# The influence of social distancing on COVID-19 mortality in US counties: Cross-Sectional Study

Phoebe Tran, Lam Tran, Liem Tran

Submitted to: JMIR Public Health and Surveillance
on: June 22, 2020

## *Table of Contents*

# The influence of social distancing on COVID-19 mortality in US counties: Cross-Sectional Study

Phoebe Tran[1] MSc, MPhil; Lam Tran[2] MSc; Liem Tran[3] PhD

[1]Chronic Disease Epidemiology Yale School of Public Health Yale University New Haven US
[2]Biostatistics Michigan School of Public Health University of Michigan Ann Arbor US
[3]Geography University of Tennessee Knoxville US

**Corresponding Author:**
Phoebe Tran MSc, MPhil
Chronic Disease Epidemiology
Yale School of Public Health
Yale University
60 College Street
New Haven
US



## *Abstract*

**Background:** Prior literature examining the impact of social distancing on US COVID-19 mortality mainly focuses on this relationship at the state level without accounting for nursing home COVID-19 deaths. This approach may obscure differences in social distancing behaviors by county in addition to the actual effectiveness of social distancing in preventing COVID-19 deaths.

**Objective:** To determine the influence of county level social distancing behavior on total COVID-19 deaths for counties throughout the US over the timespan that stay at home orders were implemented in most US states (March 2020-May 2020).

**Methods:** Using social distancing data from tracked mobile phones in all US counties, we estimated the relationship between social distancing (average % of mobile phones leaving home between March 2020-May 2020) and total COVID-19 deaths (when the state a county is in first started enacting social distancing measures and up to May 31st, 2020) with a mixed-effects negative binomial model while accounting for nursing home COVID-19 deaths as well as social distancing and COVID-19 related factors (days between when the first confirmed case was reported and May 31st, 2020, population density, social vulnerability, and hospital resource availability).

**Results:** Average mobile phone use between March 2020-May 2020, the social distancing variable in the study, was significantly associated with total COVID-19 deaths ($P < .001$). We observed that a 1% increase in average mobile phone use between March 2020-May 2020 led to an increase in total COVID-19 deaths by 6% (Incidence rate ratio (IRR): 1.06, 95% Confidence Interval (CI): 1.02, 1.10). Additionally, average mobile phone use values of 25-50% during these three months corresponded with 34.89 total COVID-19 deaths/100,000 people-155.43 total COVID-19 deaths/100,000 people.

**Conclusions:** As stay-at-home orders have been lifted in many US states, continued adherence to other social distancing measures such as avoiding large gatherings and maintaining physical distance in public are key to preventing additional COVID-19 deaths in counties across the country. Clinical Trial: Not applicable

(JMIR Preprints 22/06/2020:21606)
DOI: https://doi.org/10.2196/preprints.21606

### Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
✓ **No, I do not wish to publish my submitted manuscript as a preprint.**
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

# Original Manuscript

**Title:** The influence of social distancing on COVID-19 mortality in US counties: Cross-Sectional Study

**Authors:** Phoebe Tran, M.Sc., M.Phil.[1]; Lam Tran, M.Sc.[2]; Liem Tran, PhD.[3]

[1]Department of Chronic Disease Epidemiology, Yale University, New Haven, CT
[2]Department of Biostatistics, Michigan School of Public Health, Ann Arbor, MI
[3]Deparment of Geography, University of Tennessee, Knoxville, TN

**Corresponding and 1st Author:** Phoebe Tran, M.Sc., M.Phil.
Mailing Address: 60 College Street, New Haven, CT 06510
**Email address:** phoebe.tran@yale.edu
**Phone Number:** 404-245-7730
**Fax Number:** NA
**ORCID:** 0000-0003-1735-6903

**2nd Author:** Lam Tran, M.Sc.
**Mailing Address:** 1415 Washington Heights, Ann Arbor, MI 48109
**Email address:** lamtran@umich.edu
**Phone Number:** 770-356-6448
**ORCID:** 0000-0002-1242-3209

**3rd Author:** Liem Tran, PhD
**Mailing Address:** 1000 Phillip Fulmer Way, Knoxville, TN 37916
**Email address:** ltran1@utk.edu
**Phone Number:** 470-496-8866

**Authorship Statement:** Phoebe Tran originated the idea and the study design. Lam Tran ran the data analyses and interpreted results. Liem Tran assisted with study design. All authors contributed in writing this original article.

**Title:** The influence of social distancing on COVID-19 mortality in US counties: Cross-Sectional Study


## Abstract

Background: Prior literature examining the impact of social distancing on US COVID-19 mortality predominantly examines this relationship at the national level in addition to not separating nursing home COVID-19 deaths from total COVID-19 deaths. This approach may obscure differences in social distancing behaviors by county in addition to the actual effectiveness of social distancing in preventing COVID-19 deaths.
Objective: To determine the influence of county level social distancing behavior on COVID-19 mortality (deaths/100,000 people) for counties throughout the US over the timespan that stay at home orders were implemented in most US states (March 2020-May 2020).
Methods: Using social distancing data from tracked mobile phones in all US counties, we estimated the relationship between social distancing (average % of mobile phone use outside the home between March 2020-May 2020) and COVID-19 mortality (when the state a county is in had its first confirmed case and up to May 31st, 2020) with a mixed-effects negative binomial model while separating nursing home COVID-19 deaths from total COVID-19 deaths as well as accounting for social distancing and COVID-19 related factors (days between when the first confirmed case was reported and May 31st, 2020, population density, social vulnerability, and hospital resource availability). Results from the mixed-effects negative binomial model were then used to generate marginal effects at the mean which allowed for the separation of the influence of social distancing on COVID-19 deaths from other covariates while calculating COVID-19 deaths/100,000 people.
Results: We observed that a 1% increase in average mobile phone use outside the home between March 2020-May 2020 led to a significant increase in COVID-19 mortality by a factor of 1.18 ($p < .001$) while every 1% increase in average % of mobile phone use outside the home in February was found to significantly decrease COVID-19 mortality by a factor of 0.90 ($p < .001$).
Conclusions: As stay-at-home orders have been lifted in many US states, continued adherence to other social distancing measures such as avoiding large gatherings and maintaining physical distance in public are key to preventing additional COVID-19 deaths in counties across the country.


**Keywords:** COVID-19; mortality; social distancing; negative binomial model; marginal effects

## Introduction

With the rapid spread of coronavirus disease (COVID-19) across the US in early 2020, many states began enacting social distancing measures (stay at home orders, maintaining physical distance in public, avoiding large gatherings) beginning in March 2020 [1]. Enforcement of social distancing measures varied widely throughout the US with some states like Arkansas, Iowa, Nebraska, North Dakota, Oklahoma, South Dakota, Utah, and Wyoming opting to not implement stay at home orders at all [1]. Even with social distancing measures in place in large parts of the country, there were over 108,000 COVID-19 deaths by the end of May 2020 [2]. As stay at home orders have been lifted in most of the US and numbers of COVID-19 cases continue to rise, it is critical to assess the role of social distancing in the prevention of US COVID-19 deaths [1].

Although there are some studies that have assessed the influence of social distancing on COVID-19 mortality in the US using actual rather than simulated data, this research such as work by Medline et al. and Siedner et al. examines this relationship at the national level and foregoes examining the impact of social distancing on COVID-19 mortality at a more granular geographic scale [3, 4]. Both of these studies also do not consider mortality directly but tangentially related mortality measures (change in the mortality growth rate, time between first case and peak number of deaths) and in the case of Medline et al. does not include all US states [3, 4]. Additionally, this research does not separate nursing home COVID-19 deaths, a major source (~30%) of all COVID-19 fatalities, from total COVID-19 deaths which could potentially lead to biased findings that are not representative of the ongoing pandemic [3, 4]. These approaches may obscure the actual effectiveness of social distancing as well as differences in social distancing behaviors by county.

As a result, we sought to bridge the gap in US COVID-19 mortality literature by determining the association between social distancing and COVID-19 mortality at the county level across the US while separating nursing home COVID-19 deaths from total COVID-19 deaths as well as accounting for social distancing and COVID-19 related factors. To accomplish this, we modeled the relationship between average % of mobile phone use outside the home during the period US stay at home orders were in place (a proxy variable for social distancing) and COVID-19 mortality through the use of a mixed-effects negative binomial model and marginal effects at the mean.

## Methods

### *COVID-19 mortality*

Data on deidentified confirmed COVID-19 deaths in all 3,142 US counties was obtained from the Johns Hopkins University, Center for Systems Science and Engineering Coronavirus Resource Center [5]. This dataset contains comprehensive information on COVID-19 death counts provided by the US Centers for Disease Control and Prevention (CDC) and state health departments [5]. Additionally, we retrieved deidentified data on US nursing home COVID-19 deaths from the Centers for Medicare and Medicaid Services (CMS) [6]. ArcGIS Pro 2.5 was then used to identify all nursing home COVID-19 deaths in facilities within a county's geographic boundaries for all US counties [7]. We then defined a county's COVID-19 mortality (deaths/100,000 people) from when the state a county is in had its first confirmed case and up to May 31st, 2020 as:

$$Mortality = \frac{county's\ confimed\ COVID-19\ deaths - county's\ nursing\ home\ COVID-19\ deaths}{county's\ population} * 100,000$$

.

May 31st, 2020 was selected as the study endpoint as a majority of social distancing orders by state in the US were in effect in late March-early April until mid and late May [1]. Nursing home COVID-19 deaths were excluded from a county's total COVID-19 confirmed deaths because social distancing as measured through mobile phone tracking is not an accurate measure of the social distancing behavior of nursing home residents [8].

## *Social distancing metrics*

The county level social distancing metrics used in this study come from the SafeGraph COVID-19 Consortium which consists of social distancing data that has been gathered through anonymous global positioning system (GPS) tracking of mobile phones [8]. We used the average % of mobile phone use outside the home between March 2020-May 2020 as our social distancing measure as this timespan encompasses when stay at home orders were in effect in most of the US [1, 8]. The average % of mobile phone use outside the home in February 2020 was used to establish a baseline in which to compare the social distancing measure with since February 2020 marks when the number of US COVID-19 cases began to rapidly rise [9].

*Covariates*

We included covariates in our study that had been found to be associated with social distancing and COVID-19 from prior literature in these areas [10-13]. These covariates were days between when the first confirmed case was reported and May 31st, 2020, population density, social vulnerability, and hospital resource availability. For each US county, the days between when the first confirmed case was reported and May 31st, 2020 was determined from Johns Hopkins University's COVID-19 dataset [5]. With data on population and county size from the United States Census Bureau, we calculated a county's population density by dividing the number of individuals in a county by that county's area in square miles [14]. For population density, 100 persons per square mile were the units used as the US population density at the county level varies across five orders of magnitude (e.g., mean = 270, median = 40, min < 1, and max > 72,000) and multiples of hundreds of the average population density allows for easier interpretation of the results [15]. Social vulnerability was assessed using the 15 measures that make up the CDC's social vulnerability index (SVI): % of population below poverty, unemployment rate, per-capita income, % of population >25 years with no high school diploma, % of population ≥65 years, % of population ≤17 years, % of civilian non-institutionalized population with a disability, % of population that is a single parent household with children <18 years, % of population that is a minority, % of population >5 years who speak English "less than well", % of housing that is a structure with ≥10 units, % of housing that is a mobile home, % of occupied housing units with more people than rooms, % of households with no vehicles, % of population in institutionalized group quarters) [16]. Data on hospital resource availability was obtained from the Environmental Systems Research Institute's (ESRI) COVID-19 Resource Center [17]. Specifically, we used the number of staffed hospital beds within a county as our measure of hospital resource availability because this variable had complete data for the greatest number of counties and it is highly correlated with other measures of hospital resource availability such as number of adult intensive care unit (ICU) beds, number of licensed beds, and average ventilator use [17, 18].

### *Statistical modelling*

In order to assess the influence of social distancing on COVID-19 mortality, we created a mixed-effects negative binomial model. Our choice of a mixed-effects negative binomial model was motivated by two factors observed during the data processing phase of our study: (1) considerable variation between the average state-level mortality rates, suggesting the necessity of a state-level random intercept and (2) the variance of county-level mortality rates in each state tending to be much greater than the average state-level mortality rate, suggesting the usage of a negative binomial model. A county's COVID-19 mortality up to May 31$^{st}$, 2020 was specified as the model outcome while the social distancing exposure in the model was average % of mobile phone use outside the home between March 2020-May 2020. Covariates associated with social distancing and COVID-19 mentioned previously were also included in the model for adjustment purposes. As larger counties are more likely to have a greater COVID-19 mortality owing to their population size than smaller counties, the model included a population size offset to account for this tendency. Potential correlation resulting from counties within the same state having comparable behavioral factors, healthcare systems, and COVID-19 testing policies was dealt with by inclusion of a random intercept by state in the model. Results from the negative binomial model are reported as incidence rate ratios (IRR), calculated by exponentiating the model's coefficients. Using the same mixed-effects negative binomial model, we also ran an analysis that included nursing home COVID-19 deaths with results showing considerable differences between the model that included nursing home COVID-19 deaths and the main model that excluded these deaths (see Appendix).

With results generated from the mixed-effects negative binomial model, we examined the impact of social distancing on a county's COVID-19 mortality apart from the influence of other factors through marginal effects at the mean. The marginal effects at the mean set all covariates besides the social distancing variable (average % of mobile phone use outside the home between March 2020-May 2020) to their average value [19]. Compared to traditional regression models, marginal effects offer us the ability to more clearly isolate from other factors the influence social distancing has on COVID-19 deaths within a county [19]. Marginal effects analyses were run for values of average % of mobile phone use outside the home between March 2020-May 2020 between 25-41% since this range contains the mean (31.73%) of the social distancing variable as well as both extremes of social distancing behavior [8]. All statistical analyses were carried out in Stata 15 statistical software with testing for statistical significance conducted at $\alpha=0.05$ [20].

## Results

Including COVID-19 nursing home deaths, the total number of COVID-19 deaths during the study period was 102,958 while the mean county-level COVID-19 mortality was 13 deaths/100,000 people. Modeling results for the social distancing variable and other covariates of interest are presented in Table 1. With respect to our social distancing variable, for every 1% increase in the average % of mobile phone use outside the home between March and May 2020, the rate of COVID-19 mortality in terms of deaths/100,000 people was expected to significantly increase by a factor of 1.18 ($p < .001$). This is in contrast to the variable in February where for every 1% increase in average % of mobile phone use outside the home in February, COVID-19 mortality was found to significantly decrease by a factor of 0.90 ($p < .001$). Coefficient estimates for population density and days between when the first confirmed case was reported and May 31$^{st}$, 2020 were also significant. For every extra 100 individuals per square mile, the expected COVID-19 death rate increases by a factor of 1.02 (1.0002^100) whereas for every extra day between the first confirmed case of COVID-19 and May 31, the expected death rate increases by a factor of 1.03.

**Table 1.** Mixed-effects negative binomial model examining impact of social distancing on US COVID-19 mortality (deaths/100,000 people) between when a county had its first confirmed case and May 31st, 2020

| **Variables** | **Incidence Rate Ratio** | **95% Confidence Interval** | ***P*** |
|---|---|---|---|
| Average % of mobile phone use outside the home between March 2020-May 2020 | 1.18 | (1.12, 1.24) | <.001 |
| Average % of mobile phone use outside the home in February 2020 | 0.90 | (0.86, 0.94) | <.001 |
| Population density (100 persons per square mile) | 1.02 | (1.01, 1.04) | < .001 |
| Days between when the first confirmed case was reported and May 31st, 2020 | 1.03 | (1.02, 1.04) | < .001 |

We plotted the expected national COVID-19 mortality rates in Figure 1 for a range of mobile phone use outside the home percentages, corresponding to different stringencies of lockdowns and adherences to social distancing. The red line represents the marginal predicted national deaths per 100,000 with the shaded band a 95% confidence interval constructed around it. The black dot represents the reported national COVID-19 mortality rate by May 31st, 2020 and falls well within the predicted interval.

**Figure 1**. Predicted US COVID-19 deaths/100,000 people vs. average percent of mobile phone use outside the home between March and May 2020.

**Discussion**

Our study examined the influence of social distancing on COVID-19 mortality in each US county while accounting for nursing home COVID-19 deaths and covariates associated with social distancing and COVID-19. Social distancing as measured through average % of mobile phone use outside the home between March 2020-May 2020 but not in February 2020 was found to correspond with a significant increase in COVID-19 mortality in a county. In addition, by estimating national COVID-19 mortality rates by May 31st, 2020 for a range of mobile phone use outside the home percentages it was determined that the COVID-19 mortality rate by this date could have been even higher had there been poorer adherence to social distancing.

While our finding that increased mobile phone use outside the home during the period when stay at home orders were in place in many parts of US is linked with higher COVID mortality is somewhat expected, it is of some interest that a significant negative association between the average % of mobile phone use outside the home in February and COVID-19 mortality was observed. In fact, high percentages of mobile phone use outside the home in February 2020 happened more often in suburban and rural counties. As the first appearance of COVID-19 tended to be later in these less densely-populated regions, the February COVID-19 mortality rates in suburban and rural counties were lower than those of urban regions where the virus had time to spread, thus causing the negative association seen in February.

Given the ongoing presence and numerous ramifications of COVID-19, a few US studies such as Medline et al. and Siedner et al. have examined the impact of social distancing on COVID-19 at the national level [3, 4]. Similar to Medline et al. and Siedner et al., we also determined that social distancing was a critical factor in reducing COVID-19 deaths [3, 4]. However, it is difficult to directly compare our findings with those of Medline et al. and Siedner et al. as the method of assessing social distancing and COVID-19 mortality varies between these two studies and ours [3, 4]. For instance, Medline et al. used the time between the first COVID-19 case and implementation of social distancing order as their social distancing measure and time between the first case and peak number of COVID-19 deaths as their way to estimate COVID-19 mortality [3]. They found a positive association between social distancing and COVID-19 mortality where each additional day it took to implement a stay at home order resulted in an additional day to reach the peak number of COVID-19 deaths [3]. In the study by Siedner et al., the social distancing measure was determined by comparing the number of cases 14 days before a stay at home order was implemented and 3 days after a stay at home order was implemented while the outcome was the change in the mortality growth rate [4]. Using these measures, Siedner et al. observed that a week after stay at home orders were implemented the COVID-19 mortality growth rate significantly decreased by 2% day by day

[4]. By including all US counties and separating nursing home COVID-19 deaths from total COVID-19 deaths, our work complements that of existing literature on social distancing and US COVID-19 mortality while providing additional information on this relationship at a more granular level.

Although our study provides novel estimates of the impact of social distancing on COVID-19 deaths, it still has limitations that need to be considered. Our study used a proxy variable for social distancing (average % of mobile phone use outside the home during March 2020-May 2020) which would not be able to track social distancing when individuals leave their phones at home or turn their phone tracking off and may not encompass some aspects of social distancing such as maintaining physical distance in public areas and avoidance of large group gathering [8]. Although our metric for social distancing is incomplete, it is arguably one of the more quantitative approaches to tracking social distancing that may be prone to less bias, relative to other metrics such as stay-at-home order dates or attendance at cultural/sporting/religious events [2, 21]. Additionally, approximately 95% of the US adult population has a mobile phone, with no significant deviations in ownership by race, age, education, income, or urban/rural residence [22]. As a result, there is no evidence that any measurement error would systematically bias the social distancing coefficient either upwards or downwards and any random measurement error that does exist would not lead to any significant changes in the estimated covariate values. Finally, as with many studies on US COVID-19 mortality, our work is limited by the quality of COVID-19 mortality data available due to a combination of an insufficient number of testing kits to properly diagnose COVID-19, suspected COVID-19 deaths that were instead attributed to other respiratory illnesses such as pneumonia and influenza, and delays in COVID-19 deaths being reported to the CDC [23, 24].

The impact of social distancing on the COVID-19 pandemic has been particularly challenging to model at a national level. Different jurisdictions have imposed inconsistent policies with varying levels of restrictions, and substantial heterogeneity is present in actual adherence to these social distancing guidelines. Furthermore, there is considerable variability in the availability and quality of COVID-19 death data. In this study, we took daily, county-level mobile phone use outside the home and aggregated it into a national-level dataset to quantitatively measure how social distancing influences COVID-19 mortality. In doing so, we managed to account for within-county heterogeneity in social distancing implementation and adherence. By including both mobile data usage outside the home from February 2020 (pre-social distancing) and March-May 2020 (during/post-social distancing), our study examined the entire span of stay-at-home orders in US states and provides a comprehensive look at the role social distancing had on reducing COVID-19 deaths. For future studies, we note the need for further examination of the relationship between social distancing and COVID-19 given the relaxation of social distancing measures and the potential burden of continued COVID-19 mortality.

**Acknowledgements:** We would like to thank the Johns Hopkins' Center for Systems Science and Engineering and SafeGraph for making their data available and free to use for research purposes.

**Compliance with Ethical Standards**

**Funding:** There was no funding for this study.

**Disclosure of potential conflicts of interest:** The authors declare that they have no conflict of interest.

**Research involving human participants and/or animals:** This article does not contain any studies with human participants or animals performed by any of the authors. The COVID-19 and social distancing data used in this study come from secondary publicly available data sources that have been completely anonymized and released for public use by the Johns Hopkins' Center for Systems Science and SafeGraph.

**Informed consent:** This article is exempt from needing informed consent as no human participants were involved in the study and the data used has been completely anonymized and approved for public use by the Johns Hopkins' Center for Systems Science and SafeGraph.

**Data Access:** Both the COVID-19 and social distancing data used in this study are publicly available and can be respectively found at the Johns Hopkins' Center for Systems Science and Engineering website (https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6) and SafeGraph's website (https://www.safegraph.com/covid-19-data-consortium).

**Abbreviations**

CDC: Centers for Disease Control and Prevention
CI: confidence interval
CMS: Centers for Medicare and Medicaid Services
COVID-19: Coronavirus 2019
ESRI: Environmental Systems Research Institute
GPS: global positioning system
ICU: intensive care unit
IRR: incidence rate ratio
SVI: social vulnerability index
US: United States

**Appendix**
Results from the mixed-effects negative binomial model with nursing home Covid-19 deaths included

| Variables | Incidence Rate Ratio | 95% Confidence Interval | *P* |
|---|---|---|---|
| Average % of mobile phones leaving home between March 2020-May 2020 | 0.90 | (0.86, 0.95) | <.001 |
| Average % of mobile phones leaving home in February 2020 | 1.18 | (1.12, 1.24) | <.001 |
| Population density (100 persons per square mile) | 1.02 | (0.99, 1.03) | 0.091 |
| Days between when the first confirmed case was reported and May 31st, 2020 | 1.03 | (1.03, 1.04) | <.001 |

## References