they were in the top quartile for DL severity. Conversely, counties in the upper right cells have low DL combined with a high vulnerability index. The cells furthest from the diagonal (i.e., close to c_{14}) would be given high priority for vaccination despite being in the bottom quartiles for DLs.
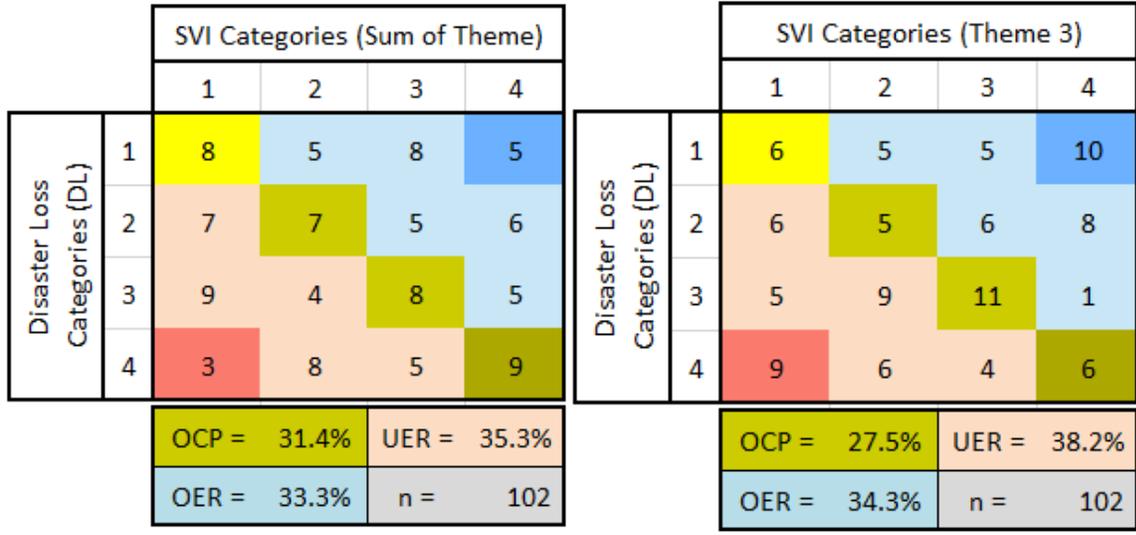
Figures 2 and 3. Illinois Counties - PHGIS performance assessment (PA) matrix with four categories applied for the summary CDC.SVI and the Theme 3 ranking of counties and their mortality rates.
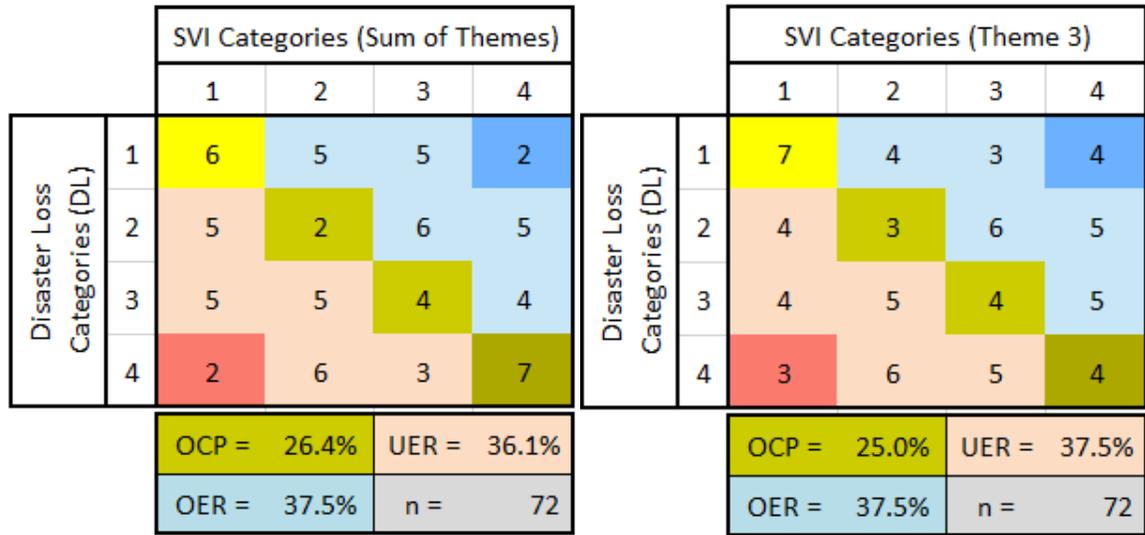
Figures 4 and 5. Wisconsin Counties - PHGIS PA matrix with four categories applied for the summary CDC.SVI and the Theme 3 ranking of counties and their mortality rates

From a statistical perspective, the measures of (mis)classification, or match/mismatch between vaccine prioritization (vulnerability index) and COVID-19 mortality rate in the counties were consistent across the four matrices, two for each state. The overall classification performance (OCP) was relatively low (ranging from 25.0% to 31.4%), while the overestimation rate (OER) and underestimation rates (UER) were in the same consistently high range (33.3% to 38.2%).

When we constructed similar matrices for three other Midwestern states (Figure 6), we found generally a lower matching classification (i.e., OCP) and higher over- and underestimation rates

(OER and UER). For example, the OCPs ranged from 14.9% to 28.3%, the UERs ranged from 31.5% to 41.0%, and the OERs ranged from 35.6% to 44.8%.

State	n	Sum of Theme rankings			Theme 3 ranking		
		OCP	UER	OER	OCP	UER	OER
INDIANA	92	28.3%	31.5%	40.2%	22.8%	35.9%	41.3%
MINNESOTA	87	25.3%	39.1%	35.6%	14.9%	40.2%	44.8%
MICHIGAN	83	24.1%	38.6%	37.3%	18.1%	41.0%	41.0%
Notes:	Themes and rankings are defined and provided by CDC.SVI Midwest States are Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. n is the number of counties with a SVI ranking by CDC.SVI						

Figure 6. Classification performance (OCP) and over and under estimation rates (UER and OER) for the CDV.SVI ranking of counties in three Midwestern states by DL categories

Use of the PHGIS PA Matrix as a Tool for Setting and Monitoring Vaccine Priorities

Our assessment suggests that the CDC.SVI may not be tapping into key factors contributing to losses from the COVID-19 pandemic. Therefore, this tool will need to be augmented to improve its value in planning for an initial vaccination rollout and for monitoring the effectiveness of vaccinations as they are deployed in a community. An effective prioritization tool will need to predict initial vulnerability to a loss as well as being able to track a community's success in mitigating against risk through vaccinations or other interventions.

The PHGIS PA matrix, which we applied in assessing the performance of the CDC.SVI, could also serve as a prioritization tool. The PHGIS-PA approach overcomes the one-dimensional aspect of SVIs. It takes into account (albeit incompletely) factors contributing to vulnerability, as well as a community's actual experience of losses.

- The matrix could be applied for vaccination planning and priority setting by determining which counties fit into each of the cells based on their combinations of SVI vulnerability and realized losses, i.e., from the beginning of the current wave up to the time of initial vaccination priority setting.
- In addition, the matrix could be used for tracking the effects of vaccinations by updating the loss categories periodically, i.e., updated weekly based on a four-week rolling average. By tracking transitions from one loss category to another, it would be possible to assess vaccination roll-out effectiveness over time.
- The matrix structure could be expanded to include additional dimensions, such as rates of vaccinations and priority populations, in order to obtain a clearer picture of effectiveness.

As an example of how the matrix could be employed for priority settings, we use the CDC.SVI, in lieu of a more-refined SVI that we have under development. We refer again to Figures 2 to 5. The

7

ordering of initial priorities would be straightforward for counties in cells along the diagonal, where vulnerability and losses line up well. The difficult cases are the counties in cells off the diagonal. Counties in the lower left cells would have high DL combined with a low vulnerability index. This mismatch could be of special concern in the cells furthest from the diagonal (i.e., close to c_{41}) because they would have a lower priority for vaccination if vulnerability were the main criterion. Yet, they were the counties with the worst losses. The vulnerability of these counties – as well as any factors that would suggest high losses continuing into the future, absent of mitigation efforts – need to be further investigated. Conversely, counties in the upper right cells would have low DL yet a high vulnerability index. The cells furthest from the diagonal (i.e., close to c_{14}) would be given high priority for vaccines from the standpoint of vulnerability, despite being in the bottom quartiles for DLs. Again, further investigation would be advised to determine if the counties should remain a high priority (because, for example, of the potential for increased losses).

Application of the PHGIS-PA matrix approach for tracking and continuous priority setting would involve an updating of the matrix as new data arrived about losses, and perhaps vaccine rollouts. The focus would be on changes in a county's cell membership between periods, which could indicate the need for re-ordering of priorities to counties experiencing increases in losses despite mitigation efforts.

Conclusions

Given the vaccine limitations and the need to ration the doses, the use of a social vulnerability index such as CDC's SVI alone as a planning tool for prioritizing vaccinations will not suffice to satisfy the multifaceted mitigation needs of a rational vaccination strategy. Assessment of this index with the PHGIS PA matrix approach found that the CDC.SVI risks assigning high priority to locations with the lowest mortality rates, and low priority to locations with the highest mortality rates.

The UIC SPH PHGIS team is proposing to use a two-dimensional approach for rationalizing the distribution of vaccinations. This approach has the potential to account for areas with high vulnerability characteristics as well as to incorporate the areas that were hard hit by COVID-19. Further research is under way to develop a planning tool with improved predictive performance that is trained on the COVID-19 experience and that incorporates the social vulnerability factors that contribute most to a community's vulnerability to the COVID-19 pandemic.

These findings could be further explored at a state and county level with the use of the Midwest Comprehensive Visualization Dashboard (MCVD), designed specifically for visualizing the spatial distribution of vulnerability and mortality at a county level throughout the Midwest. This dashboard is available at:

https://univofillinois.maps.arcgis.com/apps/MapSeries/index.html?appid=8bd3f5653abb41619b50d8c974e8a72b

References

1. National Academies of Sciences, Engineering, and Medicine. 2020. *Framework for Equitable Allocation of COVID-19 Vaccine*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25917.

2. Dooling, K, Marin, M., Wallace, M., et al. The Advisory Committee on Immunization Practices' Updated Interim Recommendation for Allocation of COVID-19 Vaccine — United States, December 2020. MMWR Morb Mortal Wkly Report 2021; 69:1657-1660. DOI: http://dx.doi.org/10.15585/mmwr.mm695152e2external icon.

3. Social Vulnerability Index CDC. Centers for Disease Control and Prevention, Geospatial Research Analysis, and Services Program (GRASP), Division of Toxicology and Human Health Sciences. Available at: https://www.atsdr.cdc.gov/placeandhealth/svi/index.html. Accessed February 15, 2021.

4. Tapsell S, McCarthy S, Faulkner H, Alexander M. (2010). Social Vulnerability and Natural Hazards. CapHaz-Net WP4 Report, Flood Hazard Research Centre – FHRC, Middlesex University, London Available at: http://caphaz-net.org/outcomes-results/CapHaz-Net_WP4_Social-Vulnerability.pdf. Accessed May 21, 2018.

5. Surgo Ventures. Vulnerable Communities and COVID-19: The Damage Done, and the Way Forward. Version 1, Published January 27, 2021. Available at: https://precisionforcovid.org/ccvi. Accessed February 10, 2021.

6. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. 2011. A Social Vulnerability Index for Disaster Management. *J Homel Secur Emerg Manage*. 8(1), 3. DOI:https://doi.org/10.2202/1547-7355.1792.

7. Arling GW, Blaser M, Cailas MD, Canar J, Cooper B, et al. 2020. A second wave of COVID-19 in Cook County: What lessons can be applied? *Online J Public Health Inform*. 12(2), e15. DOI:https://doi.org/10.5210/ojphi.v12i2.11506. PubMed

8. Blaser M, Cailas MD, Canar J, Cooper B, Geraci P, et al. Analyzing COVID-19 Mortality Within the Chicagoland Area: Data Limitations and Solutions. Research Brief No. 117. Policy, Practice and Prevention Research Center, University of Illinois Chicago. Chicago, IL. July 2020. DOI: https://doi.org/10.25417/uic.13470324.v19.   CDC COVID Data Tracker. Federal Pharmacy Partnership for Long-Term Care (LTC) Program. Centers for Disease Control and Prevention. Available at: https://covid.cdc.gov/covid-data-tracker/#vaccinations-ltc. Accessed February 18, 2021.

10. Recht H, Weber L. COVID-19: As Vaccine Rollout Expands, Black Americans Still Left Behind. Kaiser Health News. January 29, 2021. Available at: https://khn.org/news/article/as-vaccine-rollout-expands-black-americans-still-left-behind. Accessed February 18, 2021.

11. Canar J, Osiecki K, Sambanis A, Arling G, Cooper B, et al. A comprehensive analytic framework for COVID-19 mortality applicable to major metropolitan centers. *BMC Public Health*. (Submitted to).

12. Schimidlin TW, Hammer BO, Ono Y, King PS. 2009. Tornado shelter-seeking behavior and tornado shelter options among mobile home residents in the United States. *Nat Hazards*. 48, 191-201. https://doi.org/10.1007/s11069-008-9257-z

13. Blaser, M, Canar, J., Arling, G., Cailas, M. (2021): Midwest Comprehensive Visualization Dashboards: COVID-19 MCVD. University of Illinois at Chicago. PHGIS Program Report. PHGIS-WP-2.2020.12.28. https://doi.org/10.25417/uic.13650440.v1