Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal Pre-proof

Low education predicts large increase in Covid-19 mortality: The role of collective culture and individual literacy.

Jie Zhuo, Nicholas Harrigan

PII: S0033-3506(23)00205-6

DOI: https://doi.org/10.1016/j.puhe.2023.06.016

Reference: PUHE 4929

To appear in: *Public Health*

Received Date: 15 January 2023

Revised Date: 22 May 2023

Accepted Date: 13 June 2023

Please cite this article as: Zhuo J, Harrigan N, Low education predicts large increase in Covid-19 mortality: The role of collective culture and individual literacy., *Public Health*, https://doi.org/10.1016/j.puhe.2023.06.016.

Title:
Low education predicts large increase in Covid-19 mortality: The role of collective culture and individual literacy.

Author 1:
Name: Jie Zhuo
Affiliation address: Level 3, 25C Wallys Walk, Macquarie University. North Ryde NSW 2113 Australia
Email address: jie.zhuo@hdr.mq.edu.au

Author 2:
Name: Nicholas Harrigan
Affiliation address: Level 3, 25C Wallys Walk, Macquarie University. North Ryde NSW 2113 Australia
Email address: nicholas.harrigan@mq.edu.au

Corresponding author:
Name: Jie Zhuo
Affiliation address: Level 3, 25C Wallys Walk, Macquarie University. North Ryde NSW 2113 Australia
Email address: jie.zhuo@hdr.mq.edu.au
Postal address: Level 3, 25C Wallys Walk, Macquarie University. North Ryde NSW 2113 Australia
Phone: +61410385750

Low education predicts large increase in Covid-19 mortality: The role of collective culture and individual literacy.

ABSTRACT

Background: While many studies have found a high correlation between socio-economic inequalities and risk of Covid-19 deaths, there is reason to believe that much of this association is the product of differing levels of education.

Methods: We present multivariate models of fortnightly (n=60) Covid-19 deaths in 3108 US counties for the period 20 January 2020 to 10 May 2022. We model the direct (unmediated) effect of education, controlling for economy, race, geography, lack of vaccination, political orientation (vote Republican), poor health, and lack of preventative health behavior.

Findings: After controlling for correlated risk factors and indirect mechanisms that mediate education's impact on Covid-19 mortality, we find a strong direct (unmediated) correlation between low education and Covid-19 mortality (IRR [incidence rate ratios] = 1.17; 95% CI [confidence interval]: 1.15, 1.20). We theorize that this correlation reflects education's relationship with (1) collective cultures, such as norms of mask wearing, and (2) individual literacy, such as ability to engage with scientific communication.

Interpretation: Low education is strongly correlated with Covid-19 deaths, with an effect size of a university degree comparable to that of being over 65 in age. If this correlation is indeed causal, then it would imply that low education accounts for between 1 in 10 and 1 in 7 deaths in low education counties. Education should be conceptualized as a potential high-risk factor for Covid-19 death, and be taken into account when attempting to combat Covid-19 in disadvantaged communities. The effect of education cannot be reduced to its impact on vaccination or correlation with poor health or economic status, but it seems likely that low education communities have collective cultures that expose individuals to greater risks, and lack of individual literacy that limits engagement with public health messaging.

Key words: Covid-19, Covid-19 mortality, Social determinants of health, Educational inequality, Health inequality

INTRODUCTION

It is widely recognized that underprivileged communities bear an unequal burden of Covid-19 deaths.[1] Explaining what aspects of these communities make them vulnerable is crucial to formulating policies to address such inequality. This study examines the relatively neglected role of education as a predictor of Covid-19 mortality. We present evidence, from county-level data in the United States, that the direct (unmediated) effect of low education is a strong and independent predictor of increased Covid-19 mortality – even after controlling for other major mechanisms such as poor health status, economic status, geography, vaccination, preventative health behavior, and political orientation. We theorize that this strong direct effect can be explained by two unobserved mechanisms: the relationship between education and collective culture; and the relationship between education and individual literacy.

It is widely known that low education predicts worse health outcomes.[2–7] What is less clear is why and how low education leads to worse health outcomes. Part of the explanation is that low education is strongly correlated with other predictors of poor health outcomes, such as economic status, geography, age, and race.[1,8–17] Another part of the explanation, is that low education works through a series of mechanisms: people with lower levels of formal education show worse health outcomes because of greater distrust of medical professionals; lower self-awareness of personal preventive health behaviors; and a lower sense of control over lifestyle choices.[18–21]

In relation to Covid-19 mortality, similar patterns are present. Low education correlates strongly with other predictors, like economic status, and through these increases Covid-19 mortality.[8,9,20,22,23] Low education also appears to drive well known risk factors for Covid-19 mortality, like low vaccination rates.[24–27]

In this paper we argue that after controlling for correlated predictors (like economic status, age, race) and known mechanisms (like vaccination, political orientation, poor health), there remains a very large significant direct (unmediated) correlation between education on Covid-19 deaths. In arguing for this large direct correlation between education and Covid-19 deaths, our research is in line with recent findings of Chen et al[28]. Chen et al[28] have shown that a direct effect for education on Covid-19 deaths exists, even after controlling for age, race, sector, and geography, and have found that even after controlling for other factors, a person without any college education (as compared to someone with some college education) had a 3.0 to 5.7 times greater likelihood of dying of Covid-19.

This paper contributes to the literature on Covid-19 mortality by, first, providing further statistical evidence for this strong direct (unmediated) correlation between low education and Covid-19 mortality; and, second, providing a theory for the underlying mechanisms that explain this direct effect of low education. We show that the direct effect of low education is comparable in effect size to that of being aged over 65 - a known major risk factor that heavily guides public policy around Covid-19. We theorize that the direct effect of education on Covid-19 deaths is likely driven by two main mechanisms: (1) education's correlation with collective culture, such as norms around mask wearing, social distancing, and compliance with isolation rules; and (2) education's correlation with individual literacy, which impacts on ability to read, seek, and criticize new information.

In the section that follows we briefly outline previous literature on what we call the direct effect of education on Covid-19 mortality. We organize this literature review under the two main mechanisms of (1) collective culture, and (2) individual literacy. A simplified model of the relationship between (1) the direct (unmediated) effect, and what we call (2) correlated predictors; and (3) indirect mechanisms is presented in Figure 1.

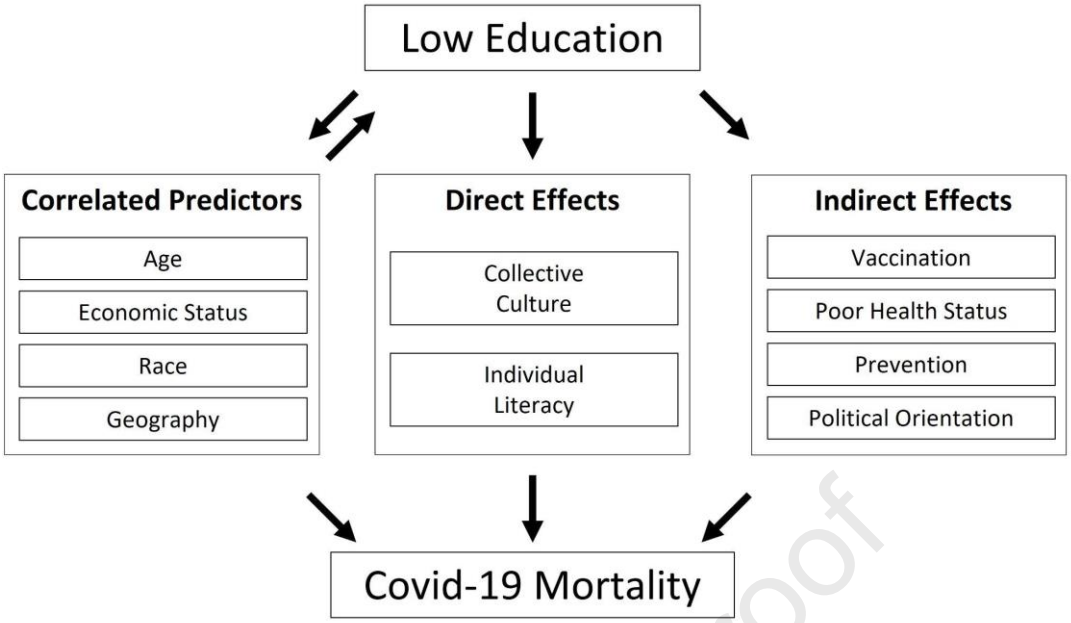

Figure 1. Theoretical model of effects of education on Covid-19 mortality

Collective culture. One potential mechanism that can explain the strong direct (unmediated) correlation of education and Covid-19 mortality is collective culture. Collective culture can be defined as a set of group level behaviors that constrain and guide individual action.[2,29,30] Collective culture can take the form of a spill-over effect, where the norms of a dominant group shaping the behavior of members of non-dominant groups immersed within the same culture.[31] An example of collective culture that likely impacts on Covid-19 deaths would be the social norm of mask wearing. The average education levels of a community might drive such a norm of mask wearing, with high education localities showing greater norms of mask wearing, with a spill-over effect where both low and high education individuals within the community embracing the community norm, and thereby reducing Covid-19 mortality. There are many studies that have found that people with higher levels of education have more positive attitudes toward preventive health behaviors, vaccination, and scientific knowledge.[24,25,32] There has not, however, been a theorizing that these attitudes will have a spill-over effect through collective culture in high/low education communities.

Individual literacy. A second potential mechanism that can explain the strong direct **correlation of** education with Covid-19 mortality is individual literacy. What we call individual literacy is the capacity to read, seek, and criticize information for a specific purpose.[19,33] In the existing literature, there are numerous examples of how higher levels of individual literacy reduces risk of Covid-19 mortality or morbidity, with high education: (1) increasing an individual's likelihood of seeking out appropriate health information;[34] (2) driving critical assessment of new information on preventive health behavior;[35] (3) increasing understanding of how medicine and prevention work;[36] and (4) correlating with better preparation of personal protective equipment.[37] We also know that certain traits associated with lower education, such as inability to identify and critique fake news, are associated with risks of Covid-19 mortality, such as lower compliance with protective public health measures.[38,39] While the idea that individual literacy is likely a risk for Covid-19 mortality is not controversial, this paper makes a contribution by theorizing that this is one of the two major mechanisms that explains the (unmediated) direct **correlation of** education **with** Covid-19 mortality.

METHODS

This paper models Covid-19 mortality in 3,108 counties, over the 60 fortnights from January 20, 2020 to May 10, 2022. The models include (1) a direct (unmediated) effect for education; (2) controls for a range of correlated predictors, such as age, race, economic status, and geography (rural area); and (3) controls for a series of indirect (mediated) effects through which education likely drives Covid-19 mortality, namely: (a) preventative health behavior, (b) vaccination rates, (c) a political orientation

(dis)trustful of mainstream scientific medical institutions (i.e. Republican vote), and (d) poor health status.

Data

Covid-19 deaths per fortnight (dependent variable). Daily Covid-19 deaths are from the New York Time (NYT) GitHub Center.[40] They cover the period January 20, 2020, and May 10, 2022. This NYT data combined some counties, which explains our final inclusion of 3108 counties. To smooth random variability in daily counts, we aggregated data into 60 fortnights. We assumed zero deaths in all counties in the period before their first reported death. There are 186,480 total observations.

Low education attainment. Educational attainment was measured with an index derived from factor analysis (see Table 1) of four measures of educational attainment, weighted by factor loading: proportion of county whose highest educational attainment was: 'less than high school', 'high school', 'college and associate's degree' and 'bachelor's degree or higher'. This data came from the 2016-2020 American Community Survey.[41]

Republican vote. Political views were measured with the percentage of the total votes in a county that voted for the Republican Presidential candidate in the November 2020 election.[42]

Vaccination. Vaccination was measured with an index created through factor analysis of three variables from the Centers for Disease Control and Prevention (CDC) 'COVID-19 Vaccinations in the United States, County' (see Table 1).[43] The index consisted of the following variables, weighted by factor loading: percentages of the relevant sub-population of a county that (a) completed the first dose, (b) completed the second dose, and (c) completed the booster, in every fortnight.

Preventative health behavior. Preventative health behavior was measured with an index created through factor analysis of four variables from the Centers for Disease Control and Prevention (CDC) 'PLACES: County Data (GIS Friendly Format), 2021 release' (see Table 1).[44] The index was comprised of the following variables, weighted by factor loading: percentages of the relevant sub-population of a county that (a) visited a doctor for routine checkup, (b) engaged in cholesterol screening, (c) had a colonoscopy (or similar screening), and (d) had a mammogram.

Poor health status. Six measures of poor health status were sourced from two CDC dataset 'Heart Disease Mortality Data Among US Adults (35+) by State/Territory and County – 2017-2019' and 'PLACES: County Data (GIS Friendly Format), 2021 release'.[44,45] We combined the six age-adjusted poor health variables (heart disease mortality, smoking, obesity, less sleeping, no physical activity, and asthma) weighted by factor loading into a poor health status index.

Demographic measures. Age, population, race, and median household income were sourced from the ACS.[41] In this paper, we use '% of over 65 year old' to measure age, and 'median household income' to measure economic position.

Geographic location. Geographic classification of the county was sourced from the 2013 Rural-Urban Continuum Codes from the US Department of Agriculture.[46] We divided geography into three categories (city: counties in metro areas; town: urban population of 2,500 and more, adjacent to a metro area; rural: completely rural or less than 2,500 urban population).

Table 1. Factor loadings of items in key indexes

Index **Items**	**Factor Loading**
Low Education Attainment	
Less than a high school diploma (%)	0.55
High school graduate (includes equivalency) (%)	0.73
College and associate's degree (%)	-0.20
Bachelor degree or higher (%)	-1.00
Poor health Status (age-adjusted prevalence)	
Heart disease mortality (per 100k)	0.73
Current smoking (%)	0.88
No leisure-time physical activity (%)	0.85
Obesity (%)	0.78
Sleeping less than 7 hours (%)	0.78
Current asthma (%)	0.65
Preventative Health Behavior	
Visits to doctor for routine checkup within the past year (%)	0.76
Cholesterol screening (%)	0.81
Fecal occult blood test, sigmoidoscopy, or colonoscopy (%)	0.59
Mammography use among women aged 50–74 years (%)	0.74
Vacciantion	
Completed dose1 (%)	0.92
Completed dose2 (%)	1.02
Completed booster (%)	0.76

Analysis

To model Covid-19 deaths per fortnight, we used three level negative binomial regression models with random effects for both the state and county levels. We also included an offset term for county population (10,000 population). Coefficients of negative binomial regression are presented as incidence rate ratios (IRR). 95% confidence intervals (95% CI) are presented. Explanatory power of the models is expressed in the (a) marginal r square; (b) conditional r square. Marginal r square expresses the explanatory power of the fixed effects in the models (with 1 = complete explanation of outcome by predictors), while conditional r square expresses the explanatory power of the whole models, including random effects. We performed all statistical analysis in R 4.2.0.

All variables except the lagged dependent variable (Covid-19 deaths) were standardized with a mean of zero and standard deviation of 1. This means that the effect sizes of each variable are interpretable from regression models, with the coefficients representing the impact on Covid-19 deaths (or other outcomes) from one standard deviation change in the independent variable. This is similar to what is called Standardized Beta in linear regression.

RESULTS

Descriptive statistics

Table 2 provides the descriptive statistics of the key variables. The data are grouped according to the quantile of low education index (Q1 = High Education; Q4 = Low Education). When compared to the highest education quartile (Q1), the lower education quartiles (Q2, Q3, Q4) show a trend of increasing death rates, decreasing vaccination rates, decreasing preventative health behavior, decreasing household income, increasing poor health status, and increasing Republican vote. We can see that the death rate (per 100k) due to Covid-19 in the lowest education quartile (Q4) is nearly twice that of the highest education quartile (Q1): 478 deaths/100k population (Q4) versus 250 deaths/100k population (Q1).

Table 2. Descriptive statistic of main variables (grouped by low education index quartile)

Low Education Index	Q1 (High Education)		Q2		Q3		Q4 (Low Education)	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Death Rate (per 100k)[1]	249.69	119.77	355.78	137.49	423.97	144.79	478.19	139.42
Total Deaths[1]	711.47	2207.22	279.02	783.86	170.35	288.53	106.31	97.72
New Deaths per Fortnight	11.98	73.12	4.68	20.31	2.85	8.70	1.79	3.61
Population (100k)	2.76	6.49	0.80	2.14	0.41	0.73	0.23	0.21
Median Houshold Income ($1000)	69.00	20.06	55.91	9.31	51.15	8.51	43.88	8.33
Republican Vote (%)	52.87	17.10	66.81	13.31	69.28	12.77	71.21	13.70
First Vac (%)[2]	67.66	16.08	57.85	13.18	55.27	11.19	51.31	10.89
Secoond Vac (%)[2]	59.71	12.95	51.60	10.59	48.80	9.85	44.79	9.39
Booster (%)[2]	47.11	11.96	45.68	10.63	43.38	9.69	39.92	7.53
Std. Vaccination Index[3]	0.15	1.15	-0.01	0.97	-0.02	0.97	-0.09	0.90
Std. Prevention Index[3]	0.46	1.05	-0.14	1.11	-0.20	0.97	-0.20	0.89
Std. Poor Health Status Index[3]	-0.82	0.69	-0.11	0.70	-0.12	0.68	0.55	0.92

Note:

The unit of analysis for all variables is the county (n = 3,108)

1. 20 Jan 2020 - 10 May 2022
2. on 10 May 2022
3. Mean = 0; Standard Deviation = 1

Table 3 provides descriptive statistics of the dataset. On average, each county has nearly 5.3 new Covid-19 deaths per fortnight. The variance of deaths per fortnight is extremely large, with a maximum value of 9382 deaths in one county in one fortnight (New York, 14 April 2020). For education attainment, the average percentage of a county's population with a bachelor degree or higher is 22.6%. This too varies a large amount. For example, the county (with more than 1,000 residents) with the lowest percentage bachelor degrees is Kusilvak Census Area (Alaska), where only 3.2% of its 8,117 residents have a bachelor degree. In contrast, the county with the largest percentage bachelor degrees is Falls Church (Virginia), where 79.14% of its 14,185 population has a bachelor degree.

Table 3. Descriptive Statistics

	Mean	Std.Dev	Min	Max
New Deaths per fortnight	5.32	38.44	0	9382
Demography				
Age 65 and over (%)	4.68	1.68	0.47	33.33
Population (100k)	10.43	35.74	0.01	1004
Median household income ($1000)	55.07	15.57	22.29	353.09
Race				
White (%)	84.80	16.94	3.45	100
Black (%)	9.28	14.75	0	87.94
Asian (%)	1.48	3.16	0	55.72
Native (%)	2.13	8.07	0	95.19
Hispanic (%)	9.52	13.87	0	98.88
Geography (dummy variables)				
City	0.37	0.48	0	1
Town	0.29	0.45	0	1
Rural	0.34	0.47	0	1
Education Attainment				
Less than a high school diploma (%)	12.40	6.04	1.39	78.15
High school graduate (includes equivalency) (%)	33.94	7.37	6.52	54.96
College and associate's degree (%)	31.06	5.34	5.93	81.82
Bachelor degree or higher (%)	22.60	9.69	0	79.14
Low Education Attainment (index)	2.79	16.39	-74.87	50.86
Poor Health Status				
Heart disease mortality (per 100k)	354.55	84.06	0	810.5
Current smoking (%)	20.4	4.16	6.50	43.00
No leisure-time physical activity (%)	30.39	5.80	12.90	51.80
Obesity (%)	35.78	4.25	16.40	51.00
Sleeping less than 7 hours (%)	36.81	3.96	25.60	49.10
Current asthma (%)	9.72	0.97	7.20	14.30
Poor Health Status (index)	365.55	71.49	85.7	720.26
Preventative Health Behavior				
Visits to doctor for routine checkup within the past year (%)	74.19	3.56	60.00	82.90
Cholesterol screening (%)	83.44	2.41	67.20	91.00
Fecal occult blood test, sigmoidoscopy, or colonoscopy (%)	61.97	4.45	39.80	74.40
Mammography use among women aged 50–74 years (%)	70.91	3.67	54.00	81.80
Prevention (index)	213.00	8.01	175.09	235.30
Vaccination				
Completed first dose (%)	22.57	25.90	0	99.90
Completed second dose (%)	19.78	25.90	0	99.90
Completed booster (%)	15.16	17.63	0	97.10
Vaccination (index)	52.46	58.70	0	238.37
Political View				
Republican vote (%)	65.04	16.03	8.73	96.18
n(fortnights)		60		
n (counties)		3,108		
n (fortnight-counties)		186,480		

Models of Covid-19 deaths

Table 4 presents the results from the multilevel negative binomial models of Covid-19 deaths. Model 1 is a baseline model which shows that low education significantly predicts more Covid-19 deaths (IRR = 1.34; 95% CI: 1.32, 1.36). Model 2 includes median household income, which is not a significant predictor (IRR = 0.99; 95% CI: 0.97, 1.00). Model 3 includes all covariates and shows that (a) low education attainment is strongly associated with more Covid-19 deaths (IRR = 1.17; 95% CI: 1.15, 1.20); (b) median household income is not significant (IRR = 0.99; 95% CI: 0.98, 1.01); and (c) poor health status, preventative health behavior, Republican vote, and vaccination rates all have a significant correlationn with Covid-19 deaths.

Table 4. Multilevel negative binomial predicting new deaths per fortnight

	Model 1		Model 2		Model 3	
Predictors	*IRR*	*95% CI*	*IRR*	*95% CI*	*IRR*	*95% CI*
(Intercept)	0.79 ***	0.74 – 0.85	0.79 ***	0.74 – 0.85	0.73 ***	0.70 – 0.77
Main Variables						
Low education attainment	1.34 ***	1.32 – 1.36	1.34 ***	1.32 – 1.36	1.17 ***	1.15 – 1.20
Median household income			0.99	0.97 – 1.00	0.99	0.98 – 1.01
Vaccination					0.60 ***	0.59 – 0.60
Poor health status					1.03 **	1.01 – 1.05
Prevention					0.97 *	0.94 – 1.00
Republican vote					1.11 ***	1.08 – 1.14
Race						
White *(ref.)*	-	-	-	-	-	-
Black	1.00	0.99 – 1.01	1.00	0.98 – 1.01	1.10 ***	1.08 – 1.13
Native	1.03 ***	1.02 – 1.05	1.03 ***	1.02 – 1.04	1.08 ***	1.06 – 1.09
Asian	0.93 ***	0.92 – 0.94	0.93 ***	0.92 – 0.95	0.96 ***	0.94 – 0.97
Hispanic	0.95 ***	0.94 – 0.97	0.95 ***	0.94 – 0.97	1.03 **	1.01 – 1.05
Age						
Age 65 and over	1.15 ***	1.13 – 1.17	1.15 ***	1.13 – 1.17	1.18 ***	1.17 – 1.20
Geography						
City (ref.)	-	-	-	-	-	-
Town	1.08 ***	1.07 – 1.09	1.08 ***	1.06 – 1.09	1.06 ***	1.05 – 1.07
Rural	1.18 ***	1.16 – 1.20	1.18 ***	1.16 – 1.20	1.14 ***	1.12 – 1.16
Time period						
Fortnight (log)	1.64 ***	1.63 – 1.65	1.64 ***	1.63 – 1.65	2.94 ***	2.90 – 2.98
New deaths last fortnight	1.69 ***	1.68 – 1.71	1.69 ***	1.68 – 1.71	1.67 ***	1.65 – 1.68
Observations	186480		186480		186480	
Marginal R^2	0.231		0.231		0.383	
Conditional R^2	0.282		0.282		0.414	

Note:

* p<0.05 ** p<0.01 *** p<0.001

All independent variables are standardlized

If the correlation of low education with Covid-19 is indeed causal, what does this imply in in terms of human lives lost? We provide three estimates of the effect size: first, the difference in yearly deaths due to Covid-19 for high and low education counties; second, a conservative estimate of the preventable deaths in the lowest education counties; and third, less conservative estimate of the preventable deaths in the bottom half of the US population (ranked by county education levels).

First, after adjusting for all other variables, we observed a difference of approximately 71 deaths per year between counties in the highest (Q1) and lowest (Q4) education quartiles when comparing their predicted Covid-19 mortality rates.

Second, our conservative model restricts our analysis to the 50% of counties with the lowest education levels (where about 50 million people live), and asks what would have happened if education levels here were the same as the median county. In this scenario, if we assume the correlations we have found are

causal, then the deaths in these counties would reduce by 32,000, saving 1 in 10 lives lost to Covid-19 in these counties.

Third, our less conservative model restricts our analysis to the counties with the lowest education levels where 50% of the USA population lives (around 150 million people). We ask what would have happened to deaths if the correlation we found is causal, and these counties had education levels of one standard deviation above the mean for the country, i.e. similar to an affluent county. In this scenario, we find that deaths in these counties would be reduced by 107,000, saving approximately 1 in 7 lives lost to Covid-19 in these counties.

DISCUSSION

This study presents evidence of the strong direct (unmediated) correlation of education with Covid-19 mortality. It shows that the correlation between education and Covid-19 deaths is still significant and strong even after controlling for age, economy, geography, vaccination, poor health status, prevention, and political vote. We put forward the theory that this direct effect of education largely operates (1) through collective culture, where the average education levels of a country has spillover effects - both positive and negative - to all members of the county, and (2) through individual literacy, which impacts on an individual's ability to read, find, and criticize relevant information.

This study is a contribution to the literature, first, because it shows that education is a significant and important correlate of Covid-19 mortality. In studies of other major social determinants of health – like economic status, race, and geography – the importance of education has been somewhat overlooked.[47–49] Second, we show that a significant part of education's impacts on Covid-19 mortality is a direct (unmediated) effect, and this correlation remains even after we control for factors like vaccination rates and differences in health behaviors of higher and lower educated populations. Third, we present a new theoretical perspective about the mechanisms driving this direct effect of education, arguing it is likely a product of both collective culture and individual literacy.

If the correlations found in our models are causal, what do these results look like for specific communities across America? We can see the differences by looking at two archetypical counties – one high in the predictors of Covid-19 mortality, and one low in the predictors of Covid-19 mortality. Walker county in Alabama has a population of 63,802 and only 13.4% have a bachelor degree. 446 members of this county died of Covid-19 over the period covered by this paper, representing 0.70% of the population. In contrast, Marin county in California has a population of 259,441 and 60.2% have a bachelor degree. 292 members of this county died of Covid-19 over the period covered by this paper, representing 0.11% of the population.

The major strengths of this study include its comprehensive dataset and its advanced statistics. The study is comprehensive in that it models deaths for most of the USA and across the period from the beginning of the pandemic until the present day. The study involves advanced modeling, including multivariate modeling, with extensive controls for vaccination, voting behavior, prevent health, poor health, economic position, age, race, and geography; negative binomial regression modeling; random effects for state and country. Another strength of this study is to solve the problem of multicollinearity with creating index and autocorrelation with statistical controls for the lagged dependent variable and time. The study tests for direct mechanisms by which low education potentially gives rise to Covid-19 mortality after controlling for four mediators (vaccination, voting behavior, prevent health, poor health) as indirect effect on Covid-19 mortality.

There are some limitations of this study. First, the most important limitation of the paper is the potential for omitted variable bias. While the dataset is longitudinal, the study is ultimately a correlational study because of the potential for unmeasured variables to have given rise to the observed correlations found in this study. We have tried, within the limits of our resources, to collect data for the most important potential alternative explanations, and we have not found any that significantly changes the results of this paper. However, the potential remains for unmeasured variables to explain the correlations presented in this paper. For this reason, the results of our models should be taken as starting points for further investigation of the relationship between education and Covid-19 deaths, not conclusive evidence of a causal relationship. Second, this study uses county level data to do the analysis which might give rise to the fallacy of inferring about individuals from collective units of analysis. Third, the deaths in big cities

are often reported as one when they covered multiple counties. Fourth, this study collects data from diverse sources which we regard as the best sources, but we did not have oversight of their collection and verification, so there is the potential for errors being introduced here. Fifth, there is the potential for future studies to collect more comprehensive control variables, such as county and state level policies like lockdown restrictions and aid packages. Finally, there is a part of the academic literature that argues that the correlations between education and health outcomes are largely spurious, and the reason such correlations appear is because of insufficient controls for (1) endogeneity, where healthier people are more likely to be better educated; and (2) unobserved variables, particularly time preference or intertemporal choice and genetic endowments which are said to predict both education and health outcomes.[7,50–52]

Our findings are largely consistent with, and extend, the findings of other studies. It is difficult to explain why some studies find no effect of education on Covid-19 mortality, including those that use very similar datasets. Despite these differences, we believe our findings are sound. With respect to studies of other social determinants of health – such as SES, race, and geography – we believe that controls for education might change some of the results of these studies.

For policy makers and medical professionals these findings suggest the potential need for tailored approaches to reach low education communities with Covid-19 interventions. For policy makers, low education – such as the percentage with bachelor degrees – could be an indicator of counties' need for priority Covid-19 educational and medical resources.[53] The findings also point to the fact that the people most in need of medical attention around Covid-19 are likely those least in contact with the medical system. Finding ways to reach these people seems crucial to addressing excess Covid-19 deaths and systematic inequalities in Covid-19 mortality (Buheji et al., 2020). For medical professionals, the results of this paper suggest ways to profile and identify patients most likely to be at risk of Covid-19. Patients from lower education backgrounds could be prioritized for support and/or focused conversations about healthy and safe practices to prevent Covid-19 infection and mortality.

In conclusion, we present evidence that education has a strong direct correlation with Covid-19 mortality. We showed that if the correlations we found are causal, then even after controlling for other factors like age, economic status, and geography around 1 in 10 to 1 in 7 of all Covid-19 deaths in "low education" counties are directly attributed to the low education profile of their citizens. These findings show that reaching such communities with effective support and engagement could be crucial to reducing both the magnitude of Covid-19 mortality, and unequal distribution of the costs of this pandemic.

REFERENCES

1. Van Dorn A, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. The Lancet. 2020;395(10232):1243–4.

2. Triandis H. Collectivism v. individualism: A reconceptualisation of a basic concept in cross-cultural social psychology. In: Cross-cultural studies of personality, attitudes and cognition. Springer; 1988. p. 60–95.

3. Subramanian SV, Huijts T, Avendano M. Self-reported health assessments in the 2002 World Health Survey: how do they correlate with education? Bull World Health Organ. 2010;88(2):131–8.

4. Davies NM, Dickson M, Davey Smith G, Van Den Berg GJ, Windmeijer F. The causal effects of education on health outcomes in the UK Biobank. Nat Hum Behav. 2018;2(2):117–25.

5. Hamad R, Elser H, Tran DC, Rehkopf DH, Goodman SN. How and why studies disagree about the effects of education on health: A systematic review and meta-analysis of studies of compulsory schooling laws. Soc Sci Med. 2018;212:168–78.

6. Clark D, Royer H. The effect of education on adult mortality and health: Evidence from Britain. Am Econ Rev. 2013;103(6):2087–120.

7. Xue X, Cheng M, Zhang W. Does education really improve health? A meta-analysis. J Econ Surv. 2021;35(1):71–105.

8. Sheikh MA, Abelsen B, Olsen JA. Role of respondents' education as a mediator and moderator in the association between childhood socio-economic status and later health and wellbeing. BMC Public Health. 2014;14(1):1–15.

9. Sheikh MA, Abelsen B, Olsen JA. Education and health and well-being: direct and indirect effects with multiple mediators and interactions with multiple imputed data in Stata. J Epidemiol Community Health. 2017;71(11):1037–45.

10. Karmakar M, Lantz PM, Tipirneni R. Association of social and demographic factors with COVID-19 incidence and death rates in the US. JAMA Netw Open. 2021;4(1):e2036462–e2036462.

11. Tan AX, Hinman JA, Magid HSA, Nelson LM, Odden MC. Association between income inequality and county-level COVID-19 cases and deaths in the US. JAMA Netw Open. 2021;4(5):e218799–e218799.

12. Dobson E, Graham C, Dodd E. When Public Health Crises Become Entwined: How Trends in COVID-19, Deaths of Despair, and Well-Being Track across the United States. Ann Am Acad Pol Soc Sci. 2021;698(1):88–110.

13. Takian A, Kiani MM, Khanjankhani K. COVID-19 and the need to prioritize health equity and social determinants of health. Vol. 65, International journal of public health. Springer; 2020. p. 521–3.

14. Monnat SM. Rural-Urban variation in COVID-19 experiences and impacts among US working-age adults. Ann Am Acad Pol Soc Sci. 2021;698(1):111–36.

15. Blagosklonny MV. From causes of aging to death from COVID-19. Aging. 2020;12(11):10004.

16. Baptiste DL, Commodore-Mensah Y, Alexander KA, Jacques K, Wilson PR, Akomah J, et al. COVID-19: Shedding light on racial and health inequities in the USA. J Clin Nurs. 2020;

17. Kim SJ, Bostwick W. <? covid19?> Social vulnerability and racial inequality in COVID-19 deaths in Chicago. Health Educ Behav. 2020;47(4):509–13.

18. Rosenkranz RR. Service-learning in higher education relevant to the promotion of physical activity, healthful eating, and prevention of obesity. Int J Prev Med. 2012;3(10):672.

19. Bijwaard GE, van Kippersluis H, Veenman J. Education and health: the role of cognitive ability. J Health Econ. 2015;42:29–43.

20. Mirowsky J, Ross CE. Education, social status, and health. Routledge; 2017.

21. Raghupathi V, Raghupathi W. The influence of education on health: An empirical assessment of OECD countries for the period 1995–2015. Arch Public Health. 2020;78(1):1–18.

22. Baradaran Motie G, Biolsi C. County-level determinants of social distancing (or lack thereof) during the COVID-19 pandemic. Contemp Econ Policy. 2021;39(2):264–79.

23. Rothwell J, Smith E. Socioeconomic Status as a Risk Factor in Economic and Physical Harm from COVID-19: Evidence from the United States. Ann Am Acad Pol Soc Sci. 2021;698(1):12–38.

24. Szilagyi PG, Thomas K, Shah MD, Vizueta N, Cui Y, Vangala S, et al. National trends in the US public's likelihood of getting a COVID-19 vaccine—April 1 to December 8, 2020. Jama. 2021;325(4):396–8.

25. Bogart LM, Ojikutu BO, Tyagi K, Klein DJ, Mutchler MG, Dong L, et al. COVID-19 related medical mistrust, health impacts, and potential vaccine hesitancy among Black Americans living with HIV. J Acquir Immune Defic Syndr 1999. 2021;86(2):200.

26. Zhang W, Wu YY, Wu B. Racial/ethnic disparities in getting COVID-19 vaccine: Do age, gender, and education matter? Health Equity. 2022;6(1):500–7.

27. Green MS, Abdullah R, Vered S, Nitzan D. A study of ethnic, gender and educational differences in attitudes toward COVID-19 vaccines in Israel–implications for vaccination implementation policies. Isr J Health Policy Res. 2021;10(1):1–12.

28. Chen YH, Matthay EC, Chen R, DeVost MA, Duchowny KA, Riley AR, et al. Excess Mortality in California by Education During the COVID-19 Pandemic. Am J Prev Med. 2022;63(5):827–36.

29. Shapiro JM. Smart cities: quality of life, productivity, and the growth effects of human capital. Rev Econ Stat. 2006;88(2):324–35.

30. Battiston F, Cencetti G, Iacopini I, Latora V, Lucas M, Patania A, et al. Networks beyond pairwise interactions: structure and dynamics. Phys Rep. 2020;874:1–92.

31. Finsterwalder J, Kuppelwieser VG. Intentionality and transformative services: Wellbeing co-creation and spill-over effects. J Retail Consum Serv. 2020;52:101922.

32. Timpka T, Nyce JM. Face mask use during the COVID-19 pandemic—the significance of culture and the symbolic meaning of behavior. Ann Epidemiol. 2021;59:1–4.

33. Barnard A, Nash R, O'Brien M. Information literacy: developing lifelong skills through nursing education. J Nurs Educ. 2005;44(11):505–10.

34. Cutler DM, Lleras-Muney A. Education and health: evaluating theories and evidence. National bureau of economic research Cambridge, Mass., USA; 2006.

35. Tsui L. Reproducing social inequalities through higher education: Critical thinking as valued capital. J Negro Educ. 2003;318–32.

36. Meng XJ, Dong GH, Wang D, Liu MM, Lin Q, Tian S, et al. Prevalence, awareness, treatment, control, and risk factors associated with hypertension in urban adults from 33 communities of China: the CHPSNE study. J Hypertens. 2011;29(7):1303–10.

37. Abrams EM, Greenhawt M. Risk communication during COVID-19. J Allergy Clin Immunol Pract. 2020;8(6):1791–4.

38. van Der Linden S, Roozenbeek J, Compton J. Inoculating against fake news about COVID-19. Front Psychol. 2020;11:566790.

39. Jones-Jang SM, Mortensen T, Liu J. Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. Am Behav Sci. 2021;65(2):371–88.

40. The New York Time. Coronavirus (Covid-19) data in the United States [Internet]. 2022. Available from: https://github.com/nytimes/covid-19-data

41. ACS. The 2016-2020 American Community Survey [Internet]. 2022. Available from: https://data.census.gov/cedsci

42. MIT Election Data and Science Lab. County Presidential Election Returns 2000-2020 [Internet]. 2021. Available from: https://doi.org/10.7910/DVN/VOQCHQ

43. CDC. COVID-19 Vaccinations in the United States,County [Internet]. 2022. Available from: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/distributing/about-vaccine-data.html

44. CDC. PLACES: County Data (GIS Friendly Format), 2021 release [Internet]. 2021. Available from: https://chronicdata.cdc.gov/500-Cities-Places/PLACES-County-Data-GIS-Friendly-Format-2021-releas/i46a-9kgh

45. CDC. Heart Disease Mortality Data Among US Adults (35+) by State/Territory and County – 2017-2019 [Internet]. 2021. Available from: https://chronicdata.cdc.gov/Heart-Disease-Stroke-Prevention/Heart-Disease-Mortality-Data-Among-US-Adults-35-by/s6p7-fvbw

46. US Department of Agriculture. 2013 Rural-Urban Continuum Codes from US Department of Agriculture [Internet]. 2020. Available from: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx

47. Gray DM, Anyane-Yeboa A, Balzora S, Issaka RB, May FP. COVID-19 and the other pandemic: populations made vulnerable by systemic inequity. Nat Rev Gastroenterol Hepatol. 2020;17(9):520–2.

48. DuPre NC, Karimi S, Zhang CH, Blair L, Gupta A, Alharbi LMA, et al. County-level demographic, social, economic, and lifestyle correlates of COVID-19 infection and death trajectories during the first wave of the pandemic in the United States. Sci Total Environ. 2021;786:147495.

49. Kamis C, Stolte A, West JS, Fishman SH, Brown T, Brown T, et al. Overcrowding and COVID-19 mortality across US counties: Are disparities growing over time? SSM-Popul Health. 2021;15:100845.

50. Behrman JR, Rosenzweig MR. Returns to birthweight. Rev Econ Stat. 2004;86(2):586–601.

51. Fuchs VR. Time preference and health: an exploratory study. National Bureau of Economic Research; 1980.

52. Behrman JR, Kohler HP, Jensen VM, Pedersen D, Petersen I, Bingley P, et al. Does more schooling reduce hospitalization and delay mortality? New evidence based on Danish twins. Demography. 2011;48(4):1347–75.

53. Christensen T, Lægreid P. Trust in government: The relative importance of service satisfaction, political factors, and demography. Public Perform Manag Rev. 2005;28(4):487–511.

Table 1. Factor loadings of items in key indexes

Index **Items**	**Factor Loading**
Low Education Attainment	
Less than a high school diploma (%)	0.55
High school graduate (includes equivalency) (%)	0.73
College and associate's degree (%)	-0.20
Bachelor degree or higher (%)	-1.00
Poor health Status (age-adjusted prevalence)	
Heart disease mortality (per 100k)	0.73
Current smoking (%)	0.88
No leisure-time physical activity (%)	0.85
Obesity (%)	0.78
Sleeping less than 7 hours (%)	0.78
Current asthma (%)	0.65
Preventative Health Behavior	
Visits to doctor for routine checkup within the past year (%)	0.76
Cholesterol screening (%)	0.81
Fecal occult blood test, sigmoidoscopy, or colonoscopy (%)	0.59
Mammography use among women aged 50–74 years (%)	0.74
Vacciantion	
Completed dose1 (%)	0.92
Completed dose2 (%)	1.02
Completed booster (%)	0.76

Table 2. Descriptive statistic of main variables (grouped by low education index quartile)

Low Education Index	Q1 (High Education)		Q2		Q3		Q4 (Low Education)	
	Mean	**Std.Dev**	**Mean**	**Std.Dev**	**Mean**	**Std.Dev**	**Mean**	**Std.Dev**
Death Rate (per 100k)[1]	249.69	119.77	355.78	137.49	423.97	144.79	478.19	139.42
Total Deaths[1]	711.47	2207.22	279.02	783.86	170.35	288.53	106.31	97.72
New Deaths per Fortnight	11.98	73.12	4.68	20.31	2.85	8.70	1.79	3.61
Population (100k)	2.76	6.49	0.80	2.14	0.41	0.73	0.23	0.21
Median Houshold Income ($1000)	69.00	20.06	55.91	9.31	51.15	8.51	43.88	8.33
Republican Vote (%)	52.87	17.10	66.81	13.31	69.28	12.77	71.21	13.70
First Vac (%)[2]	67.66	16.08	57.85	13.18	55.27	11.19	51.31	10.89
Secoond Vac (%)[2]	59.71	12.95	51.60	10.59	48.80	9.85	44.79	9.39
Booster (%)[2]	47.11	11.96	45.68	10.63	43.38	9.69	39.92	7.53
Std. Vaccination Index[3]	0.15	1.15	-0.01	0.97	-0.02	0.97	-0.09	0.90
Std. Prevention Index[3]	0.46	1.05	-0.14	1.11	-0.20	0.97	-0.20	0.89
Std. Poor Health Status Index[3]	-0.82	0.69	-0.11	0.70	-0.12	0.68	0.55	0.92

Note:

The unit of analysis for all variables is the county (n = 3,108)

1. 20 Jan 2020 - 10 May 2022
2. on 10 May 2022
3. Mean = 0; Standard Deviation = 1

Table 3. Descriptive statistics

	Mean	Std.Dev	Min	Max
New Deaths per fortnight	5.32	38.44	0	9382
Demography				
Age 65 and over (%)	4.68	1.68	0.47	33.33
Population (10k)	10.43	35.74	0.01	1004
Median household income ($1000)	55.07	15.57	22.29	353.09
Race				
White (%)	84.80	16.94	3.45	100
Black (%)	9.28	14.75	0	87.94
Asian (%)	1.48	3.16	0	55.72
Native (%)	2.13	8.07	0	95.19
Hispanic (%)	9.52	13.87	0	98.88
Geography (dummy variables)				
City	0.37	0.48	0	1
Town	0.29	0.45	0	1
Rural	0.34	0.47	0	1
Education Attainment				
Less than a high school diploma (%)	12.40	6.04	1.39	78.15
High school graduate (includes equivalency) (%)	33.94	7.37	6.52	54.96
College and associate's degree (%)	31.06	5.34	5.93	81.82
Bachelor degree or higher (%)	22.60	9.69	0	79.14
Low Education Attainment (index)	2.79	16.39	-74.87	50.86
Poor Health Status				
Heart disease mortality (per 100k)	354.55	84.06	0	810.5
Current smoking (%)	20.4	4.16	6.50	43.00
No leisure-time physical activity (%)	30.39	5.80	12.90	51.80
Obesity (%)	35.78	4.25	16.40	51.00
Sleeping less than 7 hours (%)	36.81	3.96	25.60	49.10
Current asthma (%)	9.72	0.97	7.20	14.30
Poor Health Status (index)	365.55	71.49	85.7	720.26
Preventative Health Behavior				
Visits to doctor for routine checkup within the past year (%)	74.19	3.56	60.00	82.90
Cholesterol screening (%)	83.44	2.41	67.20	91.00
Fecal occult blood test, sigmoidoscopy, or colonoscopy (%)	61.97	4.45	39.80	74.40
Mammography use among women aged 50–74 years (%)	70.91	3.67	54.00	81.80
Prevention (index)	213.00	8.01	175.09	235.30
Vaccination				
Completed first dose (%)	22.57	25.90	0	99.90
Completed second dose (%)	19.78	25.90	0	99.90
Completed booster (%)	15.16	17.63	0	97.10
Vaccination (index)	52.46	58.70	0	238.37
Political View				
Republican vote (%)	65.04	16.03	8.73	96.18
n(fortnights)	60			

n (counties)	3,108
n (fortnight-counties)	186,480

Table 4. Multilevel negative binomial predicting new deaths per fortnight

	Model 1		Model 2		Model 3	
Predictors	*IRR*	*95% CI*	*IRR*	*95% CI*	*IRR*	*95% CI*
(Intercept)	0.79 ***	0.74 – 0.85	0.79 ***	0.74 – 0.85	0.73 ***	0.70 – 0.77
Main Variables						
Low education attainment	1.34 ***	1.32 – 1.36	1.34 ***	1.32 – 1.36	1.17 ***	1.15 – 1.20
Median household income			0.99	0.97 – 1.00	0.99	0.98 – 1.01
Vaccination					0.60 ***	0.59 – 0.60
Poor health status					1.03 **	1.01 – 1.05
Prevention					0.97 *	0.94 – 1.00
Republican vote					1.11 ***	1.08 – 1.14
Race						
White *(ref.)*	-	-	-	-	-	-
Black	1.00	0.99 – 1.01	1.00	0.98 – 1.01	1.10 ***	1.08 – 1.13
Native	1.03 ***	1.02 – 1.05	1.03 ***	1.02 – 1.04	1.08 ***	1.06 – 1.09
Asian	0.93 ***	0.92 – 0.94	0.93 ***	0.92 – 0.95	0.96 ***	0.94 – 0.97
Hispanic	0.95 ***	0.94 – 0.97	0.95 ***	0.94 – 0.97	1.03 **	1.01 – 1.05
Age						
Age 65 and over	1.15 ***	1.13 – 1.17	1.15 ***	1.13 – 1.17	1.18 ***	1.17 – 1.20
Geography						
City (ref.)	-	-	-	-	-	-
Town	1.08 ***	1.07 – 1.09	1.08 ***	1.06 – 1.09	1.06 ***	1.05 – 1.07
Rural	1.18 ***	1.16 – 1.20	1.18 ***	1.16 – 1.20	1.14 ***	1.12 – 1.16
Time period						
Fortnight (log)	1.64 ***	1.63 – 1.65	1.64 ***	1.63 – 1.65	2.94 ***	2.90 – 2.98
New deaths last fortnight	1.69 ***	1.68 – 1.71	1.69 ***	1.68 – 1.71	1.67 ***	1.65 – 1.68
Observations	186480		186480		186480	
Marginal R^2	0.231		0.231		0.383	
Conditional R^2	0.282		0.282		0.414	

Note:

* p<0.05 ** p<0.01 *** p<0.001

All independent variables are standardlized

Conditional model:

	Estimate	Std. Err	z value	b*x	b*(x-std)	b*(x+std)
(Intercep	-0.31054	0.02639	1	-0.31054	-0.33693	-0.28415
loweducat	0.16021	0.01084	-1.1339	-0.18166	-0.16937	-0.19395
black	0.09918	0.01152	0.00	0	0	0
native	0.07639	0.00686	0.00	0	0	0
asian	-0.0442	0.00692	0.00	0	0	0
hispanic	0.0287	0.00889	0.00	0	0	0
`65over`	0.16873	0.00735	0.00	0	0	0
town	0.05795	0.00609	0.00	0	0	0
rural	0.13108	0.00772	0.00	0	0	0
weeklog	1.07938	0.00698	0.00	0	0	0
lastfortr	0.51208	0.00445	0.00	0	0	0
mhi	-0.00698	0.00768	0.00	0	0	0
poorhealt	0.0267	0.00883	0.00	0	0	0
preventic	-0.03082	0.01353	0.00	0	0	0
percent_r	0.10486	0.01221	0.00	0	0	0
vaccinati	-0.51262	0.00431	0.00	0	0	0
			sum(b*x)	-0.4922	-0.5063	-0.4781
			exp(sum(b	0.611279	0.602721	0.619958
			offset =	10.426	10.426	10.426
			y*offset	6.373193	6.283971	6.463682
			death/yr	165.703	163.3832	168.0557

-1.1339 1.0927

		2.319777	2.352714
High Educ	166	163	168
Low Educa	237	228	246
	71.02709	8.88049	9.226609